TRANSPORTATION RESEARCH

# RECORD



No. 1304

*Maintenance*

---

# Highway Maintenance Operations and Research 1991

*A peer-reviewed publication of the Transportation Research Board*

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

**Transportation Research Record 1304**
Price: $38.00

Subscriber Category
IIIC Maintenance

TRB Publications Staff
*Director of Publications*: Nancy A. Ackerman
*Senior Editor*: Naomi C. Kassabian
*Associate Editor*: Alison G. Tobias
*Assistant Editors*: Luanne Crayton, Kathleen Solomon, Norman Solomon
*Graphics Coordinator*: Diane L. Ross
*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.

Highway maintenance operations and research 1991.
p. cm.—(Transportation research record, ISSN 0361-1981; 1304)
ISBN 0-309-05111-8
1. Roads—Maintenance and repair. 2. Bridges—Maintenance and repair. 3. Concrete—Corrosion. I. National Research Council (U.S.). Transportation Research Board. II. Series.
TE7.H5 no. 1304
[TE220]
388 s—dc20
[625.7'6] 91-30812
CIP

**Sponsorship of Transportation Research Record 1304**

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

**Maintenance Section**
*Chairman: Jimmy D. Lee, North Carolina Department of Transportation*

Committee on Maintenance and Operations Management
*Chairman: Bernard H. Ortgies, Sverdrup Corporation*
*Secretary: Dorothy L. Andres, Consultant-Highway Maintenance Programs, Lawrenceville, New Jersey*
*Kenneth A. Brewer, Clyde A. Burke, John P. Burkhardt, Bertell C. Butler, Jr., Thomas L. Cain, R. Franklin Carmichael III, Brian E. Cox, Edward H. Crowe, Asif Faiz, John S. Jorgensen, W. M. Lackey, Kjell Levik, Michael J. Markow, Joseph A. Mickes, Dean L. Morgan, James S. Moulthrop, George R. Russell, James W. Shay, Leland D. Smithson, Theodore E. Stephenson, Marshall L. Stivers, Jerome J. Thomas, David C. Wyant*

Committee on Traffic Safety in Maintenance and Construction Operations
*Chairman: Nicholas J. Garber, University of Virginia*
*Secretary: William J. Buglass, Wilbur Smith Associates*
*R. F. Benekohal, James E. Bryden, Gerald A. Donaldson, J. R. Doughty, Conrad L. Dudek, Michael N. Gostovich, Ernest D. L. Huckaby, Mohammad M. Khan, Russell M. Lewis, Virginia M. Lorenz, James Migletz, Craig Miller, Stephen H. Richards, Robert K. Seyfried, S. C. Shah, Robert L. Shea, Frank D. Shepard, T. Darcy Sullivan*

Committee on Pavement Maintenance
*Chairman: Roger E. Smith, Texas A&M University*
*Roy R. Almendarez, Kathryn A. Cation, James G. Chehovits, Michael I. Darter, Barry H. Dunn, Robert P. Elliott, Jim W. Hall, Jr., David J. Halpenny, Kenneth R. Hoover, Starr D. Kohn, Claus C. Kuehl, Jay K. Lindly, James B. Martin, James S. Moulthrop, David E. Newcomb, Frazier Parker, Jr., Michel Ray, Michael M. Ryan, Charles F. Scholer, S. C. Shah, Mohamed Y. Shahin, Scott Shuler, T. Paul Teng, Paul E. Theberge*

Committee on Structures Maintenance
*Chairman: Robert N. Kamp, Albany, New York*
*John R. Allen, Bernard R. Appleman, William G. Byers, Ian J. Dussek, Gill M. Gautreau, Keith Giles, Ray W. James, Eldon D. Klein, Robert H. Krier, Jimmy D. Lee, Patrick McCarthy, Daniel D. McGeehan, Wallace T. McKeel, Jr., Daniel S. O'Connor, Jack W. Roberts, George P. Romack, Arunprakash M. Shirole, Lloyd M. Smith, Robert A. P. Sweeney, Alden L. West*

Committee on Winter Maintenance
*Chairman: Donald M. Walker, University of Wisconsin–Madison*
*Duane E. Amsler, Sr., James D. Arnoult, Robert R. Blackburn, Robert L. Brown, Albert Carlier, Henri De Lannoy, Charles E. Dougan, Karl H. Dunn, Joe Fedosoff, Darryl L. Hearn, David G. Ibbott, James F. Kelley, Henry W. Kirchner, Sung M. Lee, Byron Nelson Lord, Kenneth A. MacDermid, L. David Minsk, Robert J. Nowak, Kynric M. Pell, Max S. Perchanok, Rodney A. Pletan, Michael M. Ryan, Stephen F. Shober, Ronald D. Tabler, Jan Verseef, D. L. Walters*

Committee on Corrosion
*Chairman: Andrew D. Halverson, Minnesota Department of Transportation*
*John E. Bennett, Kenneth J. Boedecker, Jr., John P. Broomfield, Gerardo G. Clemena, Carl F. Crumpton, Robert J. Girard, William H. Hartt, Robert H. Heidersbach, Jr., Donald R. Jackson, Daniel P. Johnston, David G. Manning, A. P. Moser, James A. Riemenschneider, Arnold M. Rosenberg, Ellen G. Segan, Yash Paul Virmani*

**GROUP 5—INTERGROUP RESOURCES AND ISSUES**
*Chairman: William J. Harris, Jr., Texas A&M University*

Committee on Low Volume Roads
*Chairman: Ronald W. Eck, West Virginia University*
*Robert A. Cherveny, Gerald T. Coghlan, Santiago Corro Caballero, Asif Faiz, Gordon M. Fay, Gerald E. Fisher, Richard B. Geiger, Jacob Greenstein, Henry Hide, Stuart W. Hudson, Kay H. Hymas, Lynne H. Irwin, Milton L. Johnson, Thomas E. Mulinazzi, Andrzej S. Nowak, Neville A. Parker, George B. Pilkington II, Jean Reichert, Richard Robinson, Bob L. Smith, Walter J. Tennant, Jr., Harry H. Ulery, Jr., Alex T. Visser, Michael C. Wagner*

Frank N. Lisle and G. P. Jayaprakash, 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 1304

## Contents

# Foreword

Maintenance of the highway infrastructure is a critical element in sustaining and expanding the U.S. economy. This Record contains 35 papers on highway maintenance operations activities and research results intended to assist maintenance engineers in improving the efficiency and effectiveness of maintenance efforts. The Record is divided into six sections: maintenance and operations management, pavements, structures, corrosion, snow and ice control, and traffic safety.

The maintenance and operations management section contains papers that address recent developments in France, maintenance level-of-service, location of highway division offices, traffic signal record systems, earth road maintenance costs, selection of maintenance alternatives, and implementation of maintenance management systems. The papers in the pavements section present details on preventing reflective cracking and an ultrathin friction course. Experiences with rapid techniques for the repair and protection of bridge decks, integrated bridge management systems, and bridge deck condition surveys are reported in the structures section. The corrosion section presents research results on a variety of subjects, including concrete permeability, protection and rehabilitation treatments, inhibitors, cathodic protection, and electrochemical removal and measurement of chloride ions. The papers in the snow and ice control section contain details on displacement plows, deicing chemicals, passive snow control, snowplow routing, and chemical undercutting. The traffic safety section contains information on urban work zone traffic control, striping of no-passing zones, radar, flagger training, and truck-mounted attenuators.

# Urban Roads in France: Overview and Recent Developments in Maintenance Management Aids

JEAN PIERRE CHRISTORY AND PIERRE LAYE

Decentralization trends in France in recent years have made decision makers and engineers aware of the importance of having modern computer-based tools for the effective management of the major infrastructures represented by urban roads. Considerable developmental work was initiated and conducted jointly by the Association of Engineers of French Cities and the decentralized highway engineering network of the Ministry of Equipment. This has led to a range of computerized maintenance management programs designed with standard, user-friendly features for easy adaption to the everyday requirements of the highway departments of local communities. An observatory for general maintenance practices was set up, allowing the forecasting of requirements from the standpoint of the general approach and programming, as well as from the finer standpoint of diagnostics and expertise. These applications are examined and the two computer-aided tools most frequently used are described in greater detail along with the most promising technical and organizational potentialities for their distribution to engineers. What is involved is the Relational Tool for Maintenance Management (ORAGE) management aid system and the Expert System for Highway Diagnostics and Maintenance (SEVADER).

In France as in many other industrialized countries, over three-fourths of the population lives in cities, and a significant increase in this population has been observed in particular during the past three of four decades.

The importance of urban road facilities, their essential role in the life of the city, and their technical, economic, and social implications (Figure 1) lead to a growing demand for management aids. Recent decentralization measures have also brought about increased awareness of the need for better dialog between technicians and decision makers, thus calling into question some earlier practices that did not always optimize the respective roles of the technical expert, consultant, and decision maker.

These new conditions have made it possible to provide local communities with modern tools making extensive use of computer technology to monitor practices, to take better stock of existing facilities and their surroundings, and to allow better diagnostics and choices of action to be taken.

The Association of Engineers of French Cities (AIVF) and more particularly its working group on highway facilities, and the Techincal Network of the French Ministry of Equipment, through its central and regional organizations, have been working towards this objective and have pooled their efforts in the design, production, experimentation, and distribution of such computer aids. [The Technical Network of the Ministry of Equipment is headed in this area mainly by the Urban Transport Research Center (CETRA), the Road Research Laboratories (LCPC), and the Road and Freeway Engineering Research Office (SETRA).]

The AIVF, whose members include engineering technicians of regional communities, represents the profession

- Within public and semipublic organizations,
- Within private corporations, and
- Within the International Federation of Municipal Engineers, which participates actively in the various research and documentation actions relative to all the engineering tasks of regional communities.

Products are examined that are most attractive to local communities in the area of technical management of maintenance for highways and surroundings, on the basis of the criteria of the number of sites and the potential demand.

## PRACTICES

The AIVF undertook a survey in 1985 to establish an overview of maintenance practices on urban roadway pavements.

Almost half of the 400 urban communities and cities in France with over 15,000 inhabitants responded to the survey, thus allowing the establishment of a computerized data base serving as an inventory and an observatory of maintenance practices on the national level. The data base contains valuable information, from which the following characteristic figures of urban road facilities were summarized:

| *Data Item* | *Amount* |
|---|---|
| Total area of road facilities | 430 million $m^2$ |
| Average area per capita | 20 $m^2$ |
| Annual renewal of pavement | 20 million $m^2$ |
| Average renewal cycle | 17 years |
| Average annual expenditure per capita for maintenance | 72 francs (1985) |

The following remarks can be made about these summary data:

- The total surface area of road facilities in cities with over 15,000 inhabitants represents the equivalent of 60,000 km of

J. P. Christory, Ministry of Equipment, Housing, and Transport Road Research Laboratory, West Greater Paris Region, 12 Rue Teisserenc de Bort, Trappes, France 78190. P. Laye, Association of Engineers of French Cities, Mairie de Cavaillon, Place Joseph Guis, B.P. 37, Cavaillon, France 84300.

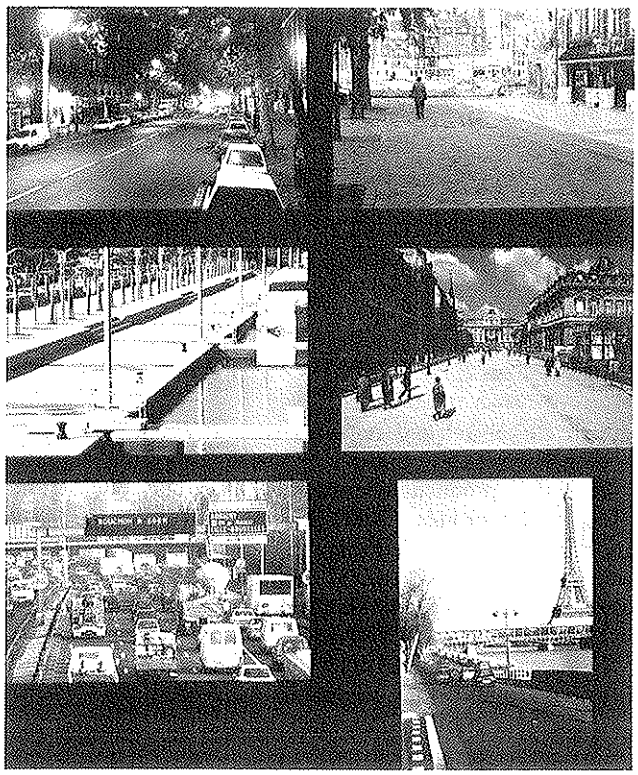


**FIGURE 1 Urban roads: a broad variety of users, utilizations, and functions.**

two-lane roads. This number is 10 times more than the motorway network and twice as much as the national highway network in France.

- There is an average of 20 m² of road per capita. This number can vary twofold from one city to another.
- The surface area contracted annually for the maintenance of urban roads in cities of over 15,000 inhabitants is approximately 20 million m² (which is the same order of magnitude as for the national roads).

On average, 5 to 6 percent of the surface area of existing road facilities is repaved annually as part of maintenance and improvements, with an average renewal rate of 17 years.

## NEEDS

In addition to knowledge of field conditions encountered daily by the partners of the AIVF and the Technical Network of the Ministry of Equipment, joint studies were conducted further upstream on the subject of decision preparation tools. As in the case of the study of practices, these investigations were based on surveys aimed at political decision makers and their technical counterparts. All the diversity of local communities (towns) and territorial authorities (departments) was represented during interviews designed to allow better understanding of what is to be expected now and in the future from the dialog between decision maker and technician.

From this work, conducted within the broadcast framework of the National Committee on Technological Forecasting for Highway Engineering, two major points were brought out:

1. Among the major issues that highway professionals will be confronted with in the next 5 to 8 years is the *growth trends in management aid requirements*.
2. Decision makers wish to have three types of aids that are clearly differentiated and categorized with respect to time:

- Aids for determining the status and quality of service of highway networks,
- Aids for defining a policy and for scheduling of highway works, and
- Aids for checking job completion against job planning.

Two clearly distinguished expectations also appear as a common denominator to these aids.

1. Better coordination of highway works and underground works on networks. The interconnection of information systems and the programming aids of the different occupants of the public domain is a major expectation of officials from technicians within the next 5 to 10 years. This challenge is now achievable by technicians owing to progress already made or expected within the stated time frame in computer systems as regards hardware, software, and networking. Highway managers must take initiatives in order to gradually achieve consistency in the tools and methods used by different players of the public domain.
2. Excellent user-friendliness of products derived from these aids, and even of the aids themselves. Decision makers wish to have clear comprehensive documents allowing immediate perception by other officials and the general population of the messages involved. Topical maps constitute excellent media in this respect, the full importance and potentialities of which have not yet been understood by technicians.

## POSSIBLE RESPONSES TO THESE EXPECTATIONS

Cooperation between the AIVF and the technical network of the Ministry of Equipment does not claim to be the only possible means of progressing with respect to the questions raised earlier. Many local or individual initiatives contribute to the improvement of knowledge. General and joint programs, nevertheless, make it possible to constantly have an overall view of the relation between needs and responses, and to apply more easily the consistency and complementarity required.

Thus, for highway network maintenance management, two approaches are considered separately, but from the standpoint of consistency (Figure 2).

1. To reply to the question as to where and when to intervene, the road network is considered as a whole. Use is made of the most sophisticated forms of management aids, optimized management models. The first-generation tool of this type most widely used in local communities is called ORAGE, the French acronym for "relational aid for highway management."
2. To reply to the question as to how to intervene, the network is no longer considered as a whole, but in terms of one or more pavements. Use is made of detailed diagnostic

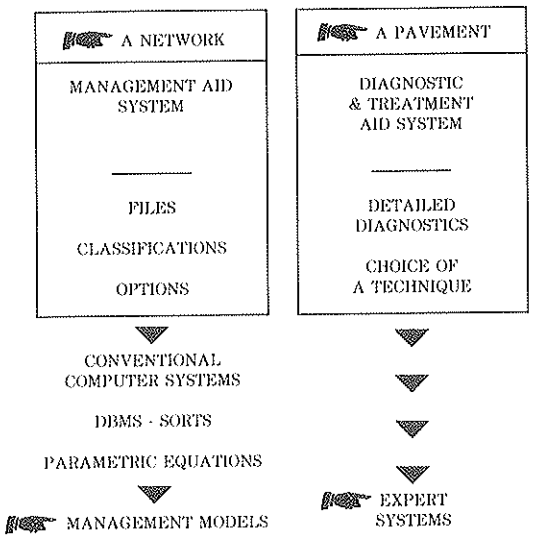



**FIGURE 2 Management logic: pavement surveying diagnostics and choice of remedies.**

methods, the most advanced examples of which are the expert systems. The tool of this type most widely used by local communities is called SEVADER, from the French for "expert system for highway engineering diagnosis and repair recommendations."

ORAGE and SEVADER may be used either independently of each other or together, because they are linked on compatible microcomputers. By mid-1990, ORAGE was set up at about 30 sites and SEVADER at 7 sites. The required equipment depends on the size of the town (number of highway segments to be managed). It is summarized in Table 1.

## DESCRIPTION OF ORAGE SYSTEM

### ORAGE and Products of Its Generation

In the family of management models, ORAGE is a first-generation tool. This designation means that it is intended primarily for setting up an order of priority for the sections to be maintained. This classification is obtained by combining,

TABLE 1 HARDWARE AND SOFTWARE RECOMMENDED FOR USE OF ORAGE MANAGEMENT MODEL AND SEVADER EXPERT SYSTEM CONSIDERED SEPARATELY OR TOGETHER

| | ORAGE ALONE | | | SEVADER ALONE | | ORAGE + SEVADER | |
|---|---|---|---|---|---|---|---|
| *Application* / *Equipment* | *Average city* *1200 seg* | *Large city* *3000 seg* | *Very large city* *20.000 seg.* | | | *Large or Very large city* | |
| Type of Micro | PC 386 > 16 MHz | PC 386 > 33 MHz | PC 386/33 SERVER + PC 386/16 Work Station | PC 386 20 MHz | PS 386 | PC 386 > 33 MHz | PS 386 > 33 MHz |
| RAM | 4 MB | 4 MB | 8 MB | 5 MB | 5 MB | 4 MB | > 8 MB |
| Screen | Color | Color | Color | Color | Color | Color | Color |
| Additional board | Coprocessor 80387 | Coprocessor 80387 | Coprocessor 80387 | no | no | no | no |
| Hard disk | > 100 MB speed < 28 ms | > 200 MB speed < 28 ms | > 320 MB speed < 28 ms | > 40 MB | > 40 MB | > 100 MB | > 100 MB |
| Printer | yes | yes | yes | yes | yes | yes | yes |
| RDBMS. | ORACLE Ver. 5.1.b. | ORACLE Ver. 5.1.b. | ORACLE Ver. 5.1.b. | | | ORACLE Ver. 5.1.b. | ORACLE Ver.5.1.b. |
| Other software | Drivers library and graphics | Drivers library and graphics | Drivers library and graphics | Windows ESCOP + le LISP | Windows ESCOP + le LISP | Windows ESCOP + le LISP | Windows ESCOP + le LISP |
| Image Bank Additional board | DVA 4000 | DVA 4000 | DVA 4000 | | | DVA 4000 | DVA 4000 |

with a certain weighting (specific to the objectives of each city and adaptable according to the type and function of road facilities), the different indicators characterizing the urban road and its surroundings. In the basic version, the basic criteria adopted are

- Deterioration of pavement surface,
- Bearing capacity of pavement structure,
- Loading,
- Condition of gutters and curbs, and
- Condition of sidewalks.

In addition to these basic criteria, multicriterion analysis sorts from descriptive files of road facilities make it possible to integrate large and diversified actions calling for an urban approach: functionality of streets, types of users, presence of underground networks, history, etc.

### ORACLE, A Good Driver for ORAGE

ORAGE is designed around the ORACLE relational data base management system (DBMS). The DBMS, of American origin, has taken on a leading role worldwide in its category. This choice is based on the following considerations:

- ORACLE is a relational data base management system supported by many systems (UNIX, MS/DOS, VM, etc.);
- The MS/DOS system is identical in minicomputer and mainframe systems;
- A warm start is obtained on the microcomputer—in the event of power failure or any other trouble, the integrity of the data base is maintained;
- File and program access confidentialities are managed on the microcomputer.

### Working Documents and Reproduction Associated with ORAGE

From the outset and throughout its development, the ORAGE system favored the quality of information reproduction, including that of intermediate data collection working documents. This feature was applied in the user-friendly design of screens for dialogs between user and machine as well as in computer graphics reproduction. A few examples of outputs from the data base generated on the site of the city of Montreuil sous Bois (of 120,000 inhabitants) in the Paris area follow:

- The dictionary of streets established in accordance with Standard AFNOR Z 20 008 of November 1976.
- Segmentation of the road network is prepared using a precise composition methodology based on rough preparation work on plan and field validation. A segment is the part of a street between two intersections. Intersections are integrated in a segment next to it.
- A preedited and personalized file for each segment contains general information on the makeup and the quantification functions for deterioration (Figure 3).
- Evaluation of road and classification of segments. The rating scale varies from 0 to 100 points, 100 being the most problematic case. The weighting is applied on the basic rating value for five adopted criteria: bearing capacity, deterioration, condition of curbs and gutters, condition of sidewalks, and traffic.
- Statistical reproduction. Two major applications are involved:
  - —Multicriterion interrogations to determine the streets and the associated quantitative data meeting specific requirements, for example, a street in bad visual condition with bus lanes and local activities consisting of businesses.
  - —The histograms of the overall or partial rating values for the segments that constitute service indicators and their evolution by comparing, year by year, the average rating values and the distribution of the ratings. In this regard, action strategies may be oriented and followed up by overall assessments.
- Topical maps (e.g., Figure 4). Operational for many years on minicomputers and color electrostatic plotters producing high-level documents, the ORAGE system is now implemented on the microcomputer in connection with intermediate working documents. The last reliability tests are under way, so that it will be possible in the near future to provide highly integrated services meeting the requirements of decision makers as well as technicians, namely the display of results, both as regards more specific issues resulting from multicriterion interrogations as well as overall assessments associated with the definition and with the follow-up of general maintenance strategies (Figure 5).

The completion of a topical map involves the following phases and options:

1. Choice of elements to be represented, such as corridors, roads, and networks;
2. Vectorizing of graphic elements to be represented (digitizing tablet and subsequent scanning);
3. Segmentation and structuring of pavements;
4. Choice of a referential (Lambert II);
5. Handling of representation tools such as color palette, lines, symbols, characters, and types of representation (centered symbol, lines, and bands);
6. Map items such as logo, legend, and title page; and
7. Graphic standards and software allowing fullest possible independence of software and hardware (e.g., DISSPLA) and graphic files.

Unlike computer-aided design (CAD), mapping does not create—it reproduces map bases and structured data that exist. The reproduction of any event occurring on these existing bases is immediate and its representation can be changed as desired.

## DESCRIPTION OF SEVADER SYSTEM

### Setup of SEVADER Operation

Following a documented survey by the Structures Working Group of the AIVF aimed at 400 French urban cities and communities of over 15,000 inhabitants, about 10 towns agreed to participate intellectually and financially in the building of

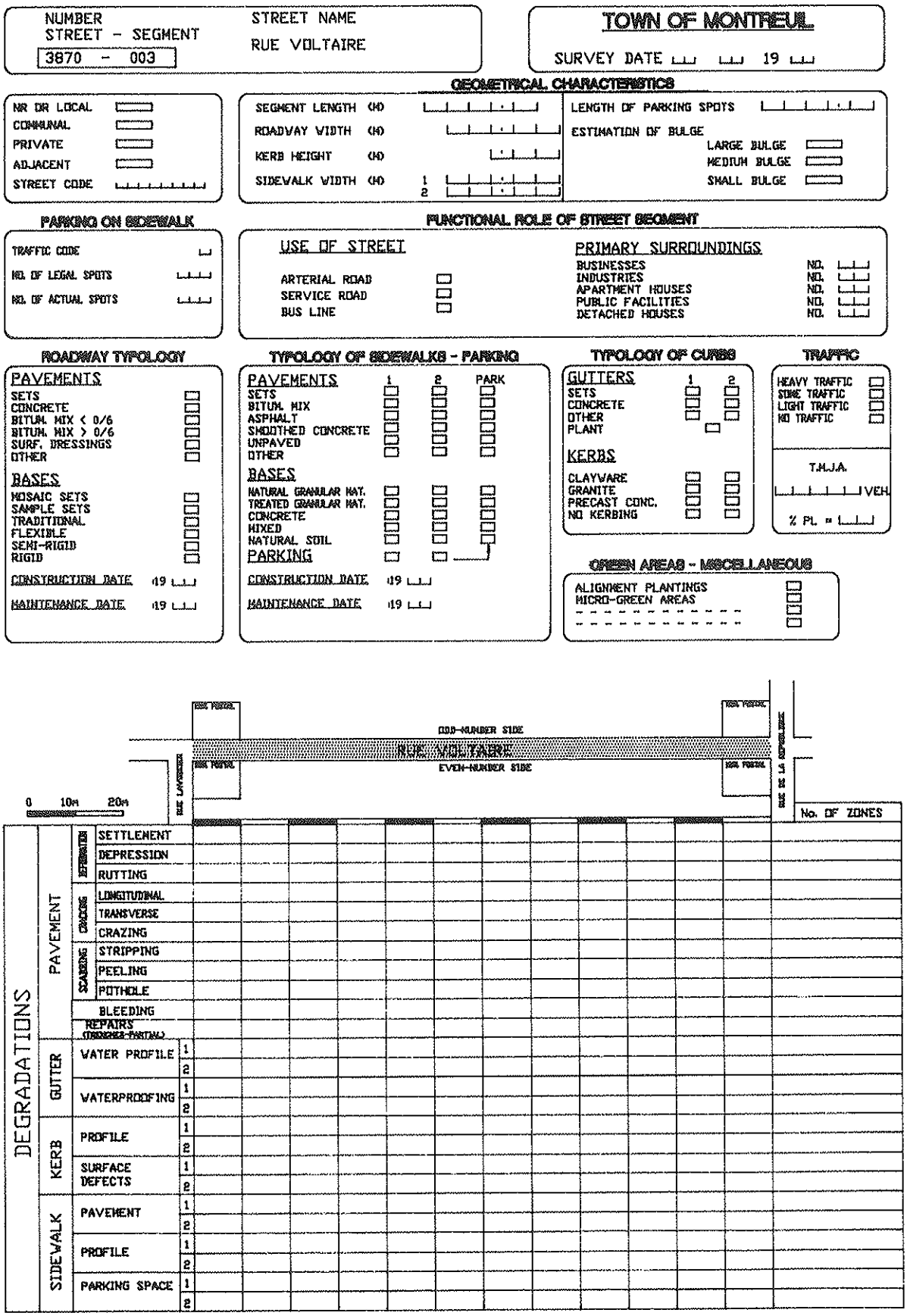



FIGURE 3 Segment file record containing typology and quantification functions for road deterioration.

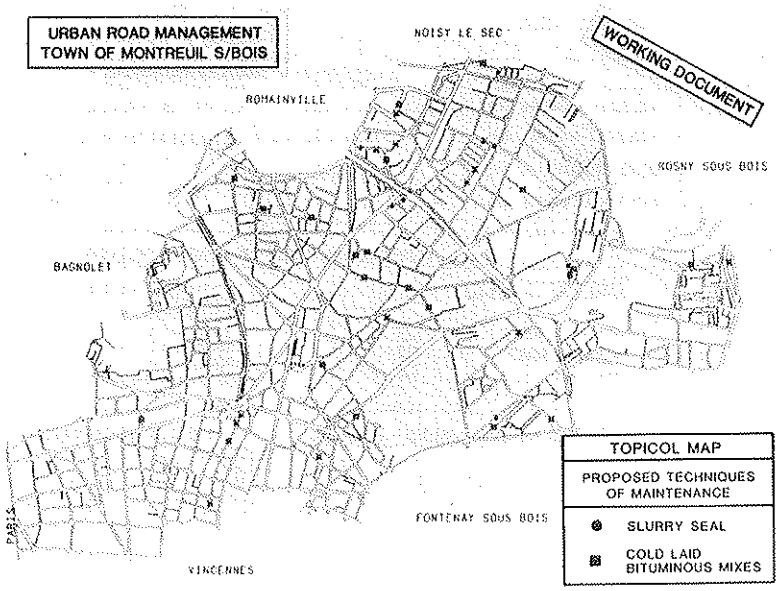



**FIGURE 4 Example of topical map.**

this expert system on the basis of a 50 percent cofinancing arrangement between the cities and the Ministry of Equipment.

The setup of the operation made it possible to distinguish

1. Pilot sites:

- Paris township—directorate of roads,
- Urban community of Lyons,
- Urban community of Lille,
- City of Nantes, and
- Laboratoire Régional de l'Ouest Parisien (LROP), Ministry of Equipment.

2. Associated sites:

- Urban community of Le Creusot-Montceau les Mines
- Urban district of Bayonne-Biarritz-Anglet
- City of Besançon
- City of Caen
- City of Reims

The latter city was the first community to be equipped with the complete system of computer aids for the technical management of highway maintenance with the SEVADER system combined with the ORAGE management system provided by the LROP.

### Originality and Strong Points of SEVADER Expert System

One of the major concerns of the practicing engineers who designed SEVADER was to meet actual field requirements by remaining pragmatic and close to the material and human conditions of the technical departments. Thus, the following choices were made (Figure 6).

1. SEVADER had to handle all existing urban road pavements: bituminous materials, hydraulic concrete, sett pavements, and slabs.
2. SEVADER had to use current computer equipment, already available in main communities and costing less than 100,000 francs. Specifically, what was involved was an AT- or XT-compatible microcomputer in which dedicated cards (AMAIA cards) were added.
3. SEVADER had to incorporate the most characteristic features of the urban environment and its complexity. The analysis relative to the choice of surfacings for overlaying and for grinding-resurfacing incorporates—in addition to classical concepts of geometric feasibility, durability, costs, etc.—such important questions for the urban community as

- Environment, surrounding population, and noise constraints;

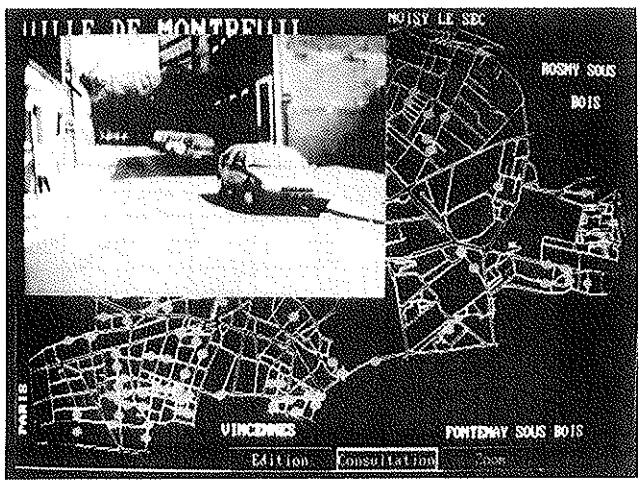



**FIGURE 5 Example of display of results from multicriterion interrogations and overall assessments.**


**FIGURE 6 Main objectives set for the design of SEVADER expert system.**

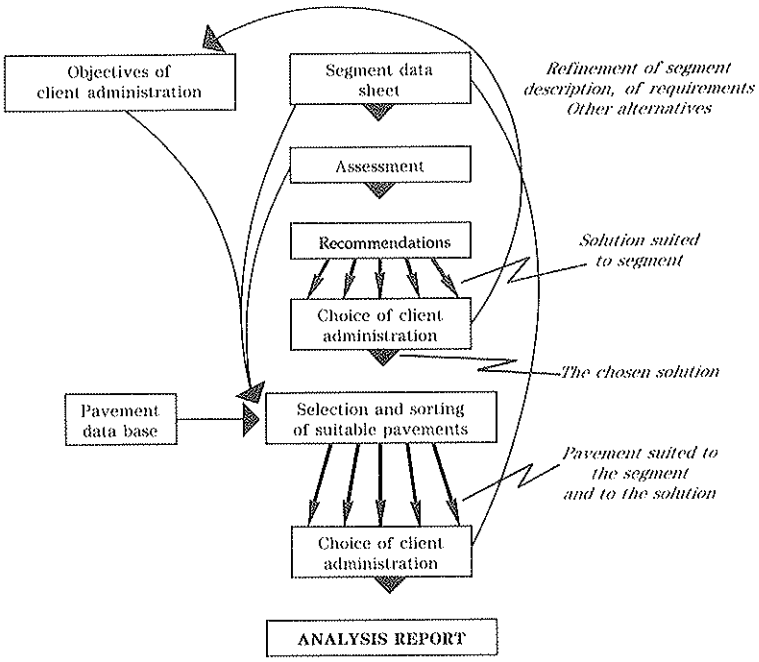



**FIGURE 7 SEVADER—a phase in the analysis of a road facility.**

- Repairability within the more general question of maintenance;
- Aesthetics and cleaning ease;
- Time required for restoral of service; and
- Workability in relation to the emergency of networks.

4. SEVADER had to contribute to a certain technical consistency and to a common experimental platform between the different players in the area concerned. For example, the knowledge base for special contractor surfacings was integrated into the SEVADER 01 version (with core common to cities) along with processes that formed technical recommendations according to the procedures set up jointly by the trade and by public authorities. This option is the foundation for an efficient maintenance procedure because of the situations that are constantly changing, sometimes within a short period.

### Description of SEVADER Expert System

As in any expert system, SEVADER distinguishes between knowledge, on the one hand, and computer logic that implements this knowledge. Thus, SEVADER is made up of

- A commercially available inference engine, ESCOP, Version 5, written in LISP programming language (specializing in the field of artificial intelligence). ESCOP was designed and developed by the ITMI Company and is in fact an expert system development generator that also delivers the knowledge expression language, the data processing system, screen multiwindowing, etc., making the application as user friendly as possible under the environmental computer constraints set.
- Three main knowledge bases implemented within the framework of a logical development of analysis and reasoning (Figure 7).
  - —Diagnostic knowledge base, which generates the assessment of the examined road facility;
  - —Action recommendation knowledge base, which feeds the search for possible solution typologies, for example, joint sealing, ultrathin surfacing, grinding-resurfacing, etc.
  - —Pavement knowledge base, which aids in determining the best choice of products within the solution family chosen after the preceding phase. This base allows a finer presentation of several choices to the user.

The implementation of these different modules and the service functions associated with them represent about 15,000 lines of program and over 600 reasoning rules. These results are included in the SEVADER documentation delivered to users. These developments indicate the logic used within each module of the SEVADER architecture.

### Description of Road Segment

Three screens, predocumented partially in the case of the SEVADER–ORAGE combination, characterize this description (Figure 8):

1. Structural and geometrical description;
2. Degradation and surveying (specific schedule by typology of pavement, bituminous pavements, concrete pavements, sett pavements, and slabs); and
3. Environment and loading of road facility.

The road segment investigated is thus characterized by 65 indications, 60 percent of which are mandatorily documented but, in most cases, this task is facilitated because what is involved is simply locating the information on the segment within a prior classification. In addition, an effective on-line aid is constantly offered to the user.

## MANAGEMENT OBJECTIVES

The concerns of the highway manager may be summarized by six main themes: safety, durability, environment, comfort, maintenance, and economy. For each project, the goals set may place greater or lesser emphasis on each of these themes. The SEVADER user can hierarchize choices and choose certain qualitative options, such as the colors for the pavement.

## RECOMMENDATIONS FOR ACTION

The logic implemented consists of the following:

- Leaving the segments within problems, i.e., having a satisfactory evaluation with regard to the objectives indicated;

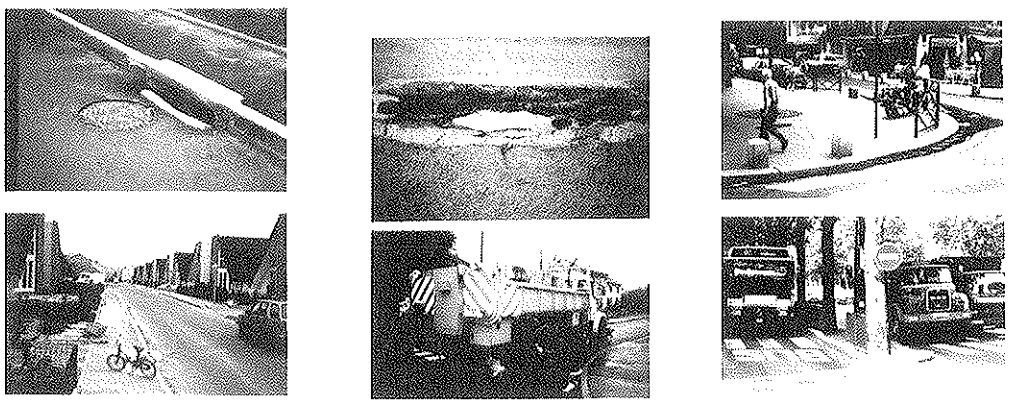


**FIGURE 8 Three families of data characterizing the urban streets and roads examined by SEVADER: (*left*) structural and geometrical description, (*center*) degradation and surveying, (*right*) environment and loading of road facility.**

- Determining what can be physically set up considering the technical and geometrical constraints;
- Determining what would be the minimum required to meet average durability objectives (life of 7 to 8 years) on the basis of condition evaluations and traffic evaluations; and
- Using the terms of the comparison between what is physically possible and the required minimum to establish several proposals able to satisfy as well as possible the hierarchy of desired objectives.

Two types of solutions are derived from this analysis: overlays covering the entire range of thicknesses, from 1 to 10 cm, and grinding-resurfacing.

## SELECTION AND SORTING OF SUITABLE SURFACINGS

The surfacings knowledge base documents the available surfacings within each solution typology in the sense defined earlier. By mid-1990, this base included about 70 products.

Each product is described quite completely. A certain number of properties are qualified by a notation system from 1 to 6 according to a hierarchy established within each solution family typology. The costs indicated in FY 1988 of course call for readjustment on the local level for each location site.

At the end of the phases, and after having interrogated the user several times on crucial points of the analysis of his choices and his options so that SEVADER may remain within its function of aid to proposals, the tool delivers the analysis report on request, summarizing the main points of the reasoning.

## SOME ASPECTS OF OTHER SOFTWARE OPERATIONAL IN CITIES

In relation with or as a complement to the two products described earlier, two systems are intended for areas oriented more towards predictive models with an overall aid in the choice of conventional solutions applied to the entire highway network, or towards detailed knowledge of the allocation of expenditures for maintenance and improvements on the road.

In the first case, what is involved is Follow-up and Maintenance of Urban Pavements (SECUR), which presently includes three main modules:

- A data base with four information levels,
- An information system, and
- A management aid system, including the calculation of quality of road from viewpoint of user (QCU) and visual quality of structure (QVS) grades and the search for conventional solutions for maintenance works.

In the second case, what is involved is Software Aid for Follow-up of Street Improvement Costs (LASCAR), which has four objectives:

- Understanding the formation of costs for the road,
- Understanding the condition of the road capital,
- Characterizing the improvements and expenditures to evaluate the implemented policy, and
- Predicting future expenses by annual follow-up.

## SUBSEQUENT ACTION TO BE CONSIDERED

The awareness of the capital and stakes represented by urban road facilities by an increasing number of officials of territorial communities has led to a growing demand for tools aiding in the management of road facilities, oriented initially toward the crucial aspect of their maintenance.

The structuring of the French highway effort and the efforts to concentrate and pool the complementary know-how between the practicing engineers of cities and the representatives of a technical network of the government, which is highly decentralized and present in the field, has made it possible to better determine the issues involved and especially the expectations of decision makers and managers.

In particular, it was possible to separate the requirements of today and of the near future on the one hand, and those of the medium term on the other. A large majority of decision makers wished to start with the setup of simple and pragmatic (first-generation) tools, structured so that they could be gradually upgraded toward more optimized systems once the community acquires the basic knowledge (for initialization and updating of the tools). The dissemination of products of this type has been initiated and is taking place satisfactorily in the urban cities and communities of France.

It is now necessary to organize a general follow-up and assessment of the adoption of these tools by local highway managers. This has already been under consideration by user clubs created in connection with partnership arrangements between AIVF engineers and the Technical Network of the French Ministry of Equipment, as in the case of the effective organization adopted for the initialization and maintenance of these software tools and the methods associated with them.

For the longer-term requirements, it is necessary to pursue and intensify research and experimentation to check the possible degrees of suitability between normal evolution in the urban area caused by increased functions (network improvement, etc.) and the maintenance laws in the more classical sense of road wear. Similarly, the cost aspect must be a permanent and major factor in the tools.

The interconnection between the different tools developed by the occupants of the public domain is still felt to be a strong requirement. Computer links between data collection systems, management models, and expert systems designed for detailed diagnostics and technical and economic choices of maintenance and rehabilitation alternatives constitute an initial phase of this chain of homogeneous production of information on highway facilities. It is up to highway managers to take the initiatives for, to initiate, and to extend this process. Engineers in France must meet many challenges in the coming decade to fulfil the expectations of the decision makers.

*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Application of Markov Decision Process to Level-of-Service-Based Maintenance Systems

SIVA GOPAL AND KAMRAN MAJIDZADEH

Systems based on level of service (LOS) currently implemented to manage highway maintenance use extensive subjective data. Collection of these data is tedious and expensive and the inherent uncertainties in the data render the results imprecise. A modified method using the Markov decision process (MDP), which overcomes most of the drawbacks of the LOS-based systems, is described. The adoption of the MDP is consistent with progressive evolution in the field of highway maintenance management. It introduces measures of performance benefits that are less subjective than those used in the NCHRP LOS model, which rely on attributes and utility functions. The modified model uses three types of key input data: transition probabilities, costs, and relative-importance weights. The transition probabilities are computed analytically using sample deterioration models and quality standards. An approach is described for computing the cost of each alternative from historical data. An analytic approach is also described for computing relative-importance weights using simple ranking of and comparison scores for the highway elements. This method was tested with 58 highway elements in 12 strata and 3 levels of service each. The resulting problem, which had 2,088 variables and 697 constraints, required less than 15 min on an IBM PC using an off-the-shelf linear programming package. The results of the test were consistent with the input data and demonstrated that the objectives set for the method were being met. Although the method was tested with mostly roadside elements, it can generally be used with any LOS-based system.

Under NCHRP sponsorship, a method was developed to assist in selecting optimum levels of service (LOS) for those highway elements that are subject to the constraints of available resources (*1*). LOS values are discrete condition state thresholds or maintenance intervals for highway elements. This method is supposed to be based on theoretically sound principles of decision analysis and to be implementable in a well-defined, step-by-step procedure. The method uses both objective data and subjective expert opinion. The subjective data are required for the following areas:

- Specification of alternate LOS values;
- Estimation of effects of alternate LOS values on various user considerations such as safety, preservation of investment, etc.;
- Assessment of individual value functions of different attributes used to measure the user considerations; and
- Assessments of relative-importance weights among the individual attributes.

Resource International, Inc., 281 Enterprise Drive, Westerville, Ohio 43081.

Further, considerable amount of subjective input is needed from experienced maintenance engineers in estimating the resource requirements for each LOS of individual elements. Availability of significant amounts of objective maintenance cost data is unlikely as records are not kept in such detail. Although deterioration models for the highway elements are not explicitly determined in this method, they certainly play a role in the experts' minds in the establishment of explicit LOS values and the corresponding resource requirements.

As discussed earlier, the implementation and operation of this model largely depend on extensive use of subjective expert opinion and management input. The potential for large variability is expected to be high in both these inputs: but the model is deterministic and does not implicitly or explicitly handle these variabilities.

Further, the integration of this model with the existing pavement management system (PMS) and bridges and structures management systems (BSMS) for the purpose of budget allocation requires top management (TM) to make adjustments to performance standards for the PMS and BSMS and budget for the LOS-based nonpavement management system (NPMS). The models can then be analyzed again; the expected costs can be summed across pavements, bridges, and structures and added to the NPMS budget; and the total can be judged for acceptability on the long-term basis. This approach to integration has the advantage that TM has direct control over the performance and budget-setting process. The disadvantage is that a large number of parameters are involved in setting all standards, and a considerable amount of adjustments to parameters may be needed to achieve an affordable solution.

Therefore, this method needs significant changes to overcome some of its inherent drawbacks:

1. Deterministic models fail to account for the variability inherent in the process.
2. Extensive dependency of the models on subjective expert opinion is a potential source of significant modeling errors.
3. Incompatibility of the models with those of other state-of-the-art management systems for integration with limited amount of adjustments by the TM in setting the performance standards.

A modified method overcomes most of these drawbacks. This method will be described with its input and output requirements. This description will then be followed by procedures to generate the input data required.

## METHODOLOGY

The development of this modified model is in keeping with the progressive evolution that has occurred in the field of highway maintenance management as LOS models have given way to a Markov decision process (MDP) model, in particular, for pavements and bridges. Although retaining some of the basic features of the LOS models, this model assumes the deterioration of the highway elements to be a Markov process.

This model introduces measures of performance benefits that are less subjective than those used in the NCHRP LOS model and that rely on attributes and utility functions. The assumption of a Markov process replaces deterministic deterioration of elements by probabilistic deterioration. Considering the overall need for an integrated highway maintenance management system, this model is compatible with other MDP-based pavement and bridges and structures management systems, making such integration feasible.

The measures of performance require definition of QS values, which are discrete ranges of condition states corresponding to the alternate LOS values. These QS values should take into account the regional, climatic, geographical, and road classification differences in the network of highways. This accounting is accomplished by stratifying the highways using some or all of these factors and defining the alternate LOS (and QS) value for each stratum. Figure 1 shows the concepts of LOS and QS for a single stratum or a group of similar strata; $n_1$, $n_2$ and $n_3$ are the corresponding maintenance intervals.

### Model Description

In order to describe the new model, some definitions and notation are needed. Let

$f_{sj}(l|t)$ = probability that an element $j$ in Stratum $s$ would reach LOS $l$, in $t$ time periods after the last action. These time periods may be in days, weeks, or months;

$T_{sjl}$ = number of time periods between actions for the element $j$ in Stratum $s$ at LOS $l$;

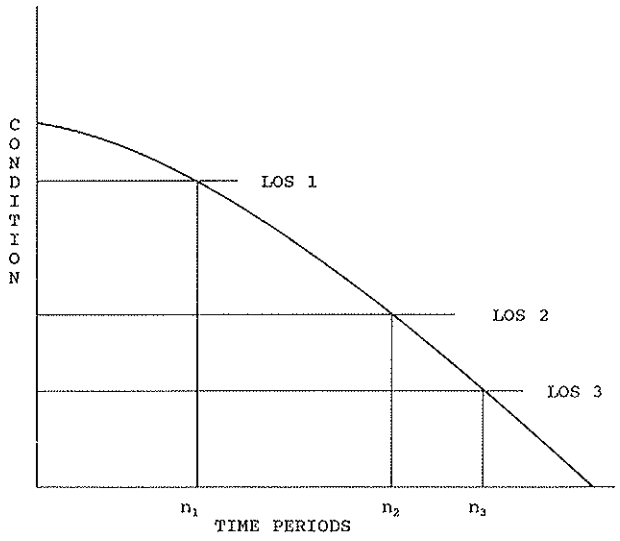



**FIGURE 1 Deterioration curve for a highway element.**

$w_{sj}(d)$ = relative-importance weight for being within a desirable QS for Element $j$ in Stratum $s$;

$w_{sj}(u)$ = relative-importance weight for not being within an undesirable QS for Element $j$ in Stratum $s$;

$D_{sjl}$ = probability that Element $j$ in Stratum $s$ will be within a desirable QS when the cycle time corresponding to LOS $l$ is selected;

$U_{sjl}$ = probability that an Element $j$ in Stratum $s$ will be within an undesirable QS when the cycle time corresponding to LOS $l$ is selected;

$\overline{N}_{sj}$ = normalized number of units of Element $j$ in Stratum $s$ (i.e., the fraction);

$X_{sjl}$ = logical variable representing the choice of cycle time (1 if selected, 0 otherwise); and

$C_{sjl}$ = annual cost of applying an action at LOS $l$ to all the units of Element $j$ in Stratum $s$.

In this notation, it is assumed that an element requiring different actions for different conditions will be considered a separate element for each condition. For example, a retaining wall with poor appearance is different from a retaining wall with damaged structure.

Given this notation, the following relations may be defined:

$$D_{sjl} = (1/T_{sjl}) \sum_{i \in d} \sum_{t=1}^{T_{sjl}} f_{sj}(i|t) \tag{1}$$

$$U_{sjl} = (1/T_{sjl}) \sum_{i \in u} \sum_{t=1}^{T_{sjl}} f_{sj}(i|t) \tag{2}$$

With these equations, the probabilities of finding an element in a desirable or undesirable QS, given a particular cycle time $T_{sjl}$, can be computed. The model to select the cycle times that maximize the performance is an assignment model with a budget constraint as follows:

$$\text{Max} \sum_s \sum_j \overline{N}_{sj} \times \left[ w_{sj}(d) \sum_l D_{sjl} X_{sjl} + w_{sj}(u) \sum_l (1 - U_{sjl}) X_{sjl} \right] \tag{3}$$

subject to

$$\sum_l X_{sjl} = 1 \qquad \text{for all } s \text{ and } j \tag{4}$$

$$\sum_s \sum_j \sum_l C_{sjl} X_{sjl} \leq B_p \tag{5}$$

$$X_{sjl} = 0 \text{ or } 1 \qquad \text{for all } s, j, \text{ and } l \tag{6}$$

$B_p$ is the budget for periodic maintenance of the nonpavement elements.

The objective function of Expression 3 is the weighted sum of transition probabilities. Equation 4 ensures that one and only one cycle time is selected for each element in each stratum, Inequality 5 is the budget constraint, and Equation 6 defines the decision variables to be (0 or 1) integer variables.

As can be seen from the objective function, the functional performance of the nonpavement elements are represented

by stochastic variables $D_{sjl}$ and $U_{sjl}$. In the NCHRP LOS-based model, the deterioration functions are implicitly deterministic and the benefit functions are utility functions based on subjective expert opinion. The stochastic variables can incorporate the variability of performance among different units of the same element.

The MDP-LOS-based model will still need subjective data for relative-importance weights $w_{sj}(d)$ and $w_{sj}(u)$, as discussed later. However, these parameters are signficantly fewer than those needed to define the user consideration attributes and the various utility functions. Further, these relative-importance weights are thought to be easily obtainable because the experts have a better feel for these weights than for the consideration attributes by LOS values.

This model is formulated as a (0 or 1) linear integer problem with a special structure. The number of generalized upper bound (GUB) type constraints, as in Equation 4, greatly exceeds the number of linear variables plus the general constraints, Inequality 5. Consequently, when the integer constraints are replaced by nonnegativity constraints, the resulting linear program has an optimal solution with mostly integer values (*2*). This approximation can provide a satisfactory solution of the original problem.

### Model Implementation

The solution of this model requires several input data, assuming that the network of highway elements are identified and stratified, and that alternative LOS and QS values are defined:

1. Number ($\overline{N}_{sj}$) of units of Element $j$ in Stratum $s$ as a fraction of the units in the network,
2. Probability $D_{sjl}$ that an element in Stratum $s$ will be within a desirable QS when maintained at LOS $l$,
3. Probability $U_{sjl}$ that an Element $j$ in Stratum $s$ will be within an undesirable QS when maintained at LOS $l$,
4. The maintenance cost $C_{sjl}$ for Element $j$ in Stratum $s$ at LOS $l$,
5. Relative-importance weights $w_{sj}(d)$ of an Element $j$ in Stratum $s$ to be within a desirable QS, and
6. Relative-importance weights $w_{sj}(u)$ of an Element $j$ in Stratum $s$ to be outside an undesirable QS.

The output of the system is an optimum set of LOS values to be implemented.

Item 1 is obtained from the inventory. Estimation of $D_{sjl}$ and $U_{sjl}$ is based on Equations 1 and 2, where the unknown function is $f_{sj}(l|t)$, the probability that an Element $j$ in Stratum $s$ would reach LOS $l$ in $t$ time periods after the last action. Methods for determining function $f_{sj}$, for estimating probabilities $D_{sjl}$ and $U_{sjl}$, and for estimating costs $C_{sjl}$, and the subjective expert input for assessing the relative-importance weights $w_{sj}(d)$ and $w_{sj}(u)$ are described in later sections.

## DETERMINATION OF TRANSITION PROBABILITIES

The approach to determining the transition probabilities requires knowledge of how the highway element deteriorates with time and the definition of the QS values. Defining the deterioration functions and the quality standards for the highway elements has some similarity to Step 1 of the NCHRP method (*1*). In this step, elements are selected, maintenance conditions are identified, and alternate LOS values are specified. These concepts are shown in Figure 1, in which the vertical axis indicates the condition and the horizontal lines represent the various LOS values for a given element. The only difference is that the deterioration of this element is indicated by the curve, which explicitly defines the condition as a function of time.

Specification of LOS values implies quality standards as ranges of condition between these LOS values. In the NCHRP method, the LOS values were defined in relation to budget levels. Alternatively, quality standards can be defined using engineering judgment with management input. For example, the full range of conditions can be divided into excellent, acceptable, barely acceptable, and unacceptable standards. Then the boundaries of these standards will be the alternate LOS values, and the intersections of these boundaries with the deterioration curve will define alternate maintenance intervals, such as $n_1$, $n_2$, and $n_3$ on the time axis shown in Figure 1.

Ideally, deterioration functions can be defined from historical maintenance data, but in reality such data are not available for all elements. Then, engineering judgment would be relied on. Assume that a group of experienced maintenance engineers could provide a series of deterioration times for a given set of conditions. These conditions may be thresholds of quality standards or some other condition values that would make it easier for the engineers to estimate the corresponding times.

Deterioration functions are obtained either by fitting these data to simple mathematical functions or by simply joining adjacent points with straight lines. Then, the maintenance times corresponding to the LOS values can be determined by inverting the functions for the conditions implied by these levels.

To compute the transition probability matrix elements (**tp**) from the deterioration functions and the quality standards, the deterioration of an element is assumed to be a Markov process in which the transition probability from the current LOS value to any of the LOS values in time period $t$ depends only on the current condition and is independent of how it got there. If the time period is chosen small enough, then it can be assumed that the element can transition only into the next level down, if at all. That is, if the probability of this element's staying in the same level is $p$, the probability of its transitioning into the next level is $(1 - p)$. Assuming four quality standards, the transition probability matrix for Element $j$ in Stratum $s$ for one period can be written as

$$\mathbf{tp} = \begin{vmatrix} p_1 & 1-p_1 & 0 & 0 \\ 0 & p_2 & 1-p_2 & 0 \\ 0 & 0 & p_3 & 1-p_3 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (7)$$

where Subscripts 1, 2, and 3 refer to the first three quality standards. Elements in the last quality standard will stay in that level with probability 1. After $n$ periods, this matrix will

be

$$\mathbf{tp}^n = \begin{vmatrix} p_1 & 1-p_1 & 0 & 0 \\ 0 & p_2 & 1-p_2 & 0 \\ 0 & 0 & p_3 & 1-p_3 \\ 0 & 0 & 0 & 1 \end{vmatrix}^n \tag{8}$$

At time $t = 0$, the element is assumed to be in the best QS. After $t = n$ periods, the probabilities $P_{.n}$ of this element's being in QS 1, 2, 3, or 4 are given by

$$(P_{1n}, P_{2n}, P_{3n}, P_{4n}) = (1, 0, 0, 0)(\mathbf{tp}^n)_{sj} \tag{9}$$

If the maintenance interval corresponding to LOS $l$ is $T_{sjl}$, then

$$D_{sjl} = \frac{1}{T_{sjl}} \sum_{i \in d} \sum_{n=1}^{T_{sjl}} P_{in} \tag{10}$$

and

$$U_{sjl} = \frac{1}{T_{sjl}} \sum_{i \in u} \sum_{n=1}^{T_{sjl}} P_{in} \tag{11}$$

The only information needed to define the probability functions is the transition probability matrix elements $\mathbf{tp}_{sj}$, in which the specific unknowns are the individual probabilities $p_1$, $p_2$, and $p_3$.

Determining these individual probabilities depends on the availability of field data. As could be expected, no historical maintenance data are available. Consequently, establishing the matrix element $\mathbf{tp}_{sj}$ for Element $j$ in Stratum $s$ is proposed to be accomplished in two steps: (a) determining initial matrices, and (b) updating these matrices from the field data. The details of these steps are discussed in the next section.

### Transition Probability Calculation

The calculation of transition probabilities is based on some form of deterioration models developed either from expert opinion or from historical maintenance data. How these probabilities are computed from the deterioration models is discussed in this section.

The structure of the **tp** matrix for $n$ periods (Equation 8) is as follows:

$$\mathbf{tp}^n = \begin{vmatrix} p_1^n & q_n & z_n & * \\ - & p_2^n & * & * \\ - & - & p_3^n & * \\ - & - & - & 1 \end{vmatrix} \tag{12}$$

where $p_1^n$, $p_2^n$, and $p_3^n$ are the $n$th powers of $p_1$, $p_2$, and $p_3$, respectively; $q_n$ and $z_n$ are unknown terms; terms denoted with an asterisk are also unknown, but their exact values are not important to this discussion. By induction,

$$q_n = (1-p_1)(p_1^{n-1} + p_1^{n-2}p_2 + p_1^{n-3}p_2^2 + \cdots + p_1 p_2^{n-2} + p_2^{n-1}) \tag{13}$$

and

$$z_n = (1-p_2)(q_{n-1} + q_{n-2}p_3 + \cdots q_2 p_3^{n-3} + q_1 p_3^{n-2}) \tag{14}$$

With these expressions, the deterioration model for an Element $j$ in Stratum $s$ shown in Figure 1 can be examined.

The deterioration curve is an expression of condition as a function of time defined from data solicited from expert maintenance engineers or from historical data. Therefore, it is conjectured that these engineers considered an average or a median unit of an element in providing this opinion. The latter means that the deterioration curves for 50 percent of the units are below the median curve and the other 50 percent are above it. With this assumption, the following equations can be written:

$$p_1^{n1} = 0.5 \tag{15}$$

$$p_1^{n2} + q_{n2} = 0.5 \tag{16}$$

$$p_1^{n3} + q_{n3} + z_{n3} = 0.5 \tag{17}$$

Equation 15 means that after $n_1$ periods the probability of a unit of element being in QS 1 is 0.5; after $n_2$ periods, the probability of this unit being in QS 1 or 2 is also 0.5; similarly, after $n_3$ periods, the probability of this unit being in QS 1, 2, or 3 is 0.5.

Equation 15 provides a value of $p_1$; a combination of Equations 13 and 16 can be solved for $p_2$; and Equations 13, 14, and 17 can be solved for $p_3$. Consequently, $(\mathbf{tp}^n)_{sj}$ can be evaluated for any $n$. Then a combination of Equations 9–11 can be used to compute the necessary probabilities to be used in the objective function in Equation 13.

### Update Considerations for the tp Matrices

Initially, the **tp** matrices are based on a deterioration model developed using subjective expert opinion. When historical data become available, these data will be used to generate more realistic deterioration models from which new **tp** matrices will be computed.

## RESOURCE REQUIREMENTS

Maintenance cost is a necessary input to the budget constraint in Equation 5 for each element in each stratum by LOS value. This cost depends on unit cost and other quantities depending on the stratum and element. This is true for the LOS value for which the unit cost and quantities are applicable. In order to generalize it for other LOS values, an intensity factor needs to be considered. This factor reflects the change in the level of effort required to accomplish the maintenance each time. In order to estimate the annual cost, this cost must be multiplied by the frequency of maintenance per year. This process can be expressed as follows:

$$C_{sjl} = U_{sj} Q_{sj} I_{sjl} F_{sjl} \tag{18}$$

where

$U_{sj}$ = unit cost of action on Element $j$ in Stratum $s$,
$Q_{sj}$ = quantities involved,
$I_{sjl}$ = intensity factor for LOS $l$ (= 1 for base case), and
$F_{sjl}$ = frequency of maintenance for the year.

$U_{sj}$, $Q_{sj}$, and $F_{sjl}$ are expected to be estimated from historical data. The intensity factor could come from engineering judgment. Alternatively, if the level effort is assumed proportional to the range of conditions over which the action restores the element, $I_{sjl}$ may be defined as

$$I_{sjl} = \frac{\text{LOS}(*) - \text{LOS}(l)}{\text{LOS}(1) - \text{LOS}(3)} \tag{19}$$

where LOS(*) is the base LOS, for which $I_{sjl} = 1.0$.

## RELATIVE-IMPORTANCE WEIGHTS

The objective function of Expression 3 uses relative-importance weights $w_{sj}(d)$ and $w_{sj}(u)$, where

$w_{sj}(d)$ = relative-importance weight for having an Element $j$ in Stratum $s$ within desirable quality standards, and
$w_{sj}(u)$ = relative-importance weight for having an Element $j$ in Stratum $s$ not within undesirable quality standards.

In defining the quality standards, the differences caused by road stratification criteria such as traffic volume and terrain have been accounted for by specifying varying QS values for different strata. The objective function in the optimization model accounts for the quantities of elements by the fraction $\overline{N}_{sj}$, which is the normalized number of units of Element $j$ in Stratum $s$. Consequently, these factors are not used in estimating the relative-importance weights $w_{sj}(d)$ and $w_{sj}(u)$ (because they are already included in the QS values), the subscript $s$ can be dropped. In the rest of the discussion, these weights are designated as $w_j(d)$ and $w_j(u)$.

The approach chosen in estimating the relative-importance weights is a general expert opinion elicitation procedure called the "analytic hierarchy procedure" (*3*). This approach consists of two sequential steps. In the first step, the nonpavement elements are divided into a hierarchy of groups of 4 to 8 elements each and a set of ranking and comparison criteria are identified. Eight is considered a reasonable maximum number of items that an engineer can be expected to rank and score among as a group. Then the experts are expected to rank each item of these groups at all levels. In the second step, a relative ratio comparison score is given to each item of these groups. These scores are then reduced to obtain the relative-importance weights.

### Concepts of Ranking and Comparison Scores

Suppose there are $n$ items $I_1, I_2, \ldots, I_n$ that need to be evaluated with respect to some Criterion $C$. An expert is asked to make judgments about these items with respect to Criterion $C$. The first step is for the expert to rank-order the items from best to worst, with respect to Criterion $C$. Next, suppose that the items have been reindexed in this order so that the index order and the rank order are the same. The third step is for the expert to compare each pair $(I_i, I_j)$ where $j > i$ to determine how much better $I_i$ is than $I_j$ with respect to Criterion $C$. Then the expert is asked to respond with ratio-scale comparisons rather than interval-scale comparisons. For example, in comparing $I_i$ with $I_j$, a response of 3 would mean that $I_i$ is three times as good as $I_j$ with respect to Criterion $C$ (rather than three additional units better). After some empirical testing and much experience, Saaty (*3*) suggests that these ratio comparisons be based on a standard scale such as the following:

| *Ratio* | *Comparison Significance* |
|---|---|
| 1 | $I_i$ and $I_j$ are about the same with respect to $C$. |
| 3 | $I_i$ is slightly better than $I_j$ with respect to $C$. |
| 5 | $I_i$ is much better than $I_j$ with respect to $C$. |
| 7 | $I_i$ is considerably better than $I_j$ with respect to $C$. |
| 9 | $I_i$ is so much better than $I_j$ with respect to $C$ that there is almost no comparison between the two. |

Values of ratios of 2, 4, 6, and 8 may be used to strike a compromise between adjacent categories.

This process may be organized in the upper part of a matrix such as the example shown in Figure 2. Such matrices are referred to as "pairwise-comparison" matrices.

In Figure 2, a value at the intersection of, say, Row $I_2$ and Column $I_4$ would represent how much better $I_2$ is over $I_4$ with respect to Criterion $C$. Because the items were first rank-ordered with respect to Criterion $C$, the numbers provided by the expert should all be greater than or equal to 1 and tend to increase from left to right.

For a given set of highway elements, a number of matrices like Figure 2 must be elicited from experts. These matrices may differ from each other in two ways. First, each deals with different types of items to be compared; second, the comparisons may be made with respect to different criteria. These items and criteria form a hierarchy the structure of which can be used to synthesize the elicited values into the relative importance weights needed to establish the optimal NPMS policy.

### First Level of the Hierarchy

At the first level, Criterion C is thought of as the importance of the considerations to the overall goal or objective of the NPMS. This overall objective may not be articulated, but the expert engineers should have an adequate sense for making these comparisons. Among others, the items that must be

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ |
|---|---|---|---|---|---|
| $I_1$ | 1 | 3 | 5 | 5 | 9 |
| $I_2$ | - | 1 | 2 | 3 | 7 |
| $I_3$ | - | - | 1 | 5 | 8 |
| $I_4$ | - | - | - | 1 | 3 |
| $I_5$ | - | - | - | - | 1 |

**FIGURE 2 Example of pairwise-comparison matrix that an expert participant might provide for five items.**

compared may include the following four considerations:

- Safety,
- Aesthetics,
- User convenience, and
- Preservation of investment in highway elements.

In order to accomplish the first step, the task asked of the expert would be to rank-order these items with respect to how important they are for the overall nonpavement system. A matrix is then set up with these items listed as row labels in the order of most important to least important, and with columns labeled in the same order. The expert then compares the items that label the rows and columns that intersect in the upper half of the table. The comparisons are scored on the scale of 1 to 9, with the score indicating how much more important for the NPMS the row item consideration is than the column item consideration. This effort results in a single pairwise-comparison matrix.

### Second Level of the Hierarchy

At the second level, the problem expands because the elicitation process needs to be applied for each of four considerations. The items to be compared are the groups of nonpavement elements that have already been identified. If the groups contain about 14 or 15 elements, they need to be split into two groups of 7 or 8 elements each. The groups are then rank-ordered with respect to one of the considerations, and then the pairwise-comparison matrix is established.

This level results in one pairwise-comparison matrix for each consideration.

### Third Level of the Hierarchy

The bulk of the effort is in establishing a pairwise-comparison matrix for the elements in each group with respect to each of the considerations.

Comparisons are not necessary between elements in different groups. The number of pairwise-comparison matrices in this level is the number of considerations, $N_c$, times the number of groups, $N_g$.

### Fourth Level of the Hierarchy

For this level, it is determined for each element and consideration whether being in a desirable QS is more or less important than not being in an undesirable QS and by how much.

As an example, one may pose the question for a given element, say, cateyes; and for consideration, suppose safety, as follows: for purposes of *safety*, is it more important for a segment of *cateyes* to be in a desirable QS than for it to not be in an undesirable QS, and by how much? The how much? answer should be expressed in the 1 to 9 scale defined earlier. The italicized words *safety* and *cateyes* are then replaced by another consideration and element, respectively, and the question is posed again. The process continues until all combinations of considerations and elements have been examined.

### Relative-Importance Weight Calculations

Each pairwise-comparison matrix is made into a square matrix. The diagonal is filled with 1s and the upper triangular part of the matrix is filled in with the evaluated elements. The area below the diagonal is filled in with the reciprocals of the numbers above the diagonal, each in the reflected position across the diagonal. For example, suppose $a_{rc}$ is an element in Row $r$ and Column $c$ and suppose that it is below the diagonal, then its value should be set to $1/a_{rc}$, which is the reciprocal of an element above the diagonal. For a numerical example, the matrix of Figure 2 is converted into the following appropriate matrix:

$$\begin{vmatrix} 1 & 3 & 5 & 5 & 9 \\ \frac{1}{3} & 1 & 2 & 3 & 7 \\ \frac{1}{5} & \frac{1}{2} & 1 & 5 & 8 \\ \frac{1}{5} & \frac{1}{3} & \frac{1}{5} & 1 & 3 \\ \frac{1}{9} & \frac{1}{7} & \frac{1}{8} & \frac{1}{3} & 1 \end{vmatrix}$$

This matrix contains some redundant information, so that the consistency of the participant expert can be checked. If the participant expert was entirely consistent, then the largest eigenvalue of this matrix would be 5 (because the matrix is $5 \times 5$). Therefore, the check is to see that the largest eigenvalue is close to 5.

Assuming that the consistency is satisfactory, the eigenvector corresponding to the largest eigenvalue is normalized so that the sum of its components equals 1. The components of this normalized eigenvector are interpreted by Saaty (*3*) to be priority weights for the elements involved.

In the case of the given matrix, the eigenvalue was 5.36, indicating a reasonable degree of consistency. The normalized eigenvector is (0.50, 0.22, 0.18, 0.07, 0.03), the components of which are the priority weights for the items $I_1$, $I_2$, $I_3$, $I_4$, and $I_5$.

Even Level 4 of the hierarchy produces a series of $2 \times 2$ matrices that have an eigenvalue of 2 and an eigenvector with two components representing priority weights one for the desirable condition and one for the not undesirable condition. Thus, each level of the hierarchy results in a set of priority weights that sum to 1 for each set of items that is directly compared at that level. The weights at each level are distributed and combined at the next level down in the hierarchy. This accomplishment results from an averaging process described in the following section.

### Weight Distribution for Analytical Hierarchy Procedure

At each level of the hierarchy, weights are developed from each matrix by calculating the eigenvector corresponding to the largest eigenvalue for the matrix. The weights are then combined as follows:

At Level 1, weights are found for the criteria. Let $\alpha_c$ be those weights for $c = 1, 2, \ldots, N_c$.

At Level 2, weights are found for groups of elements for each criterion. Let $\beta_{cg}$ be those weights for $c = 1, 2, \ldots,$

$N_c$; and $g = 1, 2, \ldots, N_g$. In this case, for each $c$, the sum over $g$ of the weights is 1.

At Level 3, weights are found for elements within groups for each criterion. Let $\gamma_{cge}$ be those weights for $c = 1, 2, \ldots, N_c$: $g = 1, 2, 3, \ldots, N_g$; and $e = 1, 2, \ldots, n_g$, where $n_g$ is the number of elements in Group $g$. These weights sum to 1 when summed over the elements of the specific group and criterion.

At Level 4, weights are found for each of two states of elements within groups for each criterion. The two states are (a) a desirable QS and (b) a not undesirable QS. Let $\delta_{ggeq}$ be those weights for which the indices $c$, $g$, and $e$ range as before, while $q = 1, 2$. As before, for fixed $c$, $g$, and $e$, the sum of these weights over $q$ is equal to 1.

In the mathematical models proposed earlier, the elements were not grouped but simply indexed sequentially. Thus, let $j$ be the element index so that $j$ assumes a value for each feasible pair $(g, e)$. The state also had index $q = d$ or $u$ instead of 1 and 2. The weights to be used in the optimization model were therefore $w_d$ and $w_u$. They were calculated as follows:

$$w_d = \Sigma\, \alpha_c \beta_{cg} \gamma_{cge} \delta_{cge1} \tag{20}$$

$$w_u = \Sigma\, \alpha_c \beta_{cg} \gamma_{cge} \delta_{cge2} \tag{21}$$

## TEST RESULTS AND CONCLUSIONS

The method described was tested with a system of 58 highway elements in 12 strata with 3 LOS values each. The resulting linear integer programming problem had 2,088 variables and 697 constraints. An off-the-shelf linear programming package (LP83) solved this problem on an IBM PC (286) in less than 15 min. Only one element in one stratum had a noninteger solution.

Both the deterioration models as well as the relative-importance weights were developed with subjective engineering judgments. Several maintenance engineers were used to collect the data. Significant variations were observed among the deterioration time estimates provided by these engineers. Improvement in the solicitation process and the use of formal analysis techniques such as Delphi to get consensus-based opinions would reduce such variabilities. Also, if a maintenance management system already exists, the development of deterioration models can use input from historical maintenance data.

The results of this test indicated that the maintenance policy generated was consistent with the input data. Detailed discussion of these results would be the subject of another paper. However, the results indicated that the objectives set for this method at the outset of this paper were being met. This method uses stochastic deterioration models and relies on relatively fewer and easier-to-get subjective data. Further, the optimization model formulated here is consistent with the other systems based on the MDP so that these models can be integrated for budget allocation exercises.

Even though this method is tested with mostly nonpavement elements, it can generally be used with any LOS-based system.

## REFERENCES

1. R. Kulkarni, F. Finn, K. Golabi, R. Johnson, and E. Alviti. *NCHRP Report 223: Maintenance Levels of Service Guidelines*. TRB, National Research Council, Washington, D.C., 1980.
2. L. S. Lasdon. *Optimization Theory for Large Systems*. McMillan, New York, 1970.
3. T. L. Saaty. *The Analytical Hierarchy*. McGraw-Hill, New York, 1980.

*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Comprehensive Study of the Location of Highway Division Offices

DANIEL S. TURNER AND NORMAN D. PUMPHREY, JR.

A study was undertaken for the Alabama Highway Department to determine whether field maintenance operations would be more efficient if one or more of the nine division offices were relocated, or if an additional division office was constructed. Historical changes in division office locations or division boundaries, previous studies performed by the department to find optimum locations of district (county-level maintenance) offices, and results of extensive interviews with department managers were examined. Additionally, the research staff obtained data on field office locations from other southeastern departments of transportation, compared division characteristics to look for more efficient arrangements, conducted two different types of modeling exercises, and performed an economic analysis. The study found that one existing division office could be moved approximately 50 mi to substantially enhance work travel patterns. The payback period for the relocation of this office would be 5 to 7 years. As a result of the study, the Alabama Highway Department has begun the relocation of the division office.

Highway agencies must make subjective decisions while selecting new locations for field maintenance offices. There appear to be few or no firm criteria that may be used to decide when a new office is needed or how to optimize the location of the new office. This study was conducted by the University of Alabama to assist the Alabama Highway Department in making such decisions.

The project work included analysis of historical changes in department field offices, interviews with department managers, examination of previous studies by the department to relocate district offices (county-level maintenance offices), survey of other southeastern state departments of transportation (DOTs), comparison of characteristics of existing department division offices, two modeling efforts, and an economic analysis.

## ORGANIZATION OF THE ALABAMA HIGHWAY DEPARTMENT

The central offices of the Alabama Highway Department are located in Montgomery, Alabama. This location provides convenient access to all of the department's field offices. For management of maintenance and construction activities, the department has divided the state into nine large geographical areas called divisions. The divisions are relatively autonomous. They are subject to the policies and funding provided by the central office, yet their managers have great leeway in directing operations. Within each division, there are three to six subdivisions called districts. Each division and each district has its own office complex. This study focused on whether new divisions were justified or whether existing divisions could be reconfigured (by realigning districts) to increase service or efficiency of operation. The division boundaries and district office locations at the time of this study are shown in Figure 1.

D. S. Turner, Civil Engineering Department, University of Alabama, P.O. Box 870205, Tuscaloosa, Ala. 35487–0205. N. D. Pumphrey, Jr., Civil Engineering Department, Louisiana Tech University, P.O. Box 10348, Ruston, La. 71272–0046.

## HISTORICAL CHANGES

The department had periodically shifted districts between divisions for the sake of efficiency and has occasionally created a new division to keep pace with growth in the state. These actions are shown by three substantial changes of the last 25 years.

Special division offices had been created in Birmingham and Montogmery to guide the development of the Alabama Interstate system. These offices were abolished in 1965. At the same time, an entirely new division was created for the Montgomery region from parts of the existing Divisions 3, 4, and 7, bringing the total to eight division offices.

Until 1973, Division 1 stretched across the entire top of the state. This width became awkward and too large to manage efficiently, so it was subdivided to create two divisions. When this occurred, all divisions in the state were renumbered from the northeast to the southwest, and at least eight districts were simultaneously shifted to new divisions. In 1980, the third major change occurred when division boundaries were again realigned. At least three districts were transferred at this time into the adjacent division for the sake of efficiency.

## PREVIOUS FIELD OFFICE LOCATION STUDIES

The department had conducted two previous studies to determine optimum locations of offices; however, both of these studies were directed toward finding the best locations for district offices within a given division. In both studies, the researchers applied analytical techniques to minimize employee travel time in reaching job sites. The linear programming technique was used to find the number and location of offices that would minimize travel time (i.e., lost work time) while employees traveled to the job site. The studies documented two important issues. First, of all the factors considered during the two analyses, employee travel time was found to be the most important in optimizing total roadway main-

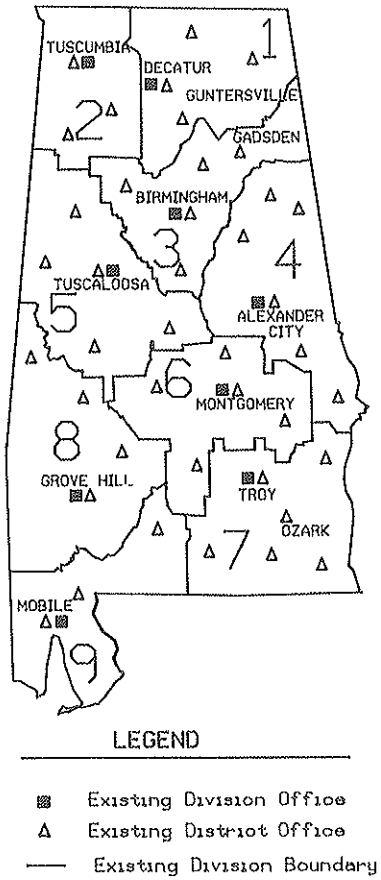



**FIGURE 1 Current district and division office locations.**

tenance costs. Second, the research found in both instances that the most efficient mode of operation was to close some existing offices. The department's efficiency increased with consolidation of operations through closing of offices. The key in minimizing maintenance expenses was consolidation, not expansion, of field offices.

## INTERVIEWS WITH DEPARTMENT MANAGERS

The research staff conducted a series of interviews with key managers to gain insight into the factors that were felt be be important in locating field offices. Twenty individuals were identified as prospective interviewees. Discussions were conducted with all nine division engineers, five of the department's most-senior district engineers (all from urban locations), and six high-level managers from the central office.

The managers were keenly aware of the need for efficiency in field maintenance operations. The topic mentioned most frequently during the discussions was minimization of travel time for employees, which translated into increased work time at the job site. The second most frequently mentioned was minimizing the creation of new field offices to minimize overhead expenses and conserve maintenance monies. Virtually all managers were aware of previous studies that targeted travel time and closing of offices.

Several managers indicated that division or district boundaries should follow county lines. When a county was divided between two districts or divisions, highway managers experienced difficulties with local politicians. It seemed that the politicians did not always know which highway manager controlled which roads, and they consequently became frustrated.

Division engineers expressed the opinion that several divisions were already too small, and that their managers had difficulty in fully using division-wide crews. They were concerned that adding a new division would further decrease the size of adjacent divisions and seriously diminish efficiency.

There were few clear thoughts among the division engineers about how to correlate growth trends in population, vehicle travel, economic development, and other factors with the need for new division offices. One important consideration was present in almost every interview. This was a strong concern for the human aspects of closing or moving division offices and putting people out of work. The managers expressed support for such actions only if they represented the best long-term interests of the department, and urged that these changes not be taken lightly.

## ANALYSIS OF OTHER STATES

A survey was conducted of a dozen southeastern state DOTs. The survey was conducted for two reasons: (a) to review the size and field office configurations of other DOTs for comparative purposes, and (b) to determine if any state had developed a model for selecting locations for division offices. The interviews were conducted by telephone. Discussions were held with the chief engineer, maintenance engineer, and other knowledgeable management officials, and requests were made that the discussion be confirmed with written materials following the telephone conversations. Nine of the states in the survey provided these written materials.

No state had developed a successful, quantifiable methodology for measuring the need for new division offices or for determining the best locations for division offices. Almost without exception, managers contacted during the telephone survey expressed a desire for such a tool.

A comparison of DOT configurations from state to state yielded useful results. Examples are shown in Figures 2 and 3. One useful piece of information shown in Figure 2 is the number of lane-miles of roadway per division. In this factor,

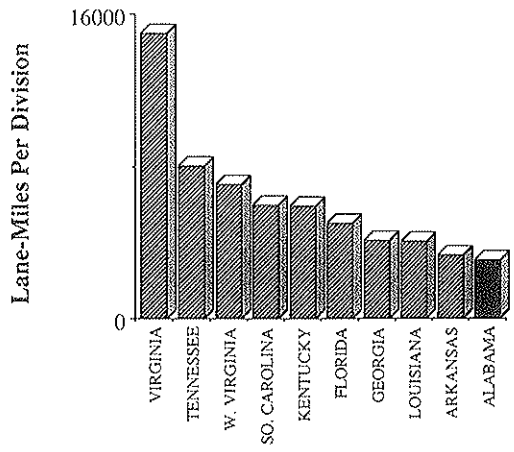



**FIGURE 2 Lane-miles per division office.**

Alabama was found to rank last. Because the department had fewer miles per division than other southeastern states, its divisions had to be smaller than those of other states. Alabama was already subdivided into smaller units than other state DOTs in the southeast. This suggested that any decision to add additional division offices had to approached carefully. Otherwise, Alabama divisions could become too small for efficient operations. Figure 3 shows similar information regarding maintenance funds. Because Alabama ranked last among the surveyed states, further subdivision would diminish, not increase, maintenance capability.

Figure 4 showed a familiar pattern. The amount of vehicle mileage per division was only relatively low in Alabama, primarily because the department had small divisions with a limited number of miles of state route in each. However, because Alabama was last in miles of road per division but above that level for vehicle miles of travel (VMT), Alabama roads were carrying more traffic per mile than some sister states.

In summary, two important facts emerged from this portion of the research. First, no other states had developed a way to predict the need for, or the optimum location of, division offices. Second, extreme care had to be used before creating any additional division offices, because Alabama divisions were already smaller than those of all other southeastern states.

## COST FOR NEW OFFICES

The creation of new district or division offices would require the one-time expenditure of funds for capital development, plus creation of a continuing annual cost for the salaries of administrative staff to run the new office. Department accounting records were screened to determine the costs associated with new division or new district offices. The values (in 1989 dollars) are summarized as follows:

| *Item* | *Amount (1989$)* |
|---|---|
| Construction of new division office | 3,948,020 |
| Annual personnel expenses for new division | 1,750,731 |
| Construction of new district office | 1,414,539 |
| Annual personnel expenses for new district | 190,000 |

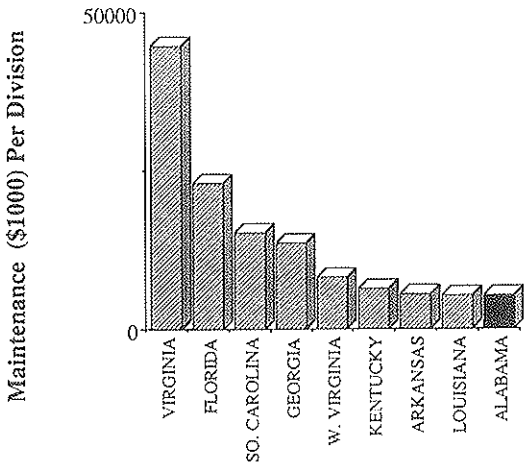



**FIGURE 3 Maintenance funds per division office.**

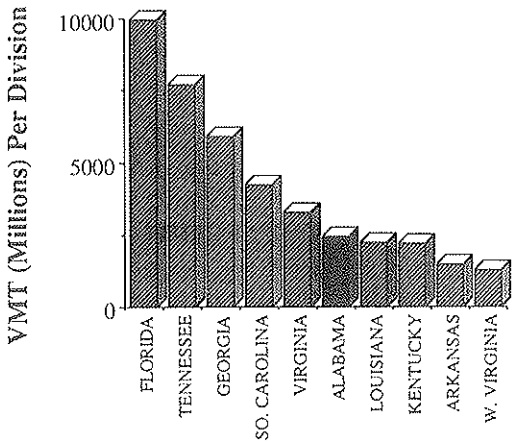



**FIGURE 4 Million VMT per division office.**

The capital costs included land, design fees, office building, office equipment, rolling stock, warehouse, shop, soil testing lab, gashouse, and assembly area. These cost values represented the department's actual expenses the last time that such facilities were constructed, adjusted for subsequent inflation. The most recent division and district offices were both small, so cost values were thought to be conservative, but appropriate for this study.

New district offices absorb some of the maintenance responsibilities of surrounding districts. Most of the employees of new district offices are thus transferred from these same offices, so the increase in payroll is small. Division employees are primarily administrative in nature and cannot be transferred from other divisions. A new division office requires a new staff of administrators, with a large (new) payroll.

## COMPARISON OF DIVISION CHARACTERISTICS

The research staff tabulated and compared several characteristics of division offices to identify parameters that might be used to predict the efficiency of field office operation. The research staff also wished to find the normal range for these parameters. The staff examined lane-miles of highway, centerline miles, vehicle-miles of travel, population, population density, maintenance costs, and economic factors. Ideally, these characteristics would be balanced between divisions; however, real-world constraints often prevent such a balance.

One of the most pertinent findings of the review was that maintenance expenses were closely related to lane mileage. The comparison is shown in Figure 5. The strong relationship between lane mileage and maintenance cost in each division was confirmed through a statistical analysis. A regression model was used to predict maintenance costs based on lane-miles per division, with strong measures of effectiveness. The $R^2$ value was 0.80 and the standard error of estimate was 231 for the following equation:

$$\text{Maintenance (in \$1,000)} = -205.1 + 1.926\ (\text{number of lane-miles})$$

Further examination of characteristics indicated that population (see Figure 6) and vehicle mileage varied widely from

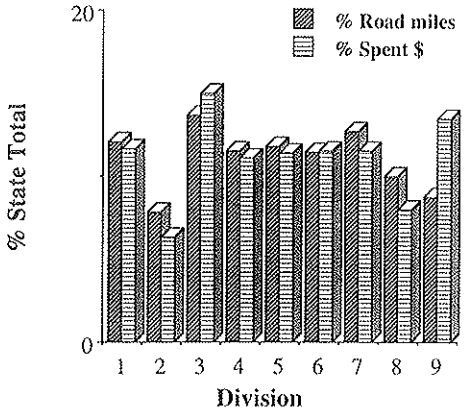



**FIGURE 5 Maintenance funds versus centerline miles.**

division to division, and that certain divisions seemed out of balance. In particular, Divisions 2 and 8 had much lower population and mileage values, whereas Division 3 had more than its share of the various parameters.

Many characteristics were tabulated, examined in this manner, then tested statistically to identify significant patterns of variations. These variations from division to division provided the initial clues to possible changes that might provide more efficient operations.

## *p*-MEDIAN STUDY

The *p*-median statistical modeling process was used to identify relationships between divisions. A brief description of the *p*-median portion of the research project was published by Turner et al. (*1*). This model determined the optimum number and locations for field offices using surrogate measures, given a fixed number of field offices and the distances between them.

The model worked by assuming trial locations of district offices, and calculating the total travel from each node in the transportation system to the closest district office. Once the amount of travel was calculated, trial locations of division offices were established and the total travel between district and division offices were calculated. The model continued trying different district and division office locations until it found the most efficient sites (for which total travel was minimized).

More than 60 scenarios were modeled using the *p*-median technique. For example, one scenario might have 9 division offices, and the next scenario might have 10 division offices. The model was adjusted between scenarios by using different weighting parameters (miles of pavement, population, economic factors, etc.) to portray the attractiveness between districts and divisions.

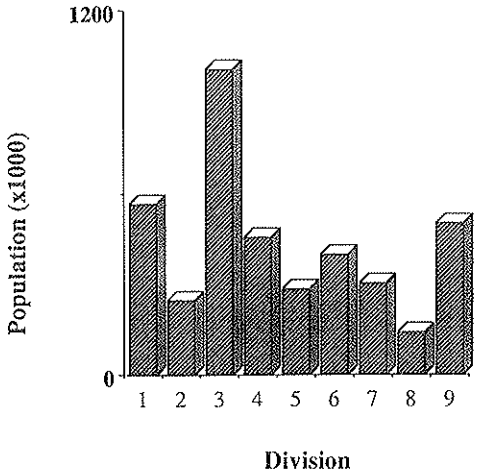



**FIGURE 6 Population per division.**

### Number of Division Offices

The *p*-median model calculated an objective value each time it was run. The objective value was an approximation of total travel time. By examining objective values for different scenarios, the best number of division offices was studied. Figure 7 shows a plot of marginal changes in objective value as the number of division offices increased. The figure indicates that for more than seven division offices there was little increase in efficiency. In other words, when the department increased from seven to nine divisions, there was only a small increase in travel efficiency because of diminishing returns (i.e., the curve on Figure 6 was getting flatter). If the department was to have more than nine offices in the future, there would be little real increase in efficiency of operation. By adding a 10th division office (and increasing overhead costs by 11.1 percent), a gain of only about 2.3 percent would be experienced in travel efficiency. This value suggested that the existing level of nine division offices was a reasonable maximum.

### Location of Division Offices

Several scenarios were evaluated to determine the appropriateness of the location of existing offices and the potential for placing new offices at other locations. The model suggested that five existing offices (located in Mobile, Montgomery, Birmingham, Tuscaloosa, and Tuscumbia) were extremely well placed. It suggested that the offices in Division 1 (Decatur), Division 4 (Alexander City), Division 7 (Troy), and Division 8 (Grove Hill) might be more effective if moved to

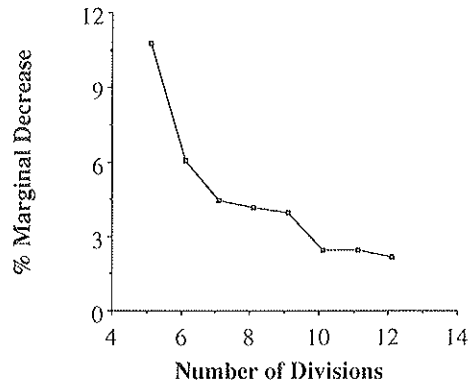



**FIGURE 7 Marginal change in objective value.**

new locations. Minor improvements could be experienced from relocating Divisions 4 and 8, with more substantial improvements from relocation of Division 1 (from Decatur to Guntersville) and Division 7 (from Troy to Ozark). The model also suggested that an east Alabama city (Gadsden) might be considered for a new office. The locations of these cities are shown on Figure 1.

The *p*-median study was not an absolute predictor of the efficiency of various office configurations because it only examined travel distance, and it used surrogate measures of effectiveness. The model was used because it gave good indications of potential efficiencies, which could be confirmed through other models.

## GEOGRAPHICAL INFORMATION SYSTEM LABORATORY MODEL

The researchers used the unique features of the University of Alabama Geography Department's geographical information system (GIS) laboratory to prepare an additional, more-specific model. This computing system was used because it was identical to the computer-aided design and drafting system operated by the Alabama Highway Department.

The GIS system allowed the construction of a model based on a graphic component (map) and an associated data base (characteristics associated with the map). In this case, the map consisted of the state roadway system, and the data base consisted of characteristics associated with it. The characteristics were type of roadway, lane-miles, traffic volumes, maintenance costs, and similar parameters.

A unique feature of the specific model allowed a fence to be drawn around any area on the digitized map. The computer would then calculate the total travel to reach a central office from each roadway segment within that fenced area. When divisions were fenced, an accessibility factor could be calculated by the computer. By moving the fences or by selecting different theoretical locations for division offices, multiple accessibility factors could be calculated and compared. The majority of the research project was used in formulating scenarios and calculating accessibility factors using the travel-specific model.

The model was used for extended investigation of about 20 different scenarios. Four of these were of primary interest to the study. They included

1. Relocate the Division 1 office from Decatur 50 mi southeastward to Guntersville (Scenario D in Figure 8)
2. Relocate the Division 1 office to Guntersville, and realign districts within Divisions 1, 2, and 3 to achieve a better balance (Scenarior G in Figure 8)
3. Relocate the Division 7 office from Troy 30 mi southeastward to Ozark (part of Scenario H in Figure 8)
4. Create a Division 10 in Gadsden, an eastern Alabama city (Scenario I in Figure 8)

The changes in accessibility factors for each of these scenarios were calculated and examined, and are presented in Table 1.

In addition to the four primary scenarios, Scenario L has been listed to illustrate the maximum savings in travel if all offices were moved to optimum locations. This last option was obviously cost-prohibitive because it would require the relocation of many offices and the construction of many new offices.

## ECONOMIC ANALYSIS

The research staff identified several factors that might affect the cost-effectiveness of new or relocated division offices. These included the following types of items:

- Expenses

1. One-time cost of new grounds, facilities, and equipment;
2. Continuing cost of new employee salaries;
3. Diminished local economy caused by lost salaries resulting from the closing or relocation of a division office;

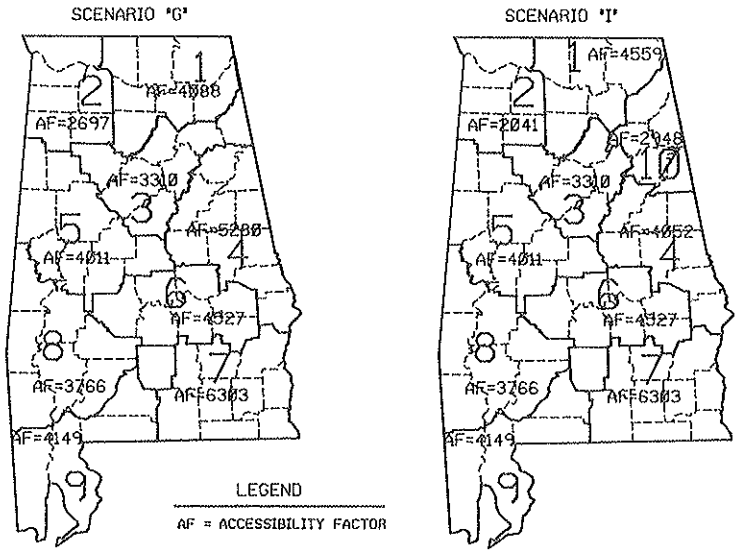



**FIGURE 8 Sample scenarios.**

4. Diminished local economy caused by lost materials and services purchases resulting from the closing or relocation of a division office;
5. Lost jobs for employees whose jobs end, and who do not transfer to a new or revised division;
6. Increased commuting costs for employees who transfer; and
7. Highway maintenance costs that often increased faster than the general rate of inflation.

- Benefits

1. Decreased travel costs,
2. Increased work efficiency for existing employees,
3. Salvage value of closed or relocated facilities,
4. Increased local economy caused by new salaries resulting from the opening or relocation of a division office,
5. Increased local economy caused by materials and services purchases resulting from the opening or relocation of a division office, and
6. Elimination of planned maintenance costs for existing buildings.

### Data Sources

The research staff used modeling studies to assess the changes in travel efficiencies. Data regarding the costs of new facilities were gathered from the fiscal files of the department (as reviewed previously). The Maintenance Bureau and Accounting Bureau supplied breakdowns of typical materials, travel, and salary expenses. Division engineers and central office managers supplied salvage values for facilities that might be sold or relocated. Division engineers supplied demographic information regarding their employees. This helped in the estimation of the number of employees that might relocate to a new office versus those that would give up their jobs rather than relocate. The division engineers also provided trip log summaries for vehicle and equipment use. Finally, the University's Center for Business and Economic Research provided background information regarding interest rates, analysis techniques, and subjective criteria relating to economic impacts of openings and closings of major facilities.

All of these economic data were applied to the scenarios that had the most promise. The results are discussed in the following paragraphs.

### Scenario D Economic Analysis

Scenario D had high relative savings in travel and relatively low costs. New facilities would be required, but salvage of existing facilities would diminish this cost. The losses to the local economy in Decatur would be offset by the gains in the local economy in Guntersville. Major portions of the existing work force could be expected to relocate to the new site.

TABLE 1 ACCESSIBILITY FACTORS FOR SELECTED SCENARIOS

| SCENARIO | DESCRIPTION | ACCESSIBILITY FACTORS FOR AFFECTED AREAS BEFORE | AFTER | CHANGE |
|---|---|---|---|---|
| "D" | Move Division 1 to Guntersville | 6,668 | 4,655 | -30.2% |
| "G" | Move Division 1 to Guntersville, Realign Divisions: | | | |
| | Division 1 | 6,668 | 4,088 | |
| | Division 2 | 2,041 | 2,697 | |
| | Division 3 | 5,245 | 3,310 | |
| | | 13,954 | 10,095 | -27.7% |
| "H" | Move Division 7 to Ozark | 6,303 | 5,389 | -14.5% |
| "I" | Best Scenario for a New Division at Gadsden | | | |
| | Division 10 | 0 | 2,948 | |
| | Division 1 | 6,668 | 4,559 | |
| | Division 3 | 5,245 | 3,310 | |
| | Division 4 | 5,280 | 4,052 | |
| | | 17,193 | 14,869 | -13.5% |
| "L" | Maximum Efficiency, Many Changes in Current Division and District Office Locations | 41,190 | 37,217 | -11.4% |

These factors were more favorable than in other scenarios. The research staff determined that moving the Division 1 office to Guntersville would have about a 10- to 12-year payback period.

There were 102 employee positions assigned to Division 1. Approximately 37 of these positions could be reassigned to other department offices in the Decatur vicinity, and about 65 positions could be transferred to the relocated office. On the basis of the geographic distribution of homes of existing employees, about one-third to one-half of the employees in the 65 positions were predicted to relocate to Guntersville (a loss of approximately 30 to 45 jobs).

### Scenario G Analysis

This scenario involved moving the Division 1 office to Guntersville, then realigning divisions. It improved the travel efficiency in two divisions, but decreased it in a third division. The net change was a better accessibility balance and a substantial travel savings. The costs and assumptions for this analysis were identical to those of Scenario D. The payback period for this scenario was 5 to 7 years. Division office personnel would experience the same loss of jobs as in Scenario D.

### Scenario H Analysis

The shift of the Division 7 office (Scenario H) from Troy to Ozark shared many of the favorable characteristics of Scenario D; however, the much smaller savings in travel efficiency made it economically unattractive.

### Scenario I Analysis

The higher one-time and continuing costs and the smaller total savings in travel caused the creation of a new Division 10 office at Gadsden to be economically nonfeasible.

## FINDINGS

While conducting the multiple studies associated with this research project, the University developed the following findings:

1. The researchers could find no existing methodology that clearly identified the need for a new division office or the optimum location for such an office.
2. Historical data showed that the department periodically adjusted the number of division offices or the location of division boundaries. Three such moves were made in the last 25 years.
3. The department had previously conducted two statistical studies to determine the optimum locations of district offices within given divisions. These district office studies found that the key variable to optimize was employee travel time to reach the job site. The cost of adding or deleting district offices was balanced against decreased or increased travel time in optimizing total expenses.
4. In both previous district office studies, optimum results were found by decreasing the number of field offices. Closing offices and consolidating existing operations minimized maintenance expenses without jeopardizing level of service.
5. Interviews with department managers indicated that they felt employee travel time would be the overwhelming consideration in establishing new division offices, or in relocating existing division offices.
6. Field office managers discouraged the opening of new field offices because they would consume overhead funding and reduce available maintenance monies.
7. Department field managers discouraged division or district boundaries that did not follow county lines.
8. Department managers could not identify any other single factor that they felt might influence the future division office location problem.
9. In comparison to other southeastern states, Alabama had the lowest number of lane-miles per division, the lowest amount of maintenance funding per division, and almost the lowest VMT per division. These factors all indicated that the Alabama Highway Department was already highly subdivided. Further subdivision would create a larger overhead expense per mile of road than other southeastern states. This suggested that divisions should not be added.
10. Transportation agency managers in other states were not aware of any deterministic method to justify a new division office.
11. A review of department records indicated the following minimum costs for adding new offices.

| *Item* | *Minimum Cost ($million)* |
|---|---|
| Construction of new division office | 3.95 |
| Division annual personnel cost | 1.75 |
| Construction of new district office | 1.41 |
| District annual personnel cost | 0.19 |

The costs were used in the economic analysis of potential new offices.

12. In comparing characteristics of the existing divisions, maintenance expenses were noted to be closely related to lane mileage within a division. This relationship was found to be strong, as confirmed by statistical testing.
13. The study of division office characteristics showed that population and lane mileage varied widely from division to division.
14. When characteristics of divisions were compared, Divisions 2 and 8 had less population and less lane mileage than other divisions, whereas Division 3 had excess population and lane mileage.
15. A *p*-median study was used to estimate the best locations for division offices using travel distances and surrogate measures of attractiveness between the offices.
16. The *p*-median technique indicated that travel efficiency increases from adding new division offices dropped sharply as the department went from seven to nine offices.
17. Any division offices added to the existing configuration would produce an extremely small improvement in efficiency of department travel. The marginal change would be a 2.3 percent increase in travel efficiencies for an 11.1 percent increase in overhead.
18. The *p*-median model showed that the department's current division offices in Birmingham, Tuscumbia, Tuscaloosa,

Montgomery, and Mobile were well placed. The offices in Decatur, Alexander City, Grove Hill, and Troy might be candidates for relocation, especially the Decature and Troy sites. In addition, the model indicated that a new division office might be considered for Gadsden.

19. A travel-specific computer model was prepared to provide a more direct measurement of the efficiency of travel between office locations.

20. More than 20 scenarios were analyzed on the travel-specific model.

21. The model indicated the following general changes in accessibility for changes to the current division office configuration:

| *Scenario* | *Accessibility Change in Affected Division* (%) |
|---|---|
| Move Division 1 to Guntersville | −30.2 |
| Move Division 1 to Guntersville, realign divisions | −27.7 |
| Move Division 7 to Ozark | −14.5 |
| Create new Division 10 | −13.5 |
| Optimum location of all offices | −11.4 |

22. Economic analysis suggested that moving the Division 1 office to Guntersville was economically feasible, with about a 10- to 12-year payback period. About 37 current employees could be reassigned to other department positions in the Decatur vicinity, and approximately 65 positions could be transferred to Guntersville. One-third to one-half of the existing employees in these 65 positions could be expected to transfer to the new location, resulting in a loss of approximately 30 to 45 jobs for current employees.

23. The most economically feasible scenarios involved moving the Division 1 office to Guntersville, moving three counties from Division 1 to Division 2, and moving two counties from Division 3 to Division 1 (Scenario G). There would be about a 5- to 7-year payback for these changes. Job losses would be the same as expected for Scenario D.

24. Neither moving the Division 7 office to Ozark nor creating a new Division 10 in Gadsden was found to be economically feasible.

25. No other changes in division office locations were found to be economically feasible.

## RESULTS OF THE STUDY

After receiving the report associated with this study, the Alabama Highway Department initiated plans to relocate its existing Division 1 office from Decatur, Alabama, approximately 50 miles southeastward to the vicinity of Guntersville, Alabama. The relocation process will take place slowly enough to give the department time to properly plan the move, and to give the affected employees the opportunity to soften the impact of the abrupt change in the location of the division office.

## ACKNOWLEDGMENTS

Funding for this project was provided by the Alabama Highway Department. The authors gratefully acknowledge the many employees of the department who provided data, technical guidance, and encouragement throughout the study. The authors are also indebted to their associates at the University of Alabama who helped conduct this research. John M. Harlin and Jami Nettles of the Department of Geography, William D. Gunther of the Center for Business and Economic Research, and Jerry R. Weaver of the Department of Management Science made invaluable technical contributions. Special thanks are extended to Judi Williams, Teresa Sikes, Nell Vice, Kelly Hall, Zachary Hall, and Julie Wheeler of the Civil Engineering Department for their support.

## REFERENCE

1. D. S. Turner, J. R. Weaver, and W. D. Gunther. Assessing Highway Field Maintenance Office Locations by the *p*-Median Model. In *Transportation Research Record 1268*, TRB, National Research Council, Washington, D.C., 1990, pp. 156–163.

*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Development of a Microcomputer-Based System for Traffic Signal Maintenance Records

DARRELL W. BORCHARDT, STEVEN Z. LEVINE, AND DARRELL D. VANOVER

The Houston district office of the Texas State Department of Highways and Public Transportation is presently responsible for the operation and maintenance of approximately 1,000 traffic signals within a six-county area. This number is expected to double within the next 5 years. The Texas Transportation Institute has begun the development of a microcomputer-based method of recording and analyzing traffic signal system maintenance records. Still under development, this system will eventually replace the manual method now used by signal maintenance personnel. The system will be based on the use of portable microcomputers and expert system software. An overview and benefits of the system as it is being developed are presented.

On completion of maintenance activities within the Houston district office of the Texas State Department of Highways and Public Transportation (SDHPT), the technician or engineer must fill out a standardized form (Figure 1). Although it is intended for this form to be filled out on completion of the activity at the specific signal location, many of these forms are not filled out until the end of the day. Therefore, some of the required items may not be accurately recorded. Accuracy in recording arrival and departure times is especially critical when responding to emergency maintenance activities. These signal maintenance records are used in tort liability, civil, and criminal cases involving the state-maintained traffic signals. (In 1989, the Houston district was involved in 28 cases. By law, such records must be maintained for 7 years.) The development of a microcomputer-based method of recording this information in the field would greatly enhance the accuracy of the records. The data would also be available at the end of each day. The developed system would save the district time and money, and would result in an enhanced and modern maintenance records system. Accurate maintenance records would be available in cases where such information is required in litigation.

This system is in its initial steps of development by the Texas Transportation Institute (TTI). An overview of the existing system as presently used and of the proposed system under development is presented. This research was sponsored by the Houston district of the Texas SDHPT under a study titled *Planning, Design, and Operation of Transportation Facilities in Houston*

D. W. Borchardt, Texas Transportation Institute, 701 North Post Oak, Suite 430, Houston, Tex. 77024. S. Z. Levine and D. D. Vanover, Texas State Department of Highways and Public Transportation, 7721 Washington Avenue, Houston, Tex. 77007.

## EXISTING SYSTEM

### Frequency of Maintenance Calls

A data base has been developed by the Houston district for all signal maintenance activities completed since 1987. The frequency of total yearly calls has remained almost constant since that time period.

| *Year* | *Total Calls* | *Number of Signals Maintained by District* | *Calls per Location per Year* |
|---|---|---|---|
| 1987 | 8,033 | 732 | 10.97 |
| 1988 | 8,120 | 763 | 10.64 |
| 1989 | 8,621 | 984 | 8.76 |
| 1990 (6 months) | 4,444 | 1,001 | 8.88 |

In 1989, SDHPT policy changed and the Houston district assumed maintenance of and operation of 25 percent more signals, which are located in cities between 15 and 50,000 population.

These data result in an average of 9.67 maintenance calls per location each year. Although each traffic signal is visited at least once each year for annual maintenance, it is not uncommon for problem locations to be called on two or more times each month. In a scenario of double the present number of signal locations and average number of visits per location, which could occur in 5 years, the management of a manual record system would become impractical. This assumed scenario results in an estimated 18,000 maintenance forms being completed each year.

### Present Procedures

Each of the existing traffic signal maintenance records follows a manual procedure that is labor-intensive and time consuming. The steps are outlined as follows:

Step 1. On completion of the required maintenance activity, a traffic signal maintenance report is filled out in the field by the technicians or the engineer. The form is filled out as completely as possible with all available information.

Step 2. The completed forms are then returned to the crew's shop at the end of the day. Any remaining information must then be added to the report. This procedure normally involves

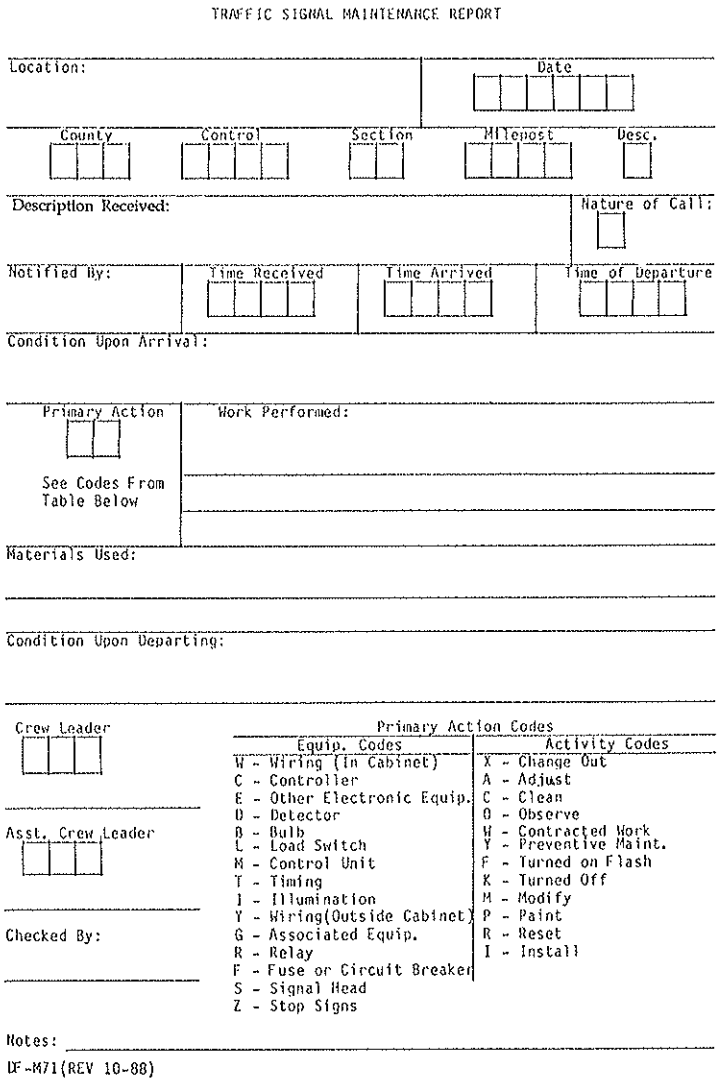



**FIGURE 1 Existing signal maintenance report form.**

matching the location (e.g., IH-10 at Mason Road) with its corresponding county, control, section, and milepost identification (e.g., 102-271-6-3.75). This location code is used as the basis for identifying the signal and is used in matching the data item with those of other data bases.

Step 3. The completed forms are then forwarded to the district office at selected time periods, usually weekly. The contents of each form are then entered into a data base using dBASE III-Plus. Although data entry is performed as accurately as possible, input errors are likely to occur. Therefore, a corresponding level of inaccuracy between the resulting data base and the actual reports exists.

Step 4. The maintenance reports then must be filed in a permanent record system. A separate file for each signal location is kept at the district office. The filing system is in numerical order according to the location code. The reports are then sorted manually according to numerical order (to make filing easier) and placed in the proper file. They must be maintained for 7 years according to the tort liability laws in Texas. An output of this process that recently began is to provide the identification of problem locations. The updated data base file is provided to TTI once all reports for the previous month have been entered. The file is then converted to standard data format (ASCII) to allow access by statistical programs. PC-SAS is used to provide any analyses as requested by the district. The present analysis provides a listing of all locations that required four or more maintenance activities during a consecutive 2-month period (Figure 2). Although the listing only provides limited information, problem locations can be identified. Signal personnel, who have responded to the same locations on multiple occasions, can also be identified.

This four-step process is somewhat time consuming and requires a lot of manual effort. The potential for errors and possible loss of valuable data increases. The turn-around time from filling out the reports in the field to entering into the data base takes approximately 14 days. If the system were redesigned to use portable microcomputers in the field, an updated and current system on a next-day basis could be provided.

DISTRICT 12 SIGNAL MAINTENANCE RECORD LISTING
4 OR MORE CALLS FOR MAY AND JUNE 1990 ONLY

COUNTY=HARRIS CONTROL=502 SECTION=1 MILEPOST=15.38

| DATE1 | MAJOR | | MINOR | NOTIFIED | ARRIVED | DEPARTED | DESC | NATURE | ECODE | ACTCODE | NAME1 | NAME2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tue, May 1, 90 | SH | 225 | SENS | 10.18 | 10.75 | 15.75 | GENOA | EMERGENCY | ASSOCIATED EQUIP. | CHANGE OUT | HJR | JHN |
| Wed, May 2, 90 | SH | 225 | SENS | 8.00 | 9.00 | 12.00 | GENOA | ROUTINE | ASSOCIATED EQUIP. | ADJUST | HJR | JHN |
| Thu, May 3, 90 | SH | 225 | SENS | 8.00 | 11.00 | 11.38 | GENOA | ROUTINE | ASSOCIATED EQUIP. | CHANGE OUT | HJR | JHN |
| Fri, May 4, 90 | SH | 225 | SENS | 6.50 | 7.50 | 8.25 | GENOA | EMERGENCY | OTHER ELECTRONIC EQUIP. | RESET | DRB | |
| Fri, May 4, 90 | SH | 225 | SENS | 8.00 | 10.10 | 12.27 | GENOA | ROUTINE | ILLUMINATION | CHANGE OUT | HJR | JHN |

COUNTY=HARRIS CONTROL=508 SECTION=1 MILEPOST=40.4

| DATE1 | MAJOR | | MINOR | NOTIFIED | ARRIVED | DEPARTED | DESC | NATURE | ECODE | ACTCODE | NAME1 | NAME2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tue, May 1, 90 | IH | 10 | BELTWAY 8 | 13.00 | 13.42 | 15.50 | DISTRICT | EMERGENCY | SIGNAL HEAD | OBSERVE | WAR | DER |
| Tue, May 8, 90 | IH | 10 | BELTWAY 8 | 14.65 | 14.67 | 15.33 | HUMBLE | ROUTINE | CONTROLLER | OBSERVE | B D | |
| Fri, Jun 1, 90 | IH | 10 | BELTWAY 8 | 12.82 | 13.50 | 14.58 | GENOA | EMERGENCY | ASSOCIATED EQUIP. | OBSERVE | HJR | |
| Fri, Jun 1, 90 | IH | 10 | BELTWAY 8 | 13.00 | 16.83 | 22.67 | DISTRICT | EMERGENCY | CONTROLLER | CHANGE OUT | WAR | DER |
| Mon, Jun 25, 90 | IH | 10 | BELTWAY 8 | 8.00 | 9.25 | 15.25 | HUMBLE | ROUTINE | CONTROLLER | ADJUST | JDM | WCH |
| Mon, Jun 25, 90 | IH | 10 | BELTWAY 8 | 8.00 | 9.33 | 12.08 | HUMBLE | ROUTINE | WIRING (IN CABINET) | INSTALL | B D | |
| Mon, Jun 25, 90 | IH | 10 | BELTWAY 8 | 13.83 | 14.17 | 15.25 | HUMBLE | ROUTINE | CONTROLLER | OBSERVE | B D | |
| Tue, Jun 26, 90 | IH | 10 | BELTWAY 8 | 8.00 | 9.08 | 12.00 | HUMBLE | ROUTINE | CONTROLLER | ADJUST | JDM | WCH |
| Thu, Jun 28, 90 | IH | 10 | BELTWAY 8 | 10.00 | 15.17 | 16.00 | GENOA | ROUTINE | CONTROLLER | ADJUST | DRS | |

COUNTY=HARRIS CONTROL=720 SECTION=3 MILEPOST=13.41

| DATE1 | MAJOR | | MINOR | NOTIFIED | ARRIVED | DEPARTED | DESC | NATURE | ECODE | ACTCODE | NAME1 | NAME2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tue, May 15, 90 | SH | 249 | FM 1960 | 8.00 | 10.75 | 12.80 | HUMBLE | ROUTINE | DETECTOR | OBSERVE | RAD | |
| Thu, May 17, 90 | SH | 249 | FM 1960 | 8.00 | 13.00 | 14.00 | HUMBLE | ROUTINE | DETECTOR | OBSERVE | RAD | |
| Mon, May 21, 90 | SH | 249 | FM 1960 | 8.00 | 10.00 | 13.00 | HUMBLE | ROUTINE | WIRING (OUTSIDE CABINET) | ADJUST | ORS | |
| Mon, May 21, 90 | SH | 249 | FM 1960 | 8.00 | 10.83 | 12.93 | HUMBLE | ROUTINE | DETECTOR | ADJUST | RAD | |
| Wed, Jun 13, 90 | SH | 249 | FM 1960 | 13.50 | 14.83 | 15.25 | GENOA | EMERGENCY | CONTROLLER | OBSERVE | ORS | |
| Fri, Jun 15, 90 | SH | 249 | FM 1960 | . | 10.58 | 10.67 | HUMBLE | SPECIAL | TIMING | ADJUST | NAA | |
| Tue, Jun 26, 90 | SH | 249 | FM 1960 | 8.00 | 10.08 | 10.25 | HUMBLE | EMERGENCY | SIGNAL HEAD | ADJUST | B D | |

**FIGURE 2 Examples of monthly listing of maintenance records.**

## MICROCOMPUTER-BASED SYSTEM

### System Concept

A system is being developed that will use portable microcomputers in the field as the basis for gathering the information presently being handwritten. This system would allow for time and date input as provided by the computer. At the end of each signal crews' shift, the maintenance report data could be forwarded to the district office. How this transmission would be accomplished depends on the capabilities of the portable microcomputer. It would most likely be completed using telephone modem hook-ups from remote locations. Once all data have been loaded on the office microcomputer, supervisors could assess a menu of programs to view and print desired information. This access of the data is much quicker than the present manual method. Daily information on field maintenance activities would also be readily available.

### Selection of Portable Computer

The initial task in designing the system was the selection of a portable microcomputer. It must have design qualities that will allow it to withstand field environmental conditions and a sufficient battery duration between recharges. Approximately 15 companies, with manufacture portable hand-held microcomputers, were contacted. Many of these companies also manufacture standard laptop computers. However, laptops were not considered for use because of power requirements and field environmental conditions. The portability or size of the unit was also a major factor in the selection process. The response from those contacted ranged from receiving literature in the mail to phone calls from sales and technical staff. All specifications were reviewed and one unit was selected to be used for the proposed system. In addition to the unit's physical requirements, the user friendliness and programming capabilities of the unit were primary considerations. The user of these units would be signal technicians and engineers.

The GRiDPad, manufactured by GRiD Systems Corporation, was selected as the portable computer to be used in the field (Figure 3). This unit is unique in that it is lightweight (4.5 lb) and that data may be entered by printing text with its attached electronic pen. In addition to its ability to run MS-DOS-compatible application software, the GRiDPad has the capability of running customized forms and user applications. Data storage is provided by removable RAM/

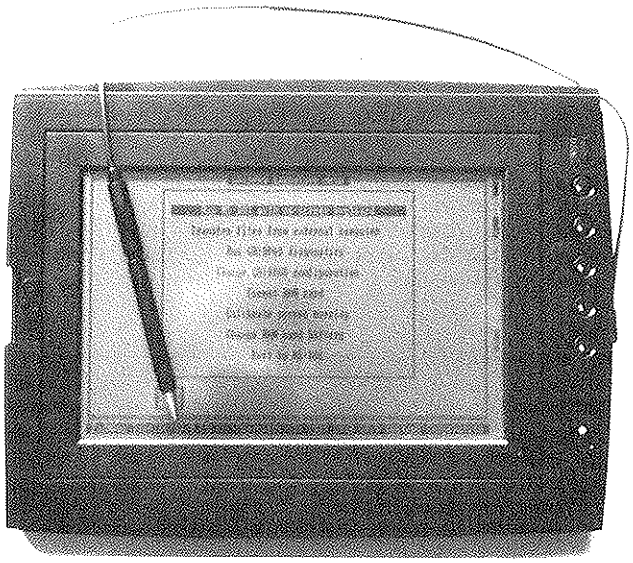


**FIGURE 3 GRiDPad portable computer and attached electronic pen.**

ROM data storage cards. Its serial port can be used to connect bar-code readers, printers, or other microcomputers. No other manufacturer presently markets a unit with these characteristics.

The use of pen-based portable computers, particularly the GRiDPad, was described in an article in *Business Week Magazine* (May 14, 1990). The Southern Pacific Transportation Company stated that billing errors because of sloppy paperwork can be avoided. Officials with the railroad estimate that setup costs for the GRiDPad-based system will be paid off in a maximum of 3 years with savings from error elimination. The Best Foods Baking Group also indicated that it could save $1.5 million annually in reduced billing errors, increased cash flow, and fewer loaves of stale bread with implementation of a GRiDPad system. The article also pointed out that Southern Pacific selected the GRiD unit partially because the unit uses the familiar MS–DOS operating system.

In the case of recording traffic signal maintenance records in the field, the GRiDPad is an ideal unit. Custom entry forms that are similar to those presently completed (Figure 1) can be developed. Therefore, less training of signal technicians on the operation of the system is required for its implementation. Initial purchase costs for software development are estimated at approximately $6,000. This amount includes purchase of all software and connection hardware necessary to develop the input form and a single GRiDPad for field testing. Depending on the amount and number of storage cards required, additional purchases of the GRiDPad only should be less than $3,000 each.

### System Development

The proposed system to be developed for the Houston district is designed for staged implementation. As each less-complicated portion of the overall system is completed, it can be implemented in the field. (This process depends on the availability of sufficient GRiDPads being purchased by the SDHPT for use by all signal maintenance crews.) The staged objectives and the methods by which each will be accomplished are described in the following sections.

The primary goal of the first stage is the development of a data collection form to be used in the field, which would be similar to the one presently used, shown in Figure 1. This process would require most of the computer programming time as well as an understanding of the capabilities of the GRiDPad. All routines for downloading the field data to the host microcomputer would be developed and field tested. The resulting software, when used in the field with the GRiDPad, would increase the efficiency of the manual procedures presently used. Implementation of this stage will eliminate the manual step of matching the location code with the intersection name. The present manual step of entering the maintenance records into a data base (or into some other acceptable format) will be completed using programming techniques. The portable computer will also provide time and date stamping as each activity is recorded in the field.

Another set of PC-based programs will be developed to manage the resulting data base. Each maintenance record will be available for screen viewing and editing of field input errors or those times when the printed information cannot be interpreted by the GRiDPad. Programs will also be developed to provide printed output for permanent filing. These listings will be produced in order according to the county, control, section, and milepoint identification. Other listings will also be provided within the PC-based routines.

The second stage would add a major component to the proposed system that would eliminate the step of inputting the signal location. A bar code reader would be attached to the GRiDPad to automatically identify the signal location. A bar code unique to each intersection would be placed within the signal cabinet and not on the controller, because the controllers are often changed out. This process would eliminate mistakes caused by misidentification of the signal location by field technicians. The bar code concept could later by expanded to include inventory control of all traffic signal equipment. The feasibility of inventory control will be examined at a later date; present emphasis is being placed on only the portion of the project pertaining to maintenance records.

Development of the second stage requires that additional software and equipment be purchased. This equipment, which would allow production of the required bar codes, would be installed within each controller cabinet. The technology for producing bar codes has improved so that labels can be printed using a standard personal computer and printer. Costs of implementing the bar code portion have not yet been estimated. The initial expense would be the purchase of a system (software and hardware) to produce the printed bar code. Minimum costs are expected to be about $750. Second-stage development would begin only after the initial system has been successfully field tested.

A third stage has also been considered that would use these hand-held computers to assist field technicians. This activity involves working with signal controller manufacturers in developing an expert system to assist field technicians in troubleshooting of malfunctioning controllers. The expert system for diagnostics would be maintained on an office PC and accessed using a cellular telephone modem connection.

### Implementation

This system is designed for implementation according to the staged development process. The first stage will be ready for implementation by the Houston district signal maintenance crews in late 1992. This time frame provides approximately 2 years for the system to be developed and field tested, and for the purchase of sufficient GRiDPad computers by the SDHPT. The second stage—use of bar codes for location, time, and data identification purposes—could begin at the same time. Expansion of the bar code system for signal inventory control could be implemented by 1993. Successful use of an expert system for troubleshooting controllers depends on the willingness of manufacturers to develop and assist in the design of such a system.

## BENEFITS OF A MICROCOMPUTER-BASED SYSTEM

The benefits of using the described microcomputer-based system for maintaining traffic signal maintenance records can be expressed in terms of operational and monetary benefits.

In terms of operations, many manual steps, which are presently used to document field activities, can be eliminated. The field information, when forwarded to the host personal computer on a daily basis, can be used as a virtual real time system for determining maintenance needs. After implementation of the second stage, the record system will be almost foolproof because of the computer-based date and time-stamping routines.

Because of the possibility of litigation, the monetary benefits from the improved accuracy cannot be adequately determined. However, any lawsuit has a cap of $280,000 potential liability to the state. Because 10 hr per week are devoted to inputting the signal record into a data base and filing of the handwritten copy, the monetary benefits of the system far outweigh the cost of implementation. After an appropriate training period, the time for the signal technician to record the information on the GRiDPad should be equal to or less than the present time to record this information in a written format.

As mentioned earlier, a change in departmental policy resulted in a 25 percent increase in the number of traffic signals maintained by the Houston district. Because the number of signal maintenance personnel was not increased, the time spent at each location became critical. Development of expert system software to assist field technicians in troubleshooting of malfunctioning controllers may reduce the repair time at each location. Each maintenance crew will then be able to provide service to more traffic signals each day at reduced stress levels of personnel.

*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Use of Field Data in Calculating Cost of Earth Road Maintenance

MAGDY ABDELRAHMAN AND ESSAM SHARAF

Earth roads will continue to form an important portion of the local road network in Egypt. The objective of the proposed study was to provide an index for road assessment and to develop a cost model in which associated repair costs could be determined. Decision making at the network level is concerned with the selection of individual sections of roads for maintenance in accordance with the availability of funds and, therefore, a matter of the assignment of priorities related to needs. For an effective maintenance program, a specified data element about the road conditions is required. Information that can be directly used in maintenance management systems must be provided. The most important step, however, is operation of its own condition assessment subsystem. The cost of periodic and routine maintenance required to renew road sections was calculated. The relation between road condition and repair cost was quantified. Also, the results indicated that environmental factors had a significant effect on road condition and, consequently, on maintenance and repair costs.

A unique aspect of pavement engineering in developing countries that has no parallel in most industrial countries is the extent to which earth roads contribute to the national road network. As a developing country, Egypt operates a road network consisting of about 19,000 km of paved roads and about 14,000 km of earth roads.

In Egypt, earth roads play a vital role in the economic and social life. The development of agriculture, the provision of health services, and many forms of communication in rural agricultural regions, all heavily depend on transport facilities. Although rail and water facilities may play important roles in certain areas, a dominant need is for a road system that can provide a reliable and yet relatively inexpensive means for the movement of people and goods, and that is what earth roads can do.

A pavement management system involves managing pavements using a maintenance managment subsystem. The prime purpose of the maintenance phase in a pavement management system is to determine the costs associated with providing various levels of serviceability for any given pavement. For effective maintenance cost determination, specific data about road conditions are required. It is important that the information provided can be directly used in a maintenance management system. For the earth road, the most important step is to develop a unique condition assessment subsystem.

The earth road condition index (ERCI) is not a new idea. In fact, it was designed to draw on past experience, through maintenance management concepts, about how roads perform and what is observed to analyze these conditions. In this work, the Sharkia Governorate earth road network was selected and the methodology described later was applied on it. The part of the network considered in this study consists of about 190 km (118 mi) covering different environmental conditions.

The development of the ERCI system and its associated distresses are described elsewhere (*1*) with slight modification. The included distresses are mainly selected to express the actual maintenance needs on these roads. An example of a distress is Road Surface Occupancy. Farmers frequently use the roadway to store and prepare imported soil, seeds, and other related materials and equipment. When they do not remove their occupancy in a complete way, it affects the road surface and requires some maintenance work, so it is a distress. The same concept was applied to all other distresses.

## METHOD

The method for rating the condition of earth roads and estimating maintenance costs includes the following steps (*1*,*2*):

- Divide the road network into sections and samples,
- Inspect the sections,
- Calculate ratings, and
- Apply the maintenance cost model on the inspected sections.

### Dividing the Road Network Into Sections and Samples

Before a road network is inspected, it must be divided into branches, sections, and sample units. A complete road length may be considered as a branch and may be divided into homogenous sections. A homogenous section consists of a number of similar samples (*3*). The process is completed by passing the road links before the condition survey.

### Sampling

Inspection of every sample unit in a road section may be necessary if exact quantities are needed for contracting. However, such inspection requires considerable effort, specially if the section is large. Because of the time and effort involved, frequent surveys of a large number of deteriorated sections may be beyond available manpower, funds, and time. However, sampling plans can reduce inspection funds and time considerably and still provide the accuracy required.

M. Abdelrahman, Zagazig University, Zagazig, Egypt. E. Sharaf, Cairo University, Giza, Egypt. Current affiliation: Department of Civil Engineering, Faculty of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia.

| DISTRESSES | LOW SEVERITY DESCRIPTION |
|---|---|
| - FAILURE & EROSION OF ROAD EMBANKMENT | WIDTH UP TO 2 M, DEPTH TO 2.5 M. |
| - FAILURE OF RETAINING WALL | NEED REBUILDING ONLY, NO STONE REPLACEMENT |
| - SURFACE HEIGHT | SURFACE ELEVATES LESS THAN 50 CM, AND DOWN TO THE GROUND LEVEL. |
| - CAMBER | DEPTH AT THE MID-POINT LESS THAN 5 CM. |
| - RUTTING | DEPTH LESS THAN 3 CM. |
| - POTHOLES | DEPTH LESS THAN 10 CM, ONE POTHOLE/ 10 $M^2$. |
| - SURFACE SATURATION | SATURATION DEPTH < 20 CM, LESS THAN HALF THE ROAD WIDTH. |
| - ILLEGAL IRRIGATION STRUCTURE | EXTENT LESS THAN 1.0 M WIDTH. |
| - PLANTS ON ROAD SURFACE | EXTENT LESS THAN 1.0 M WIDTH. |
| - ROAD SURFACE OCCUPANCY | EXTENT LESS THAN 1.0 M WIDTH. |

**FIGURE 1 Summary of low-severity cases.**

### Measurement Frequency

As a network level condition assessment, measurement frequency is arranged to reflect network condition. The number of samples are determined according to homogeneity of assessed road sections. Homogenous sections are divided into equal samples of 100 m each. As a rule, at least 0.1 of the entire section length is considered sufficient to represent the section length (*3*). One random sample of 100 m is considered to represent 1 km and 5,300 ft is considered to represent 1 mi (1.61 km = 5,300 ft) of section length. Additional samples are taken at

- Beginning and end of roads,
- Intersection with other roads,
- Inhabited villages, and
- Any significant change in the surrounding environment.

### Inspecting the Sections

Each homogeneous section is first passed and the sample locations are determined randomly. It is important that each sample unit be identified concisely, so it can be located for additional inspections, comparison with future inspections, maintenance requirements, and random sampling purposes. Sample inspection includes measuring cross-section dimensions to be used in estimating maintenance costs.

### Severity Level

The ERCI system involves two severity levels. The low-severity case (slight case) is shown in Figure 1. Any condition other than low severity will be a severe case.

### Extent Weight

Extent weight is defined as the ratio of deteriorated length to the total sample length.

### Deduct Value

An important item in the calculation of the ERCI is the deduct value. A deduct value is a number from 0 to 100, with 0 indicating that the distress has no impact on road condition, and 100 indicating an extremely serious distress that causes the pavement to fail. In fact, the case of a deduct value of 100 is a theoretical case, and requires an abnormal road condition. The concept of deduct value for any distress is based on the impact of this distress on road condition.

### Calculating Sample Unit ERCI

The procedure for calculating the ERCI for a sample unit involves four steps:

Step 1. Each sample unit is inspected and distress data are recorded on the form, as shown in Figure 2.

Step 2. The deduct values are determined by multiplying distress weight by severity level and extent weight for each recorded distress, as shown in Figure 3.

Step 3. A total deduct value is computed by summing all individual deduct values.

Step 4. The ERCI value is computed using the relation ERCI = 100 − total deduct value.

The ERCI value for the entire road length is the average of sample unit values.

| EARTH ROAD INSPECTION SHEET | | | | | |
|---|---|---|---|---|---|
| ROAD: | DISTRICT: | | | | |
| SAMPLE UNIT: | SUBDISTRICT: | | | | |
| DIRECTION FROM / TO | DATE: | | | | |
| INSPECTOR: | O: OCCASIONAL<br>F: FREQUENT<br>E: EXTENSIVE | | | | |
| DISTRESS | SEVERITY | | EXTENT WEIGHT | | |
| FAILURE & EROSION OF EMBANKMENT | Slight | Sever | O | F | E |
| FAILURE OF RETAINING WALL | Slight | Sever | O | F | E |
| SURFACE LEVEL | Slight | Sever | O | F | E |
| CAMBER | Slight | Sever | O | F | E |
| RUTTING | Slight | Sever | O | F | E |
| POTHOLES | Slight | Sever | O | F | E |
| SURFACE SATURATION | Slight | Sever | O | F | E |
| ILLEGAL IRRIGATION STRUCTURE | Slight | Sever | O | F | E |
| PLANTS ON ROAD SURFACE | Slight | Sever | O | F | E |
| ROAD SURFACE OCCUPANCY | Slight | Sever | O | E | E |
| REMARKS | | | | | |

**FIGURE 2 Earth road inspection sheet.**

### Maintenance Categories for Earth Roads

Involved distresses are arranged to measure the adequacy of three categories. Routine and recurrent maintenance activities include surface condition repair involving camber (road surface crown), potholes, rutting, mowing of plants on road surfaces, and removal of surface occupancy. Periodic maintenance activities include the repair of embankment failure, surface raising by adding soils, repair of failure of retaining wall, and repair of illegal irrigation structures. Urgent maintenance activities are mainly related to surface saturation caused by rains or damages in irrigation and drainage systems of adjacent fields.

### Level of Serviceability Concept for Earth Roads

The concept of earth road level of serviceability is based on two maintenance categories: (a) periodic maintenance applied to repair the generated distresses, and (b) routine maintenance including recurrent activities applied to preserve the surface quality with the desired serviceability. Thus, to provide a certain level of serviceability, two main associated costs are considered, periodic cost and routine maintenance cost. Table 1 indicates the extent that each maintenance category contributes to the total rating scale.

## APPLICABLE MAINTENANCE PRACTICES AND COSTS

On the basis of practical measurements, past experiences about earth road maintenance practices, and a literature survey (*4–6*), the associated maintenance activities were determined. These maintenance activities and their associated costs required for repairing the proposed distresses are presented in Table 2. Periodic maintenance activities were determined as the best available repairs for the distress types in either severity level. Periodic maintenance activities repair the combined effects both of load and environmental factors including embankment soil stability, soil type, and surrounding environment. An example of a periodic maintenance distress is the failure and erosion of road embankment, caused both by traffic loads and by the erosion effects of adjacent streams. For an example of low-severity distress, see Figure 4. The associated repair of this distress is to construct a retaining wall for the high-severity case, or to rabble (place cement-treated rip rap on) the failed-embankment, low-severity case, as shown in Figure 5.

| DISTRESSES | WEIGHTS | SEVERITY WEIGHTS | | EXTENT WEIGHTS | | | DEDUCT POINTS | RATING |
|---|---|---|---|---|---|---|---|---|
| | | SLIGHT | SEVER | OCCASIONAL | FREQUENT | EXTENSIVE | | |
| - FAILURE & EROSION OF ROAD EMBANKMENT | 20 | 0.5 | 1.0 | 0.4 | 0.8 | 1.0 | | 100 VERY GOOD 90 |
| - FAILURE OF RETAINING WALL | 12 | 0.3 | 1.0 | 0.4 | 0.8 | 1.0 | | GOOD 80 |
| - SURFACE HEIGHT | 15 | 0.5 | 1.0 | 0.4 | 0.8 | 1.0 | | FAIR |
| - CAMBER | 8 | 0.5 | 1.0 | 0.4 | 0.8 | 1.0 | | 65 |
| - RUTTING | 8 | 0.5 | 1.0 | 0.4 | 0.8 | 1.0 | | |
| - POTHOLES | 8 | 0.5 | 1.0 | 0.4 | 08 | 1.0 | | POOR |
| - SURFACE SATURATION | 8 | 0.6 | 1.0 | 0.4 | 0.8 | 1.0 | | 40 |
| - ILLEGAL IRRIGATION STRUCTURE | 7 | 0.4 | 1.0 | 0.4 | 0.8 | 1.0 | | |
| - PLANTS ON ROAD SURFACE | 7 | 0.5 | 1.0 | 0.4 | 0.8 | 1.0 | | FAILED |
| - ROAD SURFACE OCCUPANCY | 7 | 0.4 | 1.0 | 0.4 | 0.8 | 1.0 | | ZERO |

TOTAL DEDUCT=

ERCI = 100 - TOTAL DEDUCT=

**FIGURE 3 ERCI calculations and rating scale.**

TABLE 1 MAINTENANCE CATEGORIES

| MAINTENANCE CATEGORIES | INVOLVED DISTRESSES | POINTS | TOTAL POINTS |
|---|---|---|---|
| PERIODIC MAINTENANCE | - FAILURE & EROSION OF ROAD EMBANKMENT | 20 | |
| | - FAILURE OF RETAINING WALL | 12 | |
| | - SURFACE HEIGHT | 15 | |
| | - ILLEGAL IRRIGATION STRUCTURE | 7 | 54 |
| ROUTINE MAINTENANCE INCLUDING RECURRENT MAINTENANCE | - CAMBER | 8 | |
| | - RUTTING | 8 | |
| | - POTHOLES | 8 | |
| | - ROAD SURFACE OCCUPANCY | 7 | |
| | - PLANTS ON ROAD SURFACE | 7 | 38 |
| URGENT MAINTENANCE | - SURFACE SATURATION | 8 | 8 |
| | | TOTAL = | 100 |

Routine maintenance activities are determined on the basis of ride quality and associated repairs required to achieve an acceptable ride quality. An example of a routine maintenance activity is surface blading. The frequency of applying this activity depends on how soon the road surface deteriorates after each successive blading. Surface blading frequency reflects traffic level and surface soil characteristics. In fact, not all routine maintenance activities are included in this study, but only those activities that affect the level of serviceability and the associated maintenance level. Examples of these activities are sign repair, bridge guard rail repair, and routine inspection (a fixed item in routine maintenance activities)(7).

The ability to estimate the costs of various maintenance activities is an important element in an effective maintenance management system (MMS). Costing road maintenance is a critical problem, particularly for local authorities. Generally, because periodic maintenance is bid by private contractors, the costs are usually well known. As most routine maintenance activities are done by the local authorities' own force, the ability to estimate the road maintenance costs depends on maintenance cost accounting or management system in operation at the authority. For this study, determination of item costs was based on recent contracts and studies. The concept of the maintenance cost model was based on provid-

TABLE 2 MAINTENANCE PRACTICES COSTING

| DISTRESS | MAINTENANCE PRACTICES | ITEMS | UNITS | QUANTITY / L. M. | UNIT COST L.E. [d] | TOTAL COST L.E. / L.M. |
|---|---|---|---|---|---|---|
| FAILURE & EROSION OF ROAD EMBANKMENT: LOW SEVERITY | EMBANKMENT RABBLING | - RABBLING<br>- FILLING UP WITH EARTH | $M^3$ | 3.5<br>3.0 x 1.2[a] | 25.0<br>1.8 | 94.5 |
| FAILURE & EROSION OF ROAD EMBANKMENT: HIGH SEVERITY | CONSTRUCTION OF RETAINING WALL | - CONSTRUCTION OF RETAINING WALL<br>- FILLING UP WITH EARTH | $M^3$ | 8.65<br>10.25x1.2 | 25.0<br>1.8 | 238.4 |
| FAILURE OF RETAINING WALL: LOW SEVERITY | REBUILDING OF RABBLING | - REBUILDING | $M^3$ | 2.25 | 13.0 | 29.3 |
| FAILURE OF RETAINING WALL: HIGH SEVERITY | REBUILDING OF RETAINING WALL | - REBUILDING | $M^3$ | 6.25 | 13.0 | 81.3 |
| SURFACE HEIGHT: LOW SEVERITY | SURFACE RAISING: LOW CASE | - ADDING 3 LAYERS OF COMPACTED SOIL | $M^2$ | ROAD WIDTH (W) | 3 x 0.55 | 1.65 x W |
| SURFACE HEIGHT: HIGH SEVERITY | SURFACE RAISING: HIGH CASE | - ADDING 5 LAYERS OF COMPACTED SOIL | $M^2$ | ROAD WIDTH (W) | 5 x 0.55 | 2.75 x W |
| SURFACE DISTRESSES[b] : LOW SEVERITY | BLADING AND SPRINKLING ROAD SURFACE : LOW CASE | - 12 BLADING TIMES ANNUALLY<br>- 60 SPRINKLING TIMES ANNUALLY | $M^2$ | W1[c] | 0.016<br>0.011 | 0.85 x w1 |
| SURFACE DISTRESSES: HIGH SEVERITY | BLADING AND SPRINKLING ROAD SURFACE : HIGH CASE | - 18 BLADING TIMES ANNUALLY<br>- 90 SPRINKLING TIMES ANNUALLY | $M^2$ | W1 | 0.016<br>0.0111 | 1.27 x W1 |
| ILLEGAL IRRIGATION STRUCTURE: LOW SEVERITY | FILLING UP STRUCTURES | - FILLING UP WITH EARTH | $M^3$ | 0.4 | 5.0 | 2.0 |
| ILLEGAL IRRIGATION STRUCTURE: LOW SEVERITY | FILLING UP STRUCTURES | - FILLING UP WITH EARTH | $M^3$ | 1.4 | 5.0 | 7.0 |
| PLANTS ON ROAD SURFACE: LOW SEVERITY | REMOVING PLANTS: LOW CASE | - REMOVING PLANTS OF 1 M OF ROAD WIDTH | L. M. | 1.0 | 0.23 | 0.23 |
| PLANTS ON ROAD SURFACE: HIGH SEVERITY | REMOVING PLANTS: HIGH CASE | - REMOVING PLANTS OF 3 M OF ROAD WIDTH | L. M. | 1.0 | 0.47 | 0.47 |
| ROAD SURFACE OCCUPANCY: LOW SEVERITY | REMOVING OCCUPANCY: LOW CASE | - REMOVING OCCUPANCY OF 1 M OF ROAD WIDTH | L. M. | 1.0 | 0.42 | 0.42 |
| ROAD SURFACE OCCUPANCY: HIGH SEVERITY | REMOVING OCCUPANCY: HIGH CASE | - REMOVING OCCUPANCY OF 2 M OF ROAD WIDTH | L. M. | 1.0 | 0.85 | 0.85 |

[a] SHRINKAGE FACTOR.

[b] THE AVERAGE RATING OF ( CAMBER, RUTTING, AND POTHOLES).

[c] BLADING WIDTH IS ASSUMED TO BE 3, 6, OR 9 M.

[d] L. E. : EGYPTIAN POUND, AND L. M. : LINEAR METER.

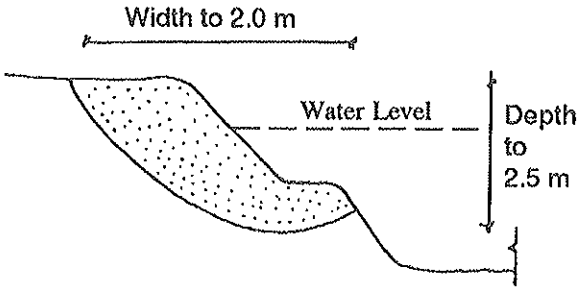



FIGURE 4 Embankment failure, low-severity.

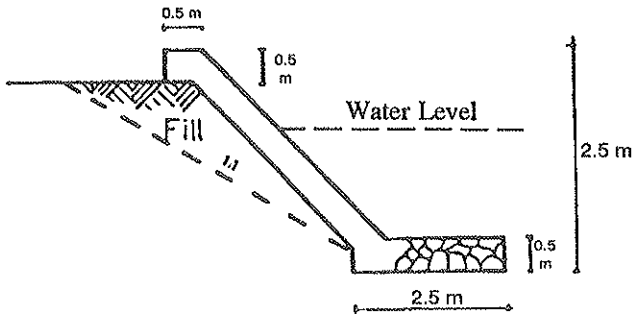



FIGURE 5 Embankment rabbling.

ing more services on roads to achieve the proposed level of serviceability.

## COST-SERVICEABILITY RELATION

On the basis of the required periodic maintenance activities and their associated costs of repairing the present distresses, the cost of bringing a section from its current ERCI level to an ERCI level of 100 can be calculated. This was done for all roads included in the sample to indicate the periodic and routine maintenance costs required to completely renew the network (ERCI = 100).

### Associated Periodic Maintenance Costs

Figure 6 shows the average ERCI value for each road rating included in the sample along with the associated cost of periodic maintenance required to bring it to an ERCI level of 100. For example, considering a road with an average ERCI value of 40 to 65 (i.e., in the poor range), the cost of bringing this road to its original ERCI of 100, using the appropriate maintenance actions to repair, remove, or eliminate the present distresses, is equal to about 117,000 Egyptian pounds (L.E.) per kilometer (U.S. $60,000/mi). On the other hand, a road with an average ERCI value of 80 to 90 (i.e., in the good range) has a corresponding renewal cost of about 43,784 L.E./km. In order to repair a section with a poor condition (on the basis of the ERCI system), it will cost about 2.5 times the cost of repairing the same section while it is in the good condition, on average. This result compares favorably with the results of several other studies done for paved roads (*8,9*).

### Associated Routine Maintenance Costs

As indicated in Table 1, the routine maintenance category contributes about 38 percent to the total rating scale and does not control the ERCI value. However, routine maintenance costs are affected mainly by the blading frequency required for a desired maintenance level. For the same maintenance level, routine maintenance costs should not control the priority scheme, because it is mainly affected by the physical measurements of the road (e.g., width). In addition, the magnitude of routine maintenance costs compared with that of the periodic maintenance is relatively small. Figure 6 justifies this statement.

### Environmental Impact

The effect of the surrounding environment is shown in Figure 7. In this figure, the average ERCI values and the associated renewal costs (as explained earlier) are shown for groups of sections, sections adjacent to drains, sections adjacent to canals, and sections penetrating the agricultural fields. The figure indicates that the impact of the surrounding environment represented by the adjacent canals or drains is significant to

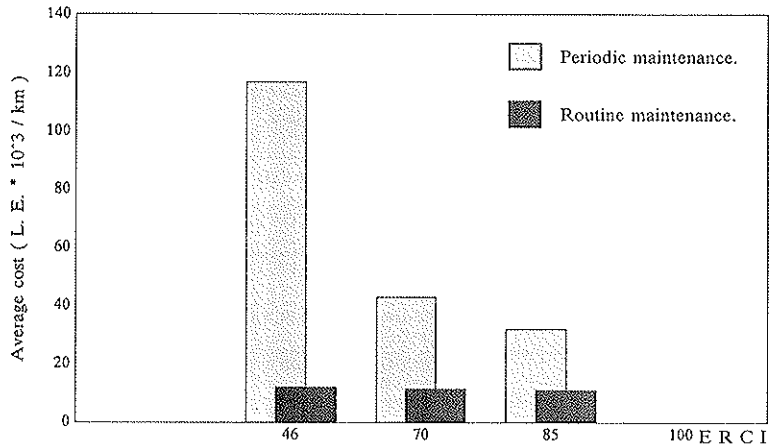



FIGURE 6 Maintenance costs versus road condition.

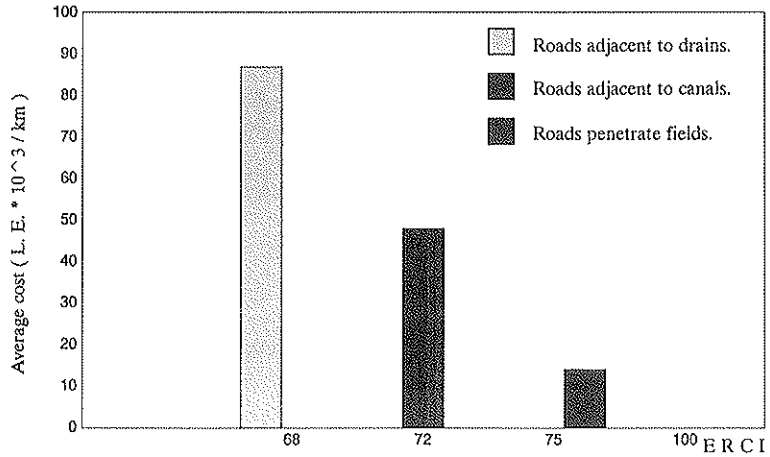



**FIGURE 7 Environmental impact on maintenance cost.**

the ERCI value and, consequently, to the repair cost. For instance, roads penetrating fields with no adjacent canals or drains have an average ERCI value of about 75 and an average renewal cost of about 12,000 L.E./km; on the other hand, roads with adjacent drains were found to have an average ERCI value of 68 and corresponding renewal costs of about 85,000 L.E./km. This result, in turn, implies that the current priority scheme, which only considers the traffic level, is not entirely correct and that other environmental factors should be considered.

## CONCLUSION

The condition assessment subsystem and its associated cost model were developed to provide the necessary information for management decisions, to be as simple and practical as possible, to be used to evaluate the identified road condition, and to provide the decision maker with data needs about maintenance budgeting and priority setting. The information developed through the ERCI system could be used along with information on traffic volume, improvements cost, and so forth, to help in making maintenance decisions.

In studying the effect of surrounding environment, adjacent canals, and drains on maintenance cost of earth roads, consideration both of the earth road and the adjacent stream under the same jurisdiction authority may result in a better maintenance management process.

## REFERENCES

1. M. Abdelrahman and E. Sharaf. Earth Road Condition Index Development. *Proc., 3rd IRF Regional Meeting*, Riyadh, Saudi Arabia, Feb. 1988.
2. M. Abdelrahman. *Low Volume Roads*. M.Sc. thesis, Zagazig University, Egypt, 1988.
3. M. Y. Shahin and S. D. Kohn. *Pavement Maintenance Management for Roads and Parking Lots*. Technical Report M294, U.S. Army Corps of Engineers, Oct. 1981.
4. C. H. Oglesby and R. G. Hicks. *Highway Engineering*, Wiley, New York, 1982.
5. J. D. Riverson. *Unpaved Road Maintenance Management in Local Highway System*. Ph.D. dissertation, Purdue University, West Lafayette, Ind., 1985.
6. W. J. Liddle. Application of AASHO Road Test Results to the Design of Flexible Pavement Structures. *First International Conference on the Structural Design of Asphalt Pavements*, University of Michigan, Ann Arbor, Aug. 1962.
7. H. Jämsä. Maintenance and Rating of the Condition of Gravel Roads in Finland. In *Transportation Research Record 898*, TRB, National Research Council, Washington, D.C., 1983, pp. 354–356.
8. E. A. Sharaf, M. Y. Shahin, and K. C. Sinha. Estimation of Maintenance Costs Using the Pavement Condition Index (PCI) Procedure. *Sixth IRF African Conference*, Cairo, Egypt, 1986.
9. R. F. Reichelt, E. A. Sharaf, and K. C. Sinha. *Development of Cost Model for the PAVER System*. Report CE-TRA-85-4, Purdue University, West Lafayette, Ind., 1985.

*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Simple Procedure for Selecting Best Maintenance Alternatives in Developing Countries

ESSAM A. SHARAF

A simple procedure for economically evaluating pavement maintenance alternatives and for selecting the most economical one is presented. The procedure includes several simple models that relate various cost and benefit items to pavement condition. Visual inspection is selected to represent pavement condition because it is the most common, and in most cases the only, method of pavement evaluation in developing countries. Cost items include (a) fixed initial cost, which depends on the type of maintenance alternative to be applied and the current unit costs of different resources, (b) variable initial cost (surface preparation) as a function of the condition of the existing pavement, and (c) annual routine (recurrent) maintenance as a function of pavement condition of the proposed maintenance alternative through its service life. Benefits, on the other hand, are considered in terms of savings in vehicle operating cost (VOC) resulting from applying maintenance alternatives. VOC is also related to pavement surface condition. A detailed example is presented to illustrate the procedure's step-by-step application. The results strongly indicate that ignoring one or more of the above-mentioned cost or benefit items can lead to considerable losses. This factor is particularly important in developing countries in which maintenance decisions are typically based on minimum initial cost without much attention given to other cost and benefit items.

A network level procedure for selecting the best maintenance alternative and calculating its associated costs and benefits for different pavement types at different levels of pavement condition is presented. This work is an extension of efforts initiated in late 1984 by Purdue University and the Construction Engineering Research Laboratory to develop a model that relates maintenance and repair costs to pavement surface condition (*1–6*).

The major addition to the original procedure is the inclusion of vehicle operating cost (VOC) in determining the best maintenance and repair alternative. This procedure is similar to the World Bank's procedure reported by NCHRP (*7*), except that it is simpler and requires fewer data.

The procedure has been initiated as a recommendation of the Expert Group Meeting of the Economic and Social Commission for Western Asia, United Nations, held in Cairo, Egypt, in 1989. The procedure, therefore, has been developed to serve the countries in this region in the first place; however, all efforts have been made to make the procedure flexible enough to allow its use in other developing countries.

Cairo University, Giza, Egypt. Current affiliation: Department of Civil Engineering, Faculty of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia.

In developing this procedure, several considerations have been taken into account:

1. The procedure should be simple so that it can be used by different management levels.
2. The procedure should be comprehensive to include different factors that may affect maintenance decisions.
3. The procedure should be applied with minimal dependence on sophisticated equipment and with maximum use of human resources, typically available in developing countries.
4. The procedure should be capable of producing results that address both costs and benefits associated with different maintenance alternatives so that rational approaches can be used when selecting the best alternative.
5. The procedure should be flexible enough to be applied in different environments.

The procedure is primarily based on the identification of pavement condition. Pavement condition can be evaluated through different techniques ranging from very simple ones to rather complicated ones that require considerable investment. In selecting the pavement condition assessment method for use in this procedure, the following two factors were considered:

1. The evaluation process should be simple and easily applied by different countries in the region.
2. The results of evaluation should provide information that address both structural and functional performance of pavements.

Therefore, visual inspection of pavement condition has been used in this procedure. From among several methods for visual pavement inspection, the pavement condition index (PCI) has been selected. This well-known method was originally developed by the U.S. Army Corps of Engineers (*8*). It depends on the detailed inspection of pavement and covers up to 19 different pavement distresses. Each distress is defined by its type, severity level (see Table 1), and extent (density) on the pavement area. The names of these distresses are presented in Table 2. The final rating of the pavement condition is based on the calculation of the PCI value. The PCI, a scale from 0 to 100 with 100 being excellent, is determined on the basis of measured type, severity, (see Table 2) and extent of different distresses. The PCI procedure is widely used in the United States, Europe, and in several countries in the local region, including Saudi Arabia, Jordan, and Egypt.

TABLE 1 DISTRESS SEVERITY LEVELS

| Severity code | Severity Level |
|---|---|
| 1 | Low |
| 2 | Medium |
| 3 | High |

The following section includes a step-by-step description of the procedure, and an example using the procedure.

## FAMILIES OF PAVEMENT SECTIONS

The network under consideration is classified into families of pavement sections. Each family includes a group of pavement sections that are subjected to the same general conditions, such as traffic, location, structural history, and pavement surface type.

As an example, agricultural roads with high traffic level and water sources (canals and drains) surrounding them should be differentiated from desert roads with low traffic level, etc. In this way, groups of nearly equivalent pavement sections are obtained. Pavement evaluation can be done through the PCI procedure by applying appropriate sampling techniques so that each family is adequately presented (*8*).

## PAVEMENT CLASSIFICATION

Pavement sections are grouped within each family on the basis of their structural type. Generally, there are numerous categories of pavement structural types, but typically they can be divided into the following types, particularly in the region under consideration:

1. Asphalt concrete;
2. Surface treatment;
3. Thin overlay (normally, less than 5 cm thick);
4. Thick overlay (normally, more than 5 cm thick); and
5. Rigid pavements.

Rigid pavements are not commonly used in paving roads in the region under consideration and therefore have been excluded from the procedure.

## MAINTENANCE AND REPAIR ACTIVITIES

Maintenance and repair activities should be grouped into a number of discrete activities that are commonly applied at different levels of pavement condition. Typical activities in the region are routine (or recurrent maintenance), surface treatment, thin overlay, thick overlay, and reconstruction.

## PCI RANGES

Because maintenance decisions are discrete in nature, pavement condition, represented by PCI values in this procedure, is also divided into discrete ranges. Typical ranges are

1. PCI = 80 to 100,
2. PCI = 60 to 80,
3. PCI = 40 to 60,
4. PCI = 20 to 40, and
5. PCI = 0 to 20.

At this point, the network is classified into families (categories) of sections. The objective is to select the most cost-effective maintenance and repair alternative for each of those categories at the various PCI ranges. This cost-effective approach is based on a comparison of alternatives using a life cycle costing calculation. Life cycle costing is typically based

TABLE 2 DISTRESS TYPES FOR ASPHALT CONCRETE PAVEMENTS

| Distress Code | Distress Name |
|---|---|
| 1 | Alligator Cracking |
| 2 | Bleeding |
| 3 | Block Cracking |
| 4 | Bumps and Sags |
| 5 | Corrugation |
| 6 | Depression |
| 7 | Edge Cracking |
| 8 | Reflection Cracking |
| 9 | Lane/Shoulder Dropoff |
| 10 | Longitudinal and Transverse Cracking |
| 11 | Patching and Utility Cut Patching |
| 12 | Polished Aggregate |
| 13 | Potholes |
| 14 | Railroad Crossing |
| 15 | Rutting |
| 16 | Shoving |
| 17 | Slippage Cracking |
| 18 | Swell |
| 19 | Weathering and Ravelling |

on two types of data: performance data and cost-benefit data. In the remainder of the procedure, these two categories of data are described for each pavement family.

## PERFORMANCE CURVES

Performance curves are those relations that describe the rate of change in pavement condition over time, under certain level of use (traffic), and subject to specific environmental factors such as rainfall, underground water, and temperature. The level of use (amount of traffic) and environmental factors are included in the original classifications of pavement section families. Therefore, this procedure suggests simple performance curves that relate PCI values to the age of pavement. This is typically done by collecting data on PCI values as well as obtaining information on pavement age (usually, from construction history files). Sets of PCI and age values can be easily obtained for sections in each family. This is followed by applying some statistical techniques (typically, regression analysis) to construct the relation between PCI and pavement age. Although several forms of such a relation can be expected, the one used in this procedure is as follows:

$$C = 100 - b * x^m \quad (1)$$

where

$C$ = PCI value,
$x$ = pavement age in months measured from the date of last application of major activity,
$b$ = slope coefficient, and
$m$ = value that controls the degree of curvature of the performance curve.

A set of statistical models, similar to that of Equation 1, can be developed for different pavement families. The importance of these model relations stem from the following:

1. They provide a tool for estimating the pavement condition in the future during the pavement's service life.
2. They can be used to estimate the service life of a pavement category. The service life is defined as the period after which the pavement reaches a certain terminal condition level described by the terminal PCI value. In this case, the model can be used in a reverse manner, that is, estimating the age of the pavement given a terminal PCI value.

## COSTS AND BENEFITS

One of the key principles of this procedure is to relate both costs and benefits to pavement condition. Unbiased comparison between different alternatives can be applied. Using this procedure, only direct costs and benefits are included. Indirect costs such as air pollution, noise, and other social costs are not included but could be easily incorporated if they were converted to the proper monetary form. Also, indirect benefits such as value of land, employment, and other social effects are not considered but could be incorporated if they were converted to a proper monetary form.

Direct costs considered in this procedure are the typical ones of initial costs (activity construction costs), and routine (or recurrent) maintenance during the service life. The salvage value is considered to be zero at the end of pavement service life. The direct benefits as considered in this procedure are those resulting from savings in vehicle operating cost (VOC). The VOC savings are presented in this procedure as follows:

VOC savings at a specific period (year)

$$= (\text{VOC})_t - (\text{VOC})_c \quad (2)$$

where

$(\text{VOC})_t$ = VOC at the terminal level of pavement condition, and
$(\text{VOC})_c$ = VOC at the current level of pavement condition.

That is, the VOC savings resulting from the application of an alternative equal the difference between VOC values when not applying the alternative and that when applying the alternative. Alternatives with higher performance (longer service life or lesser rates of deterioration) are expected to produce higher VOC savings during their service lives and may be selected in spite of their relative higher initial costs. This is particularly important because of the general tendency among highway top management to select the alternatives with the least initial cost.

### Initial Cost

Initial cost includes the following two components:

1. Surface preparation cost associated with the repairs of defective areas of the existing surface before applying the maintenance alternative itself. This cost is directly related to the condition of the existing pavement at the time of maintenance alternative application.
2. Maintenance alternative application cost that does not depend on pavement condition and only depends on current prices of different labor, material, and equipment elements used in the application of the particular maintenance alternative.

### Recurrent Maintenance Cost

Recurrent maintenance activities are day-to-day activities such as crack sealing and pothole patching. The cost of applying these activities is a function of pavement condition at the time of repair.

### VOCs

VOCs are the costs required to run a vehicle. They include depreciation, fuel, lubrication, time, maintenance, and so on. Although the mechanism for calculating these costs will not be described, VOC dependence on pavement condition is a key element in the procedure. Otherwise, it would be impossible to indicate the benefits of maintenance activities (the reduction of VOCs caused by improving pavement condition).

In order to identify these benefits, the following steps were performed during the testing of this procedure, and it is believed that they could be replicated easily in any other environment:

1. Necessary information was collected about the most common vehicle types using the considered road network—for example, representative vehicles for passenger cars, small trucks, combination trucks, and articulated trucks. Then, the necessary mechanical properties and unit costs were estimated, on an average basis. The purpose of this step was to obtain the necessary information required to run the World Bank's VOC model, which estimated the total VOC for each vehicle type at different pavement conditions. This model is based on data from several countries and has proven to be accurate enough to estimate VOCs for different vehicle types (*9*).
2. After the model was run, the following equations were obtained:

$$(VOC)_{PC} = e^{(5.624 + 0.06814 \cdot IRI)} \quad (3)$$

$$(VOC)_{ST} = e^{(6.337 + 0.06516 \cdot IRI)} \quad (4)$$

$$(VOC)_{MT} = e^{(6.465 + 0.06766 \cdot IRI)} \quad (5)$$

$$(VOC)_{AT} = e^{(6.899 + 0.05116 \cdot IRI)} \quad (6)$$

where

$(VOC)_{PC}$ = estimated vehicle operating cost for a passenger car (L.E. per 1,000 veh-km);
$(VOC)_{ST}$ = estimated vehicle operating cost for a small truck (L.E. per 1,000 veh-km);
$(VOC)_{MT}$ = estimated vehicle operating cost for a medium truck (L.E. per 1,000 veh-km);
$(VOC)_{AT}$ = estimated vehicle operating cost for an articulated truck (L.E. per 1,000 veh-km); and
IRI = International roughness index, an index that represents the degree of unevenness of a pavement section, which is highly correlated to VOC.

Equations 3–6 represent the condition of Egypt (*10*). However, similar equations can be easily developed for other countries in the region when appropriate vehicle and roadway characteristics are used.

The values of IRI are typically obtained using roughness measurement equipment. In this procedure, a simplified approach has been used to avoid using roughness equipment. Surveying measurements were used on about 60 sections, and the IRI was calculated for each section (*11*,*12*). Detailed visual inspection was also completed on the same sections, and PCI values were determined. A set of PCI, IRI data pairs was available; regression analysis was applied resulting in the following simple model:

$$IRI = 0.15(100 - PCI)$$
$$R^2 = 0.81 \quad (7)$$

Figures 1 and 2 show scatter plots for IRI and PCI values and the observed versus predicted values, respectively. The

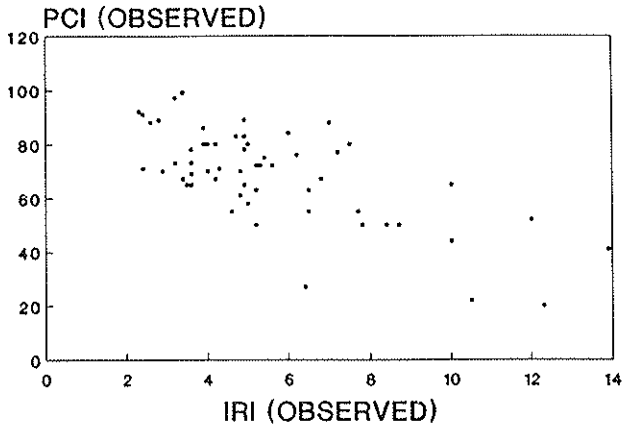



**FIGURE 1 Scatter plot for observed values of IRI and PCI measurements.**

IRI values can be easily estimated from the PCI values using Equation 7.

Calculation of VOCs can best be illustrated through another simple example, assuming the following data:

- A highway section is of 10-km length and 7.5-m width;
- Its average daily traffic (ADT) is 10,000 veh/day with 70 percent passenger cars, 10 percent small trucks, 10 percent medium trucks, and 10 percent articulated trucks; and
- Its current PCI level is 80.

In order to calculate the total VOC, Equations 3–6 are first used to determine the VOC values for passenger car (PC), small truck (ST), medium truck (MT), and articulated truck (AT), respectively. In the application of these equations, the value of IRI calculated using Equation 7 is 0.15(100 − 80) = 3. For this case,

$$(VOC)_{PC} = 339.8 \text{ L.E. per } 1{,}000 \text{ veh-km} \quad (8)$$

$$(VOC)_{ST} = 678.1 \text{ L.E. per } 1{,}000 \text{ veh-km} \quad (9)$$

$$(VOC)_{MT} = 786.8 \text{ L.E. per } 1{,}000 \text{ veh-km} \quad (10)$$

$$(VOC)_{AT} = 1{,}155.7 \text{ L.E. per } 1{,}000 \text{ veh-km} \quad (11)$$

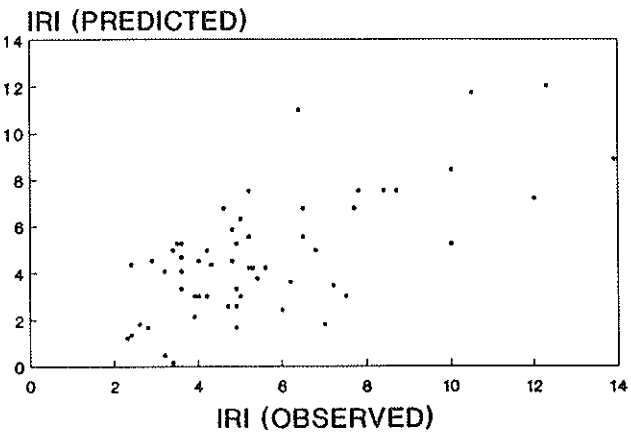



**FIGURE 2 Observed versus predicted values of IRI using simplified relation between IRI and PCI.**

The average VOC is then calculated using the percentages of the different vehicles, as follows:

$$(VOC)_{average} = 0.7 * 339.8 + 0.1 * 687.1 + 0.1 * 1{,}155.7 \text{ L.E. per } 1{,}000 \text{ veh-km}$$

$$= 500.8 \text{ L.E. per } 1{,}000 \text{ veh-km}$$

Total VOC on this section is 500.8 * total veh-km/1,000 = 500.8 * 10 * 10,000/1,000 = 50,080 L.E. per unit area of road, VOC/m² = 50,080/(10,000 * 7.5) = 0.668 L.E./m². The cost per area per year, (VOC/m²)/year = 0.668 * 360 = 240.5 (L.E./m²)/year.

The last value allows direct comparison between total VOC and maintenance cost [which is normally calculated in terms of (L.E./m²)/year].

As mentioned, both the surface preparation part of the initial cost and the routine (recurrent) maintenance costs during the service life of an alternative are dependent on PCI value. In order to determine such costs as a function of PCI, one may think of the required repair cost as the defective area multiplied by the unit area cost of repair. Thus, if at any PCI level, the average defective area and the corresponding required repair action and cost are estimated, then the surface preparation or recurrent maintenance at the PCI level can be easily estimated.

This process led to the development of the density matrix for each pavement category under consideration. A distress density is defined as the percent of section area indicating that specific distress type. The density matrix of a specific pavement category summarizes the average density values for each PCI range by distress type and severity level combination. Figure 3 shows a typical density matrix. For instance, in Figure 3, the average density of low-severity (Severity Code 1) alligator cracking (Distress Code 1) is 9.64 percent at PCI range 0 to 20, whereas the average density of high-severity (Severity Code 3) weathering and raveling (Distress Code 19) is 23.47 percent for PCI range 0 to 20. For this pavement category, at PCI range 0 to 20, there is, on the average, about 9.46 percent of a section area with low-severity alligator cracking and 23.47 percent weathering and raveling.

If the repair method for each distress and severity level is defined, then the required repair costs for the various PCI ranges can be estimated. These values, in turn, lead to the development of the surface preparation policy matrix and recurrent maintenance policy matrix. A surface preparation policy matrix contains the necessary surface preparation actions for different distress-severity level combinations, whereas the recurrent maintenance policy matrix contains the necessary recurrent maintenance actions for different distress-severity level combinations.

| Distress Code | Severity Code | Average Density (%) by PCI Range | | | | |
|---|---|---|---|---|---|---|
| | | 81-100 | 61-80 | 41-60 | 21-40 | 0-20 |
| 1 | 1 | 0.32 | 1.15 | 5.54 | 11.36 | 9.46 |
| 1 | 2 | 0.13 | 0.20 | 1.80 | 10.24 | 14.09 |
| 1 | 3 | 0.03 | 0.04 | 0.39 | 10.05 | 14.02 |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| 19 | 1 | 4.49 | 12.44 | 17.15 | 19.57 | 11.25 |
| 19 | 2 | 0.56 | 1.18 | 6.60 | 10.84 | 17.31 |
| 19 | 3 | 0.06 | 0.08 | 0.42 | 6.04 | 23.74 |

**FIGURE 3 Portion of a typical density matrix (3).**

Figure 4 shows a typical surface preparation matrix. The recurrent maintenance matrix is similar to the surface preparation matrix except that the repair policies of the recurrent maintenance matrix differ from those of the surface preparation. The repair policies associated with recurrent maintenance are relatively of lower standard if compared to those of surface preparation.

With the density matrix, surface preparation policy matrix, and annual recurrent maintenance policy matrix, the following equations are used to calculate required surface preparation or annual recurrent maintenance costs at any PCI range:

$$(SP)_k = (D_{ij})_k * (C_{ij})_{sp} \quad (12)$$

$$(RM)_k = (D_{ij})_k * (C_{ij})_{rm} \quad (13)$$

where

$(SP)_k$ = surface preparation cost at the $k$th PCI range,
$(RM)_k$ = recurrent maintenance cost at the $k$th PCI range,
$(D_{ij})_k$ = average density of the $i$th distress type with the $j$th severity level combination at the $k$th PCI range (from the density matrix of the pavement category under consideration),
$(C_{ij})_{sp}$ = unit cost of the required surface preparation action for the $j$th severity level of the $i$th distress type (from the surface preparation policy matrix), and
$(C_{ij})_{rm}$ = unit cost of the required recurrent maintenance action for the $j$th severity level of the $i$th distress type (from the recurrent maintenance policy matrix).

## SUMMARY OF PROCEDURE APPLICATION

The purpose of this procedure is to select the most cost-effective maintenance alternative at a specific pavement condition state (PCI range). In order to achieve that, different cost and benefit elements are calculated for each alternative and the one with the highest net equivalent uniform annual cost is considered to be the most cost-effective alternative. In the following section is a summary of cost and benefit elements as considered in this procedure.

### Initial Cost

#### *Fixed Cost*

Fixed cost, which has no relation to the condition of the existing pavement, can be estimated from the most current similar contracts.

| Distress Type | Severity | Preparation Method | Unit | Unit Cost |
|---|---|---|---|---|
| 1- Alligator Cracking | H | Deep Patch | SF | 2.32 |
| | M | Skin Patch | SF | 0.55 |
| 3- Block Cracking | H | Skin Patch | SF | 0.55 |
| 4- Bumps / Sags | H | Skin Patch | SF | 0.55 |
| 5- Corrugation | H | Deep Patch | SF | 2.32 |
| 6- Depressions | H | Deep Patch | SF | 2.32 |
| 7- Edge Cracking | H | Deep Patch | LF | 3.48 |
| | M | Deep Patch | LF | 2.9 |
| 9- Lane/Shoulder Dropoff | H | Add Aggregate and Grade | LF | 0.14 |
| 10- Long./Trans. Cracking | H | Shallow Patch | LF | 1.65 |
| | M | Crack Fill | LF | 1.00 |
| 11- Patching and Utility | H | Deep Patch | SF | 2.32 |
| 13- Potholes | H | Deep Patch | ea | 7.29 |
| | M | Shallow Patch | ea | 1.82 |
| | L | Pothole Filling | ea | 0.79 |
| 14- Railroad Crossing | H | Deep Patch | SF | 2.32 |
| 15- Rutting | H | Deep Patch | SF | 2.32 |
| | M | Skin Patch | SF | 0.39 |
| 16- Shoving | H | Deep Patch | SF | 2.32 |
| 17- Slippage Cracking | H | Skin Patch | SF | 0.55 |
| 18- Swell | H | Deep Patch | SF | 2.32 |

**FIGURE 4 Typical surface preparation policy matrix (*3*).**

*Surface Preparation Cost*

Surface preparation cost is directly related to the condition of the existing pavement at the time of application activity. It is estimated using Equation 12 taking into consideration the PCI range of the existing pavement.

### Annual Maintenance Cost

Annual maintenance cost is directly related to the condition of the pavement for each year within the service life of the proposed alternative. It is estimated using Equation 13 taking into consideration the PCI range during each year of the alternative service life (Equation 1 is used to estimate PCI for each year). For the alternative under consideration, the service life is calculated using Equation 1 with the appropriate terminal PCI value and the appropriate value for coefficient *b*.

### Vehicle Operating Cost

VOC is directly related to the condition of the pavement, represented in terms of the IRI value. The IRI value is calculated using Equation 7. The VOC value is then estimated using Equations 3–6.

### Savings and Benefits

Savings are the difference between the VOC value in a specific year with a specific IRI value and the VOC value at the terminal IRI value. This calculation is done for each year within the service life of an alternative. The IRI values for each year are estimated using Equation 7 after estimating the PCI value for each year using Equation 1.

## EXAMPLE OF PROCEDURE APPLICATION

A simple example of the use of the procedure is given in this section. The purpose of this example is to illustrate the step-by-step application of the procedure to select a best maintenance alternative among several alternatives. In addition, several important conclusions based on this example are discussed in the following section.

### Input Data

- Pavement section length = 30 km.
- Pavement section width = 7.5 m.
- ADT = 5,000 veh/day.
- Percentage of passenger cars = 70 percent.
- Percentage of small trucks = 10 percent.

- Percentage of medium trucks = 10 percent.
- Percentage of articulated trucks = 10 percent.
- Current PCI level = 30.
- Terminal PCI level = 30.
- Existing pavement type = thin overlay.
- Surface preparation cost of existing pavement (at current PCI) = 8.5 L.E./m² (from appropriate density matrix).
- Maintenance costs of different alternatives at different PCI ranges are as follows (from appropriate density matrices):

| | *PCI Range* | | | | |
|---|---|---|---|---|---|
| *Alternative* | *0–20* | *21–40* | *41–60* | *61–80* | *81–100* |
| Surface dressing | 7.70 | 2.20 | 0.80 | 0.50 | 0.13 |
| Thin overlay | 7.00 | 2.00 | 0.70 | 0.35 | 0.13 |
| Thick overlay | 4.00 | 1.00 | 0.60 | 0.30 | 0.07 |
| Reconstruction | 4.40 | 1.30 | 0.65 | 0.33 | 0.07 |

- Performance equations of different maintenance alternatives are as follows:

$PCI = 100 - 0.0319(age)^{1.5}$ for surface dressing.

$PCI = 100 - 0.0158(age)^{1.5}$ for thin overlay.

$PCI = 100 - 0.0129(age)^{1.5}$ for thick overlay.

$PCI = 100 - 0.0104(age)^{1.5}$ for new asphalt and reconstructed pavement.

- Inflation-adjusted discount rate = 6 percent.

### Procedure Steps

- The initial cost of each available maintenance alternative is presented in Table 3. It consists of the fixed part and the surface preparation part (which is a function of the PCI value of the existing pavement with thin overlay). The surface preparation cost is calculated using the appropriate density and surface preparation policy matrices, as described earlier.
- The expected service lives of different maintenance alternatives are calculated using the performance equations (by substituting PCI = terminal value of 30). The resulting service lives are presented in the last column in Table 4.

- Table 5 presents the calculation of the equivalent uniform annual cost (EUAC) of the initial costs of different maintenance alternatives.
- The performance equations are then used to calculate the PCI values at each year within the service life of an alternative. Several steps follow the calculation of the PCI value in a specific year. The results of these steps are presented in Table 6 for surface dressing. Similar tables are developed for other alternatives. Each table consists of 12 columns, as follows:

Column 1. Year number within the service life of the alternative.

Column 2. PCI value as calculated from the performance equation.

Column 3. Annual maintenance cost (L.E./m²). This cost is determined by first converting the PCI value to its corresponding range (0–20, 20–40, etc.). Then, the associated annual maintenance cost is determined from the table given in the input data.

Column 4. The single payment present worth factor (sppwf) corresponding to the year ($n$) under consideration.

Column 5. The present worth value of the annual maintenance cost = Column 3 * Column 4.

Column 6. EUAC of annual maintenance cost = [capital recovery factor (CRF)] * [present worth value (PWV) of annual maintenance costs].

Column 7. IRI estimated by Equation 7 as a function of PCI value (Column 2).

TABLE 3 INITIAL COSTS

| | INITIAL COST | | |
|---|---|---|---|
| ALTERNATIVE | FIXED | SURFACE PREPARATION | T O T A L |
| SURFACE TREATMENT | 1.58 | 8.5 | 10.08 |
| THIN OVERLAY | 3.76 | 8.5 | 12.26 |
| THICK OVERLAY | 5.07 | 8.5 | 13.57 |
| RECONSTRUCTION | 20.70 | 0.0 | 20.70 |

TABLE 4 PERFORMANCE MODELS AND SERVICE LIVES

| ALTERNATIVE | PERFORMANCE CURVE | SERVICE LIFE AT PCI = 30 |
|---|---|---|
| SURFACE TREATMENT | $PCI = 100 - 0.0319(age)^{1.5}$ | 14 |
| THIN OVERLAY | $PCI = 100 - 0.0158(age)^{1.5}$ | 22 |
| THICK OVERLAY | $PCI = 100 - 0.0129(age)^{1.5}$ | 25 |
| RECONSTRUCTION | $PCI = 100 - 0.0104(age)^{1.5}$ | 29 |

TABLE 5 EUAC OF ALTERNATIVE INITIAL COSTS

| | MAINTENANCE ALTERNATIVE | | | |
|---|---|---|---|---|
| | SURFACE DRESSING | THIN OVERLAY | THICK OVERLAY | RECONSTRUCTION |
| i | 6 | 6 | 6 | 6 |
| n | 14 | 22 | 25 | 29 |
| CRF | 0.1075 | 0.0830 | 0.0780 | 0.0739 |
| INITIAL COST | 10.08 | 12.26 | 13.57 | 20.70 |
| EUAC (I.C.) | 1.08 | 1.02 | 1.06 | 1.53 |

RF : Capital Recovery Factor
UAC: Equivalent Uniform Annual Cost

TABLE 6 SURFACE DRESSING

| n | PCI | A.M.C | sppwf | PWV | EUAC | IRI | VOC AT CURRENT | VOC AT PCI= 30 | VOC SAVINGS | PWV | EUA BENEFIT |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 98.67 | 0.13 | 0.943 | 0.122 | | 0.2 | 101.84 | 197.80 | 95.96 | 90.49 | |
| 2 | 96.20 | 0.13 | 0.889 | 0.116 | | 0.57 | 104.25 | 197.80 | 93.55 | 83.53 | |
| 3 | 93.10 | 0.13 | 0.840 | 0.109 | | 1.04 | 107.21 | 197.80 | 90.59 | 76.09 | |
| 4 | 89.00 | 0.13 | 0.790 | 0.103 | | 1.65 | 110.20 | 197.80 | 87.60 | 69.20 | |
| 5 | 85.00 | 0.13 | 0.750 | 0.098 | | 2.25 | 116.02 | 197.80 | 81.78 | 61.34 | |
| 6 | 80.00 | 0.50 | 0.710 | 0.360 | | 3.00 | 121.70 | 197.80 | 76.10 | 54.03 | |
| 7 | 75.00 | 0.50 | 0.670 | 0.340 | | 3.75 | 127.60 | 197.80 | 70.20 | 47.03 | |
| 8 | 70.00 | 0.50 | 0.630 | 0.320 | | 4.50 | 133.92 | 197.80 | 63.88 | 40.24 | |
| 9 | 64.00 | 0.50 | 0.590 | 0.295 | | 5.40 | 141.84 | 197.80 | 55.96 | 33.02 | |
| 10 | 58.00 | 0.80 | 0.560 | 0.450 | | 6.30 | 150.25 | 197.80 | 47.55 | 26.63 | |
| 11 | 51.62 | 0.80 | 0.526 | 0.420 | | 7.26 | 159.60 | 197.80 | 38.20 | 20.09 | |
| 12 | 44.88 | 0.80 | 0.496 | 0.397 | | 8.27 | 170.10 | 197.80 | 27.70 | 13.74 | |
| 13 | 37.84 | 2.20 | 0.468 | 1.030 | | 9.32 | 182.70 | 197.80 | 15.10 | 7.07 | |
| 14 | 30.50 | 2.20 | 0.442 | 0.970 | | 10.43 | 195.30 | 197.80 | 2.50 | 1.11 | |
| Σ | | | | 5.130 | 0.55 | | | | | 623.60 | 67.03 |

AMC : Annual Maintenance Cost
SPPWF : Single Payment Present Worth Factor
PWV : Present Worth Value
EUAC : Equivalent Uniform Annual Cost

TABLE 7 COMPARISON OF COST ELEMENTS OF ALTERNATIVES

| | MAINTENANCE ALTERNATIVE | | | |
|---|---|---|---|---|
| | SURFACE TREATMENT | THIN OVERLAY | THICK OVERLAY | RECONSTRUCTION |
| EUA (I.C) | 1.08 | 1.02 | 1.06 | 1.53 |
| EUA (AMC) | 0.55 | 0.35 | 0.23 | 0.25 |
| EUA (IC+AMC) | 1.63 | 1.37 | 1.29 | 1.78 |
| EUA (BENEFITS) | 67.03 | 71.68 | 73.18 | 75.12 |
| EUA (NET) | 65.40 | 70.31 | 71.89 | 73.34 |

TABLE 8 BEST ALTERNATIVE AND ASSOCIATED MINIMUM COSTS UNDER DIFFERENT CURRENT CONDITION CASES

| | CURRENT PCI = 30 | | CURRENT PCI = 60 | | CURRENT PCI = 80 | |
|---|---|---|---|---|---|---|
| | i=6 | i=12 | i=6 | i=12 | i=6 | i=12 |
| SELECTION BASED ON INITIAL COST | THIN OVERLAY (1.02) | SURFACE DRESSING (1.52) | THIN OVERLAY (0.89) | SURFACE DRESSING (1.20) | SURFACE DRESSING (0.42) | SURFACE DRESSING (0.50) |
| SELECTION BASED ON INITIAL + MAIN. COST | THICK OVERLAY (1.29) | THIN OVERLAY (1.86) | THICK OVERLAY (1.09) | THIN OVERLAY (1.43) | SURFACE DRESSING (0.60) | SURFACE DRESSING (0.67) |
| SELECTION BASED ON BENEFITS | RECONST (73.34) | RECONST. (78.93) | RECONST. (30.06) | THICK OVERLAY (33.38) | THICK OVERLAY (12.26) | THIN OVERLAY (13.10) |

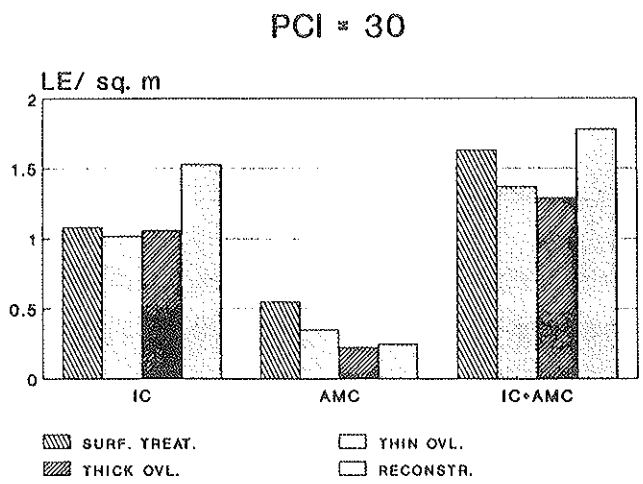



FIGURE 5 EUAC values for different maintenance alternatives under different cost consideration policies.

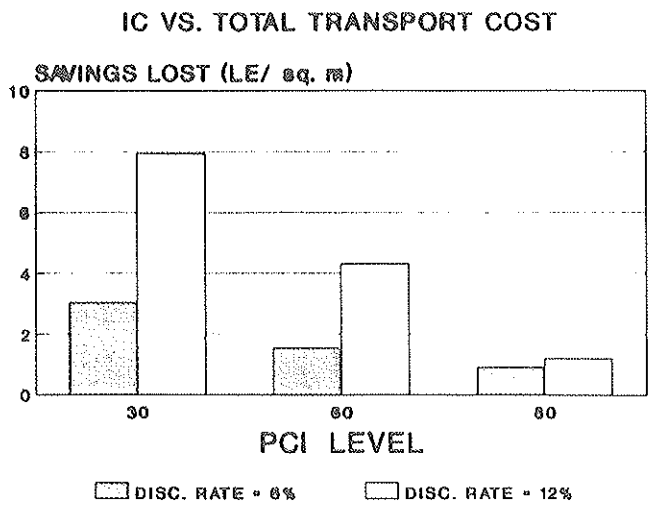



FIGURE 7 Savings lost from not considering both annual maintenance cost and VOCs.

Column 8. VOC calculated using earlier procedure, represented in units of L.E./m² to allow direct comparison of VOC and maintenance costs.

Column 9. VOC at terminal PCI value (30 in this example), using procedure of previous example.

Column 10. VOC savings (benefits) resulting from applying the alternative under consideration = Column 9 − Column 8.

Column 11. PWV of VOC savings (benefits) = Column 4 − Column 10.

Column 12. Benefits expressed in EUAC form = CRF * PWV of benefits.

## SUMMARY

A procedure was developed for helping decision makers in developing countries in selecting the best maintenance alternative among several available, taking into consideration initial and annual maintenance costs associated with each available alternative in addition to the associated VOC during the service life of each alternative. VOC savings are used to represent the benefits of applying a maintenance alternative. VOC calculations are based on the World Bank's models (*10*).

The procedure is based on relating both the agency costs (initial and maintenance costs) and user costs (VOCs) to the pavement surface condition. Pavement surface condition is presented in this procedure in terms of the PCI value. A life cycle cost analysis is then applied and the alternative with the minimum EUAC is considered to be the best one.

An example is presented that describes a step-by-step application of the procedure. The main results for this example are presented in Tables 7 and 8 and in Figures 5–8.

Table 6 and Figures 5 and 6 indicate the EUAC values associated with different maintenance alternatives under different cost item considerations. For instance, if the initial cost is the only criterion for selecting the best alternative, then thin overlay will be selected as the best alternative, with an EUAC value of 1.02 L.E./m². On the other hand, thick overlay will be selected (with EUAC of 1.29 L.E./m²) if the

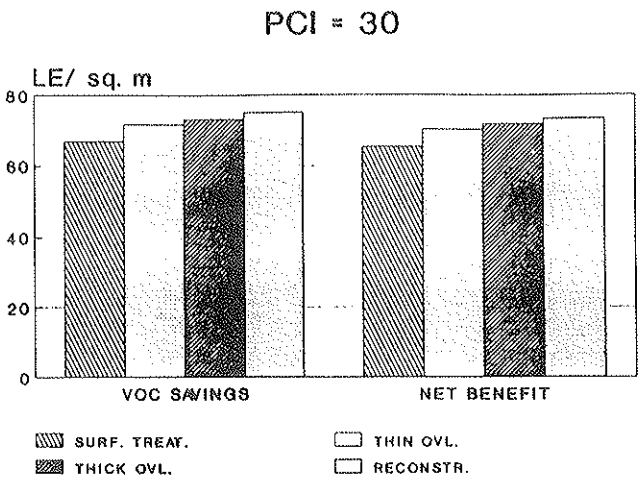



FIGURE 6 EUAC values for VOC savings and net benefits associated with the selection of different maintenance activities.

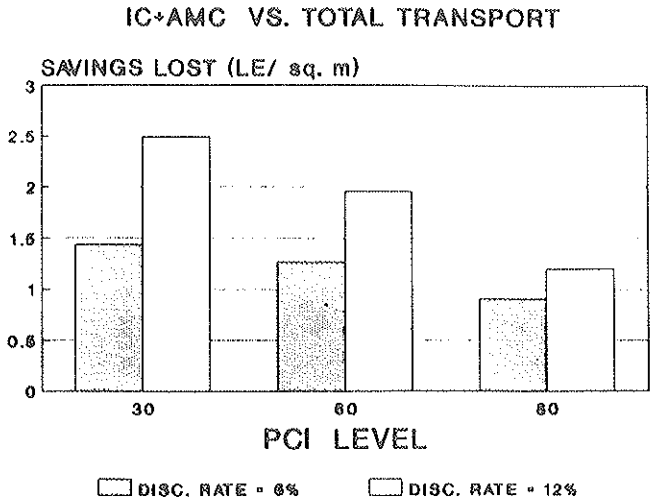



FIGURE 8 Savings lost from not considering VOCs.

criterion of selection is based on both minimum initial and annual maintenance cost; and reconstruction will be selected if the criterion is based on maximum benefits (VOC savings = 75.12 L.E./m²) or maximum net benefits (73.34 L.E./m²).

The selected alternative and its associated costs are significantly dependent on the selection criterion. This, in turn, implies that considerable savings could be lost if the selection criterion neglects one or more of the cost categories (VOC, in particular). This particular point is further analyzed and presented in Table 8 and Figures 7 and 8 where different PCI levels (current and terminal) are used (PCI = 30, 60, and 80) and two levels of the inflation-adjusted discount rate (6 and 12 percent) are used. Table 8 lists the selected best alternative under each criterion for the different cases of current PCI values and discount rates as described earlier. Figure 7 shows the corresponding savings lost after neglecting both the annual maintenance and VOC and basing the selection on the initial cost only, which is the case in most countries in this region. Figure 8, on the other hand, shows the values of savings lost after neglecting VOC only. These two figures indicate an extremely important issue, particularly in developing countries in which decisions, in most cases, are based on minimum initial cost plus some indications about the performance of different maintenance alternatives. Finally, the results strongly recommend alternatives with higher initial costs to obtain maximum benefits.

## REFERENCES


1. E. F. Reichelt, E. A. Sharaf, and K. C. Sinha. Development of a Cost Model for the PAVER System. Report CE-TRA-85-4, Purdue University, West Lafeyette, Ind., 1985.
2. E. A. Sharaf, M. Y. Shahin, and K. C. Sinha. Estimation of Maintenance Costs Using the Pavement Condition Index (PCI) Procedure. *Proc., 4th International Road Federation Conference*, Cairo, Egypt, 1986.
3. E. F. Reichelt, E. A. Sharaf, K. C. Sinha, and M. Y. Shahin. The Relationship of Pavement Maintenance Costs to the Pavement Condition Index. Interim Report M-87/02. Construction Engineering Research Laboratory, 1987.
4. E. A. Sharaf, E. Reichelt, M. Y. Shahin, and K. C. Sinha. Development of a Methodology to Estimate Pavement Maintenance and Repair Costs for Different Ranges of Pavement Condition Index. In *Transportation Research Record 1123*, TRB, National Research Council, Washington, D.C., 1987.
5. E. A. Sharaf, M. Y. Shahin, and K. C. Sinha. Analysis of the Effect of Deferring Pavement Maintenance. In *Transportation Research Record 1205*, TRB, National Research Council, Washington, D.C., 1988.
6. E. A. Sharaf, M. Y. Shahin, and M. T. Moustafa. Consequences of Delaying Pavement Maintenance on the Life Cycle Cost of Maintenance Alternatives. *Proc., 3rd IRF Meeting*, Riyadh, Saudi Arabia, 1988.
7. D. E. Peterson. *NCHRP Synthesis of Highway Practice 122: Life-Cycle Cost Analysis of Pavements*. TRB, National Research Council, Washington, D.C., 1985.
8. M. Y. Shahin and S. D. Kohn. Pavement Maintenance Managment for Roads and Parking Lots. Technical Report M-294, Construction Engineering Research Laboratory, 1981.
9. A. Dhareshwar and R. Archondo-Callao. Vehicle Operating Cost Prediction. Transportation Department, The World Bank, Washington, D.C., 1986.
10. E. A. Sharaf and D. F. Hanno. An Analysis of the Effect of Pavement Condition on Vehicle Operating Costs. *Proc., 1st Al-Azhar Engineering Conference*, Cairo, Egypt, 1989.
11. *Effect of Increased Axle Load Limits on the Egyptian Highway Network*. Final Report, National Transport Institute, Ministry of Transport, Egypt, 1990.
12. A. M. Abd-Allah. Analysis of Flexible Pavement Roughness in Egypt. Master's thesis, Department of Construction Management, Zagazig University, Egypt, 1990.


*Publication of this paper sponsored by Committee on Maintenance and Operations Management.*

# Making the Change to Maintenance Management Systems and Optimizing the Results

T. G. B. ARMITAGE

The introduction of a maintenance management system (MMS) into a roading organization is a significant organizational change. The theory behind making change is examined to establish the ground rules for successful change, highlighting the importance of *ownership* of the change. Examples are given of the application of those rules from work in Australia and New Zealand. Better management information is only one type of enhancement needed to improve organizational performance. Other work required to enhance the results of an MMS includes specifying the best maintenance methods and training people in them, and developing the technical and managerial competence of supervisors and managers to strengthen the quality of decision making and to provide appropriate management support for the MMS.

In Australia and New Zealand, there is considerable work in the development of pavement management systems (PMS) and maintenance management systems (MMS) for roads. Many organizations have had types of manual systems for years, but the advent of computers allows much more sophisticated systems to be available to everyone.

There are many advantages of an MMS, and persons involved with design and application of such a system generally understand these advantages. However, this knowledge is not necessarily shared by personnel in the field who are the ones affected by use of such systems.

Changes in government policies in New Zealand are forcing traditional roading organizations to become totally commercial, and in both countries increased productivity is becoming an important part of public corporate life. The pressure is on road maintenance organizations to improve performance and justify their share of public funding.

Roading organizations normally make changes in small increments. So for many field personnel unaccustomed to systematic planning, the introduction of an MMS may pose the most significant change of their working life; demanding a change in attitude towards work planning accompanied by the acquisition of new skills.

Most people feel uncomfortable when faced with a change from existing work practices; they tend to resist the change and return to the status quo. For this reason, the most well-intended and important change can have a limited effect and a short life, even after it has shown significant initial benefits.

It is critical to the successful introduction of an MMS that the field personnel not only know all about the system and understand its advantages for them, but also that they are willing to use it. It is equally as important that the change makers understand the basic constituents of developing and introducing successful change.

Armitage Consultancy Limited, P.O. Box 8174, Christchurch, New Zealand.

## THEORY OF ORGANIZATIONAL CHANGE

In considering organizational change, it is useful to examine the fundamental systems that operate within an organization.

Leavitt (*1*) considered organizations as "multivariate systems, in which at least four interacting variables loom especially large." He identified these four variables as

- Task. Ranging from the main reason for the organization (e.g., road maintenance) to the actual work done (e.g., repairing edgebreaks).
- People. Those carrying out the tasks.
- Technology. The technical tools and systems used in problem solving and carrying out the tasks.
- Structure. The systems of communication, authority and work flow.

Leavitt postulated that the "four are highly interdependent, so that change in any one will most probably result in compensatory (or retaliatory) change in others."

### Structural Change

In New Zealand, the creation of state-owned enterprises (SOEs) has usually been accompanied by restructuring in the affected organizations. The move to decentralize SOEs has affected people by giving them more accountability for managing their own profit centers. By concentrating on their core task, the SOEs have eliminated peripheral activities and consequently down-sized (i.e., made people redundant).

For example, the new Telecom organization restructured into five autonomous regional operating companies and several special purpose companies (e.g., international tolls). The companies have eliminated activities such as building construction, furniture manufacture, and vehicle servicing to concentrate on their core activity—telecommunications. New computerized accounting procedures (a technology change) will affect people (through redundancies and retraining), tasks (the type of work will change), and structure (communication systems will change).

Usually these changes have been introduced without consultation, against the belief of people-oriented practitioners that "human acceptance of ideas is the real carrier of change; and that emotional human resistance is the real road block" (*1*). In the case of SOEs, however, emotional human resistance has not been an impediment to change. The change itself has been so overwhelming that it has surpassed human feelings. It has involved the fundamental change of the whole organization, rather than improvements within the existing organization.

The situation is different when a manager is trying to make change within an existing organization. The Hawthorne study (*2*) found that people react well to change if it is discussed with them in advance, and subsequent experience continues to support this positive note. Peters and Waterman (*3*) found that the excellent companies owed much of their success to their people-oriented approaches.

### Technological Change

A new management system can be considered as a technological approach to organizational change. The most famous (or infamous) precedent in the area of technological approaches is Taylor's *Scientific Management* (*4*), which began early this century. Pilloried as inhuman, it created a separate planning specialist, thus removing the planning role (and with it the right to choose) from workers and their supervisors.

The perception of management systems is little different from that of scientific management. The problem is that the "technological and structural approaches tend to focus on problem-solving, sliding past the microprocesses by which new problem-solving techniques are generated and adopted" (*1*). This statement suggests that the developers of the new systems have not recognized the need to work with the end users during development, so that the system will meet end user requirements and be readily adopted by end users. Unfortunately, this negative factor continues to be the case with some developers of management systems.

### People Change

More than 50 years ago, Carnegie (*5*) considered that the relationship between changer and changee was critical to successful change. He considered that changes in feelings and attitudes were prerequisites to voluntary changes in overt behavior. His model is manipulative because he built good relationships and then bargained with them.

A more acceptable method, used by Coch and French (*6*) to introduce changed methods in a pajama factory, was to use group methods. They provided the opportunity for need satisfaction, then directed the group forces towards the desired change. In both these approaches, the power is in the hand of the manager wanting change; they are similar to the Theory X authoritarian management style outlined by McGregor (*7*).

Later researchers (*8*) placed great emphasis on the need for collaboration between changer and changee for change to take place. This attitude parallels McGregor's concept of the participative style, Theory Y; it provides a major shift from the powerful superior dealing with weak subordinates to a more equal balance of power.

In a development of Theory Y, individuals are encouraged to set their own objectives. This has been termed "management by objectives" (*9*) and demands a greater degree of leadership and competence from the individuals' managers (*10*) than the more conventional Theory X style.

### Task Change

Road maintenance organizations seldom undergo substantial changes in the tasks done. But they should have built into their task system a strategy of continually examining current methods to identify where improvements can be made. Improvements may require training (an important element in the people system) and changes in technology.

### Review

The message from the research and emerging management practices is clear. In order to ensure its success, an organizational change must

1. Be the result of collaboration between the promoters and the final users, and
2. Assist users to carry out their work, improving the quality of their decisions—not making them.

This attitude is in stark contrast with the popular perception of some MMSs, for example. Supervisors at every level and some engineers have privately expressed the fear that an MMS will increase their paper work, computerize their planning function, and remove the need for them to use experience and judgment, and that it will be developed without full consultation with them.

Unfortunately, their misgivings have some foundation in (even recent) history. Systems have been introduced with little or no input from the users and insufficient thought to the best method of implementation. In addition, some systems reduce rather than enhance the quality of decision making, because they take it away from the practitioner. Although such a system may have some relevance in a factory setting where most conditions remain constant, conditions in roading are so variable that decisions must be made by competent, experienced people.

These variations can range from political forces altering budget decisions to changes in planned repair procedures once the road is opened up. Each variation requires a person to have the authority and ability to make judgments aided (but not constrained) by a management system. The system must be considered as an aid to better decision making.

In the past, the importance of the ownership of change may have been underestimated. Researchers have concluded: "A great deal of change seems to demand participation, especially at the implementation stage" (*11*); "When ideas are a person's own, they are much more likely to be translated into meaningful practices than when they are the suggestions of an outside expert" (*12*); "A cardinal value of organizational development is the introduction of shared or distributed influ-

ence—based on the assumption that people will be more committed to objectives if they have participated in establishing them" (*13*).

## ORGANIZATIONAL CHANGE—PREVIOUS PRACTICE

In the past, the changes to management information systems have most often been "head office driven", to meet corporate objectives. Frequently, foremen and overseers have reported that they have been required to provide regular information for the latest corporate-inspired system.

Commonly they report a lack of feedback. Recently an overseer reported that he had been asked to provide daily information for a new management system. He continued this extra effort for 2 years, but received no indication that his effort was appreciated nor had he received any benefit from it. So he stopped, and no one has spoken to him about it since—either to ask him why he stopped or to ask him to continue. His reasonable perception is that his effort was of little value to the organization. Naturally, he was a reluctant starter when the organization introduced a new management system.

This experience is by no means unique. Supervisors who are affected feel that such experiences badly reflect on the organization and managers who initiate unneeded systems.

Quite reasonably, system development has concentrated on the need to provide certain management information. But even the best system is of little long-term value if the people using it (or providing information for it) do not understand its importance to them. So an important part of system design should be its introduction and long-term support.

## ORGANIZATIONAL CHANGE—THE PROCESS

The overall model for making change need not be complex. An example relevant to roading is shown in Figure 1.

1. The first step is to identify the change to be made; e.g., the organization wants to introduce an MMS.
2. The systems and procedures inherent in the change must be developed, and appropriate standards set against which the outcomes of the change can be evaluated.
3. The work affected by the change must be identified and itemized in sufficient detail that the individual activities and their constituent work methods and skills are specified. This enables work standards to be set. Changed work performance can be measured against these standards.
4. Finally an implementation process should be designed to introduce the change to the people affected by it and to ensure that the change continues as part of the new corporate culture. This process must include training to ensure that the people affected by the change will feel comfortable with it and will be able to use it to their advantage.

The mechanical processes of system design are well understood and have been practiced for many years. However, recent work has attempted to improve the human side of

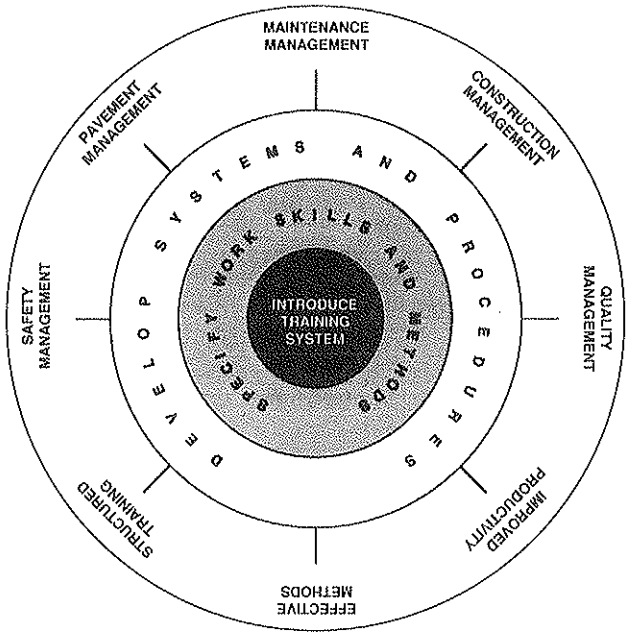



**FIGURE 1 Developing structured improvement.**

introducing the change, by influencing the people who will be affected by it and making the change process more acceptable to them.

Change in people is not as simple as changing direction in a vehicle. Change can be mentally uncomfortable, even threatening to a person. It is not a mechanical process like turning a steering wheel or programming a computer; it works on the human mind, so the human reaction to change is tempered by nonmechanical, psychological factors such as attitude and motivation.

Merely handing out the operation manual or sitting someone in front of a computer terminal may not instantly produce the desired change. People must perceive that the change will benefit them and, preferably, experience that benefit as part of their training. Then they may make the mental adjustment necessary to accept the change.

Thus, the training to support a desired change may have to concentrate more on changing attitudes than on the new system or procedure itself. The training provided is likely to be a critical part of the change process, making the difference between success and failure.

## DEVELOPING AN MMS

The Queensland Department of Transport is developing an MMS for its roads. Development is being undertaken at corporate level, but the organization is keeping its works districts involved and informed. Districts are subjecting parts of the system in trial during development to ensure they will work and are user friendly. Trial districts are gaining a sense of ownership of the final product because they have contributed to its development. The system has been described by Caldwell (*14*).

### Activity Descriptions

A basic part of the system is a series of activity descriptions—definitions the tasks used to repair defects. These are the activities that the system manages.

Each description includes the method, materials, plant, and workers required as well as other planning information such as intervention level, restoration standard, and average daily production. The descriptions set out the quality and safety standards for the work and indicate the skills needed by the workforce.

Rather than write these descriptions corporately and without reference to the final users, the organization asked a cross section of maintenance engineers and supervisors from a number of districts to do this in a 3-day workshop. Another open objective of the workshop was that participants should have a complete understanding of the MMS as it affected them, to understand its advantages to them so that they would accept ownership of it.

Subsequent evaluation indicated that the workshop met its objectives, so that a nucleus of the organization was willing to use the MMS. The workshop was designed and run using training methods.

### Work Methods

In particular, the workshop gave activity descriptions on the basis of practices generally acceptable in the districts represented. However, the corporate descriptions are only a stepping stone; each district is encouraged to review them and alter them where necessary to ensure they represent the best way of doing the work in the district. In this way, responsibility for setting and maintaining standards (within wider guidelines) devolved to the districts where the work is actually happening, and the districts develop a sense of ownership of these standards.

Most of the activity descriptions represent changes to existing methods, but nothing will change until staff are trained in the new methods. Once a district has written its own activity descriptions, it can begin to train its staff in the new work methods. This training effort alone should provide significant improvements to standards and several other benefits to the staff and organization (*15*).

The final phase of the work, introducing the system to the districts, will also involve staff traininig in the use of the system itself. It will come at a time when districts will have seen some of the benefits from the skills training. These benefits will have reinforced their belief that the system as a whole will also benefit them.

### A Champion

The development work for the MMS and the associated supervisor and worker training has been strongly supported by the districts. But even then there must be a word of caution. The initial commitment to this change must continue actively in each district, otherwise the change will have minimal long-term effect. Even though a change may have been successfully introduced, there are many forces working against its survival. For example, the champion in one district may retire, leaving no one there to take up the role; priorities alter (with a new boss or a run of bad weather), and the MMS may be placed in abeyance temporarily, but never restarted.

The best insurance against loss of momentum is the appointment of a champion, someone (in a large organization at least) whose sole function is to keep commitment to the change high even when there are many other competing changes in the organization. In New Zealand, a system of worker training was introduced into the Ministry of Works and Development (MWD) in 1986 (*16*). Since then, the MWD has been restructured into a commercial SOE and undergone subsequent restructuring.

With such extreme changes in such a short time it would be surprising indeed if the training system had survived, but it has. The MWD assigned a staff member to work on the introduction stage, and initially that person's sole responsibility was to keep the system working. This meant a continual round of visits to managers, acquainting new managers with the training, and retaining the commitment of existing managers. Training in instruction techniques has been provided for new overseers and foremen, and it is being extended into other interpersonal skills required for roading supervisors.

A similar appointment has been made in Queensland, to continue the training and motivating work needed to ensure the long-term success of this significant change.

## SUPPORTING AN MMS

On its own, an MMS enhances the quality of decision making in the organization—it cannot improve actual work performance in the field. Such improvement requires the best appropriate methods to be decided and workers to be trained in them. Supervisors must also be trained in those methods so that they can properly supervise the work. In addition, supervisors must be able to plan and organize work (where that is their responsibility) and to lead the work group on the site (if that is their role).

In Queensland, a model was developed outlining the material and training required to ensure that introduction of the system would actually improve field performance. This model is shown in Figure 2. The argument is that two things are required before there can be effective planning of work: the organization must not only specify the work activities but also train supervisors in planning and organizing work. If either the specification or the training is omitted, there cannot be effective planning.

Each activity description includes the best appropriate method for doing the work in each particular district. Supervisors require leadership and site supervision skills to control and motivate workers before those method specifications can result in effective work.

The methods are broken down into their constituent steps and skills. Separate skill instruction guides are being written to cover the large range of skills required by people in the department. Supervisors must be trained to instruct their workers so that these skill specifications can be used to improve the safety and quality of the work being done.

The result of all these steps combined will be improved productivity. These specifications and the training will also

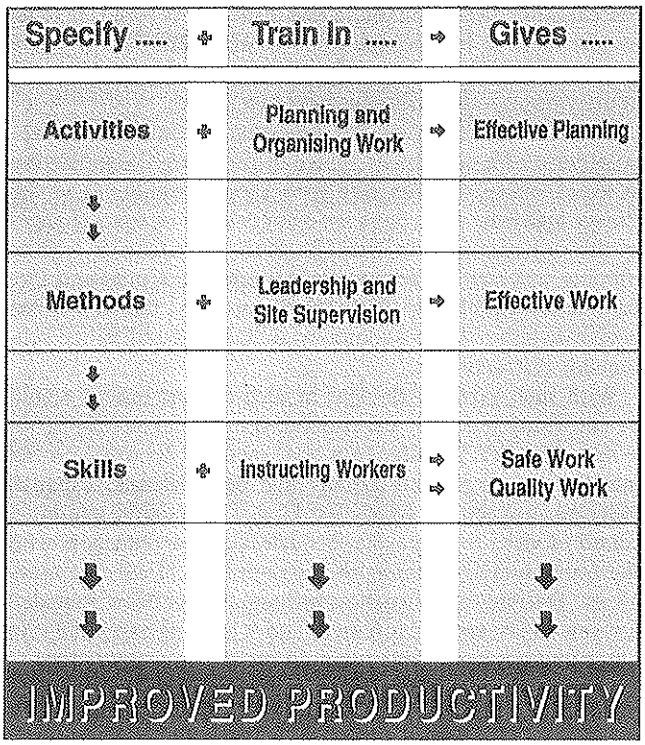



**FIGURE 2 Improvement model for Queensland.**

allow the department to meet its statutory obligations for safety under the State's Workplace Health and Safety Act and to implement a policy of total quality management.

### Best Local Methods

Part of a road maintenance worker training project in New Zealand (*17*) was designed to specify methods that represented the best practice nationally. Comments on draft methods were sought from representative samples of practitioners throughout the country before final recommendations were written.

It was recognized that there would be regional variations because of varying local conditions, maintenance standards, and available resources. So engineers and field people were encouraged to work together to develop their own local methods on the basis of the national recommendations. A similar approach has been used in Queensland. Also in Queensland, works districts have been encouraged to write skill instruction guides themselves. These are forwarded to the corporate office for checking and coordination, then distributed to all districts.

A variation is being subjected to trial in the Roads and Traffic Authority of New South Wales. Here a technical team is being formed in each works office to determine the best appropriate local method and skills. Each team has about five members, representing the works engineer, foreman, ganger, leading hand, and worker levels.

They will subject new methods or procedures to trial in one crew first (generally in a crew with a representative on the technical team so that it has some ownership of the method it is testing). When the method has been tested, all affected supervisors will be trained in it; then, they in turn will instruct their own workers.

In New South Wales, work practices must comply with the state occupational health and safety legislation, whereas a recent Commonwealth act requires employers to provide training equivalent to 1 percent of employee's time. The technical team work and the ensuing training will allow compliance with these two requirements and provide an excellent base for the MMS being introduced throughout the Roads and Traffic Authority in 1991.

### Appropriate Training for the Organization

Most training is designed to meet the needs of the individual. Its objectives are expressed in terms like " . . . each participant will be able to . . .". In other words, the training should change the individual's potental to perform, but it does not consider how that potential can be harnessed to actually change performance in the workplace. This step presumably is left to the individual's manager. Although this procedure may be good for the individual participant, it does not necessarily help the manager whose objective may be to change organizational performance.

An extreme example of how this approach can go wrong happened in a shoe factory. The owner-manager sent four of his supervisors to a personal motivation course. The course motivated them to the extent that two supervisors resigned from their job immediately and the other two within a year. Unfortunately for the owner, the course was successful on an individual basis, but it was not related to the requirements of his organization. He was unable to change his organization to take advantage of the supervisor's new level of competence.

So what should the owner have done? The second time around, he engaged a training consultant who analyzed the supervisory performance problems in the factory, designed an approach to overcome those problems, then worked with all the supervisors and managers to implement the approach. This approach included training that improved individual performance and attitudes and built a stronger management team. The team openly discussed long-standing problems within the factory. Most of these related to poor communication between supervisors and managers, and the stronger team quickly identified solutions for them.

This example indicated that a trainer and manager could work together to provide in-house training, designed to make specific improvements within the organization. In fact, of the several shoe factories that had operated in that city in the early 1980s, the factory just described became the sole survivor of their common economic problems.

Training can be most powerful when it is used to support planned changes in the organization. This could include training someone for promotion and training a group in new methods or procedures. It is particularly powerful when used as a basis for the attitude and skill changes required for the successful introduction of an MMS.

### Appropriate Individual Training

Training must be presented in a manner appropriate to the individual learners. In one organization, engineers reported

their apprehension at having to present lectures on "foremanship" to supervisors. The lectures (called "seminars") were presented to a strict format and timetable with the use of slides or overhead transparencies. The engineers were nervous about standing in front of a group and lecturing to them.

What the engineers did not realize was that the supervisors were equally as uncomfortable. As a general rule, practical, outdoor people do not enjoy listening to lectures for an hour or more inside a room. Participative training is the most appropriate form of learning for adults. The lecture method, by which most engineers have been taught, is probably the least appropriate. Certainly, changes in attitude and manual skills are best developed by actual experience.

Learning to drive a motor vehicle is a good example. People can learn the theory—a knowledge of the road traffic rules—by reading a book; they have to be in a vehicle to learn and develop the necessary driving skills. Also, driving experience and practice are the only way people can fully appreciate the application of the road rules and become confident, safe drivers. Unfortunately, the supervision of drivers is too loose to ensure that people continually maintain standards and attitudes conducive to safe driving.

Exactly the same procedure should be followed when learning the skills and developing appropriate attitudes to work productively. The only difference should be the standard of supervision after training.

The highly participative approach normally used by the author comes as a shock to many participants, who are used to trainer-driven training. In this approach, the participants set the times for the training, spending as long discussing a subject or doing an exercise as they need; they make their own decisions on matters under consideration. The trainer's role is to guide the learning along logical paths, to stimulate discussion, to set exercises suitable for the learning, and to ensure all participants are gaining from the experience.

### Testing the Training

Another important feature that receives continual commendation from training participants is a follow-up assessment visit. The training provider visits participants and their managers after the training to assess the participants' competence in their workplace. Typically, this involves a formal assessment of the participant using the new skill. This follow-up visit allows the trainer to assess more clearly that the training objective has been achieved, it allows individual coaching where needed and gives managers and participants the opportunity of discussing how the new skill can be used in the workplace.

It allows the provider to assess if the training was appropriate for the organization and the individual learners.

Field people in particular gain visibly from this visit. For most of them, it is the first time they have been tested at work. Often their record at school was one of failing, so passing the assessment gives them a real sense of achievement and confidence in their ability to learn. It may be remembered that some school examination results are scaled so that 50 percent of examinees will fail. The objective of training should be that all will pass; if they do not, the training, not the learner, should be reexamined.

Most participants enjoy this positive approach to relevant, participative, and tested training; and they begin to request more of it.

## FIELD STAFF TRAINING

Training for field staff can be divided into four areas: worker training, operator training, supervisor technical training, and supervisor interpersonal training. Each has different requirements, requiring different approaches to ensure success.

### Worker Training

The worker training system developed in New Zealand (*17*) was designed specifically to meet the requirements of road maintenance. Instruction is given on the job (not in a classroom), using actual work; instruction is by the workers' own supervisors; and the methods taught have been previously selected by local supervisors and engineers. This process ensures that workers (and their supervisors) learn the best methods appropriate to their conditions, in a setting in which they are at ease and from a person with whom they are in daily contact.

Instruction of workers by their own supervisors improves the relationships between them, builds a better work team, and removes the chances of misunderstandings introduced by a third party. The related principle is that supervisors who do not have the ability or time to instruct workers also do not have the ability or time to properly supervise their work.

A significant advantage of this system is that it uses existing resources within the organization, so it is exceptionally cost-effective. Other benefits as perceived by users have been recorded (*16*), but no quantitative benefits have been studied. In 1989, the training system was introduced into the Department of Transport (Queensland) and Roads and Traffic Authority (New South Wales). In recent work, Thomas and Anderson (*15*) have quantified the effects of improved methods and training in pothole repair. Their study found that introducing a standard repair procedure decreased the repair cost per ton by 44 to 48 percent when taken over the life of the repair. Training then decreased the cost of the standard procedure by 45 percent. The total savings was nearly 70 percent on the original "throw-and-go" method.

The longevity of the repairs improved from 60 to 899 days (average). This increase not only yields dramatic reduction in cost but also correspondingly boosts workers' sense of self-worth. The defect is repaired only once, not six times a year.

In addition, the study found an exceptional saving of resources, with the amount of material placed in potholes each year reduced by 62 percent in 5 years. The study concludes that "Training programs and the proper selection and standardization of equipment can significantly reduce overall costs. . . . The factors that have the greatest influence on total repair costs are repair longevity (procedures), daily production, and crew deployment practices."

This argument for better methods and work skills, combined with better management, is powerful. With such savings to be made, the question is not "How can we afford training and MMSs?", but "How soon can we get started?" The answer must be "Now."

### Operator Training

One of the important features of the worker training in New Zealand (*17*) was that workers should be trained by their own supervisors. This principle cannot be translated to operator training. The operation of plant equipment such as graders, computerized bitumen sprayers, large rollers, and others, requires specialists. It is unreasonable to expect that supervisors should be skillful operators of the wide range of plant equipment used in road maintenance. However, they must know the standards of work expected from a competent operator and the correct methods of using the plant equipment to achieve those standards.

In the past, there have been few successful attempts in New Zealand to provide operator training. This fact is rather strange considering the replacement and running costs of the plant equipment and its potential to create or ruin a good road. Although all bitumen sprayers used on state-subsidized sealing work in New Zealand must be certified every second year, the operator whose ability influences the standard of the work faces no such requirement.

Operator training is best carried out by a specialist operator who also understands the principles of instruction. The person could train operators on the job, ensuring that both they and their plant equipment come up to acceptable standards. The other way is for the operators to come together for a training workshop. In this way, they can compare techniques and learn from each other as well as the specialists.

Several years ago, a Gold Grader Competition and Training Workshop was conducted. The event took place over 2 days, with training not only for beginners but also for experienced operators in operating skills, machine maintenance, safety, and correct use on the road. As a follow-up, the participants requested (and received) a 1-day workshop on maintenance. Most participants stated that it was the first organized training they had ever received. No attempt was made to evaluate the success of the training, but the participants displayed an increase in skill level during the workshop.

Once again, the training was designed to meet the requirements of the participants and their organizations.

### Supervisor Technical Training

In addition to the basic worker skills, road maintenance supervisors must have a second level of knowledge and skill. They must understand the basic requirements of good road maintenance and be able to apply these in the field. An important part of this second level is the inspection of roads to identify defects and their causes, then to specify appropriate remedial measures.

In New Zealand, technical workshops for supervisors are typically run over 2 days. The workshop content is 50 percent discussion and 50 percent field work. It has four parts: discussion of some basic principles, making a road inspection and deciding the work to be done, doing that work, and a final summary session. The most frequently run workshops are Maintaining Sealed Pavements, Maintaining Unsealed Roads, and Maintaining Drainage for Roads. They are based on National Roads Board publications (*18–20*).

At an individual assessment visit later, the trainer observes each participant inspecting a road and specifying repair work. Then there is a discussion with participants and their managers on the results of the assessment and on ways to ensure that the new skills will be used. Informal evaluation of results (subsequent discussions with participants and their managers) indicates that the workshops provide a useful learning experience and that the learning is transferred to actual performance in the field.

A similar approach has been adopted for the training of engineers, technicians, and senior supervisors in chip seal design and supervision.

### Supervisor Interpersonal Training

Unfortunately, the project to provide interpersonal skills training for road maintenance supervisors in New Zealand has made slow progress since mid-1986. The concept was that training would be provided through a series of supervisor development workshops on the subjects of supervising workers, planning and organizing work, personnel management, and safety. Participants would leave each workshop with exercises that they would do with their managers' supervision. A coaching workshop was proposed so managers could better understand and carry out their coaching role for the workshops.

If the managers are not involved in this way, there is no guarantee that the new competences will be exercised usefully for the organization. Indeed, one organization has reported that its supervisors (who have been trained by an internal consultant) have gained interpersonal skills and understanding superior to their managers, thus weakening (rather than strengthening) the relationships between them. In the worst possible situation, people were encouraged to leave the organization because they had outgrown it (as in the shoe factory noted previously).

The concept of the manager as a coach is part of modern management philosophy (*21*). This concept was used as part of the system design for the supervisor development workshops. Also, it was expected that by being involved managers would better understand (and themselves gain) the competences learned at the workshops. This understanding would enable them to make better use of those competences in the workplace.

It is critical for the success of interpersonal training that managers are involved, not only in developing training that is appropriate to their staff, but also in working with their staff to ensure that the new competences are actually used.

### Management Training for Engineers

In New Zealand during 1989–1990, local councils (which maintain much of the country's roading system) underwent massive change. Over 230 existing councils were amalgamated into just 75. Positions such as Works Engineer and Roading Engineer were typically redesignated Works Manager and Roading Manager. The emphasis in such management positions is now on management more than technical qualifications and experience.

Several organizations in New Zealand provide training for managers. When it decided to develop industry-specific training for supervisors and workers, the Local Government Training Board considered that management training need not be specifically designed. Management skills are more generic than industry-specific, and management requires people with greater conceptual skills than supervision, so they must be able to transfer skills learned in the context of one industry to another.

Management of change is becoming an increasingly demanded skill for managers; it is particularly important in relation to the introduction of an MMS. Managers must be able to create an atmosphere in which people will change entrenched attitudes towards work management and people (often unused to training) will learn new skills of planning, organization, and leadership.

### Technical Training for Engineers

Publications and conferences provide opportunities for roading engineers to improve their technical knowledge and glean the results of research and innovative practices on a reasonably frequent basis.

However, one of the difficulties with research (and conference proceedings) is converting them into actual changes in the workplace. Often, there are long delays between an idea or the results of research and their application. Lay (*22*) noted "It has commonly taken about a decade to develop it (an idea) to a useful form."

The National Roads Board (now part of a new Transit New Zealand) has recently encouraged researchers and others to present training workshops or seminars that are based on their work. This policy means that the information is disseminated more positively and faster than before. It also means that there can be two-way communication between researchers and a wide range of practitioners, perhaps improving the knowledge of both groups.

In 1988, the National Roads Board supported the design and presentation of training workshops on the maintenance of unsealed roads. This policy was proposed to disseminate the information published (*19*) as a result of research into acceptable practices. The project was undertaken by Ferry, Armitage, and National Roads Board staff.

The research had recommended fundamental changes in the design, construction, and maintenance practices commonly used throughout New Zealand for unsealed roads. Two workshops were designed to cover the range of personnel levels identified as contributing to the decision to change organizational practices. There was a management workshop for policymakers, managers, and technical staff, and a maintenance workshop for field staff. As part of the project, a video was made to help change attitudes towards unsealed roads practice.

The management workshop was 1 day of group discussions and practical exercises to get normally desk-bound people involved in the consequences of their decisions. The management workshop was run about 2 to 4 weeks before the maintenance workshop in each particular location so that managers could confirm and discuss an action plan for changing their practices before they sent field staff to the second workshop.

The maintenance workshop was for supervisors and grader operators. It followed the format established for similar technical workshops. On the first day, participants discussed principles, then inspected some roads for faults; on the second day each group worked on a length of road, bringing it up to an acceptable standard.

The final stage was a follow-up visit by the researcher. He assessed the competence of the field staff in inspecting a road and specifying remedial work for it. He then discussed their action plans and proposed practices with the field staff, engineers, and managers. It is too early to judge the success of this approach in changing existing practices. The initial reaction was positive, but it is the long-term results that are important. A workshop for engineers titled "Economic Analysis" has also been based on recent research work (*23*).

This type of approach can decrease the time delay between research and its effect on practice. It is a positive approach for providing technical training for engineers and changing field practice.

## CONCLUSION

Change can present physical and emotional challenges to the people affected by it. Increasingly, managers are recognizing the importance of involving those people in the development and introduction of the change so they will develop a feeling of ownership of the change rather than resent it as an imposition.

Management information systems will enhance the quality of decision making if properly developed. However, better decision making with a road MMS does not unequivocally lead to improved performance on the road. First, the best methods, appropriate to the circumstances, must be determined; then workers and supervisors must be trained in those methods; supervisors must be trained in site leadership and planning (according to their particular responsibilities); and managers must be able to support the required changes in personnel responsibilities and attitude before the full benefits of an MMS will be realized.

All of these processes certainly do not happen overnight—they require long-term effort, with strong management support to persevere.

The research and experience considered indicate that training can be used to advantage in a roading organization: to support organizational changes such as installation of MMSs, to improve the technical and managerial competence of supervisors and managers, and to give workers the basic skills needed for the organization to improve its work performance.

## REFERENCES

1. H. J. Leavitt. Applied Organizational Change in Industry: Structural, Technological and Humanistic Approaches. In J. G. March, *Handbook of Organizations*. Rand McNally, Chicago, 1964, pp. 1144–1145.
2. F. J. Roethlisberger and W. J. Dickson. *Management and the Worker*. Harvard University Press, Cambridge, Mass., 1939.
3. T. J. Peters and R. H. Waterman. *In Search of Excellence*. Harper & Row, New York, 1982.
4. F. W. Taylor. *Scientific Management*. Harper & Row, New York, 1947.

5. D. Carnegie. *How to Win Friends and Influence People.* Simon and Schuster, New York, 1936.
6. L. Coch and J. R. P. French, Jr. Overcoming Resistance to Change. *Human Relations*, Vol. 1, 1948.
7. D. M. McGregor. *The Human Side of Enterprise.* McGraw-Hill, New York, 1960.
8. R. Lippitt, J. Watson, and B. Westley. *The Dynamics of Planned Change.* Harcourt, Brace & World, New York, 1958.
9. P. F. Drucker. *The Practice of Management.* Heinemann, London, 1954.
10. D. M. McGregor. The Human Side of Enterprise. In *Adventures in Thought and Action: Proc., 5th Anniversary Convocation of the School of Industrial Management*, Massachusetts Institute of Technology, Cambridge, Mass., 1957, pp. 23–30.
11. R. M. Kanter. *The Change Masters—Corporate Entrepreneurs at Work.* Unwin, London, 1985.
12. D. Katz and R. L. Kahn. Organizational Change. In J. M. Thomas and W. G. Bennis, *The Management of Change and Conflict.* Penguin Books, Middlesex, England, 1972.
13. G. L. Lippitt. *Implementing Organizational Change.* Jossey-Bass, San Francisco, 1985.
14. G. Caldwell. Development of a State Wide Maintenance Management System for Queensland, Australia. *Proc., 6th REAAA Conference*, 1990.
15. H. R. Thomas and D. A. Anderson. Pothole Repair: You Can't Afford Not To Do It Right. In *Transportation Research Record 1102*, TRB, National Research Council, Washington, D.C., 1986.
16. T. G. B. Armitage. Training for Improved Productivity in Road Maintenance. *Proc., 14th Australian Road Research Board Conference*, Part 5, 1988.
17. T. G. B. Armitage. Training for Road Maintenance Workers. *Proc., New Zealand Roading Symposium*, 1983, pp. 187–202.
18. A. G. Ferry. *A Manual of Pavement Repairs.* National Roads Board, Wellington, 1981.
19. A. G. Ferry. *Unsealed Roads. A Manual of Repair and Maintenance for Pavements.* National Roads Board, Wellington, 1986.
20. T. G. B. Armitage. Drainage. Part E in a series *Maintenance Manuals for Workmen.* National Roads Board, Wellington, 1988.
21. A. X. Deegan. *Coaching: A Management Skill for Improving Individual Performance.* Addison-Wesley, Reading, Mass., 1979.
22. M. G. Lay. Australia's Road Inheritance. *Proc., 14th Australian Road Research Board Conference*, Part 1, 1988.
23. I. H. Bone. *Economic Appraisal of Roading Improvement Projects.* National Roads Board, Wellington, 1986.

*Publication of this paper sponsored by Committee on Low-Volume Roads.*

# New System for Preventing Reflective Cracking: Membrane Using Reinforcement Manufactured on Site (MURMOS)

J. SAMANOS, J.-C. ROFFE, H. TESSONNEAU, AND J.-P. SERFASS

The first experiments on the totally new process discussed in this paper were conducted in France in 1988. Since then, almost 1 000 000 m² have been successfully laid. The technique involves in situ creation of a reinforced membrane composed of a layer of elastomer binder onto which continous threads are sprayed immediately after it is laid. The threads interweave to create the reinforcement. This paper describes the process, gives the mix designs used, and introduces the machine that was especially designed to manufacture and lay the reinforced membrane. The results of laboratory tests conducted by one of the laboratories of the French administration and the Belgian Road Transport Research Center are also presented and discussed. These tests show that this process is one of the most efficient processes for preventing reflective cracking. Some examples of its application are mentioned. Its principal advantages over processes using shop-manufactured materials (geotextiles or geogrids) are outlined.

The first experiments on this treatment process were conducted in France in 1988. Since then, almost 1 000 000 m² have been successfully laid. The technique involves in situ creation of a reinforced membrane composed of a layer of elastomer binder onto which continuous threads are sprayed as it is laid. The threads interweave to create the reinforcement. This composite system is laid on the cracked pavement by a special machine before laying of the wearing course.

## PROBLEM

Cracking of pavements is a problem that has plagued roadbuilders for many years. Two types of cracking are particularly worrying.

- Fatigue cracking. Fatigue cracks appear primarily in flexible or conventional pavements and are caused by repeated tensile stresses in the bituminous materials as a result of traffic. They start at the base of these materials and gradually rise to the road surface where they form disorderly network (alligator cracking). Fatigue cracks are a sign of advanced breakdown caused by inadequate design, or a symptom betraying the fact that the road has reached the end of its service life. Maintenance or reinforcement is therefore necessary.
- Reflective cracking. Reflective cracks affect semirigid pavements and originate as soon as the cement or pouzzolanic binders used begin to set. The speed at which they rise to the surface depends on the type and thickness of the materials that form the upper layers, the amplitude of temperature variations, mix designs, etc. Reflective cracks generally form regular patterns (transverse cracks every 5 to 8 m), and although they are not a sign of serious deterioration, they nonetheless call for curative treatment.

To these cracks must be added the joints between concrete slabs of rigid pavements, and the fine but numerous cracks that run in all directions across continuously reinforced concrete pavements.

Whatever their origin, these cracks are waterproofing disorders. Water may infiltrate into the ground and significantly shorten the lifetime of pavements on soils that are not naturally self-draining or that may be affected by frost heave. Furthermore, traffic on cracked semirigid pavements can aggravate cracking, which eventually produces an unsightly wearing course. On concrete pavements, improper treatment of the joints can lead to relative movement of slabs, and eventually to rupture.

The problem is so important that in France, in 1983, the Ministry of Transport launched a contest for innovative techniques designed to prevent reflective cracks. Such processes were classified priority No. 1 by the Department of Road Transport in 1988 (*1*).

These processes concern a wide market, that of all semirigid roads built to date (curative applications) or currently being built (preventive application), as well as concrete pavements, particularly on densely trafficked roads. In France, the potential market for this process is approximately 3 000 000 m² per year for new roads, and about 30 000 000 m² per year for existing pavements.

## FRENCH EXPERIENCE IN TECHNIQUES FOR PREVENTING REFLECTIVE CRACKING

Current French know-how in techniques for preventing reflective cracking was inventoried recently (*2*).

Several processes slow down the rate at which existing cracks rise to the surface:

- Sealing before laying of chip seals or asphalt overlays. A binder, generally modified, is applied along the crack and to about 10 cm on either side, and is then gritted with sand before being covered with chip seal or asphalt overlay.

J. Samanos and J.-C. Roffe, SCREG Routes, B.P. 100 Guyancourt, St. Quentin en Yvelines, France. H. Tessonneau, SCREG Sud-Est, B.P. 85, Venissifux, France. J.-P. Serfass, La Recherche Technique d'Entreprise, B.P. 43, Bonneuil sur Marne, France.

- Stress-absorbing membrane interlayer (SAMI). A binder is applied to the entire cracked pavement, then chippings are spread, and finally an asphalt overlay is provided as a wearing course. The binder is generally modified with polymers or rubber powder (from tires).
- Geotextiles impregnated with bituminous binder. This process involves rolling a manufactured geotextile onto the cracked surface, to which it is made to adhere with ample quantities of binder, and then applying an asphalt overlay.
- Geogrids impregnated with bituminous binder. As in the previous case, the geogrid is made to adhere with ample quantities of binder and is surfaced with a wearing course.
- Double-layer overlay. Double-layer overlays contain a 1.5- to 2.0-cm-thick first course of 0/4-mm sand in a 9 to 10 percent bitumen mix (the bitumen may be modified), topped with a 4- to 6-cm-thick course of asphaltic concrete (possibly using a modified binder).

Several years of experience in France have given the following results:

- Sealing of cracks is slow and laborious, the final appearance is not pleasing, and impermeability is only temporary. The wearing course near the cracks may soon start to deform.
- Thick SAMIs have been deemed to give not very satisfactory results on the jobsites where they have been applied. (Laboratory results are nevertheless good.)
- Impregnated goetextiles reinstate the impermeability of the pavement, but cracks nevertheless rise quickly through overlays. (These techniques have been little developed because of they are so slow to apply, being particularly difficult in bends. Folds and overlapping of strips do nothing to improve the efficiency of the system.)
- Geogrids have been judged to have given rather negative results;
- Double-layer overlays containing a binder-rich mortar covered with a thin asphalt overlay are currently the most efficient solution, but are relatively expensive.

There is therefore seen to be room for a process that combines the qualities of efficiency, moderate cost, and high speed of application.

## DESCRIPTION OF PROCEDURE

### Principle

The principle (see Figures 1 and 2) consists of a composite system formed by simultaneously spraying a bituminous binder and continuous organic threads that interweave to create reinforcement, and then spreading chippings onto this membrane before applying the wearing course.

### Formula

The binder is an elastomeric bituminous emulsion applied at a rate of 0.9 to 1.8 kg/m² (equivalent to a residual binder content of 0.6 to 1.2 kg/m²).

The polyester threads are sprayed on at high velocity. The multistrand threads ranging in size from 140 to 500 decitex (1 decitex is the weight in grams of 10 000 m of thread) are generally spread to a density of 100 g/m², with variation between 40 and 120 g/m² possible.

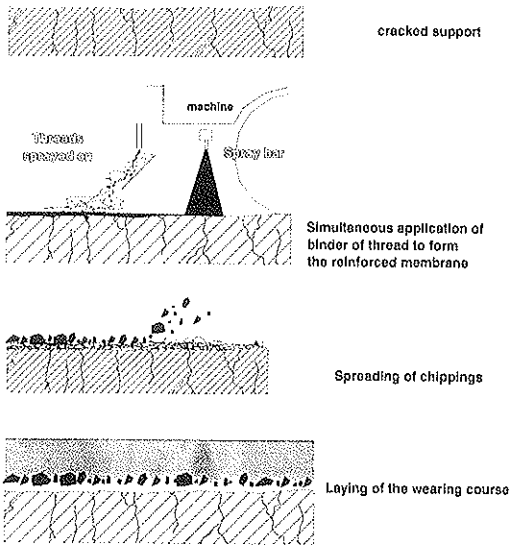



**FIGURE 1 Steps in MURMOS application.**

Chipping (6/10 or 10/14 mm) coverage is about 5 to 10 L/m².

There are several options for the wearing course: e.g., Glg-type chip seal (course chippings, binder interlay, and fine chippings) or conventional asphalt concrete, but the most frequent wearing course is fiber-based asphalt, which, because of its high binder content, further improves the performance of the membrane.

### Filaflex Machine

This new MURMOS process required the development of a special machine. Once all the parameters for application of this concept had been defined with a sprayer-drawn prototype that sprayed threads under pressure in a 1-m-wide strip, two high-throughput machines were built. They were in fact slightly different, but the following equipment was common to them both:

- A roadgoing trailer with two steering axles that carries all the subunits required to make the membrane;
- A power source for the hyraulics units and the pump feeding the thread injectors;

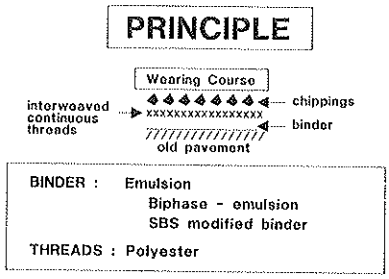



**FIGURE 2 Schematic of MURMOS technique.**

- A built-in binder spray bar (extendable from 1.25 to 3.70 m) for application of a uniform membrane at a constant rate, onto which the threads are sprayed immediately before spreading of chippings and laying of the wearing course;
- A 6 000-L emulsion tank (on one machine only);
- Two independent thread-spraying units, each including
  - —A 900-kg reel of thread on a 1.5-m-long, telescopic arm that can be extended overboard to obtain a total width of application of 3.70 m (with both reels);
  - —A tensioning system that unwinds the 900 threads from the drum; and
  - —A thread spray bar (seven injectors on one machine, nine on the other) placed immediately behind the binder nozzles and connected to the pump by flexible hoses. As for the binder spray bar, the thread spray bar can be extended laterally to vary the width of application anywhere between 1.25 and 3.70 m.

The layout of the Filaflex machine is shown in Figure 3.

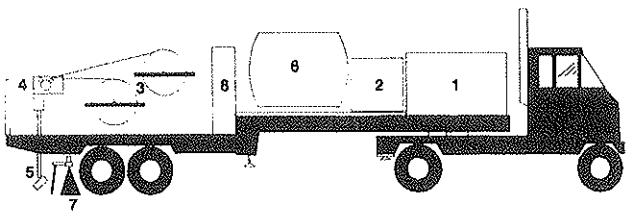
## LABORATORY ASSESSMENT

Different structures were tested on a variety of devices to identify the most suitable systems.

### Tests Conducted at the Regional Laboratory of Ponts et Chaussées in Autun, France

*Principle Behind the Test*

The tests (*3*) involved examining the speed at which a crack rises through a composite pavement including (a) a system intended to delay the reflection of cracks and (b) a wearing course (*4*,*5*).




1. Engine driving hydraulics unit.

THREAD UNIT
2. Pump unit generating continuous flow of compressed air.
3. Two 900 kg reels, each wound with 900 threads and mounted on 4 bearings; they are gradually braked as the quantity of thread on the reel diminishes; this system is controlled by an ultrasound detector.
4. Tensioning-roller system for unwinding the threads: 2 smooth stainless-steel rollers, and 1 hydraulically-controlled rubber roller whose rotation speed is linked to the forward speed of the unit.
5. The thread is sprayed onto the road by a series of 18 polished stainless-steel nozzles with a tubular extension and deflector plate. The nozzle system can be moved laterally.

BINDER UNIT
6. 6,000 litre binder tank.
7. Built-in spray bar. Can be extended outboard.
8. Crane for loading reels of thread.

Dimensions: L = 16 m (trailer and tractor unit) W = 2.5 m H = 4 m

**FIGURE 3 Description of Filaflex II machine.**

TABLE 1 RESULTS OF LABORATORY TESTING OF MURMOS SAMPLES OF VARIOUS COMPOSITION

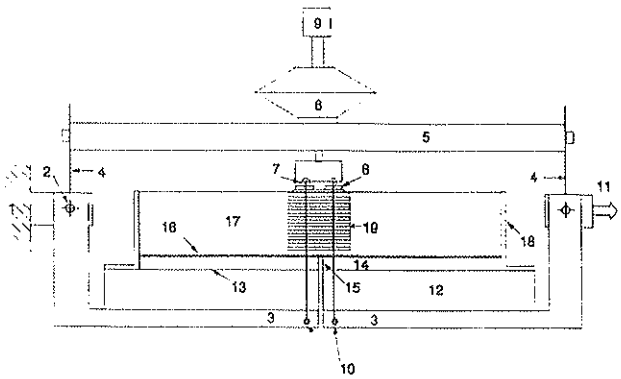
| Test | Binder | Binder content ($kg/m^2$) | Thread content ($g/m^2$) | Crack starting time (minutes) | Rupture time (minutes) |
|---|---|---|---|---|---|
| 1 | Biphase elastomer bituminous emulsion | 1.2 | 50 | 110 | 450 |
| 2 | " | 1.2 | 100 | 160 | 570 |
| 3 | " | 1.6 | 50 | 200 | 520 |
| 4 | " | 1.6 | 100 | 260 | 630 |
| 5 | SBS bitumen | 1.0 | 80 | 270 | 540 |

Each test sample was representative of the system and was submitted to the following two simultaneous loads at a constant temperature (5°C):

- Slow, sustained, longitudinal tensile stress to simulate thermal contraction; and
- Cyclic vertical bending at 1 Hz to simulate traffic.

The progress of the crack was monitored by means of a series of electrical conductors.

The test apparatus is shown in Figure 4. The test made it possible to assess various characteristics linked to the efficiency of the system studied (reflection of crack, speed of development, and time for full cracking of the system).




1. Frame of the apparatus
2. Hinge axis of L-shaped plates
3. L-shaped plates
4. Flexible blades
5. Hydraulic cylinder thrust bar
6. Hydraulic cylinder (adjustable stroke)
7. Rollers
8. Bearing plates
9. Adjustment of hydraulic cylinder stroke
10. Load transmission lines.
11. Recirculating-ball thread tensioning jack
12. Base plates bolted to L-shaped plates (thickness varies with that of samples).
13. Glue (sample glued to base plates)
14. 1.5 cm thick layer of precracked asphaltic concrete to simulate the old cracked pavement.
15. Crack inducer (cardboard)
16. Interface system (geotextile, membrane, sealant) (in some cases)
17. Sample
18. Glued foil to prohibit vertical movement of the ends of the test sample whilst allowing horizontal movement.
19. Crack-detection system (electric wires).

**FIGURE 4 Contraction and bending test apparatus of the regional laboratory at Autun.**

*Results*

Five MURMOS systems were studied. The differences between them concerned

- Binder content (per square meter),
- Type of binder, and
- Thread content (per square meter).

The compositions of the system are presented in Table 1, along with the experimental results.

In all cases, the wearing course was 4-cm-thick fiber-based asphalt with the following characteristics:

- Continuous grading curve: 0 to 10 mm.
- Binder content: 7.2 percent of 60/70 pen
- Fiber content: 1.5 percent.

This table, as well as information on other existing solutions assessed with the same test procedure, indicates that the best performance was obtained with the system containing 1.6 kg/m² of biphase elastomeric bituminous emulsion and 100 g/m² of thread (Figure 5).

### Tests Conducted by the Belgian Road Transport Research Center

The Belgian Road Transport Research Center has developed a tensile and compressive test device that simulates the movement of a cracked concrete support as a result of thermal stresses. This apparatus was described at the RILEM Congress in Liège in March 1989 (*6*). It complements a mathematical model based on the finite difference method, which uses the mechanical properties of the materials (fatigue and modulus of elasticity) as its principle input.

*Principle of the Test*

The test simulates thermal stresses by means of the apparatus shown in Figure 6.

The outstanding feature of this test device is that it simulates crack opening and closing at a speed close to that observed

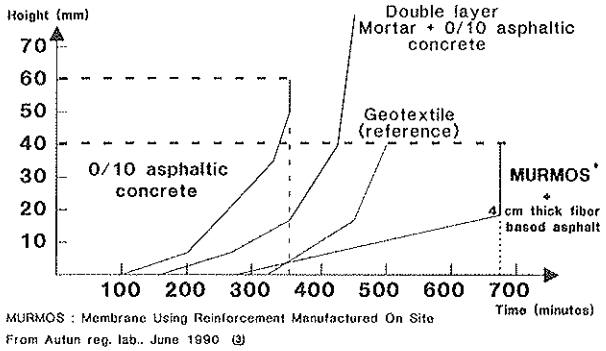



**FIGURE 5 Results of contraction and bending tests on various systems for preventing reflective cracking.**

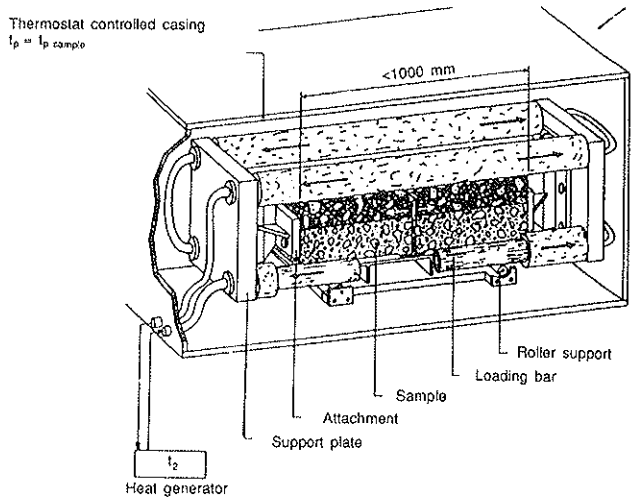



**FIGURE 6 Apparatus for simulating thermal cracking.**

in situ as a result of daily temperature variation, i.e., several 10ths of a millimeter per hour.

With this apparatus, deformation is induced by the expansion or contraction of the frame to which the sample is attached. Such slow and perfectly controllable deformation is obtained by periodically changing the temperature of a liquid circulating in the bars of the frame. The sample is kept at a constant $-5$°C or $-10$°C throughout the test. The roadbase lies on a bed of steel balls that allow free and practically frictionless movement. The apparatus is fitted with sensors for measuring the stresses applied, the opening of the crack, and the displacement of the pavement relative to the cracked support. Visual observation of the phenomenon is facilitated by a video system: filming is controlled by the data acquisition and verification system.

The system to be tested is glued to the road base, a block of precracked concrete. It is then covered with a wearing course. The test samples measure 60 cm in length, and 7 cm in width, and the concrete is 7 cm thick. The speed at which cracks open and close is 15 μm/min, i.e., 1 mm/hr.

The test is computer-controlled, with continuous measurement of (a) the opening of the crack in the support and in the asphalt overlay, (b) the displacement of the wearing course relative to the support, and (c) the applied stress. The progress of the crack is video recorded.

The maximum strain (tensile stress) applied at the edges of the crack inducer is about 25 percent (i.e., 1 mm of opening during a 2-hr cycle for a starting gap of 4 mm).

Figures 7 and 8 show the resulting variation of crack opening and the evolution of the measured forces with time. The results were summarized by Vecoven (*6*): "In general, the interface systems with the highest tensile strengths crack rapidly whereas those comprising emulsions of low-viscosity bitumen can withstand high relative displacement between the overlay and the cracked support without cracking or separation of the various layers of the structure (*5*)."

*Results*

Given the results of the MURMOS tests performed at the Regional Laboratory in Autun, only systems containing 100

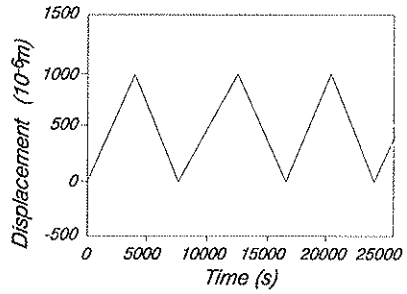



**FIGURE 7 Thermal cracking test: variation of crack opening with time.**

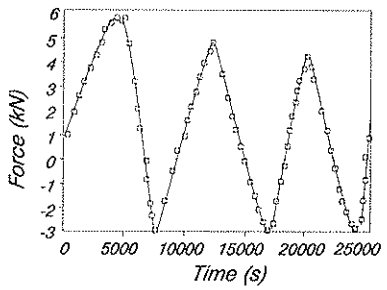



**FIGURE 8 Thermal cracking test: variation of force with time.**

g/m² of thread and 1.6 kg/m² of emulsion were tested. Table 2 presents descriptions of the structures and the results obtained.

It appears that none of the MURMOS systems tested cracked after 20 cycles at −5°C. The same result was obtained with the test at −10°C.

At −5°C, the tested systems using geotextiles and geogrids indicated forces in concrete three or four times higher than that obtained with test sample RTE 3.1. At the time this paper was written, the results of testing at −10°C indicated that no system other than that described in Table 2 withstood cracking. This excellent performance was ascribed to a decoupling effect caused by the thread reinforcement in the membrane that considerably reduced the maximum stress induced in the asphalt overlay when the crack opened.

To sum up the results of the laboratory tests performed by the French Public Service in Autun and by the Belgian Road Transport Research Center, it appears that the composite overlay presented in this paper is the best of the crack-inhibiting courses known to date.

## PRACTICAL EXPERIENCE

As of mid-1990, almost 1 million m² of MURMOS systems had been applied.

The first test sections were laid at the end of 1988 and early 1989. These sections enabled definition of the various application parameters, which led to the development of the Filaflex machine, and they also gave the opportunity for long-term testing of several types of composite systems under different conditions:

- Different types of support;
  - —Alligator-cracked flexible pavement,
  - —Cracked semirigid pavements,
  - —Cracked rigid pavements;
- Different circumstances;
  - —Densely trafficked roads T0, T1;
  - —Urban roads, e.g., rue St. Ferjus, Gallieni bridge; and
  - —Surfaces with special stresses—mountain bends, carparks, etc.

TABLE 2 TESTS PERFORMED AT THE BELGIAN ROAD TRANSPORT RESEARCH CENTER

| Test Sample | R.T.E.1 | R.T.E.2 | R.T.E.3-1 | R.T.E.3-2 |
|---|---|---|---|---|
| System for preventing reflective cracking | M.U.R.M.O.S. with biphase elastomer-bitumen emulsion 1.6 kg/m² | M.U.R.M.O.S. with biphase elastomer-bitumen emulsion 1.6 kg/m² | M.U.R.M.O.S. with biphase elastomer-bitumen emulsion 1.6 kg/m² 6/10 chippings | M.U.R.M.O.S. with biphase elastomer-bitumen emulsion 1.6 kg/m² |
| Wearing course | Fiber based asphalt | Type 1 overlay (7 cm) | Type 1 overlay (7 cm) | Type 1 overlay (7 cm) |
| Test temperature (°C) | -5°C | -5°C | -5°C | -10°C |
| OBSERVATIONS | | | | |
| Duration | 20 cycles | 26 cycles | 20 cycles | 20 cycles |
| Cracking | None | None | None | None |
| Maximum elongation of support | 1.012 mm | 1.021 mm | 1.10 mm | 1.014 mm |
| MEASUREMENTS (cycle) | 1st cycle | 1st cycle | 1st cycle | 1st cycle |
| - Max. force in concrete | 2,745 N | 4,595 N | 1,501 N | 3,346 N |
| - Max. elongation in overlay | 0.180 mm | 2.225 mm | 0.051 mm | 0.084 mm |
| - Max. stress in overlay | 0.89 MPa | 0.97 MPa | 0.30 MPa | 0.68 MPa |
| - Max. stress at interface system | 0.125 MPa | 0.210 MPa | 0.069 MPa | 0.159 MPa |
| MEASUREMENTS (cycle) | 20th cycle | 20th cycle | 20th cycle | 20th cycle |
| - Max. force in concrete | 530 N | 2,216 N | 504 N | 1,452 N |
| - Max. elongation in overlay | 0.098 mm | 0.179 mm | 0.025 mm | 0.039 mm |
| - Max. stress in overlay | 0.17 MPa | 0.47 MPa | 0.10 MPa | 0.30 MPa |

These test sections have been followed up, and they confirm the good performance of the MURMOS system. Some details of actual applications are given in the following sections.

### RN 86 in the Ardèche

A 20 000 m² stretch of national highway RN 86 in the center of France (Ardèche) was upgraded in 1988. The MURMOS innovative technique was selected as part of the contest organized by the regional administration.

The technique consisted of two processes:

- Preventing cracks rising to the surface of the old semirigid pavement. The solution proposed replaced the basic solution, which consisted of 12 cm of conventional asphalt concrete.
- Preventing cracks rising to the surface of the new pavement or the existing pavement reinforced with a cement-treated layer.

To date, i.e., almost 2 years later, no crack has reached the surface.

The wearing course is a 4-cm-thick fiber-based asphalt concrete.

### A55 Highway

The A55 highway is a highly cracked thick flexible pavement. A 30 000-m² stretch was upgraded in August 1989. The T0+ traffic corresponds to 1,200 trucks per day in each direction.

The wearing course is a 2-cm-thick fiber-based asphalt concrete.

### A6 Highway

A 100 000-m² stretch of highway was upgraded in May 1990. Traffic there is also more than 1,200 trucks per day in each direction.

The wearing course is a 4-cm-thick fiber-based porous asphalt.

In this case, the process was applied to three types of support:

- Continuous reinforced cement concrete,
- Californian cement slabs stabilized with connectors (in order to prevent relative vertical movement), and
- Thick cement concrete slabs.

The MURMOS system proposed for all three supports was

- 1.6 kg/m² of binary elastomer bituminous emulsion,
- 100 g/m² of threads,
- 6 litres/m² of 6/10 chippings, and
- 80 kg/m² of porous asphalt.
  - —Gradation: 0 to 14 mm,
  - —Binder: 6 percent of 60/70 pen., and
  - —Fibers: 1 percent.

## ADVANTAGES OF MURMOS PROCESS COMPARED WITH PARALLEL PROCESSES

The MURMOS process has three types of advantage:

- Technical advantages,
- Advantages linked to application, and
- Economic advantages.

### Technical Advantages

Although applied relatively recently (2 years ago), the MURMOS test strips laid in 1988 and at other job sites since then total more than 1 000 000 m². The efficiency of the process has been demonstrated because, to date, no cracks have risen to the surface on the sites monitored.

In addition, the experiments performed by the Autun laboratory, part of the Ponts et Chaussées technical network, specialized in the study of processes for preventing reflective cracking, as well as those performed at the Belgian Road Transport Research Center in Liège, have led to selection of the most efficient formulas that provide some of the best performances for processes known to date.

### Advantages of MURMOS System Application

The advantages that the MURMOS system has over techniques using shop-manufactured materials are as follows:

- System parameters can be adjusted to meet needs;
- The system can be fully adapted to the longitudinal profile of the road (no folds in curves, no overlap problem); and
- High daily application rates (close to 30 000 m² per day) can be obtained.

### Economic Advantages

In the MURMOS process, the form in which the raw materials are used (reel of threads instead of ready-manufactured geotextiles), together with the high daily application rate, permits entirely competitive costs. According to the laboratory test results, the application rate must nevertheless be near the following mean values: binder 1.6 kg/m² (1.4 to 1.8 kg/m²), threads 100 g/m² (80 to 100 g/m²).

## CONCLUSIONS

First experimented with in 1988, a new system for preventing reflective cracking has been developed. MURMOS consists of manufacturing a reinforced membrane in situ by simultaneously spraying a layer of generally modified binder and continous threads that interweave to form the reinforcement. This composite system is then covered with a wearing course.

After nearly 3 years of experiments, the main conclusions are as follows:

- Laboratory tests indicate that the performance of the MURMOS system is among that of the best;
- With the specially designed machine, daily laying rates are high; they can reach 30 000 $m^2$ per day;
- In distinction to systems using shop-manufactured reinforcement, the MURMOS process presents no problem with respect to bends or road width (no folds in curves or problem of overlapping);
- Although relatively recent, the upgrading work performed with the MURMOS process up to mid-1990, amounting to almost 1 000 000 $m^2$ on a wide variety of roads (cracked flexible pavements, semirigid pavements, and concrete roads), demonstrates the excellent performance of the process in the field; and
- The MURMOS process is felt to be an economic solution for cracking problems on all types of road (flexible, semirigid, and rigid).

## REFERENCES

1. *Priorités de Développement des Innovations dans le Système des Chaussées et terrassement*. French Highway Research Department, French Ministry of National Public Works, Housing, National Development and Transport, Road Transport Division, Feb. 1988.
2. *Techniques pour Limiter la Remontée des Fissures à la Surface des Chaussées Semi-Rigides*. Leaflet 57, French Highway Research Department, Road Transport Division, March 1990.
3. C. Le Noir and M. Baillie. *Etude de Système anti Remontée de Fissures*. Procédé FILAFLEX. French Ministry of National Public Works, Housing, National Development, and Transport, Road-Transport Division, CETE, Ponts et Chaussées Regional Laboratory at Autun, 1990.
4. *Fissuration de Retrait des Chaussées à Assises Traitées aux Liants Hydrauliques*. Ponts et Chaussées Laboratory Bulletin Nos. 156 and 157, June and Aug. 1988.
5. T. H. Vecoven. *Méthode d'Essai de Systèmes Limitant la Remontée des Fissures dans les Chaussées*. Ponts et Chaussées Regional Laboratory in Autun, RILEM Congress on Reflective Cracking in Pavements: Assessment and Control, Liège, Belgium, March 8–10, 1989.
6. C. Clauwaert and L. Franken. Etude et Observation de La Fissuration Réflective au Centre de Recherches Routières Belge. Belgian Road Transport Research Center, RILEM Congress on Reflective Cracking in Pavements: Assessment and Control, Liège, Belgium, March 8–10, 1989.

*Publication of this paper sponsored by Committee on Pavement Maintenance.*

# New Type of Ultrathin Friction Course

J. P. SERFASS, P. BENSE, J. BONNOT, AND J. SAMANOS

A new kind of ultrathin surfacing, intermediate between chip seal and very thin asphalt hot mix, has appeared recently. It basically consists of a layer of hot precoated aggregate, spread over a binder spray application. A specifically designed machine spreads both binder and coated aggregate in a single pass and smooths the course. The machine includes a receiving hopper, conveyors, binder storage tanks, a variable-width spray bar, and a light heating screed unit with extensions. It operates at high speed, currently about 20 m/min. The binder is a modified emulsified asphalt. Its application rate varies between 0.6 and 1.2 kg/m$^2$. The friction course itself consists of an aggregate, mostly 6 to 10 mm in size, coated with a mortar made of sand, filler, and asphalt cement. The latest content ranges from 5 to 5.5 percent. The course thickness is between 1 and 2 cm. Its application rate of total mix is between 23 and 28 kg/m$^2$. The resulting surfacing is homogeneous. It offers a rough macrotexture and high levels of skid resistance. It has shown some capacity of leveling existing small irregularities of the support. Its rolling noise is significantly lower than that of chip seals made with aggregate of the same size. To date, the behavior of this friction course appears to be promising, as no deterioration of any kind has been observed after 2 years, even under heavy traffic.

Chips seals provide widely recognized advantages, such as waterproofing and skid resistance, but they also present serious drawbacks such as no reshaping capacity and high rolling noise. Moreover, their success always depends on the climatic conditions, not only at the time of application, but also during the first months of trafficking. Adverse weather or inadequate application procedures inevitably result in whip-off or bleeding.

In the field of surface treatment, there has consequently been a need for a new technique that would eliminate the defects in chip seals while keeping the main qualities of skid resistance and impermeability.

The new type of surfacing described here is the outcome of a continuous evolution in French asphalt wearing courses. In the 1970s, classical hot-mix wearing courses were ranging between 6 and 10 cm in thickness; they were used chiefly for structural reasons. In the second half of the 1970s, "thin" hot mixes (3 to 5 cm) appeared, mainly designed to meet maintenance requirements; a tack coat was then systematically applied to ensure complete bonding of the thin course to its support. In the early 1980s, because maintenance requirements were mostly concerning surface service quality, "very thin" hot mixes (2 to 3 cm) were developed and are now extensively utilized—50 000 000 m$^2$ in the 1984 to 1989 period.

J. P. Serfass, SCREG Routes, 1 Avenue Eugène Freyssinet, 78065 St. Quentin en Yvelines, Cedex, France. P. Bense, SCREG Routes, 2 Rue Virginie Mauvais, 54015 Nancy, Cedex, France. J. Bonnot, laboratoire Central des Ponts et Chaussées, Orly Sud No. 155, 94396 Orly Aerogare, Cedex, France. J. Samanos, SCREG Routes, 1 Avenue Eugène Freyssinet, 78065 St. Quentin en Yvelines, Cedex, France.

Such very thin asphalt is always placed on a generous tack coat, the amount of which can reach 0.7 to 0.8 L/m$^2$ of emulsion in some cases.

The ultimate step in this evolution has been a new decrease in the thickness of hot mix and a simultaneous increase in the tack coat application rate. A new type of friction course has resulted, as described in the following sections, intermediate between very thin hot-mix asphalt and chip seal. It is being developed under a cooperative agreement between the Laboratoire Central des Ponts et Chaussées (LCPC) and SCREG Routes.

## THE CONCEPT

The technique involves three indissociable elements:

- Materials,
- Design, and
- Equipment and application method.

The friction course consists of a layer of hot precoated chippings, spread over a binder spray application. A machine, specially designed for this process, spreads both binder and aggregate in one pass and smoothes the course, at high speed.

Aggregates, prepared in a hot-mixing plant, are coated with mortar consisting of a bituminous film considerably thickened by fine mineral materials. They are hot-applied and spread on the binder layer sprayed a few seconds before.

Such a coating ensures a strong bonding between the chippings, which can be superposed in depressions and therefore level distorted surfaces to a certain extent. As a result of their immediate application to the binder, chippings are perfectly held in place, whip-off is totally eliminated, and the working site remains clean all over.

The machine finishes the application by ironing the chippings layer. This operation provides a good riding quality. Moreover, because the aggregates are laid flat (Figure 1), the rolling noise is reduced.

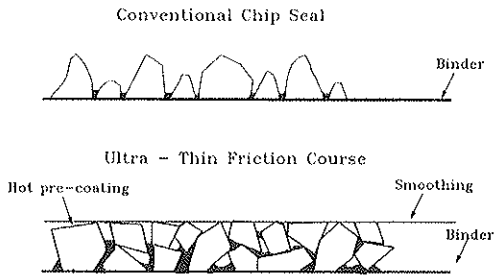



**FIGURE 1 Comparison of conventional chip seal with ultrathin friction course.**

## DESIGN

### Components

Aggregates are in compliance with the French Directive for Surface Dressing. To date, 6 to 10 mm has been the most commonly used size, but in some cases 4 to 6 or 10 to 14 mm grading sizes are preferred. Typical aggregate gradations are the following:

| | *Aggregate Size Range (mm)* | | |
|---|---|---|---|
| | *10–14* | *6–10* | *4–6* |
| *Sieve Size (mm)* | *Percentage Passing (by Weight)* | | |
| 20 | 100 | – | – |
| 14 | 83 | – | – |
| 12.5 | 53 | 100 | – |
| 10 | 7 | 89 | – |
| 8 | 1 | 44 | 100 |
| 6.3 | – | 8 | 95 |
| 5 | – | 1 | 51 |
| 4 | – | – | 5 |
| 2 | – | – | 1 |

Aggregates are coated with a mastic that usually contains pure bitumen. However, for roads with heavy traffic, the use of a modified binder can be justified.

This mortar is usually made with sand of 0 to 2 mm, a typical gradation of which follows:

| Sieve size (mm) | 4 | 2 | 1 | 0.5 | 0.315 | 0.2 | 0.08 |
|---|---|---|---|---|---|---|---|
| Percentage passing | 100 | 93 | 62 | 44 | 35 | 28 | 19 |

The coating mortar represents 20 to 25 percent (by weight) of the total mix. The bitumen content in the total (aggregate plus mortar) ranges from 5 to 5.5 percent.

The binder applied on the support is either a latex- or elastomer-modified emulsion. The use of classical cationic CRS emulsion of plain asphalt cement should be possible for low traffic, but certainly not for medium or high traffic, because of bleeding hazards.

### Application Rates

The application rate of the binder layer depends on traffic and support condition. For the jobs carried out so far, emulsion rates have ranged from 0.6 to 1.2 kg/m² (most of them bewteen 0.7 and 1 kg/m²), that is, from 0.4 to 0.7 kg/m² of residual binder.

For aggregates, average site values vary between 23 and 27 kg/m² of coated 6- to 10-mm aggregate.

Overall application rates are close to those of conventional chip seal:

| *System* | *Total Asphalt Cement (kg/m²)* | *Aggregate (kg/m²)* |
|---|---|---|
| Single-seal with 6–10 mm | 1.3 | 14 |
| Single-seal with double-chipping application 10–14/4–6 mm | 1.6 | 20 |
| Double seal 10–14 + 4–6 mm | 2.0 | 25 |
| Ultrathin friction course | 2.0 | 26 |

## EQUIPMENT AND APPLICATION

In order to execute all operations required for application of the product, new equipment had to be designed capable of providing the various functions and integrating the specific constraints stemming from the new material used, including the use of hot, self-bonding chippings, in place of cold, free-running chippings.

Another constraint was introduced by the comparatively low spray rate of the binder, approximately half that for a conventional chip seal. The binder film must be protected from any form of vehicular traffic, to guarantee the continuous impermeability of the road structure and to ensure impeccably clean job sites.

Moreover, application has to be completed by ironing of a more or less monogranular layer of aggregate, performed at high speed, three to four times that of a conventional finisher. For this reason, particular attention was paid to the smoothing device to be used.

A machine shown in Figure 2 has been designed to complete the three operations of spreading the binder, applying the hot precoated chippings, and smoothing the course in a single pass in a short time. Figure 3 shows a schematic of this machine. At present, five machines of this type have been constructed and are in operation.

Schematically, this equipment comprises the following components:

- A receiving hopper for precoated chippings, with a self-locking hook for the supplying truck;
- A scraper-type conveyor;
- A chippings storage chamber with appropriate thermal insulation and a total capacity of 5 m³;
- Several binder tanks, thermally insulated, with a capacity of 12 m³ (more than a half-day of work);
- A conveyor transferring chippings to the screed unit;
- A variable-width spray bar;
- Two spreading screws; and
- A variable-width heating screed unit.

Special attention was paid to the design of the spray bar. It includes wide-angle nozzles, whose delivery is slaved to the road speed of the machine. This ramp has a variable width, and the constant application rate for variable widths is ensured by an original system.

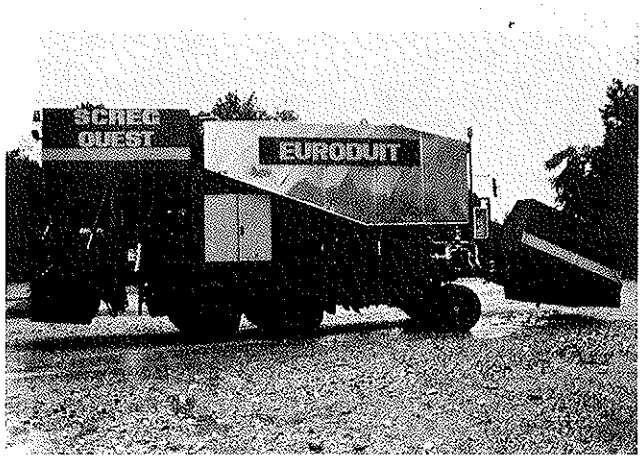



**FIGURE 2 Photograph of single-pass machine for rolling Novachip friction course.**

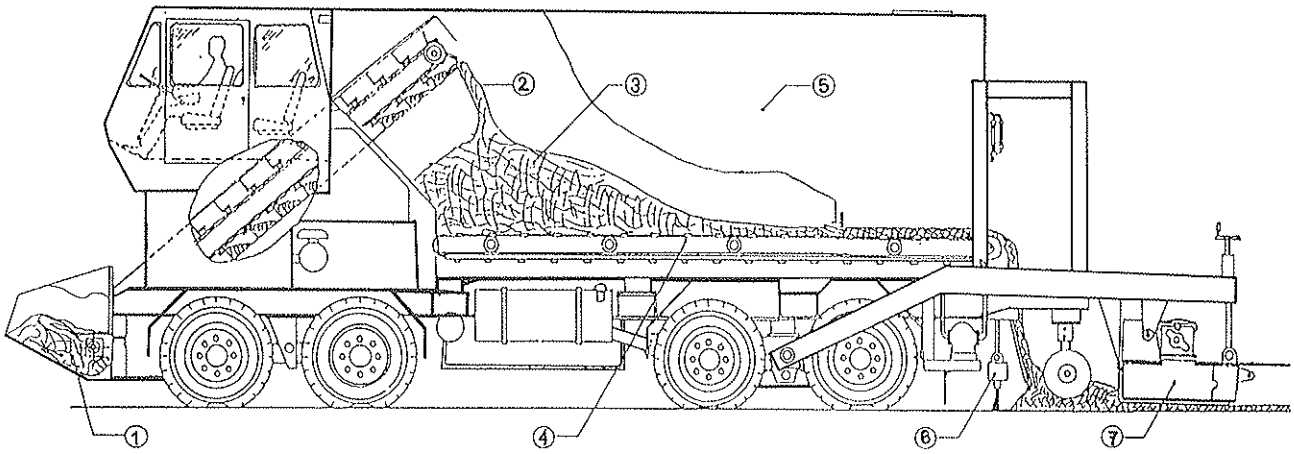



**FIGURE 3 Schematic of single-pass machine for rolling Novachip friction course.**

The adjustable width of the spreading and smoothing system varies from 2.50 to 4.20 m.

The friction course is applied at high speed, always >10 m/min, and able to reach 20 to 25 m/min.

The three operations of spraying the binder, applying the precoated chippings, and smoothing the course are all completed in a few seconds. This quick application results in

- Good bonding of coated chippings to the binder and between themselves, even when the weather is cool or unsettled;
- Clean site (no plant or truck traffic on the binder layer); and
- Short construction time, thus reduced inconvenience to road users.

Finally, the friction course is rolled (Figure 4) and can be open to traffic as soon as it is cooled.

**FIGURE 4 Photograph of rolling operation of ultrathin course.**

## IN-PLACE CHARACTERISTICS—BEHAVIOR UNDER TRAFFIC

### Projects Completed

The first trial sections took place in the fall of 1988. In 1989, more than 800 000 m² were applied from May to October, in varied regions and under different traffic ranging from medium-low (T3) to very high (T0). French traffic classes are defined as follows:

| Traffic classes | T5 | T4 | T3 | T2 | T1 | T0 | TE |
|---|---|---|---|---|---|---|---|
| Average daily number of heavy vehicles per lane | 0–25 | 25–50 | 50–150 | 150–300 | 300–750 | 750–2,000 | >2,000 |

Reference is made to the average daily number of heavy vehicles traveling in one direction on the most heavily trafficked lane. Heavy vehicles are defined as those vehicles having a payload of at least 5 metric tons. According to the French legislation, the maximum axle load is 13 metric tons.

In 1990, the surface area treated will exceed 3 000 000 m². All French regions have been involved, with a variety of climates, types of support, and traffic intensities. Several jobs were also carried out in Sweden and Belgium.

### In-Situ Properties of the Surfacing

A multiyear assessment program has been set up and funded jointly by SCREG Routes and LCPC to evaluate the effects of mix parameters and to monitor the in situ behavior of the friction course. These parameters are presented in Table 1.

#### *Effect on Riding Quality*

Three sets of measurements have been analyzed, as presented in Table 2.

TABLE 1 LCPC–SCREG MONITORING PROGRAM

| Parameter studied | Measurement | | | | | |
|---|---|---|---|---|---|---|
| | Profile Analyzer Coefficient (PAC) | | Surface texture depth (sand patch and laser texture meter) | Longitudinal Friction Coefficient | In place Permeability | Noise |
| | Before | After | | | | |
| Binder spray rate Aggregate size | X | X | 5 sections | 5 sections | 5 sections | 3 sections |
| Asphalt coating content | X | X | 2 sections | 2 sections | 2 sections | |
| Mortar content | | | 3 sections | 3 sections | 3 sections | |
| Aggregate type | | | 2 sections | 2 sections | 1 section | |

TABLE 2 EFFECT OF FRICTION COURSE ON RIDING QUALITY

Friction course = 23 kg/m2

| PAC* | < 6 (%) | < 13 (%) | < 16 (%) | Job |
|---|---|---|---|---|
| Before | 13 | 62 | 76 | CD 312 (Maine et Loire) |
| After | 19 | 85 | 96 | |

Δ = ≈ + 6

Friction course = 28 kg/m2

| PAC | < 6 (%) | < 13 (%) | < 16 (%) | Job |
|---|---|---|---|---|
| Before | 23 | 90 | 94 | RN 23 (Maine et Loire) |
| After | 39 | 97 | 97 | |

Δ = ≈ + 16

Friction course = 25 kg/m2

| PAC | < 6 (%) | < 13 (%) | < 16 (%) | Job |
|---|---|---|---|---|
| Before | 36 à 53 | 96 à 98 | 98 à 99 | RN 18 |
| After | 59 à 72 | 98 à 99 | 99 à 100 | Meuse |

Δ = ≈ + 20

*PAC : Profile Analyzer Coefficient

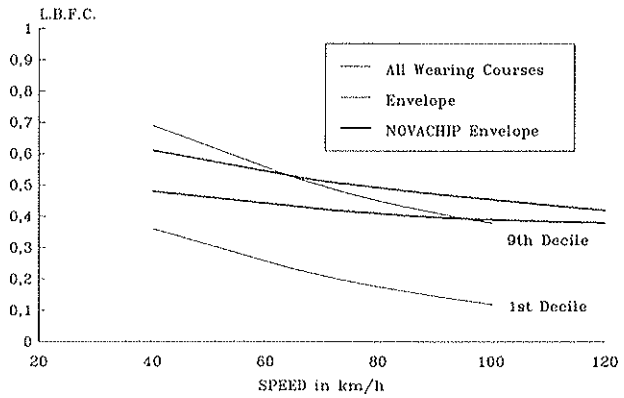



**FIGURE 5 LBFC versus traffic speed—comparison of envelopes for all wearing courses and Novachip.**

Figure 5 shows the results obtained on one of the sites. (NOVACHIP is the trade name of the ultrathin friction course developed by SCREG Routes, in cooperation with LCPC).

These results prove that such an ultrathin friction course has, somehow unexpectedly, a limited but real leveling capacity. Indeed, it appears that precoated aggregates can be superposed and stand up to traffic effects in this position. It is estimated that, in case of light traffic (up to T3), the total thickness in depressed areas can reach up to 4 cm.

### *Surface Texture Depth*

A considerable number of measurements have been obtained with the sand patch test (SPT) method or with the laser Mini

TABLE 3 TEXTURE DEPTH MEASUREMENTS BY SAND PATCH TEST

| Aggregate size (mm) | Sand Patch Depth (mm) | |
|---|---|---|
| Size range 6-10 | Minimum 1,1<br>Mean 1.9<br>Maximum 3,2 | (50 measures) |
| Size range 4-6 | Minimum 1<br>Mean 1.3<br>Maximum 1.4 | (8 measures) |
| Size range 10-14 | Minimum 2.1<br>Mean 3.8<br>Maximum 3.2 | (12 measures) |

TABLE 4 TEXTURE DEPTH MEASUREMENTS BY MINI TEXTURE METER

| Aggregate size (mm) | Sensor Measured Texture Depth (mm) | |
|---|---|---|
| Size range 6-10 | Minimum 0.65<br>Mean 0.85<br>Maximum 1.01 | (145 measures) |
| Size range 4-6 | Minimum 0.43<br>Mean 0.52<br>Maximum 0.64 | (100 measures) |
| Size range 10-14 | Minimum 1.24<br>Mean 1.75<br>Maximum 2.04 | (130 measures) |

TABLE 5 VARIATION OF LBFC FRICTION WITH TRAFFIC SPEED

- At 60 km/h

| | T2 Traffic | | T1-T0 Traffic | |
|---|---|---|---|---|
| | CD 1 | CD 12 | RN 23 | RN 24 |
| After 5 months | 0.53 | 0.52 | 0.46 | 0.48 |
| After 8 months | | | | |
| After 16 months | 0.53 | | | |

- At 90 km/h

| | T2 Traffic | | T1-T0 Traffic | |
|---|---|---|---|---|
| | CD 1 | CD 12 | RN 23 | RN 24 |
| After 5 months | 0.47 | 0.45 | 0.40 | 0.43 |
| After 8 months | | | | |
| After 16 months | 0.40 | | | |

- At 120 km/h

| | T2 Traffic | | T1-T0 Traffic | |
|---|---|---|---|---|
| | CD 1 | CD 12 | RN 23 | RN 24 |
| After 5 months | 0.40 | 0.39 | | |
| After 8 months | | 0.38 | | |

Texture Meter (MTM), on job sites between 1 and 10 months old. The results are as follows:

- Manual SPT. Seventy-six tests were executed on nine different job sites. The statistical profile is presented in Table 3.
- Mini Texture Meter. The number of measurements gathered here amounts to almost 400, and yields the sensor-measured texture depth (SMTD) results presented in Table 4.

### *Skid Resistance*

Consistent results were obtained after 4 to 5 months on three jobs, with 6- to 10-mm aggregate. Values of the longitudinal braking force coefficient (LBFC) were as follows:

| *Speed km/hr* | *LBFC* |
|---|---|
| 40 | 0.61 |
| 60 | 0.53 |
| 90 | 0.45 |
| 120 | 0.39 |

Measurements carried out on some of the 1988 sections contained little variation between 5, 8, and 16 months, as presented in Table 5.

The levels of skid resistance obtained generally classify this type of friction course as among the most effective surface treatments available (Figure 6). This effectiveness is clearly because of the rough and open macro texture and, to a lesser extent, the hardness and polishing resistance of the aggregate selected.

### *Rolling Noise*

Rolling noise measurements were performed in accordance with the French-German method, i.e., using three different vehicles (with engine running) with different types of tires at various speeds. This methods establishes a correlation between acoustical pressure $Lp_{max}$ and vehicle speed. Figures 7 and 8 show an example of rolling noise recoding.

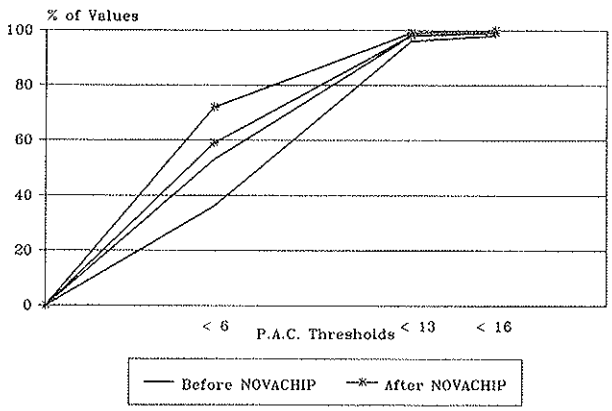



**FIGURE 6 Profile analyzer coefficient values for Novachip RN 18 (Meuse, France).**

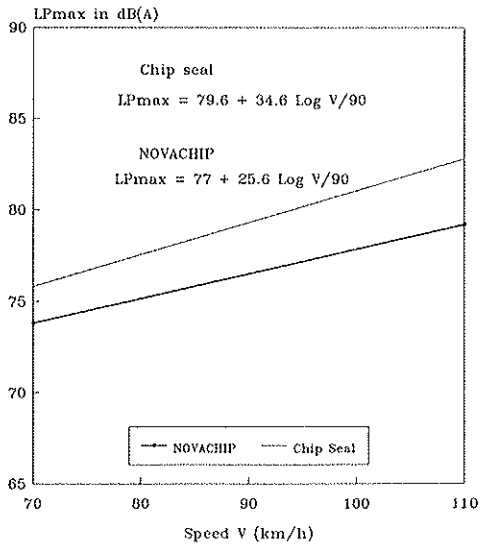



**FIGURE 7 Regression analysis for rolling noise versus traffic speed.**

The noise values recorded indicated that ultrathin friction courses cause lower rolling noise than conventional chip seal using the same aggregate size.

The reduction amounted to 2 to 3 decibels within the usual vehicle speed range. This benefit was significant for neighboring residents and user comfort, because acoustical energy was decreased by 30 to 40 percent.

## Overall Behavior

To date, all visual observations and in-place assessments indicate that this ultrathin friction course performs satisfactorily.

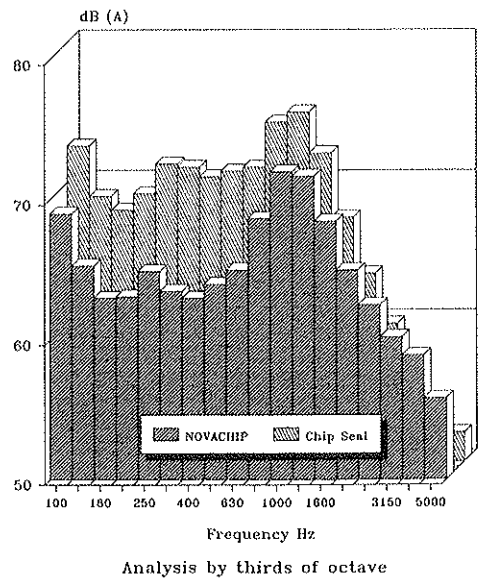



**FIGURE 8 Spectrum analysis of rolling noise at 90 km/hr.**

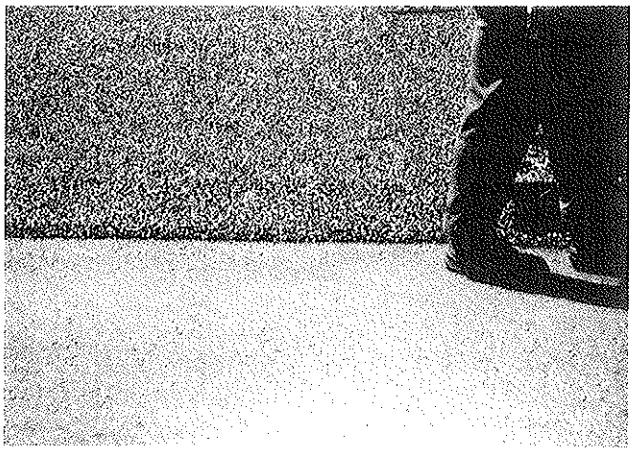


**FIGURE 9 Photograph of surface aspect of ultrathin friction course.**

**FIGURE 10 Photograph showing open surface texture of ultrathin friction course.**

- Its surface aspect is homogeneous (see photograph in Figure 9),
- No aggregate dislodgment has been noticed,
- No bleeding nor fatting up has occurred,
- The surface texture is open (see photograph in Figure 10)

and appears to be stable even under heavy traffic (up to T0).

Consequently, no traffic volume limitation can so far be determined or suggested for application.

## CONCLUSION

The first results, as well as the favorable answer of road authorities and users, are encouraging.

It has been established that ultrathin friction courses can readily be applied at spread rates well below 30 kg/m$^2$ and at high speed (>20 m/min).

The resulting coarse-graded courses have a thickness ranging from 10 to 20 mm and provide high skid resistance and satisfactory waterproofing. As compared with conventional chip seals, major improvements have been achieved, including virtual elimination of aggregate dislodging, better riding quality, and lower rolling noise.

At present, there is no evidence for any traffic volume limitation.

*Publication of this paper sponsored by Committee on Pavement Maintenance.*

# Rapid Techniques for the Repair and Protection of Bridge Decks

MICHAEL M. SPRINKEL, RICHARD E. WEYERS, AND ANGELA R. SELLARS

Bridges that are candidates for rapid repair techniques have peak-hour traffic volumes that are so high it is not practical to close a lane to repair the deck or to install a deck protection system except during off-peak traffic periods. Results of the first 25 months of a 55-month project (Task 4 of Strategic Highway Research Program Project C103) to investigate rapid techniques for the protection, rehabilitation, and replacement of bridge decks are summarized. A review of the literature and responses to questionnaires sent to state departments of transportation (DOTs), Canadian provinces, selected turnpike and thruway authorities, technology transfer centers, and material suppliers was conducted. Techniques being used by the DOTs are identified and compared from the standpoint of frequency of use, performance characteristics, time demands, service life, maintenance, initial cost, and life cycle cost.

The Strategic Highway Research Program (SHRP) awarded contract SHRP C103 to Virginia Polytechnic Institute and State University on September 22, 1988, to conduct a 55-month study entitled *Concrete Bridge Protection and Rehabilitation: Chemical and Physical Techniques* (*1*). The objective of Task 4 of SHRP C103 was to develop technically and economically feasible methods of deck protection, rehabilitation, and replacement that could be used where contruction must be rapid. The objective would be accomplished by a progression through six activities. The state-of-the-art review, data reduction and analysis, and comparison of alternatives (Activities 1 and 2) are summarized herein. This paper is based on reviews of the literature and of the responses to three questionnaires. Additional details can be found in Interim Report 1 (*2*). Rapid repair techniques are compared from the perspective of frequency of use, performance characteristics, time demands, service life, and cost.

## CRITERIA FOR RAPID REPAIR TECHNIQUES

For this study, rapid repair is not defined in terms of repair rate, such as surface area per unit of time, because repair rate is a function of manpower and equipment. Rates at which repairs are done can best be controlled by contract requirements with incentives and penalties to promote rapid rates of repair. Contractors can then invest in additional manpower and equipment to accelerate the rate of repair.

M. M. Sprinkel and A. R. Sellars, Virginia Transportation Research Council, P.O. Box 3817, University Station, Charlottesville, Va. 22903. R. E. Weyers, Department of Civil Engineering, 204 Patton Hall, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061.

For this study, rapid repair is defined in terms of suitability for stage construction. To be considered a rapid-repair technique, the repair system must be suitable for installation during off-peak traffic periods and suitable for traffic during peak traffic periods.

A flow diagram for rapid repair techniques for bridge decks is shown in Figure 1. Lane closure and surface preparation are necessary first steps for any rapid technique. Lane closure can be accomplished using cones or other temporary portable barriers. All unsound concrete must be removed in preparation for new repair materials.

If there is insufficient time to install and cure a protection system or repair material, temporary materials should be placed to maintain a traffic-bearing surface. Otherwise, the repair should continue with the installation of a protection system, a rapid-curing concrete repair material, or a precast replacement section. The materials are allowed to cure to the required strength to receive traffic. After necessary temporary materials are installed, the lane is opened to traffic. If needed, a rapid deck protection system is installed following deck replacement or rehabilitation.

A bridge deck that must be repaired using a rapid-repair technique will usually have one of four maximum lane closure time conditions that require the use of one of four rapid-repair techniques as follows:

- <56 hr—semirapid (e.g., Friday at 9:00 p.m. to Monday at 5:00 a.m.);
- <21 hr—rapid (e.g., 6:30 p.m. to 3:30 p.m.);
- <12 hr—very rapid (e.g., 6:00 p.m. to 6:00 a.m.); and
- <8 hr—most rapid (e.g., 9:00 p.m. to 5:00 a.m.).

A repair system must follow the flow diagram (see Figure 1) within the lane closure constraints of <56, <21, <12, or <8 hr to qualify as part of a rapid-repair technique.

## QUESTIONNAIRE RESPONSE

Three questionnaires on rapid-repair techniques for bridge decks were prepared and distributed in 1989 to obtain state-of-the-art information. Questionnaire 1 was sent to state department of transportation (DOT) coordinators, SHRP Canadian provincial coordinators, and selected turnpike and thruway authorities. Questionnaire 2, a condensed 1-page version of Questionnaire 1, was sent to the directors of the technology transfer centers for publication in their newsletters. Questionnaire 3, an expanded 14-page version of Questionnaire 1, was designed to obtain detailed data on the properties

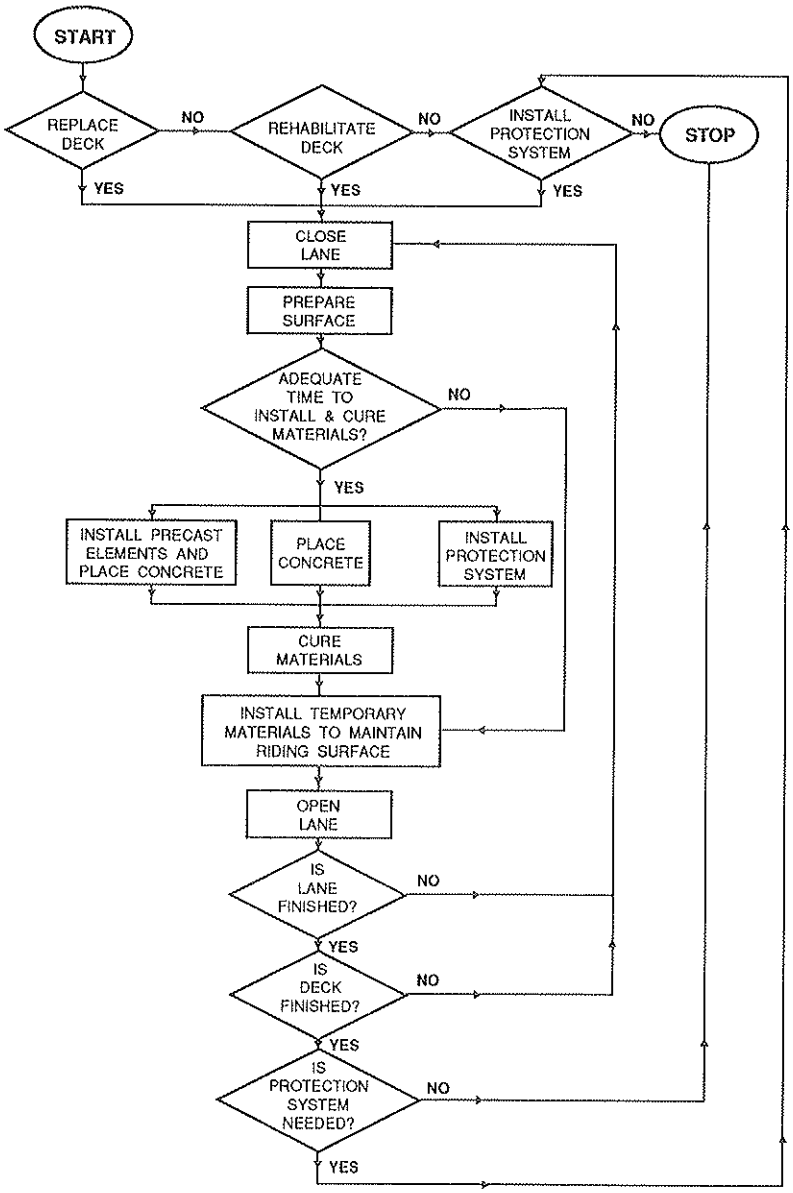



**FIGURE 1 Flow diagram for rapid repair techniques for bridge decks.**

of materials. It was sent to selected material suppliers. The questionnaires were distributed and returned as follows:

| *No.* | *Sent to* | *Date Mailed* | *No. Mailed* | *No. Returned* |
|---|---|---|---|---|
| 1 | SHRP state DOT coordinators | March 8 | 55 | 49 |
| 1 | CSHRP provincial coordinators | March 8 | 12 | 10 |
| 1 | Selected turnpike and thruway authorities | May 30 | 44 | 9 |
| 2 | Directors of technology transfer centers | April 26 | 58 | 8 |
| 3 | Selected material suppliers | June 7 | 276 | 31 |

## FREQUENCY OF USE

Table 1 presents the frequency of use of rapid-repair systems on the basis of the responses to Questionnaires 1 and 2. The respondents were requested to list the three most frequently used techniques for the rapid protection, rehabilitation, and replacement of bridge decks. The rehabilitation of a deck usually requires crack repair, joint repair, patching, and the application of a protective system. In order to simplify the reporting of data, protective systems are not recorded as part of rehabilitation systems. The systems most often used are the bituminous concrete overlay for rapid protection (35 responses), the high-early-strength portland cement concrete patch for rapid rehabilitation (30 responses), and no rapid replacement technique (43 responses).

## PERFORMANCE CHARACTERISTICS

The most important performance characteristics of rapid protection and rehabilitation systems for bridge decks are the condition of the temporary surfaces, minimum curing time,

TABLE 1 FREQUENCY OF USE OF RAPID REPAIR SYSTEMS

| Protection System | No. Users | Rehabilitation System | No. Users | Replacement System | No. Users |
|---|---|---|---|---|---|
| Bituminous Concrete Overlay | 35 | Crack Repair and Sealing | 3 | Precast Concrete Slab Span | 0 |
| Coating | 3 | Joint Repair | 0 | Precast Concrete Box Beam | 0 |
| Portland Cement Concrete Overlay | 9 | Bituminous Concrete Patch | 11 | Precast Concrete Channel and Tee Beam | 0 |
| Penetrating Sealer | 9 | Portland Cement Concrete Patch | 30 | Precast Concrete Deck Panel | 5 |
| Polymer Overlay | 13 | Polymer Concrete Patch | 3 | Permanent Forms with Site Cast Concrete | 0 |
| Other Hydraulic Concrete Overlay | 1 | Other Hydraulic Concrete Patch | 11 | Site Cast Portland Cement Concrete | 9 |
| None | 33 | Steel Plate over Concrete | 3 | Site Cast Polymer Concrete | 0 |
| No Reply | 13 | None | 31 | Other Site Cast Hydraulic Concrete | 3 |
| | | No Reply | 10 | None | 43 |
| | | | | No Reply | 20 |

bond strength, permeability to chloride ion, skid resistance, and wear. With two exceptions, the same performance characteristics apply to rapid-replacement systems. Bond strength is not important unless a protective overlay will be applied, and permeability to chloride ion is less important because the rebar in new decks is usually coated with epoxy.

### Temporary Surfaces

A major requirement for a rapid-repair system is a temporary surface that is suitable for traffic during peak-hour traffic periods. The temporary surface is the disturbed surface between the original surface of the deck and the completed surface. For bridges whose entire deck surface can be repaired during one off-peak traffic period, there is no temporary surface. The surface should provide a satisfactory ride when the lane is open to traffic. Typical surface elevations for the rapid protection systems are presented in Table 2 and shown in Figure 2. When patching, bituminous concrete, steel plates, or timber plank can be used to provide a temporary riding surface if the patching materials cannot be placed and cured properly before opening the surface to traffic.

### Minimum Curing Time

One of the most important properties of a rapid protection, rehabilitation, or replacement system is the strength of the materials at the time they are first subjected to traffic. Materials that do not have adequate strength can be damaged by traffic and fail prematurely as a result of a failure of the matrix or the bond interface. Obviously, a material must be relatively free of cracks and must be adequately bonded to the substrate to protect the deck and provide skid resistance. With the exception of bituminous concrete, sealers, and coat-

TABLE 2 TYPICAL SURFACE ELEVATIONS FOR RAPID PROTECTION SYSTEMS

| Protection System | System Thickness (in) | Surface Preparation Depth (in) | Change in Elevation (in) | Effect on Ride Quality |
|---|---|---|---|---|
| Bituminous Concrete Overlay on Membrane | ≥1.6 | ≤0.1 | >1.6 | Major |
| Coating | ≤0.1 | ≤0.1 | ≤0.1 | Negligible |
| Portland Cement Concrete Overlay | ≥1.3 | ≥0.5 | ≥0.8 | Medium |
| " " | ≥2.0 | ≥0.5 | ≥1.5 | Major |
| Penetrating Sealer | ≤0.1 | ≤0.1 | ≤ 0.1 | Negligible |
| Polymer Overlay | ≥0.3 | ≤ 0.2 | ≥0.1 | Negligible |
| " " | ≥0.5 | ≤0.2 | ≥ 0.3 | Minor |
| Other Hydraulic Concrete Overlay | ≥1.3 | ≥0.5 | ≥0.8 | Medium |
| " " | ≥2.0 | ≥0.5 | ≥1.5 | Major |

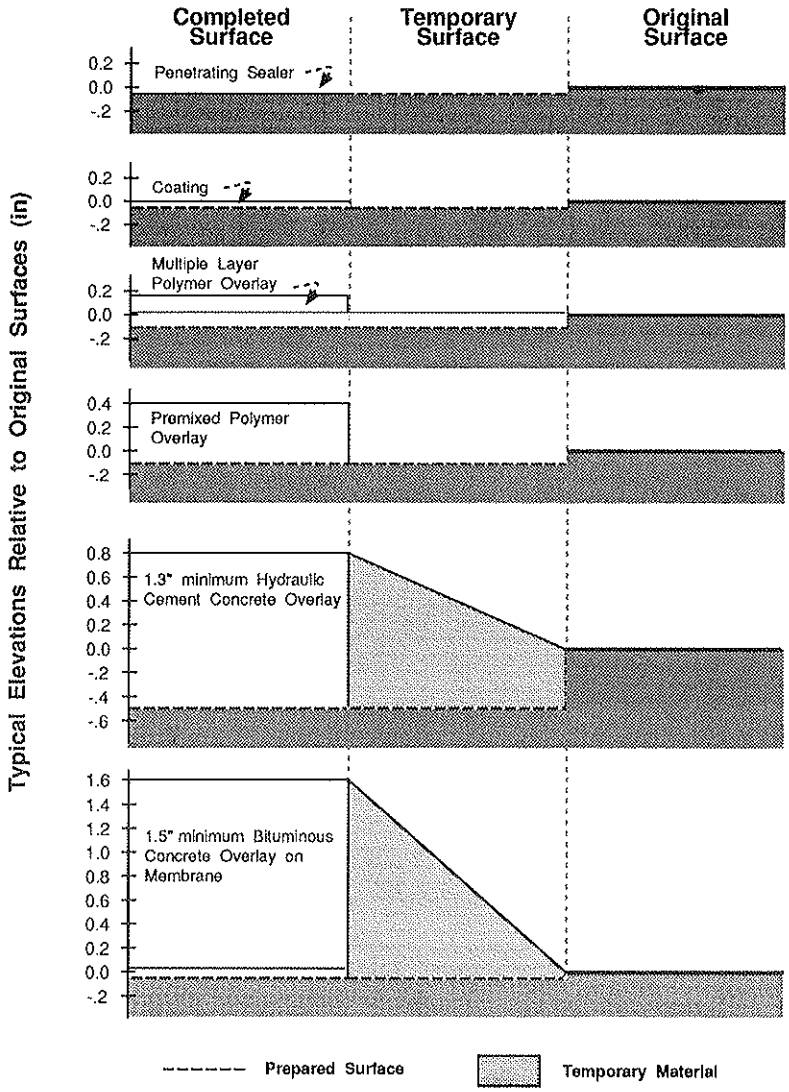



**FIGURE 2 Typical surface elevations for rapid protection systems.**

ings, the most convenient indicators of strength are the compressive strengths of 4 × 8-in. cylinders of concrete and 2-in. cubes of mortar. Hydraulic cement concretes and polymer concretes are usually required to have a compressive strength of 2,500 to 4,000 psi before being subjected to traffic (*3*). Guillotine shear bone strengths of at least 200 to 400 psi are usually obtained at these compressive strengths when concrete substrates are properly prepared (*4,5*). Tensile adhesion strengths greater than 100 psi are also indicative of satisfactory performance (*6,7*). Coatings and sealers must be tack-free at the time they are subjected to traffic. Membranes must be tack-free before being overlaid with bituminous concrete, which is then allowed to cool to 150°F before being opened to traffic (*3*). Patches that can be protected with a steel plate can be opened to traffic once the plate is in place. Minimum curing times do not apply to precast members because they have adequate strength when installed. However, site-cast materials used to connect the members must have adequate strength. Site-cast concrete used for deck replacement should have a minimum compressive strength of 4,000 psi when subjected to traffic (*3*).

Table 3 presents estimates of the minimum curing times needed before subjecting protective systems to traffic without causing major damage to them. The estimates are based on compressive and bond strength data, tack-free times, and cooling rate data for bituminous concrete obtained from the literature and the responses to the questionnaire sent to the materials suppliers (*3,7–13*). Curing time is a function of the curing temperature of the material, which is a function of the mixture proportions, mass, air and substrate temperature, and degree to which the material is insulated. The values in Table 3 are reported as a function of air temperature for typical installations. Minimum curing times can be reduced by increasing the rate of reactions by adjusting the mixture proportions, applying insulation, and increasing the mass of the application. Bituminous concrete cools more rapidly when placed in thin lifts, and sealers become tack-free sooner when the application rate is reduced.

TABLE 3 MINIMUM CURING TIMES OF RAPID PROTECTION SYSTEMS (HOURS)

| System | Installation Temperature (°F) | | | | References |
|---|---|---|---|---|---|
| | 40 | 55 | 75 | 90 | |
| Bituminous Concrete Overlay on Membrane | NA | 2 | 2 | 2 | 3, 8 |
| Coating | NA | 9 | 3 | 1 | 7, 9 |
| Portland Cement Concrete Overlay | 8 | 6 | 4 | 4 | 10, 11 |
| Penetrating Sealer | 4 | 3 | 2 | 1 | 7 |
| Polymer Overlay | 2* | 6 | 3 | 2 | 7, 12 |
| Other Hydraulic Cement Concrete Overlay | 1* | 1* | 1 | 1 | 10, 13 |

NA: Not applicable since materials are not usually placed at indicated temperature.
* Special cold weather formulation used.

### Permeability to Chloride Ion

A rapid permeability test (AASHTO T277) can be used to measure the permeability to chloride ion of 4-in.-diameter by 2-in.-thick specimens prepared in the laboratory or 4-in.-diameter by 2-in.-thick slices of cores obtained from bridge decks. The results are usually reported in coulombs, which have the relationship to permeability indicated in the footnote to Table 4.

Table 4 presents the permeability to chloride ion of cores taken from decks to which rapid protection systems had been applied and of specimens prepared in the laboratory (*5,7,9,14–17*). Results for specimens tested at early and later ages are reported where data are available to provide an indication of how the permeability changes with age. To properly rank the protective systems, the permeability over the life of the systems needs to be considered. Typically, unprotected bridge deck concretes have a moderate-to-high permeability. The materials used to rehabilitate a deck should have a low permeability to chloride ion unless a protective system will be placed following the crack repair or patching.

### Skid Resistance and Wear

To be used on traffic-bearing surfaces, a protection system must have an adequate skid resistance. Corrective action is required when smooth tire numbers (ASTM E524) are <20 and treaded tire numbers (ASTM E501) are <37. Table 5 presents skid numbers for the protection systems at <1 year of age and at 5 years of age to provide an indication of how the skid resistance changes with age (*5,7,14,18*). As indicated by Table 5, unacceptable skid numbers can be obtained when coatings and some penetrating sealers are applied to screeded concrete surfaces. Coatings and sealers can usually be applied to tined and grooved surfaces as long as the material does not fill the grooves. Freshly placed hydraulic cement concretes can be tined and grooves can be sawcut in the hardened concrete to ensure proper skid resistance. Silica aggregate can be broadcast onto polymer materials to provide good skid numbers.

### Subjective Rating

Subjective ratings of the most rapid protection systems based on performance characteristics, as presented in Table 6, can be used to select the optimum system. As indicated by Table 6, typically the best most rapid protection system (lowest total) is the polymer overlay, and the least desirable system (highest total) is the high-early-strength portland cement concrete overlay. Although the results presented in Table 6 would not necessarily be applicable to every situation, the application of a polymer overlay or penetrating sealer is typically

TABLE 4 PERMEABILITY TO CHLORIDE IONS OF RAPID PROTECTION SYSTEMS

| System | Laboratory Specimens | Cores at Indicated Age | | | References |
|---|---|---|---|---|---|
| | | ≤1 yr | 5 yr | 10 yr | |
| Bituminous Concrete Overlay on Membrane | — | N | — | — | 14 |
| Coating | — | L | — | — | 7,9 |
| Portland Cement Concrete Overlay | L | L | VL | VL | 5, 15, 16, 17 |
| Penetrating Sealer | — | L, M | L, M | — | 7 |
| Polymer Overlay | N | N | VL, L | VL, L | 7, 14 |
| Other Hydraulic Cement Concrete Overlay | VL | — | — | — | 15 |

| Permeability | | Coulombs |
|---|---|---|
| H = High | > | 4,000 |
| M = Moderate | = | 2,000 – 4,000 |
| L = Low | = | 1,000 – 2,000 |
| VL = Very Low | = | 100 – 1,000 |
| N = Negligible | < | 100 |

TABLE 5 SKID NUMBERS AT 40 mph FOR RAPID PROTECTION SYSTEMS

| System | Texture | Smooth Tire | | Treaded Tire | | References |
|---|---|---|---|---|---|---|
| | | ≤1 yr | 5 yr | ≤1 yr | 5 yr | |
| Bituminous Concrete Overlay on Membrane | Compacted | 26 | 28 | 46 | 41 | 14 |
| Coating | Screeded | 7 | — | 7 | — | |
| | Tined | 36 | — | 47 | — | 7, 9 |
| Portland Cement Concrete Overlay | Screeded | — | 28 | 61 | 51 | |
| | Tined | 41 | — | 44 | — | 5, 18 |
| Penetrating Sealer | Screeded | 23 | 34 | 36 | 51 | |
| | Tined | 45 | 45 | 46 | 45 | 7 |
| Polymer Overlay | Tined | 38 | 45 | 45 | 48 | |
| | Sand broadcast | 63 | 36 | 64 | 45 | 7, 14 |

desirable because acceptable skid resistance and permeability to chloride ion can be obtained with negligible effect on ride quality and with short curing times. Also, in situations where traffic begins to back up, these protective systems can be open to traffic in short times to relieve conjestion. On the other hand, bituminuous overlays and high-early-strength portland cement concrete overlays do not lend themselves to use where the most rapid repairs are desired because of the major effect on ride quality and the effort required to remove installation equipment and apply temporary materials to prepare the surface for traffic. Bituminous overlays and portland cement concrete overlays become more desirable as longer times are allowed for lane closure. These systems are much better suited for rapid installations and are particularly well suited for semirapid installations.

## TECHNIQUE TIME DEMANDS

The responses to Questionnaires 1 and 2 concerning the time required to set up and remove traffic control, prepare the surface, and place and cure materials are presented in Table 7 along with the average deck area (in square yards) for which the time estimates were made.

The technique time demands for three of the most used rapid protection systems and three of the most used rapid patching systems are shown in Figures 3 and 4, respectively. Figures 3 and 4 and the data in Table 7 should be useful to bridge engineers when planning rapid repairs for bridge decks.

No time requirement data for precast concrete slab spans, box beams, and channel and tee beams were obtained from the responses to the questionnaires. However, these members can be used for rapid deck replacement when the spans are shorter than 100 ft (*19*).

## SERVICE LIFE AND MAINTENANCE

The responses to Questionnaires 1 and 2 provided sufficient information to estimate the service life of most of the rapid repair systems (see Table 8). The times until minor repairs (maintenance) are required are also presented in Table 8. Service life data obtained from a review of the literature are presented in Table 9 (*7,14,20–34*). Site-cast portland cement concrete decks can be constructed to last 50 years with maintenance in the form of an overlay applied at 25 years of age (*35*). The maintenance and service life estimates were used to determine the life cycle cost for each repair system. It is anticipated that in SHRP Contract Year 4, the influence of rate of corrosion on repair life and the influence of a repair on the service life of a deck will be determined so that more accurate life cycle costs can be computed in Contract Year 5.

TABLE 6 SUBJECTIVE RATING OF MOST RAPID PROTECTION SYSTEMS

| System | Temporary Surfaces | Minimum Curing Time | Permeability | Skid No. | Total | Rank |
|---|---|---|---|---|---|---|
| Bituminous Concrete Overlay on Membrane | 4 | 2 | 1 | 3 | 10 | #5 |
| Coating | 1 | 2 | 3 | 3 | 9 | #4 |
| High Early Strength Portland Cement Concrete Overlay | 3 | 3 | 2.5 | 2 | 10.5 | #6 |
| Penetrating Sealer | 1 | 1 | 3.5 | 2 | 7.5 | #2 |
| Polymer Overlay | 1 | 2 | 2 | 1 | 6 | #1 |
| Other Hydraulic Cement Concrete Overlay | 3 | 1 | 2 | 2 | 8 | #3 |

1 – excellent
2 – very good
3 – good
4 – fair

TABLE 7 TECHNIQUE REQUIREMENTS

| System | Average Area, ($yd^2$) | Average Time Requirements (hr) | | | | Number of Responses Indicating Total Time | | |
|---|---|---|---|---|---|---|---|---|
| | | Traffic Control | Surface Preparation | Placing and Curing | Total | ≤ 8 hr | > 8 ≤ 12 hr | > 12 hr ≤ 21 hr |
| Bituminous Concrete Overlay on Membrane | 587 | 2.5 | 3.7 | 6.5 | 12.7 | 5 | 8 | 12 |
| Coating | 519 | 2.0 | 1.8 | 5.7 | 9.5 | 0 | 3 | 0 |
| Portland Cement Concrete Overlay | 1181 | 0.9 | 2.3 | 5.6 | 8.8 | 2 | 3 | 0 |
| Penetrating Sealer | 673 | 1.5 | 2.2 | 3.4 | 7.1 | 6 | 1 | 0 |
| Polymer Overlay | 481 | 1.2 | 4.0 | 4.7 | 9.9 | 3 | 8 | 1 |
| Other Hydraulic Concrete Overlay | 452 | 0.9 | 4.0 | 3.1 | 8.0 | 1 | 0 | 0 |
| Crack Repair and Sealing | 700[a] | 2.0 | 1.3 | 4.0 | 7.3 | 1 | 1 | 0 |
| Bituminous Concrete Patch | 5 | 0.9 | 0.4 | 0.7 | 2.0 | 6 | 0 | 0 |
| Portland Cement Concrete Patch | 9 | 1.7 | 3.3 | 2.6 | 7.6 | 14 | 9 | 0 |
| Polymer Concrete Patch | 202 | 2.1 | 1.9 | 5.2 | 9.2 | 1 | 2 | 0 |
| Other Hydraulic Concrete Patch | 43 | 1.5 | 2.2 | 3.1 | 6.8 | 6 | 4 | 0 |
| Steel Plate over Concrete | 2 | 0.8 | 1.7 | 2.2 | 4.7 | 1 | 1 | 0 |
| Precast Concrete Deck Panel | 1291 | 1.4 | 4.6 | 5.1 | 11.1 | 1 | 2 | 1 |
| Site-Cast Portland Cement Concrete | 4 | 3.2 | 2.6 | 5.6 | 11.4 | 0 | 3 | 0 |
| Other Site-Cast Hydraulic Concrete | 3 | 1.9 | 2.5 | 3.9 | 8.3 | 2 | 1 | 0 |

[a]Linear feet.

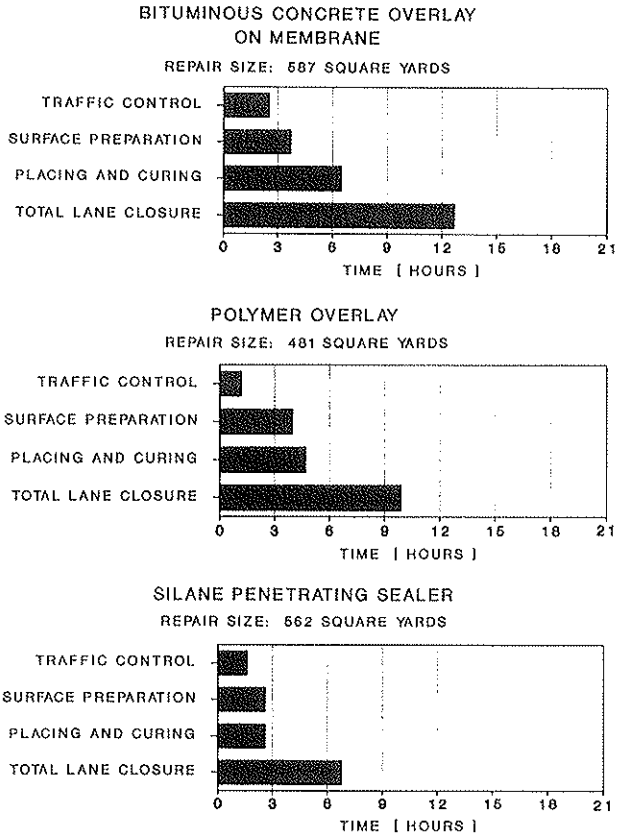



FIGURE 3 Technique time requirements for the three most frequently used rapid protection systems.

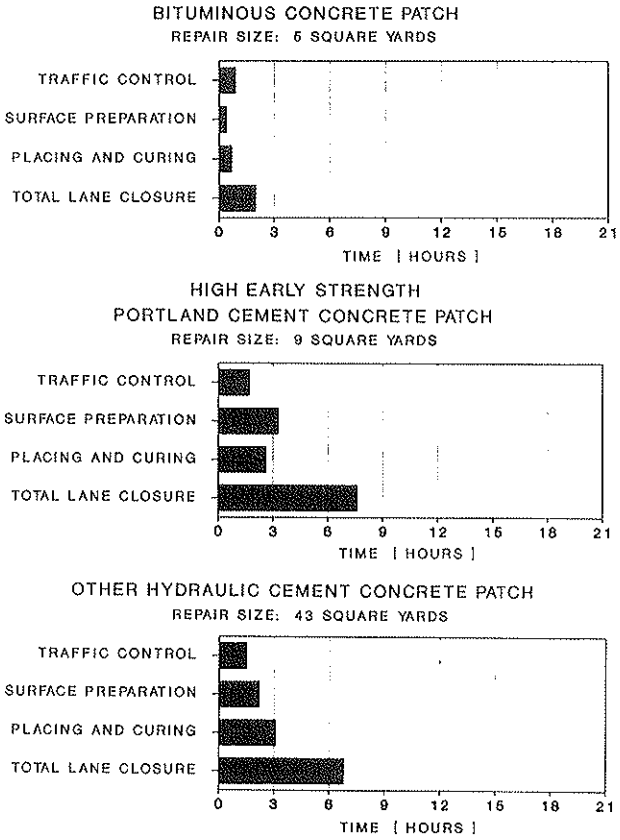



FIGURE 4 Technique time requirements for the three most frequently used rapid patching systems.

TABLE 8 SERVICE LIFE AND MAINTENANCE ON THE BASIS OF QUESTIONNAIRE RESPONSE (YEARS)

| System | Time until Maintenance | | | Service Life | | |
|---|---|---|---|---|---|---|
| | Average | Low | High | Average | Low | High |
| Bituminous Concrete Overlay on Membrane | 5.1 | 1.0 | 10.0 | 11.8 | 4.5 | 20.0 |
| Coating | 5.2 | 2.8 | 10.3 | 10.3 | 5.5 | 20.0 |
| Portland Cement Concrete Overlay | 8.3 | 5.3 | 11.9 | 15.5 | 10.0 | 22.5 |
| Penetrating Sealer | 6.8 | 4.0 | 10.1 | 16.5 | 10.0 | 25.0 |
| Polymer Overlay | 6.4 | 3.0 | 10.0 | 12.7 | 6.0 | 25.0 |
| Crack Repair and Sealing[a] | 7.5 | 5.0 | 10.0 | 15.0 | 10.0 | 20.0 |
| Bituminous Concrete Patch | 0.3 | 0.1 | 0.8 | 1.7 | 1.0 | 3.0 |
| Portland Cement Concrete Patch | 2.8 | 0.3 | 7.0 | 5.9 | 1.8 | 10.0 |
| Polymer Concrete Patch | 10.0 | — | — | 20.0 | 15.0 | 25.0 |
| Other Hydraulic Concrete Patch | 6.3 | 1.0 | 10.0 | 11.9 | 2.0 | 20.0 |
| Steel Plate over Concrete | 10.0 | — | — | 15.0 | — | — |
| Precast Concrete Deck Panel | 20.0 | 12.5 | 30.0 | 38.8 | 30.0 | 50.0 |
| Site-Cast Portland Cement Concrete | 6.2 | 4.0 | 8.0 | 11.7 | 7.5 | 15.0 |
| Other Site-Cast Hydraulic Concrete | 2.0 | — | — | 5.5 | 5.0 | 6.0 |

[a]($/linear foot).

TABLE 9 SERVICE LIFE AND INITIAL COST OF RAPID REPAIR SYSTEMS ON THE BASIS OF LITERATURE REVIEW

| System | Service Life (yrs.) | | | Initial Cost ($/yd$^2$) | | | References |
|---|---|---|---|---|---|---|---|
| | Average | Low | High | Average | Low | High | High |
| Bituminous Concrete Overlay on Membrane | 9.7 | 3.7 | 15.0 | 50.84 | 15.53 | 136.44 | 7, 20, 21, 22, 23 |
| Coating | — | — | — | — | — | — | — |
| Portland Cement Concrete Overlay | 17.9 | 13.6 | 25.0 | 83.21 | 11.19 | 287.75 | 20, 21, 22, 23, 24, 25, 26 |
| Penetrating Sealer | 5.0 | — | — | 5.45 | 2.58 | 9.84 | 7, 23, 27, 28, 29 |
| Polymer Overlay | 10.0 | — | — | 43.55 | 7.03 | 100.08 | 7, 14, 23, 24, 25, 30, 31, 32 |
| Other Hydraulic Concrete Overlay | — | — | — | 6.08 | — | — | 24 |
| Crack Repair and Sealing[a] | 10.0 | — | — | — | — | — | 23 |
| Joint Repair[a] | 3.7 | 3.5 | 3.9 | 78.23 | 77.73 | 78.72 | 21 |
| Bituminous Concrete Patch | 0.6 | 0.1 | 1.0 | 40.57 | 20.01 | 72.24 | 21, 23, 33, 34 |
| Portland Cement Concrete Patch | 14.8 | 4.3 | 35.0 | 202.17 | 164.71 | 239.63 | 20, 21, 23 |
| Polymer Concrete Patch | 5.5 | — | — | 247.07 | — | — | 21 |
| Other Hydraulic Concrete Patch | 3.8 | — | — | 235.16 | — | — | 21 |
| Steel Plate over Concrete | — | — | — | — | — | — | — |
| Precast Concrete Box Beam | 44.1 | — | — | 967.44 | — | — | 21 |
| Precast Concrete Channel and Tee Beam | — | — | — | — | — | — | — |
| Precast Concrete Deck Panel | 25.3 | 24.5 | 26.1 | 852.35 | 822.58 | 882.11 | 21 |
| Site Cast Portland Cement Concrete | 34.8 | 29.6 | 40.0 | 482.39 | 468.84 | 495.93 | 20, 21 |
| Other Site Cast Hydraulic Concrete | 12.5 | — | — | 686.64 | — | — | 21 |

[a]($/linear foot).

## INITIAL COST AND LIFE CYCLE COST

The responses to Questionnaires 1 and 2 provided initial costs for traffic control, surface preparation, placing and curing materials, and other items as presented in Table 10. It was assumed that the cost data were accurate for 1988. Costs obtained from a review of the literature were inflated at the rate of 5 percent per year to provide reasonable values for 1988 (see Table 9).

The information in Tables 8 and 10 was used to estimate the initial cost and life cycle costs for the rapid repair systems presented in Table 11. In order to compute the life cycle costs presented in Table 11, it was assumed that maintenance and system replacement occurred at the time intervals presented in Table 8. The data from Table 9 were used to estimate the life cycle costs presented in Table 12. Because maintenance intervals were not obtained from the literature review, maintenance costs were not included in the life cycle costs presented in Table 12. Present values were calculated for a period of 50 years because present value data based on a 50-year period are available for new decks, and present values calculated for longer than 50 years are not much higher (*35*). Present values were also calculated for a 25-year period because a deck with a high rate of corrosion would not likely be repairable for more than 25 years. In Figure 5, present-value life cycle costs of repair systems based on the surveyed literature are compared with averaged questionnaire responses. Several systems shown in Figure 5 have a present-value life cycle cost based only on one source. It is anticipated that, in SHRP Contract Year 4, more accurate values and precise conclusions will be available as the result of more studies of repair materials and techniques are added to the data base.

## INTERIM CONCLUSIONS

1. Most transportation agencies do not use rapid-repair techniques.
2. The most-used rapid-protection systems are bituminous concrete overlays on membranes, polymer overlays, high-early-strength portland cement concrete overlays, and penetrating sealers.
3. The most-used rapid-patching systems are high-early-strength portland cement concrete patches, bituminous concrete patches, and other hydraulic cement concrete patches.
4. The most-used rapid deck replacement systems are site-cast high-early-strength portland cement concrete and precast concrete deck panels.
5. Most of the rapid-repair techniques can be done with lane closures of 8 hr or less.
6. On the basis of the life cycle cost analysis, the most cost-effective protection system is the application of a penetrating sealer. The most cost-effective patching system is patching

TABLE 10 INITIAL COST OF RAPID REPAIR SYSTEMS ON THE BASIS OF QUESTIONNAIRE RESPONSE (DOLLARS PER SQUARE YARD)

| System | Traffic Control | Surface Preparation | Placing and Curing | Other | Average Total | Low Total | High Total |
|---|---|---|---|---|---|---|---|
| Bituminous Concrete Overlay on Membrane | 3.73 | 3.09 | 15.28 | 2.52 | 24.62 | 1.95 | 44.00 |
| Coating | 0.11 | 4.39 | 11.95 | 0.00 | 16.45 | 6.95 | 24.41 |
| Portland Cement Concrete Overlay | 19.31 | 21.39 | 38.02 | 8.73 | 87.45 | 77.28 | 95.60 |
| Penetrating Sealer | 0.67 | 0.46 | 1.57 | 0.07 | 2.77 | 1.36 | 4.55 |
| Polymer Overlay | 0.73 | 5.68 | 31.35 | 0.64 | 38.40 | 4.00 | 92.99 |
| Other Hydraulic Concrete Overlay | 0.36 | 46.80 | 53.30 | 0.00 | 100.46 | — | — |
| Crack Repair and Sealing | 0.15 | 5.28 | 4.05 | 0.00 | 9.48 | 6.95 | 12.00 |
| Bituminous Concrete Patch | 63.42 | 7.54 | 39.57 | 0.63 | 111.16 | 7.00 | 250.00 |
| Portland Cement Concrete Patch | 30.93 | 108.34 | 119.74 | 7.12 | 266.13 | 15.00 | 611.43 |
| Polymer Concrete Patch | 0.11 | 18.00 | 48.75 | 0.00 | 66.86 | — | — |
| Other Hydraulic Concrete Patch | 32.84 | 31.26 | 102.92 | 14.30 | 181.32 | 3.96 | 527.47 |
| Steel Plate over Concrete | 9.00 | 6.00 | 9.00 | 60.00 | 84.00 | — | — |
| Precast Concrete Deck Panel | 149.37 | 176.29 | 288.55 | 162.44 | 776.65 | 741.94 | 800.00 |
| Site-Cast Portland Cement Concrete | 33.14 | 33.77 | 74.65 | 0.00 | 141.56 | 34.32 | 249.00 |
| Other Site-Cast Hydraulic Concrete | 271.67 | 94.33 | 297.33 | 0.00 | 663.33 | 249.00 | 980.00 |

[a]($/linear foot).

TABLE 11 INITIAL COST AND LIFE CYCLE COST ON THE BASIS OF QUESTIONNAIRE RESPONSE (DOLLARS PER SQUARE YARD)

| Code Number | System | Initial Cost | Present Value Total Cost* | |
|---|---|---|---|---|
| | | | 25-Yr Evaluation Period | 50-Yr Evaluation Period |
| IA | Bituminous Concrete Overlay on Membrane | 24.62 | 42.84 | 55.40 |
| IB | Coating | 16.45 | 31.69 | 41.03 |
| IC | High-Early-Strength Portland Cement Concrete Overlay | 87.45 | 127.08 | 160.77 |
| ID | Penetrating Sealer | 2.77 | 3.90 | 4.90 |
| IE | Polymer Overlay | 38.40 | 63.03 | 81.53 |
| IF | Other Hydraulic Cement Concrete Overlay | 100.46 | — | — |
| IIA | Crack Repair and Sealing** | 9.48 | 14.08 | 17.86 |
| IIC | Patching with Bituminous Concrete | 111.16 | 1,453.69 | 1,884.92 |
| IID | Patching with High-Early-Strength Portland Cement Concrete | 266.13 | 815.22 | 1,057.85 |
| IIE | Patching with Polymer Concrete | 66.86 | 81.36 | 104.88 |
| IIF | Patching with Other Hydraulic Concrete | 181.32 | 312.20 | 403.78 |
| IIG | Temporary Steel Plate over Conventional Concrete Patch | 84.00 | 123.77 | 157.14 |
| IIID | Replacement with Precast Concrete Deck Panel | 776.65 | 724.35 | 874.72 |
| IIIF | Replacement with Site-Cast High Early Strength Portland Cement Concrete | 141.56 | 247.03 | 319.35 |
| IIIH | Replacement with Other Site-Cast Hydraulic Concrete | 663.33 | 2,334.08 | 3,017.19 |

* Parameters: 10% interest rate; 5% inflation rate; maintenance cost 10% of initial cost.
** ($/linear foot).

TABLE 12 INITIAL COST AND LIFE CYCLE COST ON THE BASIS OF LITERATURE REVIEW (DOLLARS PER SQUARE YARD)

| Code Number | System | Initial Cost | Present Value Total Cost* | |
|---|---|---|---|---|
| | | | 25-Yr Evaluation Period | 50-Yr Evaluation Period |
| IA | Bituminous Concrete Overlay on Membrane | 50.84 | 95.90 | 123.21 |
| IB | Coating | — | — | — |
| IC | High-Early-Strength Portland Cement Concrete Overlay | 83.21 | 103.13 | 130.96 |
| ID | Penetrating Sealer | 5.34 | 17.74 | 22.98 |
| IE | Polymer Overlay | 43.55 | 80.27 | 102.96 |
| IF | Other Hydraulic Cement Concrete Overlay | — | — | — |
| IIA | Crack Repair and Sealing** | — | — | — |
| IIB | Joint Repair** | 78.23 | 334.16 | 432.49 |
| IIC | Patching with Bituminous Concrete | 40.57 | 991.02 | 1,283.63 |
| IID | Patching with High-Early-Strength Portland Cement Concrete | 202.17 | 281.82 | 360.28 |
| IIE | Patching with Polymer Concrete | 247.07 | 742.20 | 958.46 |
| IIF | Patching with Other Hydraulic Concrete | 235.16 | 980.81 | 1,268.66 |
| IIG | Temporary Steel Plate over Conventional Concrete Patch | — | — | — |
| IIIB | Replacement with Precast Concrete Box Beam | 967.44 | 843.71 | 1,006.87 |
| IIID | Replacement with Precast Concrete Deck Panel | 852.35 | 849.37 | 1,098.63 |
| IIIF | Replacement with Site-Cast High Early Strength Portland Cement Concrete | 482.39 | 442.27 | 547.01 |
| IIIH | Replacement with Other Site-Cast Hydraulic Concrete | 686.64 | 1,059.77 | 1,372.73 |

* Parameters: 10% interest rate; 5% inflation rate; maintenance cost 10%of initial cost.
** ($/linear foot).

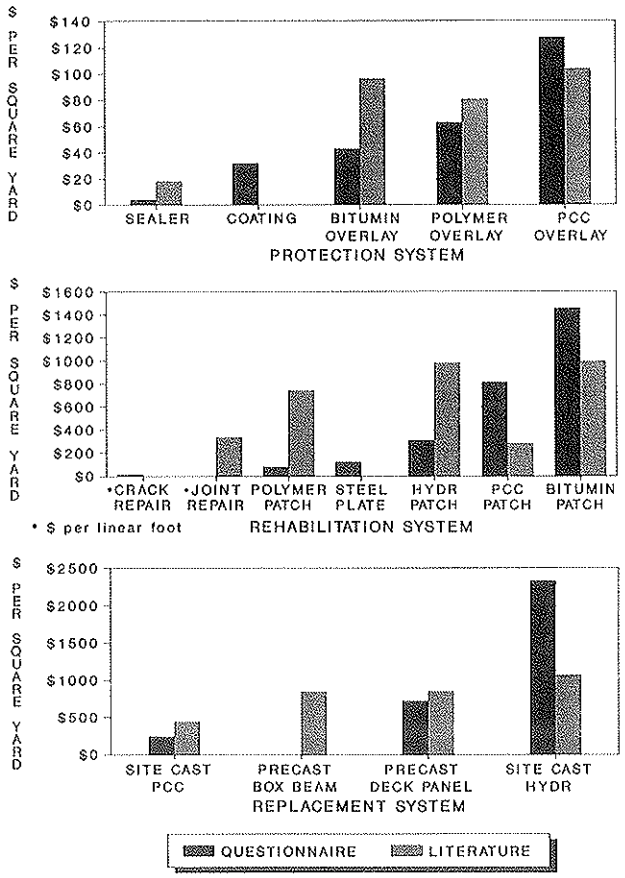



**FIGURE 5 Present value life cycle cost on the basis of a 25-year evaluation period.**

with polymer concrete (based on the questionnaire responses) and patching with high-early-strength portland cement concrete (based on the literature review). The most cost-effective replacement system is site-cast high-early-strength portland cement concrete. High-early-strength portland cement concrete overlays are the most expensive protection systems, and patching with bituminous concrete is the most expensive patching system. Other site-cast hydraulic concrete is the most expensive replacement system. The analysis of some systems was based on a limited data and results can change as more data become available.

7. Information on the effect of the repairs on the service life of a deck and the effect of the rate of corrosion of the rebar in a deck on repair life is needed to make an accurate assessment of life cycle costs.

## REFERENCES


1. R. E. Weyers, N. S. Berke, P. D. Cady, W. P. Chamberlain, J. G. Dillard, P. C. Hoffman, F. Sebba, M. M. Sprinkel, R. E. Swanson, and M. C. Vorster. Proposal, *Strategic Highway Research Program* Contract C103, *Concrete Bridge Protection and Rehabilitation: Chemical and Physical Techniques*, Virginia Polytechnic Institute and State University, Blacksburg, April 8, 1988, Revised July 29, 1988.
2. M. M. Sprinkel, R. E. Weyers, and R. Sellars. *Rapid Repair Techniques for Bridge Decks*. Interim Report 1, *Virginia Polytechnic Institute and State University*, Blacksburg, Sept. 1990.
3. *Road and Bridge Specifications*. Virginia Department of Transportation, Richmond, Jan. 1987.
4. L. I. Knab, M. M. Sprinkel, and O. J. Lane, Jr. *Preliminary Performance Criteria for the Bond of Portland Cement and Latex-Modified Concrete Overlays*. NISTIR 89-4156, National Institute of Standards and Technology, Gaithersburg, Md., Nov. 1989, p. 97.
5. M. M. Sprinkel. High-Early-Strength Latex-Modified Concrete Overlay. In *Transportation Research Record 1204*, TRB, National Research Council, Washington, D.C., 1988, pp. 42–51.
6. E. J. Felt. Resurfacing and Patching Concrete Pavements with Bonded Concrete, *TRB Proc.*, Vol. 35, 1956, pp. 444–469.
7. M. M. Sprinkel. *Comparative Evaluation of Concrete Sealers and Multiple-Layer Polymer Concrete Overlays*. VTRC 88-R2. Virginia Transportation Research Council, Charlottesville, Sept. 1987.
8. *Proceedings of the Association of Asphalt Paving Technologists*, Vol. 39, Feb. 9–11, 1970.
9. M. M. Sprinkel. *Evaluation of the Use of High-Molecular-Weight Methacrylate Monomers to Seal Cracks in Decks on I-81 over the New River*. VTRC 91-R13, Virginia Transportation Research Council, Charlottesville, Sept. 1990.
10. M. A. Temple, R. D. Ballou, D. W. Fowler, and A. H. Meyer. *Implementation Manual for the Use of Rapid Setting Concrete*. Research Report 311-7F, University of Texas, Austin, Nov. 1984.
11. R. L. Carrasquillo and J. Farbiarz. *Pyrament 505 Program 1 Project Progress Report*, University of Texas, Austin, May 8, 1987.
12. L. Kukacka and J. Fontana. *Polymer Concrete Patching Materials: Vol. II Final Report*. Implementation Package 77-11. FHWA, U.S. Department of Transportation, 1977.
13. S. Popovics and N. Rajendran. Early Age Properties of Magnesium Phosphate-Based Cements Under Various Temperature Conditions, In *Transportation Research Record 1110*, TRB, National Research Council, Washington, D.C., pp. 34–95.
14. M. M. Sprinkel. *Evaluation of the Construction and Performance of Multiple-Layer Polymer Concrete Overlays*. Interim Report 2, VTRC 87-R28, Virginia Transportation Research Council, Charlottesville, May 1987.
15. A. Bradbury. *Laboratory Evaluation of Concrete Patching Materials*. Ministry of Transportation of Ontario, Toronto, Nov. 1987.
16. R. L. Carrasquillo. *Permeabilities and Time to Corrison of Pyrament Blended Cement*. Pyrament, Houston, Tex.,
17. C. Ozyildirim. Experimental Installation of a Concrete Bridge Deck Overlay Containing Silica Fume, In *Transportation Research Record 1204*, TRB, National Research Council, Washington, D.C., 1988, pp. 36–41.
18. S. S. Tyson. *Two-Course Bonded Concrete Bridge Deck Construction—Condition and Performance After Six Years*. VHTRC 81-R50. Virginia Transportation Research Council, Charlottesville, May 1981.
19. M. M. Sprinkel. *NCHRP Synthesis of Highway Practice 119: Prefabricated Bridge Elements and Systems*. TRB, National Research Council, Washington, D.C., Aug. 1985.
20. *Monolithic Bridge Deck Overlay Program*. New York State Department of Transportation Bridge Preservation Board, New York, 1986.
21. R. E. Weyers, P. D. Cady, and J. M. Hunter. *Cost-Effectiveness of Bridge Repair Details and Procedures—Part 1: Final Report*. Report FHWA-PA-86-025. Pennsylvania Department of Transportation, Harrisburg, 1987.
22. G. Malasheskie, D. Maurer, D. Mellott, and J. Arellano. *Bridge Deck Protective Systems*. Report FHWA-PA-88-001+85-17. Pennsylvania Department of Transportation, Harrisburg, 1988.
23. W. Chamberlin, P. Hoffman, and R. E. Weyers. *Concrete Bridge Protection and Rehabilitation: Chemical and Physical Techniques, Task 1 Field Survey: First Annual Report*. Virginia Polytechnique Institute and State University, Blacksburg, 1989.
24. P. D. Krauss. *New Materials and Techniques for the Rehabilitation of Portland Cement Concrete*. Report FHWA-CA-TL-85. California Department of Transportation, Office of Transportation Laboratory, Sacramento, 1985.
25. M. M. Sprinkel. *Polymer Concrete Overlay on Beulah Road Bridge: Interim Report 1*. VTRC Report No. 83-R28. Virginia Transportation Research Council, Charlottesville, 1982.

26. D. Bunke. ODOT's Experiences with Silica Fume (Microsilica) Concrete. Presented at 67th Annual Meeting, TRB, National Research Council, Washington, D.C., 1988.
27. D. W. Pfeifer and M. J. Scali. *NCHRP Report 244: Concrete Sealers for Protection of Bridge Structures*. TRB, National Research Council, Washington, D.C., 1981.
28. T. S. Rutkowski. *Evaluation of Penetrating Surface Treatments of Bridge Deck Concretes*. WisDOT Report 81-5. Wisconsin Department of Transportation, Madison, 1988.
29. P. D. Carter and A. J. Forbes. *Comparative Evaluation of the Waterproofing and Durability Performance of Concrete Sealers*. Report ABTR-RD-RR-86-09. Alberta Transportation, Edmonton, 1986.
30. *Technical Report on Flexiolith Epoxy Overlay*. Steinman, Boynton, Gronquist, and Birdsall, New York, 1987.
31. Polymer Concretes Protect Bridge Decks. *Better Roads*, May, 1989, p. 34.
32. P. D. Krauss. Status of Polyester-Styrene Resin Concrete Bridge Deck and Highway Overlays in California. *Proc., 43rd Annual Conference of the Society of the Plastics Industry*, 1988, pp. 1–7.
33. Transportation Research Board. *NCHRP Synthesis of Highway Practice 45: Rapid-Setting Materials for Patching of Concrete*. TRB, National Research Council, Washington, D.C., 1977.
34. *Bituminous Patching*. Report FHWA-TS-78-220. FHWA, U.S. Department of Transportation, 1980.
35. K. Babaei and N. M. Hawkins. *NCHRP Report 297: Evaluation of Bridge Deck Protective Strategies*. TRB, National Research Council, Washington, D.C., Sept. 1987.

*Publication of this paper sponsored by Committee on Structures Maintenance.*

# Optimization Enhancements for an Integrated Bridge Management System

WILLIAM V. HARPER AND KAMRAN MAJIDZADEH

The optimization methodology used in a modular bridge management system (BMS) is described. The optimization module minimizes cost subject to top management's performance objectives. It is a Markovian-based system that stratifies the bridge network to improve degradation predictions. Graphical displays are given that illustrate the results of the BMS, providing an optimal path from current conditions to the desired goals. The system is solved using linear programming on the network level. The use of Lagrange methods and parametric programming allows an efficient integrated solution of the large network optimization in this BMS. This BMS is part of an overall highway maintenance management system, which integrates a pavement management system, a nonpavement management system, and a bridges and structures management system (B&SMS). The B&SMS includes optimization of bridges, tunnels, and culverts. The BMS is the bridge portion of the B&SMS.

Bridges are constructed of one or more spans that vary in length and width from bridge to bridge and can exhibit considerable variations in condition from span to span. Bridges can be rated and modeled in segments (a superstructure span with an abutment or pier). The many components of a bridge may be individually rated on a span-by-span basis and modeled as three structural elements (desk, superstructure, and substructure) at the network level in a bridge management system (BMS). Functional deficiencies such as inadequate load capacity and insufficient deck width may also be included.

The BMS is not dependent on having network data on spans, but the system is designed to accommodate such data. If information is only available for the entire deck, superstructure, and substructure (a condition that is common in the United States), the BMS will operate with this level of information at the network level.

Figure 1 shows the modules that this BMS comprises. The modules have been described by Harper et al. (*1,2*). The condition module uses surveyed condition rating data to derive condition states that characterize the overall condition of each bridge segment. Condition modeling begins with the surveyed condition rating (SCR) values assigned to the bridge components on a span-by-span basis. The following scale is being used by the Kingdom of Saudi Arabia; however, any similar scale, such as the 0 to 9 scale used by the FHWA, can be accommodated.

| *Rating* | *Definition* |
|---|---|
| 7 | Like new |
| 6 | Good condition |
| 5 | Insignificant deterioration |
| 4 | Minimum adequacy |
| 3 | Not functioning as designed |
| 2 | Structurally inadequate |
| 1 | Potentially hazardous |
| 0 | Beyond repair |

The SCR values of the components of the deck (e.g., deck surface and deck structure), superstructure (e.g., primary and secondary members), and substructure (e.g., pedestals and capbeams) are used to derive composite condition index (CCI) values for each structural element; CCI values are, in turn, translated into condition levels. The various configurations of condition levels are used to construct the core condition states. Equations convert the SCR values to the CCI values. User-defined thresholds can be incorporated so that when a feature is rated at, say, less than 4, the CCI value would be modified by or assigned the value of the lowest SCR. Different equations are used for certain bidge types (*1,3*).

The various combinations of CCI ratings for the structural elements making up each segment are used to define core condition states that represent the overall condition of that segment. The CCI values are translated into one of four condition level descriptors for each of the three elements, according to the following scheme:

| *Range of CCI Values* | *Condition Level* |
|---|---|
| From 6 to 7.00 | Good |
| From 4 to 5.99 | Fair |
| From 2 to 3.99 | Poor |
| From 0 to 1.99 | Critical |

Core condition states are defined as possible combinations of condition levels for the elements that make up the structural segment. There are 64 ($4^3$) core condition states. Additional parameters are added depending on the needs of the organization implementing the BMS. Typical examples include element-age parameters (e.g., superstructure age) or various functional deficiency parameters (e.g., insufficient deck width).

The maintenance and rehabilitation (M&R) scopes module contains 40 M&R scopes. The impact of each M&R scope depends on the current condition state of the segment. This is modeled in the prediction module described in following paragraphs. The M&R scopes provide input to the prediction, cost, optimization, packaging, and comparator modules.

The prediction module for structural degradation develops and updates estimates of transition probabilities that are defined as the probability that a structural segment in Condition State $i$ in Stratum $s$ will be in Condition State $j$ in 1 year, given M&R Scope $a$. These estimates are updated each year with new survey data for each stratum using a Bayesian up-

Resource International, 281 Enterprise Drive, Westerville, Ohio 43081.

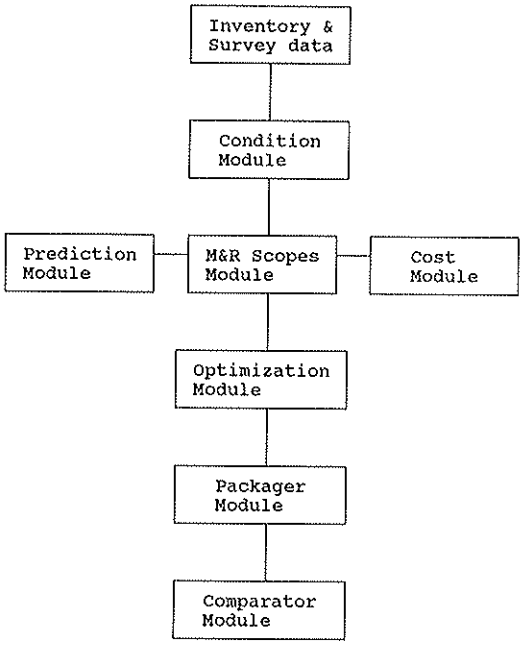



**FIGURE 1 Structure of BMS.**

dating procedure (*2,3*). The cost module uses historical cost data, condition states, M&R scopes, and other inputs to calculate M&R scope costs.

The optimization module has three network level models based on Markovian decision models using linear programming techniques (*2,3*). Subject to the desired performance goals in the optimization, it is the interaction of cost and transition probabilities that determines the optimal (minimal cost) policy. The three models are as follows:

1. A steady-state model to establish steady state minimum cost goals for each stratum,
2. A multiyear model to optimize expenditures within a desired time horizon leading to steady state for each stratum, and
3. A financial exigency model to force the total network (all strata) to meet a specified budget if sufficient funds are not available.

These models provide their results in terms of the proportions of each stratum that should receive various maintenance scopes for different condition states. The first two models are similar to the Arizona pavement models (*4*) in mathematical structure. Strata have been developed to group bridges that exhibit similar degradation patterns, and that have approximately the same M&R scope costs.

The packaging module packages the optimized network solutions into individual work projects. In the project level analyses by the packager, maintenance costs identified by the optimizer are more accurately assessed using actual material quantities and contractor prices.

The comparator module serves as quality control on the performance of the BMS. It provides necessary comparisons of both cost and predictive capabilities of the models against actual experience when the BMS solutions are implemented.

## ANNUAL USE OF BMS

This section briefly describes the steps in the annual usage of BMS. To supplement Figure 1, Figure 2 shows the flow from one optimization model to another. Management input, cost parameters, transition probabilities, and condition survey data are necessary to run the suite of models. This process is iterative. Looping backwards may be necessary if satisfactory results are not obtained. The BMS annual usage scenario has the following steps (*2,3*):

1. Perform the condition survey and update the transition probabilities.
2. Make policy decisions regarding performance objectives.
3. Run the steady state optimization model for all strata. If the resulting steady state budget is acceptable, the output of the steady state model is used to develop constraint equations for the multiyear model. If the steady state budget is not acceptable, then management has to lower the performance objectives set in Step 2 and rerun the steady state optimizer. The final result of this model becomes a goal to be reached in the last year of the planning horizon for the multiyear model.
4. The multiyear model is solved for all strata to determine the optimal maintenance policy and the expected expenditures for each year in the planning horizon.
5. If the budgeting requirements from the multiyear model are too high, then the financial exigency model is run. The financial exigency solution yields the optimal first-year maintenance policy that stays within the first-year budget while at the same time computing the resulting additional expenditures needed to successfully achieve the performance objectives for the remaining years of the planning horizon.
6. Run the packager module to divide the M&R scopes into the detailed maintenance actions that are necessary for the selected bridge projects.
7. At the end of the fiscal year, the comparator module is run to provide feedback on the performance and implementation of the BMS.

## BMS NETWORK OPTIMIZATION MODULE

As described earlier, the three optimization models are the steady state model, the multiyear model, and the financial exigency model. The steady state model is used to establish the long-term goals that provide targets for the multiyear and financial exigency models. The multiyear model addresses the year-by-year maintenance needs for the planning horizon. The steady state and multiyear models solve separate linear programs for each stratum. The financial exigency model imposes a network-wide budget constraint across all strata if insufficient budget is available to satisfy the sum of the individual stratum multiyear models.

The optimization models are used to develop a set of maintenance plans for a bridge system over the desired planning horizon. The steady state model described by Harper et al. (*2,3*) provides the goals for the final year of the multiyear model. Results from the steady state model were provided by Harper et al. (*5*). Similarly, the complete mathematical description for the multiyear model is provided by Harper et

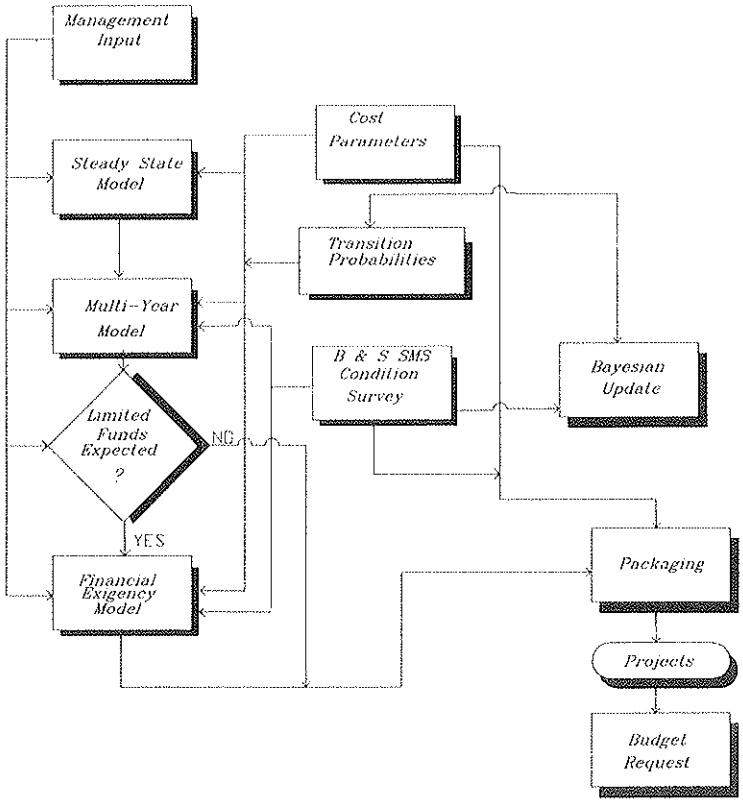



**FIGURE 2 Interconnection of optimization models.**

al. (*2,3*). Selected parameters from the multiyear model that are important in the subsequent financial exigency section are as follows:

$I$ = $(1, 2, \ldots, n)$, index set of condition states.
$S$ = $(1, 2, \ldots, m)$, index set of bridge strata.
$M_i$ = $(a_1, a_2, \ldots, a_m)$, index set of feasible maintenance Scopes $a$ for bridge segments in Condition State $i$.
$C_{ia}(s)$ = average cost of applying maintenance Scope $a$ to one bridge segment in Stratum $s$ and Condition State $i$.
$N(s)$ = number of segments in Stratum $s$.
$r$ = discount rate (= 0.0 in this paper) for computing net present value.
$w^t_{ia}(s)$ = proportion of the segments in Stratum $s$ that is in condition State $i$ and should receive maintenance Scope $a$ in Year $t$. These are the optimization output decision variables.
$E^t(s)$ = expected expenditures in Year $t$ in Stratum $s$.

The multiyear optimization model for Stratum $s$ in which $T$ (= 6 for the runs in this paper) represents the year in which the steady state goals are met follows:

$$\text{Minimize} \sum_{t=1}^{T-1} \sum_{i \in I} \sum_{a \in M_i} (1 + r)^{1-t} w^t_{ia}(s) C_{ia}(s) \qquad (1)$$

The objective function given in Expression 1 minimizes the average present cost per segment of maintenance over the time horizon of interest. To get $E^t(s)$ (the necessary budget for Stratum $s$ for a given Year $t$), the following calculation is necessary:

$$E^t(s) = N(s) \sum_{i \in I} \sum_{a \in M_i} w^t_{ia}(s) C_{ia}(s) \qquad (2)$$

Selected condition states are designated desirable or undesirable. Top management sets goals (lower bounds for desirable and upper bounds for undesirable) for each year of the planning horizon. The multiyear model results may be summarized by the optimal desirable and undesirable percentages that are predicted from the model. Figures 3 and 4 show the desirable and undesirable percentages that resulted from a typical multiyear optimization for a given stratum.

## FINANCIAL EXIGENCY MODEL

When the multiyear model is run for all strata, the sum of the first year budgets from each stratum may exceed the available network budget. When this condition occurs, the financial exigency model is used. The financial exigency model links all strata together using Lagrange methods on a first-year network budget constraint. It would not be efficient (or feasible with some linear programming packages) to jointly solve the network optimization problem by pooling all the separate stratum linear programs. The use of Lagrange methods allows this problem to be solved in an efficient, straightforward manner.

The purpose of the financial exigency model is similar to that of the multiyear model, but it also incorporates the net-

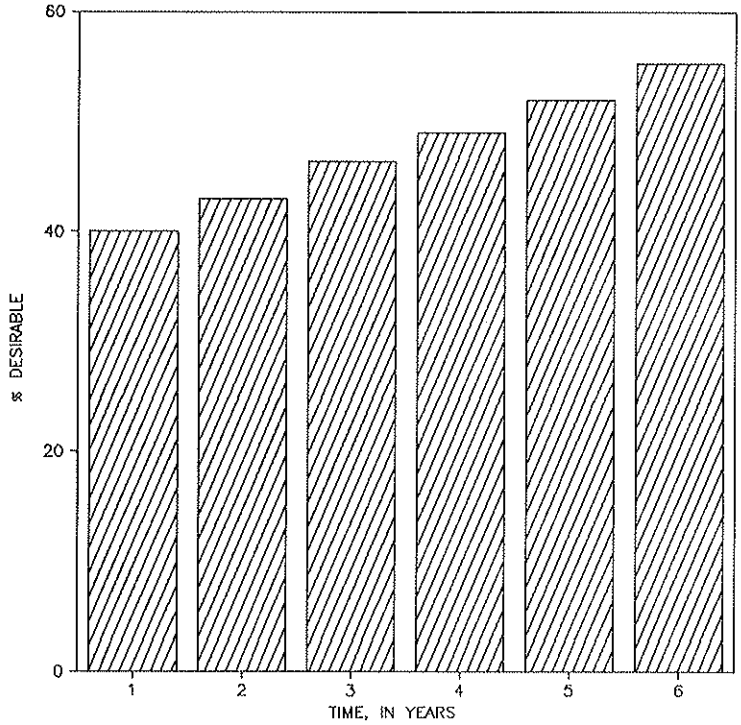



**FIGURE 3** Desirable percentage versus time.

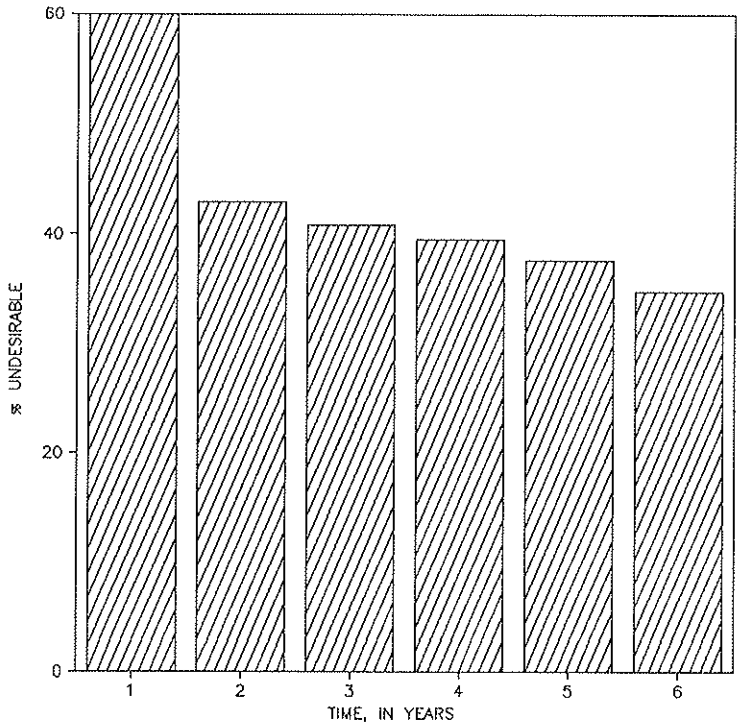



**FIGURE 4** Undesirable percentage versus time.

work budget constraint for the first year of the planning horizon. Although the financial exigency model combines all strata together through the budgetary constraint, it decomposes the overall problem into linear programming problems for individual strata. The financial exigency model also allows the relaxation of the second-year goals if necessary to meet the first-year budget target. The financial exigency model objective function in Expression 3 is solved by determining the optimal value for the Lagrange multiplier $\alpha$ (*6*). If necessary, there are three phases (A, B, and C) of the financial exigency model that can be used to find an optimal solution that meets the available first-year budget. The objective function for this model is as follows:

Minimize

$$\sum_{s\in S} N(s) \sum_{i\in I} \sum_{a\in M_i} \left[ \sum_{t=2}^{T-1} (1 + r)^{1-t} w_{ia}^t(s) C_{ia}(s) + \alpha w_{ia}^1(s) C_{ia}(s) \right] \quad (3)$$

subject to the constraints of the multiyear model for all $s \in S$, and with $\beta^1$ = available budget for the first year.

Different values of $\alpha$ will yield solutions that expend different amounts in year one. If for a given $\alpha$, the solution prescribes a policy that expends too much money in the first year, a new solution can be obtained for a larger value of $\alpha$ that will expend a smaller amount in Year 1. For $\alpha = 1$, this objective function is identical to the multiyear model.

The value of $\alpha$ that produces the solution in which the total of all first-year expenditures among the strata is as close to (but less than or equal to) the first-year budget, $\beta^1$, results in a solution that is a globally optimal for the original financial exigency model (*6,7*). The first-year budget is a monotonically decreasing function of $\alpha$.

Parametric programming on the objective function allows the financial exigency problem to be solved with minimal computational burden. In order to make the financial exigency objective consistent with parametric programming features found in some linear programming packages, the objective function in Expression 3 may be rewritten as Expression 4:

Minimize

$$\sum_{s\in S} N(s) \sum_{i\in I} \sum_{a\in M_i} \left[ \sum_{t=2}^{T-1} (1 + r)^{1-t} w_{ia}^t(s) C_{ia}(s) + (\alpha_{\min} + \Theta) w_{ia}^1(s) C_{ia}(s) \right] \quad (4)$$

subject to multiyear constraints for all $s \in S$.

Thus, $\alpha$ has been replaced by $(\alpha_{\min} + \Theta)$. For Phases A and B, $\alpha_{\min} = 1.0$, whereas $\alpha_{\min} < 1.0$ for Phase C. Then $\Theta$ will range from 0.0 to $\Theta_{\max}$ ($\Theta_{\max}$ may be different for each phase) for the financial exigency runs.

Before describing the financial exigency algorithm, a brief summary of each phase is as follows:

- Phase A. The Year 2 goals are the same as the multiyear model.
- Phase B. Relaxes the Year 2 goals so that the current percentages desirable and undesirable (on the basis of the condition survey) are maintained.
- Phase C. Completely removes the Year 2 goals and attempts to spend as much money as possible while meeting the network level budget.

The goals referred to are the percentages desirable and undesirable that were set by top management for the multiyear model. One of the advantages of the three phases is that it is not necessary to go back to top management and request revised goals. The goals specified by top management are assumed to have had Year 2 goals that improved on the current conditions found in the stratum. If this is not the case, then Phase B may be skipped.

The algorithm used for the financial exigency problem applies to all phases. $\Theta_{\max}$ is determined from initial runs of the BMS or may be set to an arbitrarily large number. Using parametric programming, the entire continuum is spanned. For any given level of $\Theta$, there is a total first-year BMS budget, $B_\Theta^{\text{tot}}$ that is calculated as follows:

$$B_\Theta^{\text{tot}} = \sum_s E_\Theta^1(s) \quad (5)$$

The only difference between the three phases is in the second-year performance goals as described earlier. All three phases use the following algorithm to find the optimal solution ($\Theta_{\text{opt}}$, $B_{\Theta_{\text{opt}}}^{\text{tot}}$):

For $\Theta = 0$ to $\Theta_{\max}$, compute $B_\Theta^{\text{tot}}$. If $B_\Theta^{\text{tot}} \leq \beta^1$ (available first-year network budget), output $\Theta_{\text{opt}} = \Theta$, $B_{\Theta_{\text{opt}}}^{\text{tot}} = B_\Theta^{\text{tot}}$. This is the optimal solution. Stop.

It is possible that $B_{\Theta_{\max}}^{\text{tot}}$ may not satisfy the desired first-year budget constrain for Phase A. In this case, Phase B changes the second-year performance goals to match the desirable and undesirable proportions in the current survey. Thus, instead of endeavoring to improve the second-year performance as it is anticipated will be the case for the multiyear model (and Phase A), the stratum desirable and undesirable percentage goals are set to maintain the existing stratum conditions. Then the preceding algorithm is used to search for an optimal solution to this modified set of Year 2 goals for Phase B.

If Phase B cannot find a solution that meets the available budget, then more drastic measures are necessary. Phase C completely removes the Year 2 goals and will spend as much money as possible while still meeting the first-year budget. In Phase C, the first-year M&R scope costs vary from inexpensive [start with $\alpha_{\min} C_{ia}(s)$] to more expensive [$\alpha_{\max} C_{ia}(s)$]. At the low end of this range, the first-year expenditures will be high because of the apparent inexpensive M&R scope costs. As $\alpha$ increases, the first-year expenditures will decrease until finally the budget goal is met. Top management will have to examine the resulting performance and decide if additional funds should be requested.

Figure 5 is the data flow diagram for the financial exigency model. The open-ended boxes are data base tables from the ORACLE relational data base that ties the entire highway maintenance management system (HMMS) together. The HMMS consists of a bridges and structures management system (B&SMS), of which this BMS is a part, linked not only

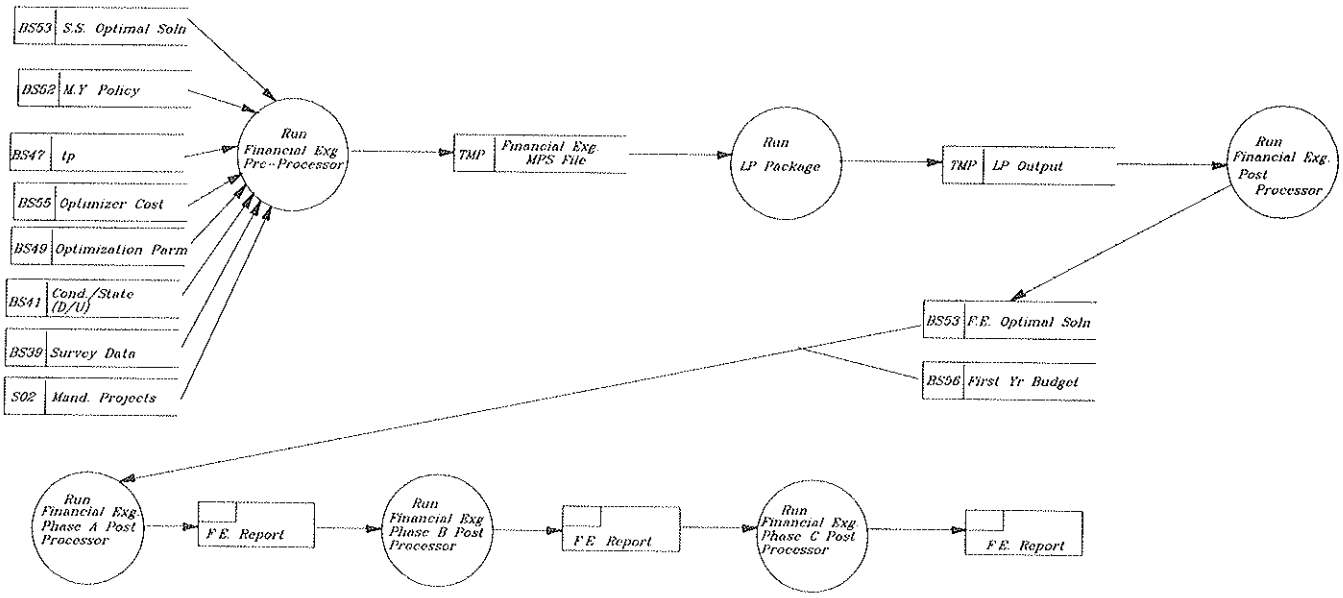



**FIGURE 5 Data flow diagram for financial exigency model.**

to pavement but also to nonpavement management systems. It is an integrated system that allows the optimal allocation of resources across the entire network of highways, bridges, culverts, tunnels, and all related nonpavement elements. The circles represent the processes that are part of the financial exigency model, and the closed rectangles represent various outputs. A complete description of this data flow diagram was provided by Harper et al. (*5*).

The left hand side of the data flow diagram in Figure 5 shows the following major inputs to the financial exigency preprocessor:

1. Steady-State Optimal Solution—provides the target for the final year of the financial exigency model.
2. Multiyear Policy—provides the yearly desirable and undesirable goals.
3. Transition Probabilities—provide the degradation estimates.
4. M&R Scope Costs (Optimizer Cost)—provide the scope costs as a function of the current condition state.
5. Optimization Tolerance Parameters—provide the tolerances on how closely the steady state solution must be met at the end of the planning horizon.

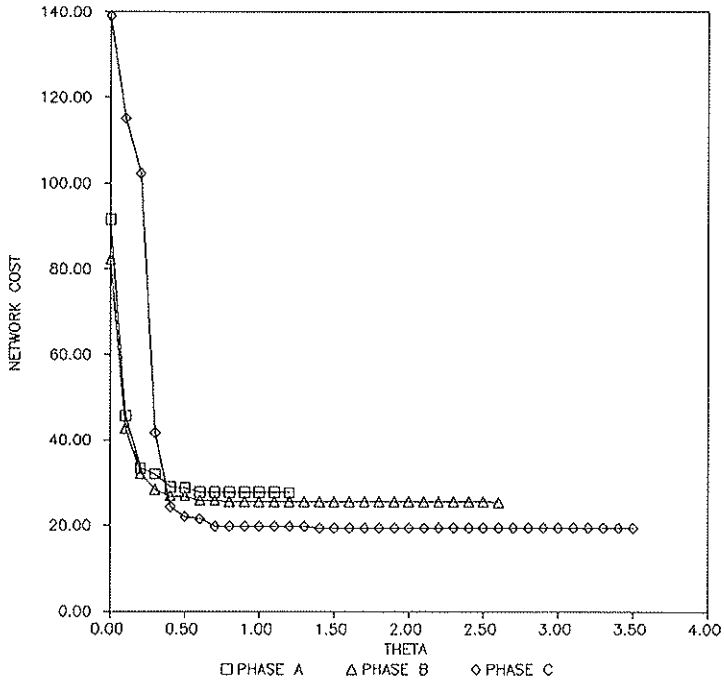



**FIGURE 6 Network cost versus Θ.**

6. Desirable/Undesirable Condition State—defines each condition state as desirable, undesirable, or neither.
7. Survey Data—provides the first-year boundary conditions giving the proportions of the stratum that are in the various condition states.
8. Mandatory Projects—these projects must be implemented and are not subject to change by the optimization.

The inputs are used to create temporary input files in the MPS program by the financial exigency preprocessor. The MPS files are input to the commercial linear programming (LP) package that produces a temporary output file. The postprocessor uses the LP output to generate the various optimal solutions that will be examined in Phases A, B, and C of the financial exigency model. The first-year budget constraint is used to select the solution set across all strata that meets the available budget. Each phase of the financial exigency model produces a report that indicates how the network budgetary needs change as a function of $\Theta$.

Figure 6 shows example results for Phases A, B, and C. As the Lagrange multiplier increases, the network cost decreases rapidly. It is highly probable that Phase A will be the only financial exigency phase needed to find a network budget that satisfies the available first-year budget. If Phase A cannot reduce the network budgetary needs enough, then Phases B and C may be used to find the optimal solution that will meet the available budget.

## CONCLUSION

A BMS was described that integrates separate strata into an overall network optimal solution. Similar techniques are used to link the BMS both with pavement and nonpavement management systems. The use of Lagrange methods and parametric programming allows an efficient solution to the multistrata Markov decision process.

## REFERENCES

1. W. V. Harper, A. Al-Salloum, S. Al-Sayyari, S. Al-Theneyan, J. Lam, and C. L. Helm. Selection of Ideal Maintenance Strategies in a Network Level Bridge Management System. In *Transportation Research Record 1268*, TRB, National Research Council, Washington, D.C., 1990, pp. 59–67.
2. W. V. Harper, J. Lam, A. Al-Salloum, S. Al-Sayyari, S. Al-Theneyan, G. Ilves, and K. Majidzadeh. Stochastic Optimization Subsystem of a Network Level Bridge Management System. In *Transportation Research Record 1268*, TRB, National Research Council, Washington, D.C., 1990, pp. 68–74.
3. W. V. Harper, A. Al-Salloum, J. Lam, S. Al-Sayyari, and S. Al-Theneyan. *B&SMS Conceptual Framework*. Report D3.1, Highways Maintenance Associates (Resource International), Columbus, Ohio, April 1990.
4. K. Golabi, R. B. Kulkarni, and G. B. Way. A Statewide Pavement Management System. *Interfaces*, Vol. 12, Dec. 6, 1982, pp. 5–21.
5. W. V. Harper, G. Vinokur, C. L. Helm, and K. Majidzadeh. *B&SMS Submodel Development*. Report D3.2, Highway Maintenance Associates (Resource International), Columbus, Ohio, Feb. 1991.
6. C. A. Mount-Campbell. *PMS Conceptual Framework*. Report C1, Highway Maintenance Associates (Resource International), Columbus, Ohio, March 1990.
7. H. Everett. Generalized Lagrange Multiplier Methods for Solving Problems for Optimum Allocation of Resources. *Operations Research*, Vol. 11, 1963, pp. 399–417.

*Publication of this paper sponsored by Committee on Structures Maintenance.*

# Bridge Deck Condition Surveys Using Radar: Case Studies of 28 New England Decks

KENNETH MASER

Repair and replacement of deteriorated bridge decks represent a major expense to many state highway agencies. Current techniques for assessing deck condition have limited the effectiveness of efforts to program, order by priority, and estimate maintenance and rehabilitation (M&R) projects. A research program sponsored by 5 New England states led to the development of ground penetrating radar as a rapid and accurate means for deck deterioration assessment. The program involved surveys of 32 asphalt-overlaid decks in the region, 28 of which were studied during maintenance for deterioration quantities. Before maintenance, radar was collected on all of these decks and analyzed. Analysis techniques were developed to predict the concrete deterioration from the variations in the concrete dielectric constant as computed directly from the radar waveforms. The computation was used to predict overall deterioration for each deck and each major span. This prediction was then correlated with the actual deck deterioration determined when the asphalt overlay was removed, and the bare concrete was visually examined and chain-dragged. Correlations were carried out both at a detailed project level (100 percent coverage) with underside survey, and at a network level (30 percent coverage). The project-level correlation produced a good fit ($R^2 = 0.83$), with standard error of $\pm 4.1$ percent of the deck area. The network-level correlation produced a reasonable fit ($R^2 = 0.72$) with standard error of $\pm 5$ percent of the deck area. Both project and network survey methods have subsequently been implemented at highway speed, at costs comparable to traditional survey methods.

Repair and replacement of deteriorated bridge decks represent a major expense to many state highway agencies. During the life of a bridge, the deck is typically replaced once and repaired frequently. Bridge deck deterioration is primarily caused by two mechanisms: (a) Freeze-thaw damage to the concrete (punky concrete), and (b) corrosion-induced delamination resulting from infiltration of chlorides introduced by winter road salting operation or by a saline environment (see Figure 1).

One of the major problems with bridge deck deterioration is that its severity and extent are difficult to assess. The mechanisms of deterioration occur below the surface, and their manifestations are not readily seen in visual inspections. This difficulty is particularly true for overlaid decks, on which both delamination and freeze-thaw damage can occur without visual manifestations. Consequently, agencies are forced to program, order by priority, and budget the repair and replacement of many structures whose condition is virtually unknown. This situation has led to major surprises during construction, and to overruns and overrepairs.

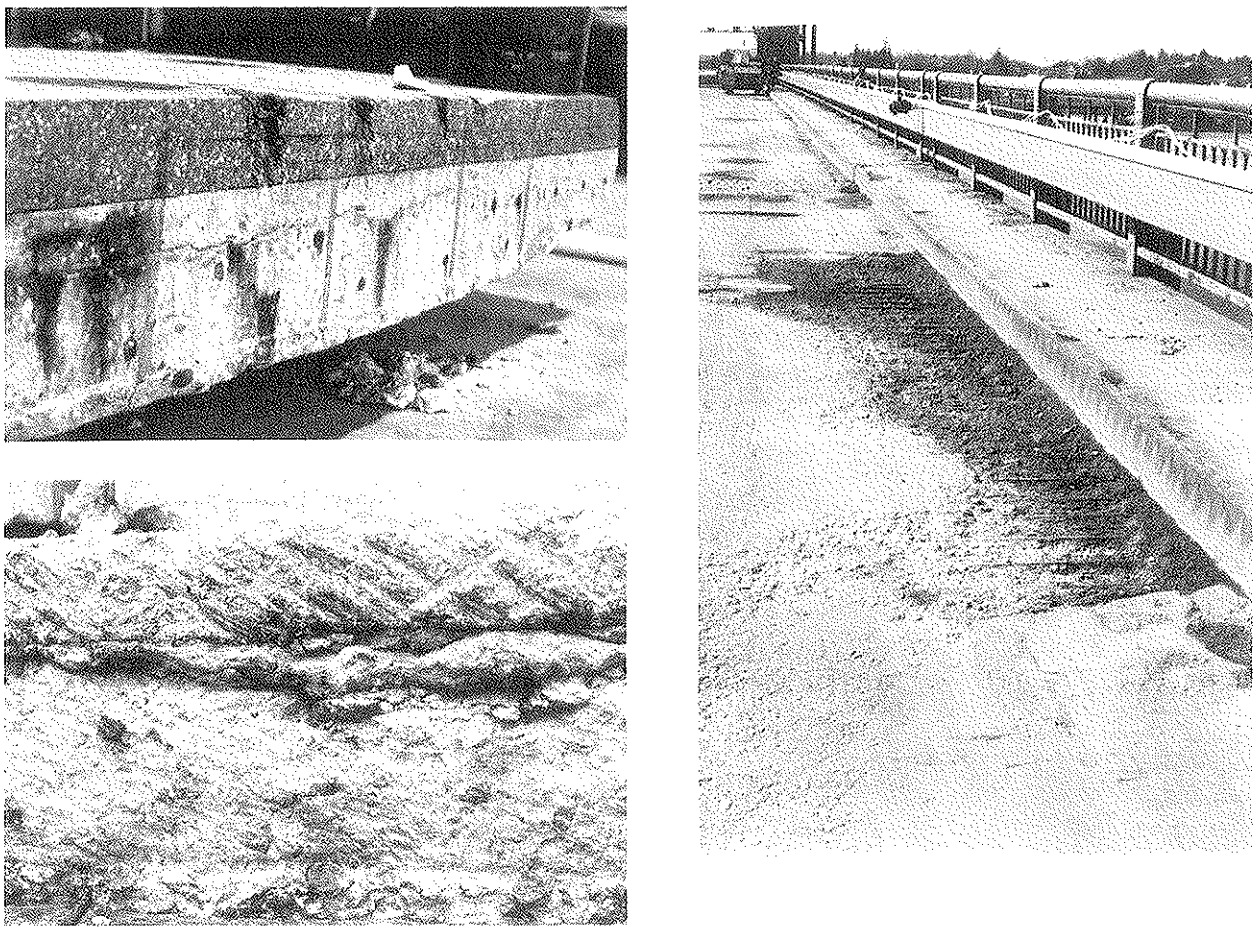
Current techniques for condition assessment of overlaid bridge decks are slow, labor intensive, intrusive to traffic, and unproductive of accurate estimates of quantity of deteriorated concrete. These techniques, which include core sampling, corrosion (half-cell) potentials, and chloride ion measurements, are well documented (*1*). Corrosion potentials and chloride ion measurements infer corrosion, but do not address unseen freeze-thaw damage. A more reliable technique, the chain drag, does not work either with asphalt overlays or in heavy traffic conditions with high ambient noise.

In recognition of this problem, a group of five New England States (New Hampshire, Vermont, Maine, Rhode Island, and Massachusetts), under the New England Transportation Consortium (NETC), sponsored a program carried out by the Massachusetts Institute of Technology (M.I.T.) to investigate the potential of new technology for bridge decks. These states almost exclusively have asphalt overlays, so the problem of deck assessment is shared. The participants in this program reflected a growing concern over their ability to keep up with the bridge deck deterioration problem in the future. The major concern was not the badly deteriorated decks, because these decks already displayed obvious surface spalling or severe underside indications. Rather, the problem was with the decks in the grey zone, where there were limited visual indications, and the degree of deterioration could vary from 0 to 40 percent of the deck area. A large population of decks built in the Interstate construction period fall into this area of concern.

The specific objective of the NETC program was to investigate two new technologies for bridge deck assessment, ground penetrating radar and infrared thermography. The approach was to conduct radar and infrared surveys on a group of asphalt-overlaid decks that were scheduled for maintenance. During maintenance, the asphalt was removed and the concrete surface observed and chain-dragged to determine deterioration quantities for removal. These quantities were correlated with the predictions from the radar and infrared surveys. These field correlations were carried out on 28 decks in the New England area. The field study was complemented by theoretical studies (*2*,*3*) and by laboratory studies on deck slabs recovered from the field.

The results of this study led to the establishment of a radar-based technique that produced accurate correlations with observed deterioration. Results for infrared thermography were

Infrasense, Inc., 19-R Brookline Street, Cambridge, Mass. 02139.



(a) (b)

**FIGURE 1 Bridge deck deterioration: (a) delamination, (b) punky concrete.**

less favorable. The remainder of this paper will focus on the radar technology, and will include principles of radar for deck evaluation, deck site selection, survey procedures, data analysis, correlation with field observations, and current survey techniques that have evolved from this project. For the results of the infrared studies, the reader is referred to the project final report (*2*).

## PRINCIPLES OF GROUND-PENETRATING RADAR

Ground-penetrating radar operates by transmitting short pulses of electromagnetic energy into the pavement using an antenna attached to a survey vehicle (see Figure 2). These pulses are reflected back to the antenna with an arrival time and amplitude that are related to the location and nature of dielectric discontinuities in the material (air-asphalt or asphalt-concrete, reinforcing steel, etc). The reflected energy is captured and may be displayed on an oscilloscope to form a series of pulses that are referred to as the radar waveform. The waveform contains a record of the properties and thicknesses of the layers within the deck, as shown schematically in Figure 3. Figure 4 shows a typical set of bridge deck waveforms collected during the NETC project.

Bridge deck deterioration can be inferred from changes in the dielectric properties of the concrete (*3*). Concrete that has high moisture and chloride contents, as associated with corrosion damage and punky concrete, will produce a large reflection at the asphalt-concrete boundary (Reflection 2 in Figure 3). This reflection is caused by the higher dielectric permittivity produced by the moisture and chloride.

Figures 5 and 6 show the results of a numerical study of the sensitivity of the concrete reflectivity to moisture and chloride content. The study was carried out using electromagnetic models for predicting radar waveforms from concrete and asphalt material properties (*4*). The ratio, R2, is defined as the amplitude of the reflection from the top of the concrete normalized by the amplitude of the reflection from the top of the asphalt (*5*). The figures clearly show that both moisture and chloride content increase the reflectivity of the concrete.

Other indicators of concrete deterioration have been proposed in other investigations (*5–9*). These indicators, however, are all sensitive to the cross-sectional geometry of the deck, including asphalt thickness and rebar spacing and depth.

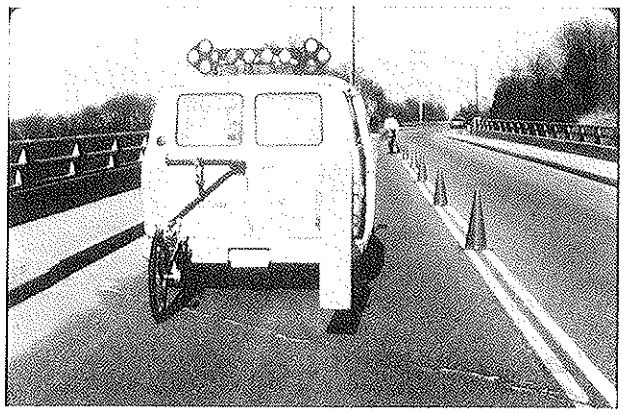


**FIGURE 2 Radar data collection.**

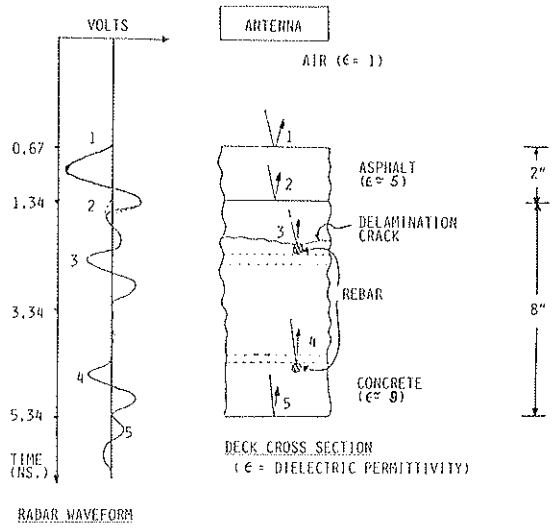



**FIGURE 3 Radar bridge deck model.**

The geometric sensitivity of these indicators decreases their usefulness as indicators of deterioration, because geometry will vary within a deck and from deck to deck. The concrete reflectivity, however, is unaffected by cross-sectional geometry except when interference occurs because of thin asphalt (less than 2 in.) or shallow rebar cover (less than 1 in.). Under these circumstances, signal processing techniques are required to reveal the true value of the concrete reflectivity.

The deterioration determination is made by computing the concrete dielectric constant, $\varepsilon_c$, from the reflectivity, R2. This computation is based on the reflection coefficient between the asphalt and the concrete. The reflection coefficient between any two layers (1 and 2) is defined as the ratio of the amplitude of the incoming wave to the amplitude of the reflected wave. The reflection coefficient is related to the contrast in dielectric properties between the two layers, as follows:

$$\text{Reflection Coefficient}\,(1\text{–}2) = (\varepsilon_1^{1/2} - \varepsilon_2^{1/2})/(\varepsilon_1^{1/2} + \varepsilon_2^{1/2}) \qquad (1)$$

where $\varepsilon$ is the dielectric constant, and Subscripts 1 and 2 refer to the successive layers. The dielectric constant of the asphalt can be determined at the air-asphalt interface by recognizing that the dielectric constant of air is 1. Using the reflection from a metal plate on the pavement surface to represent the incident wave (because the metal plate reflects 100 percent), Equation 1 can be rearranged to yield the asphalt dielectric constant, $\varepsilon_a$, as follows:

$$\varepsilon_a = [(1 + A/A_{pl})/(1 - A/A_{pl})]^2 \qquad (2)$$

where

$A$ = amplitude of reflection from asphalt (see Figures 3 and 4), and
$A_{pl}$ = amplitude of reflection from metal plate (= negative of incident amplitude).

Equations 1 and 2 can now be combined with some additional manipulation to yield the dielectric constant of the concrete, $\varepsilon_c$, as follows:

$$\varepsilon_c = \varepsilon_a\,[(F - \text{R2})/(F + \text{R2})]^2 \qquad (3)$$

where

$$F = (4\varepsilon_a^{1/2})/(1 - \varepsilon_a)$$

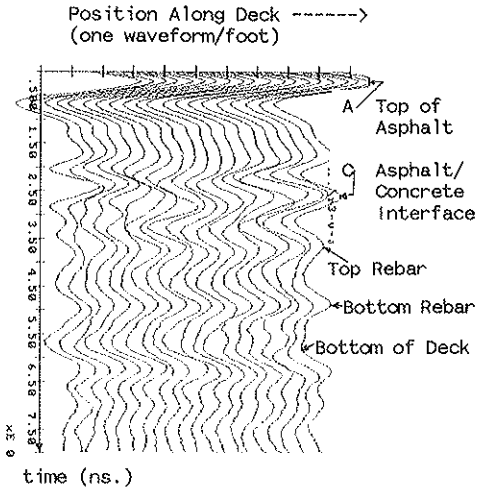



**FIGURE 4 Typical waveforms.**

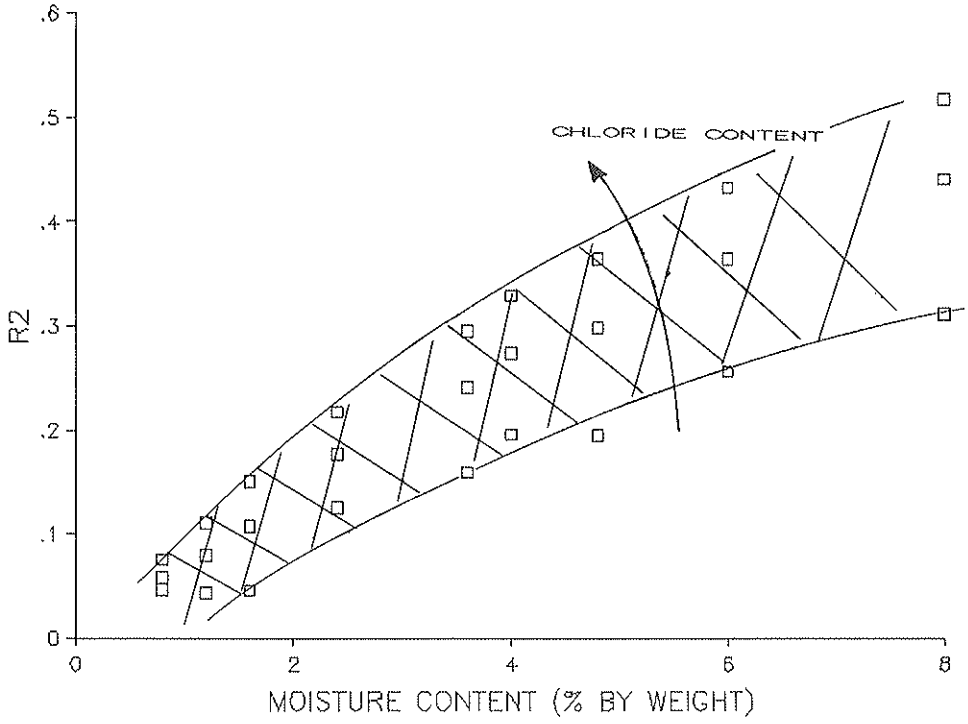



**FIGURE 5** Concrete reflectivity (R2) versus moisture content.

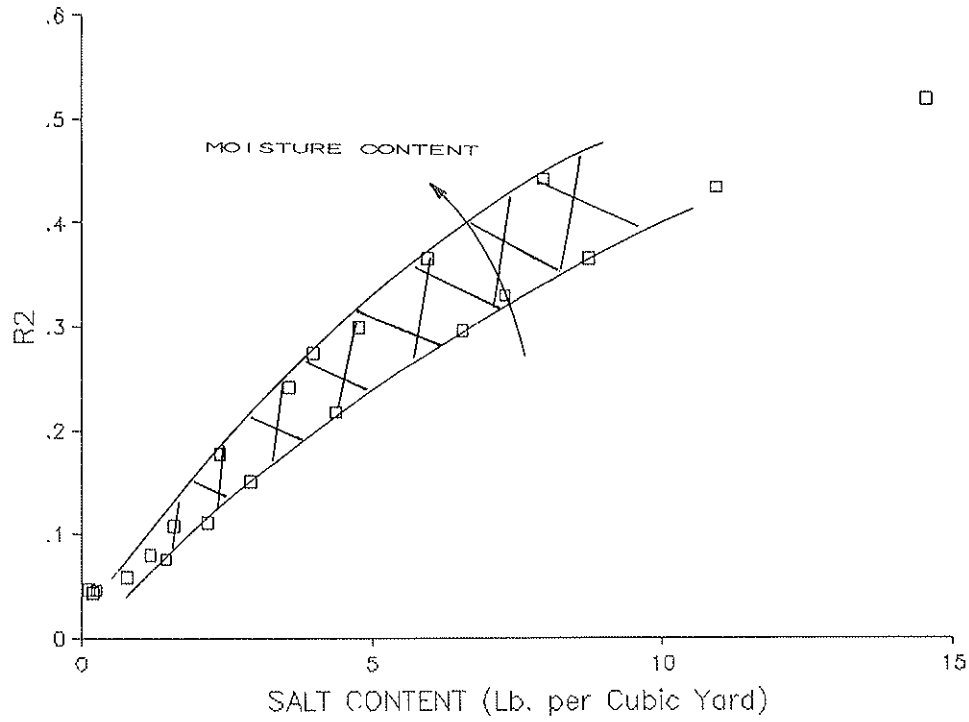



**FIGURE 6** Concrete reflectivity (R2) versus chloride content.

The computed value of $\varepsilon_c$ is used as an indicator of concrete deterioration. Normal concrete dielectric constant values range from 7.5 to 10.5, depending on air content, moisture content, and aggregate type. Deteriorated concrete has a higher value of dielectric constant. The quantity of deterioration is then inferred from the percentage of deck area exceeding a threshold value of dielectric constant.

Because the radar pulse has a width, the asphalt layer has to be sufficiently thick for the reflections from each layer to be clearly resolved. This minimum thickness can be calculated from the radar pulse width (in nanoseconds) and the radar velocity in the medium. For the horn antennas commonly in use for this application, this thickness is approximately 2 in. Ground-coupled antennas commonly used for geotechnical applications have transmit pulses that are two to three times longer because of ringing, and cannot resolve the concrete dielectric properties.

The analytical techniques described served as the basis for data analysis carried out during the NETC study.

## DESIGN AND CONDUCT OF THE TEST PROGRAM

### Identification of Deck Sites

Each participating state was asked to identify asphalt-overlaid bridge decks suitable for investigation during the research program. The selected decks were those that were scheduled for rehabilitation during the project period sometime after radar and infrared data could be collected. Deck rehabilitation involved asphalt removal, condition assessment of the concrete surface, and removal and repair of deteriorated concrete. Where possible, additional data describing deck condition were obtained either from previous inspections, or through planned testing during the rehabilitation process. These data included results of chloride content tests and corrosion potential measurements.

### Collection of Radar Data

Radar data were collected initially by the M.I.T. and subsequently by commercial vendors. Radar data were collected over the entire deck by carrying out a series of longitudinal passes spaced transversely at 1.5 to 2 ft. This method allowed for 100 percent coverage of the deck. For the two antenna systems shown in Figure 2, with the antenna booms separated by 6 ft, this method was implemented as follows:

1. Position the van in the center of the lane;
2. Swing the antenna booms so that the right-side antenna is 2 ft from the curb, and the left-side antenna is 8 ft from the curb;
3. Acquire radar data with both antennas while they are traveling longitudinally down the deck at about 5 mph;
4. Return to the beginning of the deck, and swing the antennas laterally so that they are now at 4 and 10 ft from the curb, respectively;
5. Repeat longitudinal radar data collection;
6. Repeat the preceding process until the entire lane is surveyed; and
7. Repeat the preceding steps on the next lane.

For the single-antenna system, it was necessary to conduct twice the number of longitudinal passes, repositioning the van laterally for each pass and using a similar procedure to that just described. Traffic control was provided by state personnel to allow for the conduct of these surveys on in-service decks. In some cases, the decks were already closed with Jersey barriers in preparation for the planned repairs, and no additional traffic control was required.

Radar data were observed on an oscilloscope during the survey. Radar data were collected as continuous analog radar waveforms representing the reflections at the radar pulses off of the various interfaces within the bridge deck. These data were acquired for each continuous radar pass and stored on magnetic tape, along with fifth-wheel data that could be used to locate the radar data on the deck surface for subsequent processing and interpretation. (The use of analog tape has since been replaced by directly acquiring the data digitally using a data acquisition system on a personal computer).

### Underside Surveys

Underside surveys were carried out both by state and M.I.T. personnel. General guidelines were established on what types of conditions to highlight in the underside survey. These included areas with rust stains, efflorescence, discoloration, and nonuniform dampness. Examples of these conditions are shown in Figure 7.

The underside survey was carried out by using the underside framing (girders and diaphragms) as a rectangular grid. With this grid, it was relatively easy to locate areas of interest within a particular rectangle. Surveys conducted both by state and M.I.T. personnel found a great deal of consistency in the location of areas of interest.

### Chloride and Corrosion Potential Measurements

Chloride content measurements were made by M.I.T. project personnel and combined with data previously collected by state personnel and consultants. The numbers of chloride samples per span generally ranged from 3 to 10. Corrosion potential data were obtained either as corrosion potential values at selected locations or as a complete corrosion potential survey on a 5-ft grid. All of the corrosion potential data collected during this project were obtained from measurements directly on the bare concrete.

### Chain-Drag and Material Removal Surveys

After the asphalt was removed, the concrete deck was normally chain-dragged to identify delaminated areas. The delaminated areas were marked with spray paint, as were areas in which concrete deterioration could be directly observed from the surface. Figure 6 shows a chain-drag survey in progress. Normally, the chain-drag survey was carried out by the

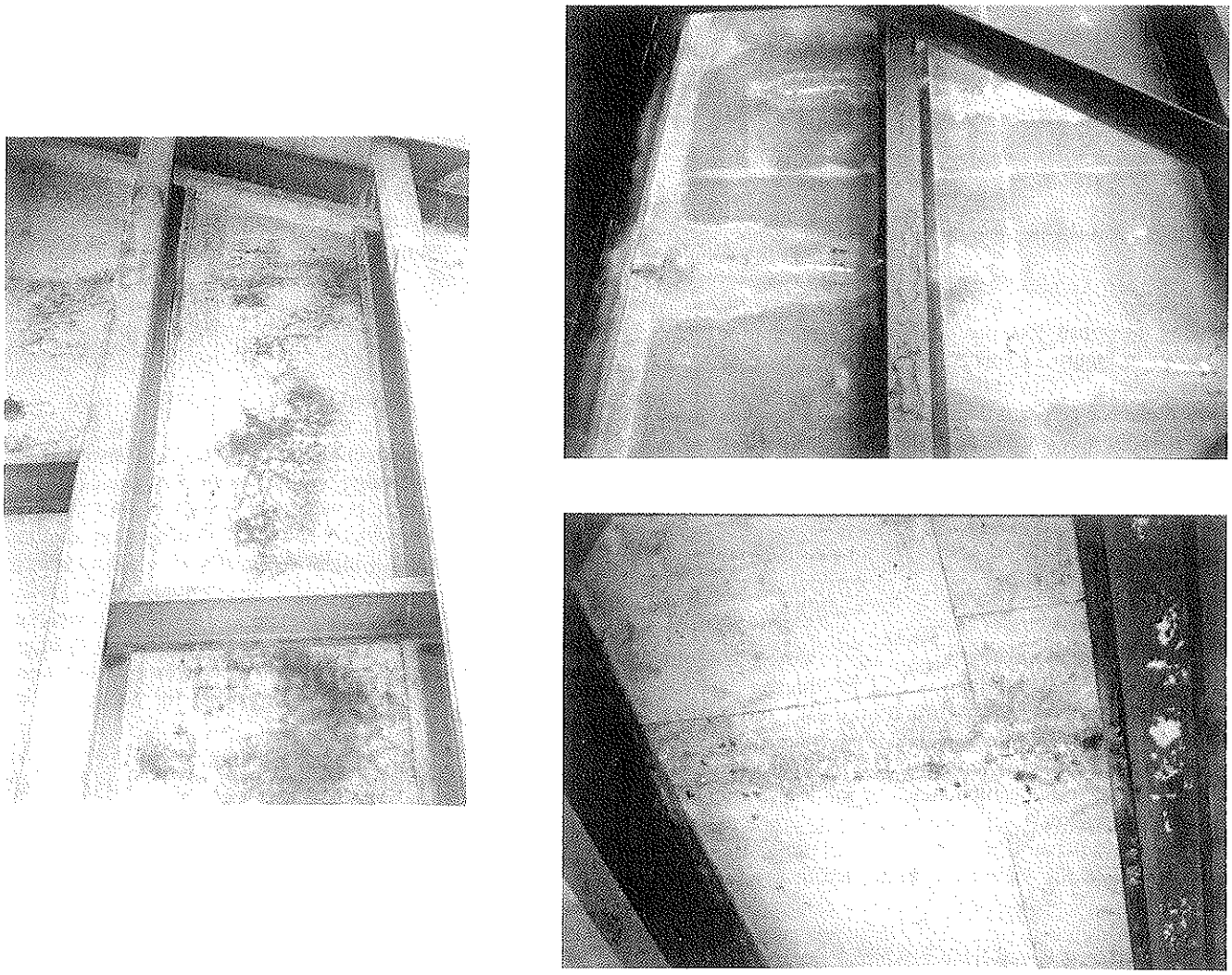


**FIGURE 7 Examples of conditions noted in underside surveys.**

resident engineer at the construction site, or by personnel from the state's materials or testing division. The results of this survey were then used by the contractor to identify areas where concrete was to be removed and replaced.

Maps of material removal were also obtained along with the chain-drag and surface deterioration surveys. The similarity between these maps depended on the contractor and on the state policy. In some cases, where the contractor followed exactly the outline of the chain-drag survey marks, they were identical. In other cases, the material removal was significantly greater than the marked deterioration because of squaring off and removal of additional concrete until corrosion was no longer seen.

## DATA ANALYSIS

### Data Entry

All of the data collected were directly transmitted to M.I.T. and entered into a computer data base for further analysis. The tape-recorded analog radar data were digitized using an analog-digital converter card installed in a personal computer. The radar equipment generated 50 waveforms per second. At the slow survey speeds used in this study, this ability resulted in many more waveforms than necessary for analysis. Therefore, the data were subsampled so that one waveform was digitized per longitudinal foot of travel. The digitized waveforms consisted of 500 points, representing 10 nsec of data.

The underside, chain-drag, and material removal data maps were digitized to scale using a digitizing tablet and AUTOCAD.

### Computation of Deterioration from Radar Data

The method for radar data analysis discussed earlier was applied to the data collected as described. Deterioration predictions were programmed to be carried out automatically using numerical computations on the digitized waveforms. Typical output of the computer analysis is shown in Figure 8. This figure shows a plot of concrete dielectric constant versus longitudinal distance along the bridge deck for a single radar pass. The dielectric constant values for multiple radar passes

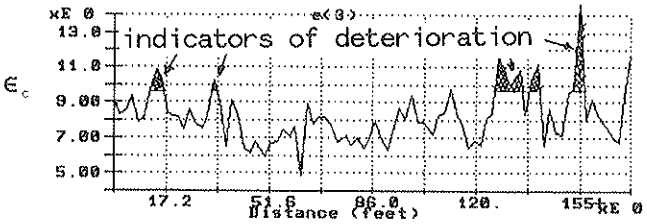



**FIGURE 8 Plot of computed dielectric constant versus distance.**

covering the entire deck can be displayed as a contour plot, as shown in Figure 9. This contour map can be used as an approximate indication of the locations of deck deterioration.

The dielectric constant data for all radar passes were used as a basis for the overall deck deterioration prediction. A number of alternative means for deterioration quantity estimations were investigated on the basis of these data, as described in the following section.

### Data Interpretation

The analysis method was investigated using a variety of regression techniques to determine the best fit with the data collected on the 28 bridge decks surveyed during this project. The regressions considered a number of options, including use of radar data alone, spatial overlays of radar and underside maps, and linear combinations of radar and underside data. The regression studies also considered variations in the radar threshold percentage.

The results of the regression studies led to the following relationship.

$$\text{Deterioration} = \text{K1} + \text{K2}(U) + \text{K3}(R_N) \quad (4)$$

where Kn are constants of regression, $U$ is the percentage of deck area noted in the underside survey, and $R_N$ is the percentage of deck area determined in Step 5 of the method using a threshold of $N$ percent above the mean. This equation fit 26 bridge decks with an $R^2$ value of 0.83 and a standard error of 4.07 percent of the total deck area. The results of this analysis are presented in Table 1 and Figure 9. The two decks that were not included in this regression were identified as having asphalt thicknesses significantly thinner than 2 in., and thus not suited to the analysis implemented in this study.

## APPLICATION OF RESULTS

The results have led to the introduction of a technique for assessing the deterioration in overlaid decks that is far more accurate than any other currently available method. On the basis of the results of this study, surveys are now being implemented with radar equipment operating at highway speed. Surveys are being conducted at various levels of detail. Project-level surveys seek 100 percent coverage of the deck area and include the results of an underside survey. These results provide sufficient detail for budgeting and scoping repair and rehabilitation projects. Network-level surveys seek 30 percent coverage of deck area, and do not necessarily include the underside survey. These results provide sufficient accuracy for network-level planning and for project priority setting.

Radar surveys can be carried out under all environmental conditions except for rain and subfreezing temperatures, and can be carried out day or night. Rain that produces standing water on the pavement surface distorts the dielectric constant computation and affects the deterioration prediction. Subfreezing temperatures alter the dielectric properties of the concrete in a manner that may obscure the detection of deterioration.

The following paragraphs describe current project- and network-level applications in further detail.

### Surveys at Highway Speed

The possibility of conducting surveys at highway speed arose because the radar equipment generated 50 waveforms per second, enough data to provide one waveform per foot at 35 mph. The question that has been raised in the past is whether or not the results of a highway-speed survey would match those of a slow-speed survey. A recent study of pavement layer evaluation using radar (*10*) clarified this issue by yielding identical results for surveys conducted at 5, 15, and 40 mph. The data analysis techniques used in that pavement study were identical to those presented in this paper for application to bridge decks.

The province of Alberta, Canada, has recently conducted a pilot test of overlaid deck surveys conducted at highway speed. The surveys were conducted by Infrasense, Inc., of Cambridge, Massachusetts. The survey vehicle made continuous round trips at normal driving speed, crossing the decks at different transverse positions until complete coverage was achieved. Two or more decks in the same vicinity were covered in the same round trip. The radar survey vehicle was followed by a chaser vehicle during the survey, and no traffic control or lane closures were required. Data acquisition began before and ended after each bridge was crossed. The bridge deck data were isolated from the pavement data subsequently during office processing.

The surveys carried out as described were provided at a cost of approximately \$0.10/ft$^2$ of deck area. The decks were distributed throughout the province, with 3 in Edmonton, 14 along Highway 2 near Red Deer, and 7 in Calgary. Costs for other surveys could be more or less, depending on the relative locations of the decks and the logistics of the survey.

### Network Level Surveys

The discussion has focused on developing a technique that provided a detailed condition assessment for decks that were being priority ordered and programmed for maintenance. There was also considerable interest among the New England states in developing a data base that included input of deck condition data for all decks in the state. Such a data base would represent an essential element of a bridge management system. The highway speed survey capability discussed could be exploited for this network-level application.

In this scenario, each lane of each surveyed deck would be covered by only two passes. In order to test out this network

TABLE 1 COMPARISON OF RADAR PREDICTIONS TO ACTUAL DETERIORATION (PERCENT OF DECK AREA)

| Decks | Radar Prediction | Actual Deterioration | Error |
|---|---|---|---|
| Vermont | | | |
| #1 | 10 | 11 | 1 |
| #2 | 12 | 9 | 3 |
| #3 | 14 | 11 | 3 |
| #4 | 21 | 12 | 9 |
| #5 | 6 | 7 | 1 |
| #6 | 6 | 7 | 1 |
| #7 | 6 | 4 | 2 |
| #8 | 6 | 12 | 6 |
| #9 | 15 | 12 | 3 |
| #10 | 5 | 3 | 2 |
| Maine | | | |
| #1 | 16 | 20 | 4 |
| #2 | 10 | 5 | 5 |
| #3 | 6 | 4 | 2 |
| #4 | 10 | 15 | 5 |
| #5 | 11 | 5 | 6 |
| Rhode Island | | | |
| #1 | 28 | 34 | 8 |
| #2 | 38 | 40 | 2 |
| #3 | 7 | 8 | 1 |
| #4 | 7 | 5 | 2 |
| #5 | 3 | 7 | 4 |
| #6 | 2 | 2 | 0 |
| #7 | 2 | 4 | 2 |
| #8 | 5 | 3 | 2 |
| New Hampshire | | | |
| #1 | 18 | 20 | 2 |
| #2 | 2 | 0 | 2 |
| Massachusetts | | | |
| #1 | 4 | 9 | 5 |
| | | AVERAGE ERROR | 3% |

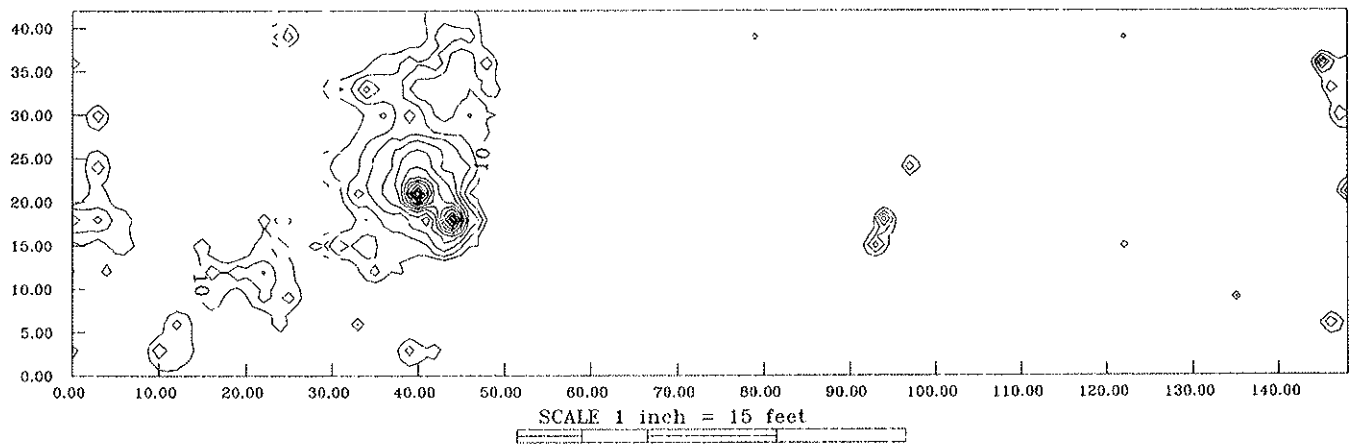



**FIGURE 9 Surface contour plot of concrete dielectric constant, Abbott Bridge, westbound.**

concept, the radar data collected during the NETC research program was analyzed as if it had been collected in the manner described for a high-speed network survey. For each lane of each surveyed deck, two passes, located approximately in the wheelpaths, were selected for analysis. The analysis used all the procedures described earlier, but treated the radar as the only source of information.

A linear regression model was developed for this data as follows:

$$\text{Deterioration} = K1 + K2(NR) \quad (5)$$

where NR represents the radar deterioration estimate from the two passes per lane. Equation 5 was fitted to the data from the same 26 decks discussed earlier, yielding $R^2 = 0.72$, $S_e = 5.0$, K1 = 0.61, and K2 = 2.04. This result indicated that with only two passes per lane, the radar results correlated reasonably well with observed deterioration. The fit was not as good as that for the more detailed survey, but the results were accurate enough for network planning and priority setting.

On the basis of this these results, the New Hampshire Department of Transportation conducted a pilot program to evaluate the network survey concept. In a network-level survey, one envisions a radar van traveling continuously along the highway logging data for every bridge deck that it crosses. For a given round trip, the van would make one pass on each lane. With such a procedure, 20 to 40 decks could be surveyed in a day, a rate that would allow for complete coverage of a typical state bridge inventory in 50 to 300 working days. In addition, the automated nature of the radar processing would allow relatively efficient analysis of this quantity of data.

Results from the New Hampshire project have demonstrated the feasibility of conducting the network surveys. In two pilot network surveys, 24 decks were covered in 1 day and 9 in another day. Decks that were scheduled for repair were investigated to correlate radar predictions with the deterioration directly observed after asphalt removal. The radar predictions were consistent with these observations.

## CONCLUSIONS

This work was motivated by a need to obtain more accurate information regarding the condition of bridge decks. The information was needed: to plan and budget overall maintenance, repair, and rehabilitation programs; to order maintenance and repair projects by priority; to select the optimum maintenance and rehabilitation approach, and to make repair and replace decisions; and to properly scope and budget maintenance and repair projects. Some of these needs require detailed estimates of amounts of deck deterioration, whereas others require reasonable estimates with less precision.

On the basis of the research described herein, a new method has been developed that accurately fits deterioration of overlaid bridge decks using processed radar data regressed against known deterioration from 26 decks. An additional outcome of this work has been the development of a high-speed deck survey concept applicable both to detailed project- and to network-level surveys. The detailed survey concept involves 100 percent radar coverage of the deck, and incorporates the results of an underside survey. The network concept involves one radar pass per wheelpath (30 percent coverage). Results for such a survey method have been simulated from the detailed data collected during this program. They have been shown to compare reasonably well with observed deterioration, but with less detail than in the detailed survey method.

## ACKNOWLEDGMENT

The research described in this paper was supported by the New England Transportation Consortium through a contract between AASHTO and M.I.T. The author would like to acknowledge the contribution of the following members of the NETC Bridge Deck Project Technical Committee for their valuable contributions throughout this work: Everett Barnard, Warren Tripp, Richard Swanson, Richard Kalunian, Gaylon Finnemore, Richard Todd, Alan Rawson, and Lou De Franzo. The author would also like to thank Ron Frascoia, Don Hamilton, and Mona Manachi for assistance in providing field data on the decks evaluated during this project.

## REFERENCES

1. *NCHRP Synthesis 57: Durability of Concrete Bridge Decks*. TRB, National Research Council, Washington, D.C., 1979.
2. K. R. Maser. *New Technology for Bridge Deck Assessment*. Phases I and II, Final Report. New England Transportation Consortium, Center for Transportation Studies; Massachusetts Institute of Technology, Cambridge, 1989.
3. K. R. Maser and W. M. K. Roddis. Principles of Radar and Thermography for Bridge Deck Assessment. *ASCE Journal of Transportation Engineering*. Vol. 116, No. 5, Sept./Oct., 1990.
4. U. B. Halabe, K. Maser, and E. Kausel. *Condition Assessment of Reinforced Concrete Using Electromagnetic Waves*. Final Technical Report. U.S. Army Research Office; Center for Construction Research and Education, Massachusetts Institute of Technology, Cambridge, 1989.
5. C. R. Carter, T. Chung, F. B. Holt, and D. G. Manning. Automated Signal Processing System for the Signature Analysis of Radar Waveforms from Bridge Decks. *Canadian Electrical Engineering Journal*. Vol. II, No. 3, 1986.
6. D. Manning and F. B. Holt. Detecting Deterioration in Asphalt Overlaid Bridge Decks. In *Transportation Research Record 899*, TRB, National Research Council, Washington, D.C., 1984.
7. C. P. F. Ulricksen. Application of Impulse Radar to Civil Engineering. Ph.D. dissertation, Department of Engineering Geology, Lund University, Lund, Sweden, 1982.
8. A. V. Alongi, T. R. Cantor, C. P. Kneeter and A. Alongi, Jr. Concrete Evaluation by Radar a Theoretical Analysis. In Transportation Research Record 853, TRB, National Research Council, Washington, D.C., 1982.
9. G. Clemeña. Non-Destructive Inspection of Overlaid Bridge Decks with Ground Penetrating Radar. In *Transportation Research Record 899*, TRB, National Research Council, Washington, D.C., 1984.
10. K. R. Maser and T. Scullion. Automated Pavement Subsurface Profiling Using Radar—Case Studies of Four Experimental Field Sites. Presented at the 70th Annual Meeting of the *Transportation Research Board*, Washington, D.C., Jan. 1991.

*Publication of this paper sponsored by Committee on Structures Maintenance.*

# Preliminary Tests of a New Surface Airflow Device for Rapid In Situ Indication of Concrete Permeability

D. WHITING

The development of a prototype surface airflow device for concrete is described. The method is based on measurement of the rate of airflow through a vacuum plate placed on a concrete surface under a vacuum of approximately 25 in. of Hg. Effective depth of measurement is demonstrated to be approximately 0.5 in. below the surface. For calibration of the method, concretes were cast using a variety of water-to-cement ratios as well as admixtures such as latex and silica fume. Results were found to correlate well with chloride diffusion constants derived from 90-day ponding tests, as well as with true air permeabilities measured using a pulse decay technique. Design of a field prototype based on the surface airflow technique is also presented. The field instrument is designed to obtain readings at the rate of approximately one per minute, and be powered by rechargeable battery packs. Although surface moisture has an influence on test results, simple techniques for rapid drying of small test areas have been developed. Plans for future testing are presented.

There is a growing awareness of the important role of permeability with regard to the long-term durability of concrete structures. Much of this interest relates to the realization that durability may be directly related to permeability of the concrete in question. If an aggressive substance, be it water, sulfate or chloride ions, or other materials, can be kept out of concrete by virtue of low permeability, then associated problems, such as freeze-thaw deterioration, corrosion of reinforcement, and formation of expansive components, may be mitigated. Therefore, there has been an interest not only in determining permeabilities of conventional concretes but also in developing improved concretes having low permeabilities.

The need for data on concrete permeability dates from the early 1930s, when designers of large hydroelectric structures required information on rates of passage of water through concrete under the influence of relatively high hydraulic heads. In such instances, flow of water through concrete can be adequately described by Darcy's law (*1*),

$$Q = \frac{KA}{\mu} \cdot \frac{dp}{ds} \tag{1}$$

where

$Q$ = volume outflow (cm³/sec),
$A$ = area (cm²),
$\mu$ = viscosity (centipoise),
$dp/ds$ = pressure gradient (atmospheres per centimeter), and
$K$ = permeability constant (Darcy).

Permeabilities to such fluids as air (*2*), liquid nitrogen (*3*), methane (*4*) and oil (*5*) have also been measured. In general, permeabilities to gases are from one to two orders of magnitude greater than those for water.

Under conditions other than those of saturated fluid flow, transport of substances through concrete can occur by a variety of different mechanisms. These may include (a) capillary attraction, (b) vapor transmission, or (c) ionic diffusion. Data developed by Wing (*6*) and Dunagan (*7*) indicate that movement of water into dry concrete by capillary attraction can be rapid. An initial surface absorption test (ISAT) (*8*) has been developed using these principles.

There is currently much concern with corrosion of reinforcing steel promoted by chloride ions that penetrate through the concrete cover and eventually reach the reinforcement. It is generally believed that such migration of chloride ions occurs primarily through diffusion processes. Various studies (*9,10*) indicate that the coefficient of diffusion is of the order of $10^{-7}$ to $10^{-8}$ cm²/sec. Diffusion will generally follow Fick's second law,

$$\frac{\partial C}{\partial t} = D \cdot \frac{\partial^2 C}{\partial x^2} \tag{2}$$

where

$C$ = concentration at distance $x$ (cm) from a boundary,
$t$ = time (sec), and
$D$ = effective diffusion coefficient (cm²/sec).

Improved laboratory techniques that can rapidly assess the permeability of concrete to water, gases, ions, or other substances have been developed. These include techniques based on high pressures (*11*), transient gas flow (*12*), or ionic conductance (*13*). Although these methods offer an improvement in testing speed over previous steady-state techniques, these still require somewhat time-consuming specimen preparation, and the necessity to remove a core from the structure makes testing a destructive process.

Recent emphasis has been on nondestructive field tests that can rapidly assess the permeability of in-place concrete. A review carried out for the Strategic Highway Research Program in 1988 (Whiting, unpublished report) indicated that

Materials Research and Consulting, Construction Technology Laboratories, Inc., 5420 Old Orchard Road, Skokie, Ill. 60077.

about 20 such tests had been developed, (some only slight variants of others). The most widely used tests included the ISAT (*8*) and the vacuum decay and capillary flow tests devised by Figg (*14*). Most of these tests involved either drilling a small hole into the concrete in which the measurement is carried out, or affixing a test apparatus to the surface of the concrete using clamps or a catalyzed resin seal. An exception was the vacuum decline test developed by Schonlin and Hilsdorf (*15*), which used a vacuum seal on the concrete surface.

In spite of existing tests, a rapid, practical permeability test was needed that provided large numbers of readings on site for given concrete structures in reasonable periods of time. The development of such a test is described in the following sections.

## OBJECTIVES AND SCOPE

The objective of the project was to develop a rapid field test for concrete permeability that would correlate reasonably well with existing laboratory techniques. It was recognized that any rapid, nondestructive field test would most likely not yield a true permeability (or diffusion) value but would afford an empirical result that could be taken as a *measure* of relative permeability. Such a device would, ideally, be rugged, portable, battery powered, and simple to operate. The intent was to develop a laboratory prototype first, ensure that results correlated with laboratory techniques, and then construct a workable field instrument. The influences of external variables (such as moisture content, temperature, and surface finish) were also to be investigated.

## DEVELOPMENT OF THE METHOD

### Basis of the Technique

The surface air flow (SAF) technique is based on the rate of air flow through a concrete surface, more permeable concretes being characterized by higher air flow rates. Rather than force air into the concrete, a procedure that would require a direct pressure seal to be created at the surface, this technique relies on creation of a vacuum at the surface, which causes air to flow out of the concrete along the pressure gradient created by a vacuum pump. The SAF technique is similar in principle to that developed by Schonlin (*15*); however, Schonlin used decay of vacuum as a measure of permeability, and therefore test times were lengthy for low-permeability concretes. The SAF technique affords a more rapid measurement once the desired pressure differential has been achieved.

The SAF technique is designed to give an indication of permeability for a wide variety of concrete materials. These include not only low-permeability concretes, such as those produced using silica fume or polymer modifiers, but also high-permeability concretes created by use of high water-to-cement ratios or poor curing practices. The effective test area is chosen as a circle 2¾ in. in diameter, offering a representative test on a small area of most structural concretes. As the test was designed to be rapid, a practical goal was to be able to obtain readings at 4-ft centers over one lane of a typical bridge deck within a normal working day. The test was to be applicable to support and substructural elements as well, necessitating that the device also be operable in vertical and overhead nodes.

### Laboratory Bench Prototype

A schematic diagram of the laboratory bench prototype is shown in Figure 1. The steel vacuum plate of 4-in. diameter is fitted with a ⅛-in.-thick soft rubber ring, allowing for a circular flow area of 2¾ in. in diameter. Flow is initially directed through Valve A until the system reaches the operating pressure of 120 to 140 mm of Hg as measured on the U-tube Hg manometer. Valve A is then closed and the air stream directed through Valve B from the rotameter, which is calibrated to read in units of milliliters per minute over a range of 0 to 280 ml/min. The total test time is approximately 1 min, after which the pump is turned off and air is allowed to reenter the system through Valve C.

### Preparation of Test Specimens

After initial system checkouts were complete, a series of concrete test specimens offering a range of permeabilities was designed and constructed. Concrete mixtures are presented in Table 1. The forms were designed to contain (a) rebar mats at ½ and 1½ in. to evaluate effects of reinforcing steel on permeability techniques, (b) brass tubes (5/32-in. diameter) at ¼, ½, and 1 in. from the surface to be used to measure extent of depth of influence of vacuum, and (c) thermocouples at ¼, ½, 1, and 2 in. for temperature studies.

Slabs from Mixtures A through D were cured under wet burlap and polyethylene overnight, then transferred to heavy plastic bags for 28 days of curing. They were then coated on four sides with epoxy and placed in the following environments:

- Continuous air drying at 73 ± 3°F and 50 ± 5 percent relative humidity (air-dry series), and
- Cycling of 1 day wet at 75°F to 80°F and 6 days dry in the same environment (moist series).

Mix E was moist-cured for 28 days to develop a high level of impermeability, then exposed to air-drying. Mix F was exposed to air-drying immediately after demolding, to develop maximum curing of the latex admixture.

## PRELIMINARY TESTING

### Air-Dried Test Series

The first series of tests was performed on those slabs that had been placed in an air-drying environment for approximately 3 months. A series of 15 individual flow readings was obtained across the cast and finished faces of each slab.

A summary of test statistics is presented in Table 2. Lowest readings were obtained for the silica fume concrete (E), highest for the concrete with $w/c = 0.6$ (D). Readings for latex concrete (F) were somewhat higher than for silica fume, as would be expected. Conventional concretes exhibited an in-

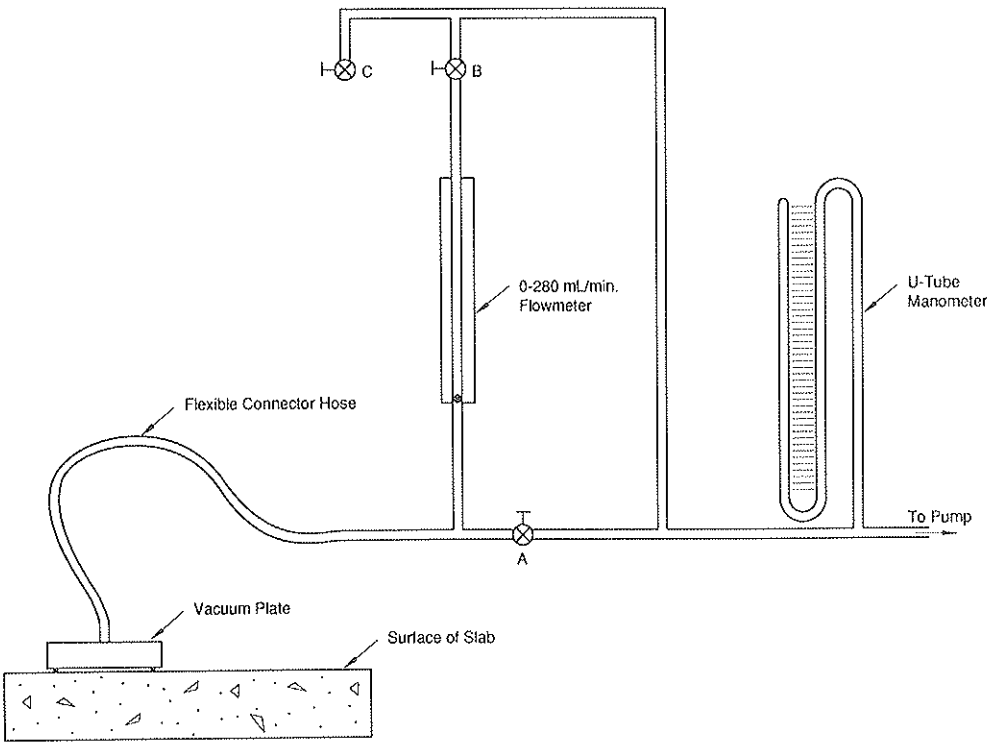



**FIGURE 1 Schematic drawing of the laboratory surface airflow device.**

creasing air flow with increase in *w/c*, at least on the finished face. There was little difference in air flow between concretes A, B, and C on the cast face, perhaps representing an increased densification at the bottom of the slabs, or some absorption of the mix water by the formwork. On the finished face of Slab C–1, it was not possible to achieve satisfactory readings. Most readings were off scale (i.e., >280 ml/min), and those that fell on scale were unstable. Because this effect might have been caused by leakage around the gasket, a silicone vacuum grease was applied to the gasket and some readings were repeated. This procedure stabilized some readings but most were still off scale. On inspection of the surface that had been coated with grease, a fine network of surface microcracks in the thin surface paste layer was visible. It is likely that this microcracking led to leakage under the gasket and high measured air flows. This surface was then given a light sandblast to remove the surface paste. However, repeat testing still indicated microcracks, which were now even more visible on the newly created surface. This slab may have been prematurely finished or may have suffered from some plastic shrinkage cracking.

The statistics presented in Table 2 were obtained from readings taken at 15 different locations on each slab. Therefore, the variance estimates represent a combination of variability in concrete properties from point to point on the surface as well as inherent test variability. Variability is greater for the finished surfaces, perhaps reflecting differences across the slab caused by hand finishing. In order to obtain a better

TABLE 1 CONCRETE TEST MIXTURES

| | Quantities - lb/cu yd | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mix | Cement | Sand | Gravel | Water | Water/ Cement Ratio | Admixtures[a/] | Slump (in.) | Air (%) |
| A | 659 | 1263 | 1772 | 252 | 0.38 | -- | 2.6 | 5.0 |
| B | 550 | 1375 | 1776 | 237 | 0.43 | -- | 3.2 | 5.3 |
| C | 453 | 1477 | 1760 | 227 | 0.50 | -- | 3.0 | 5.8 |
| D | 374 | 1563 | 1720 | 224 | 0.60 | -- | 3.8 | 6.4 |
| E | 665 | 1300 | 1680 | 239 | 0.36 | Silica Fume 67 lb. | 5.0 | 5.0 |
| F | 685 | 1579 | 1311 | 185 | 0.27 | Latex 25 gal. | 6.9 | 5.6 |

a/ Other than air-entraining agents.

TABLE 2 SURFACE AIRFLOW MEASURED ON AIR-DRIED SLABS

| Slab | Readings | Face | Air Flow (mL/min) Mean | Minimum | Maximum | Std. Dev. | % C.V. |
|---|---|---|---|---|---|---|---|
| A-1 | 15 | Cast | 67 | 55 | 85 | 8.38 | 12.5 |
| A-1 | 15 | Finished | 50 | 40 | 65 | 6.97 | 13.7 |
| B-1 | 15 | Cast | 68 | 55 | 90 | 8.80 | 12.9 |
| B-1 | 14 | Finished | 87 | 60 | 155 | 26.29 | 30.2 |
| C-1 | 15 | Cast | 64 | 50 | 95 | 11.01 | 17.3 |
| C-1 | a/ | Finished | -- | -- | -- | -- | -- |
| D-1 | 15 | Cast | 152 | 125 | 175 | 13.66 | 8.9 |
| D-1 | 15 | Finished | 128 | 85 | 175 | 29.56 | 23.1 |
| E-1 | 15 | Cast | 19 | 14 | 22 | 2.23 | 11.7 |
| E-1 | 10 | Finished | 16 | 12 | 23 | 3.14 | 19.4 |
| F-1 | 15 | Cast | 22 | 17 | 30 | 3.54 | 15.8 |
| F-1 | 10 | Finished | 34 | 22 | 65 | 12.54 | 36.9 |

a/not tested due to surface crazing.

estimate of test reproducibility, a series of measurements was performed consisting of five replicates on each of three positions on cast and finished surfaces of each slab. Summary statistics are presented in Table 3 for the cast faces. The variability of successive measurements at one position is much less than the variability across the slab. This difference indicates that the method was actually detecting differences in permeability across the surface of the slab, most likely caused by small differences in consolidation and finishing across the surface. Reproducibility statistics are not presented for Slab F–1 (latex concrete) because each reading appeared to be successively lower, indicating that air was being evacuated faster than it was being replaced in the time interval between each measurement.

The slabs cast with embedded brass tubes were used to measure the effective depth of penetration of the vacuum into the slabs. During casting, a solid brass rod was extended through the tube. When it was removed, a 1-in.-long cavity was created at the ½ and 1-in. depths. A microflowmeter and a manometer were attached to each well to measure air flow and vacuum level at each depth (Table 4). Results indicate that for low-*w/c* concrete (Slab A–2) only small flow rates could be measured, even at ¼ in. below the surface. For slightly higher *w/c* ratio (Slab B–2), flow at ¼ in. below the surface was approximately 20 percent of the surface reading. For concretes with high *w/c* ratio, flow was much greater at ¼ and ½ in., and was even detectable at 1 in. below the surface. Effects of steel placement and cover on air-flow permeability readings were also investigated. Readings were taken directly above and between No. 5 bars having ½ and 1½ in. of cover. Results are presented in Table 5. There is little significant difference between readings taken above or between reinforcing bars, indicating that reinforcing steel (and presumably other embedments) has little effect on the test, provided cover depth is ½ in. or greater.

### Effects of Moisture Content

Testing was performed on slabs that had been subjected to a weekly wetting cycle (1 day soak and 6 days air-dry) since an age of 28 days. The top (finished) surface of each slab was subjected to the soaking. The bottom (cast surface) was protected from direct wetting during the soak but was in a damp environment. Tests were performed on both cast and finished surfaces approximately 1 hr after removal of soaking blankets.

Test results are presented in Table 6. For the finished surfaces, lower readings were encountered for the wetted slabs, the differences between mean dry and wet readings ranging from about 15 units for Batch A to 30 units for Batch D.

TABLE 3 REPRODUCIBILITY MEASUREMENTS

| Slab | Face | Position | Air Flow (mL/min) Mean | St. Dev. | % C.V. |
|---|---|---|---|---|---|
| A-1 | Cast | 1 | 52 | 2.09 | 4.0 |
| | | 6 | 75 | 2.74 | 3.6 |
| | | 15 | 54 | 2.19 | 4.1 |
| B-1 | Cast | 1 | 58 | 2.50 | 4.3 |
| | | 5 | 68 | 2.73 | 4.0 |
| | | 11 | 50 | 2.74 | 5.4 |
| C-1 | Cast | 1 | 50 | 3.06 | 6.1 |
| | | 7 | 53 | 2.33 | 6.1 |
| | | 15 | 89 | 2.24 | 2.5 |
| D-1 | Cast | 6 | 121 | 6.52 | 5.4 |
| | | 7 | 153 | 2.50 | 1.6 |
| | | 11 | 161 | 4.54 | 2.8 |
| E-1 | Cast | 1 | 16 | 1.79 | 11.0 |
| | | 6 | 21 | 1.12 | 5.5 |
| | | 11 | 23 | 1.04 | 4.5 |

TABLE 4 EFFECTIVE DEPTH OF VACUUM PENETRATION

| Slab | Well Depth (inches) | Air Flow (mL/min) | Vacuum |
|---|---|---|---|
| A-2 | Surface | 40 | -635 mm |
| | 1/4 | 1.0 | nd |
| | 1/2 | 0.5 | nd |
| | 1 | 0 | nd |
| B-2 | Surface | 65 | -635 mm |
| | 1/4 | 12.6 | -180 mm |
| | 1/2 | 2.1 | nd |
| | 1 | 0 | nd |
| D-2 | Surface | 130 | -635 mm |
| | 1/4 | 100 | -600 mm |
| | 1/2 | 13.5 | -160 mm |
| | 1 | 0.4 | nd |

nd/ no vacuum detected

TABLE 6 COMPARISON OF SURFACE AIRFLOW TESTS ON DRY AND WET SLABS

| Batch | Slab | Face | Condition[a/] | Air Flow (ml./min) Mean | Minimum | Maximum |
|---|---|---|---|---|---|---|
| A | A-1 | Cast | Dry | 67 | 55 | 85 |
| | | Finished | Dry | 50 | 40 | 65 |
| | A-4 | Cast | Damp | 60 | 45 | 70 |
| | | Finished | Wet | 45 | 25 | 75 |
| B | B-1 | Cast | Dry | 68 | 55 | 90 |
| | | Finished | Dry | 87 | 60 | 155 |
| | B-4 | Cast | Damp | 65 | 48 | 95 |
| | | Finished | Wet | 60 | 40 | 80 |
| C | C-1 | Cast | Dry | 64 | 50 | 95 |
| | C-4 | Cast | Damp | 71 | 55 | 115 |
| D | D-1 | Cast | Dry | 152 | 125 | 175 |
| | | Finished | Dry | 128 | 85 | 175 |
| | D-4 | Cast | Damp | 154 | 85 | 215 |
| | | Finished | Wet | 96 | 65 | 155 |
| E | E-1 | Cast | Dry | 19 | 14 | 22 |
| | E-2 | Cast | Damp | 22 | 15 | 27 |
| F | F-1 | Cast | Dry | 22 | 17 | 30 |
| | F-2 | Cast | Damp | 18 | 9 | 32 |

a/ Slabs surface-dry at start of test.

Differences for cast surfaces were less, as the cast surfaces were not directly exposed to liquid water. In order to examine the effects of wetting on cast surfaces, the cast surfaces were placed upright, put through three weekly cycles of wetting, then retested. As with the finished surfaces, there were considerable differences caused by wetting of the test surfaces. These differences were greatest for the high-*w/c* concrete (D–4), and were insignificant for low-permeability concretes (E–2 and F–2). A summary of airflow measurements on the various sets of test slabs under differing conditions of moisture is shown in Figure 2. In spite of the differences induced by varying surface conditions and moisture levels, a general trend of increasing airflow with decreasing quality of concrete is noted. The concretes designed to have low permeabilities (Batches E and F, silica fume and latex concretes) have the lowest airflows. Conventional concretes typical of those specified for structural and paving applications (Batches A, B, and C) reveal a slight upward trend for airflow as *w/c* ratio is increased going from Batches A to C. The poorest-quality concrete produced in this test series (Batch D), exhibits the highest airflow readings.

A preliminary classification of permeability based on airflow measurements is suggested by Figure 2. Readings less than 30 ml/min are generally associated with low-permeability concretes. Readings between 30 and 80 ml/min are associated with moderately permeable concrete (*w/c* from about 0.4 to 0.5). Readings above 80 ml/min are associated with high-permeability concretes. Readings in borderline regions may be ambiguous; for instance, very wet, high-*w/c* concretes may on occasion yield readings close to the moderate zone.

In an attempt to devise a practical means of reducing the effect of near surface moisture content on test results, various drying stategies were carried out. A 12- × 12- × 4-in. louvered sheet metal box was mounted on a slab so that an embedded thermocouple tree (of ¼-, ½-, 1-, and 2-in. thermocouple depths) was at the approximate center of the area. A heat gun of 1,500-watt capacity was then placed in a tube mounted on the top center of the box and activated. Temperature profiles for Slab D–2 (*w/c* = 0.6) are shown in Figure 3. Although surface temperatures reach nearly 300°F, temperatures at ¼, ½, and 1 in. peak at approximately 140°F, 110°F, and 95°F after 5 min of heating. The slab was then allowed to cool for an additional 10 min until all temperatures fell below 90°F. At this time, airflow tests were repeated. Results presented in Table 7 indicate that the surface heating was able to restore flow values close to those obtained on initially dry surfaces. This approach, at least in principle, could offset the interference that moisture appears to have on the test. The 1,500-watt heat gun requires a high energy expenditure, and would necessitate that a relatively large generator be transported to the test site. It was decided to investigate the use of a small hand-held infrared heater fired

TABLE 5 EFFECTS OF STEEL PLACEMENT AND CLEAR COVER

| Slab | 1/2 Inch Cover above bars | 1/2 Inch Cover between bars | 1-1/2 Inch Cover above bars | 1-1/2 Inch Cover between bars |
|---|---|---|---|---|
| A-2 | 50 | 50 | 40 | 40 |
| B-2 | ns | ns | 60 | 65 |
| D-2 | 155 | 145 | 135 | 130 |

ns/ no stable readings obtainable (crack in concrete above bar).

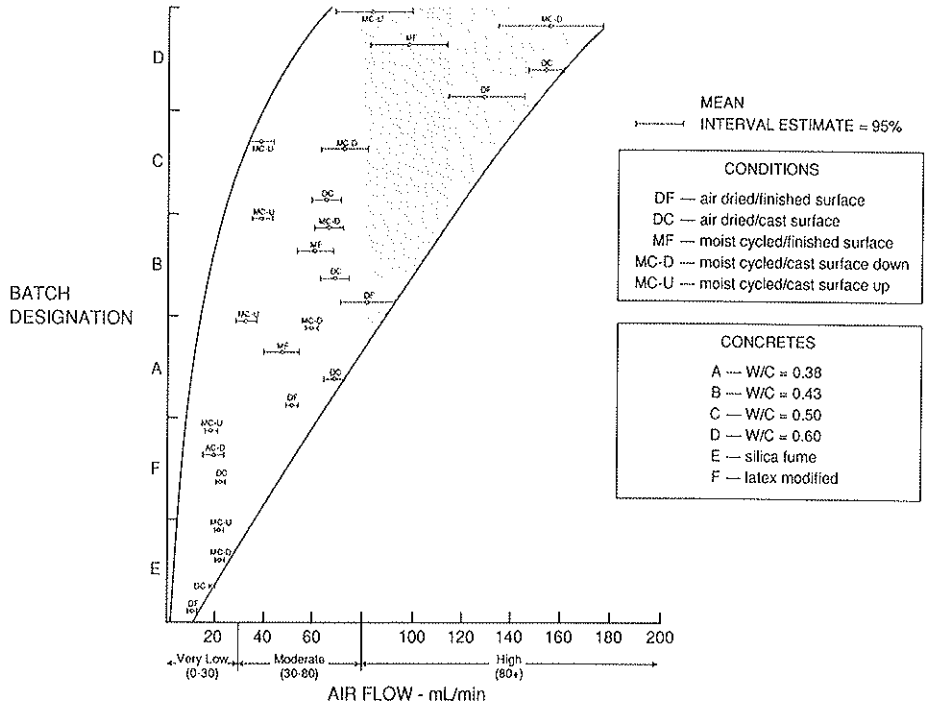



**FIGURE 2 Summary of surface airflow measurements.**

by liquid propane (such as is used for stripping paint). The heater was held at a height of 4 in. above the desired test areas for periods of 5 min. Heating profiles are shown in Figure 3. Results compared favorably with those for the 1,500-watt heat gun previously tested. After cooling, SAF measurements were made on the heated areas. Results presented in Table 8 indicate that the infrared method is capable of restoring the test area essentially to the initial air-dry condition.

Examination of Figure 3 indicates that essentially the same temperatures at any given depth are reached after 3 min using the infrared heater, as opposed to 5 min using the heat gun. In order to investigate the possibility of using shorter heating times, the slabs were resaturated and the tests carried out using 3 min of heating with the infrared heater. Results presented in Table 9 were again favorable, indicating that a 3-min heating time could be used. Because further tests at

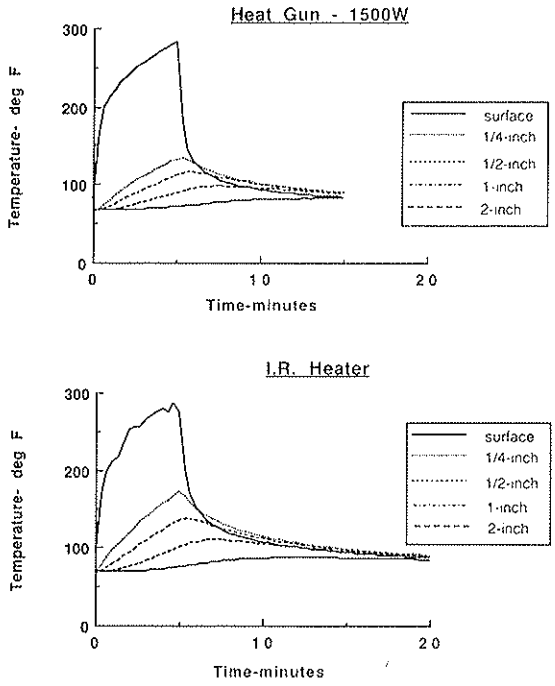



**FIGURE 3 Comparison of temperature profiles generated using infrared heater and 1,500-watt heat gun.**

TABLE 7 EFFECTS OF SURFACE DRYING USING 1,500-WATT HEAT GUN

| Slab | w/c | Initial Flow Air Dry [a] (mL/min) | Flow After Soaking [b] (mL/min) | Flow After Heat Drying [c] (mL/min) |
|---|---|---|---|---|
| A-2 | 0.38 | 45/50 | 30/35 | 40/45 |
| B-2 | 0.43 | 60/70 | 45/50 | 65/70 |
| D-2 | 0.60 | 130/140 | 95/105 | 125/135 |

a/ Stored in air at 73±3°F and 50±5% RH for 10 months after curing.
b/ Soaked once weekly for 7 weeks, then continuously for 3 days.
c/ Dried 5 minutes using 1500 W heat gun at 4 inches above surface.

TABLE 8 EFFECTS OF SURFACE DRYING USING INFRARED HEATER—5-min HEATING

| Slab | w/c | Initial Flow Air Dry [a] (mL/min) | Flow After Soaking [b] (mL/min) | Flow After Heat Drying [c] (mL/min) |
|---|---|---|---|---|
| A-2 | 0.38 | 45/50 | 30/35 | 40/50 |
| B-2 | 0.43 | 60/70 | 45/50 | 65/70 |
| D-2 | 0.60 | 130/140 | 110/120 | 130/140 |

a/ Stored in air at 73±3°F and 50±5% RH for 10 months after curing.
b/ Soaked once weekly for 15 weeks, then continuously for 3 days.
c/ Dried 5 minutes using I.R. heater at 4 inches above surface.

TABLE 9 EFFECTS OF SURFACE DRYING USING INFRARED HEATER—3-min HEATING

| Slab | w/c | Initial Flow Air Dry a/ (mL/min) | Flow After Soaking b/ (mL/min) | Flow After Heat Drying c/ (mL/min) |
|---|---|---|---|---|
| A-2 | 0.38 | 45/50 | 30/35 | 45/50 |
| B-2 | 0.43 | 60/70 | 45/50 | 65/75 |
| D-2 | 0.60 | 130/140 | 70/80 | 120/130 |

a/ Stored in air at 73±3°F and 50±5% RH for 10 months after curing.
b/ Soaked once weekly for 16 weeks, then continuously for 3 days.
c/ Dried 5 minutes using I.R. heater at 4 inches above surface.

shorter heating times were not as promising, a 3-min heating time using the infrared heater at 4 in. above the test area was adopted.

### Correlation Studies

In order to develop information on the relation between the SAF method and the more conventional techniques, a series of correlation studies was carried out. A set of comparison air-dry slabs was subjected to surface ponding on the finished surfaces with 15 percent sodium chloride for a period of 90 days. After exposure was complete, powder samples were obtained from the slabs at ¼-in. increments from the surface using procedures described in AASHTO T260. Using the procedure developed by Weyers (*16*), diffusion coefficients for chloride ions were obtained from the data. In addition, a series of cores of 4-in. diameter was obtained from selected air-dried slabs used for the initial SAF readings. The top 2 in. of each core was removed and used for the test specimens. The slices were dried to constant weight at 140°F, then tested for air permeability using a pulse-decay technique (*12*). Relationships between SAF readings, chloride diffusion coefficients, and air permeability are shown in Figure 4. The good correlations demonstrate that SAF can be used as a measure of true diffusion coefficients or permeabilities of the surface layer of concrete.

The relations shown in Figure 4 are not meant to imply that SAF testing can be used as a substitute for actual permeability measurements. If actual permeability values are desired, then the corresponding physical tests must be carried out.

## DESIGN AND CONSTRUCTION OF A FIELD DEVICE

### Design Criteria

In order to obtain a workable field technique, it was necessary to package the components into a fieldworthy prototype instrument. This field prototype required certain characteristics that would not necessarily have been included in the laboratory version of the device. Those characteristics were as follows:

- Ruggedness. It had to be water-resistant, resistant to accidental impact, and capable of transport in typical construction field vehicles.

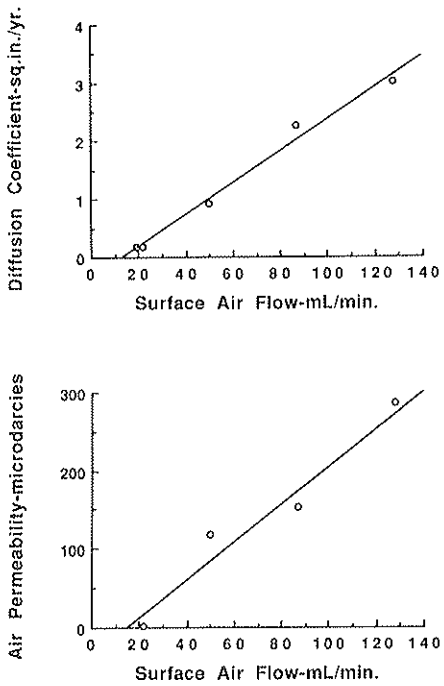



**FIGURE 4 Relations between surface airflow readings, chloride diffusion coefficients, and air permeability.**

- Portability. It had to be carried by one man, with total weight preferably less than 15 lb, and be capable of operating in horizontal, vertical, and overhead modes.
- Ability to operate from on-board power sources. A rechargeable battery source capable of powering the device for one working day was desirable. Although an on-board source was preferable, if this would degrade the portability of the device, an external source linked by cable to the device might be used.
- Rapidity of test and data collection. Although total test time (excluding setup or conditioning) for the laboratory device was only 1 min, it could be decreased even further if some of the manual operations, including timing of sequences such as opening and closing of valves, were automated. On-board data acquisition was also considered; however, aside from the increased weight, and power requirements, the developers decided that this might be an unnecessary complexity to add to what was, in effect, a first-generation field instrument. The simple expedient of recording each data point on grid paper keyed to the structure (much as is done for most half-cell corrosion potential surveys) was considered to be adequate.

### Construction of a Field Device

A conceptual drawing of a prototype field device is shown in Figure 5. The device is designed to use a dc vacuum pump powered by a rechargeable 12-volt Ni-Cad power supply. The flowmeter is an electronic mass flowmeter that utilizes the change in rate of heat flow through a heated tube to sense

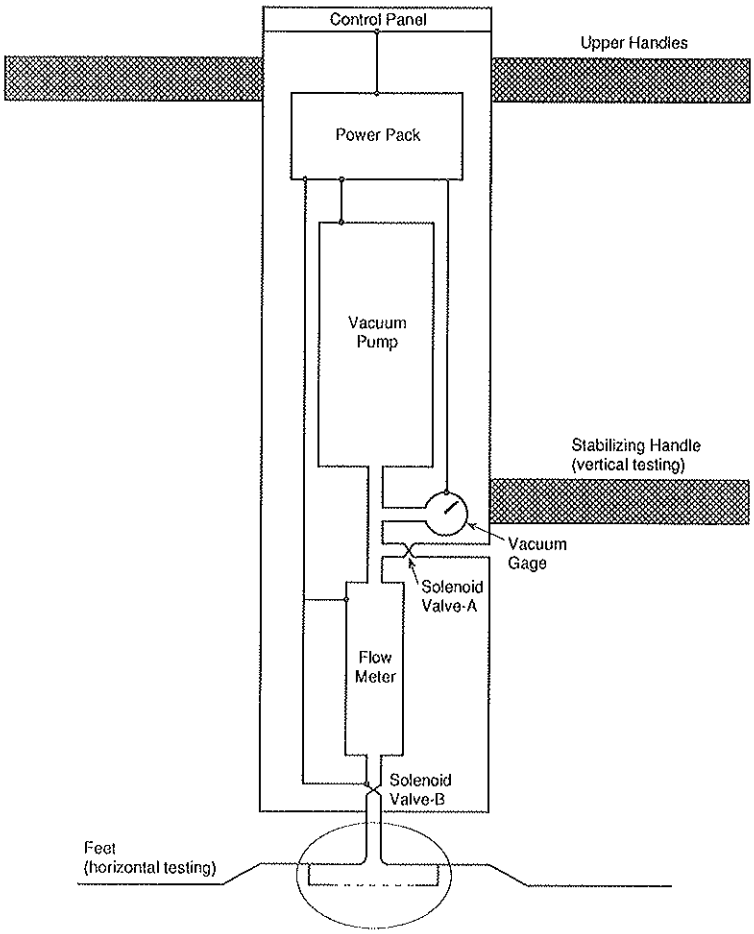



**FIGURE 5 Conceptual drawing of a prototype field surface airflow device.**

the mass flow through the sensor. By use of solenoid valves and electronic timers, the device is designed for one-man operation, and it is estimated that data could be gathered at the rate of approximately one test per minute. The field prototype was under construction in the summer of 1990 and field trials were planned after verification that the unit was operating properly.

## ACKNOWLEDGMENTS

This research was funded by the Strategic Highway Research Program, National Research Council, through a subcontract from the Pennsylvania State University. The author would like to acknowledge the participation of staff at Texas Research Institute, Austin, in design and construction of the field prototype.

## REFERENCES

1. H. D'Arcy. Les Fontaines Publiques de la Ville de Dijon. 1856.
2. R. M. Mantialay. Air Permeability of Concretes. Monografia 332, Instituto Eduardo Torroja de la Construccion y del Cemento, Madrid, Spain, Dec. 1975.
3. A. Hanaor and P. J. E. Sullivan. Factors Affecting Concrete Permeability to Cryogenic Fluids. *Magazine of Concrete Research*, Vol. 35, No. 124, Sept. 1983, pp. 142–150.
4. L. Chou Chen and D. Katz. Diffusion of Methane Through Concrete. *Proc., American Concrete Institute*, Vol. 75, No. 12, Dec. 1978, pp. 673–679.
5. A. J. Watson and C. C. Oyeka. Oil Permeability of Hardened Cement Pastes and Concrete. *Magazine of Concrete Research*, Vol. 33, No. 15, June 1981. pp. 85–95.
6. G. Wiley and D. C. Coulson. A Simple Test for Water Permeability of Concrete. (Discussion by S. P. Wing.) *Proc., American Concrete Institute*, Vol. 34, Jan.-Feb. 1938, pp. 76–1 to 76–4.
7. W. M. Dunagan. Methods for Measuring the Passage of Water through Concrete. *Proc., ASTM*, Vol. 39, 1939, pp. 866–880.
8. M. Levitt. The ISAT—A Non-Destructive Test for the Durability of Concrete. *British Journal of Non-Destructive Testing*, Vol. 13, No. 4, July 1971, pp. 106–112.
9. R. D. Browne. Design Prediction of the Life for Reinforced Concrete in Marine and Other Chloride Environments. *Durability of Building Materials*, Vol. 1, 1982, pp. 113–125.
10. H. Nagaro and T. Naito. Application of Diffusion Theory to Chloride Penetration into Concrete Located in Splashing Zones. *Transactions of the Japanese Concrete Institute*, Vol. 7, 1985, pp. 157–164.
11. A. Bisaillon and V. M. Malhotra. Permeability of Concrete Using a Uniaxial Water-Flow Method. In *Permeability of Concrete*, ACI SP–82 (D. Whiting and A. Walitt, eds.), 1988, pp. 175–194.

12. T. Bourbie and J. Walls. Pulse Decay Permeability: Analytical Solution and Experimental Test. *Journal of the Society of Petroleum Engineers*, Dec. 1982, pp. 719–721.
13. B. Mobasher and T. M. Mitchell. Laboratory Experience with the Rapid Chloride Permeability Test. In *Permeability of Concrete*, ACI SP–82 (D. Whiting and A. Walitt, eds.), 1988, pp. 117–144.
14. J. W. Figg. Methods of Measuring the Air And Water Permeability of Concrete. *Magazine of Concrete Research*, Vol. 25, No. 85, 1973, pp. 213–219.
15. K. Schonlin and H. K. Hilsdorf. Evaluation of the Effectiveness of Curing of Concrete Structures. In *Concrete Durability*, ACI SP–100 (J. Scanlon, ed.), 1987, pp. 207–226.
16. R. E. Weyers and D. G. Smith. Chloride Diffusion Constant for Concretes. *Proc., Structures Congress '89: Performance of Structural Materials*, ASCE, San Francisco, May 1989, pp. 106–116.

---



*Publication of this paper sponsored by Committee on Corrosion.*

# Protection and Rehabilitation Treatments for Concrete Bridge Components: Status and Service Life Opinions of Highway Agencies

W. P. CHAMBERLIN AND R. E. WEYERS

As part of Task 1 of Strategic Highway Research Program Project C–103, "Concrete Bridge Protection and Rehabilitation: Chemical and Physical Treatments," state and provincial highway agencies in the United States and Canada were surveyed in early 1989 by mailed questionnaire on the status and service life of protective and rehabilitative treatments applied to concrete components of bridges in their jurisdictions. Responses were received from 47 states and 9 provinces. Respondents indicated that patching with rigid mortar or concrete (portland cement, quick-set, or polymer) is more widely accepted as a standard practice than any other deck treatment category (71.4 percent of agencies). Some treatments were judged by more agencies to be experimental rather than standard and were associated with generally lower acceptance frequencies. With the exception of cathodic protection, treatments for substructure and superstructure concrete were judged to be far less experimental than those for decks, and the standard acceptance frequencies more uniform. Opinions on the service life of treatments were generally widely scattered. Median responses for deck treatments varied from 1 year for asphalt concrete patching to >20 years for micro-silica overlays; and for nondeck treatments from 5 to 10 years for sealers to 20 years for cathodic protection. Questionnaire responses have been used to focus the study of service life expectancy in Task 1 on those treatments considered to be in the mainstream of current practice.

The purpose of Strategic Highway Research Program (SHRP) Project C–103 is to develop cost-effective, nonelectrical methods of protecting and rehabilitating salt-contaminated decks and other concrete bridge components subject to corrosion. The object of Task 1 of C–103 is to determine costs and service lives of those treatment methods that are in *current* use (*1*).

In connection with Task 1, a short questionnaire was mailed to each of the 50 SHRP state coordinators and to each of the 12 Canadian provincial coordinators. The questionnaire was designed to serve several purposes:

1. To provide an opportunity for the client agencies to influence the direction of the Task 1 study by indicating which of the current treatments were considered to be standard and which were experimental (Question 1);
2. To obtain an indication of which of the treatments were most important to the agency's program (Question 2); and
3. To elicit a body of informed opinions on the average service lives of the treatments (Question 3).

W. P. Chamberlin, 292 Washington Avenue Extension, Albany, N.Y. 12303. R. E. Weyers, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061.

For each question, respondents were asked to provide answers for each of 26 combinations of treatment methods and bridge components. The treatment methods included were those thought to represent the common current alternatives. Cathodic protection (CP) was included for information, at the request of SHRP main office staff, even though it was not within the scope of Project C–103. Respondents were invited to add new treatments to the matrix reflecting practices not otherwise included.

Bridge components were grouped in the questionnaire matrix on the basis of assumptions regarding differences in chloride exposure, i.e.,

1. Those subject directly to chloride-laden runoff such as pedestals and pier cap beams; and
2. Those subject to spray or splash such as pier columns adjacent to a roadway, and elements in a marine environment.

The introduction that accompanied the questionnaire instructed that its intent was to "identify treatments that are expressly used as long-term solution rather than an immediate reaction to an existing problem," and the SHRP coordinators were asked to direct the questionnaire to the persons in their agency best qualified to respond.

## QUESTIONNAIRE RESULTS

Responses were received from 47 states and 9 Canadian provinces. One state, Mississippi, and one province, Prince Edward Island, responded but declined to complete the questionnaire. Mississippi does not use chloride deicing chemicals and Prince Edward Island, until recently, has built mostly timber bridges.

Four treatments not included in the original matrix were added voluntarily by respondents:

1. A proprietary rubber-modified asphalt concrete (AC) overlay by Pennsylvania (experimental);
2. In situ polymerization with methacrylate of high molecular weight by California (standard);
3. Bituminous chip seal by Kansas (experimental); and
4. An epoxy-coated deck seal by Wisconsin (experimental).

Of those, two (in situ polymerization as practiced by California and epoxy-coated deck seal) fall within the existing

treatment category of thin polymer overlay. The other two were added to the list of treatments in the matrix.

Individual questionnaire responses were discussed by Chamberlin (2).

## STATUS OF TREATMENTS (QUESTION 1)

In the discussion that follows, the terms "acceptance" and "frequency of acceptance" are used in referring to the number of agencies reporting a particular treatment as standard or experimental in their jurisdiction. These terms should not be confused with use or frequency of use, which are the number of applications of that treatment actually applied. For instance, a treatment may be accepted as a standard by an agency, but not widely used because of economic, logistic, or other reasons. This work deals with acceptance, not use, and reports the status of treatments as of the date the questionnaire was mailed, February 1989.

### Decks

The status of deck treatments is shown in Figure 1, which is a graphical representation of the survey responses, and presented in Table 1, which correlates results of this survey with results of two similar surveys, one conducted by TRB in 1977 (*3*), and another conducted by the New Mexico State Highway Department in 1984 (*4*,p.33)

For the purpose of this discussion, treatments are divided into two groups, topical and areal. Topical treatments are those used to repair damage at specific locations on a deck (i.e., patching and crack filling). Areal treatments are those typically applied to the entire deck surface at one time. On the basis of the relative extent and nature of their acceptance as revealed by this survey, areal treatments are further divided into those that are widely enough accepted to be considered conventional and those that are still highly experimental.

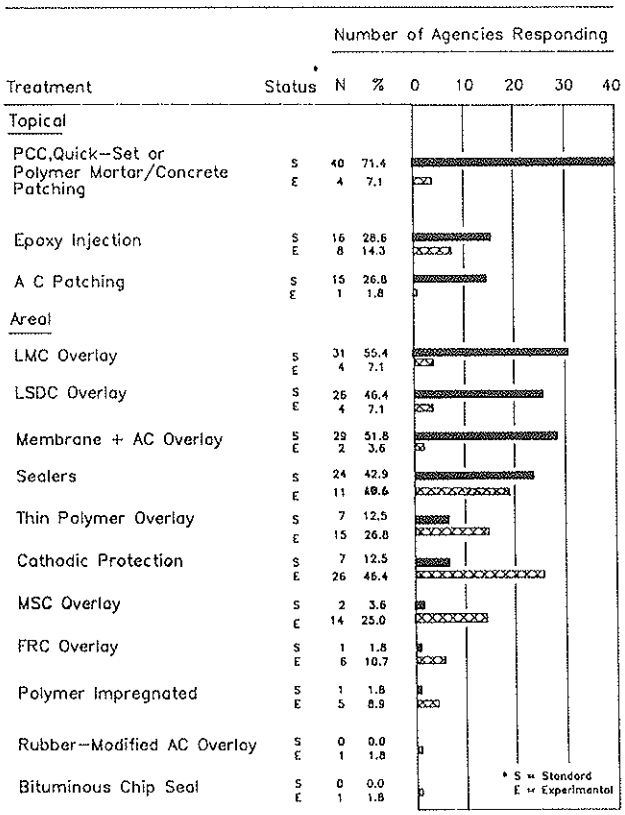



**FIGURE 1 Status of bridge deck treatments.**

#### *Topical Treatments*

Of the three treatments in the topical category, the various mortar and concrete patching materials were reported as standard at a frequency more than twice that of each of the other two (Figure 1). The remaining two, epoxy injection and AC patching, are still not widely accepted as standards for deck repair, but their acceptance has grown significantly since 1977 (from 6.2 to 34.0 percent of states and from less than 10.4 to 29.8 percent of states, respectively).

#### *Conventional Areal Treatments*

Of the 11 treatments in the areal category, four latex-modified concrete (LMC) and low-slump dense concrete (LSDC) overlays (membranes and sealers) were identified as standard with a high enough frequency, and as experimental with a low enough frequency, to set them apart from the other seven (Figure 1).

Each of the three overlay systems represents a technology that has been available in its present form for at least 15 years and, with the exception of membranes, has grown substantially in acceptance during that period (Table 1). Membrane systems, being a somewhat older approach to deck rehabilitation than concrete overlays, appear to have maintained a relatively constant level of acceptance, at least since 1977 (Table 1).

Forty-nine of the 56 responding agencies accept as standard at least one of these three overlay systems, an indication of the popularity of the overlay concept. The Venn diagram of Figure 2 shows the frequency with which the responding agencies have accepted as standard treatment more than one of these options. These associations appear to be random (chi square analysis at the 0.95 significance level), that is, acceptance of one of the treatments by an agency has not predisposed it to acceptance of another.

Although sealers are included with areal treatments, they differ in that they are used on existing decks only for protection, not for rehabilitation. They also have a higher ratio of experimental to standard acceptance than the other three. This situation undoubtedly reflects the continuing introduction of new proprietary materials to the market and the difficulties in evaluating these products, as well as the shift in interest away from surface sealers such as linseed oil to so-called "penetrating" sealers such as the salines. Also, what appears in Table 1 to be aggressive growth in the acceptance of sealers between 1977 and the present (8.3 versus 44.7 percent of states, respectively) may be explained merely by the failure of either survey to distinguish clearly among surface sealers, penetrating sealers, and coatings.

The geographic distribution of acceptance of the four conventional treatments within the United States is shown in Figure 3.

TABLE 1 CHANGES IN REHABILITATION PRACTICES FOR U.S. BRIDGE DECKS AS REFLECTED IN NATIONAL SURVEYS

| | Percentage of Respondents Indicating Use* | | | | | |
|---|---|---|---|---|---|---|
| | TRB (1977) (2) | | NM (1984) (3) | | SHRP (1989) | |
| Treatment | Std. | Exper. | Extive. | Lmtd. | Std. | Exper. |
| Sealers | 8.3 | 10.4 | Not Included | | 44.7 | 19.2 |
| Silane Treatment | Not Included | | 4.6 | 7.0 | Not Included | |
| Epoxy Injection | 6.2 | 4.2 | Not Included | | 34.0 | 12.8 |
| Polymer Impregn. | Not Included | | 7.0 | 0.0 | 2.1 | 10.6 |
| Patch with AC | 10.4 | 12.5 ** | Not Included | | 29.8 | 2.1 |
| Patch with PC Mortar/Concrete | | | Not Included | | 76.6 | 4.2 |
| Membrane + AC Overlay | 48.8 | 12.5 | 37.2 | 39.5 | 46.8 | 4.2 |
| Thin Polymer Overlay | Not Included | | Not Included | | 12.8 | 29.8 |
| Normal Slump PCC Overlay | 8.3 | 0.0 | Not Included | | Not Included | |
| LSDC Overlay | 31.2 | 6.2 | 32.6 | 39.5 | 51.1 | 8.5 |
| LMC Overlay | 35.4 | 14.6 | 30.2 | 44.2 | 63.8 | 6.4 |
| MSC Overlay | Not Included | | Not Included | | 2.1 | 29.8 |
| FRC Overlay | Not Included | | Not Included | | 0.0 | 12.8 |
| Cathodic Protection | 0.0 | 18.8 | Not Included | | 10.6 | 53.2 |

* Number of Respondents:
TRB . . . . . 48
New Mexico. . . . . 43
SHRP C-103. . . . . 47
** The 1977 survey did not distinguish among patching materials.

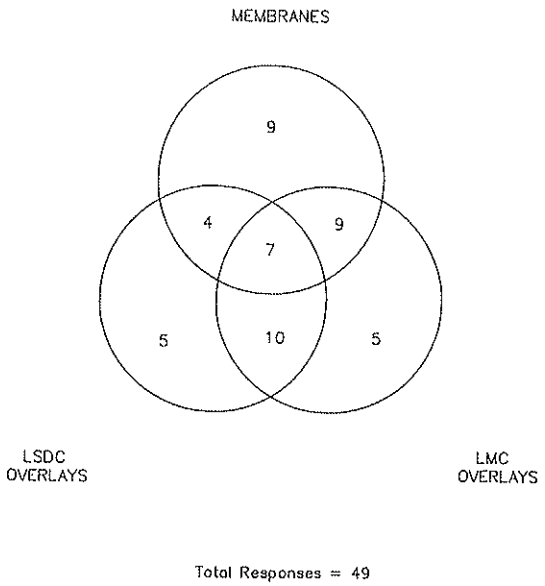



FIGURE 2 Acceptance frequency for conventional overlay treatments.

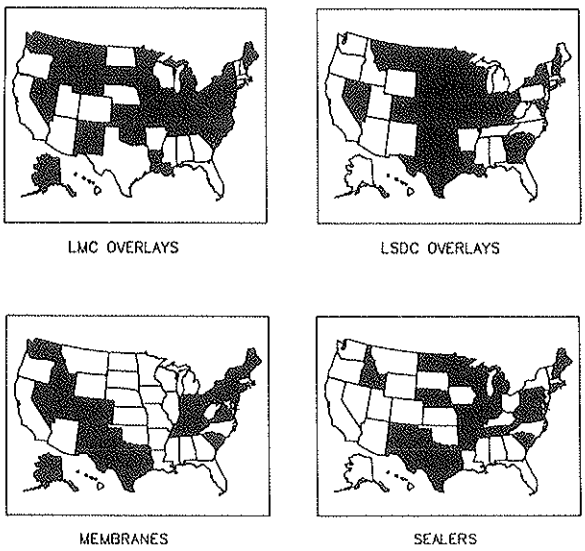



FIGURE 3 Geographic distribution of acceptance for conventional areal treatments.

*Experimental Areal Treatments*

The remaining seven areal treatments are each characterized by a substantially lower acceptance frequency, and, except for the two added by respondents, have an experimental to acceptance ratio that well exceeds 1. CP, micro-silica concrete (MSC) overlays, and thin polymer overlays appear to have experienced higher overall acceptance than the other four.

CP, the only widely promoted system to date that will arrest corrosion, has been an option for bridge deck application for approximately the same period of time as have LMC and LSDC overlays. Yet, its acceptance as a standard in the United States is still only a modest 10.6 percent of states compared to 63.8 and 51.1 percent for the other two, respectively (Table 1). This is in contrast to the strongest experimental interest indicated for any of the 14 deck treatments included in the survey. There are probably a variety of reasons for CP's lagging acceptance for decks and many of the reasons may be nontechnical (*5*).

In contrast to CP, MSC overlays are relatively new. The first experimental bridge deck overlay with MSC was placed in 1984 (*6*). Yet, such overlays are rapidly gaining in acceptance, in part because of their low reported permeability and their compatibility with conventional concrete overlay techniques.

Thin-polymer overlays have become particularly popular in those instances where minimizing dead load is a factor and where it is advantageous not to have to raise joints or approaches. The relatively strong experimental components to the responses for this treatment reflect the proprietary nature of many of the overlay systems that are on the market and the difficulties inherent in their evaluation.

### Elements Other Than Decks

The status of treatments applied to concrete bridge elements other than decks is shown in the bar charts of Figures 4 and 5.

With the exception of CP, all of the nondeck treatments were identified as standard by a substantial number of respondents (roughly comparable to that for the conventional areal deck treatments), with the various mortar and concrete patching materials having the greatest acceptance. None had a strong experimental component.

No consistent difference was indicated between treatments applied to elements subject to runoff and those subject to spray or splash. Where the same treatment was applicable both to deck and nondeck elements, a higher level of acceptance in favor of the former was indicated for mortar and concrete patching and for CP.

## IMPORTANCE OF SERVICE-LIFE AWARENESS (QUESTION 2)

In Question 2, agencies were asked to identify whether a knowledge of the service life of each of the treatments they had identified in Question 1, as either standard or experimental, was of primary or secondary importance. Responses to this question were taken as a measure of respondent's interest in service life awareness. The purpose of this information was to assist in setting priority for the task of collecting service life and cost data.

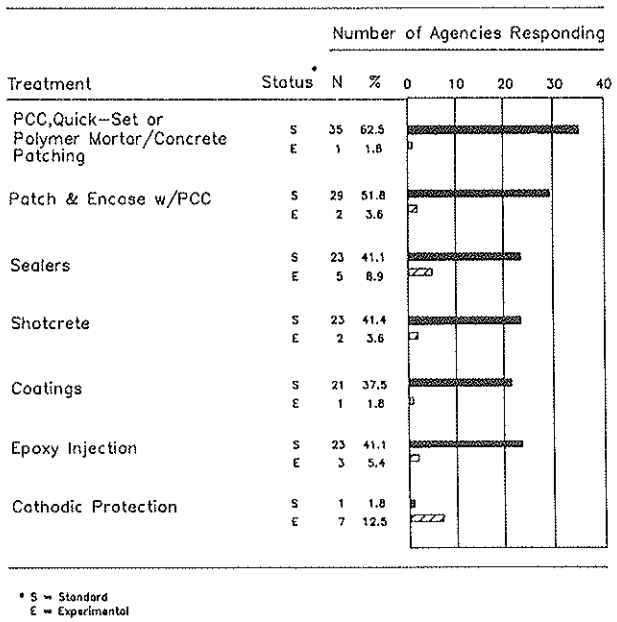



**FIGURE 4 Status of nondeck treatments for bridge elements subject to chloride-laden compounds.**

Accordingly, information on the frequency of treatment acceptance from Question 1, including both standard and experimental use, was plotted against the frequency of primary interest from Question 2, expressed as a percent of the former (see Figure 6). In Figure 6, the response field has been divided into four quadrants using the mean responses as boundaries. The identities of the treatments in each of the response quadrants are presented in Table 2 in the order listed in the questionnaire.

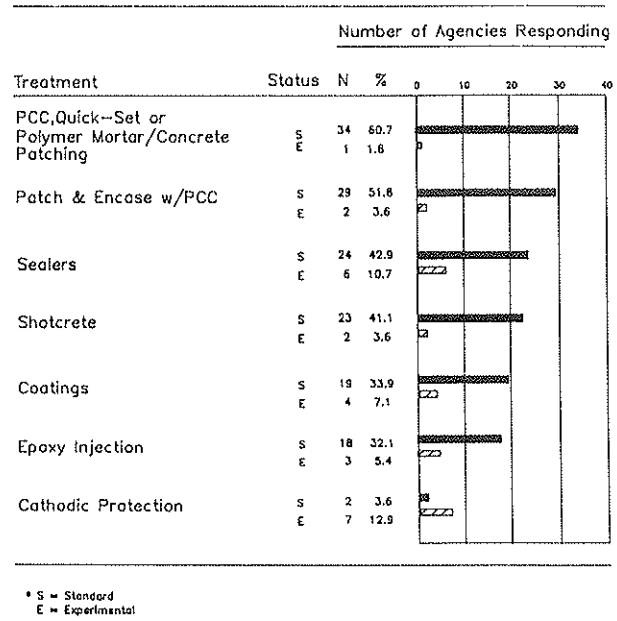



**FIGURE 5 Status of nondeck treatments for bridge elements subject to chloride-laden spray and splash.**

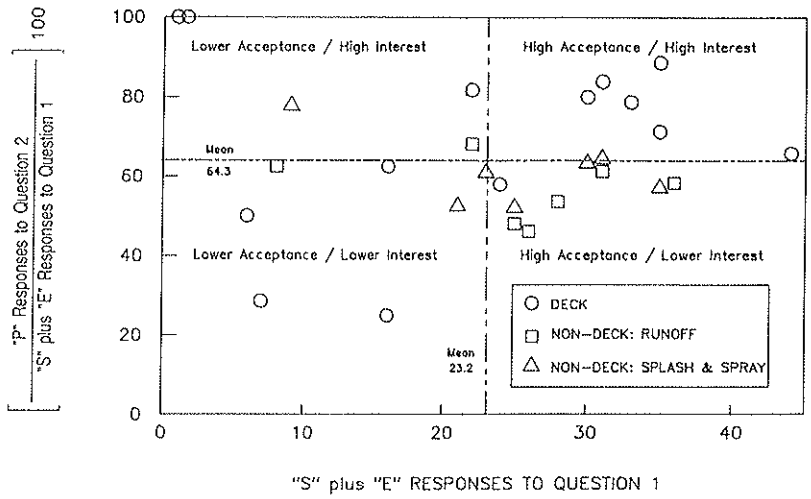



**FIGURE 6 Treatment acceptance versus service life awareness interest.**

Bridge deck treatments, as a group, rank high both in frequency of acceptance and frequency of primary interest. Also, most of the condition and performance data that exist relate to bridge deck treatments, in contrast to treatments for non-deck elements.

Except for bituminous concrete seal coats and rubber-modified AC overlays, for which there was only one response each, three anomalies stand out in Figure 6 in what is otherwise an approximately straight-line relation between acceptance and interest.

1. CP is associated with a high level of interest despite its relatively low acceptance. This is consistent with its high ratio of experimental to standard acceptance, noted earlier, and is taken as further evidence of the reluctance of highway agencies to make widespread use of this technology.

2. AC patching is associated with low to moderate acceptance relative to use, but there seems to be little interest in knowing its service life. The modest experimental interest in this treatment has already been noted, and it will be seen later that there is a strong respondent's perception that its service life is also modest, on the order of 1 year.

3. In contrast, thin polymer overlays, which are also associated with a moderate level of acceptance, have one of the highest levels of respondent interest measured. This is taken

TABLE 2 TREATMENTS IN EACH OF THE RESPONSE QUADRANTS

| Quadrant | Treatment | Bridge Element* |
|---|---|---|
| High Acceptance/High Interest | Sealers | Decks |
| | Mortar/Concrete Patching | Decks |
| | Membranes + AC Overlay | Decks |
| | LSDC Overlay | Decks |
| | LMC Overlay | Decks |
| | Cathodic Protection | Decks |
| High Acceptance/Lower Interest | Sealers | Non-Deck (R) |
| | Sealers | Non-Deck (S/S) |
| | Epoxy Injection | Decks |
| | Epoxy Injection | Non-Deck (R) |
| | Mortar/Concrete Patching | Non-Deck (R) |
| | Mortar/Concrete Patching | Non-Deck (S/S) |
| | Patch & Encase | Non-Deck (R) |
| | Patch & Encase | Non-Deck (S/S) |
| | Shotcrete | Non-Deck (R) |
| | Shotcrete | Non-Deck (S/S) |
| Lower Acceptance/High Interest | Coatings | Non-Deck (R) |
| | Thin Polmer Overlays | Decks |
| | Cathodic Prtection | Non-Deck (S/S) |
| | Bituminous Chip Seal | Decks |
| | Rubber-Modified AC Overlay | Decks |
| Lower Acceptance/Lower Interest | Coatings | Non-Deck (S/S) |
| | Epoxy Injection | Non-Deck (S/S) |
| | Polymer Impregnation | Decks |
| | AC Patching | Decks |
| | Cathodic Protection | Non-Deck (R) |
| | MSC Overlay | Decks |
| | FRC Overlay | Decks |

* R = elements subject to chloride-laden runoff
S/S = elements subject to chloride-laden splash and spray

as a reflection of the niche (noted earlier) that these materials occupy among rehabilitation options and of the potential for product development in this area.

## ESTIMATED SERVICE LIFE (QUESTION 3)

Question 3 asked for the respondent's best estimate of the average useful life (service life expectancy) of each of the combinations of treatments and bridge components listed in the response matrix.

Most of the answers to Question 3 were in the form of discrete numbers. However, some were given as ranges and others as numbers followed by a plus sign. One response simply said, "indefinite." For these reasons, values of the median and mode (rather than the arithmetic mean) were reported as measures of central tendency in summarizing the responses. Minimum and maximum values, as well as interquartile ranges, were used as measures of scatter.

The responses for bridge deck treatments are presented in Table 3 where they are arranged in order of increasing median value of the service life estimate, within the categories of topical and areal. Responses for nondeck treatments are presented in Table 4, also arranged in order of increasing median value of the service life estimate, but without further categorization.

### Deck Treatments

Considering the deck treatments first (Table 3), those typically considered more in the nature of repair or protection, as distinguished from major rehabilitation, are associated with the expectation of a shorter service life. The median life expectancies for both forms of patching, epoxy injection and sealers, were judged to be between 1 and 10 years. The single most consistent response for any of the treatments was that for AC patching, with a median estimated life of 1 year.

In contrast, those treatments that are of a more rehabilitative nature were associated with median service lives estimated at between 10 and 20 years. Several observations follow.

1. Apparently, most respondents perceive LSDC and LMC overlays as roughly equivalent in terms of their expected life. This has been traditional as long as these materials are placed in thicknesses that reflect the presumed difference in their inherent chloride permeability. In this regard, LSDC has usually been placed at a nominal thickness ½ to ¾ in. greater than LMC, reflecting its higher permeability. However, several recent studies have indicated that much of the LSDC in actual use is more permeable to chlorides than laboratory studies have indicated (*7–9*). Awareness of this information, which would seem to have a performance implication, does not appear to be reflected in current opinion regarding the relative service lives of these two overlays.

TABLE 3 ESTIMATED SERVICE LIVES FOR DECK TREATMENTS (IN YEARS)

| | | Average* | | Scatter* | | |
|---|---|---|---|---|---|---|
| Treatments | N | Median | Mode | Min. | Max. | Inter-Quartile Range |
| Topical | | | | | | |
| AC Patching | 30 | 1 | 1 | 0 | 25 | 1-2 |
| Mortar/Concrete Patching | 45 | 5 | 5 | 1/2 | 35 | 4-10 |
| Epoxy Injection | 28 | 10 | 10 | 4-5 | 50 | 10-20 |
| Conventional Areal | | | | | | |
| Sealers | 32 | 4-5 | 5 | 1 | 25 | 2-10 |
| Membranes + AC Overlay | 39 | 15 | 15 | 5 | 60 | 10-15 |
| LMC Overlay | 40 | 15-20 | 20 | 0 | 60 | 10-20 |
| LSDC Overlay | 36 | 20 | 20 | 0 | 50 | 10-20 |
| Experimental Areal | | | | | | |
| Thin Polymer Overlay | 29 | 10 | 10 | 2 | 25+ | 6-12 |
| Polymer Impregnation | 7 | 15 | NA | 1-10 | 30+ | NA |
| FRC Overlay | 7 | 15 | NA | 0 | 25 | NA |
| Cathodic Protection | 28 | 20 | 20 | 1 | Indef. | 15-30 |
| MSC Overlay | 13 | 20+ | NA | 10 | 60 | 20-25 |

* Where "NA" appears in the "Mode" column, it indicates either that the number of responses is too few for the mode to have meaning or that the distribution of responses is multimodal. Where it appears in the "Interquartile Range" column, it indicates that the number of responses is too few to identify a meaningful interquartile range.

TABLE 4 ESTIMATED SERVICE LIVES FOR NONDECK TREATMENTS (IN YEARS)

| | | | Average* | | Scatter* | | |
|---|---|---|---|---|---|---|---|
| Treatments | Element | N | Median | Mode | Min. | Max. | Inter-Quartile Range |
| Sealers | R | 30 | 5-10 | 10 | 2 | 25+ | 3-10 |
| Sealers | S/S | 29 | 5-10 | 10 | 2 | 35 | 3-10 |
| Concrete Patching | R | 36 | 10 | 10 | 2 | 35 | 5-10 |
| Concrete Patching | S/S | 37 | 10 | 10 | 2 | 35 | 5-10 |
| Coatings | R | 29 | 10 | 10 | 4 | 25 | 7-15 |
| Coatings | S/S | 31 | 10 | 10 | 2-3 | 25 | 5-10 |
| PCC Patch & Encase | R | 29 | 10 | 10 | 5 | 40 | 10-20 |
| PCC Patch & Encase | S/S | 33 | 10 | 10 | 2 | 40 | 10-20 |
| Shotcrete | R | 33 | 10+ | 10 | 0 | 40 | 10-15 |
| Shotcrete | S/S | 31 | 10+ | 10 | 0 | 40 | 10-15 |
| Epoxy Injection | R | 29 | 15 | NA | 4-5 | 50 | 10-25 |
| Epoxy Injection | S/S | 25 | 15 | 10 | 0-1 | 50 | 10-20 |
| Cathodic Protection | R | 10 | 15-20 | NA | 10 | 50 | NA |
| Cathodic Protection | S/S | 11 | 20 | NA | 5 | 50 | NA |

* Where "NA" appears in the "Mode" column, it indicates either that the number of responses is too few for the mode to have meaning or that the distribution of responses is multimodal. Where it appears in the "Interquartile Range" column, it indicates that the number of responses is too few to identify a meaningful interquartile range.

2. Similarly, MSC, which has been shown in laboratory studies to be inherently less permeable than either LSDC and LMC, is also rated as equivalent in respondents' perceptions of its service life in overlays.

3. All three of the concrete overlay systems (LSDC, LMC, and MSC) were associated with lower estimated service lives than expected. This may reflect the fact that individuals' perception of service life is probably influenced more by the age at which the first items in the population fail than the average age at which they fail. This is likely to be particularly true for treatments that have been in use for a period of time less than their average service life.

4. Membranes were found to have a level of acceptance roughly equivalent to LMC and LSDC overlays (Figure 1) and a greater perceived service life (15 years) (*3*) than has generally been assumed, at least as reflected by the median estimated value.

## Nondeck Treatments

Among the treatments applied to elements other than decks, there was essentially no difference in respondents' opinions regarding the service life of those elements subject to direct chloride-laden runoff and those subject to chloride-laden splash or spray (Table 4). In retrospect, it may have been more informative to have categorized the nondeck treatments in terms of the nature of the surface of repair (e.g., vertical, horizontal, and irregular).

For those treatments applicable both to deck and nondeck elements, respondents' opinions favored a longer service life for the nondeck applications, except for CP (Figure 7). This undoubtedly reflects the perception that decks generally present a more severe environment for the durability of rehabilitative treatments than do other elements of the bridge.

## General Comments

Considering these responses as data, they are in the realm of expert opinion collected under highly unstructured conditions. Thus, any single response taken by itself should probably be looked on more as representing the individual respondent than the agency by which that respondent is employed. This judgment is supported by the surprising divergence observed between two completed questionnaires prepared and submitted by different persons in the same agency, each apparently without knowledge of the other.

The information on service lives for both deck and nondeck treatments offers the only concensus judgment to date of the relative service lives of the different treatment options in common use. Clearly, it oversimplifies. For instance, the effectiveness of any of the treatments is dependent on a complex interaction along material characteristics, condition, and pre-

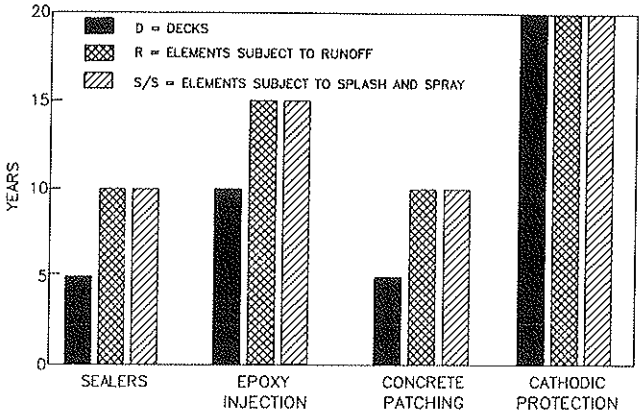



**FIGURE 7 Comparison of service life estimates for deck and nondeck treatments.**

treatment of the bridge element being enhanced; the techniques and skill of those actually doing the work; and the environment of service. Some of the factors contributing to this performance variance are identifiable, others are not. Notwithstanding, it is believed that there is a base of common experience that permits a general ranking of the treatments with respect to one another and that that experience is reflected in the median service life estimates reported here. Likewise, the scatter, which in the case of some treatments was surpassingly large, is believed to reveal useful information about variance in the perception and experience of the respondents.

No geographical influence was detected in the responses to Question 3, suggesting that the respondents, as a group, do not perceive a relationship between treatment service life and severity of the climate in which the agency operates.

## CONCLUSIONS

The following general conclusions have been drawn from results of the questionnaire on the status and service life of protection and rehabilitation treatments for concrete bridge components.

1. Among deck treatments, rigid mortar and concrete patching; overlays (including LMC, LSDC, and membranes with AC); sealers; epoxy injection; and AC patching are considered by more highway agencies to be standard practices than experimental, and in that order of decreasing frequency.
2. Acceptance by highway agencies of cathodic protection, thin polymer overlays, MSC overlays, fiber-reinforced concrete overlays, polymer impregnation, rubber-modified AC overlays, and bituminous chip seals as experimental treatments is more common than their acceptance as standard; and in that order of decreasing frequency.
3. Among treatments for bridge elements other than decks, only CP is considered to be experimental by more agencies than consider it to be a standard treatment.
4. There is a general and positive correlation between the frequency of acceptance as standard and the frequency of interest in service life awareness (as reflected in the responses to Question 1).
5. Opinions regarding the service life of treatments are generally widely scattered and bear little relationship to the severity of the climate in the responding agency's jurisdiction.
6. No difference was found between estimates of service life for nondeck treatments applied to elements subject to direct runoff and those subject to splash or spray.
7. Where the same treatment was applicable both to decks and other bridge elements, longer service lives were estimated for the nondeck elements.

## REFERENCES

1. R. E. Weyers. *Concrete Bridge Protection and Rehabilitation: Chemical and Physical Techniques*. Research Plan for SHRP Contract C-103, Virginia Polytechnic Institute and State University, Blacksburg, Aug. 5, 1988.
2. W. P. Chamberlin. Summary of the Field Survey Questionnaire. Strategic Highway Research Program, TRB National Research Council, Washington, D.C., June 1989.
3. *NCHRP Synthesis of Highway Practice 57: Durability of Concrete Bridge Decks*, TRB, National Research Council, Washington, D.C., May 1979.
4. R. M. Tachau and R. B. McPherson. *A Study of New Mexico Bridge Deck Protective Systems*. The New Mexico State Highway Department, Sante Fe, July 1984.
5. V. Fairweather. Clearing the Decks. *Civil Engineering*, Sept. 1985, pp. 56–59.
6. M. D. Luther. Silica Fume (Microsilica) Concrete in Bridges in the USA. Presented at the 67th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 1988.
7. W. P. Chamberlin. Performance and Service Life of Low-Slump Concrete Bridge Deck Overlays in New York State. *Proc.: Corrosion/87 Symposium on Corrosion of Metals in Concrete*, National Association of Corrosion Engineers, 1987. (Also published in *Corrosion*, Vol. 44, No. 6, June 1988, pp. 397–403).
8. D. Whiting and L. A. Kuhlmann. Curing and Chloride Permeability. *Concrete International*, April 1987.
9. D. Whiting and W. Deziedzic. Chloride Permeabilities of Rigid Concrete Bridge Deck Overlays. Presented at the 68th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 1989.

*Publication of this paper sponsored by Committee on Corrosion.*

# Surface Characterization of Reinforcing Steel and the Interaction of Steel with Inhibitors in Pore Solution

J. G. DILLARD, J. O. GLANVILLE, T. OSIROFF, AND R. E. WEYERS

Studies of rebar surfaces following cleaning in various ways and following exposure to corrosive solutions in the presence and absence of inhibitors have been carried out using the surface-sensitive technique: x-ray photoelectron spectroscopy (XPS). After cleaning in hexane, rebar specimens were exposed to simulated pore solution with or without corrosion inhibitors under specified conditions. Subsequent XPS analysis of the rebar specimens indicated that the corrosion inhibitors sodium nitrite, sodium molybdate, sodium dihydrogen phosphate, sodium monofluorophosphate, and sodium tetraborate produced changes in rebar surface chemistry that could be associated with corrosion inhibition. The principal changes were (a) alteration in surface iron content, (b) reduction in the surface hydroxide concentration, and (c) increase in surface oxide oxygen concentration. The results are interpreted to indicate that these inhibitors promote the formation of surface oxides at the expense of hydroxide functionality. Results from the study of sodium tetraborate reveal that this inhibitor produces a coating on the rebar surface.

The deterioration of reinforced concrete structures caused by corrosion of reinforcing steel has long been recognized (*1,2*). For highway bridges, chloride ions from deicing salts interacting with steel have been implicated in accelerating the degradation of steel and spalling of concrete structures (*3*). The initiation of chloride-induced corrosion of steel in concrete occurs via localized attack or pitting corrosion (*3,4*). Treatment of corroded structures with inhibitors offers a practical solution to the corrosion of steel in bridges (*5–7*). Corrosion inhibitors that have shown promise and are of current interest include calcium or sodium nitrite (*8–13*), stannous chloride (*12*), and sodium benzoate (*14*). As a part of a program to evaluate the effectiveness of inhibitors for the repair of reinforced bridge structures, to identify and test potential new inhibitors, and to investigate the surface chemistry of reinforcing steel (rebar) following aqueous treatment, a surface-sensitive analytical technique, combining especially electron spectroscopy for chemical analysis (ESCA) and x-ray photoelectron spectroscopy (XPS), has been used (*15,16*). The analyses were carried out to determine the chemical nature of inhibitor constituents on the rebar surface, to evaluate the surface concentration of inhibitor elements, and to correlate the results with corrosion test experiments. It was reasoned not only that surface analysis measurements could aid in understanding the role and mechanism of corrosion inhibitor action, but also that such measurements would be valuable for determining the effectiveness of inhibitors in short-term screening tests.

J. G. Dillard, J. O. Glanville, and T. Osiroff, Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0212. R. E. Weyers, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0105.

## EXPERIMENTAL SECTION

Rebar rods were obtained from Roanoke Electric Steel Co., Roanoke, Virginia. The rod composition was formulated to be similar to material used 20 to 30 years ago. Bulk analysis of the rod material provided by the vendor is presented in Table 1. The rebar rods were ½ in. in diameter and 6 ft long. Test specimens were prepared by first cutting the bar in half longitudinally and then 1-in. specimens were cut from the split bar. Rebar cleaning experiments were carried out so that throughout the study a common pretreatment designed to remove grease and dirt would be used, and thus a kind of standard surface would be studied. To this end, organic solvents (hydrocarbon, alcohol, and ketone) and an aqueous acid solution were investigated to discover which treatment altered the as-received rebar surface to the smallest extent. Solvents used for tests to select a cleaning solvent included hexane, isopropanol, and acetone. Rods were also cleaned in a 50 weight percent (wt.%) sulfuric acid:distilled water solution for 1 min at room temperature, rinsed three times with distilled water, and dried at 110°C. To facilitate the preparation of samples for surface analysis, a notch was cut in the 1 in. specimens at approximately ¾ in. from one end of the specimen. After immersion in the inhibitor test solution, the ¾ in. portion of the treated bar was separated from the 1 in. specimen and analyzed. By using this procedure, the integrity of the treated rebar surface could be maintained, in that no cutting of the samples was required following treatment. The curved, outer portion of the rebar specimen was analyzed.

The test solutions were simulated pore solution [KOH (0.600 M); NaOH (0.300 M); saturated with $Ca(OH)_2$]; pore solution containing NaCl (3.5 wt.%); and pore solution containing NaCl (3.5 wt.%) and inhibitor (0.3 M). The inhibitors studied included sodium nitrite ($NaNO_2$), sodium molybdate ($Na_2MoO_4$), sodium dihydrogen phosphate ($NaH_2PO_4$), sodium monofluorophosphate ($Na_2PO_3F$), sodium tetraborate ($Na_2B_4O_7$), and three commercial reagents; Dequest 2000 [aminotri(methylene phosphoric acid), 50 percent active aqueous solution]; Dequest 2010 [1-hydroxyethylidene-1,1-diphosphonic acid, 60 percent active aqueous solution]; and

TABLE 1 BULK (wt.%) AND SURFACE ANALYSIS (at.%) OF REBAR MATERIAL

| | | | Rebar Treatment atomic % (measured; XPS) | | | |
|---|---|---|---|---|---|---|
| Element | wt% | calc. at. % | as received | alcohol | hexane | sulfuric acid |
| C | 0.22 | 1.0 | 58.4 | 53.7 | 56.4 | 51.9 |
| O | not analyzed | | 28.7 | 36.7 | 31.6 | 30.3 |
| N | not analyzed | | 1.41 | 1.10 | 1.09 | 0.96 |
| Fe | 97.2 | 96.5 | 3.68 | 3.97 | 4.64 | 7.54 |
| P | 0.018 | 0.32 | <0.1 | <0.1 | <0.1 | <0.1 |
| S | 0.036 | 0.062 | 0.85 | 0.81 | 0.62 | 0.58 |
| Si | 0.59 | 1.16 | 3.26 | 3.27 | 2.56 | 2.04 |
| Na | not analyzed | | 0.97 | <0.1 | 0.68 | 0.78 |
| Ca | not analyzed | | 0.45 | 0.43 | 0.59 | <0.1 |
| Mn | 1.00 | 1.00 | <0.1 | <0.1 | <0.1 | <0.1 |
| Al | 0.006 | 0.01 | <0.1 | <0.1 | <0.1 | <0.1 |
| Cu | 0.26 | 0.23 | 2.06 | 1.02 | 1.70 | 6.01 |
| Zn | not analyzed | | 0.18 | <0.1 | 0.10 | <0.1 |
| Cd | not analyzed | | <0.1 | <0.1 | <0.1 | <0.1 |

Dequest 2054 [hexapotassium hexamethylenediaminetetra(methylene phosphonate), 35 percent active aqueous solution]. The Dequest compounds obtained from Monsanto Chemical Co. had the following compositions:

Dequest 2000

$N[CH_2P(O)(OH)_2]_3$

Dequest 2010

$CH_3\text{-}C\text{-}(P(O)(OH)_2)_2$
$|$
$OH$

Dequest 2054

$[(HO)_2(O)PCH_2]$
$|$
$[(KO)_2(O)PCH_2]N\text{-}CH_2CH_2N\text{-}[CH_2\text{-}P(O)(OK)_2]_2$

The test solutions were aerated for at least 1 hr before rebar samples were introduced into the solutions. Exposure time was 8 days. Samples were maintained at 60°C. The exposure procedure was to run five replicate samples at each exposure to provide data for statistical analysis of the results of exposure for individual samples.

Two kinds of experiments were carried out with respect to rebar immersion in pore solutions. In experiments termed "initial inhibition," hexane-cleaned rebar was immersed for 8 days at 60°C in pore solution containing inhibitor (0.300 M) and NaCl (3.5 wt.%). At the end of the exposure, rebar was removed, rinsed with distilled water, and characterized by XPS. For experiments indicated as "delayed inhibition," hexane-cleaned rebar was immersed for 8 days at 60°C in pore solution containing NaCl (3.5 wt%). At the end of this period the specimens were then immersed for 8 days at 60°C in pore solution containing NaCl (35 wt.%) plus inhibitor (0.300 M). At the end of this exposure time, the rebar samples were removed from solution, washed with distilled water, and the surface chemistry evaluated with XPS. The test temperature of 60°C was selected to accelerate the rate of inhibitor interaction with rebar. XPS results for experiments carried out at room temperature (23°C), indicated that the oxides formed were equivalent to those produced at 60°C.

Surface analysis measurements were carried out using a PHI Perkin-Elmer 5300 photoelectron spectrometer (*16*). Photoelectrons were generated using Mg $K_\alpha$ radiation [$h\nu$ = 1,253.6 electron volts (eV)]. Ejected photoelectrons were analyzed using a hemispherical analyzer and the electrons were detected using a position-sensitive detector. In the presentation of elemental analysis results, photoelectron spectral peak areas were measured and subsequently scaled to account for ionization probability and an instrumental sensitivity factor to yield results that were indicative of surface concentration in atomic percent. The precision for the concentration evaluations was determined from measurements on five different rebar specimens and the results are given in the tables in parentheses. The binding energy (BE) scale was calibrated by setting the C 1s hydrocarbon peak BE value at 285.0 eV (*17*). At least two different measurements on two different rebar samples were made and the average results are given.

## RESULTS AND DISCUSSION

The surface analysis results following the treatment of rebar in selected solvents and solutions are presented in Table 1. Hexane cleaning was the treatment selected for the following reasons:

1. The chemical content and the chemical nature of the surface elements on the rebar surface are not altered significantly.
2. Residual solvent on the treated surface is minimal and less than that found following treatment with alcohol, acetone, or other organic solvents.

Following alcohol cleaning, surface concentrations of oxygen increase and surface concentrations of copper and zinc decrease. In addition, the chemical nature of carbon is altered to about 25 atomic percent (at.%) for -COR (for ether, R = alkyl; for alcohol, R = H) functionality, whereas the concentration of this group on as-received rebar is about 10 at.%. For rebar treated with sulfuric acid, significant surface con-

centration changes are noted for the metals. Iron and copper concentrations increase, whereas the surface contents for calcium and zinc decrease below the detection level (<0.1 at.%). Thus, cleaning of rebar with sulfuric acid results in a significant change in surface chemical content. Because the purpose of the present experiments is to simulate as closely as possible long-term exposure to corrosive conditions of rebar used in the construction or repair of bridges, the severe alterations caused by sulfuric acid cleaning are not desirable for this study. The most representative surface is that of rebar that has been cleaned in hexane and dried in an oven.

These studies were made primarily on rebar exposed to corrosion at 60°C. This temperature had been selected for related accelerated corrosion studies (see Dressman et al., a companion paper in this Record). Compared with studies at room temperature, no difference was found between surface oxides formed at 23°C and 60°C. It appears from these findings that the rate of corrosion changes with temperature but the mechanism does not change.

The surface analysis results following the immersion of rebar in pore solution containing NaCl are compared with the corresponding data presented for hexane-cleaned rebar in Table 2. The principal alterations in surface chemistry as a result of immersion of hexane-cleaned rebar in pore solution containing NaCl are increases in oxygen, iron, silicon, sodium, potassium, and chlorine; and decreases in carbon, nitrogen, and copper. The presence of calcium, potassium, and chlorine on the treated rebar may arise from adsorption of these elements on the oxide surface. Associated with the alteration in the oxygen atomic concentration is a change in the shape of O 1s photopeak. The spectra shown in Figure 1c exhibit features attributed to oxide oxygen (BE ≅ 529 eV) as the dominant peak for hexane-cleaned rebar. In the spectra in Figure 1b for rebar immersed in pore solution, contributions by hy-

TABLE 2 SURFACE ANALYSIS RESULTS (IN at.% WITH STANDARD DEVIATIONS IN PARENTHESES) FOR REBAR SPECIMENS FOLLOWING INTERACTION WITH INHIBITORS

| Treatment | Element | C | O | N | Fe | Si | Na | Ca | Cl | Inhibitor element |
|---|---|---|---|---|---|---|---|---|---|---|
| Hexane Cleaned | | 60.3 (5.3) | 28.3 (3.5) | 1.83 (0.67) | 4.07 (1.04) | 1.33 (0.78) | 1.67 (0.9) | 0.21 (0.2) | --- | |
| 8 days chloride-contg. pore sol. 60° | | 37.4 (2.6) | 43.0 (3.2) | 0.57 (0.23) | 6.43 (0.89) | 2.69 (0.15) | 3.72 (0.46) | 4.24 (0.28) | 1.33 (0.08) | --- |
| 0.300M $NaNO_2$ | Initial Inhibition | 49.8 (8.3) | 33.2 (4.9) | 1.11 (0.96) | 5.23 (2.27) | 2.08 (0.38) | 3.90 (1.03) | 1.78 (1.01) | 1.10 (0.22) | --- |
| | Delayed Inhibition | 57.3 (3.7) | 30.7 (2.8) | 0.47 (0.12) | 5.02 (0.58) | 2.39 (1.02) | 1.92 (1.11) | 1.27 (0.22) | 0.44 (0.13) | --- |
| 0.300M $Na_2MoO_4$ | Initial Inibition | 41.6 (3.7) | 39.6 (2.2) | 1.13 (0.21) | 7.29 (1.08) | 3.31 (0.56) | 4.07 (0.83) | 0.81 (0.27) | 1.01 (0.09) | Mo 0.46 (0.15) |
| | Delayed Inhibition | 30.6 (0.74) | 45.4 (2.7) | 1.92 (0.43) | 9.13 (1.92) | 5.07 (1.98) | 3.83 (0.39) | 1.67 (0.83) | 1.06 (0.51) | Mo 0.52 (0.07) |
| 0.300M $NaH_2PO_4$ | Initial Inibition | 35.1 (3.84) | 42.3 (0.41) | 0.32 (0.40) | 5.94 (1.42) | 1.67 (0.36) | 10.2 (2.6) | 0.14 (0.03) | 1.81 (0.55) | P 1.54 (0.44) |
| | Delayed Inibition | 20.1 (7.4) | 48.8 (4.4) | 0.22 (0.17) | 12.0 (1.3) | 2.5 (0.31) | 11.5 (3.3) | 1.21 (1.6) | 1.85 (0.49) | P 2.27 (0.55) |
| 0.300M $Na_2PO_3F$ | Initial Inibition | 48.5 (9.9) | 32.6 (5.7) | 0.38 (0.87) | 5.82 (3.2) | 2.04 (0.55) | 6.60 (2.80) | 0.06 (0.09) | 1.54 (1.05) | F 0.27 (0.27) P 0.71 (0.29) |
| | Delayed Inhibition | 35.5 (4.6) | 40.3 (3.9) | 0.32 (0.23) | 8.39 (0.86) | 3.01 (0.78) | 6.15 (1.83) | 0.50 (0.23) | 1.55 (0.05) | F 0.77 (0.18) P 1.20 |
| 0.300M $Na_2B_4O_7$ | Initial Inhibition | 61.1 (4.9) | 29.4 (4.1) | 0.19 (0.27) | <0.1 | 1.03 (0.76) | 3.23 (1.58) | 3.15 (1.16) | 1.20 (0.51) | B 5.03 (3.81) |
| | Delayed Inhibition | 39.7 (2.3) | 38.8 (4.8) | 1.30 (1.13) | 0.29 (0.51) | 0.73 (0.35) | 7.13 (1.82) | 1.95 (0.30) | 4.77 (2.12) | B 6.14 (1.50) |
| 0.300M Dequest 2000 | Initial Inhibition | 18.9 (3.9) | 48.9 (2.8) | 4.44 (0.33) | 5.73 (1.97) | 0.55 (0.35) | 5.97 (0.68) | 0.5 (0.33) | 0.08 (0.12) | P 13.2 (0.9) |
| | Delayed Inhibition | 18.7 (1.4) | 45.6 (1.9) | 4.39 (0.17) | 6.60 (0.42) | 0.78 (0.23) | 5.75 (1.82) | 0.10 (0.18) | 0.64 (0.72) | P 13.6 (0.3) |
| 0.300M Dequest 2010 | Initial Inhibition | 17.9 (1.4) | 50.5 (0.7) | <0.1 | 14.6 (1.5) | 4.56 (1.21) | 4.12 (1.35) | 0.34 (0.11) | 2.20 (0.81) | P 2.72 (0.45) |
| | Delayed Inhibition | 21.3 (9.9) | 48.6 (4.9) | 0.27 (0.38) | 6.20 (1.51) | 2.53 (0.73) | 5.70 (1.69) | 0.10 (0.10) | 1.40 (0.34) | P 2.59 (0.78) |
| 0.300M Dequest 2054 | Initial Inhibition | 47.4 (2.8) | 33.7 (3.0) | 1.50 (0.37) | 4.06 (1.50) | 1.56 (0.30) | 2.28 (0.49) | 2.74 (1.02) | 0.66 (0.36) | P 3.06 (0.58) |
| | Delayed Inhibition | 44.3 (8.7) | 36.9 (5.0) | 1.08 (0.16) | 3.80 (0.88) | 1.69 (0.88) | 2.07 (0.90) | 2.07 (0.90) | 0.63 (0.22) | P 2.69 (1.08) |

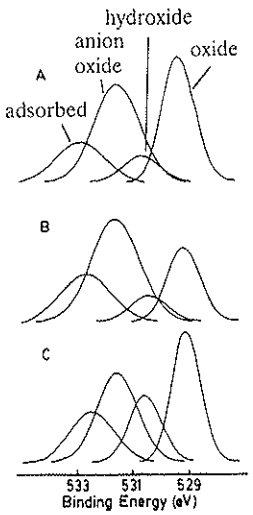



**FIGURE 1 Oxygen 1s photoelectron spectra and curve-resolved O 1s spectra for rebar specimens: (a) rebar treated for 8 days at 60°C in chloride-containing pore solution and $NaNO_2$ (0.300 M); (b) rebar treated for 8 days at 60°C in chloride-containing pore solution; and (c) hexane-treated rebar. (The BE scale has been corrected by reference to a hydrocarbon standard.)**

droxide dominate (BE ≅ 530 eV), although the concentration of oxide oxygen remains at a significant level. Additional oxygen photopeaks in the curve-resolved spectrum are attributed to oxygen in silicon-containing species and in adsorbed water.

In the discussion of the analysis for rebar treated in the inhibitor solutions, results obtained for hexane-cleaned rebar and rebar immersed in chloride-containing pore solution are compared. Following that, the results for material from initial and delayed inhibition experiments are compared.

## Sodium Nitrite

The surface analysis results following the treatment of rebar with sodium nitrite in pore solution containing chloride are presented in Table 2. The findings for the initial inhibition experiments indicate an increased surface content only for carbon. The surface concentrations of oxygen and calcium decrease, whereas the surface contents of nitrogen, iron, silicon, and sodium remain essentially unchanged.

Of particular interest are the results for nitrogen. The percent nitrogen and the N 1s BE of 399.0 eV for treated rebar are similar (within experimental error) to the results found for rebar treated in pore solution containing chloride. That a nitrogen-containing species with a binding energy not characteristic of nitrite (N 1s BE in sodium nitrite = 404.1 eV) is detected indicates that nitrite is not chemisorbed on the rebar surface.

Alterations in the oxygen photopeak indicate that a chemical change has taken place on the rebar surface as a result of immersion in nitrite-containing pore solution. Thus, any nitrogen-containing reaction product must be released into solution, or it is the adsorbed nitrogen exhibiting a binding energy at 399.0 eV. The decrease in the oxygen concentration may at first appear surprising in view of the fact that nitrite appears to alter the surface chemistry of rebar. Nevertheless, the change that occurs is an alteration in the distribution of oxygen surface groups.

The oxygen 1s photopeak was curve-resolved (Figure 1a) into contributions from oxide oxygen (BE = 529 to 530 eV); hydroxide oxygen associated with metals (BE = 530 to 531 eV); oxide oxygen for alkali and alkaline earth metal compounds and silicon oxide species (BE = 531 to 532 eV); and adsorbed water (BE = 532 to 533 eV). As a result of the treatment, the $OH^-$ oxygen (BE = 530 eV) decreases, whereas the relative percent for the transition metal oxide oxygen (BE = 529 eV) increases for the nitrite-treated sample (see Table 3). Oxygen associated with other functionalities remains unchanged. These findings can be interpreted to suggest that the probable role of nitrite in the inhibition process is to increase oxide surface concentration. Because sodium nitrite is an effective corrosion inhibitor (*8, 9*), the surface analysis results suggest that one of the characteristics of useful inhibitors would be to increase the surface concentration of metal oxide functionality, especially iron oxide content (see also Dressman et al., a companion paper in this Record).

The results for delayed inhibition samples indicate little or no change (within the error limits) in elemental composition compared to the results for initial inhibition materials. The principal alteration is the increase in the concentration of metal oxide oxygen from 23.6 to 35.5 percent. It is likely that the increase is associated with the formation of additional passive iron oxide at the surface. This interpretation is similar to that presented earlier for other characterization studies (*8, 9*). In screening tests (Dressman et al.), sodium nitrite was an effective corrosion inhibitor.

## Sodium Molybdate

The interaction of sodium molybdate with rebar either by initial or delayed inhibition experiments produced an increased surface concentration of iron and associated iron oxide. Molybdenum was detected on the rebar surface at a concentration of about 0.5 at.% and a corresponding increase in the oxygen associated with Mo(VI) was noted. A comparison of the Mo $3d_{5/2}$ binding energy for $Na_2MoO_4$ (BE = 232.4 eV) with that for molybdenum from the two rebar exposure experiments (BE = 232.2 and 232.3 eV for initial and delayed inhibition experiments, respectively) indicates that molybdenum as molybdate [Mo(VI)] is adsorbed on the rebar surface.

A decrease in calcium and an increase in silicon concentrations were found, compared to the data obtained for rebar exposed only to chloride-containing pore solution. Within experimental error, the concentration of other elements did not change. The increase in iron and oxide oxygen and the detection of molybdenum as Mo(VI) indicate that molybdenum

TABLE 3 CURVE-RESOLVED O 1s RESULTS FOR REBAR AND REBAR IMMERSED IN PORE SOLUTIONS WITH AND WITHOUT INHIBITOR

| Sample Treatment | Metal Oxide BE=529-530 | Metal Hydroxide BE=530-531 | Silicon-oxygen oxy anion-oxygen BE=531-532 | adsorbed-oxygen (water) BE=532-533 |
|---|---|---|---|---|
| hexane cleaned | 31.5 | 15.2 | 27.7 | 25.6 |
| pore solution (8 days) | 21.7 | 7.1 | 50.2 | 21.0 |
| pore solution + 3.5% NaCl (8 days) | 16.0 | 3.5 | 61.7 | 18.9 |
| pore solution + 3.5% NaCl (8 weeks) | 15.0 | 4.3 | 56.9 | 23.8 |
| initial inhibition* (ii) $NaNO_2$ | 23.6 | 3.6 | 45.6 | 27.2 |
| delayed inhibition* (di) $NaNO_2$ | 35.5 | 9.2 | 34.1 | 21.2 |
| ii; $Na_2MoO_4$ | 32.9 | <2. | 37.4 | 29.8 |
| di; $Na_2MoO_4$ | 35.9 | <2. | 43.6 | 20.6 |
| ii; $NaH_2PO_4$ | 38.8 | <2. | 42.8 | 18.4 |
| di; $NaH_2PO_4$ | 42.1 | <2. | 41.6 | 16.3 |
| ii; $Na_2PO_3F$ | 34.5 | <2. | 46.7 | 18.8 |
| di; $Na_2PO_3F$ | 31.2 | <2. | 48.8 | 20.0 |
| ii; $Na_2B_4O_7$ | 3.1 | 8.7 | 59.2 | 29.0 |
| di; $Na_2B_4O_7$ | 2.2 | 2.1 | 72.7 | 23.0 |

*Dequest oxygen spectra not curve resolved, only interpreted in a qualitative manner in the text.

affects the rebar surface through an adsorption process such that oxide constituent contributions are increased at the surface. If molybdate acts as an oxidizing agent, the surface analysis results cannot identify the reduced molybdenum product. The molybdenum photopeak was characteristic only of Mo(VI), i.e., no reduced molybdenum species were detected at the surface. Lack of detection of reduced molybdenum could occur if the reduced product is not adsorbed on the rebar surface or if the concentration of reduced molybdenum is too small to contribute significantly to the Mo 3d photoelectron signal. However, screening corrosion tests (Dressman et al.) indicate that molybdate is a relatively poor inhibitor.

### Sodium Dihydrogen Phosphate (DHP)

The interaction of sodium DHP with rebar surfaces results in little or no significant change in oxygen, nitrogen, silicon, or iron concentrations, whereas the concentrations of sodium and phosphorous increase dramatically in the initial inhibition experiments. In these experiments, the calcium content is reduced significantly compared to that found for rebar treated in chloride-containing pore solution. A comparison of the initial and delayed inhibition results indicates a significant increase in iron content

Accompanying the change in surface concentrations for oxygen in both initial and delayed exposures is an alteration in the surface distribution of oxygen species. The O 1s photopeak was resolved into three components (see Table 3) that are characteristic of metal oxide (BE = 529 to 530 eV), metal hydroxide and phosphate oxygen (BE = 531 to 532 eV); and adsorbed oxygen, probably water (BE = 532 to 533 eV). Compared with rebar treated in chloride-containing pore solution, the concentration of oxygen surface species for metal oxide increases following phosphate treatment. The concentration of oxygen attributable to $OH^-$ is insignificant (<2 at.%) following phosphate treatment. The dominant contribution to the O 1s photopeak in the range 531 to 532 eV is oxygen bound to phosphorus. The detection of phosphate phosphorus on the treated surface combined with the decrease in hydroxide group content may indicate a surface acid-base reaction as the process promoting rebar surface changes that relate to inhibition. The adsorption of phosphate may also aid in corrosion inhibition by passivating potentially active corrosion sites on the rebar surface. Screening tests (Dressman et al.) indicate that DHP is only a modest corrosion inhibitor.

### Sodium Monofluorophosphate (MFP)

Inhibition experiments carried out with MFP permitted comparison of the results with those found for DHP. The results are presented in Table 2. The principal differences are that the oxygen, iron, sodium, and phosphorus concentrations are lower on MFP-treated rebar. The results indicate that MFP interacts with rebar to a lesser extent than DHP does. The oxygen functionality distribution is also consistent with this finding in that metal oxide content and phosphate oxygen concentration are both lower for the rebar surface treated with MFP. The expected 1:1 phosphorus to fluorine atomic ratio for $PO_3F^{2-}$ is not found on the rebar surface. The P/F ratio for the initial inhibition samples is 2.6 and that for the

delayed inhibition samples is 1.6. That this ratio is not unity suggests the loss of fluorine for adsorbed phosphate inhibitor. A potential process to account for this observation is hydrolysis of MFP (partial or complete) at the rebar surface or in solution. Either partial or complete hydrolysis would liberate fluoride and phosphate into solution and might result in subsequent adsorption of fluoride and phosphate. The phosphorus 2p binding energy for phosphorus adsorbed on rebar from MFP is equivalent to that for phosphate ($PO_4^{3-}$) in phosphate salts. The equivalence of BE values is consistent with the proposed hydrolysis process. Corrosion screening tests (Dressman et al.) demonstrate that MFP is a good corrosion inhibitor.

### Sodium Tetraborate

The reaction of sodium tetraborate with rebar produced a unique result. The oxygen photopeak is characteristic of oxide oxygen from borate and the boron 1s BE value is equivalent to that for pure sodium tetraborate. Boron is detected at 5 and 6 percent on initial and delayed inhibition rebar specimens, respectively. No iron was detected (<0.1 at.%) in the measurement of the Fe 2p photoelectron spectra for borate-treated rebar. The fact that iron is not detected at the surface, whereas oxygen and boron photopeaks characteristic of borate are detected, suggests that borate reacts under the chosen experimental conditions to produce a coating on the rebar. The behavior of forming a coating on rebar is unlike the modes of interaction found for other inhibitors studied.

On comparison of initial versus delayed inhibition, the atomic concentrations vary as noted; sodium and chlorine increase, whereas calcium decreases. The increase in sodium is consistent with the increase in borate concentration and may indicate adsorption of sodium on the borate coating. The findings for borate treatment suggest that such a coating could function as a barrier layer on the rebar to inhibit chloride-induced corrosion. Sodium tetraborate exhibits good corrosion inhibition in screening tests (Dressman et al.).

### Dequest 2000

Dequest 2000 is a trialkylphosphate amine. There are no significant differences in atomic composition on comparing initial and delayed inhibition results. The important surface composition changes for Dequest-treated rebar compared to rebar treated in chloride-containing pore solution are increases in nitrogen, phosphorus, and sodium, and a decrease in calcium. The P/N surface ratio in these samples is 3:1—a result indicative of the presence of adsorbed Dequest active component—trialkylphosphate amine. The oxygen spectra (Table 3) indicate contributions from iron oxide, but the principal contribution is from the phosphate functional group. The phosphorus atomic composition (13 at.%) indicates significant adsorption on rebar samples.

### Dequest 2010

The adsorption of this phosphate material on rebar is noted by the appearance of phosphorus in the spectra for initial and delayed inhibition samples. The concentrations of the respective individual elements are equivalent when comparing initial and delayed treatments, except for iron and silicon. For these latter elements, the concentration is greater following the initial inhibition treatment. The oxygen photopeak could be resolved (Table 3) to indicate contributions from metal oxide and adsorbed oxygen (probably water). However, the principal contribution is from the phosphate oxygen species. The phosphorus 2p binding energy data are indicative of the adsorption of the phosphorus component without change in chemical nature, i.e., no measurable or detectable change in the oxidation state of phosphorus occurs. On the basis of the percent of surface phosphorus, the adsorption of Dequest 2010 is less favorable by at least a factor of four (on a mole percent basis) compared to the adsorption of Dequest 2000 on rebar specimens.

### Dequest 2054

The interaction of this ethylenediaminetetraalkylphosphate with rebar does not produce any significant differences in the surface composition when comparing initial and delayed tests. On the basis of the amount of phosphorus present, the quantity of this material present on rebar is at least less than half that for the active component in Dequest 2010. The oxygen functionality includes contributions principally from iron oxide and phosphate from the inhibitor.

The corrosion inhibition performance (Dressman et al.) of Dequest materials is not superior to that of the simple metal salts discussed earlier. The inhibiting activity of Dequest materials was greater than that for molybdate, but less than that for phosphate.

## SUMMARY

The mode of inhibitor interaction with rebar samples can be grouped into three classes on the basis of the surface analysis results:

1. Nitrite interacts, leading to the formation of an iron oxide surface, but the inhibitor itself is not adsorbed as nitrite.
2. Tetraborate interacts to form a coating on the rebar surface rendering substrate iron undetectable by surface sensitive analytical measurements.
3. Other inhibitors interact by adsorption on rebar, leading to enhancement of oxide oxygen surface functionalities. In some instances the oxide functionalities could be associated with iron oxide.

From the surface analysis data alone, no inhibitor is significantly the most effective in the delayed inhibition experiments. The findings are consistent with the known corrosion inhibition behavior of nitrite caused by the formation of an oxide surface layer. The surface analysis results suggest that the formation of iron oxide surface components may indicate a desired reaction for potentially active inhibitors. In addition, another potentially beneficial corrosion inhibition process has been revealed in studies of sodium tetraborate solutions in which a coating is produced on the rebar surface.

## ACKNOWLEDGMENTS

Thanks for support of this work are due to Frank Cromer, who helped in the surface analysis measurements; to the National Science Foundation and the Commonwealth of Virginia for surface analysis equipment grants; and to SHRP, Project C-103.

## REFERENCES

1. *Durability of Concrete Bridge Decks*. Report 5. Portland Cement Association, Skokie, Ill., and U.S. Bureau of Public Roads, 1969, 46 pp.
2. *Durability of Concrete Bridge Decks*. Final Report. Portland Cement Association, and U.S. Bureau of Public Roads, 1970, 33 pp.
3. *Corrosion of Steel in Concrete*. Report of Technical Committee 60-CSC RILEM, The International Union of Testing Research Laboratories for Materials Structures, (P. Schiessl, ed.), Chapman and Hall, London, 1988, 102 pp.
4. K. Tuutti. *Corrosion of Steel in Concrete*. Swedish Cement and Concrete Research Institute, Stockholm, 1982, pp. 17-101.
5. M. G. Fontana and N. D. Greene. *Corrosion Engineering*. McGraw-Hill, New York, 1967, pp. 198-204.
6. H. H. Uhlig. *Corrosion and Corrosion Control*. John Wiley, New York, 1971, pp. 257-271.
7. J. M. West. *Electrodeposition and Corrosion Processes*. Van Nostrand-Reinhold, New York, 1971, pp. 138-148.
8. J. T. Lundquist, A. M. Rosenberg, and J. M. Gaidis. Calcium Nitrite as an Inhibitor of Rebar Corrosion in Chloride Containing Concrete. *Materials Performance*, Vol. 18, 1979, pp. 36-40.
9. A.M. Rosenberg and J. M. Gaidis. The Mechanism of Nitrite Inhibition of Chloride Attack on Reinforcing Steel in Alkaline Aqueous Environments. *Materials Performance*, Vol. 18, 1979, pp. 45-48.
10. B. B. Hope and A. K. C. Ip. Corrosion Inhibitors for Use in Concrete. *ACI Materials Journal*, Vol 86, 1989, pp. 602-608.
11. N. S. Berke and P. Stark. Calcium Nitrite as an Inhibitor: Evaluating and Testing for Corrosion Resistance. *Concrete International: Design and Construction*, Vol. 7, 1985, pp. 42-47.
12. N. S. Berke. The Effects of Calcium Nitrite and Mix Design on the Corrosion Resistance of Steel in Concrete (Part I). *Proc., Corrosion 85*, Paper 273, National Association of Corrosion Engineers, Houston, Tex., 1985.
13. N. S. Berke. The Effects of Calcium Nitrite and Mix Design on the Corrosion Resistance of Steel in Concrete (Part 2, Long-Term Results). Corrosion of Metals in Concrete. *Proc., CORROSION/87: Symposium on Corrosion of Metals in Concrete*, National Association of Corrosion Engineers, Houston, Tex., 1987.
14. J. M. Lewis, C. E. Mason, and D. Brereton. Sodium Benzoate in Concrete. *Civil Engineering and Public Works Review*, London, Vol. 51, 1956, pp. 881-882.
15. K. Siegbahn, C. N. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S.-E. Karlsson, I. Lindgren, and B. Lindberg. *Electron Spectroscopy for Chemical Analysis (ESCA): Atomic, Molecular, and Solid State Structure Studies by Means of Electron Spectroscopy*. Almqvist and Wiksells, Uppsala, Sweden, 1967.
16. J. M. Epp and J. G. Dillard. Effect of Ion Bombardment on the Chemical Reactivity of Gallium Arsenide (100). *Chemistry of Materials*, Vol. 1, 1989, pp. 325-330.
17. *Practical Surface Analysis by Auger and X-Ray Photoelectron Spectroscopy*. (D. Briggs and M. P. Seah, eds.), John Wiley, New York, 1983.

---

*The opinions, findings, and conclusions are those of the authors and not necessarily those of the sponsoring agencies.*

*Publication of this paper sponsored by Committee on Corrosion.*

# Migration of Inhibitors in Aqueous Solution Through Concrete

J. G. DILLARD, J. O. GLANVILLE, T. OSIROFF, L. A. WEBSTER, AND R. E. WEYERS

A study of the migration of corrosion inhibitors in aqueous solution through concrete disks (of ≈ 6-mm thickness) has been undertaken. Concrete disk specimens were placed between glass vacuum flanges and aqueous solutions of salts and a vacuum applied on one side of the disk caused inhibitors to penetrate the disk. The contact times of the aqueous solution with the disks were varied, and the quantities of solution penetrating the disk were evaluated. The depth of solute migration was measured by surface-sensitive analysis (x-ray photoelectron spectroscopy, XPS) of selected portions of fractured disks. XPS measurements were used to establish whether chemical changes had occurred during solute migration and to determine whether differential migration of solute cation and anion had taken place. The implications of the migration of salts and inhibitors for bridge repair are discussed.

The effectiveness of an inhibitor in bridge decks is influenced by its deliverability to the concrete-rebar interface. The study of a related phenomenon, diffusion of chloride in concrete, has attracted attention as a part of the effort to understand the corrosion of steel reinforcing rods in concrete structures. Investigators have reported on chloride diffusion under various conditions to which concrete is exposed (*1–4*). Some studies of inhibitor diffusion through concrete have been reported. For example, nitrite ion has been found to diffuse downward from a bonded concrete overlay doped with calcium nitrite (*5*).

The migration of inhibitors in aqueous solutions is examined using sample test specimens designed to permit accelerated migration of inhibitor solution. The principal objectives of the work are to evaluate the rate of inhibitor migration, to determine whether any change in the chemical nature of the inhibitor occurred as a result of interaction with concrete, and to inquire whether the anion-cation concentration ratio was affected by migration through the concrete disks. If an inhibitor is effective in corrosion screening tests (see Dressman et al., a companion paper in this Record) and is also able to migrate under the present test conditions, then it is a good candidate for larger-scale testing. The use of thin concrete disks permitted acquiring information regarding the migration of inhibitors in a shorter time frame. The combination of short experiment times and sensitive surface analysis methods permitted rapid evaluation of inhibitor migration and determination of the chemical nature of the migrating solutes.

J. G. Dillard, J. O. Glanville, T. Osiroff, and L. A. Webster, Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0212. R. E. Weyers, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0105.

## EXPERIMENTAL

The concrete test specimens were sand-modified concrete disks. Concrete cylinders, 3.0 cm in diameter and approximately 10 cm long, were cast. Portland cement was used and the water-cement ratio was 0.47. The concrete mix contained fine sand at a volume equivalent to that had aggregate been used. The concrete was mixed in a small mixer. A vibration table was used to compact the concrete cylinders. The cylinders were cured in a humidity chamber for 7 days.

The cured concrete cylinders were sliced to form disks approximately 0.6 ± 0.1 cm thick. The edges of the disks were sealed using an epoxy resin, leaving an effective penetration diameter of 1.5 cm. Rubber O-rings were placed on the epoxy surfaces while the epoxy was still tacky to ensure good adhesion between the O-rings and the concrete surface. A schematic representation of the concrete disk specimen is shown in Figure 1. The disks were placed between two glass flanges and connected to a vacuum source.

The concrete disks were pumped on a vacuum line to detect leaks caused by any large voids or defects in the epoxy sealing or in the disks themselves. This pump-out process also provided conditioning and drying of the samples. After pumping on the vacuum line, the disks were attached to a vacuum flask and evacuated using a water aspirator. In this arrangement, one side of the disk was exposed to the vacuum and the other side was exposed to air at atmospheric pressure (see Figure 1). A measured quantity of solution was placed on the air side of the system and a rubber stopper was placed on the open tube to eliminate any solvent loss because of evaporation. Vacuum was maintained on the samples for various periods of time, and the fluid remaining in the reservoir was measured. When surface analyses were to be carried out, the disk was carefully removed from the apparatus, dried at 60°C for 15 to 20 min, and then analyzed using x-ray photoelectron spectroscopy (XPS).

Solutions that were studied as a part of the penetration of inhibitors through concrete were: 0.300 M RbCl; chloride-containing pore solution [composition NaOH (0.300 M) KOH (0.600 M), saturated with $Ca(OH)_2$] containing 3.5 weight percent (wt.%) NaCl; and chloride-containing pore solutions with inhibitor (0.300 M). In the preparation of the inhibitor solutions, sufficient inhibitor was added to pore solution to achieve a concentration of 0.300 M. The inhibitors of interest in this study were sodium tetraborate ($Na_2B_4O_7 \cdot 10H_2O$) and sodium monofluorophosphate ($Na_2PO_3F$) (MFP).

XPS analyses were performed on selected portions of the specimens using a PHI Perkin-Elmer 5300 electron spectrom-

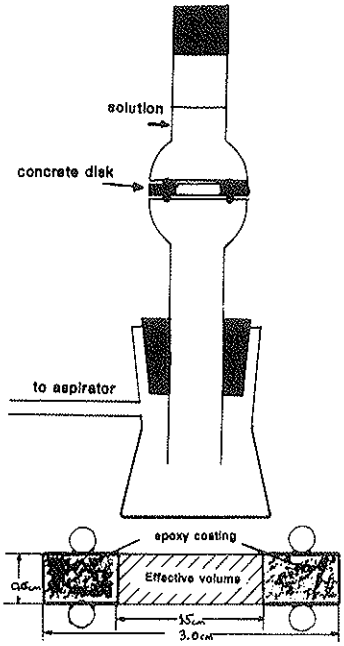



**FIGURE 1 Schematic diagram of the solution penetration-aspirator vacuum system. The vacuum could be maintained either with a water aspirator or a conventional vacuum pump. The shaded area of the disk is the region that was coated with epoxy resin. The nominal disk thickness was 0.6 cm. The effective volume exposed to the inhibitor-pore solution is 1.77 $cm^2$ × 0.6 cm.**

eter modified for small-spot measurements (*6*). Because of the porous nature of the concrete and the amount of water retained in the specimens (even after heating), the samples were maintained at liquid nitrogen temperature (−150°C) for the XPS analysis. The analysis by XPS allows identifying elements that originate from the inhibitor and determining their surface concentrations—thus, the penetration time of solute through the disk. Analysis of the specimens was carried out for upper and lower (vacuum side) portions of fractured disk samples. The spot size for the analysis was 1 × 3 mm. The results are presented in units of atomic percent (at.%) (*6*). The binding energy (BE) scale was calibrated by setting the C 1s hydrocarbon (contamination) photopeak binding energy at 285.0 eV (*7*). The BE data were used to determine the chemical nature of inhibitor elements.

## RESULTS AND DISCUSSION

The study of the migration of inhibitor solute species through concrete disks was investigated by (a) determining the volume of solution passing through the disks, and (b) analyzing the upper and lower disk surfaces after exposing the disks to the solution for designated periods of time. The results of the volume transport measurements are shown in Figure 2, in which the volume change is plotted versus the square root of time. In the presentation of the results, the volume of solution

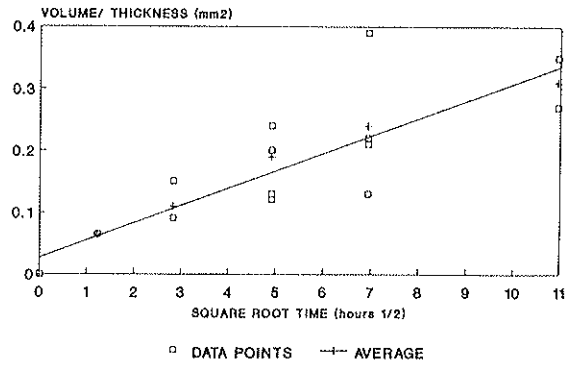



**FIGURE 2 Washburn plot of pore solution migration versus square root of time.**

migrating into the disks is normalized by dividing the volume transported by the thickness of the disk. In the figure, the data points correspond to the migration of solutions for rubidium chloride, sodium chloride in pore solution, and sodium tetraborate in chloride-containing pore solution. The data points are not distinguished in the figure because the volume change measured for each solution followed the same behavior. The volume change is a linear function of the square root of time. This finding is in agreement with the Washburn equation (*8*). The significant advantage of the present sample configuration is that the chemical nature of the diffusing species can be determined by subsequent surface analysis of fractured or whole concrete disks.

In the surface characterization measurements, the composition of a concrete (cement + sand) disk was determined to evaluate which elements, at which concentrations, were present on the disk surface, and thus which species could be studied without interference in the migration process. The average elemental composition from measurements on a representative group of disks is presented below. The results are atomic percents.

- C: 20.6,
- O: 56.1,
- Si: 13.9,
- Ca: 6.11,
- S: 1.24,
- K: 1.07,
- Na: .82, and
- Cl: <.2.

The surface chemistry is characterized by silicon, calcium, sodium, potassium, oxygen, sulfur, and carbon. Chlorine was detected at the detection level (<0.2 percent) in some samples. The chemical nature of the metals, inferred from BE measurements, corresponds to that expected for alkali (+1) and alkaline earth (+2) metals. Silicon is present as silicon-oxygen species and sulfur exists as sulfate in these specimens (*9*). The concentrations of inhibitor elements of interest are low or below the detection limits, and thus detection of inhibitor that had migrated through the disk was possible.

The results for the migration experiments involving aqueous solutions of RbCl are presented in Table 1. Solutions of RbCl were contacted with two disks of different thicknesses, 4.1 and 5.1 mm, for 48 and 120 hr, respectively. The concentrations of rubidium on the top and bottom portions of the disks

TABLE 1 XPS ANALYSIS FOR RbCl PENETRATION THROUGH CONCRETE: SOLUTION PENETRATION RESULTS FOR RbCl (0.300 M)

| Disk # | Thickness (mm) | Time (hr) | Surface | C | O | Si | Ca | Cl | Rb |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 4.1 | 48 | Top | 45.5 | 40.0 | 7.1 | 5.0 | 0.3 | 2.1 |
| | | | Bottom | 32.7 | 48.4 | 9.4 | 6.7 | 0.5 | 2.2 |
| 2 | 5.1 | 120 | Top | 30.3 | 49.3 | 12.1 | 5.0 | 0.4 | 2.1 |
| | | | Bottom | 46.1 | 37.5 | 7.2 | 2.9 | 0.2 | 3.8 |

are approximately equal, indicating that after each time period, the salt had migrated through the disk. Within the experimental error, the chlorine content is the same for top and bottom parts of the disks. However, the chlorine content is significantly less than the rubidium concentration. If equivalent migration had taken place, equal atomic concentrations of rubidium and chlorine should be detected. The migration of rubidium appears, from the present results, to occur more rapidly. To maintain electroneutrality, the migration of rubidium with another anion, as a cation-anion pair, could have taken place, leading to a lower chloride concentration. The identity of another potential anion is not revealed from the present XPS results.

Because the migration experiments involving inhibitors were conducted using chloride-containing pore solution, the migration of pore solution containing 3.5 wt.% NaCl was also investigated. The surface analysis data are presented in Table 2. The analytical results for sodium and chlorine reveal that the cation has migrated through the disk within 8 or 24 hr. The respective sodium concentrations on the top and bottom portions of the two disks are equal. The chloride concentration, either on the top or bottom of the specimen, is less than that for sodium. It appears that an anion exchange process must be taking place with species in the concrete specimen. Elements that could be associated with the cation, including halide ions, were not detected in the XPS measurements. Although sulfate sulfur was detected in the spectra, the concentration of sulfur does not change sufficiently or in a consistent manner such that it could be associated with cation migration. Migration of hydroxide ion along with the cation could take place and provide an explanation for the observed results. An examination of the O 1s photoelectron spectra does not reveal any significant increase in hydroxide oxygen. However, an increase associated with hydroxide migration would be difficult to distinguish because the concentration of oxygen from other chemical species, especially silicon-oxygen- and calcium-oxygen-containing entities, in the concrete is high.

TABLE 2 XPS ANALYSIS FOR CHLORIDE-CONTAINING PORE SOLUTION MIGRATION THROUGH CONCRETE: SOLUTION PENETRATION RESULTS FOR PORE SOLUTION + 3.5 wt.% NaCl

| Disk # | Thickness (mm) | Time (hr) | Surface | C | O | Si | Ca | Cl | Na |
|---|---|---|---|---|---|---|---|---|---|
| 7 | 5.5 | 8 | Top | 49.7 | 34.1 | 6.1 | 2.7 | 1.9 | 5.5 |
| | | | Bottom | 43.2 | 41.4 | 5.8 | 5.1 | 0.4 | 4.1 |
| 8 | 6.7 | 24 | Top | 46.3 | 35.0 | 7.2 | 4.6 | 2.1 | 4.8 |
| | | | Bottom | 52.9 | 34.0 | 4.9 | 4.4 | 0.4 | 3.3 |

The results for penetration into concrete by sodium borate in chloride-containing pore solution are presented in Table 3. In the table, surface analysis results are presented as a function of time for the upper and lower surfaces of the disks following penetration of the solute. Penetration of borate to the bottom of the disk occurs by at least 24 hr as evidenced by the large boron percent (6.1 at.%) for that sample. This percentage is to be compared with the value of 0.8 at.% boron for the bottom surface of the specimen examined after 1.5 hr. The chlorine and sodium concentrations increase for the 24-hr compared with the 1.5-hr sample. This finding is related to the migration of sodium chloride from the pore solution.

In addition, the silicon and calcium contents are smaller for those surfaces on which inhibitor is present. This result is attributed to the fact that small crystalline particles of inhibitor material were visible on the disk surface, thus blanketing silicon and calcium in concrete. Because the inhibitor solution was in contact with the upper surface of the disks for the duration of these experiments, it is surprising that the boron content for the upper surfaces is not constant. The values range from 5.4 at.% (48 hr) to 1.5 at.% (25 hr). The variation in the percent may be a result of the heterogeneity of concrete and the relatively small spots ($1 \times 3$ mm$^2$) that were analyzed on the disks.

The BE values for the B 1s level in the salt that appeared on the upper and lower disk surfaces, BE = 192.2 and 192.3 eV, respectively, were equivalent to the value, BE = 192.2 eV, measured for 0.300 M sodium borate pore solution frozen on the XPS sample probe. Thus, interaction of aqueous borate with concrete did not result in degradation or decomposition of the inhibitor. These results indicate also that introduction of borate inhibitor to the rebar-concrete interface should be possible.

The experiments with MFP yielded the surface analysis results presented in Table 4. Before considering the results,

TABLE 3 XPS ANALYSIS FOR SODIUM TETRABORATE MIGRATION THROUGH CONCRETE: SOLUTION PENETRATION RESULTS FOR PORE SOLUTION + 3.5 wt.% NaCl + 0.300 M $Na_2B_4O_7 \cdot 10H_2O$.

| Disk # | Thickness (mm) | Time (hr) | Surface | C | O | Si | Ca | Cl | B | Na |
|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 6.2 | 0.17 | Top | 38.4 | 41.3 | 4.8 | 5.7 | 0.3 | 3.0 | 6.5 |
| | | | Bottom | 44.7 | 39.1 | 7.6 | 6.5 | 0.2 | 0.66 | 1.2 |
| 10 | 5.6 | 0.5 | Top | 46.7 | 36.1 | 7.5 | 5.4 | 0.2 | 1.7 | 2.4 |
| | | | Bottom | 45.9 | 39.0 | 8.4 | 4.7 | 0.2 | 0.2 | 1.5 |
| 11 | 5.5 | 1.5 | Top | 36.5 | 43.1 | 6.4 | 6.1 | 0.2 | 3.6 | 4.2 |
| | | | Bottom | 35.6 | 43.9 | 9.9 | 5.8 | 0.2 | 0.8 | 3.7 |
| 12 | 6.3 | 25 | Top | 37.2 | 42.2 | 8.3 | 6.1 | 0.4 | 1.5 | 4.5 |
| | | | Bottom | 33.4 | 39.4 | 2.9 | 2.0 | 3.1 | 6.1 | 13.2 |
| 4 | 4.9 | 48 | Top | 32.7 | 45.0 | 6.3 | 5.3 | 0.4 | 5.4 | 5.0 |
| | | | Bottom | 39.9 | 37.1 | 1.0 | 0.9 | 3.6 | 5.8 | 11.7 |

TABLE 4 XPS ANALYSIS FOR SODIUM MFP MIGRATION THROUGH CONCRETE: SOLUTION PENETRATION RESULTS FOR PORE SOLUTION + 3.5 wt.% NaCl + 0.300 M MFP

| Disk # | Thickness (mm) | Time (hr) | Surface | C | O | Si | Ca | Na | Cl | F | P |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 17 | 7.2 | 8 | Top | 45.7 | 29.5 | 1.6 | 0.8 | 13.9 | 2.1 | 3.1 | 3.3 |
| | | | Bottom | 39.0 | 41.5 | 10.8 | 6.4 | 1.2 | 0.3 | 0.5 | 0.3 |
| 20 | 5.8 | 24 | Top | 34.4 | 43.4 | 6.2 | 5.4 | 5.2 | 0.2 | 0.9 | 4.3 |
| | | | Bottom | 34.5 | 45.6 | 10.4 | 5.8 | 2.9 | 0.2 | 0.2 | 0.4 |
| 21 | 7.7 | 48 | Top | 31.4 | 44.6 | 4.3 | 7.1 | 5.0 | 0.3 | 0.8 | 6.5 |
| | | | Bottom | 32.8 | 46.5 | 10.1 | 7.2 | 2.4 | 0.3 | 0.3 | 0.4 |
| 22 | 6.2 | 72 | Top | 27.5 | 46.9 | 4.0 | 7.9 | 5.0 | <0.2 | 1.3 | 7.4 |
| | | | Bottom | 41.6 | 42.7 | 8.0 | 5.7 | 1.1 | 0.4 | 0.3 | 0.2 |
| 23 | 7.2 | 114 | Top | 41.6 | 35.8 | 5.1 | 4.9 | 6.0 | 0.3 | 1.1 | 5.2 |
| | | | Bottom | 45.2 | 39.2 | 7.4 | 6.1 | 1.2 | 0.4 | 0.2 | 0.3 |

it is informative to recall the elements that exist in the inhibitor solution. The solution is composed of NaOH, NaCl, and $Na_2PO_3F$, [plus KOH and $Ca(OH)_2$], yielding respective total concentrations of Na, 1.500 M; Cl, 0.600 M; P, 0.300 M; and F, 0.300 M—or the atomic ratios of Na:Cl:F:P of 5:2:1:1. Analysis of the top surface of the disk having been in contact with MFP-pore solution for 8 hr exhibits ratios of Na:Cl:F:P of 5:0.75:1.1:1.2. This result is in good agreement with the expected ratios except that the chlorine content is significantly lower than expected. Examination of the results for longer exposure times revealed that the chlorine percent remained at a level of about 0.3 ± 0.1 at.%, a level near the detection limit for chlorine. It is reasonable to suggest that chlorine as chloride is adsorbed below the surface of the disk. Following 24 hr of exposure and up to 114 hr, the concentration of sodium at the upper disk surface remains at a concentration of 5 to 6 at.%, whereas the concentration on the lower surface is in the range 1 to 3 at.%. If it is recognized that sodium from NaCl and NaOH represents 60 at.% of the total sodium present, and if the initial sodium content for the 8-hr top surface is taken as representative of the concentration for sodium for solute that has not migrated sufficiently into the disk, then the expected concentration for sodium if it had penetrated the disk at longer times would exceed 8 at.%. That the sodium content on the lower disk surface is 1 to 3 at.%, indicates that sodium has not penetrated the disks in these experiments.

The concentration of phosphorus on the top disk surface remains at about 6 ± 1 at.%, whereas that on the lower portion of the disk is 0.3 ± 0.1 at.%. The concentration of fluorine does not change significantly on the lower disk surfaces. These data indicate that phosphorus and fluorine and thus MFP do not appear to migrate through the disks under the present experimental conditions.

The surface analytical results for MFP elements indicate that at early times the P:F ratio was the expected 1:1, within experimental error. On the other hand, the P:F ratio on the upper surface does not remain at 1:1 for exposure times beyond 8 hr. The P:F ratio is in the range of 4.7 to 8.1, with an average value of 5.8 for four samples. This result indicates that chemical alterations must be occurring for MFP as a result of its interaction in pore solution with concrete.

The photoelectron spectra in the F 1s region for frozen (−150°C) MFP (0.300 M) pore solution and for the top portions of disks following migration experiments are shown in Figure 3. For MFP dissolved in pore solution, only one fluorine photopeak is detected; with F 1s BE of 687.5 eV. This BE value is consistent with values measured for fluorine bonded to phosphorus in other fluorophosphates (*9*). The F 1s photoelectron spectrum for the top portion of a disk following MFP interaction with concrete for 8 hr reveals two fluorine peaks with BE values at 688.0 and 685.3 eV. The photopeak of higher BE value is associated with fluorine attached to phosphorus in MFP. The fluorine photopeak at the lower binding energy is assigned to fluoride as in metal fluorides such as sodium or as adsorbed fluoride (*9*). In the fluorine spectra for concrete specimens that had been exposed to MFP for longer time periods (24 and 72 hr, see Figure 3), only one F 1s peak is noted, with BE ≅ 685 eV. These findings support the notion that interaction of MFP with concrete leads to chemical changes in MFP and that the changes occur within 24 hr. The surface analysis results are consistent with the formation of fluoride and phosphate species in the hydrolysis reaction. Assuming that fluoride is produced, the surface analytical results, which show a greater concentration of phosphorus on the disk surface, indicate that fluoride is not as strongly adsorbed on the concrete surface as is phosphate or

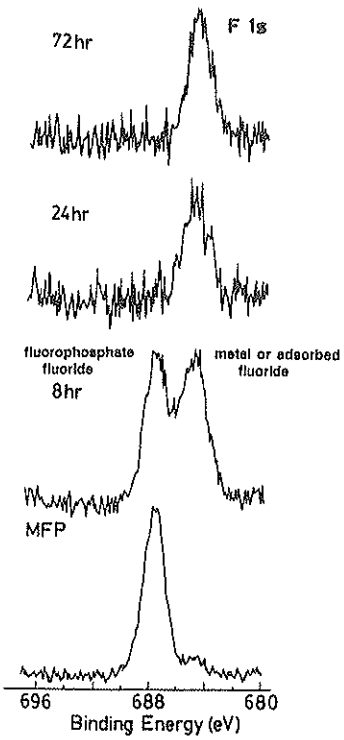



**FIGURE 3 F 1s photoelectron spectra, illustrating the change in fluorine chemistry that takes place as a result of interaction with the concrete disk.**

that fluoride diffuses into the concrete disk at a faster rate than phosphate. These findings demonstrate that MFP itself may not be an effective inhibitor, but that its hydrolysis products may be the active species delivered at the rebar-concrete interface.

## CONCLUSIONS

The migration of inhibitor constituents through concrete disks has been studied, and the migration varied with the square root of time. The migration time was determined by measuring the change in volume for solutions in contact with concrete disks while one side of the disk was maintained under vacuum. Analysis of the disk surfaces was used also to measure the time required to penetrate the disks. The chemical nature of the emerging solute was evaluated with respect to elemental concentrations and chemical form from BE measurements using photoelectron spectroscopy. The transport of sodium tetraborate through the disk resulted in no chemical change in the inhibitor. Migration of sodium MFP, however, led to hydrolysis of the salt. It is suggested that the hydrolysis of MFP produces fluoride ion and phosphate species in solution. The effect of these ions on the integrity and strength of concrete was not investigated.

## ACKNOWLEDGMENTS

Thanks are due to Frank Cromer, who helped in the surface analysis measurements; to the National Science Foundation and the Commonwealth of Virginia for surface analysis equipment grants; and to SHRP, Project C-103, for support of this work.

## REFERENCES

1. J. A. Stilwell. *Concrete in the Oceans.* Technical Report 8, 1983, p. 59.
2. K. Tuutti. Corrosion of Steel in Concrete. Swedish Cement and Concrete, Stockholm, 1982, p. 469.
3. R. D. Browne and M. P. Geighegan. *Proc. Symposium on Corrosion of Steel Reinforcement in Concrete Construction*, London, 1978, p. 79.
4. P. S. Mangat and K. Gurusamy. Chloride Diffusion in Steel Fibre Reinforced Marine Concrete. *Cement Concrete Research*, Vol. 17, 1987, pp. 385–396.
5. G. P. Jayaprakash, J. E. Bukovatz, K. Ramamurti, and W. J. Gilliland. Electro-Osmotic Techniques for Removal of Chloride from Concrete and for Emplacement of Concrete Sealants. Report FHWA-KS-82-2, Kansas Department of Transportation, Topeka, 1982.
6. J. M. Epp and J. G. Dillard. Effect of Ion Bombardment on the Chemical Reactivity of Gallium Arsenide(100). *Chemical Materials*, Vol. 1, 1989, pp. 325–330.
7. D. Briggs. In *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (D. Briggs and M. P. Seah, eds.), Wiley, New York, 1983, pp. 359–396.
8. E. W. Washburn. The Dynamics of Capillary Flow. *Physical Review*, Vol. 177, 1921, p. 273.
9. C. D. Wagner, L. H. Gale, and R. H. Raymond. Two-Dimensional Chemical State Plots: A Standardized Data Set for Use in Identifying Chemical States By X-Ray Photoelectron Spectroscopy. *Analytical Chemistry*, Vol. 51, pp. 466–482.

# DISCUSSION

RÉJEAN BEAUDOIN
*Domtar Research Centre, P.O. Box 300, Senneville, Québec, Canada H9X 3L7.*

The authors present the results of a study of the penetration of various ionic species in a simulated pore solution, through 7-day cured mortar. Sodium monofluorophosphate (MFP) was included in this study. The results failed to indicate any penetration of MFP through the mortar disks. It is known that MFP will hydrolyze in alkaline solution that contains an excess of calcium hydroxide. The experimental conditions used in this study were precisely alkaline solutions, saturated with calcium hydroxide, forced through freshly made mortar. These are the best conditions that cause hydrolysis of MFP.

The results presented in the paper clearly demonstrate that MFP did not penetrate the mortar under these conditions. Furthermore, the mortar prevented the penetration of sodium and chloride ions because of the blocking effect of the hydrolysis products of MFP. This property should have a positive effect on corrosion prevention of reinforcing steel in concrete. Unfortunately, the volume of solution that passed through the mortar disks for the MFP experiment is not reported.

However, the conditions of this study do not reflect the conditions encountered in aged concrete. In this case, many published studies have shown that the pore solution contains a limited amount of calcium hydroxide; the concrete might also be carbonated, reducing its alkalinity. Hydrolysis of MFP is therefore reduced or eliminated in old concrete. Partial drying of the concrete (naturally or artificially) also greatly enhances the penetration of an ionic species in solution. In field treatment of a structure with a corrosion inhibitor, neutral solutions should also be used, again diminishing the possibility of hydrolysis for MFP.

In many laboratory experiments, it has been demonstrated that MFP can easily penetrate through concrete at a depth of 3 cm or more. Multiple treatment with partial drying of concrete also enhances the penetration.

# AUTHORS' CLOSURE

We appreciate the comments of Mr. Beaudoin. In this investigation, it was our objective to study the interaction of solute species under chemical conditions similar to those in concrete—thus we used $Ca^{2+}$ saturated pore solution. The results demonstrate that monofluorophosphate (MFP) hydrolyzes under these experimental conditions. We have not studied nor have we determined or demonstrated specifically whether or not the hydrolysis product(s) or MFP itself produces a blocking effect. We have only demonstrated that under the experimental conditions, MFP hydrolyzes, producing at the concrete surface fluoride species attributed to fluoride associated with metal ions or adsorbed fluoride. The time variation of the XPS spectral features also demonstrates that, under our experimental conditions, the hydrolysis reaction is not instantaneous.

In addition, it is known that pore solution of aged concrete is a saturated calcium hydroxide solution along with potassium and sodium hydroxides. Also, as cement continues to hydrate,

calcium hydroxide is produced along with calcium silicate hydrates. Under normal field conditions, concrete in a temperate environment, such as the northeast portion of the United States, does not dry out to any significant degree below a depth of 10 mm. Concrete in the United States does not carbonate over a depth of 10 mm in 50 years because of the relatively low water-cement ratios that are used. Thus, hydrolysis of MFP in concrete under the field conditions described would occur.

*The opinions, findings, and conclusions reported herein are those of the authors and not necessarily those of the sponsoring agency.*

*Publication of this paper sponsored by Committee on Corrosion.*

# Screening Test for Rebar Corrosion Inhibitors

S. DRESSMAN, T. OSIROFF, J. G. DILLARD, J. O. GLANVILLE, AND R. E. WEYERS

The chloride-induced corrosion of steel reinforcing bars (rebars) in concrete structures can be vitiated by the presence of chemical compounds that serve as corrosion inhibitors. A simple, quick procedure has been developed to test for the relative effectiveness of inhibitors. The test is based on the observable corrosion of multiple test specimens under specified conditions. Details of the test procedure are reported along with the results for a range of inhibitors.

This paper is one of four related reports that describe the results of basic research studies aimed at answering the question of how corrosion inhibitors can be applied to mitigate bridge component reinforcing steel corrosion if the concrete remains in place during the rehabilitation process. This paper attempts to answer the question as to what corrosion inhibitors might work and how can they be tested in the laboratory. Other papers in this series deal with the questions of the mechanism of action of the corrosion inhibitors (Dillard et al., a companion paper in this Record), the mobility of corrosion inhibitors through concrete (Dillard et al., another companion paper), and electrochemical studies of inhibition phenomena (Webster et al., a companion paper in this Record). This last paper includes a comparison of the results of corrosion screening tests, surface analytical studies, and electrochemical studies of a series of inhibitors.

According to pessimistic estimates, as a consequence of their contamination by chloride salts, as many as one-half of all highway bridges in the United States (*1*) are deteriorating from reinforcing bar (rebar) corrosion. A similar situation prevails in the United Kingdom (*2*). Rebar corrosion in bridges is an outcome of the repeated wintertime application of deicing salts such as sodium chloride and calcium chloride. Reviews of the problem are available (*3,4*). Chloride-induced corrosion destroys the rebar and causes formation of corrosion products. The corrosion products occupy a greater physical volume than the rebar itself and cause internal expansion that leads to cracking and spalling of the concrete cover. Once cracking and spalling have occurred, the rebar is accessible to further chloride-induced corrosion—and so conditions deteriorate rapidly.

S. Dressman, T. Osiroff, J. G. Dillard, and J. O. Glanville, Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0212. R. E. Weyers, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0105.

Using corrosion-inhibiting agents is a traditional approach to preventing or slowing the corrosion of steel. Most corrosion inhibitors for steel find use in acidic or neutral conditions—where uninhibited attack may be rapid; corrosion of steel under alkaline conditions experienced by rebar in concrete is slow. However, because of the long life required of reinforced bridges and because corrosion is induced by chloride contamination, corrosion inhibitors are useful (*5*). At least two corrosion inhibitors are currently in use for the protection of rebar. The use of calcium nitrite as an admixture has become fairly widespread during the past decade, and concrete made with added calcium nitrite has considerable resistance to chloride-induced corrosion (*6,7*). A different approach to inhibition involves the use of a corrosion additive as a minor component of the road salt. For this purpose, sodium monofluorophosphate (MFP) has been used (Domtar Corp., Mississauga, Ontario).

Substances were sought that could be applied to existing, chloride-contaminated bridge components that would penetrate the concrete, and that would, on arrival at the rebar, stop or inhibit corrosion. The advantage of inhibitor substances is obvious: their use would allow the relatively inexpensive treatment of rebar corrosion in existing bridge components without the need for removal of concrete. Proving the efficacy of any proposed inhibitor is ultimately a task for field trials and practical testing. However, screening tests can be conducted under laboratory conditions to test (a) inhibitor effectiveness, and (b) inhibitor penetration through concrete. The use of a visual estimation method to judge inhibitor effectiveness is described in a later section.

Corrosion of rebar in chloride-contaminated concrete occurs where aqueous solutions within the pores of the concrete contact the rebar. The existence of such pore solutions is necessary to provide a conduit by which chloride ion may diffuse from the surface of the concrete to the rebars, several inches below the surface of the deck. In the work reported here, a synthetic pore solution (*8*) was used to simulate the practical situation.

In order to measure the relative effectiveness of inhibitors, a high concentration of sodium chloride was added to the simulated pore solution, and the corrosivity of the chloride-doped simulated pore solution was estimated by the amount of visual corrosion produced in multiple rebar specimens exposed under standardized conditions. The procedure used was modified from ASTM Standard G46–76, *Standard Practice for Examination and Evaluation of Pitting Corrosion*. The selection of materials to be tested was strongly guided by what was available commercially (*9*).

## OBJECTIVES

The main objective in undertaking this work was to develop a rapid, inexpensive screening test to study inhibiting agents for potential corrosion of rebar in concrete. Related studies have focused on the related questions of chemical surface changes (Dillard et al., a companion paper in this Record), of the relative mobility of inhibitors through concrete paste (Dillard et al., a companion paper in this Record), and on the electrochemical potential change during the process of corrosion and its inhibition (Webster et al., a companion paper in this Record). It was envisioned that candidate inhibitors identified by means of the test would be carried foward into a program of scaled-up testing and eventually taken to full-scale field testing.

A second objective was to use the test to conduct a survey of possible materials. As part of the investigation, numerous technical representatives of companies that offer corrosion inhibitors were contacted. Thus, the selection of candidate materials was based on the industrial state of the art.

## EXPERIMENTAL

Test rebar was obtained from a single heat produced by a local electric furnace from a scrap recycle source. Rebar was cleaned by rinsing with hexane (commercial solvent grade, of mixed isomers) but otherwise used in an as-received condition. Corrosion test pieces were prepared in the form of half-cylinders by longitudinally cutting approximately 1-in-long sections of rebar. Synthetic pore solution was prepared by weighing on a laboratory balance using reagent grade laboratory chemicals. The composition of the synthetic pore solution was chosen to to be sodium hydroxide (0.300 M), potassium hydroxide (0.600 M), and saturated calcium hydroxide in distilled water. Chloride-doped pore solution contained an additional 3.5 weight percent (wt.%) of sodium chloride. Before use, all solutions were air-saturated by passing air through them for 1½ to 2 hr.

Rebar test specimens were prepared in replicate (usually 5 or 10 replicates per test) by placing approximately 10 mL of the aerated test solution in a small plastic vial and putting a single rebar test piece into the solution. The rebar test pieces were completely submerged. The vials were loosely capped and placed in a laboratory oven and maintained at 60°C. Solutions were replenished periodically and replaced every 2 weeks.

From time to time, each test specimen was carefully examined with a 5× hand magnifying lens and graded on the basis of an estimate of the percentage of the surface corroded. Corrosion was taken to be any surface modification; no attempt was made to distinguish between corrosion of varying morphology, i.e., between area or pitting corrosion. The recorded corrosion value was then taken as the ratio of the altered surface to total surface, multiplied by 100, and reported as percent corrosion. The corrosion observations were made on the original (curved) surface of the rebar (not the freshly cut surface).

For experiments to test the comparative effect of inhibitors, accurately weighed quantities of the inhibitors were added to the chloride-doped simulated pore solution. Control solutions were chloride-doped pore solution without added corrosion inhibitor. Corrosion inhibitors were obtained as reagent grade chemicals from laboratory supply houses. Commercial products were obtained as manufacturers' samples. Their manufacturer and chemical nature are summarized in the following list.

- Monsanto Chemical Co.
  - –Dequest 2000, amino tris(methylene phosphonic acid), 50 percent active aqueous solution
  - –Dequest 2010, hydroxy-ethylidene diphosphonic acid, 60 percent active aqueous solution
  - –Dequest 2054, hexapotassium hexamethylene diamine (methylene tetraphosphonate), 35 percent active aqueous solution.
- Alox Chemical Co.
  - –Alox 901
  - –Alox 502A
  - –Alox 2291
  - –Alox 319F
  - –Alox 350
  - –Alox 2162
  - –Aqualox 2268
- Angus Chemical Co.
  - –Alkaterge T-IV, oxazoline compound
  - –Amine CS-1135, oxazoladine blend
- Miranol Inc.
  - –Miramine TOC, substituted imidazoline of tall oil fatty acid
  - –Monocor BE, borate ester
- Mona Industries Ltd.
  - –Monacor 39, imido ester carboxylic acid derivative
- Witco Corp.
  - –Witcamine PA 78-B, salt of fatty imidazoline
  - –Witcamine PA 60-B, slat of fatty imidazoline

In order to calculate molarities of commercial samples of proprietary composition and, hence, unknown molecular weight, a molecular weight of 250 g/mol was assumed.

## RESULTS AND DISCUSSION

### Preparation of Rebar Samples

Because corrosion effects are highly dependent on surface phenomena, it was recognized that considerable thought should be given to the way in which the test pieces were prepared. Various treatments with acids were investigated and x-ray photoectron spectroscopy (XPS) studies made. The surface analytical composition (obtained from surface chemical analysis with XPS) is reported in detail elsewhere (Dillard et al., a companion paper in this Record). The surface composition of the rebar is considerably different from the bulk composition, with high surface concentrations of copper and zinc reflecting the origin of the steel as recycled from shredded automobiles.

Cleaning with a hydrocarbon solvent (hexane) was the preparation method most likely to yield good results in accelerated, laboratory corrosion tests. However, cleaning with hexane does not significantly alter the surface composition. The

rebar in the aged bridge components to be protected by the method described here is likely to be virgin steel, not recycled steel. Thus, the question of surface catalytic effects on corrosion is open. The results reported are valid on the basis of the assumption that the steel used yields a fair relative assessment of the various corrosion inhibitors.

During the studies, most of the corrosion occurred on the gridiron regions of the specimens, i.e., on the raised regions of the deformations formed into the rebar to enhance bonding between the concrete and the steel. Weighing the test specimens as a method of estimating corrosion was considered, but the possible improved accuracy was judged to not outweigh the considerable extra effort involved.

### Development of the Screening Test

At the beginning of this study, all of the basic test parameters were open to investigation. Preliminary experiments implied that testing would have to be conducted at the elevated temperature if results were to be achieved after a reasonably short time. The final choice of 60°C is sufficiently high for rapid testing, yet sufficiently low that the needs for solution replenishment and replacement are kept within manageable limits.

The baseline data by which the extent of corrosion in chloride-doped pore solution at 60°C was judged are presented below. Under the conditions finally selected for the screening tests, the controls suffer surface corrosion to the extent of about 1 percent a day. The results are based on 10 replicate specimens and the standard deviations of the average percent corrosion are indicated.

| *Time (days)* | *Percent Corrosion* |
|---|---|
| 35 | 26 ± 8 |
| 56 | 52 ± 22 |

Depite the fact that this test is based solely on visual estimates of the extent of corrosion, the estimated standard deviations indicate that with a sufficient number of replicate samples a good estimate of corrosion can be reliably obtained. Overall, a relative standard deviation of 20 to 40 percent for a wide range of inhibitors is found. On the basis of this criterion, the screening test may be judged to be reasonably successful.

Experiments using the known corrosion inhibitors sodium nitrite and sodium MFP were useful in choosing how long to conduct the tests, at what temperature to maintain the oven, and when to either replenish or change solutions.

During the preliminary phase of the investigation, a number of studies were made of the effect of specific inhibitor concentration on the extent of corrosion. Table 1 presents the effect of varying sodium nitrite concentration on the percent corrosion; Table 2 presents the effect of varying sodium molybdate concentration. Sodium nitrite is one of the best available inhibitors under the test conditions, and Table 1 indicates that it is effective even at concentrations as low as 0.00200 M. Conversely, sodium molybdate is one of the worst inhibitors under the test conditions. Corrosion in sodium molybdate (0.00200 M) environment is indistinguishable from that in the control solutions. However, Table 2 indicates that sodium molybdate has a modest inhibiting effect at high concentration.

On the basis of these and other concentration studies, an inhibitor concentration of 0.00200 M was used in the screening test. Other final parameters in the screening test are shown in the following list summarizing the recommended procedures.

| *Parameter* | *Specification* |
|---|---|
| Test solution | sodium hydroxide (0.300 M)<br>potassium hydroxide (0.600 M), saturated with calcium hydroxide<br>3.50 at.% sodium chloride, aerated for minimum of 90 min |
| Test specimen | Half-cylinder of rebar, 1 in. long<br>Rebar prepared by hexane cleaning |
| Test conditions | 60°C for 30 days with bi-weekly replacement of the test solution<br>One rebar test specimen submerged in 10 mL of test solution in a small vial |
| Samples | Minimum of five replicates |
| Evaluation | Visual estimation per ASTM G46–76. |

### Selection and Performance of the Corrosion Inhibitors

The known corrosion inhibitors for rebar in steel were among those first tested. As presented in Table 3, both sodium nitrite and sodium MFP rank high on the basis of the results obtained. These results are encouraging in that they demonstrate that the rapid screening test has practical value. Although by themselves they do not validate the test, had these inhibitors performed badly the test itself would be suspect.

A number of criteria were used in selecting materials to be tested. Known inhibitors in alkaline systems, such as sodium mestasilicate, were natural candidates. The phosphonic acid

TABLE 1 CORROSION INHIBITION BY SODIUM NITRITE

| Sodium nitrite concentration mol/L | Percent Corrosion | Number of replicates |
|---|---|---|
| 0.00200 | 5.4 ± 2 | 10 |
| 0.0100 | 6.2 ± 2 | 10 |
| 0.0500 | 4.8 ± 2 | 10 |
| 0.100 | 3.8 ± 1 | 10 |
| 0.500 | 3.4 ± 1 | 10 |
| control (at 28 days) | 21% | |

NOTE: Corrosion of rebar samples in chloride-doped pore solution at 60° C. Reported values are the average of 10 replicate samples at each concentration. Exposure time: 28 days.

TABLE 2 CORROSION INHIBITION BY SODIUM MOLYBDATE

| Sodium molybdate concentration mol/L | Percent Corrosion | Number of replicates |
|---|---|---|
| 0.00200 | 25 ± 14 | 10 |
| 0.0110 | 23 ± 9 | 10 |
| 0.0500 | 13 ± 7 | 10 |
| 0.100 | 11 ± 3 | 10 |
| 0.500 | 8 ± 2 | 10 |
| control (at 33 days) | 25% | |

NOTE: Corrosion of rebar samples in chloride-doped pore solution at 60° C. Reported values are the average of ten replicate samples at each concentration. Exposure time: 33 days.

salts (Dequests) were selected because of their high solubility in alkaline systems and potential ability to form surface films. Consulting a standard listing of commercially available inhibitors (*10*) provided many sources of up-to-date information, and industry experts were asked for their recommendations whenever possible.

Eventually, a wide range of potential inhibitors were tested. Details of the chemical composition (where known) and the sources of the inhibitors are described in the experimental section. The results of these studies are presented in Table 3. Because the test is relatively inexpensive to conduct, many commercial materials could be investigated in a relatively short period of time.

The mode of action of the various inhibitors can be studied using surface sensitive analytical techniques (Dillard et al., a companion paper in this Record). These studies show that sodium tetraborate forms a surface film on the rebar. There is also evidence that borate ion is more mobile than chloride in concrete (Dillard et al., a companion paper in this Record). Further studies are planned using the screening test. In particular, pursuing the inhibiting potential of different borates would be of interest.

TABLE 3 SUMMARY RESULTS OF INHIBITOR SCREENING TESTS

| Inhibitor | Exposure (days) | Percent Corrosion |
|---|---|---|
| Alox 901 | 34 | 4.8 |
| Sodium nitrite | 28 | 5.4 |
| Sodium monofluorophosphate | 28 | 5.4 |
| Aqualox 2268 | 28 | 5.6 |
| Sodium tetraborate | 28 | 5.6 |
| Alox 350 | 34 | 6.4 |
| Alox 2162 | 34 | 6.4 |
| Miramine TOC | 34 | 6.4 |
| Alox 600 | 34 | 6.8 |
| Monacor 39 | 34 | 7.2 |
| Sodium nitrate | 34 | 8.0 |
| Sodium silicate (1:3.22, $Na_2O$:$SiO_2$) | 34 | 9.8 |
| Sodium metasilicate | 34 | 10.0 |
| Sodium carbonate | 34 | 10.0 |
| Alox 502 A | 34 | 10.2 |
| Witco PA 78B | 34 | 10.4 |
| Alox 2291 | 34 | 10.6 |
| Witco PA 60B | 34 | 10.8 |
| Sodium dihydrogen phosphate | 28 | 12.6 |
| Amine CS-1135 | 34 | 12.8 |
| Alox 319F | 34 | 12.8 |
| Potassium dichromate | 34 | 14.0 |
| Dequest 2010 | 21 | 15.0 |
| Potassium nitrate | 34 | 15.0 |
| Dequest 2000 | 28 | 15.0 |
| Monocor BE | 34 | 15.4 |
| Calcium borate | 34 | 19.0 |
| Dequest 2054 | 21 | 20.8 |
| Calcium sulfate | 34 | 22.4 |
| Sodium molybdate | 33 | 25.0 |
| Control (no inhibitor) | 35 | 26.0 |

NOTE: Values reported are the averages of five replicate samples exposed to chloride-doped pore solution at 60° C for the number of days stated. The inhibitor concentration was 0.00200 M.

## CONCLUSIONS

A useful, rapid screening technique has been developed to test the effectiveness of corrosion inhibitors for rebar corrosion in steel. Known corrosion inhibitors, recommended for this use, such as sodium nitrite and sodium MFP, perform well in the test. These results tend to confirm the utility and validity of the test.

Under the conditions chosen, a visually estimated corrosion of greater than 9 percent is sufficient to exclude a particular inhibitor from further testing. Highly promising corrosion inhibitors are those that exhibit less than 7 percent corrosion under the test conditions. Further studies using the test should be made.

Many organic commercial inhibitors also perform well, as do certain borate salts. Tests of the practical efficacy of borate salts in the treatment of chloride-contaminated bridge decks where the chloride-contaminated concrete remains in place are warranted because borate salts are inexpensive. Certain of the organic corrosion inhibitors may find use in treating the rebar in situations where bridge deck rehabilitation involves concrete removal and exposure of the rebar.

## ACKNOWLEDGMENTS

The authors thank Frank Cromer for his assistance with surface analysis measurements. The work reported here was supported under a grant from the Strategic Highway Research Program Project C–103.

## REFERENCES

1. *Seventh Annual Report to Congress—Highway Bridge Replacement and Rehabilitation Program.* Office of Engineering, Bridge Division, FHWA, U.S. Department of Transportation, 1986, 59 pp.
2. *The Performance of Concrete Bridges—A Survey of 200 Bridges.* Her Majesty's Stationary Office, London, 1989.
3. Chloride Corrosion of Steel in Concrete. ASTM 629, (D. E. Tonini and S. W. Dean, eds.) ASTM, Philadelphia, Pa., 1977.
4. D. G. Manning and J. Ryell. *Durable Bridge Decks.* Ontario Ministry of Transportation and Communications, Toronto, April 1976.
5. N. E. Hammer. Inhibitors for use on Reinforcing Steel in Concrete. *Corrosion Inhibitors* (C. C. Nathan, ed.), 1973, pp.190–195.
6. N. S. Berke. The Effects of Calcium Nitrite and Mix Design on the Corrosion Resistance of Steel in Concrete. Paper 132, presented at the Annual Meeting of the National Association of Corrosion Engineers, San Francisco, 1987.
7. Y. P. Virmani. *Time-to-Corrosion of Reinforcing Steel in Concrete Slabs, Volume VI: Calcium Nitrite Admixture.* FWHA-RD-88-165. FHWA, U.S. Department of Transportation, 1988.
8. A. Moragues, A. Macias, and C. Andrade. Equilibria of the Chemical Composition of the Concrete Pore Solution. *Cement and Concrete Research*, Vol. 17, 1986, pp. 173–182.
9. *McCutheon's Functional Materials: Corrosion Inhibitors.* Manufacturing Confectioner Publishing, Glen Rock, N.J., 1989.

---



*Publication of this paper sponsored by Committee on Corrosion.*

*Abridgment*

# Innovations Deserving Exploratory Analysis: Research on Structures

JOHN P. BROOMFIELD

Strategic Highway Research Program projects are described that have been carried out under the Innovations Deserving Exploratory Analysis (IDEA) program and that relate to structures research. These projects included the following examples of technology: laser-induced ultrasonic wave measurements, ac impedance spectroscopy, conducting polymer anode beds, and electroacoustic wave technology.

The Strategic Highway Research Program (SHRP) is a $150 million, 5-year effort to accelerate the development of new technology in specific areas of highway research. The structures area research, amounting to approximately $10 million, is concerned with evaluation and repair of reinforced concrete structures suffering from corrosion.

The contracts in this part of the program have been discussed elsewhere (*1*,*2*), and are described in this publication. The Innovations Deserving Exploratory Analysis (IDEA) program uses 2 percent of the budget to explore research related to the overall aims of SHRP but outside the scope of the main contracts. IDEA contracts are usually of 1-year duration and are funded in the $50,000 to $100,000 range. Because SHRP research is supported by the state highway agencies (SHAs), it must address state highway problems and produce useful products that SHAs can use. Projects are selected on the basis of their innovation, their potential for producing useable products, and anticipated benefits to the highway industry if the project is successful.

The projects discussed were all selected for their applicability to the problem of corrosion of steel in concrete. Some have been completed and reported on, some are still underway, and others have been approved for additional SHRP IDEA funding to advance the work from the research laboratory to the field testing of a prototype. This paper will address individual projects and their expected results.

## LASER-INDUCED ULTRASONICS FOR NONDESTRUCTIVE TESTING (ID002)

This contract was based on a proposal submitted by the National Nondestructive Test Centre, Harwell, United Kingdom. The work was started in September 1988 and was completed about 1 year later. The final report has been submitted (*3*), and will be published soon.

The goal was to develop a laser-induced ultrasonic wave measuring system for rapid, noncontact inspection of concrete and asphalt (including pavements) for detection of flaws in the structures.

The technique used for the laser ultrasonic testing is shown schematically in Figure 1. The system has a laser (Nd–YAG), which produces a 20-nsec pulse, at a power density of $2 \times 10^{11}$ W/m$^2$. This power vaporizes a tiny amount of the surface, generating a broadband, ultrasonic pulse. The pulse is detected by a low-power laser interferometer. The interferometer compares a beam of light reflected from the oscillating surface with a reference beam. The output of the interferometer is proportional to the phase change and can therefore be used to determine the movement of the surface.

All stimulation of the concrete and detection of pulses is by lasers. Therefore none of the signal generation or detection equipment is in physical contact with the specimen, i.e., this is a noncontact technique. In principle, it can rapidly scan bridge decks or substructures.

A complete description of the apparatus and results of the work can be found in the final report when available. Figure 2 shows a typical spectrum. The residual R wave is the surface wave between the point of creation of the pulse and the point of detection. The 2P wave is the compression wave reflected from the delamination, the 2S wave is the shear wave from the same delamination, and finally there is a 2P wave from the base of this all-asphalt specimen. Results indicated that the system could detect the bottom surface of a concrete block, and a delimination created in the block. The system also detected delaminations and debonding in aged asphalt. The main problem was difficulty in finding the peak representing a defect in the other noise in the spectrum.

The Harwell research project concluded that the system works in principle. Higher-powered lasers will be needed both for the interferometer and the laser that stimulates the signal. The work carried out for SHRP was conducted with available test equipment on a stable optical bench, which thereby differed from a vehicle-mounted system traveling along a roadway. The system may have more promise as a noncontact substructure nondestructive test (NDT) method. It could be mounted on a stable platform, and the laser fired at concrete surfaces to detect delaminations.

## AN ELECTROCHEMICAL TECHNIQUE (ID005)

Cortest of Columbus, Ohio, proposed and carried out this research using ac impedance spectroscopy to detect whether a cathodic protection system was effective in protecting all above-ground elements of a reinforced concrete structure suffering from chloride-induced corrosion. Figure 3 shows the

Strategic Highway Research Program, 818 Connecticut Avenue, N.W., Suite 400, Washington, D.C. 20006. Current address: HMA House, 78 Durham Road, Wimbledon, London SW20 0TL, England.

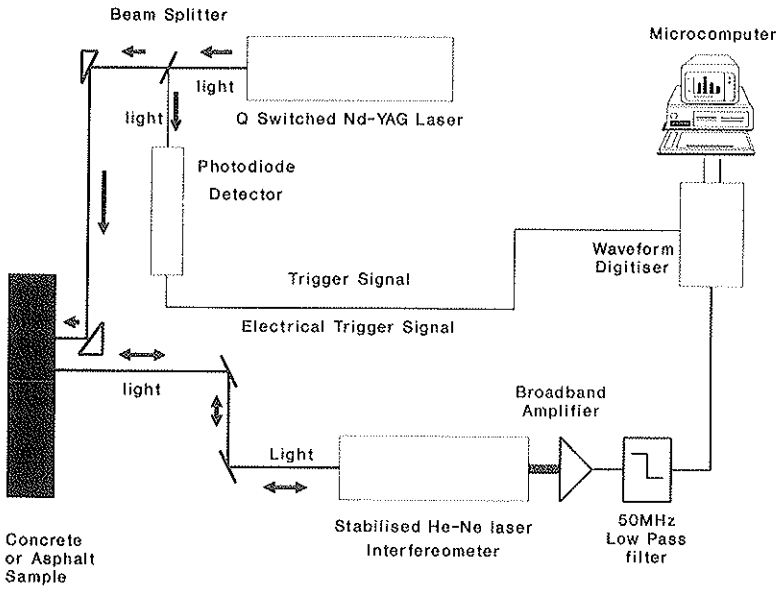



**FIGURE 1 Schematic layout of laser system for noncontact inspection of pavement.**

experimental set up. The corrosion cell can be represented by an equivalent electric circuit, and ac circuit theory can be applied. The response of the circuit can be represented by Nyquist plots in the complex plane. The shape of the plot should change as the equivalent electrical circuit switches from one representing a corroding cell to one representing a cathodically protected system.

Although there were definite differences between corroding and cathodically protected small specimens, these differences were not apparent on large specimens containing adjacent corroding and noncorroding areas, because macrocell or averaging effects severely reduced the differences. Further work may be funded to overcome the macrocell problem. Because the results of this work have implications for other SHRP projects, a workshop was held (*3*) to share the results of this research with others working in the field and with SHRP's researchers on Projects ID008 and C–101.

## ULTRALOW-FREQUENCY AC IMPEDANCE SPECTROSCOPY (ULFACIS) (ID008)

A proposal from SRI International was approved for funding. The objective was to explore the low-frequency end of the ac

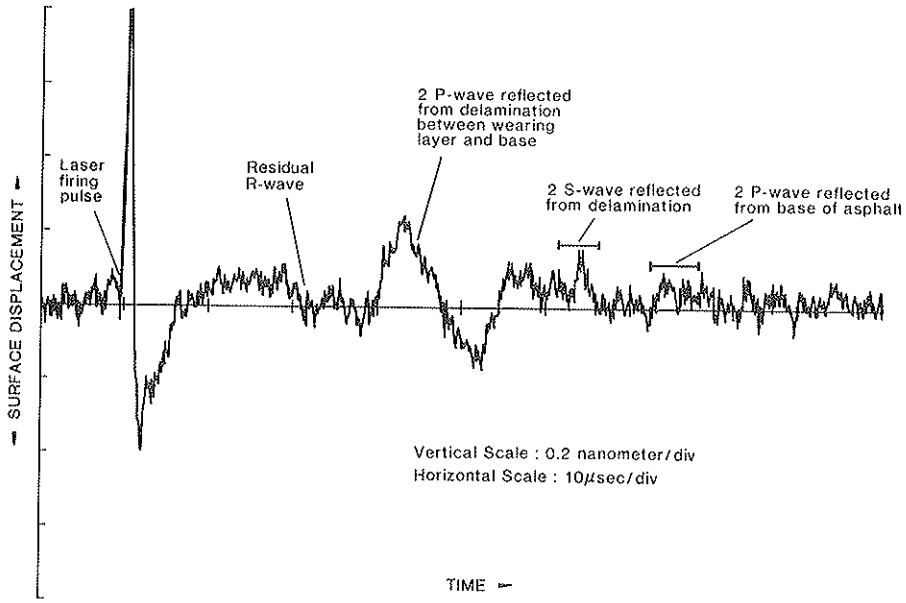



**FIGURE 2 Laser-induced ultrasonic waveform from aged asphalt. (Compression wave velocity = 3.0 mm/μsec; separation between lasers = 40 mm; position of lasers near visible delamination.)**

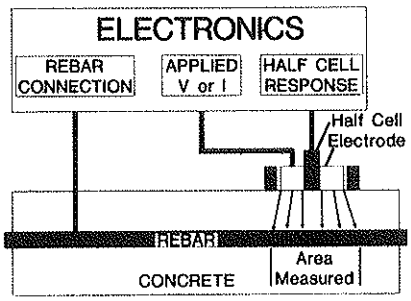



**FIGURE 3 Schematic experimental setup for ac impedance measurement.**

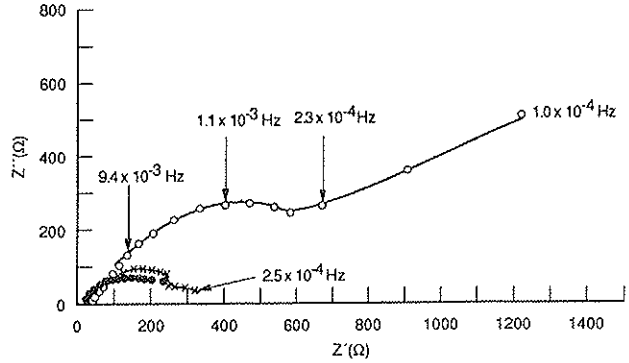



**FIGURE 4 Nyquist plots of noncorroding (open circles) and corroding (solid circles and crosses) (SRI).**

impedance spectrum to determine how accurately corrosion can be measured and located, given the slow response time of the reinforced-concrete system. The speed of response of the system determines the ac frequency range used, the time taken for a single measurement, and its susceptibility to error caused by rapid changes in local conditions (*3*).

The system of a corroding piece of reinforcing steel in moist concrete can be represented by an electrical circuit consisting of resistances, capacitances, and other electronic impedances that (unlike a simple resistance), change their values with the frequency of the alternating current, creating a phase lag in the circuit. This effect can be displayed as a spectrum of frequency shift and wave amplitude as the applied ac frequency varies. From these plots, usually complex plane plots of the impedances, the resistance of the electrolyte (the concrete) and the interfacial resistance at the steel-concrete interface can be determined. The latter is inversely proportional to the corrosion current, and therefore the corrosion rate can be calculated from Faraday's law. This law relates electrical charge passed with the amount of a metal consumed or deposited at an electrode.

Figure 4 shows examples of the spectra generated, with the frequencies at several points. Figure 5 shows how the spectra are generated. The corroding system can be regarded as an electrical circuit. Because of the presence of the capactance and the Warburg impedance (which represents diffusion control of the corrosion reaction), the circuit resistance appears to change with the ac frequency applied. The phase of the ac is also changed by the circuit. In ac circuit theory, this effect is often represented in a Nyquist plot of the real versus the complex (phase-related) resistance or impedance ($Z$ in the figure). Figure 5 shows two possible equivalent circuits and their corresponding Nyquist plots.

During the research, software for processing ac impedance spectra, locating corrosion, and determining corrosion rates was developed. A second stage of work is now underway, with Caltrans sharing the cost by providing manpower and sites. SRI International is developing prototype hardware for the system, and will be doing field tests on real structures.

## CONDUCTING POLYMER ANODE BEDS (ID014)

SHRP's research on electrochemical injection of inhibitors (*4*) showed that large, positively charged molecules can be diffused into concrete under the influence of an electric field. This IDEA project proposed by SRI is investigating the de-

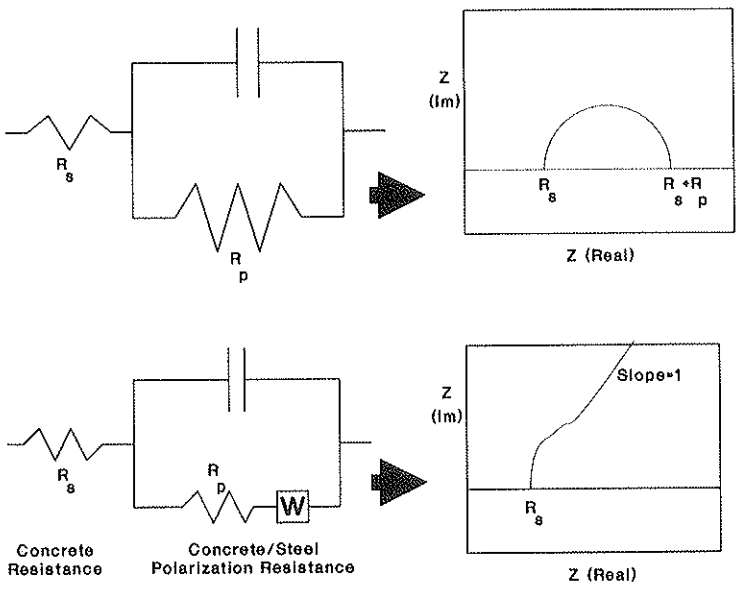



**FIGURE 5 Equivalent circuits and their ac impedances.**

velopment of an anode for cathodic protection, created by electrical injection of a charged monomer that is polymerized in situ to form a conductive layer in the concrete. It is important that a positively charged molecule is used, as the chloride ion is negatively charged, and would be attracted to the rebar by the field, thus risking accelerating corrosion.

This contract is still underway. Results so far have shown that the monomer can be injected electrically on small specimens. Work is underway to optimize the injection procedure, determine the conductivity of the finished product and to ensure that the injection can be controlled to avoid short circuits between the anode and the rebar. More realistic specimens will be used in the ongoing test program.

## ELECTROACOUSTIC TECHNOLOGY (ID019)

Combining an electric field with acoustic vibration has been found to move water around in slurries. Battelle Laboratories, Columbus, Ohio, proposed investigating the efficacy of this process not only in hardened but also in plastic concrete for controlling the water-cement ratio in the wet mix, and for dewatering and corrosion control of aging structures.

Results so far have indicted that the combination of acoustic vibration with an electric field is more effective than either alone in removing chloride ions. Results also suggest that the electroacoustic combination also will accelerate polymer impregnation and inorganic inhibitor injection. Further work will be undertaken to optimize the technology and to apply it to larger, more realistic concrete specimens.

## CONCLUSIONS

New technologies being investigated under the SHRP IDEA program may provide the bridge rehabilitation engineer with new tools for attacking the corrosion problems caused by salt ingress on bridge decks and substructures.

The technologies are innovative, and none are yet ready for field use. However, field tests will be conducted on the most promising techniques, and some products will be available by the completion of the program in 1993.

## ACKNOWLEDGMENT

The author wishes to thank SHRP and the National Research Council for permission to publish this paper. The author also would like to thank the researchers whose results have been quoted in this report.

## REFERENCES

1. *Concrete and Structures: Progress and Products Update*. Strategic Highway Research Program, National Research Council Washington D.C., Nov. 1989.
2. J. P. Broomfield. *SHRP Structures Research. Strategic Highway Research Program—Sharing the Benefits*, The Institute of Civil Engineers, London, Oct. 1990.
3. J. P. Broomfield. *Corrosion Workshop Taps Worldwide Expertise*. FOCUS, March 1990.
4. S. Hettiarachchi, A. T. Gaynor, and M. F. Asaro. *Electrochemical Injection of Synergistic Corrosion Inhibitors*. Final Report, Strategic Highway Research Program, National Research Council, Washington D.C., 1990.



*Publication of this paper sponsored by Committee on Corrosion.*

# Cathodic Protection of Prestressed Members: An Update

JOHN WAGNER, JR., WALTER T. YOUNG, AND SCOTT T. SCHEIRER

Research is being conducted in use of cathodic protection of highly stressed steel tendons, both prestressed and posttensioned, embedded in concrete bridge structures. These tendons are subject to embrittlement by hydrogen generated by cathodic protection under certain conditions. Research indicates that hydrogen penetrates steel and causes ductility reduction at potentials equal to or more negative than those normally considered necessary for the thermodynamic stability of iron. Other criteria commonly proposed for protection of steel reinforcement produce potentials considerably less negative than the potential required to generate hydrogen at the steel-concrete interface. Because of the critical dependence of hydrogen evolution on potential, cathodic protection must be potential controlled and the detection circuit must be free of IR effects. The use of freely corroding iron as a reference is suggested. Ongoing research includes studies on cathodic protection to answer questions about surface anode behavior, current distribution within bridge members, and use of cathodic protection on posttensioned and segmented bridge construction.

The deterioration of concrete bridge components is a major problem facing highway agencies in the United States. An important factor related to the deterioration of concrete is corrosion of embedded steel reinforcement. This corrosion results in cracking of the concrete, spalling, and eventually an actual weakening of the structure. In addition to reinforced decking, other structural members may include standard embedded reinforcement, concrete members with pretensioned or posttensioned high-strength steel tendons in embedded ducts, and segmented bridge construction.

When high-strength, pretensioned and posttensioned components are involved, there is a special problem. High-strength steels are subject to a phenomenon known as hydrogen embrittlement, which occurs when atomic hydrogen is released by electrochemical action at the steel interface with the environment. Ordinarily, in an alkaline environment such as cement mortar, this does not take place; however, once corrosion begins, loss of the alkaline environment permits direct release of hydrogen from the steel acid reaction into the steel matrix. Therefore, embrittlement can occur as a result of corrosion.

Hydrogen can also be formed at the steel-concrete interface through application of cathodic protection. Under certain conditions, electrolysis of moisture into atomic hydrogen can occur. The hydrogen can then enter the metal matrix and cause embrittlement. Embrittled high-strength steels exhibit reduced ductility and strength. Therefore, failure without warning caused by brittle fracture may occur in such members at strength levels well within design limits.

The FHWA Turner-Fairbank Research Center is sponsoring research into the use of cathodic protection for corrosion control of prestressed and posttensioned concrete structures. The objectives of this research are to define the effects of hydrogen on the high-strength steel used in bridge members and to define appropriate cathodic protection criteria. The research also includes studies to improve application and performance of surface applied anodes used in cathodic protection of concrete bridge members. Results that have been obtained to date in the hydrogen embrittlement studies are described.

PSG Corrosion Engineering, Inc., 610 Brandywine Parkway, West Chester, Pa. 19380.

## BACKGROUND

Details of the corrosion mechanism of steel in concrete and common methods used to protect steel are provided elsewhere in the literature (*1,2*). Parkins et al. (*3*) performed stress corrosion tests on notched and precracked prestressing steel strands in calcium hydroxide solutions varying in pH and chloride ion concentration using the constant extension rate (slow strain rate) technique (CERT). Enhanced cracking occurred at potentials lower than −0.900 volt [saturated calomel electrode (SCE)], irrespective of solution pH, and there was a second regime of cracking at potentials higher than −0.600 volts. In the intermediate region (between −0.900 and −0.600 volt), cracking was either absent or less severe. At potentials below −0.900 volt, the amount of quasi-cleavage observed on fracture surfaces increased when the potential became more negative and failure was caused by hydrogen embrittlement. On the other hand, failure in the second regime appeared to be dissolution related, as cracking occurred at potentials higher than the pitting potential. One of the few stress corrosion experiments incorporating an actual concrete environment was performed by Treadaway (*4*). He conducted tests on cold-drawn prestressing steel wires stressed to 70 to 80 percent of the tensile strength in concrete with and without calcium chloride. The specimens were exposed both outdoors in an industrial environment for periods from 2 to 27 months and in hot sodium hydroxide (0.02 N) solutions containing different concentrations of sodium chloride at pH 10.2 to 11.6. No evidence of stress corrosion cracking of the steel either in concrete or in the hot alkaline chloride solution was observed. Tensile testing of the steel wires subsequent to exposure resulted in ductile fracture. When corrosion had occurred, the fracture was initiated at a pit or an area where the wire suffered severe corrosion. However, as pointed out by Slater

(5), no applied potentials were used in these experiments, and the exposure time of 27 months in the concrete environment might have been too short to be conclusive.

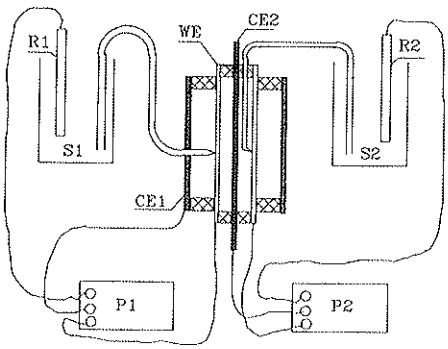
Recently, Scannell and Hartt (6) conducted a two-phase experiment to study the susceptibility of prestressing tendons to hydrogen embrittlement when cathodically overprotected. First, a three-point bending test capable of detecting hydrogen embrittlement was developed. In the second-phase experiment, a wrought, ferritic-pearlitic prestressing steel tendon was embedded in concrete and cathodically polarized to −1,300 mvolt (SCE) for 36 days. Specimens were then recovered from the concrete, notched, and tested by three-point bending. However, no evidence of hydrogen embrittlement was observed.

On the basis of the above discussion, a comprehensive research plan was developed to investigate the susceptibility of prestressing steel to hydrogen embrittlement under conditions applicable to cathodic protection in concrete. This involved a three-phase program consisting of a phase to characterize and measure hydrogen formation followed by a three part testing phase which included (a) CERT, (b) constant stress testing (CST), and (c) ripple effect testing (RET). The third phase consists of a program to evaluate the findings of the laboratory testing and cathodic protection criteria on full-scale prestressed members; this phase is now in progress.

## EXPERIMENTAL

### Hydrogen Formation Measurement

The hydrogen embrittlement studies were done in two phases. The first phase was designed to detect the potential at which hydrogen would be generated at the concrete-steel interface. The method used is described by McBreen (7). The method involves electrochemical detection of hydrogen that has passed through a steel membrane as the result of hydrogen production on the other side of the membrane. Figure 1 shows the equipment constructed for this study. It consisted of a carbon steel tube (WE) used as a membrane. The tube dimensions were 2.54 cm (1 in.) in diameter, 0.094 cm (0.037 in.) in wall thickness, and 14 cm (5.5 in.) in length. The wall thickness of the tube was on the order of the reduced section radius of the test specimen used in the mechanical testing phase—0.064 cm (0.025 in.). The interior of the tube was used as the working surface at which hydrogen would be produced. The exterior of the tube provided the surface for detection of hydrogen passing through the tube wall.




| | | | |
|---|---|---|---|
| S1 | 1N Sodium Hydroxide | S2 | Test Solution |
| R1 | Sat. Calomel Ref. | R2 | Sat. Calomel Ref. |
| WE | Working Electrode | CE2 | Counter Electrode 2 |
| CE1 | Counter Electrode 1 | P1 & P2 | Potentiostats |

**FIGURE 1 Hydrogen detection test.**

A platinum counter (CE2) electrode was inserted through rubber stoppers in the ends of the tube to permit electrochemical tests, and the tube was filled with test solution (S2) and lined with mortar. The entire working tube was then inserted through rubber stoppers into an outer steel tube (CE1) 5 cm (2 in.) in diameter, which served as a counter electrode and electrolyte chamber for the hydrogen detection cell. Deaerated sodium hydroxide (1 N) (S1) was used as the detector electrolyte. The potential of the inner surface of the tube was monitored by a reference SCE (R2) through a Lugin probe. The potential was controlled with a potentiostat (P2). The potential of the outer detector surface of the tube was monitored with a second reference SCE (R1). A second potentiostat (P1) held the potential of the outer surface of the steel tube (WE) at +100 mvolt with respect to reference R1

When hydrogen generated on the inside surface is absorbed by the steel and exits the outer surface, the outer surface potential changes, causing the detector potentiostat (P1) to change its output current in an effort to control the preset potential. Detector output current is a measure of the quantity and duration of hydrogen flow through the metal and was continuously measured during the test with a strip chart recorder.

The effects of different potentials such as might be encountered in the presence of corrosion products or carbonated concrete were evaluated. The detector was proven able to detect hydrogen generation at potential and pH values consistent with chemical thermodynamic theory. Tests to determine hydrogen evolution potential were run at pH values of 9.0 and 12.4, and with a tube lined with cement mortar. The effect of pulsed cathodic potentials, such as might be encountered during testing or erroneous operation, were tested using the hydrogen detector.

### Mechanical Testing for Hydrogen Embrittlement

The second phase consisted of a series of mechanical tests designed to measure the relative effects of hydrogen on prestressing steel when subjected to various test environments. These tests consisted of CERT, CST, and RET.

All specimens were machined from the central straight strand of 1.5 cm (0.600 in.) 1.86 × $10^9$ Pa (270 ksi) LOLAX prestressing wire conforming to ASTM A416-88B. The diameter of the central strand is 0.508 cm (0.200 in.). Notches were machined into the specimens using two configurations. In one configuration, a 90-degree notch with a 0.013-cm (0.005-in.) radius was used, and in the other a reduced section of 2.54-cm (1-in.) radius was used (smooth specimen configuration). The nominal diameter of the reduced section was 0.127 cm (0.050 in.). Some specimens were prepared so that the notches were encased in portland cement. The mortar used was either chloride free or contained 1 × $10^4$ ppm of chloride and completely encased the notch.

TABLE 1 CERT MATRIX

| | Cathodic Protection Potential Level (mV) (SCE) | | | | | | | Test Temperature (°C) | | | | NUMBER OF TESTS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| F.C. | -600 | -700 | -800 | -900 | -1000 | -1100 | -1200 | 0 | 20 | 40 | 60 | |
| Laboratory Air Environment: | | | | | | | | | | | | |
| X | | | | | | | | X | | | | 2 |
| X | | | | | | | | | X | | | 5 |
| X | | | | | | | | | | X | | 1 |
| X. | | | | | | | | | | | X | 2 |
| Ca (OH)2 Environment: | | | | | | | | | | | | |
| | | | X | | | | | X | | | | 1 |
| | | | | X | | | | X | | | | 1 |
| | | | | | X | | | X | | | | 1 |
| | | | | | | X | | X | | | | 2 |
| X | | | | | | | | | X | | | 4 |
| | | X | | | | | | | X | | | 2 |
| | | | X | | | | | | X | | | 2 |
| | | | | X | | | | | X | | | 2 |
| | | | | | X | | | | X | | | 3 |
| | | | | | | X | | | X | | | 2 |
| | | | | | | | X | | X | | | 3 |
| | | | X | | | | | | | | X | 1 |
| | | | | X | | | | | | | X | 1 |
| | | | | | X | | | | | | X | 1 |
| | | | | | | X | | | | | X | 1 |
| Cl-Free Cemented Notch: | | | | | | | | | | | | |
| X | | | | | | | | | X | | | 1 |
| | X | | | | | | | | X | | | 1 |
| | | | X | | | | | | X | | | 1 |
| | | | | X | | | | | X | | | 1 |
| | | | | | X | | | | X | | | 1 |
| | | | | | | | X | | X | | | 1 |
| $10^4$ ppm Cl-Cemented Notch: | | | | | | | | | | | | |
| X | | | | | | | | | X | | | 1 |
| | X | | | | | | | | X | | | 1 |
| | | | X | | | | | | X | | | 1 |
| | | | | X | | | | | X | | | 1 |
| | | | | | X | | | | X | | | 1 |
| | | | | | | | X | | X | | | 1 |
| $Na_2B_4O_7 \cdot 10H_2O$ Environment: | | | | | | | | | | | | |
| | X | | | | | | | | X | | | 3 |
| | | X | | | | | | | X | | | 2 |
| | | | X | | | | | | X | | | 2 |
| | | | | X | | | | | X | | | 2 |
| | | | | | X | | | | X | | | 2 |
| | X | | | | | | | | X | | | 3 |
| | | X | | | | | | | X | | | 2 |
| $CaSO_4$ Environment: | | | | | | | | | | | | |
| | | | X | | | | | | X | | | 3 |
| | | | | X | | | | | X | | | 2 |
| | | | | | X | | | | X | | | 1 |

### Slow Strain Rate Tests

The CERT method was used to determine the effects of temperature, pH, and cathodic protection level on the hydrogen embrittlement tendency of the prestressing wire. These tests were performed at temperatures of 0°, 20°, 40°, and 60°C in environments of laboratory air and air-saturated solutions of calcium hydroxide (pH 12.4), borax (pH 9.2), and calcium sulfate (pH 6.8). Cathodic protection potentials ranged from freely corroding to −1,200 mvolt (SCE). Additional tests, performed with a nitrogen purge of the solutions, were allowed to stabilize 4 days before testing. Table 1 presents the different conditions of the CERT test. Cathodic potentials were developed after the stabilization period using a custom-designed and custom-built potentiostat, a platinum-coated niobium counter electrode, and a reference SCE.

The SCE was placed in a separate vessel and joined to the test cell with a salt bridge and Lugin probe. Slow strain rate CERT tests were conducted at a preselected rate of $4.0 \times 10^{-6}$ cm ($1.5748 \times 10^{-6}$ in.) per second on an electrically powered CERT test machine.

Hartt et al. at Florida Atlantic University (FAU) also performed CERT tests as part of this project. One of the purposes of the FAU testing was to explore the degree of surface roughness that determines whether a specimen behaves as a smooth specimen or a notched specimen. Another objective was to investigate mechanical property recovery after excessive hydrogen generating potentials have been applied and released as might occur during testing or improper cathodic protection application. The tests were conducted using a constant extension rate of $4.7 \times 10^{-6}$ cm/sec ($1.85 \times 10^{-6}$ in./sec) in a nitrogen-purged saturated calcium hydroxide solution (pH 12.5). The effect of notch severity was studied by determining fracture stress as a function of notch geometry. The geometry considered was specimen diameter, notch diameter, notch angle (45 and 60 degrees) and notch root radius (0.001, 0.020, 0.25, 1, and 2 mm).

Specimens used to determine the effect of notch angle were charged at between −700 and −1,500 mvolt in 200-mvolt steps. Specimens used to evaluate the effect of root radius were charged at −900 and −1,300 mvolt (SCE). Notched specimens tested to determine the effect of precharging time

on embrittlement were charged at −900 mvolt (SCE) for 0 to 16 days. Smooth specimens used to evaluate the effect of rest time were precharged for 9 hr at −1,300 mvolt (SCE).

### Constant Strain Testing

CST was conducted to determine the effects of short- and long-term exposures to hydrogen-generating cathodic potentials. Both notched and smooth specimens were used. All tests were conducted in a saturated solution of calcium hydroxide at 20°C and all but one test were conducted with a nitrogen purge. The test specimen was strained to a predetermined load before testing. The specimen was exposed to the test solution 4 days before application of the cathodic potential of −1,200 mvolt. Table 2 presents the test matrix.

### Ripple Effect Testing

RET was performed to determine the relationship between the static-load environment cracking threshold and the corrosion fatigue threshold, and how this relationship is influenced by cathodic protection of the prestressing steel. All tests were performed in a 20°C calcium hydroxide solution.

Table 3 presents the test matrix for the RET. After test specimen stabilization, the specimen was loaded to the mean stress level and cycling commenced at a frequency of 1 Hz and stress amplitude of ±5 percent. RET was carried out on a computer-controlled servohydraulic mechanical test machine. Testing continued until the specimen failed or until a total of $10^6$ cycles accumulated.

## DISCUSSION OF RESULTS

### Hydrogen Formation Tests

There was a delay between the time the potential was applied above the hydrogen evolution potential and the time that the detector current began to change, indicating the presence of a depolarizing agent, i.e., hydrogen, on the surface. The delay time is the result of the time it takes the hydrogen to diffuse through the steel membrane. The detector current continued to increase after the hydrogen generating potential was removed and several hours were required before the current returned to initial values. The potential at which hydrogen flow from the cement-coated steel surface was first detected in the detector at pH 12.4 was −974 mvolt (SCE). No hydrogen generation was detected at −900 mv (SCE). Tests at the other pH levels indicated that hydrogen was generated on the steel tube surface and penetrated the tube wall at potential levels consistent with thermodynamic considerations.

There was no difference in the potential for hydrogen evolution between the pH 12.4 solution adjusted with calcium hydroxide and the pipe coated with cement mortar, except that the response time for the cement-coated steel was slower. This slowness might mean that the concentration of hydrogen entering the metal is lower in alkaline environments; however, it was not explored further.

TABLE 2 CST MATRIX FOR TESTS CONDUCTED IN 20°C $Ca(OH)_2$ SOLUTION AT CATHODIC PROTECTION POTENTIAL OF 1,200 mvolt (SCE)

| CEMENTED NOTCH | CONSTANT STRAIN 70% UTS | 70% NFS(1) | N2 PURGE | SOAK TIME (DAYS) | HOLD TIME AFTER PULSE (HOURS) | COMMENTS |
|---|---|---|---|---|---|---|
| | X | X | NO | 4 | 24 | |
| | X | | YES | 2 | 24 | |
| | | X | YES | 4 | 0 | SEE NOTE 2 |
| | | X | YES | 4 | 0 | SEE NOTE 3 |
| | | X | YES | 4 | 24 | SEE NOTE 4 |
| | | X | YES | 4 | 0 | SEE NOTE 5 |
| X | | X | YES | 4 | 0 | SEE NOTE 2 |
| X | | X | YES | 4 | 0 | SEE NOTE 3 |
| X | | X | YES | 4 | 24 | SEE NOTE 4 |
| X | | X | YES | 4 | 0 | SEE NOTE 5 |
| (6) | | X | YES | 4 | 0 | SEE NOTE 2 |
| (6) | | X | YES | 4 | 0 | SEE NOTE 3 |
| (6) | | X | YES | 4 | 24 | SEE NOTE 4 |
| (6) | | X | YES | 4 | 0 | SEE NOTE 5 |

NOTES:

1 - NOTCH FRACTURE STRESS BASED ON THE AVERAGE OF THE 20°C AIR TESTS RUN AT $1.5748 \times 10^{-6}$ INCHES/SECOND)

2 - AFTER SOAK PERIOD ESTABLISH -1200mV AND HOLD 3 TO 5 DAYS. AFTER HOLD PERIOD COMMENCE STRAINING AT $1.5748^{0}10^{-6}$ INCHES/SECOND.

3 - COMMENCE STRAINING AT $1.5748 \times 10^{-6}$ INCHES/SECOND AFTER SOAK PERIOD.

4 - COMMENCE STRAINING AT $1.5748 \times 10^{-6}$ INCHES/SECOND AFTER HOLD PERIOD.

5 - COMMENCE STRAINING AT $1.5748 \times 10^{-6}$ INCHES/SECOND IMMEDIATELY AFTER 60 SECOND CATHODIC PROTECTION PULSE.

6 - TEST PERFORMED WITH SMOOTH SPECIMEN CONFIGURATION.

TABLE 3 RET MATRIX

| CEMENTED NOTCH | CATHODIC PROTECTION LEVEL (mV) (SCE) | C.P. PULSES AT -1200 (mV) (SCE) | COMMENTS |
|---|---|---|---|
| | FREELY CORRODING | | |
| | -600 | | |
| | -1200 | YES | SEE NOTE 1 |
| | -600 | | |
| X | FREELY CORRODING | | |
| X | -600 | | |
| X | -1200 | | |
| X | -600 | YES | SEE NOTE 1 |

ALL TESTS CONDUCTED IN $N_2$ PURGED 20°C Ca $(OH)_2$ SOLUTION

CYCLIC FREQUENCY - 1 Hz

MEAN STRESS - 70% OF NOTCH FRACTURE STRESS FOR TEST CONDUCTED AT 20°C IN AIR

ALTERNATING STRESS - ±5% OF MEAN STRESS

CYCLE SPECIMEN TO FAILURE OR TOTAL CYCLE COUNT OF $10^6$

NOTES:

1 - SPECIMEN TO BE PULSED TO -1200 mV (SCE) CATHODIC PROTECTION LEVEL FOR 60 SECONDS EVERY 24 HOURS.

Pulsed-current flow at potential levels sufficient to generate hydrogen resulted in detection of hydrogen passing through the steel. A 60-sec pulse at −1,200 mvolt resulted in measurable hydrogen on the detector side of the steel membrane after about 50 min. The detector current at the steel membrane surface reached a peak after several hours before it began to decline towards initial values. This behavior indicates that even short-duration exposure to cathodic potentials of sufficient magnitude can produce hydrogen in the metal.

### Hydrogen Embrittlement Tests

#### *CERT*

Figures 2–4 show the results of the CERT testing. All normalized notch fracture stress calculations are based on specimen fracture load and the notch diameter at failure. Normalized notch fracture stress is the ratio of the notch fracture stress under the test conditions to the average notch fracture stress in air.

Figure 2 shows the CERT test results for the specimens in calcium hydroxide at 20°C. The figure plots normalized notch fracture stress against potential for plain specimens and specimens with cement encasement (with and without chloride in the cement). These tests indicate that as the potential becomes more negative than the hydrogen evolution potential, the normalized notch fracture stress decreases. There is approximately a 30 percent decrease in the normalized fracture stress over the range of potentials from about −950 to −1,200 mvolt (SCE). For the potential range of −300 through −950 mvolt (SCE), the normalized notch fracture stress remains relatively constant. This effect is consistent with previous research (*8*,*9*), which indicated that the effect is a result of hydrogen embrittlement.

Results indicate that the test results for 12.4 pH solution faithfully represent the behavior in cement-encapsulated specimens. There is no apparent additional effect of chloride contamination alone on the environmental cracking behavior of LOLAX steel wire. Corrosion pitting was not observed in the notch area of the cement-encased specimens containing chloride.

The temperature range selected for testing represents typical temperatures to which a bridge structure might be exposed. Data from the current work indicate that temperature has little effect on the normalized notch fracture stress through the potential levels investigated. All tests over the range of temperatures tested, 0°C to 60°C, exhibit similar behavior above and below the hydrogen evolution potential (see Figure 3).

Figure 4 shows the results of testing in pH environments of 12.4, 9.2, and 6.8 in a nitrogen-purged environment. Each plot indicates that there is a corresponding reduction in the normalized notch fracture stress when the potential is more negative than the hydrogen evolution point for the given solution. Hydrogen embrittlement can occur at increasingly positive potentials as the environmental pH decreases below that found in sound cement.

The FAU tests found a strengthening effect in the specimen diameter range 1.27 to 2 mm when comparing specimen diameter to fracture stress at the notch diameter. This was attributed to possible notch strengthening. No discernable change in fracture stress was observed for diameters greater than 2 mm. The notch diameter was held constant for all tests, meaning that those with specimen diameters in the affected range had a rather shallow notch. The unaffected specimens had deeper notches (greater percentage of the cross section). No discernable trend was observed when comparing either notch base diameter or notch radius with fracture stress. Specimens charged at −900 mvolt (SCE) for varying time periods indicated first an increase in fracture strength for short precharging times (less than 5 days), then a decrease in strength for longer charging times.

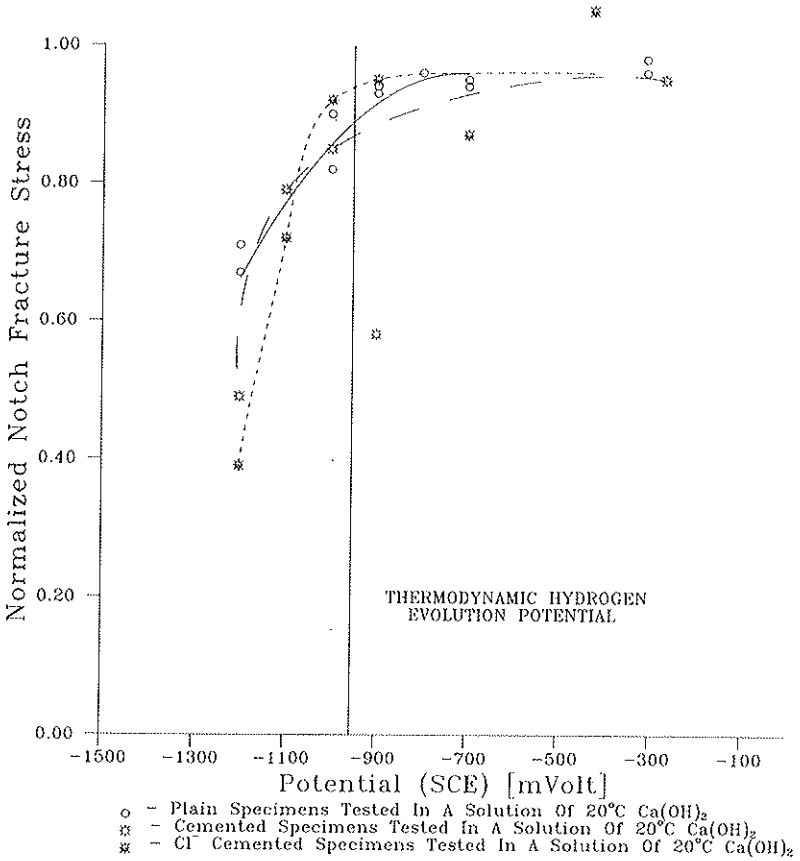



**FIGURE 2 Effects of cemented notch on CERT results.**

Specimens precharged for 9 hr at −1,300 mvolt, then allowed to rest for varying times before testing, displayed an increase in fracture load initially, followed by a return to lower fracture stress levels. The reduction in area for these specimens displayed a significant decrease for the first 50 hr, then a return to higher values.

*CST*

CST data indicate that short-term exposures of up to 60 sec to potentials more negative than the hydrogen evolution potential pose no danger to unnotched prestressing steel. Notched specimens did exhibit decreased normalized notch fracture stress when exposed for 2 hr to hydrogen charging. The notch in one of these specimens was encased in cement that might have prevented the hydrogen from escaping, resulting in a shorter time to failure than the nonencased specimen. The smooth specimen did not fail after a 5-day exposure to hydrogen charging and was strained to failure after the 5 days, with no reduction in normalized notch fracture stress. Some of the specimens exposed for 60 sec were held for 24 hr before straining, whereas other specimens were strained immediately. Notched specimens strained immediately exhibited a 20 percent reduction in normalized notch fracture stress, but the smooth specimens exhibited no noticeable reduction in normalized notch fracture stress. All specimens held for the 24-hr period following exposure did not exhibit a reduction in normalized notch fracture stress.

Cement-encased specimens did not exhibit a stress relaxation over the hold period. The unencased specimens did experience a stress relaxation over the hold period that lasted 2 to 5 days. One explanation might be that the bond between the specimen and the cement was sufficient for the cement to carry some of the load; however, the cement used was relatively weak and a strengthening effect of the cement appears unlikely. This effect was not investigated further.

*RET*

At the time of this paper, only the smooth specimen testing was completed. All of the tested specimens failed as a result of circumstances not related to the conditions. Failures occured because power line failures caused overload conditions. The specimens were able to withstand the imposed stresses without failure even while being charged at potential levels well above the hydrogen evolution point. On a few occasions, the hydrogen being generated in the test cell displaced the liquid in the salt bridge. When the salt bridge circuit was severed, the reference SCE probe was removed from the control circuit, resulting in an uncontrolled application of cathodic protection at a level more negative than the targeted

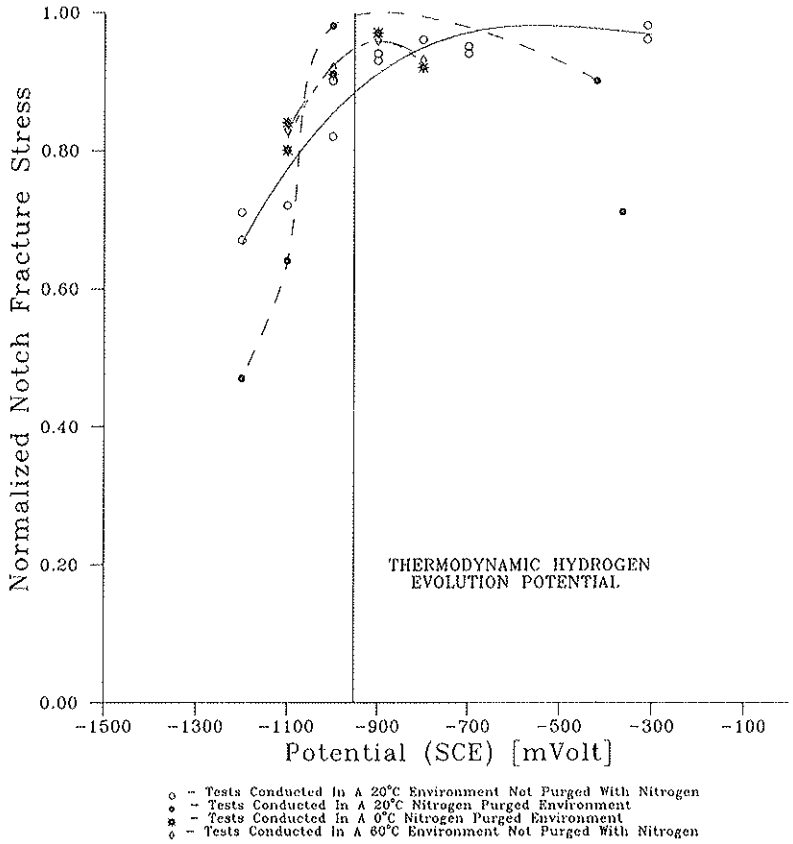



**FIGURE 3 Temperature effects on CERT results.**

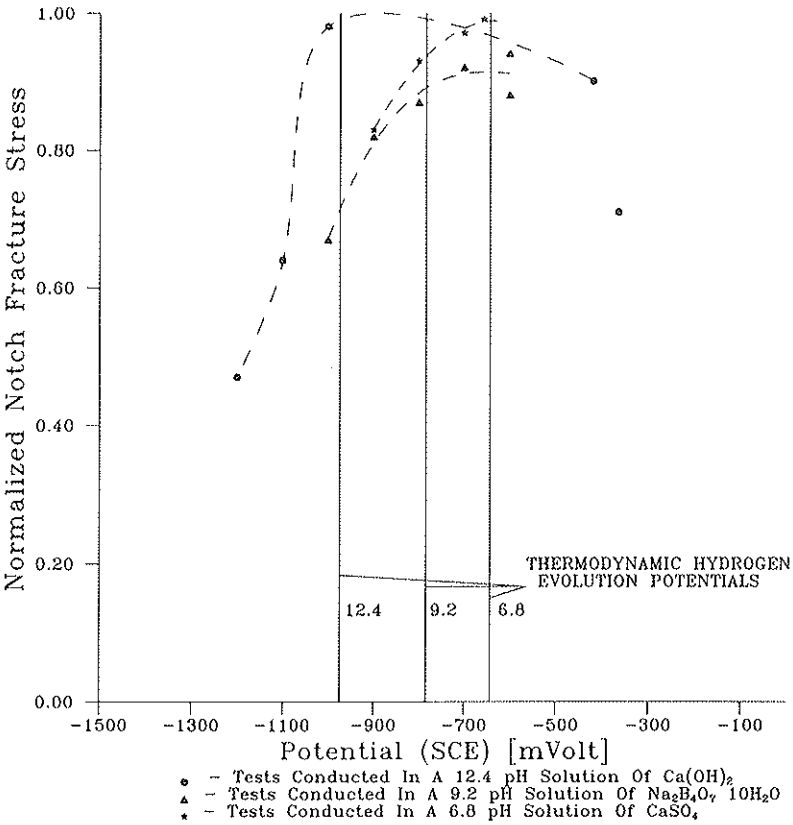



**FIGURE 4 Effects of pH and $N_2$ purging on CERT results.**

−1,200-mvolt value. No apparent decrease in cracking resistance was observed, even under these conditions.

The data indicate that there is no short-term reduction in the resistance of prestressing steel to environmental cracking as a result of small cyclic stresses and cathodic protection.

## INTERIM CONCLUSIONS

1. Hydrogen is generated at the surface of highly stressed steel members embedded in concrete at potential levels consistent with thermodynamic considerations. That is, the potential at which hydrogen is generated can be predicted from the pH of the environment through the use of Pourbaix diagrams. The potential at which hydrogen flow from the cement-coated steel surface was first detected in the detector at pH 12.4 was −974 mvolt (SCE). No hydrogen generation was detected at −900 mvolt (SCE).
2. Lower pH values surrounding steel embedded in concrete are possible in areas where carbonation has occurred, at cracks in the concrete, and beneath corrosion products. This effect means that hydrogen can be generated at potentials less negative than −974 mvolt (SCE). At a pH of 9.0, the potential for hydrogen generation is −750 mvolt (SCE). Lower pH results in embrittlement at less negative potentials than required in sound cement.
3. There was no difference in the potential for hydrogen evolution between the pH 12.4 adjusted with calcium hydroxide and the cement mortar coating, except that the response time for the cement-coated steel was slower.
4. Pulsed-current flow at potential levels capable of generating hydrogen was sufficient to produce detectable hydrogen passing through the steel. This effect indicates that even short-duration exposure to cathodic potentials of sufficient magnitude can produce hydrogen in the metal.
5. Cathodic potentials more negative than the hydrogen evolution potential sustained for durations greater than 2 hr will result in a reduction in the dynamic load-carrying capabilities of notched steel tendons. There was approximately a 30 percent decrease in the normalized notch fracture stress over the range of cathodic potential levels from about −950 to −1,200 mvolt (SCE). The normalized notch fracture stress in the range −300 to −950 mvolt (SCE) remained relatively constant. These potentials are consistent with the laboratory hydrogen detection experiments.
6. Smooth specimens did not exhibit increased susceptibility to environmentally assisted cracking (hydrogen embrittlement) when subjected to cyclic or static stresses. These specimens retained their machined finish throughout the test period; there was no pitting.
7. The notch severity necessary to cause an increase in cracking susceptibility is not evident from the data collected so far. The data are too scattered.
8. Temperatures in the range 0°C to 60°C and the presence of chloride ions alone do not significantly affect the material's susceptibility to environmental cracking from hydrogen embrittlement.
9. Effects of short-term exposures (less than 60 sec) were as follows:

- Potential levels more negative than the hydrogen evolution potential should not result in a reduction in the static load-carrying capability of unnotched prestressing tendons.
- Potential levels more negative than the hydrogen evolution potential result in a reduction in the dynamic load-carrying capability of severely notched prestressing steel tendons.
- Severely notched prestressing tendons exposed to short-term potential levels more negative than the hydrogen evolution potential should be held in a static stress condition for at least 24 hr, to permit some hydrogen to dissipate and to allow a restoration of ductility.

## WORK IN PROGRESS

The following work is being conducted:

1. Three different cathodic protection criteria are being evaluated on three full-sized prestressed beams. The 14- × 18-in. × 30-ft-long beams were specially fabricated and instrumented. The criteria are 100-mvolt depolarization, *E* Log *I*, and potential control to the most anodic potential [−700 mvolt (SCE)]. Each beam has salt-contaminated concrete in known areas of the prestressing wire and each of the strands can be included or excluded from the cathodic protection system. The beams will be destructively evaluated for the effectiveness of the criteria after the test period. Cathodic protection is being supplied by potentially controlled rectifiers feeding current to conductive paint anodes on one surface.
2. The three criteria are below the hydrogen evolution potential at the normal pH of the concrete encasement. A fourth beam will be overprotected at −1,200 mvolt (SCE). The beam will be destructively examined at the end of the test period. All four beams are located in a marine environment, are sprayed with seawater daily, and are being flexed in a bending mode at 1 Hz to a peak load of 4 kips.

## ACKNOWLEDGMENT

The authors would like to thank William Hartt and his associates at FAU, John Kulicki and Dennis Mertz of Modjeski and Masters for their part in this project, and Paul Virmani of the FHWA for providing the opportunity to present this work.

## REFERENCES

1. A. P. Crane. *Corrosion of Reinforcement in Concrete Construction.* John Wiley, New York, 1983.
2. *Corrosion of Rebars in Concrete.* (V. Chaker, ed.) STP 906, ASTM, Philadelphia, Pa., 1985.
3. R. N. Parkins, M. Elices, V. Sanchez-Galvez, and L. Caballero. Environment Sensitive Cracking of Prestressing Steels. *Corrosion Science,* Vol. 22, No. 5, 1982, p. 379.
4. K. W. J. Treadaway. Corrosion of Prestressed Steel Wire in Concrete. *British Corrosion Journal,* Vol. 6, March 1971, p. 66.
5. J. E. Slater. Roll of Cathodic Protection in Preventing Corrosion of Prestressing Steel in Concrete Structures. *Proc., Corrosion/83,* National Association of Corrosion Engineers, Paper 177, Anaheim, Calif., 1983, April 18–22.

6. W. T. Scannell and W. H. Hartt. Applicability of Cathodic Protection in Prevention of Corrosion Damage to Steel Tendons in Prestressed Concrete. Paper 127, *Proc., Corrosion/87,* National Association of Corrosion Engineers, San Francisco, March 9–13, 1987.
7. J. McBreen, L. Nanis, and W. Beck. A Method for Determination of the Permeation Rate of Hydrogen Through Metal Membranes. *Journal of Electrochemical Society.* Vol. 113, No. 11, Nov. 1966.
8. W. H. Hartt et al. Cathodic Protection and Environmental Cracking of Prestressing Steel. *Proc., Corrosion/89,* Paper 382, National Association of Corrosion Engineers, Houston, Tex., 1989.
9. W. H. Hartt, C. C. Kumria, and R. J. Kessler. Influence of Chlorides, pH and Precharging Upon Embrittlement of Cathodically Polarized Prestressing Steel. *Proc., Corrosion/90,* Paper 322, National Association of Corrosion Engineers, Houston, Tex., 1990.



*Publication of this paper sponsored by Committee on Corrosion.*

# Electrochemical Removal of Chloride Ions from Reinforced Concrete: Initial Evaluation of the Pier S19 Field Trial

D. G. MANNING AND F. PIANCA

The initial evaluation of a commercial process for the electrochemical removal of chloride ions from concrete is reported. A trial was conducted over an 8-week period in 1989 on a section of Pier S19 of the Burlington Skyway. The evaluation included visual examination, corrosion potential, and rate-of-corrosion measurements on the structure, and petrographic examination and measurement of chloride ion profiles on samples removed from the structure. Data up to 13 months after treatment of the concrete are reported. The treatment was successful in moving chloride ions away from the reinforcing steel and in removing a substantial proportion from the concrete without apparent damage to the concrete. The long-term effectiveness of the treatment is the major unknown factor.

The feasibility of using electrochemical techniques to remove chloride ions from concrete, and thereby arrest or reduce corrosion of embedded reinforcement, was investigated in two major studies in the 1970s (*1–3*). The process consists of applying an anode and an electrolyte to the surface of a reinforced concrete structure and passing current between the anode and the reinforcing steel, which acts as a cathode. The technique is similar to the application of cathodic protection but differs in two important respects: the anode is temporary, and the current density applied is approximately 100 times that used in most cathodic protection installations. However, if successful, the technique has the important advantage over cathodic protection in that it does not require regular monitoring and maintenance.

Although the studies performed by Battelle Columbus Laboratories (*1,2*) and the Kansas Department of Transportation (*3*) indicated that electrochemical migration of chloride ions was a promising rehabilitation technique, there was little further research in North America until the Strategic Highway Research Program (SHRP). This was largely because of concerns for undesirable effects on the concrete and steel such as an increase in permeability, reduction in bond of the reinforcement, migration of alkalis to the reinforcement, and the possibility of cracking because of high temperatures developed in the concrete. A project was defined within SHRP to examine these effects, determine the feasibility of electrochemical removal techniques, and, if feasible, develop procedures for use in the field (*4*). SHRP Contract C-102A, "Electrochemical Chloride Removal and Protection of Bridge Components" was awarded in May 1988 with a completion data of March 1993. Preliminary findings indicated that the technique is feasible (*5*).

Work in Europe has led to the development of an electrochemical technique for the treatment of carbonated concrete and the removal of chloride ions. The treatment is known commercially as the NORCURE™ process. The purpose of the current study was to evaluate this process. The study was funded by the Ontario Ministry of Transportation but, because it was the first application in North America and was closely related to the SHRP study, the SHRP contractor was invited to audit the application (*6*). The treatment of a section of Pier S19 at the Burlington Skyway in Ontario, during an 8-week period from July to September 1989, and the results of condition surveys up to October 1990 are described.

## FIELD PROCEDURES

### Test Site

The test site was the lower 4 m of the west column of Pier S19 of the Burlington Skyway. The plan dimensions of the pier were approximately 1.7 × 3.0 m and the area treated was approximately 30.8 m$^2$. The north face was left untreated to provide a control area. The column was relatively lightly reinforced, the surface area of reinforcement being approximately 0.55 m$^2$ per square meter of concrete surface for the east and west faces, and 0.79 m$^2$ per square meter of concrete surface for the north and south faces.

The piers of the Burlington Skyway were constructed in 1955. Roadway drainage, which contained salt in the winter months, flowing through open deck joints and over portions of the piers had caused corrosion of the reinforcement even though the cover was generally in excess of 75 mm. The concrete contained approximately 330 kg/m$^3$ of cement, a crushed dolomitic limestone coarse aggregate, and a natural sand fine aggregate. The water-cement ratio was about 0.45. The dolomitic limestone was high in chloride ion content, with the result that the background concentration of chloride ion in the concrete was measured to be 0.042 percent by mass (using the acid-soluble extraction procedure). If the threshold value for corrosion of embedded reinforcement is taken to be 0.20 percent by mass of cement, the threshold value for the concrete in Pier S19 can be calculated to be 0.070 percent by mass of concrete.

Before the electrochemical treatment, a condition survey made of the test site included

Research and Development Branch, Ministry of Transportation, 1201 Wilson Avenue, Downsview, Ontario, Canada M3M 1J8.

- A visual survey of the concrete surface,
- A delamination survey by sounding,
- Measurement of concrete cover,
- Verification of electrical continuity of the reinforcement,
- Measurement of corrosion potentials (in accordance with ASTM C876-87),
- Corrosion rate measurements (using a commercial 3LP device), and
- Removal of cores for petrographic examination and measurement of chloride ion content.

Four small graphite probes and two thermocouples were installed at the depth of the steel. A 220-volt, 30-A temporary electrical service was provided.

### Installation

A connection was made to the reinforcement and insulated to prevent corrosion. Wooden strips, approximately 40 mm wide by 20 mm thick, were fastened vertically to the concrete using insulated anchor screws. Cellulose fiber, moistened with lime water to act as an electrolyte, was sprayed on the concrete to the same thickness as the battens, as shown in Figure 1. A 100- × 100-mm steel mesh of 5-mm wire diameter was attached to the wooden strips with staples as shown in Figure 2, and a second layer of moist fiber was applied to cover the mesh to a thickness of about 25 mm.

An electrical connection was made to the mesh on each face and insulated to reduce corrosion. The wires from the mesh were connected to the positive terminal of the rectifier and the wire from the reinforcement to the negative terminal. The power was switched on, and after fluctuations during the first few days the current density was 0.77 A/m$^2$ of concrete surface.

### Treatment Parameters

The treatment was continued for a period of 8 weeks. The fiber was sprayed with water daily to minimize the circuit resistance. The rectifier output was limited to a nominal value of 40 V.

During treatment, the following measurements were made:

- The voltage and current at the rectifier were measured daily.
- The potentials of the embedded reference electrodes were measured periodically.
- The chloride concentration of the electrolyte was also measured periodically. Because some of the electrolyte drained

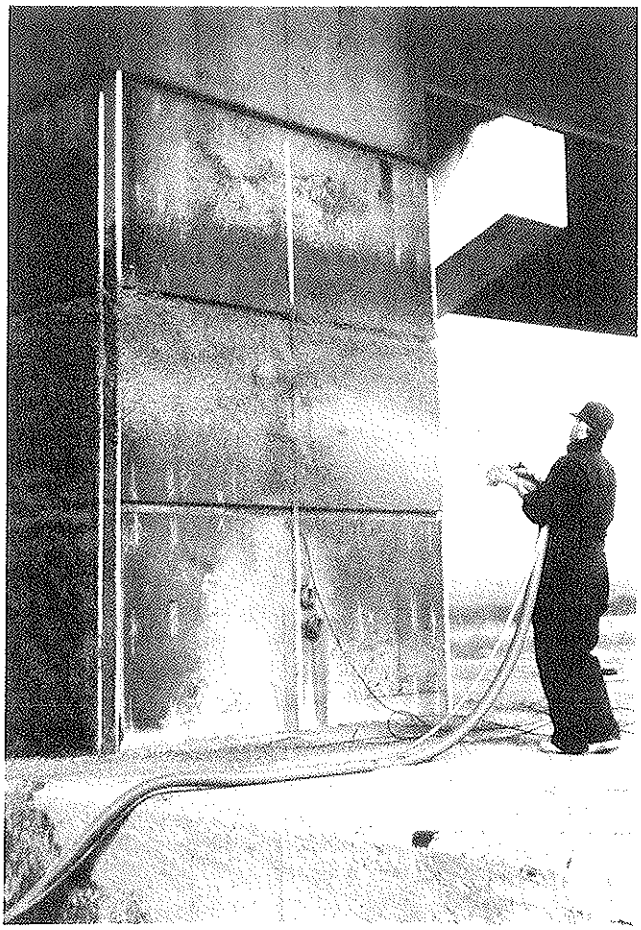


FIGURE 1 Spraying of cellulose fiber on concrete.

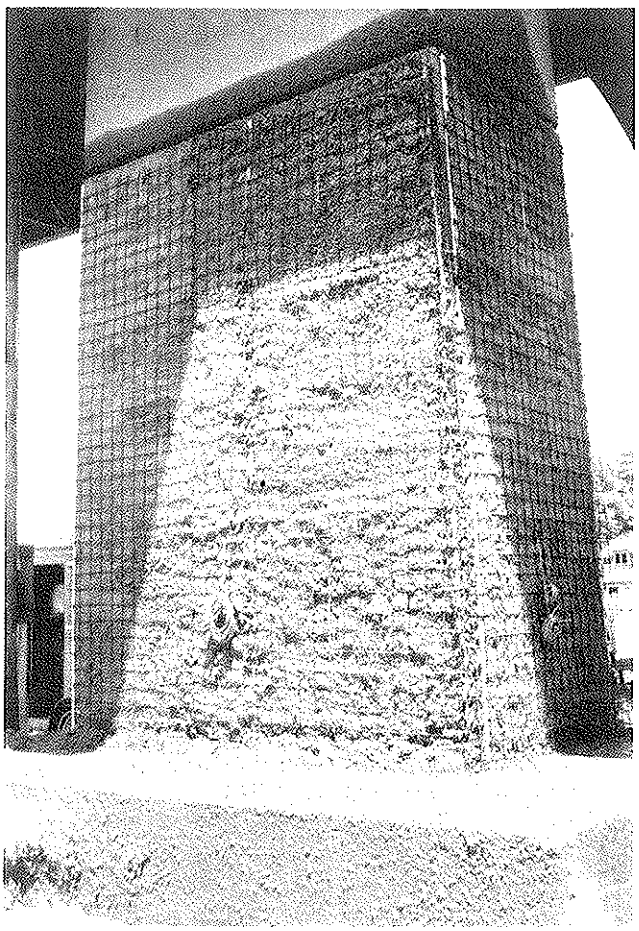


FIGURE 2 Covering of wire mesh with moist cellulose fiber.

away, these measurements were used only to provide an indication of the progress of the chloride removal.

- Cores were taken after 3, 7, and 8 weeks to establish the chloride ion profiles in the concrete.

After treatment, the following actions were taken:

- Potentials of the embedded reference electrodes were measured periodically.
- Surface corrosion potentials and rate-of-corrosion measurements were made 6 weeks, 6 months, and 13 months after treatment. Surface corrosion potentials were also measured 10 months after treatment.
- Cores were taken 6 weeks after treatment for petrographic examination.

## RESULTS

### Appearance

The cellulose was light brown in color and, immediately after installation, was unobtrusive. After 24 hr, a rust-stained outline of the mesh was visible, and within about 3 weeks the cellulose had a fairly uniform red-brown appearance as the result of corrosion of the mesh, as shown in Figure 3. The excess water that drained from the cellulose following the daily wetting also contained rust that stained the column plinth. After the 8 weeks of treatment, the surface of the column and parts of the plinth were heavily stained by rust. The staining was removed by blast cleaning.

### Power Requirements

The voltage and current throughout the treatment period are shown in Figure 4. The voltage and current fluctuated during the first 3 days. During the remainder of the 8 weeks, there was a slight increase in voltage, but a substantial decrease in current to about one-third its initial value. The total charge passed was approximately 610 A-hr/m² of concrete surface.

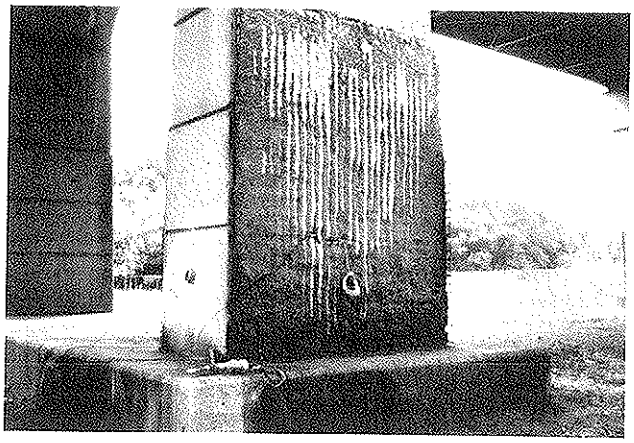


**FIGURE 3 Corrosion of the mesh.**

The power level used was not sufficient to heat the concrete. There was no measurable difference between thermocouples embedded at the level of the steel in the south and north faces.

### Chloride Ion Content

Typical chloride ion profiles at various stages in the treatment process are shown in Figure 5. The profiles were established by cutting cores into 10-mm-thick slices, pulverizing each slice, and measuring the acid-soluble chloride ion content. Because the flow of drainage water over the column was uneven, the initial chloride ion content of the concrete varied considerably over each face. In order to minimize these variations and permit a valid comparison of the progress of the desalination, cores were taken vertically below each other on each face. Two sets of cores were taken on the west face, one set directly over a vertical reinforcing bar and the other set midway between adjacent vertical bars.

All the results exhibited a similar pattern: chloride ion reduction was most rapid in the early stages of treatment, there

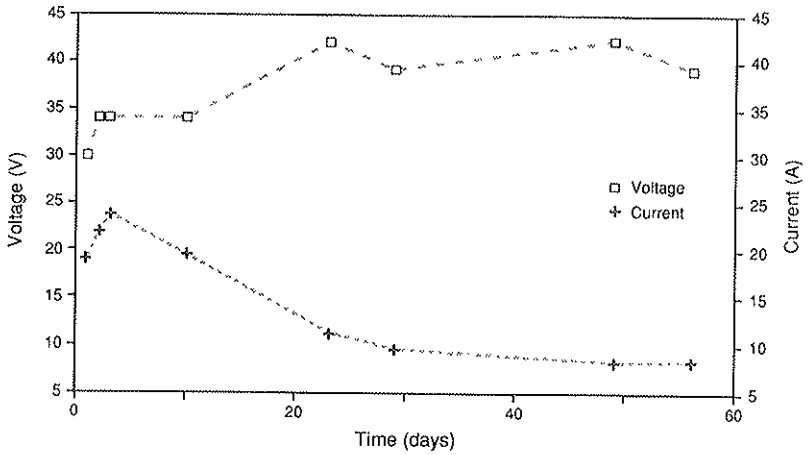



**FIGURE 4 Rectifier output during treatment period.**

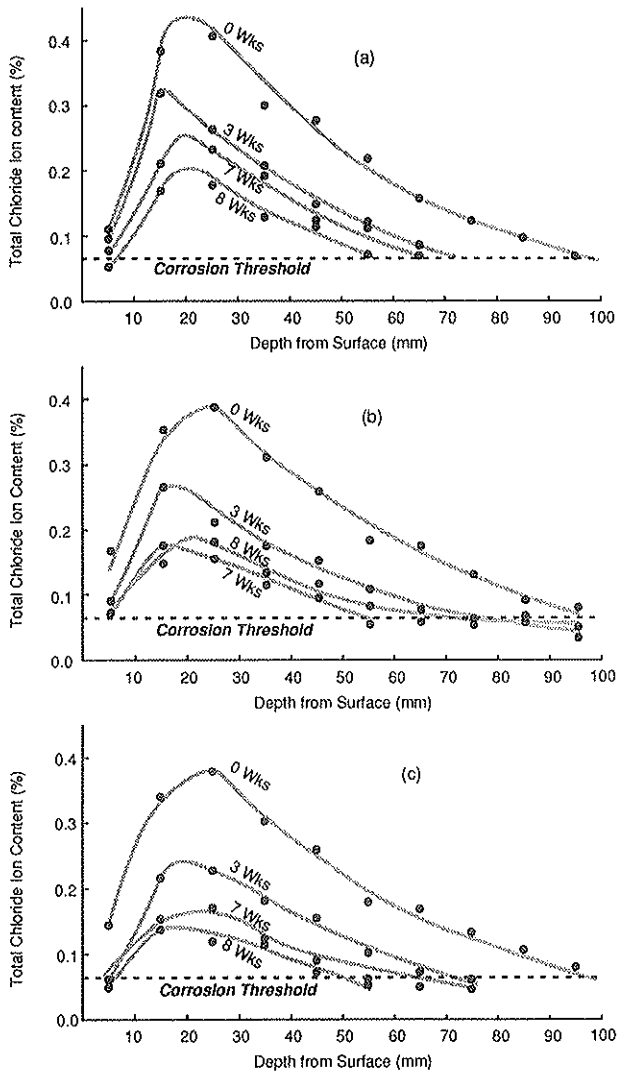



**FIGURE 5 Chloride ion profiles during treatment period: (a) south face between rebars, (b) west face between rebars, and (c) west face over rebar.**

was a reduction in the peak chloride ion content and a gradual shift in the peak value towards the concrete surface, and there was little difference between the results after 7 and 8 weeks. In all cases, the chloride ion content at the level of the reinforcement was at or slightly below the calculated corrosion threshold value for the concrete. From Figures 5b and 5c, it can be seen that proximity to the reinforcement had only a small effect, though the use of 100-mm-diameter cores would tend to mask differences in the immediate vicinity of the reinforcement.

The proportion of chloride ion in excess of the corrosion threshold value removed by the treatment was calculated from the area under the curves in the chloride ion profiles. The results are presented in Table 1. The range represents the values calculated after 7 and 8 weeks. In some cases, as for example in Figures 5b and 5c, the values were less after 7 weeks than after 8 weeks, though this effect is thought to be the result of the method of sampling.

### Corrosion Potential Measurements

The potential of the embedded reference electrodes was measured periodically, and the rate of depolarization is shown for the case of the west face in Figure 6. Measurements were made on an embedded graphite electrode. The original static potential was reached after about 60 days. The steel continued to depolarize and zero potential was measured about 8 months after treatments. This value was confirmed by surface measurements.

The first set of surface potential measurements were made 6 weeks after treatment because of the impending onset of winter, but the values were influenced by the fact that the reinforcement had not completely depolarized. Additional readings were taken 6, 10, and 13 months after treatment when the air temperatures were 15°C, 27°C, and 12°C, respectively. The percentages of readings falling within each of the three ranges normally associated with passive, uncertain, and active corrosion activity are presented in Table 2. The potential measurements for the treated faces were sensibly the same at 6 months and later ages, and indicated a substantial reduction in corrosion activity. However, the potential measurements taken on the untreated north face also showed a shift in the positive direction, which fluctuated with time and temperature.

### Corrosion Current Measurements

The corrosion currents were measured using a commercial three-electrode linear polarization device. The results 6 weeks, 6 months, and 13 months after treatment are presented in Table 3 and indicate a substantial reduction in corrosion activity. The average values recorded on the north face exhibit the fluctuation commonly observed on reinforced-concrete structures.

### Petrographic Examination

Two cores, 120 mm in diameter, were examined petrographically by Lankard Materials Laboratory, Inc., to apply the same techniques as were developed for use in SHRP Contract C-102A. One core was taken from the east face and the other from the untreated north face. Both cores contained a No. 11 rebar about 100 mm from the surface.

The cores were subjected to the following tests: a petrographic examination on lapped and fracture surfaces; scanning electron microscopy (SEM) on fracture surfaces; energy dispersive X-ray analyses (EDS) to provide elemental chemical

TABLE 1 PERCENTAGE OF CHLORIDE ION REMOVED BY TREATMENT

| FACE | LOCATION | % $Cl^-$ REMOVED |
|---|---|---|
| West | Over rebar | 76-87 |
| | Between rebars | 74-77 |
| South | Between rebars | 65-79 |
| East | Between rebars | 42-55 |

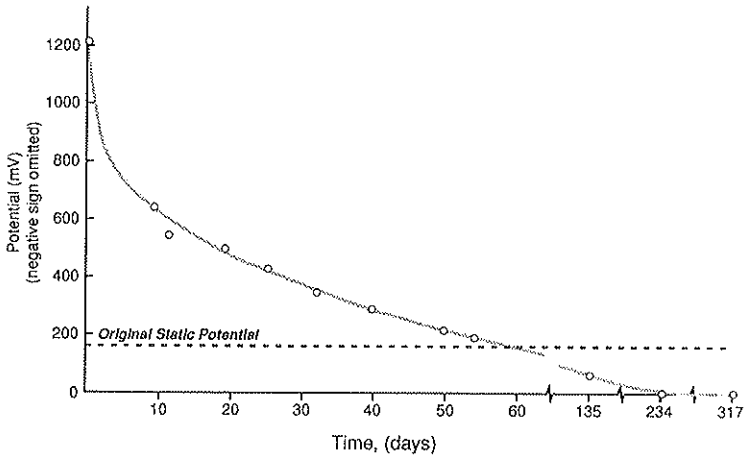



**FIGURE 6 Depolarization of the west face measured on an embedded graphite electrode.**

TABLE 2 CORROSION POTENTIALS IN MILLIVOLTS RELATIVE TO Cu-$CuSO_4$, NEGATIVE SIGN OMITTED

| | NORTH | | | WEST | | | SOUTH | | | EAST | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | <200 | 200 to 350 | >350 | <200 | 200 to 350 | >350 | <200 | 200 to 350 | >350 | <200 | 200 to 350 | >350 |
| Original | 0 | 85 | 15 | 0 | 96 | 4 | 0 | 96 | 4 | 0 | 84 | 16 |
| 6 weeks after treatment | 0 | 81 | 19 | 38 | 42 | 20 | 22 | 33 | 45 | 29 | 31 | 40 |
| 6 months after treatment | 63 | 37 | 0 | 96 | 4 | 0 | 100 | 0 | 0 | 98 | 2 | 0 |
| 10 months after treatment | 30 | 70 | 0 | 96 | 4 | 0 | 100 | 0 | 0 | 98 | 2 | 0 |
| 13 months after treatment | 41 | 59 | 0 | 98 | 2 | 0 | 100 | 0 | 0 | 100 | 0 | 0 |

TABLE 3 AVERAGE CORROSION CURRENTS ($\mu A/cm^2$)

| | NORTH | WEST | SOUTH | EAST |
|---|---|---|---|---|
| Before | 1.24 | 0.71 | 0.60 | 1.61 |
| 6 weeks after treatment | 0.90 | 0.15 | 0.11 | 0.16 |
| 6 months after treatment | 0.81 | 0.22 | 0.10 | 0.17 |
| 13 months after treatment | 1.24 | 0.15 | 0.10 | 0.18 |

analyses on mortar fracture surfaces; and mercury porosimetry on mortar taken from adjacent to the rebar to determine pore volume and distribution.

The concrete was assessed to be of good quality, and no features of mechanical or chemical distress that could be attributed to the electrochemical treatment were observed in the petrographic examination.

The results of the SEM and EDS examinations on samples removed from the top 75 mm of the cores indicated that the electrochemical treatment had no significant effect on the concrete chemistry or microstructure other than the intended purpose of reducing the chloride ion concentration. In the electrochemically treated concrete, a slight discoloration of the concrete adjacent to the rebar was observed, but the paste and aggregate were judged to be of the same quality as the other sites examined in the core. A redistribution of potassium ions occurred in the treated concrete, and high levels were found in isolated pockets of mortar within 6 mm of the rebar. It is likely that the concentration of sodium ions also increased at these locations. The porosity of the cement paste phase was 25 to 30 percent higher in the treated concrete, though the increased porosity was primarily in the pore size range less than 1 μm and was not expected to have a significant effect on the permeability of the concrete.

No evidence of cracking or softening of the concrete resulted from the treatment. It was concluded that at the time

of examination the electrochemical treatment had not had an adverse effect on the concrete (7).

## DISCUSSION OF RESULTS

The electrochemical treatment removed a large proportion of the chloride ions from the concrete without apparent damage to the pier and with only minor effects on the microstructure of the concrete. The progress of the removal was much as anticipated; relatively rapid at first with a noticeable decline in efficiency as the circuit resistance increased and the transference number decreased. The transference number, $t_{Cl^-}$, where $t_{Cl^-} = Q_{Cl^-}/Q_{TOTAL}$, represents the quantity of electric charge carried by the chloride ions as a proportion of the total charge passed. The transference number is directly related to the chloride concentration but is independent of current density and only slightly affected by temperature (5). Although no attempt was made to determine the optimum treatment time, there was little change in the chloride ion profiles between 7 and 8 weeks. Treatment times could be reduced by increasing the current density, although this would have required a larger power source and increased the risk of damage to the concrete. The current density on the concrete decreased from an initial value of 0.77 A/m$^2$ to 0.26 A/m$^2$ after approximately 6 weeks, with the result that the operating current density was substantially less than the 2 A/m$^2$ suggested as the upper limit for treatment without damage to the concrete (5).

The corrosion potential measurements and, more particularly, the rate of corrosion measurements indicated that the corrosion activity had been substantially reduced. These measurements were consistent with the chloride ion contents measured at the level of the reinforcement. However, a change in corrosion activity would be expected as a result of the large degree of polarization of the steel, and three major questions that will determine the technical and economic feasibility of the method remain to be answered:

1. For how long is the treatment effective?
2. What happens to the chloride ions that remain in the concrete?
3. What happens to the chloride ions in the corrosion products and pits in the reinforcement?

The answer to the first question requires long-term monitoring, which is planned for Pier S19 at Burlington. If the treatment was used as a routine operational procedure, it would be expected that, as a minimum, the surface of the concrete would be sealed against further ingress of chloride ions. Because the application procedures for most concrete sealers require that the surface be prepared by blast cleaning, the rust staining associated with the use of the steel mesh anode is not considered a major disadvantage compared with the use of an inert anode. The surface of the concrete in this project was not sealed because the source of the chloride ions has been eliminated as the result of the construction of a new bridge deck in 1988 that was continuous over Pier S19. This will enable changes in the pier to be monitored without the complicating factor of the role of a sealer. In the longer term, it is possible that the service life of chloride ion removal treatments will be extended by the migration of cationic corrosion inhibitors into the concrete, either concurrently with the removal process or as a separate posttreatment.

Although the treatment removed a substantial proportion of the chloride ions from the concrete, 13 to 58 percent of the original chloride ion content in excess of the threshold for corrosion remained in the concrete at the conclusion of the treatment. The reason for the much lower proportion of chloride ions removed from the east face is uncertain because the current density applied to each of the treated faces was sensibly the same. Although chloride ion measurements indicated that the value at the reinforcing steel was at or below the anticipated threshold value for corrosion, the EDS analyses detected chloride ions at the concrete-rebar interface and in close proximity to the rebar in the treated concrete. However, the chloride ions were associated with sulfur and aluminum ions and may have been chemically combined as chloroaluminates or chlorosulfoaluminates. The biggest unknowns are then the extent to which chloride ions contained in the steel corrosion products and in the concrete will initiate further corrosion and the period of time during which this may happen, because these factors determine the practicality of the process. In the case of Pier S19, the chloride ion content below the reinforcing steel was quite low, but in a heavily contaminated member, the ability of the process to remove chloride ion from behind the reinforcing steel would also require investigation. This situation is likely to represent the more typical case because of the large amount of cover in Pier S19.

In addition to the above uncertainties, additional work is required to define optimum treatment parameters, particularly with respect to zone area, current density, and time. However, even with optimization, treatment times are likely to be weeks rather than days, making the process more suitable for application to substructure components than decks. Finally, the costs of the treatment have not been well established. Mobilization and the requirement for daily wetting and periodic monitoring have a major impact on the cost of small-scale field trials. There has been insufficient experience to determine reliable costs for full-scale applications, but the costs are expected to be competitive with other rehabilitation methods providing that electrochemical removal will offer a similar extension of service life.

## CONCLUSIONS

1. Chloride ions were moved away from the reinforcing steel and a significant proportion were removed from the concrete without apparent mechanical or chemical damage to the concrete.
2. The corrosion activity of the steel was reduced substantially.
3. The long-term effectiveness of the treatment is the major unknown factor.

## REFERENCES

1. D. R. Lankard, J. E. Slater, W. A. Hedden, and D. E. Niesz, *Neutralization of Chloride in Concrete*. Report FHWA-RD-76-60. FHWA, U.S. Department of Transportation, 1975, 143 pp.

2. J. E. Slater, D. R. Lankard, and P. J. Moreland. Electrochemical Removal of Chlorides from Concrete Bridge Decks. In *Transportation Research Record 604,* TRB, National Research Council, Washington, D.C., 1976, pp. 6–15.
3. G. L. Morrison, Y. P. Virmani, F. W. Stratton, and W. J. Gilliland. *Chloride Removal and Monomer Impregnation of Bridge Deck Concrete by Electro-Osmosis*. Report FHWA-KS-RD-74-1. FHWA, U.S. Department of Transportation, 1976, 41 pp.
4. Strategic Highway Research Program Research Plans. Final Report, Strategic Highway Research Program, National Research Council, Washington, D.C., 1986.
5. J. E. Bennett, and T. J. Schue. Electrochemical Chloride Removal from Concrete: A SHRP Contract Status Report. Paper 316, *Proc., Corrosion 90,* National Association of Corrosion Engineers Tex., 1990.
6. SHRP Evaluates Chloride Removal Systems for Corroded Bridges. *Focus, Strategic Highway Research Program,* National Research Council, Washington, D.C., 1989, pp. 4–5.
7. *Characterization of Concrete Cores Taken From Pier S-19, Burlington Skyway.* Report to the Ontario Ministry of Transportation, Lankard Materials Laboratory, Inc., May 17, 1990.

*Publication of this paper sponsored by Committee on Corrosion.*

# Cathodic Protection of the Concrete Piers of Two Bridges in Virginia Using a Water-Based Conductive Coating

GERARDO G. CLEMEÑA AND DONALD R. JACKSON

There is a need for a simple and inexpensive anode for use in the impressed-current cathodic protection (CP) of inland concrete piers that are deteriorating because of salt-induced corrosion of rebars. In search of such an anode, a new water-based conductive coating was used recently on the cathodic protection of some concrete piers in Virginia. Further, as a possible means of eliminating the need for regular site visits to inspect and ensure that the CP is functioning properly (a disadvantage common to existing CP systems), a microprocessor-based data acquisition device that facilitates remote monitoring was tested with the system. The design, the installation, and the performance of the CP system during its first year of operation are described.

There are three strategies for the rehabilitation of concrete piers that are deteriorating because of salt-induced corrosion of rebars: (a) remove the deteriorated concrete, clean the rebars, and patch the excavated areas; (b) in addition to these procedures, replace the structurally sound concrete that is already contaminated with high amounts of deicing salt with new concrete; or (c) in addition to these procedures, apply cathodic protection (CP) to the piers.

Although the least costly, Alternative 1 eventually leads to cycles of deterioration and repair because the new concrete in the patches creates new electrochemical imbalances in the structure. Alternative 2 could prevent these cycles from occurring for probably 10 to 15 years, after which time enough salt would eventually accumulate again in the new concrete to induce corrosion of the rebars. Further, this alternative can be prohibitively expensive if some load-bearing concrete has to be replaced.

Alternative 3 is, theoretically, the ideal solution because CP eliminates the need to remove contaminated concrete, and it deals with the underlying problem directly by halting further corrosion of the rebars. In addition, the cost of a CP system is becoming relatively economical.

In the cathodic protection of any concrete substructures, the selection of a suitable anode system is critical, because the anodes are vital components of a CP system (see Figure 1). In general, anodes must be

- Electrically conductive,
- Reasonably chemically inert,
- Inexpensive,
- Easy and safe to install or apply, and
- Reasonably easy to maintain.

Small-scale testing of different anode materials on concrete piers, including conductive coatings and zinc metallized coatings, has been reported (*1–3*). It appears from these and several unpublished reports that conductive coatings have the best prospect of fulfilling these requirements. In 1988, the most extensive use of a conductive coating was made in a CP system for 93 concrete pier caps in Richmond, Virginia (*4*). The early observations made on this system confirmed that this type of anode indeed holds promise for effective use on inland concrete piers.

Conductive coatings are paints made electrically conductive by the addition of finely dispersed carbon particles; they are applied on the surface of concrete with brushes or rollers. However, in the past, their use required extreme precaution because all available conductive coatings contained organic compounds—such as xylene, propylene glycol, and monoethyl ether—which are considered potentially hazardous.

The recent introduction of a proprietary water-based conductive coating has provided a safer alternative. This coating consists of a blend of specially treated carbon dispersed in an acrylic resin with various properties that are presented in Table 1. Recent testings in a laboratory and an exposure yard have indicated that this water-based coating is as durable as the best organic-based conductive coating (*5*).

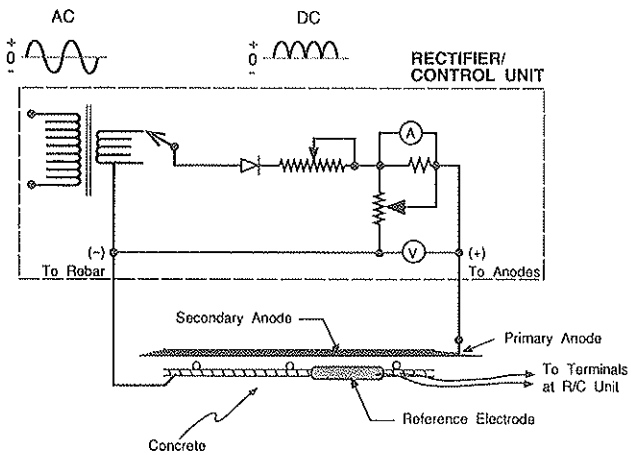



**FIGURE 1 Components of a typical CP system for reinforced concrete substructures.**

G. G. Clemeña, Virginia Transportation Research Council, Box 3817 University Station, Charlottesville, Va. 22903–0817. D. R. Jackson, Demonstration Projects Division, FHWA, 6300 Georgetown Pike, McLean, Va. 22101.

TABLE 1 PROPERTIES OF THE WATER-BASED CONDUCTIVE COATING

| | |
|---|---|
| Pigment | specially treated non-graphite carbons |
| Binder | acrylic |
| Color | black |
| Carrier | water |
| Density | @ 1.56 g/cc (13 lbs/gal)% |
| Solids (by wt.) | 73% |
| Solids (by vol.) | 67 |
| Viscosity | 6,000-10,000 cps (Brookfield RVT) |
| pH | 8.5 |
| Flash point | none |
| Linear resistance | 5.9-7.9 ohm/cm (15–20 ohms/in), point-to-point, at 10 mils dry film thickness |
| Recommended thickness | 0.254-0.381 mm (10–15 mils) |
| Theoretical coverage | 26.2 sq m/l at @ 0.025 mm thick (1072 sq ft/gal at @ 1 mil) |
| Actual coverage | 2.44 sq m/l (100 sq ft/gal) |

Obtained from manufacturer's provisional product data sheet

These developments prompted the recent experimental use of this water-based conductive coating in a CP system for the 10 concrete piers of two twin bridges in Virginia. The design and installation of the system are described and up-to-date observations on the system's performance are reported.

## DESIGN OF THE CP SYSTEM

The coating was used on the concrete piers of the two bridges (Structures 2014 and 2015) that were built in 1966 to carry the northbound and southbound lanes, respectively, of Interstate 81 over Route 698 in Shenandoah County, Virginia. During an inspection in 1987, it was found that all 10 concrete piers were exhibiting all the symptoms related to salt-induced corrosion of rebars, which is a consequence of deicing salt coming through faulty deck joints located directly above the piers. By sounding the concrete, its deterioration was estimated to range from 0.1 to 18.0 percent of the surface area of the piers.

In accordance with the geometry of the piers, it was specified that a set of 6 separate PT-Nb-Cu wires of 0.79-mm diameter be installed on each pier at locations shown in Figure 2 to serve as the primary anodes. These six anode wires would provide adequate redundancy to prevent complete failure in a circuit should any of the wires become accidentally disconnected in the future. Each anode wire was taped in place with adhesive mesh tapes and then covered with a conductive anode paste (see Figure 3). Then, two coats of the conductive coating, which would serve as a secondary anode, were applied on each pier with rollers or brushes. Although not absolutely necessary, the black conductive coating was covered with a light exterior acrylic coating to ensure the durability of the conductive coating.

In order to avoid electrical interference between the piers, the six anode wires on each pier were connected to an independent circuit in a common rectifier-controller (R/C) unit; hence, the R/C unit had to have 10 independent adjustable circuits. The maximum capacity of each circuit was 10 amperes at 20 volts. The R/C unit was operated in a constant-current mode and connected to a 220-volt ac utility line.

The locations of two system ground (negative) connections, one embedded graphite reference electrode, and eight test windows that would allow for measuring the rebar potential with an external Cu-$CuSO_4$ electrode are shown in Figure 2.

In order to ensure that a CP system is functioning properly, its R/C unit has to be inspected regularly (at least once a month) by obtaining readings of the current and voltage outputs of each circuit. When many systems have to be monitored, such requirements can become a burden. Fortunately, this difficulty can be avoided by the use of remote monitoring technology to eliminate the need for bridge visitation to inspect the CP systems.

In order to test this new technology, the R/C unit was equipped with a microprocessor-based data acquisition system

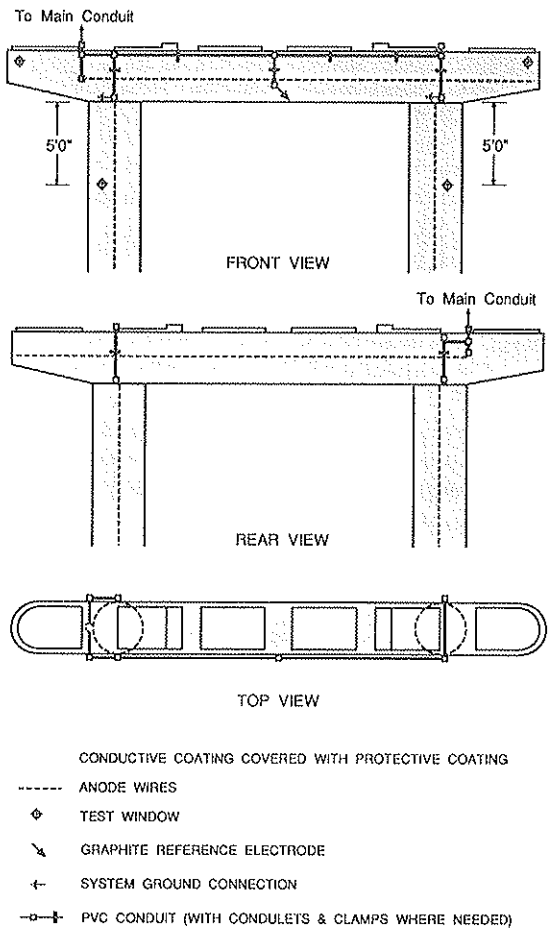



**FIGURE 2 Layout of CP system components on each pier.**

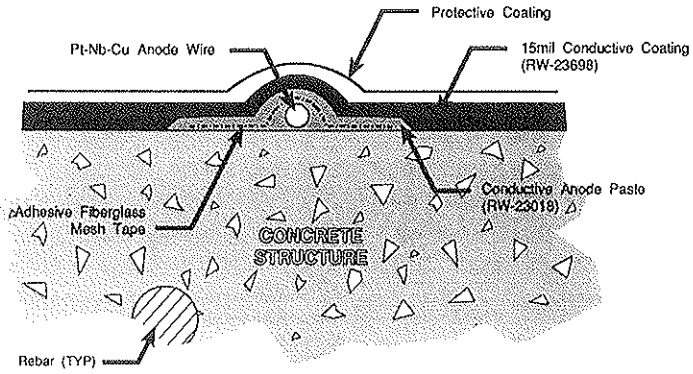



**FIGURE 3 Taping of anode wire and application of anode paste, conductive coating, and protective coating.**

(DAS) and a modem, which was connected to a telephone line at the bridge site to allow access to the unit from a remote office through another modem and a microcomputer. The 30-channel DAS facilitated the remote monitoring of the current output, the driving voltage of each of the 10 circuits, and the response of the rebars in each pier (i.e., the rebar potential as measured by the embedded graphite electrode).

## INSTALLATION PROCEDURES AND ENERGIZATION OF THE SYSTEM

The installation procedures were as follows:

1. The deteriorated concrete was removed to 25 mm below the rebars, and the corroded rebars were sandblasted in accordance with VDOT specifications (*6*).
2. Tests were made for electrical continuity between the rebars in each pier to ensure that no rebars were electrically isolated from the rest, thus leaving them unprotected by the CP system. This important test was conducted by measuring the dc and ac resistances between the rebars at several locations (usually at the ends of the cap and at the top and bottom of the columns). Any electrically isolated rebar (identified by abnormally high resistance) was then electrically bonded to a nearby rebar by thermite welding an insulated copper wire between them.
3. All metallic appurtenances (such as drains, anchor bolts, etc.) were connected electrically to the rebars.
4. Two system ground connections and a graphite reference electrode were installed in each pier at locations shown in Figure 2.
5. The excavated areas were prepared and then patched with pneumatically applied mortar in accordance with VDOT specifications (*6*). If an excavated area was too large, it was necessary to tie a small metal mesh to the rebars to ensure that the mortar would stay in place. In such cases, only ungalvanized steel mesh was allowed, and then the contractor had to ensure that the edge of the mesh was more than 3.8 cm below the finished surface.
6. All small metal wires exposed at the surface of the concrete were masked with vinyl ester resin sealant. This important procedure had to be carried out methodically to prevent any wire (particularly chairs and tie wires at the underside of a pier cap, where the concrete cover tends to be thin) from being left exposed to come in contact with the conductive coating, thereby creating a short in the circuit.
7. Several 51- by 51-mm concrete areas in each pier had to be masked with duct tape to prevent them from being coated with the conductive coating. These areas would serve as test windows.
8. The edges of metallic appurtenances and the concrete area (to 75 mm) surrounding these appurtenances had to be masked to prevent their being coated by the conductive coating and thereby creating shorts in the circuits.
9. Six Pt-Nb-Cu primary anode wires were installed on each pier at specified locations shown in Figure 2 and covered with conductive anode paste.
10. Two coats of the conductive coating were applied to the concrete with rollers after cleaning the concrete with light sandblasting. Figure 4 shows a pier after the conductive coating was applied.
11. An acrylic coating was applied over the dried conductive coating. Figure 5 shows a pier after the completion of the top coating.
12. All wirings were routed through PVC conduits to the common R/C unit located at the south end of the bridges.

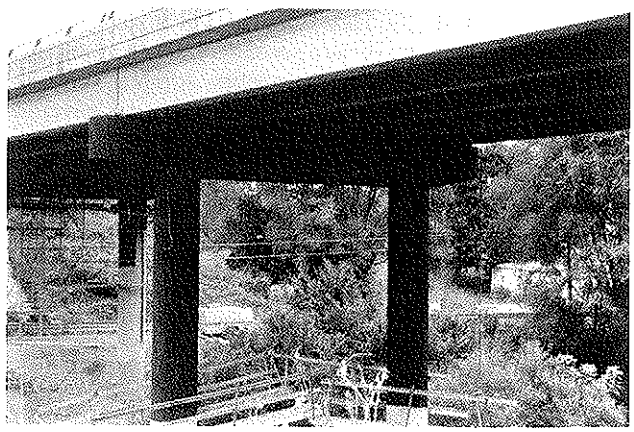


**FIGURE 4 Pier 3 of Structure 2014 after application of the conductive coating.**

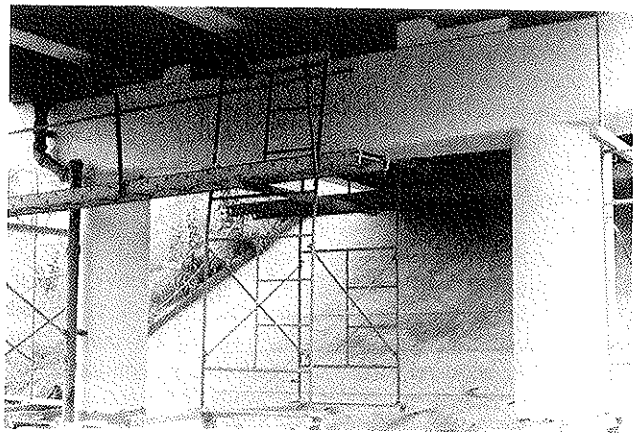


**FIGURE 5 Pier 1 of Structure 2014 after the application of top coating.**

After the installation, the wirings were inspected to verify that they were properly connected at the rectifier unit and in good working order. The rectifier unit was then energized and tested to determine the proper cathodic protection current required for each pier or circuit. These tests, which included measurements of the static potential of each reference electrode and of the $E$ versus log $I$ characteristic of each zone (circuit), will also provide data for future monitoring of the CP system. On the basis of the tests and some engineering judgment, the current level of each circuit was then adjusted and left to operate in constant-current mode for approximately 30 days. Thereafter, a depolarization test was conducted on each circuit to ascertain that its current level was providing sufficient polarization of the rebars in the pier. This test involves disconnecting the power to each circuit and recording the decay in the rebar potential (as measured with each embedded electrode). If the resulting potential decay curve shows a positive shift from the instant off value (after 4 hr of deenergization) of at least 100 mvolt, which is a criterion recommended by the National Association of Corrosion Engineers, the rebars in the pier are considered adequately protected from further corrosion.

## DISCUSSION OF RESULTS

### Cost of the CP System

In the three lowest bids received, the quoted total costs of the system (excluding the cost of concrete repair) ranged from \$85,546 to \$123,292. With a total concrete area of 1,040 m$^2$ to be protected, the lowest unit cost of the CP system came to be \$82.56/m$^2$. This cost was lower than that of the James River Bridge's CP system, which was approximately \$129/m$^2$ for a total concrete area of 7,520 m$^2$.

### Quantities of Coatings Used

Because of the relatively high solid content of the conductive coating, which was 73 percent by weight, it required appreciable effort to apply the coating on the concrete piers. With this exception, the application process proceeded without any problem. The total quantity of the various coatings used and their respective effective coverage are presented in Table 2. The effective coverage of the anode coating was estimated to be approximately 1.96 m$^2$/L, which was 20 percent less than the manufacturer's original estimate of 2.45 m$^2$/L (Table 1). However, the manufacturer has since revised its estimate to reflect this result.

### Cathodic Protection Current Outputs

On the basis of the $E$ versus log $I$ curves obtained during the postinstallation testing, the current output from the R/C unit to each pier was set to the level presented in Table 3, with

TABLE 2 QUANTITIES OF COATINGS USED

| Coating | Quantity Used (liter) | | | Actual Coverage (sq m/l) |
|---|---|---|---|---|
| | Str. 2014 | Str. 2015 | Total | |
| Anode Coating | 257 | 272 | 529 | 1.96 |
| Anode Paste | 23 | 23 | 46 | |
| Top Coating | 114 | 114 | 228 | 4.54 |

TABLE 3 STATIC REFERENCE ELECTRODE POTENTIAL AND INITIAL DC POWER SETTINGS ON THE RECTIFIER

| Zone | Structure | Pier | 11/21/89 | 11/22/89 | | | |
|---|---|---|---|---|---|---|---|
| | | | Static Potential (v) | Current (amp) | Voltage (volt) | Potential (volt) | dE (volt) |
| 1 | 2014 | 1 | −0.134 | 252 | 3.7 | −0.500 | −0.366 |
| 2 | | 2 | −0.050 | 140 | 1.8 | −0.417 | −0.367 |
| 3 | | 3 | −0.056 | 176 | 3.8 | −0.514 | −0.458 |
| 4 | | 4 | −0.073 | 312 | 4.8 | −0.526 | −0.453 |
| 5 | | 5 | −0.040 | 302 | 5.1 | −0.485 | −0.445 |
| 6 | 2015 | 1 | −0.101 | 540 | 5.3 | −0.458 | −0.357 |
| 7 | | 2 | −0.048 | 200 | 4.9 | −0.494 | −0.446 |
| 8 | | 3 | −0.073 | 80 | 8.5 | −0.485 | −0.412 |
| 9 | | 4 | −0.056 | 220 | 4.7 | −0.600 | −0.544 |
| 10 | | 5 | −0.142 | 594 | 6.3 | −0.579 | −0.437 |
| | | | minimum | 80 | 1.8 | | |
| | | | maximum | 594 | 8.5 | | |
| | | | average | 282 | 4.9 | | |

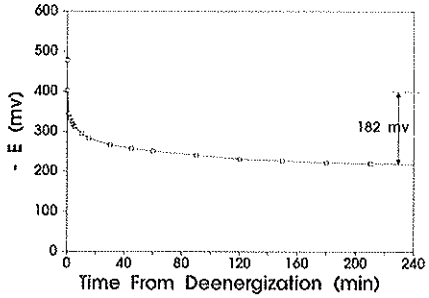



**FIGURE 6 Depolarization testing of Zone 1 (Pier 1 of Structure 2014).**

the goal of establishing rebar polarization of at least 100 mvolt. These current settings ranged from 80 to 594 mamp, with an average of 282 mamp. Expressed in terms of current density, these correspond to 0.7 to 5.9 mamp/m$^2$ of concrete area.

### Depolarization Tests

After these current outputs were maintained for 27 days, a depolarization test was performed in December 1989 on each circuit to ensure that its current output was sufficient to protect the rebars. The resulting decay curve for Zone 1 (Figure 6) indicated that the potential shifted from the instant-off value of −402 to −220 mvolt after 3.5 hr, indicating that there would be a depolarization of at least 182 mvolt after 4 hr. The extent of depolarization observed in all the piers ranged from 182 to 319 mvolt, with an average of 257 mvolt (Table 4). These results indicate that, by the criterion of the National Association of Corrosion Engineers (NACE) of minimum 100-mvolt depolarization shift, the rebars in all 10 piers were more than sufficiently cathodically protected from further corrosion. These results also indicated that the current applied to several piers can be decreased to levels that may provide an optimum balance between adequate protection of the rebars and prolonged service life for the conductive coating. Therefore, the current outputs of several circuits were decreased.

After more than 4 months of operation, a second set of depolarization tests was conducted in early May 1990. These tests indicated that the extent of depolarization shifts ranged from 126 to 288 mvolt, with an average of 205 mvolt (see Table 4). The most recent depolarization tests, conducted on August 1990 after 8 months of operation, indicated depolarization shifts ranging from 67 to 259 mvolt, with an average of 191 mvolt.

With the exception of Pier 1 of Structure 2014, all piers appeared to be sufficiently polarized (see Table 4). The cathodic current for Pier 1 will be increased if the results of the next depolarization test, which is planned for December 1990, indicate again that the pier is not sufficiently polarized.

### Operational Characteristics of the System

Since it started operation 12 months ago, the system has been monitored remotely with the aid of a desktop computer and a modem. The use of a remote monitoring device completely eliminated the need for on-site reading of the rectifier outputs; although difficult to estimate, the resulting savings in labor could be substantial. In addition, the device will also substantially reduce the downtime of the system if there is a

TABLE 4 DEPOLARIZATION TESTS

| Structure | Pier | Depolarization Shift (mv) 12/19/89 | 05/02/90 | 08/21/90 |
|---|---|---|---|---|
| 2014 | 1 | 182 | 126 | 67 |
| | 2 | 206 | 201 | 176 |
| | 3 | 236 | 189 | 178 |
| | 4 | 265 | 194 | 209 |
| | 5 | 262 | 226 | 190 |
| 2015 | 6 | 248 | 214 | 254 |
| | 7 | 297 | 209 | 214 |
| | 8 | 268 | 227 | 181 |
| | 9 | 319 | 288 | 259 |
| | 10 | 286 | 179 | 182 |
| minimum | | 182 | 126 | 67 |
| maximum | | 319 | 288 | 259 |
| average | | 257 | 205 | 191 |

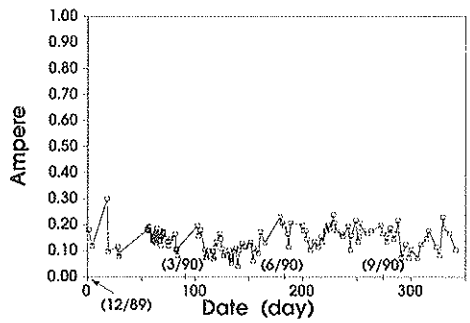



**FIGURE 7 Circuit current of Zone 9 (Pier 4 of Structure 2015) from Day 1.**

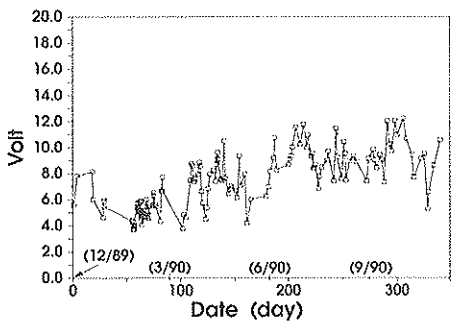



**FIGURE 8 Circuit voltage of Zone 9 (Pier 4 of Structure 2015) from Day 1.**

malfunction because it allows such a situation to be brought to the attention of the user automatically.

Figures 7–9 show what the cathodic current, the driving voltage, and the rebar potential, respectively, for Zone 9 (Pier 4 of Structure 2015) were during that period. Figure 7 shows that the R/C unit maintained an average current of approximately 0.14 amp for that pier. Responding to the varying resistance of the concrete in the pier, the driving voltage fluctuated between 3.7 and 12.2 volts—with an average of approximately 7.3 volts (see Figure 9). Figure 9 shows the response of the rebars to the cathodic current, as manifested in the recorded rebar potentials. These illustrations reflect the general behavior of the other nine zones. Table 5 presents the 10 average characteristics of the system during the first 11 months of operation.

The circuit resistances of the various anode coatings that have been tested or used on concrete piers in Virginia are presented in Table 6. Caution must be exercised when comparing these data, because those of the zinc spray and the polymer spray were obtained from testing only one concrete pier per type of coating. In contrast, the data for the organic-based coatings were obtained from 93 pier caps, and those for the water-based coating were derived from 10 piers. The water-based coating exhibited higher circuit resistances than the other coatings. Nevertheless, it appeared that zinc spray offered the lowest resistance. However, zinc is likely to degrade faster than carbon, which is the component that makes the other coatings conductive. It is not certain whether there is any difference between the resistances of the polymer spray and the organic-based coating. Consistent with relatively higher linear resistance quoted by its manufacturer (see Table 1), the water-based coating yielded higher circuit resistance. Although this resistance should not present any problem, adjustment must be made in the design of any future CP system that will use the water-based coating to counterbalance it.

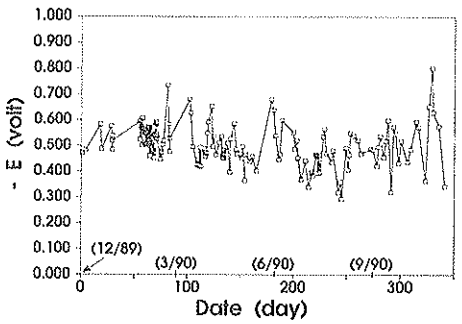



**FIGURE 9 Concrete-to-rebar potential measured by graphite electrode embedded in Zone 9.**

### Visual Inspection of the Coating

Visual examination of the coating after 12 months of operation did not reveal general deterioration. However, some localized carbonization of the protective coating was observed beside two of the four steel drain pipes that are attached to the piers (see Figure 10).

This carbonization was indirectly the result of cracks that were found at the elbow sections of these two pipes, which allowed excessive leakage of rainwater on the nearby protective coating. This caused some temporary localized discharge of the direct current from the nearby primary anode to the protective coating, which led to the carbonization of the latter. This conclusion is supported by the fact that the other two drain pipes were in good condition and the surrounding coating did not show any carbonization. This problem can be avoided by proper maintenance of the drain pipes and the

TABLE 5 AVERAGE OPERATIONAL CHARACTERISTICS OF THE CP SYSTEM

| Zone | Structure | Pier | Current (amp) | Voltage (volt) | Current Density (ma/sq m)** |
|---|---|---|---|---|---|
| 1 | 2014 | 1 | 0.233 | 6.07 | 2.9 |
| 2 | | 2 | 0.228 | 5.88 | 2.3 |
| 3 | | 3 | 0.174 | 5.88 | 1.7 |
| 4 | | 4 | 0.232 | 9.07 | 2.3 |
| 5 | | 5 | 0.194 | 9.68 | 2.0 |
| 6 | 2015 | 1 | 0.272 | 6.55 | 3.2 |
| 7 | | 2 | 0.133 | 8.84 | 1.1 |
| 8 | | 3 | 0.256 | 8.71 | 2.1 |
| 9 | | 4 | 0.143 | 7.28 | 1.2 |
| 10 | | 5 | 0.213 | 5.75 | 1.9 |
| minimum | | | 0.133 | 5.75 | 1.1 |
| maximum | | | 0.272 | 9.68 | 3.2 |
| average | | | 0.208 | 7.37 | 2.1 |

* for the first 11 months of operation
** per concrete area

TABLE 6 CIRCUIT RESISTANCES OF VARIOUS ANODE COATINGS INSTALLED IN VIRGINIA

| Coating System | Resistance (ohm) Low | High | Time of Observation |
|---|---|---|---|
| Zinc Spray (Norfolk, Va.) | 3.4 | 8.0 | During first 8 months |
| Polymer Spray (Norfolk, Va.) | 6.7 | 19.0 | During first 8 months |
| Organic-Based Coating (Richmond, Va.) | 1.4 | 38.6 | At start of operation |
| | 0.6 | 45.6 | After 12 months |
| Water-Based Coating (Shenandoah County, Va.) | 18.2 | 70.7 | At start of operation |
| | 26.9 | 77.5 | After 11 months |

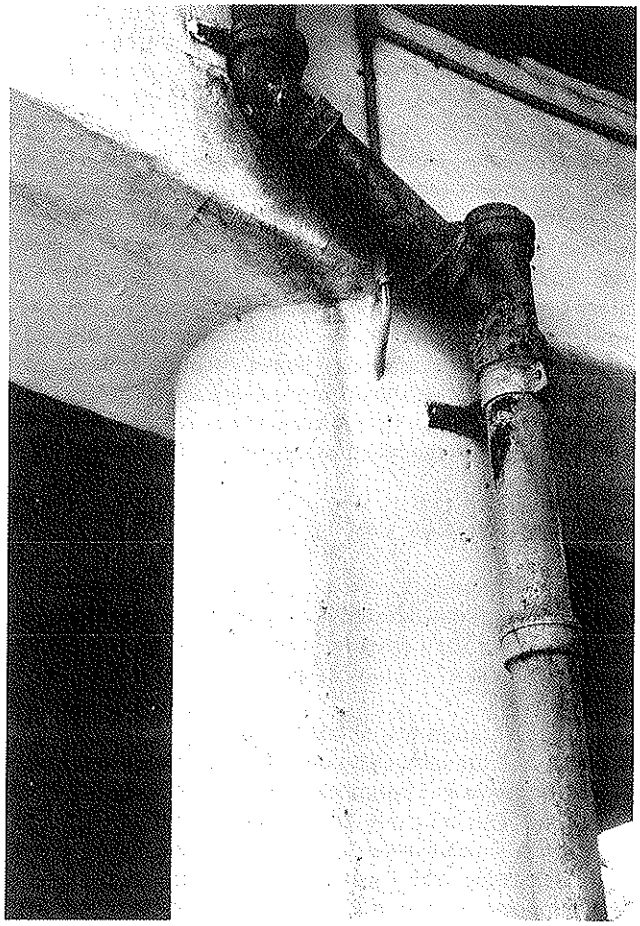


**FIGURE 10 Carbonization of protective coating beside a leaky drain pipe.**

placement of the primary anodes away from the vicinity of a pipe.

## CONCLUSION

On the basis of the observations made so far, the water-based coating still holds promise as an effective alternative anode coating for cathodic protection of inland concrete piers. Its relative ease of application and the fact that it is free from the health hazards that the other anode coatings present during application make the anode coating an attractive alternative.

As expected, the remote monitoring device allows convenient monitoring of the conditions of the electrical components of the CP system from anywhere a phone, a modem, and a personal computer are available.

## ACKNOWLEDGMENT

Appreciation is extended to John M. Bouwens of the Acheson Colloids Company and Brian J. McGrath of Matcor, Inc., for their cooperation. Special appreciation should go to Larry L. Misenheimer of the Staunton District of the Virginia Department of Transportation for his support of the project.

## REFERENCES

1. W. J. Swiat and J. W. Rog. *Further Improvements in Cathodic Protection of Bridge Structures.* FHWA-RD-87-062. FHWA, U.S. Department of Transportation, 1987.
2. W. J. Swiat and P. E. Bushman. *Further Improvements in Cathodic Protection.* FHWA-RD-88-267. FHWA, U.S. Department of Transportation, 1989.
3. H. G. Schell and D. G. Manning. Research Direction in Cathodic Protection for Highway Bridges. *Materials Performance,* Vol. 28, No. 10, 1989, pp. 11–15.
4. *James River Bridge Cathodic Protection System.* Norton Corrosion Engineers, Bothell, Wash., 1988.
5. *Laboratory Testing of an Acheson Conductive Coating Formulated for Exterior Concrete Surfaces.* Kenneth C. Clear, Inc., Sterling, Va., 1987.
6. *Road and Bridge Specifications.* Virginia Department of Transportation, Richmond, Va., 1987.

*Publication of this paper sponsored by Committee on Corrosion.*

# Electrochemical Studies of Rebar Corrosion and Inhibition in Simulated Pore Solution

L. A. WEBSTER, T. OSIROFF, J. G. DILLARD, J. O. GLANVILLE, AND R. E. WEYERS

Measurements of electrochemical potential have been used to evaluate relative inhibitor effectiveness in controlling the corrosion of rebar in chloride-doped simulated pore solution. The rates at which inhibitors bring the corrosion potential into the passive region provide a comparison of inhibitor effectiveness and serve as a screening test. Degreased rebar specimens were immersed in aerated solutions and their electrode potential was measured against a standard calomel electrode over several months. Measurements were made of corrosion potentials as a function of time both in the presence and in the absence of inhibitors and at various inhibitor concentrations. Results of the test procedure are described and comparisons are made with parallel tests of inhibition using visual methods and surface analytical techniques.

Spalling has been recognized as a major contributing factor to the deterioration of reinforced concrete structures (*1*,*2*). Spalling occurs when an iron oxide (or rust) on the steel reinforcement (rebar) occupies a volume greater than the metallic iron from which it is formed (*3*). Rebar corrosion occurs at an accelerated rate in concrete structures exposed to deicing salts or sea water (*3*,*4*). Typically, chloride-induced corrosion takes place at local points along the length of the bar, forming severe pitting (*3–5*). Thus, the load-carrying capacity of the bar is decreased at that point and, along with spalling, leads to gross deterioration of the concrete structure.

The treatment of corroded structures with inhibitors is a possible solution to the corrosion problem (*4*,*6–8*). In order to identify, evaluate, and compare the corrosion protection afforded by various inhibitors, rapid and nondestructive techniques are necessary (*9–11*). In this study, the measurement of electrochemical potential has been examined as a method of quickly evaluating relative inhibitor effectiveness for rebar specimens immersed in simulated pore solution. Rebar samples whose surfaces had been cleaned to remove surface contaminants and rebar specimens whose surfaces had been prepared to contain chloride-induced corrosion products were exposed to simulated pore solutions containing chloride ions and inhibitors. Electrochemical potential measurements were made as a function of time both in the presence and in the absence of inhibitors and at various concentrations. Materials that are found to be effective inhibitors under these experimental test conditions would be candidates for treatment of bridge decks without removal of concrete. The electrochemical measurement results are presented and comparisons are made with the findings from corrosion studies in which the effectiveness of inhibitors was determined using surface analytical techniques and visual inspection procedures (Dillard et al. and Dressman et al., companion papers in this Record).

L. A. Webster, T. Osiroff, J. G. Dillard, and J. O. Glanville, Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0212. R. E. Weyers, Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061–0105.

## EXPERIMENTAL

Electrochemical measurements were made to determine the corrosion potential of rebar immersed in test solutions. For all experiments, a standard calomel electrode was used as the reference electrode. The potentials were measured using a Fisher Accumet Model 910 pH/voltmeter and Fisher calomel electrode. ASTM Standard C876–87 relates potential ranges [relative to a copper sulfate electrode (CSE)] to the probability of corrosion. These values were converted from a CSE scale to a saturated calomel electrode (SCE) scale. The potential ranges (relative to SCE) and the probabilities are presented in Table 1.

Cylindrical reinforcement steel rods 6 ft long and ½ in. in diameter were obtained from Roanoke Electric Steel Co., Roanoke, Virginia. The composition of this material was formulated to be similar to that used 20 to 30 years ago. Bulk and surface analyses (Dillard et al., a companion paper in this Record) of the rods are presented in Table 2. Test specimens were prepared by cutting the rods into 3½-in. segments. One end of each segment was drilled and tapped to permit making electrical connections. The segments were cleaned in hexane and allowed to dry. Tru-Bond TB–700 epoxy paste was used to cover the wire connections. The cut end of the rebar was also covered with epoxy paste.

These cylindrical rebar samples were immersed in simulated pore solution, pH = 13.9 ± 0.1 [KOH (0.600 M); NaOH (0.300 M); saturated with $Ca(OH)_2$], containing NaCl (3.5 wt percent) (control); and pore solution containing NaCl (3.5 wt percent) and the following concentrations of inhibitor: 0.00200, 0.0100, 0.0500, and 0.100 M. The inhibitors studied were sodium nitrite ($NaNO_2$), sodium monofluorophosphate ($Na_2PO_3F$) (MFP), and sodium tetraborate ($Na_2B_4O_7$). The test solutions were aerated for 2 hr before rebar immersion. Samples of rebar were maintained at 60°C. In order to min-

TABLE 1 CORROSION SCALE (ASTM C876–87)

| Probability of Corrosion | E (corr); SCE Scale |
|---|---|
| greater than 90% | more negative than -290 mV |
| uncertain | -140 mV to -290 mV |
| less than 10% | more positive than -140 mV |

imize evaporation of solution, the bottles were covered. The test solutions were replaced every 2 to 3 weeks to maintain proper aeration, volume, concentration, and pH. Potential measurements were taken at regular time intervals.

Flat stock steel (A36, 1 in. wide × ¼ in. thick), with a composition similar to that of cylindrical rebar, was cut from material obtained from Roanoke Electric Steel Co., Roanoke, Virginia. The bulk and surface analyses for this material are presented in Table 3. The bars were drilled and tapped in one end to enable electrical connection. Bars were cleaned in hexane and allowed to dry. Additional bars were cleaned in a 1:1 (v/v) solution of concentrated sulfuric acid and deionized water. The latter bars were also scrubbed with a ScotchBrite pad, rinsed with deionized water, allowed to dry, cleaned with hexane, and allowed to dry. Two coats of Nybco epoxy paint were applied to each end of the rebar and allowed to cure according to the manufacturer's specifications.

These flat stock rebar specimens were immersed in simulated pore solution, pore solution containing NaCl (3.5 wt percent), pore solution containing NaCl (3.5 wt percent) and $NaNO_2$, (0.300 M), and pore solution containing NaCl (3.5 wt percent) and $NaNO_2$ (0.670 M). The test solutions were aerated for 2 hr before use. Samples were maintained in solution at 60°C and the samples bottles were covered to prevent evaporation of solution. The test solutions were replaced every 2 weeks. Potential measurements were taken at regular time intervals.

In experiments in which corroded rebar samples were prepared to investigate the effectiveness of compounds to inhibit active corrosion, flat stock rebar was maintained at 60°C in a solution of deionized water containing NaCl (10 wt percent) for a time sufficient to achieve an active corrosion potential. These corroded bars were then placed in the test solutions: pore solution, pore solution containing NaCl (3.5 wt percent), and pore solution containing NaCl (3.5 wt percent), and sodium tetraborate inhibitor at the concentrations 0.00200, 0.0100, 0.0500, and 0.100 M. The solutions were aerated for 2 hr before rebar immersion. The test samples were prepared in duplicate and maintained at 60°C. The test solutions were replaced every 2 weeks. Potential measurements were taken several times weekly and then at selected intervals.

In studies to evaluate the effectiveness of inhibitors toward active corrosion at different pH values, flat stock rebar specimens were cleaned in hexane, dried, and placed in aerated deionized water containing NaCl (10 wt percent) at pH 8.5 to 8.6 and 60°C for a time sufficient to obtain an active corrosion potential. The bars were subsequently placed in the solutions: NaCl (10 wt percent) (control), and NaCL (10 wt percent) containing 0.100 M inhibitor. The inhibitors were sodium nitrite, sodium tetraborate, MFP, and tetra-n-butylphosphonium bromide $\{[CH_3(CH_2)_3]_4PBr\}$. The pH values of the solutions were adjusted to 8, 10, and 12, respectively, with HCl, NaOH, or $NaHCO_3$. The samples were maintained at 60°C.

Potential and pH measurements were taken after 1, 3, 6, 9, and 18 days. The solutions were replaced after 18 days. At this time (18 days), the NaCl concentration was decreased to 1.75 wt percent and the inhibitor concentration was increased to 0.600 M, except for the concentration of tetra-n-butylphosphonium bromide, which was maintained at 0.100 M. At

TABLE 2 ESCA ANALYSIS OF REBAR ROD BEFORE AND AFTER HEXANE CLEANING AND MANUFACTURER'S BULK ANALYSIS

| | ESCA Element (atomic percent) | | Bulk Analysis (visible spectroscopic) | |
|---|---|---|---|---|
| | as received | after hexane cleaning | mass % | atomic % |
| C | 58.4 | 56.4 | 0.22 | 1.00 |
| O | 28.7 | 31.6 | nd | nd |
| Fe | 3.68 | 4.64 | 97.2 | 95.8 |
| Si | 3.26 | 2.56 | 0.59 | 1.16 |
| Cu | 2.06 | 1.70 | 0.26 | 0.23 |
| N | 1.41 | 1.09 | nd | |
| Na | 0.97 | 0.68 | nd | |
| S | 0.90 | 0.62 | 0.036 | 0.062 |
| Ca | 0.45 | 0.59 | nd | |
| Zn | 0.18 | 0.10 | nd | |
| Al | <0.1 | <0.1 | 0.006 | 0.01 |
| P | <0.1 | <0.1 | 0.018 | 0.032 |
| Mn | nd | nd | 1.00 | 1.00 |
| Cr | nd | nd | 0.50 | 0.52 |
| Ni | nd | nd | 0.11 | 0.10 |
| V | nd | nd | 0.04 | 0.04 |
| Mo | nd | nd | 0.02 | 0.01 |
| Total | 100.0 | 100.0 | 100 | |

ESCA = Electron spectroscopy for chemical analysis - surface composition
nd = not determined

TABLE 3 ESCA ANALYSIS OF REBAR (FLAT) BEFORE CLEANING, AFTER HEXANE CLEANING, AFTER ACID WASHING, AND MANUFACTURER'S BULK ANALYSIS

| | ESCA Element (atomic percent) | | | Bulk Analysis (visible spectroscopic) | |
|---|---|---|---|---|---|
| | as received | after hexane cleaning | after acid wash | mass % | atomic % |
| C | 39.5 | 58.0 | 54.4 | 0.11 | 0.51 |
| O | 45.0 | 28.0 | 25.2 | nd | |
| Fe | 6.67 | 4.09 | 4.17 | 98.0 | 97.5 |
| Si | 5.05 | 1.45 | 1.04 | 0.24 | 0.47 |
| Cu | nd | 0.68 | 0.07 | 0.31 | 0.27 |
| N | 0.82 | 2.34 | 2.32 | nd | |
| Na | 0.79 | 2.21 | 0.51 | nd | |
| S | 1.37 | 0.99 | 1.01 | 0.036 | 0.062 |
| Ca | 0.76 | 0.29 | 0.23 | nd | |
| Zn | nd | nd | 0.16 | nd | |
| Al | nd | nd | 2.25 | 0.004 | 0.008 |
| P | <0.1 | 0.39 | <0.1 | 0.022 | 0.004 |
| Mn | <0.1 | nd | 8.74 | 0.65 | 0.66 |
| Cr | nd | nd | nd | 0.25 | 0.27 |
| Ni | nd | nd | nd | 0.26 | 0.25 |
| V | nd | nd | nd | 0.003 | 0.003 |
| Mo | nd | nd | nd | 0.05 | 0.29 |

ESCA = Electron spectroscopy for chemical analysis; nd = not determined

this time (18 days), additional rebar specimens and solutions [NaCl (1.75 wt percent) containing 0.600 M inhibitor] were prepared to permit studies at pH 12.5 and 13.0. Potential and pH measurements were made after 3 and 6 days at the new conditions, and the pH value of each solution was adjusted to maintain a fixed pH value.

## RESULTS AND DISCUSSION

The first series of experiments was carried out using rebar to determine the manner in which the electrochemical potential changed for specimens immersed in chloride-containing pore solution containing inhibitor anions that are known to be effective agents, namely nitrite (*9–11*) and monofluorophosphate (Domtar Corp., Mississauga, Ontario) salts, and to compare the results with other potential inhibitors, e.g., borate and alkylphosphonium salts. The potential measurement results for cylindrical rebar, maintained in chloride-containing pore solution (control) or in chloride-containing pore solution including the inhibitors sodium nitrite, MFP, or sodium tetraborate, each at 0.0100 M concentrations, are shown in Figure 1. At the beginning of the experiments, the measured potential for all specimens is approximately −450 mvolt, a potential that indicates active corrosion.

As time passes, the potential for the control sample remains in the active corrosion region. In contrast, the potentials for rebar in solutions containing inhibitors become more positive with time and reach values of about −312 mvolt, nitrite; −315 mvolt, MFP; −282 mvolt, tetraborate; compared with −420 mvolt, control; after 16 weeks. The important result is that these three inhibitors are effective in making the corrosion potential more positive. Under the experimental conditions, the time at which the potential begins to become more positive (see Figure 1) is nitrite, 12 weeks; MFP, 10 weeks; and borate, 4 weeks.

For studies in which the concentration of inhibitor was greater than 0.010 M, the respective maximum values for the potentials were not greater than those measured for the 0.0100 M solutions, but the potentials began to increase at shorter times. The times at which the potential increased at concentrations of inhibitor other than 0.0100 M were also in the relative order borate < MFP < nitrite.

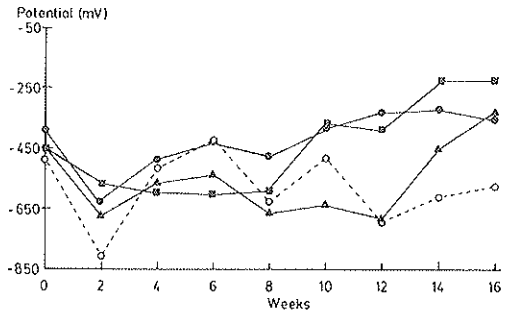



**FIGURE 1 Electrochemical potential (SCE) as a function of time for hexane-cleaned cylindrical rebar immersed in chloride-containing pore solution [NaCl (3.5 wt%)], and in chloride-containing pore solution [NaCl (3.5 wt%)] and 0.0100 M inhibitor solutions.**
**○—Control solution: NaCl (3.5 wt%) in pore solution**
**Test solutions: NaCl (3.5 wt%) in pore solution containing 0.0100 M inhibitor.**
**△—Sodium nitrite ($NaNO_2$)**
**▨—Sodium monofluorophosphate ($Na_2PO_3F$)**
**◎—Sodium tetraborate ($Na_2B_4O_7$)**

The preceding discussion indicates that potential measurements can be used to screen potential inhibitors; however, there are aspects of the approach that are undesirable for a *rapid* screening test; the effect of inhibitors at concentrations less than 0.0100 M is revealed only after several weeks into the test and if there is significant variation (±50 mvolt) in the measured potential from week to week. Furthermore, over this period of time complications were introduced by deterioration of rebar under the epoxy coating, and by degradation of the electrical connections and the epoxy that covers them. It is likely that chemical heterogeneity among the rebar specimens accounts for the variability in potential measurements.

In an effort to combat these problems, A36 flat rebar was used in all subsequent experiments and the method of making connections was modified. To determine the electrochemical behavior of the flat stock material, experiments were carried out (a) to determine whether flat stock rebar exhibited less variability in potential measurements than cylindrical rebar; (b) to examine the influence of an effective inhibitor, sodium nitrite, at selected concentrations, on electrochemical potential; and (c) to determine whether the method of cleaning (acid-washing plus cleaning in hexane compared with cleaning only in hexane) influenced the potential measurements using flat stock material.

That flat stock rebar (cleaned in acid and washed in hexane) exhibited less variability in potential is evident from the plot of potential (Figure 2) as a function of time for rebar immersed in pore solution containing no chloride. The effect of inhibitor on electrochemical potential is also shown in Figure 2. The potential for rebar maintained in chloride-containing pore solution containing $NaNO_2$ (0.300 M) becomes more negative during the first 2 weeks of immersion, indicating that severe corrosion occurs within 2 weeks of initiating the test. However, the potential increases for measurement taken at 4 weeks and continues to increase during the 12-week test, approaching a potential characteristic of reduced corrosion. The effect of increased $NaNO_2$ concentration is also shown in Figure 2, in which more positive potentials are obtained in a shorter time for the more concentrated inhibitor solution. Although not shown in the figure, the magnitude of the measured potential and the change with time for rebar cleaned only in cyclohexane was equivalent to that for the acid-cleaned, hexane-washed rebar.

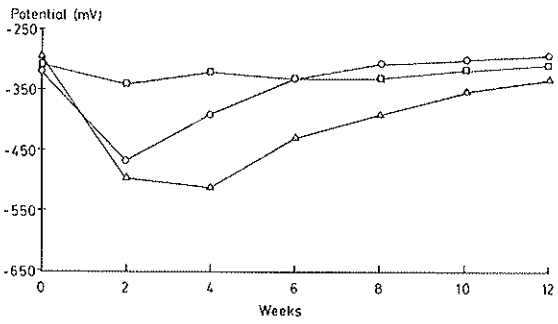



**FIGURE 2 Electrochemical potential (SCE) as a function of time for acid-washed, hexane-cleaned flat stock rebar immersed in pore solution, and in chloride-containing pore solution [NaCl (3.5 wt%)] and containing $NaNO_2$ (0.300 and 0.670 M).**
**□—Pore solution**
**△—Chloride-containing pore solution—$NaNO_2$ (0.300 M)**
**○—Chloride-containing pore solution—$NaNO_2$ (0.670 M)**

The principal results from these measurements with flat stock rebar are as follows: (a) flat stock rebar gives less variability in potential measurements than cylindrical rebar, (b) the method of rebar cleaning does not appear to influence the measured potential, (c) the higher concentration (0.670 M) of sodium nitrite produces a slightly more positive potential than that measured for sodium nitrite (0.300 M), and (d) the time to recovery of a corroded rebar surface is more rapid for more concentrated solutions.

Once it was demonstrated that measurement of electrode potential could be used to evaluate the performance of a given inhibitor, the change of potential for rebar immersed in sodium tetraborate, a potentially effective inhibitor, could be examined. Corroded A36 flat rebar specimens were immersed in pore solution containing NaCl (3.5 wt percent) and sodium tetraborate (0.00200, 0.0100, 0.0500, and 0.100 M). The potentials, measured as a function of time, are shown in Figure 3 for the borate (0.0100 M) solution. The variation of potential with time for rebar in other concentrations of borate was similar to that shown in Figure 3. Also plotted in Figure 3 is the variation in potential for a flat rebar sample immersed in pore solution containing NaCl (3.5 wt percent).

The potentials for the specimens maintained in the chloride-containing pore solution plus borate increased dramatically after only 12 hr in the test solution, whereas the potential for

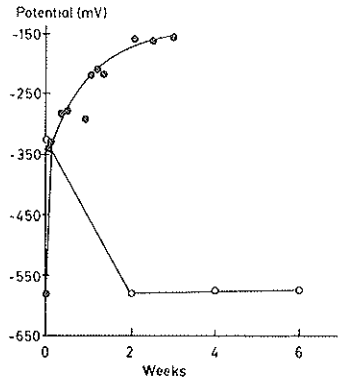



**FIGURE 3 Electrochemical potential (SCE) as a function of time for acid-washed, hexane-cleaned flat stock rebar immersed in chloride-containing pore solution [NaCl (3.5 wt%)], and for corroded flat stock rebar immersed in chloride-containing pore solution [NaCl (3.5 wt%)] and 0.0100 M $Na_2B_4O_7$.**
**○—Chloride-containing pore solution**
**◉—Corroded rebar in chloride-containing pore solution—$Na_2B_4O_7$ (0.0100 M)**

rebar in the solution containing no borate remained in the active corrosion region. By the end of the experiments, the potentials for all samples immersed in chloride-containing pore solution *with* borate had increased to values in the range −190 to −160 mvolt. On comparing potentials at different borate concentrations, no significant difference was noted. The potentials for the bars in pore solution containing chloride and no inhibitor remained in the active corrosion region, about −550 mvolt. The findings indicate that the potential for actively corroding rebar specimens can be increased to be in the region corresponding to uncertain probability of corrosion (see Table 1). From the data shown in Figure 3, prolonged immersion of rebar samples in borate inhibitor could eventually result in a potential more positive than −140 mvolt, the region of inactive corrosion, i.e., <10 percent probability of corrosion.

The question of the ability of inhibitors to function at pH values less than highly alkaline was of interest for the situations in which rebar might be exposed to more acidic environmental conditions. Corroded flat rebar samples were immersed in NaCl (10 wt percent) and 0.100 M inhibitor at selected pH values for 18 days and then subsequently placed in NaCl (1.75 wt percent) and 0.600 M inhibitor for additional time. The results are presented in Table 4. On comparing the control and inhibitor solutions during the first 18 days, there was no substantial difference in potential. In order to permit an evaluation of the performance of the inhibitors, the salt concentration was lowered and the inhibitor content was increased. With decreased salt concentration and increased inhibitor concentration, differences in inhibitor performance were revealed, as presented in Table 4. This procedure enables the parameters of the screening test to be optimized.

The performance of the various inhibitors is indicated by comparing the potential for test specimens maintained in chloride-containing solutions *with* inhibitor to those for control specimens, i.e., rebar immersed in chloride-containing solutions *without* inhibitor. An inhibitor is taken as being effective if the potential, $\Delta V$, is more positive by at least 90 to 100 mvolt when comparing potentials for inhibitor-containing solutions with those for solutions containing no inhibitor. The comparison indicates that borate is ineffective at pH 8; marginally effective at pH values 10, 12, and 12.5; and significantly effective at pH 13. Nitrite is effective at all pH values investigated. MFP is active at the intermediate pH values but is marginally effective at pH 8 and ineffective at pH 13. Tetra-n-butylphosphonium bromide is effective only at pH 13.

TABLE 4 ELECTROCHEMICAL POTENTIAL MEASUREMENTS FROM SCREENING TESTS

| | A | B | del V* | C | del V* |
|---|---|---|---|---|---|
| control | | | | | |
| pH 8 | -634 | -486 | | -431 | |
| pH 10 | -649 | -520 | | -508 | |
| pH 12 | -642 | -500 | | -428 | |
| pH 12.5 | -627 | | | -341 | |
| pH 13.0 | -651 | | | -652 | |
| sodium tetraborate | | | | | |
| pH 8 | -628 | -473 | +13 | -442 | -11 |
| pH 10 | -649 | -462 | +58 | -409 | +99 |
| pH 12 | -647 | -424 | +76 | -334 | +94 |
| pH 12.5 | -651 | | | -246 | +95 |
| pH 13.0 | -639 | | | -211 | +441 |
| sodium nitrite | | | | | |
| pH 8 | -645 | -481 | +5 | -27 | +404 |
| pH 10 | -630 | -445 | +75 | -61 | +447 |
| pH 12 | -646 | -406 | +94 | -79 | +349 |
| pH 12.5 | -648 | | | -114 | +227 |
| pH 13.0 | -621 | | | -266 | +386 |
| sodium monofluorophosphate | | | | | |
| pH 8 | -635 | -467 | +19 | -344 | +87 |
| pH 10 | -645 | -452 | +68 | -308 | +200 |
| pH 12 | -636 | -459 | +41 | -290 | +138 |
| pH 12.5 | -635 | | | -161 | +180 |
| pH 13.0 | -647 | | | -619 | +33 |
| tetra-n-butyl phosphonium bromide | | | | | |
| pH 8 | -635 | -482 | +4 | -443 | -12 |
| pH 10 | -632 | -549 | +29 | -502 | +6 |
| pH 12 | -636 | -517 | -17 | -505 | -77 |
| pH 12.5 | -623 | | | -367 | -26 |
| pH 13.0 | -631 | | | -317 | +335 |

* del V = potential of control minus potential of inhibitor under equivalent solution conditions.

A = Potential of corroded specimen before immersion in inhibitor solution.

B = Potential after 18 day immersion in 10(wt)% NaCl including 0.100M inhibitor.

C = Potential after 6 day immersion in 1.75(wt)% NaCl including 0.600M inhibitor of samples from B. (phosphonium bromide = 0.100M)

In summary, sodium nitrite, containing the anion common to other corrosion inhibitors (*9–11*), performed the best over the wide range of pH values. It produced the highest increase in potential for pH values 8, 10, 12, and 12.5. MFP, another known corrosion inhibitor (Domtar Corp., Mississauga, Ontario), performed second best, except at pH 13.0, where it ranked relatively poorly. Sodium tetraborate exhibited excellent behavior and ranked as the best inhibitor at pH 13.0, which most accurately simulates the macrolevel conditions in concrete.

During the experiments, there were noticeable changes in pH (1 to 2 units) that required addition of reagents to restore the solution to constant pH conditions. For the sodium nitrite solutions, there was a general increase in pH value for initial pH values 8 and 10. This can possibly be attributed to the formation of $OH^-$ anions in solution from hydrolysis involving $NO_2^-$ with $H_2O$. For MFP solutions, there was a decrease in pH value for solutions at initial pH values of 10 and 12. A possible process to account for this observation is hydrolysis of MFP in solution to produce HF and hydrogen phosphate species. No significant change was noted in the pH value of the borate solutions during the experiments.

## COMPARISON OF VISUAL INSPECTION WITH SURFACE ANALYSIS

In related studies using a screening test based on the visual estimation of corrosion (Dressman et al., a companion paper in this Record), it was found that sodium nitrite, MFP, and sodium tetraborate all exhibit good corrosion-inhibiting properties for rebar in chloride-doped pore solution. The twin procedures of visual and electrochemical screening provide complementary results, and together suggest that rapid laboratory screening of candidate inhibitors is a worthwhile endeavor. Surface analysis measurements (Dillard et al., a companion paper in this Record) suggest that sodium tetraborate reacts to form a coating on the rebar surface. This coating could possibly function as a barrier layer to inhibit chloride-induced corrosion, thereby increasing the electrochemical potential values as noted in the present experiments. Surface analysis of MFP-treated rebar suggests that hydrolysis of MFP is a likely cause of the resultant phosphorus-to-fluorine atomic ratio of nonunity on the rebar surface. This idea is supported by the decrease in pH value noted for specimens at initial pH values 10 and 12.

## CONCLUSIONS

The effect of corrosion inhibitors for rebar under highly alkaline conditions can be directly studied by electrochemical methods. Specifically, side-by-side measurements of the electrochemical potentials of rebar in chloride-doped pore solutions exhibit the protective effect of added inhibitors compared with uninhibited solutions. The effect is manifested by the corrosion potential becoming progressively more positive over a period of weeks when inhibitor is present, or has been added.

Therefore, controlled electrochemical studies under specified conditions can serve as screening tests for candidate corrosion inhibitors. This latter conclusion is substantially validated by the fact that known corrosion inhibitors performed well in the test procedure. Sodium nitrite and MFP caused the corrosion potential to become more positive under the selected test conditions. Sodium tetraborate performs similarly—a significant result, because it has recently been demonstrated by visual inspection corrosion evaluation techniques (Dressman et al., a companion paper in this Record) that sodium tetraborate is a promising inhibitor.

Having two complementary screening tests available for studies of candidate inhibitors means that the selection of the most promising materials can be expedited. Thus, sodium tetraborate is a good candidate as a practical corrosion inhibitor for rebar corrosion. The coating-forming properties of sodium tetraborate (Dillard et al., a companion paper in this Record) on rebar may provide the mechanism by which it works. Furthermore, this inhibitor can be expected to diffuse through concrete cover to rebar at depth in a reinforced structure (Dillard et al., a companion paper in this Record).

The present results confirm the well-known fact that corrosion potentials are dependent on the solution pH value. Beyond that, different inhibitors have different relative effectiveness at varying pH values. Thus, sodium nitrite is an effective inhibitor over the pH range 8 to 13. In contrast, sodium tetraborate is most effective at pH 13, and there is some evidence to suggest that it is the best of the tested inhibitors at that pH value.

The precise properties of the rebar, either its configuration or chemical composition, are not critical in comparing inhibitors. Providing side-by-side testing is carried out with all parameters except inhibitor held constant, a fair measure of relative effectiveness can be obtained.

The results demonstrate that corrosion inhibition is a function of the inhibitor concentration. Experiments at higher inhibitor concentrations indicate that the electrochemical potential more rapidly reverts to the voltage range of low corrosion probability.

This work deserves extension to a wide range of inhibitors at various concentrations. Scaled-up tests of promising materials have already been initiated on the basis of the data reported in this work.

## ACKNOWLEDGMENTS

Thanks are expressed to Frank Cromer, who helped in the surface analysis measurements; to the National Science Foundation and the Commonwealth of Virginia for surface analysis equipment grants; and to the Strategic Highway Research Program, Project C–103, for support of this work. The authors are grateful to a referee who carefully read and provided valuable comments on this manuscript.

## REFERENCES

1. *Durability of Concrete Bridge Decks*. Report 5. Portland Cement Association and U.S. Bureau of Public Roads, 1969, 46 pp.
2. *Durability of Concrete Bridge Decks*. Final Report. Portland Cement Association and U.S. Bureau of Public Roads, 1970, 33 pp.
3. *Corrosion of Reinforcement in Concrete Construction*. (Alan P. Crane, ed.) Ellis Horwood Ltd., Chichester, England, 1983, 437 pp.

4. *Corrosion of Steel in Concrete*. Report of the Technical Committee 60-CSC, RILEM (The International Union of Testing Research Laboratories for Materials Structures), P. Schiessl, ed., Chapman and Hall, London, 1988, 102 pp.
5. K. Tuutti. *Corrosion of Steel in Concrete*. Swedish Cement and Concrete Research Institute, Stockholm, pp. 17–101, 1982.
6. M. G. Fontana and N. D. Greene. *Corrosion Engineering*. McGraw-Hill, New York, 1967, pp. 198–204.
7. H. H. Uhlig and R. W. Revie. *Corrosion and Corrosion Control*. John Wiley, New York, 1985, pp. 263–277.
8. J. M. West. *Electrodeposition and Corrosion Processes*. Van Nostrand–Reinhold, New York, 1971, pp. 138–148.
9. J. T. Lundquist, A. M. Rosenberg, and J. M. Gaidis. Calcium Nitrite as an Inhibitor of Rebar Corrosion in Chloride Containing Concrete. *Materials Performance*, Vol. 18, 1979, pp. 36–40.
10. A. M. Rosenberg and J. M. Gaidis. The Mechanism of Nitrite Inhibition of Chloride Attack on Reinforcing Steel in Alkaline Aqueous Environments. *Materials Performance*, Vol. 18, 1979, pp. 45–48.
11. N. S. Berke and P. Stark. Calcium Nitrite as an Inhibitor: Evaluating and Testing for Corrosion Resistance. *Concrete International: Design and Construction*, Vol. 7, 1985, pp. 42–47.

---

*The opinions, findings, and conclusions herein are those of the authors and not necessarily those of the sponsoring agencies.*

*Publication of this paper sponsored by Committee on Corrosion.*

# Analysis of Energy Dissipation Caused by Snow Compaction During Displacement Plowing

ANDREW C. HANSEN

Energy dissipation during displacement plowing of snow caused by snow compaction in front of the plow is estimated. Depending on the speed of the plow, a plastic wave will form in front of the plow, thereby compressing the snow before removal. This compression wave dissipates energy provided by the prime mover. An approximate solution of the compression zone was developed to determine the significance of the energy dissipation caused by a plastic wave. The results indicate that energy losses may be substantial during high-speed plow operations. A parametric analysis of parameters involving plow geometry, operating speeds, and snow properties was conducted. Conditions under which energy losses caused by compression may be either reduced or eliminated are clearly identified.

The power requirements for displacement plowing of snow may be roughly divided into two categories: (a) power required to cast the snow off to the side, and (b) power dissipated as a result of snow compaction in front of the plow. The latter requirement has received considerably less attention in the literature as the primary objective of plowing snow is the actual removal of material from the roadway. Furthermore, these problems are fundamentally different in that the first is essentially a hydrodynamic problem in which the snow is in a fluidized state, whereas the second requires a detailed knowledge of high-rate mechanical properties of snow.

Power is dissipated when plowing snow as the result of compaction of the snow caused by a plastic (nonlinear) wave running out in front of the plow. The presence of a plastic wave in snow resulting from impact loading has been documented analytically and experimentally. Wakahama and Sato (*1*) conducted drop weight impact tests of snow using a 1-kg plate with impact velocities of 2.5 to 5 m/sec. High-speed photography, in conjunction with streak lines on the sample, was used to determine the wave speeds. The photography clearly indicates a nonlinear wave extending out from the impact plate. For instance, for an impact velocity of 4.3 m/sec, the plastic wave velocity was found to range from 6.5 m/sec for an initial snow density of 200 kg/m$^3$ to 12 ± 2 m/sec for an initial density of 400 kg/m$^3$.

On the other end of the loading spectrum, Napadenski (*2*) conducted stress wave experiments in snow by explosively loading flier plates into a snow sample. Results of these data show plastic wave speeds reaching as high as 170 m/sec. Brown (*3–5*) has done a substantial amount of theoretical work aimed at studying nonlinear waves in snow using a volumetric constitutive law. The theoretical results for wave speeds were found to be in agreement with data presented by Napadenski.

The dynamics of displacement plowing of snow can be considered similar to those of drop weight impact testing as well as explosively induced stress wave experiments. The main difference is that there is a continuous energy supply available to the driver plate, i.e., the plow itself. Furthermore, the operating speeds of a displacement plow will lie somewhere in between those of the low-velocity impact tests of Wakahama and Sato (*1*) and the high-rate explosive loadings of Napadenski (*2*). Therefore, it is reasonable to assume the existence of a plastic wave running in front of a displacement snow plow.

Under the assumption of a plastic wave does exist, estimating the amount of energy dissipated by compression of the snow as the wave passes through the material is desirable. For analysis of the compression zone, an analytical formulation has been developed on the basis of an assumed one-dimensional deformation.

## ANALYSIS OF THE COMPRESSION ZONE

Consider a displacement plow moving forward with velocity $V_0$ as shown in Figure 1. The region to be analyzed is taken from the compression zone boundary forward into the undisturbed snow. This region is defined by the Lagrangian (reference) coordinate $X$. Figure 1a is defined at time $t = 0$. Hence, the material in the positive $X$ direction is in its undisturbed state.

At a later time, $t^*$, the plow has reached the snow particles originally located at $X = 0$, Figure 1b. The total zone of compression referenced to Lagrangian coordinates is given by $X_c$.

### Modified Sheet Model

First, a modified sheet model analysis is performed to predict the forward pressure applied to the snow by the plow. The sheet model in general neglects the effects of snow compaction and surface friction or bonding. These assumptions produce a forward pressure on the plow given by

$$p = \rho_0 * V_0^2 [\cos^2 \beta(1 - \cos \phi)] \quad (1)$$

Department of Mechanical Engineering, University of Wyoming, P.O. Box 3295, University Station, Laramie, Wyo. 82070.

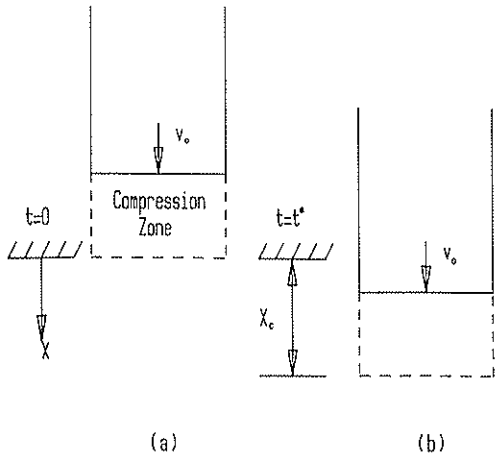



**FIGURE 1 Schematic of the compression zone leading a displacement plow.**

where $\beta$ is the swivel angel, $\phi$ is the cutting angle, and $\rho_0$ is the initial density, as shown in Figure 2.

Now it is assumed that a compression wave is running in front of the plow, thereby compacting the snow before removal. The compression wave imparts a forward velocity to the snow before the arrival of the plow. The sheet model may still be used to predict the forward pressure acting on the plow by modifying the density and the velocity. Let $V$ denote the velocity of the plow with respect to the snow after compression, and $\rho$ the corresponding density of the snow. From conservation of mass,

$$\rho_0 V_0 = \rho V \tag{2}$$

Hence, allowing for compression, the velocity of the snow with respect to the plow is given by

$$V = (\rho_0/\rho)V_0 \tag{3}$$

On substitution of the relative velocity of Equation 3 and the corresponding density into Equation 1, the forward pressure predicted by the sheet model becomes

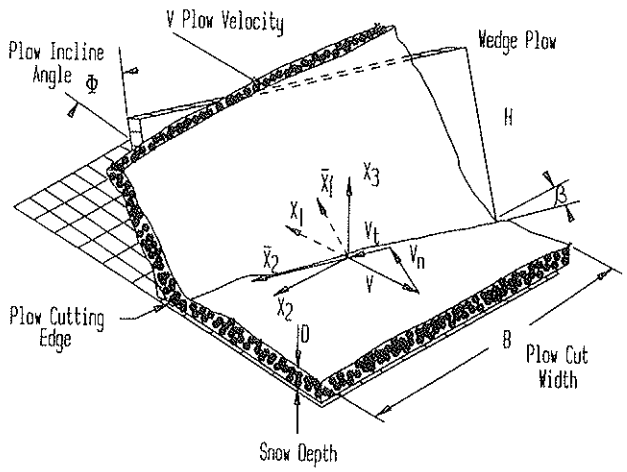



**FIGURE 2 Simplified plow model.**

$$p = \rho(\rho_0/\rho)^2 V_0^2 [\cos^2 \beta(1 - \cos \phi)] \tag{4}$$

Therefore, from Equation 4 for a given plow velocity, $V_0$, the forward pressure acting on the snow may be computed as a function of the density, as shown in Figure 3. In this figure, the forward pressure applied to the snow by the plow actually decreases with increasing density caused by a compression. This decrease may be attributed to the decrease in the relative velocity of the plow with respect to the snow.

### Power Requirements for Compressing Snow

Now consider the power dissipated by compacting snow in the compression zone. The volumetric property of snow is expressed in terms of a dimensionless density ratio given by

$$\alpha = \rho_m/\rho \tag{5}$$

where $\rho_m$ is the density of the matrix material (ice) and $\rho$ is the density of the snow.

As shown in Figure 1a, the deformation is assumed to be strictly one dimensional. Hence, the Cartesian components of the deformation gradient are given by

$$F_{iJ} = \begin{vmatrix} \frac{\partial x}{\partial X} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{6}$$

where $x$ represents the deformed coordinate and $X$ represents the reference coordinate.

The Lagrangian form of conservation of mass is given by

$$\frac{\rho_0}{\rho} = \det \mathbf{F} = \frac{\partial x}{\partial X} \tag{7}$$

Equation 7 may be expressed in terms of the density ratio as

$$\frac{\alpha}{\alpha_0} = \frac{\partial x}{\partial X} \tag{8}$$

The stress power (power dissipation) per unit *undeformed* volume is given by $\mathbf{T} \cdot \cdot \dot{\mathbf{F}}$ where $\mathbf{T}$ is the First Piola-Kirchhoff stress tensor. This tensor is related to the Cauchy (engineer-

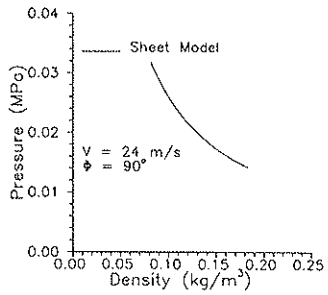



**FIGURE 3 Pressure versus density at the cutting edge of a plow on the basis of a modified sheet model analysis.**

ing) stress tensor by

$$\mathbf{T} = \frac{\alpha}{\alpha_0}\mathbf{F}^{-1}\mathbf{t} \tag{9}$$

where $\mathbf{t}$ is the Cauchy stress. Noting Equations 7 and 8, the components of $\mathbf{T}$ are

$$T_{ij} = \begin{vmatrix} t_{11} & t_{12} & t_{13} \\ \frac{\alpha}{\alpha_0}t_{12} & \frac{\alpha}{\alpha_0}t_{22} & \frac{\alpha}{\alpha_0}t_{23} \\ \frac{\alpha}{\alpha_0}t_{12} & \frac{\alpha}{\alpha_0}t_{22} & \frac{\alpha}{\alpha_0}t_{33} \end{vmatrix} \tag{10}$$

For the assumed deformation, the axial stress is to a good approximation equal to the hydrostatic pressure, $p$. Therefore, differentiating Equation 6 with respect to time and noting Equations 8 and 10, the stress power may be expressed as

$$P = p(\dot{\alpha}/\alpha_0) \tag{11}$$

The total energy dissipated per unit volume in compressing the snow is obtained by integrating the stress power over time, i.e.,

$$E = \int_0^t p\,\frac{\dot{\alpha}}{\alpha_0}\,ds \tag{12}$$

Finally, the power dissipated by the compression wave is obtained by multiplying Equation 12 by the volumetric flow rate of snow removal.

In order to close the equations, a volumetric constitutive law must be specified providing a relation between the pressure $p$ and the density ratio $\alpha$. The form taken is

$$p = F \ln(\dot{\alpha}/A_1) \tag{13}$$

where

$$F = \frac{C_1\rho_0(\alpha/\alpha_0)^{C_2+C_3\alpha_0}}{B_1\alpha}$$

and $A_1$, $B_1$, and $C_i$ $(i = 1, 2, 3)$ are constants. The constitutive law is a simplified version of a microphysical model developed by Brown (*6*) and works well for snow of low density.

### Approximate Solution

An exact solution of the energy dissipation given by Equation 12 requires knowledge of the volumetric rate of deformation $\dot{\alpha}$ as the pressure wave advances through the material. This information may only be determined by carrying out a complete solution of the wave propagation problem in the compression zone. Such an analysis is currently in progress and is based on a two-dimensional finite difference wave propagation code for finite deformation of snow.

In order to obtain some immediate quantitative results, an approximate solution has been formulated by assuming a constant value for the volumetric rate of deformation. In reality, $\dot{\alpha}$ varies in a continuous fashion as the compression wave advances. Fortunately, the constitutive behavior of snow is only weakly rate dependent, as demonstrated by Equation 13 in which the pressure is shown to vary with the natural log of $\dot{\alpha}$. As a result, an order of magnitude error in the value of $\dot{\alpha}$ through the compression zone results in only a 10 percent difference in the total energy dissipated in the approximate solution. Hence, the results predicted by assuming $\dot{\alpha}$ is constant throughout the deformation are felt to be a good measure of the exact analytical solution.

The approximate solution of the compression problem is begun by assuming values for the plow velocity, plow cutting angle, and initial snow density. The modified sheet model is then used to determine a curve for pressure versus density similar to that shown in Figure 3. Next, a curve for pressure versus density predicted by the constitutive law is generated using Equation 13 and superimposed on the sheet model analysis, as shown in Figure 4. The energy dissipated per unit volume is then determined from Equation 12, in which the integration is carried out until the pressure and density predicted by the constitutive law agree with the corresponding values predicted by the modified sheet model. These values are determined by the crossover point shown in Figure 4.

The rate of deformation $\dot{\alpha}$ was determined by assuming

$$\dot{\alpha} = \alpha_0\,\dot{g} \tag{14}$$

where $\dot{g}$ is the ratio of the plow velocity divided by the length $X_c$ of the compression zone. Physically, Equation 14 corresponds to the deformation rate for a spatially homogeneous uniaxial compression test. The length $X_c$ of the compression zone was taken to be 0.5 m for all calculations; this assumption in turn fixes $\dot{\alpha}$, because of Equation 14. As stated previously, the assumed value for $\dot{\alpha}$ has little effect on the energy dissipation.

There are conditions under which the two curves shown in Figure 4 do not intersect. This situation occurs when the initial pressure predicted by the constitutive law is greater than that of the sheet model. Therefore, the critical pressure needed to cause compression is not achieved and the snow behaves in a sheet-like manner during removal.

## RESULTS

In what follows, estimates of power requirements for plowing snow caused by the compression wave leading the plow are

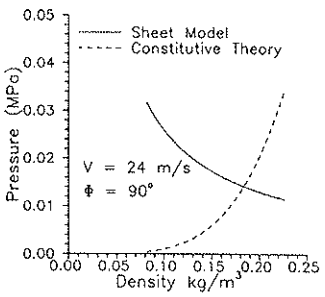



**FIGURE 4 Pressure versus density as predicted by the sheet model and the constitutive law.**

described. Power is expressed in units of horsepower, the most familiar unit applied to vehicles. Parameters that are varied are limited to the initial density, velocity, and plow cutting angle.

In the following data, the plow width was assumed to be 3.33 m with a swivel angle of $\beta = 33$ degrees, Figure 2. The plow velocities considered were allowed to vary from 9 to 30 m/sec (20 to 66 mph). This range covered the entire spectrum of low-speed city plowing up to high-speed highway operations. The snow depth is assumed to be 15 cm. A range of initial snow densities was considered from 60 to 160 kg/m³. This range represents low densities that may be encountered in a fresh mountain snowfall up to higher values associated with wet snow or snow that has been allowed to sinter for some time. For instance, on February 13, 1990, approximately 20 cm of snow fell at Laramie, Wyoming. Laramie is located at an elevation of 2190 m and has an extremely dry climate. Density measurements taken for the snowfall were found to range from 76 to 108 kg/m³ for three different samples.

Figure 5 shows a plot of horsepower versus velocity for an initial density of 60 kg/m³ and cutting angles of 90 and 70 degrees, respectively. The figure shows that for either case, the power lost to compression is negligible for velocities under 9 m/sec (20 mph). However, as the velocity increases beyond 9 m/sec the power lost to compression increases significantly. For instance, at 24 m/sec (54 mph) the total power dissipated reaches 17 hp for the vertical cutting angle. The lost power represents a significant portion of the total power available to the prime mover. The figure also shows a reduction in the energy dissipation is achieved by laying the cutting angle back to 70 degrees. For instance, the horsepower drops from 17 to 13 hp at 24 m/sec, representing a reduction of 24 percent in the power requirements for these particular high-speed plowing conditions.

Figures 6–8 show plots similar to Figure 5 for initial densities of 80, 120, and 160 kg/m³. The figures indicate trends similar to those of Figure 5. In general, there is little or no energy dissipated for plowing speeds up to 9 m/sec (20 mph). Also, the benefits from the reduced cutting angle become more significant as the initial density increases. For instance, at 24 m/sec the power decreases from 17 to 13 hp for low-density snow of 60 kg/m³. In contrast, for an initial density of 160 kg/m³, the power drops from 57 to 41 hp, representing a reduction in energy requirements of nearly 28 percent.

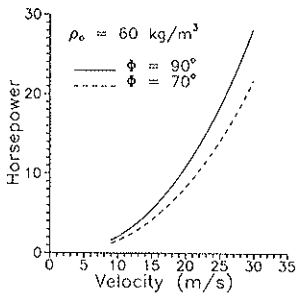



**FIGURE 5 Horsepower versus velocity for cutting angles of 70 and 90 degrees and an initial density of 60 kg/m³.**

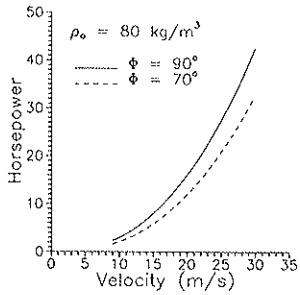



**FIGURE 6 Horsepower versus velocity for cutting angles of 70 and 90 degrees and an initial density of 80 kg/m³.**

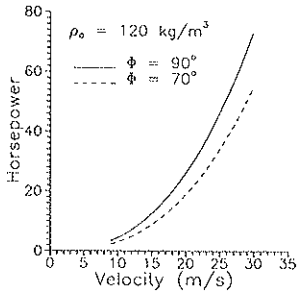



**FIGURE 7 Horsepower versus velocity for cutting angles of 70 and 90 degrees and an initial density of 120 kg/m³.**

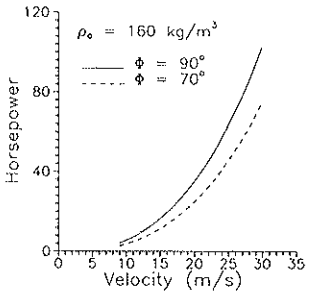



**FIGURE 8 Horsepower versus velocity for cutting angles of 70 and 90 degrees and an initial density of 160 kg/m³.**

Figure 9 shows a parametric plot of power requirements versus velocity for initial densities of 60, 80, 120, and 160 kg/m³, respectively. The figure dramatically shows the effect of initial density on the energy requirements.

Figure 10 shows a parametric plot of power versus velocity for cutting angles ranging from 50 to 90 degrees at an initial density of 80 kg/m³. The results indicate that significant energy reductions are possible by laying the plow angle back to 50 degrees. For example, at 24 m/sec (54 mph) the power consumed by compression decreases from 25 hp for $\phi = 90$

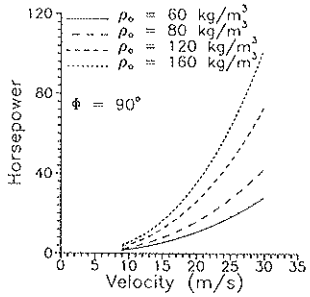



**FIGURE 9 Parametric plot of horsepower consumed by compression versus velocity for various initial densities.**

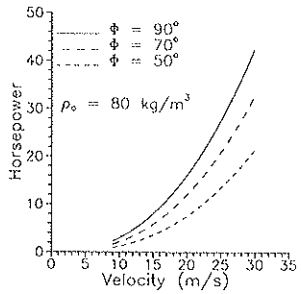



**FIGURE 10 Parametric plot of horsepower consumed by compression versus velocity for various cutting angles.**

degrees to 12 hp for $\phi = 50$ degrees, representing a 52 percent reduction in energy dissipation.

Figure 11 shows a plot of the final density versus velocity for initial densities ranging from 60 to 160 kg/m³. The density changes shown in the figure represent measurable changes caused by the advancing compression wave. Hence, it may be possible to verify these changes experimentally.

## CONCLUSION

The results of the approximate solution discussed in the previous section indicate that energy dissipation caused by a compression wave is negligible at speeds below 9 m/sec (20 mph). However, energy dissipation may be significant at higher speeds typically found in highway plow operations.

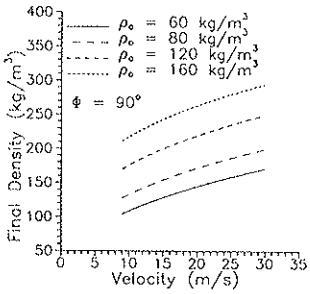



**FIGURE 11 Parametric representation of the final density as a function of velocity for various initial densities.**

The significant parameters affecting energy dissipation caused by compression are the initial snow density, plow cutting angle, and plow velocity. Of these parameters, only the cutting angle is readily varied without affecting plow operations. The data suggest that by laying the cutting angle back, significant reductions in energy dissipation are possible. Furthermore, in addition to energy dissipation, the compression wave leading the plow adversely effects the cast distance as the exit velocity of the snow is reduced.

The primary adverse effect of laying the cutting angle back lies in safety. In particular, proper tripping of the blade when the plow encounters hidden obstacles becomes increasingly difficult as the angle is laid further back.

A similar analysis is currently underway to estimate compression losses for plowing snow that has been precompacted by vehicle traffic. The primary variable that must be changed in the analysis is the initial density. For this case, densities typically will lie in the range of 350 to 450 kg/m³. For instance, initial densities and compressed densities were taken for a snowfall in Laramie, Wyoming, on February 13, 1990. The data for the vehicle-compacted snow ranged from 354 to 400 kg/m³. The densities resulting from vehicle compaction are consistent with results from plate sinkage tests for a deep snow cover obtained by Hansen (*7*).

Work is continuing on developing a fundamentally sound experimental procedure to measure the density changes caused by a compression wave leading the plow. An approach currently being investigated is to use high-speed photography, in conjunction with streak lines placed on the snow, to observe the advancing compression wave. The density and pressure changes caused by the wave may then be determined from conservation of mass and momentum principles in the form of the Rankine-Hugoniot relations (*1*).

## ACKNOWLEDGMENT

This research was sponsored by the Strategic Highway Research Program Project H-206.

## REFERENCES

1. G. Wakahama and A. Sato. Propagation of a plastic Wave in Snow. *Journal of Glaciology*, Vol. 19, No. 81, 1977, pp. 175–183.
2. H. Napadenski. Dynamic Response of Snow to High Rates of Loading. Research Report 119, Cold Regions Research and Engineering Laboratory, Hanover, N.H., 1964.
3. R. L. Brown. Pressure Waves in Snow. *Journal of Glaciology*, Vol. 25, No. 91, 1980, pp. 99–107.
4. R. L. Brown. An Analysis of Non-Steady Pressure Waves in Snow. *Journal of Glaciology*, Vol. 25, No. 92, 1980, pp. 279–287.
5. R. L. Brown. Propagation of Stress Waves in Alpine Snow. *Journal of Glaciology*, Vol. 26, No. 96, 1980, pp. 235–243.
6. R. L. Brown. A Volumetric Constitutive Law Based on a Neck Growth Model. *Journal of Applied Physics*, Vol. 51, No. 1, 1980, pp. 161–165.
7. A. C. Hansen. *A Constitutive Theory for High Rate Multiaxial Deformation of Snow*. Ph.D. dissertation, Montana State University, Bozeman, 1985.

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Influence of Wind, Temperature, and Deicing Chemicals on Snow Accretion

E. E. ADAMS, R. G. ALGER, AND J. P. BECKWITH

Tests run for natural snow deposition on asphalt pavement treated with highway and airport runway deicing chemicals demonstrate that, for conditions of strong wind and cold temperatures, these chemicals can have a deleterious effect. Friction is assumed to be the primary criterion for judging chemical efficacy. Statistical correlations with meteorologic data support observations that some conditions that produced scouring on nonchemically treated pavement caused snow to accumulate on areas that had been treated with deicing chemicals.

The use of chemicals is a vital aspect of snow and ice control in many nations of the world. In addition to the direct cost, however, the adverse environmental effects of these chemicals on the transportation infrastructure is a serious concern. Since 1975, an estimated 10 million tons of road salt has been used annually in the United States alone, causing corrosion of cars, roads, and bridges; adverse influences on vegetation; and contamination of water supplies. Noncorrosive, environmentally safe alternative deicers have higher initial prices, although this may be offset by the long-term full cost of salt. In any case, efficient, effective application of chemicals is essential from the standpoint of monetary and environmental concerns.

The use of measured friction values as an objective parameter for judging the efficacy of chemicals in deicing and anti-icing field tests in Michigan produced some seemingly anomalous results in an objective statistical ranking of the effectiveness of the independent test sections. However, further analysis using meteorologic data exhibited a correlation between wind speed, temperature, and the effectiveness of deicing chemicals. This result supports observations made in this study and mentioned in the literature (*1*) that the use of chemicals in conjunction with wind-transported snow can induce snow accumulation on pavement that would otherwise have been scoured bare.

## OBJECTIVES

The objective of the field study was to compare the efficacy of deicing and anti-icing chemicals in a realistic, but controlled, well-monitored environment. Chemicals were applied to large asphalt test sections under naturally occurring snow conditions. This work considered proprietary research materials, as well as commercially available deicing products. In one aspect of the study, reported here, an interaction of wind and temperature on the deicing chemicals is apparent. Because the subject phenomenon was observed among all the test deicers, it is not necessary for the purposes of this discussion to identify specific chemicals. This conclusion differs somewhat from the observations of Fromm (*2*) who conducted studies on wetting of NaCl with $CaCl_2$ solutions. He observed that the wetted salt caused blowing snow to stick, but that the untreated salt allowed the road to dry.

Because surface friction is critically important to vehicle traction, handling, and safety, it was used as the primary measure of chemical efficacy. This paper addresses only the influence of wind speed and temperature with the use of deicing chemicals.

## TEST PROCEDURE AND EQUIPMENT

Two testing areas designed to simulate highway and airport operations were established on an unused portion of asphalt runway at the Houghton County Airport, Houghton, Michigan. Each test area contained two separate parallel lanes to allow for different maintenance operation tests to be made during a single test. All test sections were 30.5 × 7.3 m (100 × 24 ft). In order to mitigate cross contamination, these sections were separated from each other along the path of vehicle motion by a 61.0-m (200-ft) buffer zone. A map view of these sections reveals a checkered pattern with four test sections used for vehicle motion in one direction and three sections staggered and offset along the 61.0-m (200-ft) buffer zones for returning traffic. This checkered pattern is apparent in the background shown in Figure 1.

On the highway course, the two separate lanes were divided into a lane on which traffic was simulated using a roller built for this purpose, while the other lane and the airport lanes were left undisturbed except for the occasional passage of the Saab Friction Tester. The friction tester measured values by means of a rotating fifth wheel with a constant 12 percent slip relative to the vehicle speed. The measuring wheel was connected by a chain drive to the freely rotating rear axle of a Saab automobile. A computer recorded a friction measurement taken every 1 m (3 ft) and presented an average of these values taken over each 30.5-m (100-ft) test section. An attempt was made to run the Saab on approximately a 15-min interval. From these friction measurements, the effectiveness of each test section could be viewed in comparison with the others.

The application rates for highway chemicals were chosen to compare with highway maintenance standards. Road salt per application is frequently applied at a rate of 23 g/m$^2$ (300

E. E. Adams and R. G. Alger, Keweenaw Research Center, and J. P. Beckwith, Mathematics Department, Michigan Technological University, Houghton, Mich. 49931.

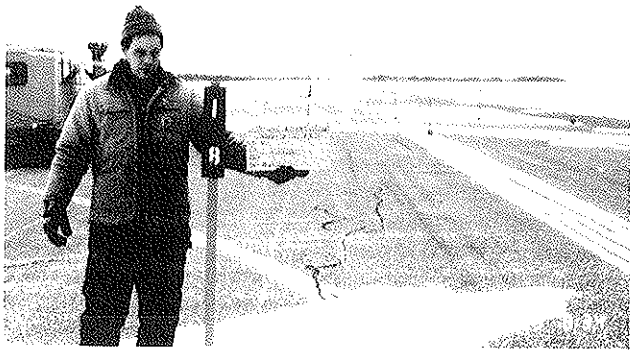


**FIGURE 1 Foreground, Test Section 8 was nontreated section. Background, chemically treated sections are offset rectangular patches. Snow accumulation is apparent on chemically treated test sections, whereas untreated pavement has been scoured.**

lb per linear mile) in actual highway maintenance operations, so rates similar to this were used for the test. The airport test area was used to simulate chemicals applied in airport maintenance activities. The major difference between this procedure and the highway scenario was that traffic was not simulated and sweepers were sometimes used for snow removal. Rates were chosen for these chemicals that were more typical of airport operations. Application rates were up to three times that used on the highway.

Chemicals were applied simultaneously by using individual walk-behind spreaders. This application was carried out either in anticipation of a storm, in an attempt to evaluate anti-icing properties, or immediately after plowing any accumulated snow. The sections were prepared for each test by plowing any accumulated snow. This plowing was performed to determine the effect of chemicals applied before an anticipated precipitation event, or to observe residual chemicals present from previous tests. Immediately after plowing, measurements were made using the Saab friction tester. The chemicals were usually applied on the test sections after the initial plowing, although this was sometimes delayed or not done, to observe the residual chemical effect after the overburden was removed. Throughout testing, plowing was performed as needed. When reapplication of chemicals was deemed prudent, it was always carried out immediately after plowing. In any event, the Saab was used to bracket any maintenance operation.

Meteorological data collected during testing were ambient air temperature, snow surface temperature, bare pavement temperature (when exposed), wind speed, wind direction, relative humidity, incoming infrared (long wavelength, 4- to 50-μm) and both incoming and reflected solar (short wavelength, 0.28- to 2.8-μm) radiation. Ambient air temperature and relative humidity were measured at approximately 3 m, and wind data at 4 m above the ground, on the roof of the mobile laboratory. Ambient temperature was obtained using a shielded thermocouple; the snow and pavement temperatures were measured using a shielded thermocouple probe. Data from the airport station, collected 24 hr a day, were used to complement these data.

## DISCUSSION OF RESULTS

Friedman statistical tests (*3*) were conducted to analytically check for differences in the distributions of the friction measurements. These statistical tests were conducted on the results of 22 days for which data were gathered during January through March of 1990. The Friedman test is a nonparametric two-way analysis of variance (ANOVA) test. It treats each test time within a given day as a block in which it ranks the friction coefficient values. The ranks of a treatment over that day are then summed and compared to those of the other treatments. The analysis is based on the hypothesis that there is no difference in the distributions and that it determines the probability that the hypothesis is correct. This test was repeated for each test day.

Using this method, there were a number of days for which the section on which no chemical was applied resulted in superior performance. The fact that the nonchemically treated section worked, as well as or better than all of the treated sections (in some instances) was at first somewhat disconcerting, because it did not coincide with the intuitively expected results. Meteorologic conditions, however, afford an explanation and should in fact lead to methods that help optimizing the application of deicing chemicals.

Figure 1 shows a dramatic example of an instance in which higher relative friction values were recorded on the nonchemically treated sections. Keeping in mind the pattern in which chemicals were applied and that the clear test section shown in the foreground was a nonchemically treated area, note that the checkered pattern of snow accumulation is on the chemically treated surfaces. During the field testing, this situation was often associated with windy conditions in which the nonchemically treated sections and in fact the runway in general were wind-scoured and bare, whereas areas that were chemically treated tended to exhibit snow accumulation, a fact that had been noted by others as well (*1*).

### Physical Motivation

Although wind is a dominant mechanism, it alone does not determine when an untreated pavement would necessarily result in higher frictional values than those of the chemically treated sections. Other interacting conditions are required as well and the situation is more involved than that discussed here. An obvious requirement is that there be snow or ice particles available to accumulate. This mass supply may be in the form of precipitating, drifting, and saltating (i.e., ice grains bouncing along the surface) snow or often a combination of these. In the case of redeposited snow, there is a threshold wind velocity necessary to produce a shear stress on the snow surface sufficiently large to cause particles to be transported. This threshold velocity is highly dependent on the texture of the snow surface, which in turn is dependent on antecedent meteorologic conditions.

In addition to wind velocity and previous environmental conditions, another parameter that greatly influences the snow scouring-deposition relation is temperature. In the presence of a cold, dry environment, saltating ice particles will act in a manner similar to the drifting of dry sand. This behavior is shown in Figures 2 and 3, in which these conditions prevailed.

**FIGURE 2 A "river" of wind-transported snow, flowing along the asphalt road surface and obscuring the pavement.**

**FIGURE 3 When drift intensity has momentarily decreased below that shown in Figure 2, snow is not accumulating on pavement.**

Regions of scouring and deposition in this situation are governed predominantly by wind velocity and the terrain geometry. At warmer temperatures, however, at which the ice is relatively close to its temperature of phase change, grains that impact the untreated pavement are more likely to adhere to and accumulate on the surface. Extreme cases of this would be freezing rain and sleet.

In contrast, when the situation was examined in which scouring was occurring on the untreated pavement surface, an obvious residual effect of the chemicals was apparent on the treated sections, as evidenced by the snow accumulation pattern in Figure 1. It is important to note that this condition occurred even when the treated sections had not been recently treated and were bare at the start of the precipitation event. These chemicals are freezing point depressants, and therefore, act to effectively lower the temperature of phase transition for the ice. This effect results in some melting and adhesion to the pavement surface. If the rate of melting is insufficient to keep pace with the quantity of snow being transported, a net accumulation results.

This effect was essentially the same for all of the chemicals tested. A graphic demonstration of this wind-driven snow accumulation on chemically treated pavements is shown by the photographs in Figures 4 and 5. When this scouring and accumulating situation occurred, friction on the nonchemically treated sections, naturally, was better than on the chemically treated pavements.

As stated previously, tests include instances for which chemicals were placed on the pavement during the particular storm event and for those in which no new chemicals were added. Results indicate that residual chemicals dissolved on the pavement surface during previous events were sufficient to cause snow accumulation.

The surface appearance of the accumulating snow displayed an orientation in which sharp features developed and grew into the prevailing wind. The snow was rather slush-like and all but a thin film of ice grains could be easily removed with a standard snow plow cutting edge. This thin film often presented the impression of a relatively bare pavement. However, measurements using the friction tester revealed a surface of low friction. When the causal conditions persisted, accumulation again resulted. A polyurethane cutting edge was tested in an attempt to remove this poorly bonded film, but to no advantage. The use of a sweeper was more effective but did not halt reaccumulation.

**FIGURE 4 An example of considerable snow transport occurring on test course.**

**FIGURE 5 Accumulation is taking place preferentially on chemically treated rectangular sections, as can be seen when visibility improved briefly from that of Figure 4.**

This snow scouring and accumulating condition should not be misconstrued to imply that the studies indicated that the use of chemicals in general has an adverse effect. As the photograph of Figure 6 demonstrates, the chemicals often worked more in the way desired when the type of conditions discussed here were not dominant. The chemically treated sections produced clear, wet pavement, with no melting taking place on the untreated test section or in the general vicinity.

### Statistical Analysis

To date, and thus far in this presentation, the phenomena under discussion have been based on observational information. At this point, using the data collected and the Friedman analysis, it is possible to give some statistical basis to the snow scouring and accumulating effects of the wind-temperature interaction. The results of the Friedman tests were further analyzed to see if there is a relationship between the rankings of the untreated sections and the wind speed.

Because the snow scouring and accumulating effects were apparent for chemicals applied before a specific test, it was considered likely that, in some cases, the effect of the scouring was the result of antecedent wind conditions. It was also felt that a sustained wind, rather than a maximum peak value, was more relevant. For these reasons, three different averaged wind speeds were used for the analysis. Wind speed was considered to be the maximum of the 3-hr average for the following periods: (a) 24 hr before testing—*W1*, (b) 12 hr before testing—*W2*, and (c) during the test—*W3*. These average values ranged from 0 to 10 m/sec (0 to 22 mph). The influence of temperature was examined by considering all tests and then excluding days when the average air temperature was warmer than approximately −4°C. This condition includes some instances of freezing rain.

A correlation analysis was done on the rank of the nonchemical treatment with the wind speeds during and before the test periods, as described earlier. Sample correlation coefficients are used to do this. These sample coefficients are presented in Table 1 for the four test lanes and the three different wind speed time periods. They were computed using data for 11 days of testing on the highway section and 11 days of testing on the airport section. These results are shown under the column headed Full. As an example, consider the airport inner lane, using the maximum 3-hr average of the 24 hr before the test (i.e., W1)—the sample correlation coefficient is 0.334.

The *p*-value (*4*) is the probability of observing a particular sample correlation value when there is actually zero or no correlation between the two variables. For example, on the airport inner lane, the *p*-value of 0.164 in the table is the probability that there is no correlation between friction rankings and wind speed *W1*. Next, the correlations were recomputed after removing days with relatively warm temperatures ($\geq -4$°C), 3 days from the highway and 2 days from the airport. These values are under the column headed Reduced with their *p*-values in the last column; *p*-values close to their maximum of 1 are indicated by two asterisks in the table. Except for the highway outer section when using all the test days, there are significant correlations between the rank of the nontreated section and the wind speed during at least one of the three periods. These correlations are positive values, which indicates that the nontreated section ranked higher as the wind speed increased, but it does not yield information on the actual magnitude of the velocity.

If observations are removed from the data at random, the sample *p*-value of the correlation coefficient would in general tend to increase, indicating that the relation between wind speed and ranking would be nonsignificant. In order to check for a temperature effect, data for days with relatively high temperatures were removed from the correlation. The resulting sample correlation coefficients are generally higher with smaller *p*-values. This effect is consistent with the claim that lower temperature and higher wind speed can result in better friction on untreated pavement.

Another correlation analysis (Pearson) was made using the friction values versus the weather parameters. Correlations were computed for the 7 days on which additional chemicals were applied on the highway and 8 days on which additional chemicals were applied on the airport section. Days on which only residual effects were examined were excluded. This procedure permitted the correlations to be matched by the times from the application of chemicals and by days. The statistically significant correlations for the three maximum wind speed averages for each of the treatments are presented in Table 2. Only those correlation values that had *p*-values $<0.05$ are presented in the table, except as noted by the asterisk; *p*-values $>0.05$ are generally not considered statistically significant. This table combines the airport results, but the highway is broken down into traffic and nontraffic sections.

In Table 2, the nonchemically treated section for the highway test is denoted by *T1* and for the airport by *T8*. The other *T* values denote the various chemical deicers. The correlations for the chemically treated sections with *W3* are all statistically significant and all negative; the only positive correlations occur in the nontreated sections. This result means that friction tended to decrease for each of the chemically treated sections as the wind speed increased during the period of the testing. On the other hand, although some negative correlation did occur for the nonchemically treated sections, it was only the nonchemically treated sections that had a significant positive correlation with wind speed. These cases occurred either dur-

**FIGURE 6 Example of test in which chemicals were effective in achieving desired results.**

TABLE 1 FRIEDMAN TESTS RESULTS—WIND EFFECTS (*W1* = MAXIMUM 3-hr AVERAGE FOR PREVIOUS 24 hr; *W2* = MAXIMUM 3-hr AVERAGE FOR PREVIOUS 12 hr; *W3* = MAXIMUM 3-hr AVERAGE DURING TESTING)

| Airport Inner Lane | | | | |
|---|---|---|---|---|
| Wind | Full | P-Value | Reduced | P-Value |
| W1 | 0.334 | (0.164) | 0.315 | (0.211) |
| W2 | 0.450 | (0.085) | 0.686 | (0.020) |
| W3 | 0.534 | (0.046) | 0.539 | (0.069) |
| **Airport Outer Lane** | | | | |
| Wind | Full | P-Value | Reduced | P-Value |
| W1 | 0.016 | (**) | 0.233 | (0.281) |
| W2 | 0.640 | (0.017) | 0.722 | (0.013) |
| W3 | 0.764 | (0.002) | 0.865 | (0.000) |
| **Highway Non-traffic Lane** | | | | |
| Wind | Full | P-Value | Reduced | P-Value |
| W1 | 0.328 | (0.166) | 0.584 | (0.068) |
| W2 | -0.075 | (**) | 0.080 | (**) |
| W3 | 0.391 | (0.122) | 0.616 | (0.055) |
| **Highway Traffic Lane** | | | | |
| Wind | Full | P-Value | Reduced | P-Value |
| W1 | 0.518 | (0.051) | 0.773 | (0.011) |
| W2 | 0.133 | (**) | 0.393 | (0.179) |
| W3 | 0.146 | (0.340) | 0.355 | (0.203) |
| ** = p-value near maximum value = 1 | | | | |

TABLE 2 CORRELATION OF FRICTION WITH WIND

| **Highway Treatment - Traffic** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Wind | T1 | T2 | T3 | T4 | T5 | T6 | T7 |
| W1 | 0.45 | - | - | - | - | - | - |
| W2 | - | - | - | - | - | - | - |
| W3 | -0.43 | -0.40 | -0.60 | -0.55 | -0.49 | -0.50 | -0.66 |
| **Highway Treatment - No Traffic** | | | | | | | |
| Wind | T1 | T2 | T3 | T4 | T5 | T6 | T7 |
| W1 | 0.77 | - | - | - | - | - | - |
| W2 | -0.60 | - | - | - | - | - | -0.39 |
| W3 | | -0.31 | -0.47 | -0.40 | -0.31* | 0.39 | -0.79 |
| **Airport Treatment** | | | | | | | |
| Wind | T8 | T9 | T10 | T11 | T13 | T14 | |
| W1 | - | -0.31 | -0.35 | - | -0.23 | -0.37 | |
| W2 | 0.33 | -0.40 | - | -0.30 | - | -0.36 | |
| W3 | 0.23 | -0.33 | -0.35 | -0.42 | -0.45 | -0.36 | |
| *=confidence interval of 7% | | | | | | | |

ing the 24-hr period before testing for the highway or during the 12 hr before testing for the airport section. The trend in this table further substantiates observations of the snow scouring and accumulation effect.

The implication of these results are that the effect of scouring on the nonchemically treated sections was generally dominated by wind speed which occurred during a period before testing, because it takes some time for the effect of the wind to clear the surface. On the other hand, the accumulation of snow onto the chemically treated areas was most closely related to wind speed during the removal operation. If the wind transport condition had lessened by the time of the actual testing, the snow that had accumulated because of the chemicals during the earlier high-wind period could more likely be removed.

## SUMMARY

Saab friction values were used as the primary standard for evaluating the efficacy of deicing chemicals during the winter field program. These values were examined using the Friedman test, a statistical method of ranking the relative performance of the different chemicals.

There were a number of occurrences in which the nonchemically treated test section had an overall higher performance rating. This effect was demonstrated to be statistically correlated to wind speed and temperature—which supports previous observations. The events at higher wind speeds caused decreases in measured friction values for the chemically treated sections, but scouring of snow tended to increase the friction values for the nontreated surfaces.

As a practical consideration, these data indicate that for sections of highway prone to high wind and low temperatures, chemical control methods may in fact yield a negative effect. In this situation, the consideration of snow fences might well prove to be a practical alternative as a complementary means of maintenance. It also indicates that, for specific sections of highway, chemicals should be used judiciously.

Additional work of this sort, based on quantitative measurements, may lead to specific recommendations for highway and airport maintenance operations. This work would include correlating the magnitude of the wind speed to chemical effectiveness in combination with site location, temperature, and other pertinent meteorological parameters. The final outcome of such work would be included in manuals that recommend conditions that best suit the use of chemicals before, during, and after a storm event.

## ACKNOWLEDGMENT

The authors wish to acknowledge the Deicing Technology Group of Chevron Chemical Company for support of this work.

## REFERENCES

1. J. H. Keyser. Chemicals and Abrasives for Snow and Ice Control. In *Handbook of Snow: Principles, Processes Management and Use*, (D. M. Gray and D. H. Male, eds.) Pergamon, Toronto, Ontario, Canada, 1981.
2. H. J. Fromm. *Winter Maintenance Study for Reduced Salt Usage*. Technical Report ME-82-02. Ministry of Transportation and Communications, Downsview, Ontario, Canada, Oct. 1982.
3. W. J. Conover. *Practical Nonparametric Statistics*, 2nd ed. John Wiley, New York, 1980.
4. R. Hogg and J. Ledolter. *Engineering Statistics*. MacMillan, New York, 1987.

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Goal-Oriented Design of an Improved Displacement Snowplow

ROBERT L. CRANE, MIKE H. DAMSON, AND KYNRIC M. PELL

The design process used in the development of an improved displacement plow is described. A summary of the methods used to establish the goals of a new design are outlined, as well as the design strategy developed to reach these goals. Preliminary results of the research and experimentation carried on as part of this strategy and the relation to the specific goals of the design process are outlined. The development of an experimental snowplow that serves as a platform to test the design improvements resulting from the research is described in detail.

The design process and progress in the development of an improved displacement snowplow are described. This project was initiated in October 1988 as part of the Strategic Highway Research Program Project H-206. The primary goal of this research, as presented in the project statement, is to design a displacement snowplow with 20 percent or greater energy savings over designs currently in general use.

Several methods used to obtain additional design goals that relate specifically to the end user of the snowplow included the following:

1. User Requirements Survey. Plow operators, foremen, and state engineers were surveyed in New York state in the Watertown District, in portions of Colorado, and in Wyoming. The primary concern of those surveyed was the visibility problem associated with plowing snow. Specific areas of concern included snow entrained in the wake of passing vehicles, snow/ice accumulation on the outside of the truck windows and lights, frost/ice accumulation on the inside windows of the truck, and light scattered off airborne snow by headlights. In addition, the vision of other drivers is impaired by the snow which is cast by the plow and/or the snow entrained in the wake of the truck. Equipment problems related to the low temperatures encountered during plowing operations were also mentioned. Freezing of sanders, spreaders, and hitches as well as sluggish hydraulic controls were the most common complaints. Operators mentioned that a near vertical cutting edge of the plow was necessary if ice removal was required. The operators indicated that they preferred a larger lay back angle when plowing snow. They stated that a plow with more than 10 or 15 degrees of lay back angle would ride up over ice, defeating the ice removal function of the plow.

2. State Highway Department Equipment Specifications. A survey was sent to all 50 states asking for information regarding their snowplow and truck specifications. Thirty-six replies from state departments of transportation were received and reviewed. Information regarding plow dimensions, design, and type were entered in a data base. Six state provided specifications for v-type plows, 28 states provided specifications for reversible plows, and 15 states provided specifications for one-way plows. Five weather parameters thought to be significant in snow removal were identified and representative values for each state were also put into the database. The five parameters were as follows:

- Mean relative humidity for the month of January,
- Mean annual total snowfall,
- Mean annual number of days with more than 1 in. of snowfall,
- Mean annual number of days with ice pellets, and
- Mean annual number of days with freezing rain.

Attempts to correlate the types of equipment used in a particular state with any of the listed parameters were unsuccessful.

3. Snowplow Manufacturer's Literature. Specifications of currently available snowplows, hitches, and accessories were obtained from manufacturers. This included 22 American and two foreign manufacturers.

4. Snowplow Rides. Members of the research team at the University of Wyoming spent approximately 100 hr riding with equipment operators during a winter storm. Both night and daylight rides took place. Observations made by the research team, and discussions with the operators were compiled after the rides. The visibility and equipment problems described in the user requirements survey were experienced first hand by the research team during these rides. Another result of these rides was the observation that the tripping mechanism safety feature of the snowplows was defeated by the operators. The tripping mechanism of these plows were of the full moldboard tripping style. The reason given for chaining this feature off was that the inertial forces encountered in the event of the moldboard tripping were sufficient to cause loss of vehicle control. The loss of vehicle control was evidently a larger safety concern for the operators than the ability of the plow to trip when encountering an obstacle.

5. Patent Search. A data base search of U.S. patents relating to snowplows was conducted and reviewed. US Patent Abstracts Weekly, U.S. Patents Abstracts, and World Patents Index data bases were searched for Chemical, Electrical, and Mechanical patents relating to snowplows for the dates from January 1963 through August 1989.

6. Literature Search. An extensive literature search was conducted by the research team at the University of Wyoming. A library of material relating to the project was collected and entered in a data base. Both U.S. and foreign literature was reviewed. One result of this process was the translation by

Department of Mechanical Engineering, University of Wyoming, P.O. Box 3295, University Station, Laramie, Wyo. 82071.

the Department of Mechanical Engineering at the University of Wyoming of the text *Snowplows Construction, Theory and Design*, by D. A. Shalman, published by Mashinostroenie, Leningrad, the Soviet Union, in 1973.

## CURRENT RESEARCH AND DESIGN CONSIDERATIONS

### Fluid Mechanics of Snow

The process of plowing snow produces a flow that can conveniently be separated into four distinct zones as shown in Figure 1. Zone I is the undisturbed snow upstream of the plow. Zone II is the compression zone in front of the plow. Zone III is the moldboard flow where snow is in a fluidized state. Zone IV represents the exit plume.

#### *Compression Region*

This research is primarily concerned with the energy absorbed in compressing the snow in front of the plow. In a companion paper in this Record, Hansen covers these studies in detail. The preliminary results of his research indicate that at plowing speeds greater than approximately 15 m/sec (34 mph), energy dissipation caused by compaction of snow can be a considerable portion of the energy available to plow the snow. The variables that influence this energy dissipation are the snow density, plowing speed, and layback angle of the blade relative to the road surface. Of these, the only parameter that is controllable in the design of a snowplow is the blade angle relative to the road surface. Hansen (*1*) indicates that for freshly fallen snow [density of 80 kg/m$^3$ (5 lb/ft$^3$)] at a plowing speed of 24 m/sec (54 mph) the horsepower consumed by compression drops from 83 kJ/sec (112 hp) for $\phi = 90$ degrees to 7.5 kJ/sec (10 hp) for $\phi = 50$ degrees, where $\phi$ is the angle between the blade and the road surface.

#### *Moldboard Flow*

Empirical methods have dominated snowplow design in the past. Often, full-scale testing of a design is the only method used to determine if modifications are necessary.

Analytical methods used in the past primarily involved modeling the snow as an inextensible sheet traveling over the moldboard. Developable surfaces for the moldboard were used because this should minimize losses caused by snow deformation (*1*). That is, as the snow sheet deforms, it does not tear or buckle, but flows smoothly along the moldboard. Analytical models producing design parameters for various efficiencies have been developed for a number of elementary geometries including the wedge plow (*2*), the cylindrical plow (*1*), the conical plow (*1,3*), and the modified cylindrical plow (*3*). In general, these models give good qualitative results but the theoretical values can vary from experimental values by as much as 50 percent for force and 25 percent for cast distance (*1*).

All of these analytical approaches ignore the effect of friction of the snow along the moldboard and the energy required to lift the snow against gravity. The effects of friction on the energy dissipated in high-speed plowing operations may be significant. Shalman (*4*) has reported that performance increases on the order of 25 to 33 percent have been observed with the use of low-friction moldboards.

John Nydahl of the University of Wyoming is currently incorporating the affects of friction and gravity into a new theoretical model. He is also including more generalized developable surfaces than the elementary geometries studied to date. A thin cantilevered plate theory as developed by Simmonds and Libai (*4,5*), Darmon (*6*), and Darmon and Benson

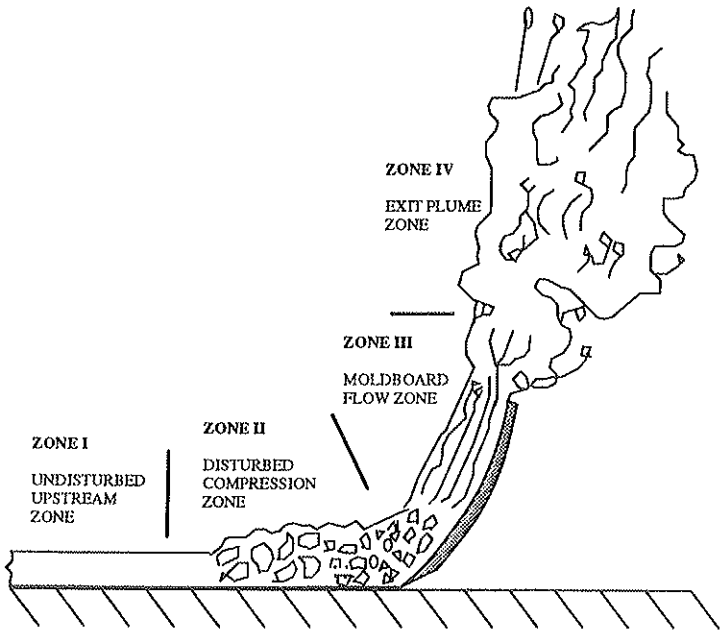



**FIGURE 1 The four flow zones.**

(*7,8*) has been investigated from which developable shapes can be numerically calculated. A case for a simple cylindrical moldboard was completed that was verified by Shalman's (*3*) work. A significant loss of velocity occurred at the exit of the moldboard that depended on the sliding coefficient of friction. These results also exhibited a significant influence of gravity in the solution.

The properties of snow vary considerably with time and temperature. Densities range from approximately 100 kg/m$^3$ (6.2 lb/ft$^3$) for freshly fallen snow to almost 500 kg/m$^3$ (31 lb/ft$^3$) for compacted snow. Friction coefficients for snow also vary widely for different temperatures and densities (*3*).

One consequence of these variable properties is that no single test can be used to determine the best plow design, because the best design will change with different snow properties. A further extension of this notion leads to the conclusion that the most favorable plow design would be one in which the geometry of the moldboard could be changed to suit changing snow conditions.

Other arguments may also be made in favor of a flexible or geometrically variable snowplow. For example, no matter which one design parameter or combination of parameters a snowplow design is optimized on, there will be plowing conditions for which the optimized plow, as far as the equipment operator is concerned, would not be optimum. For instance, the method in which the location of snow deposition if controlled in most current designs is through control of the snowplow vehicle speed. A variable geometry could allow for the snowplow vehicle to control the cast location through adjustment of the plow geometry, rather than adjustment of the plowing speed.

In addition, many relatively complex developable moldboard shapes can be obtained through use of a thin flexible plate. These developable surfaces may yield higher performance than those found in the elementary geometric shapes studied to date.

One serendipitous result of looking for a moldboard material capable of having a flexible geometry is that many of the materials that are suitably flexible also have relatively low friction coefficients. A popular material already in use in moldboard designs by several snowplow manufacturers is ultra high molecular weight polyethylene (UHMWPE). UHMWPE has excellent abrasion resistance (better than stainless steel), while also having a low coefficient of friction, high impact resistance for a plastic, good chemical resistance, and negligible water absorption (D. Walrath, informal communication). Compression-molded sheets of this material are relatively inexpensive and available in several sizes from various suppliers.

### *Exit Plume*

This research is discussed in detail by Lindberg and Petersen in a companion paper in this Record. These studies use a water tow tank containing a negatively buoyant jet apparatus that is towed at a constant velocity along the tank. Multiple experiments in which a range of experimental parameters were studied, including jet inclinations and velocities, have been run from which photographic records were analyzed.

This research will allow prediction of cast distance and cast height as functions of plow geometry and plow speed. Knowledge gained from these experiments can be applied to the variety of snow plowing conditions found operationally to gain an understanding of actual performance. Understanding of the effect on cast distances, visibility, and plume trajectories when plowing in cross-flow conditions may be used in snowplow design considerations.

### *Investigations of the Plow-Truck Airflow Field*

The investigation of the gross airflow features close to the plow-truck combination was carried out through both field experiments on a full size plow-truck using yarn tufts, and numerical two-dimensional analysis using the PHOENICS computational fluid dynamics package developed by CHAM Ltd. (*9*).

Both the field test results and the results of the two-dimensional numerical model show the locations of several recirculation zones at highway plowing speeds. Reversed flow near the surface of the hood of the vehicle, in front of the windshield, and behind the cab of the truck are evident. Also, two counter-rotating vortices extended behind the truck.

The placement of turning vanes at the top of the cab to fill in the recirculation behind the cab, and at the rear of the truck to damp the counter-rotating vortices in the trucks wake would likely improve visibility problems caused by the snow particles entrained behind the truck. Placing the vanes only at the rear of the truck would have little success, because it would be directly in the recirculation vortex caused by the cab of the truck. A deflector might also be useful on the hood of the truck to reduce some of the recirculation occurring in front of the windshield, thus reducing entrained snow, which reduces the forward visibility of the driver.

## Mechanical Aspects

This section concentrates on the progress in the mechanical design of a suitable tripping mechanism, control arms for a variable geometry moldboard, and stability of the plow and truck.

### *Plow-Truck Stability*

Computer models using the ADAMS commercial computer program (Mechanical Dynamics Inc.) are currently under development (*10*). These models investigate tripping mechanism dynamics and vehicle stability when the truck or plow encounter an obstacle during the plowing operation. Output from these models should provide information as to the forces arising on the snowplow and truck.

### *Tripping Mechanism*

The surveys of operators, manufacturers, and state departments of transportation showed that tripping mechanisms cur-

rently in use on snowplows fell into two general categories: those designs in which the entire moldboard assembly trips, and those in which only the cutting edge, or a section of the edge, is allowed to trip.

The large inertial forces arising in the case of tripping of the entire moldboard assembly tripping mechanism may cause severe safety hazards caused by potential loss of vehicle control. The coil springs used in some existing full moldboard tripping designs can have dangerous failure modes. Lack of containment of these springs during a failure could cause potentially lethal debris being projected in the path of vehicles and pedestrians. The above problems have eliminated this type of design from further consideration.

Problems with many of the cutting-edge tripping mechanisms also exist, however. If the pivot of these mechanisms is located behind the plow, the cutting edge of the plow must either dig into the road surface, or lift the entire mass of the plow, or both, to clear an obstacle. This condition can cause severe gouging of roadways, especially on open graded surfaces. Large forces are imposed on the truck, degrading control. This process can also cause increased wear on the cutting edge surface.

## INCORPORATION OF RESULTS INTO EXPERIMENTAL SNOWPLOW

An experimental plow that will serve as a test unit for the application of the goals of the project and the verification of research results is currently under development. The main design features of the experimental plow are a unique tripping edge, a front edge snow scoop, and a variable-geometry moldboard. These features are being added to a surplus drive frame acquired by the research group from the Wyoming Highway Department.

### Variable Geometry Moldboard

The moldboard material chosen was a 4- × 12-ft, ⅜-in.-thick, sheet of UHMWPE. The sheet is controlled by arms adjustable through the use of large turnbuckles. The attachment and control of this sheet are such that the inextensible cantilevered plate theory is directly applicable. This theory can be used to predict and compare the actual shapes achieved by the moldboard under certain loading conditions. Field studies of snow flow over these various shapes will allow for comparison to analytical models being developed.

### Tripping Edge Design

After reviewing current tripping edge designs, the research team decided to try a novel approach. The basic components of the tripping edge are a sectioned cutting edge, the front edge snow scoop, and a compressed air cylinder in place of a conventional spring.

One advantage of using an air cylinder is that the air pressure may be varied to allow different tripping strengths under different plowing environments. For instance, this feature would allow the snowplow operator to set a lower tripping force in a municipal setting where the vehicle is plowing at a lower speed, then increase the pressure if part of the route was in a highway setting.

Another advantage incorporated in the tripping edge design is nonlinear characteristics caused by the kinematic design that reduce the vertical forces transferred to the vehicle during tripping.

Three separate sections were designed into this tripping edge. Most obstacles encountered during plowing operations will only trip one or two of these sections. In concert, these design features reduce the potential for loss of vehicle control by decreasing the inertial forces and maintaining the friction forces between the front tires and the road, as compared to most other tripping designs.

The location of the pivot for the tripping edge is located in line with the edge itself. This feature allows a blade section to trip without lifting the weight of the plow when the cutting edge is in a vertical orientation. Castor assemblies have been designed that will keep the plow from dropping onto the road surface during a trip that should reduce, or eliminate, gouging of the road surface.

### Front-Edge Snow Scoop

As noted previously, a layback angle of almost 50 degrees with the road surface would be preferable for the high-speed plowing of snow. Also, conversations with plow operators, foremen, state engineers, observations of current plowing practices, and recent investigations into ice fracture mechanics, lead to the conclusion that a layback angle of 90 degrees with the road surface is preferable when cutting ice.

A possible solution to this dichotomy is the concept of the front-edge scoop. This is a flexible deflector attached to the front of a conventional, near-vertical, cutting edge. The deflector has been designed to be rigid enough to allow for the weight of snow carried up its surface, but to be flexible enough to endure the tripping of the cutting edge. The most recent design consists of a 1-in.-thick sheet of UHMWPE mounted on webs at about a 50 degree layback angle. Gaps between the flexible sheet and the cutting edge allow for a path for the material collected by the vertical cutting edge to migrate onto the moldboard, while the bulk of the snow will be carried up the snow deflector face onto the moldboard. Figure 2 shows a detailed view of a single 4-ft section of this snow deflector.

## FUTURE GOALS

The experimental snowplow is approximately 90 percent complete at this time. Figure 3 shows the arrangement of the previously discussed items. This plow will be used during the winter of 1990–1991 to obtain data necessary for the refinement of the design to be used for the two prototypes deliverable on the project. The innovations that have been incorporated in this experimental plow are direct outcomes of research on snow and snowplows, as well as reviews of requirements and current practices. Additional modifications

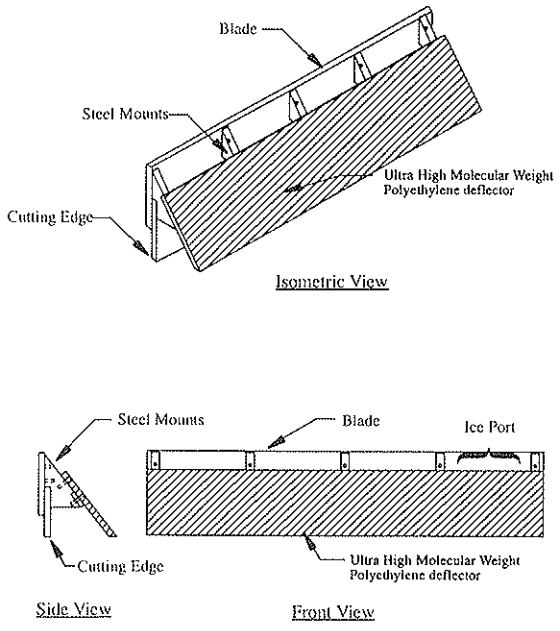



**FIGURE 2 Snow deflector geometry.**

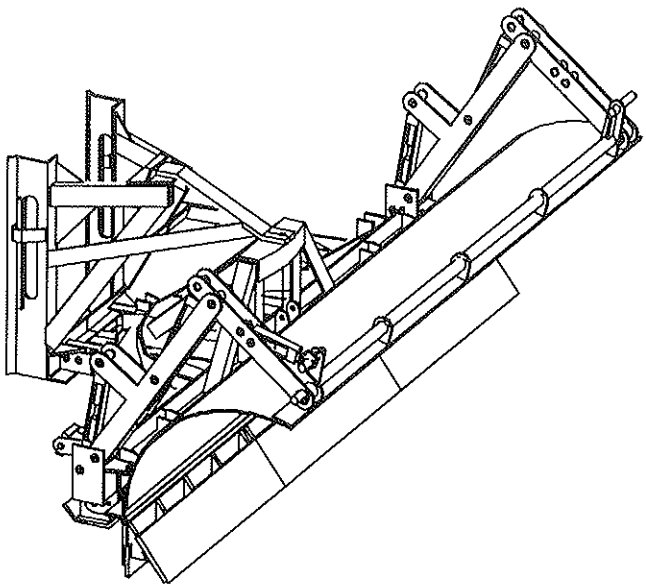


**FIGURE 3 Experimental snowplow.**

and innovations may arise from the field experiments and research that are continuing.

## ACKNOWLEDGMENTS

The research described herein was supported by the Strategic Highway Research Program (SHRP). SHRP is a unit of the National Research Council that was authorized by section 128 of the Surface Transportation and Uniform Relocation Assistance Act of 1987. The authors would like to thank the rest of the research and design team at the University of Wyoming, the manufacturers, state highway departments, and the SHRP staff for their continuing support and critical review of this program.

## REFERENCES

1. B. Kihlgren. *Snowplow Investigations.* (Translated by the U.S. Army Material Command Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire.) Report 38, The Swedish Road Research Institute, Stockholm, Sweden, 1961.
2. K. Croce. Der Heutige Stand der Schneepflugtechnik. (The Present Status of Snowplow Technology.) *Forchungsareiten aus dem Strassenwesen*, Band 31: Winterdienst auf Strassen und Reichsautobahnen, Berlin, 1941, pp. 62–76.
3. D. A. Shalman. *Snowplow Construction, Theory and Design.* (Translated by the Department of Mechanical Engineering at the University of Wyoming, Laramie.) *Mashinostroenie*, Leningrad, Soviet Union, 1973.
4. J. G. Simmonds and A. Libai. Exact Equations for the Large Inextensional Deformation of Cantilevered Plates. *ASME Journal of Applied Mechanics*, Vol. 46, Sept. 1979, pp. 631–636.
5. A. Libai and J. G. Simmonds. Nonlinear Elastic Shell Theory. *Advances in Applied Mechanics*, Vol. 23, Academic Press, New York, pp. 335–337.
6. P. Darmon. Large Inextensional Deformation of Thin Orthotropic Inhomogeneous Cantilevered Plates with Distributed Loads. Ph.D. dissertation, University of Rochester, Rochester, New York, 1985.
7. P. Darmon and R. C. Benson. Large Inextensional Deformation of Orthotropic Cantilevered Plates with Distributed Loads. *ASME Journal of Applied Mechanics*, Vol. 52, June 1985, pp. 385–388.
8. P. Darmon and R. C. Benson. Numerical Solution to an Inextensible Plate Theory with Experimental Results. *Transactions of the ASME*, Vol. 53, Dec. 1986, pp. 886–890.
9. *PHOENICS Instruction Course, Lecture Notes.* CHAM of North America Inc., Huntsville, Ala., 1988.
10. *ADAMS User's Manual, Applications Manual, Tire User's Manual.* Mechanical Dynamics, Inc., Ann Arbor, Mich., 1989.

*This paper represents the views of the authors only, not necessarily those of the National Research Council, SHRP, or SHRP's sponsors. The results reported here are not necessarily in agreement with the results of other SHRP research activities. They are reported to stimulate review and discussion within the research community.*

*Publication of this paper sponsored by the Committee on Winter Maintenance.*

# PASCON: An Expert System for Passive Snow Control on Highways

DARRELL F. KAMINSKI AND SATISH MOHAN

Blowing and drifting snow is a common occurrence on roadways in cold regions that cause reduced visibility and snowdrifts on the roadway, resulting in hazardous road conditions and partial or total road closure. Consequences include longer travel time, greater maintenance and snow control costs, and more vehicle accidents involving property damage, personal injury, and, in extreme cases, loss of life. Passive snow control is the name given to methods offering some control over where wind-driven snow will or will not be deposited. Passive snow control techniques include snow fences, shelterbelts, and design of aerodynamic roadway sections. Currently, no widely accepted algorithmic methods exist for passive snow control on highways. The main objective of the project was to provide a tool for highway design and maintenance personnel to use in evaluating snow problem locations and identifying possible solutions, without requiring an extensive knowledge of passive snow control methods. To this end, an expert system, PAssive Snow CONtroller (PASCON), was developed on an IBM PC microcomputer. PASCON incorporates knowledge from a nationally recognized expert in passive snow control and from the literature. PASCON includes five external programs for design procedures, computations, and graphics. Several consultations with the expert system yielded results that agreed with the domain expert and with solutions worked out manually.

Blowing and drifting snow is a common occurrence on roadways in snow regions, causing reduced visibility and snowdrifts on the roadway, resulting in hazardous road conditions and partial or total road closure. Consequences include longer travel time, increased maintenance and snow control costs, and more motor vehicle accidents involving property damage, personal injury, and, in extreme cases, loss of life. Most state highway departments consider snow control primarily a maintenance responsibility, with little attention given to snow-related problems during the highway design process. Also, there are currently no widely accepted preventive methods of snow control. Passive snow control is the name given to methods offering some control over where wind-driven snow will (or will not) be deposited. This method contrasts with mechanical methods of snow control by plowing and deicing that are in predominant use today.

Passive snow control techniques include snow fences, shelterbelts, and the design of aerodynamic roadway sections. Although this technology to mitigate or even eliminate many of the problems created by blowing snow has been available for many years, it is seldom put into practice. One reason is the fact that expertise in passive snow control is virtually nonexistent in most areas of the world. In addition, snow control measures are often not used because of reluctance resulting from past experience with improper designs, lack of information on proper techniques, inadequate right-of-way, insufficient funds, or absence of a passive snow control policy to address these problems. This paper presents an expert system, PAssive Snow CONtroller (PASCON), that incorporates domain knowledge available in the literature and the experience and knowledge of a leading expert on the subject. It is intended that this system will prove to be an effective mode for transfer of technology from those who possess the knowledge to those who need it.

D. F. Kaminski, New York State Department of Transportation, 125 Main St., Buffalo, N.Y. 14203. Current address: 962 Klein Road, Amherst, N.Y. 14221. S. Mohan, Department of Civil Engineering, State University of New York at Buffalo, Buffalo, N.Y. 14260.

## PASSIVE SNOW CONTROL TECHNIQUES

Although passive snow control techniques have been in use for more than a century, engineered passive snow control is a relatively new technology. Modern techniques have been in existence only since the early 1970s (*1,2*). This new era began with successful installation of engineered snow fences along a 77-mi section of Interstate 80 in southeast Wyoming in 1971. This new highway section was closed 10 times during its first winter in service because of severe problems with blowing and drifting snow. In an effort to alleviate this hazardous and also somewhat embarrassing situation, the Wyoming State Highway Department was willing to install several miles of snow fence designed almost exclusively from untested research studies (*3*). The success of these fences has greatly aided development and acceptance of passive snow control as an attractive and economical alternative. There are three basic categories of passive snow control—drift-free roadway design, snow fences, and shelterbelts.

### Road Design

The idea of preventing snow drifting on roadways by providing an aerodynamic cross section was pioneered by E. A. Finney in the 1930s. One of his most significant findings was the conclusion that the length of a snowdrift was 6.5 times the embankment height (or cut depth) for heights (or depths) of 2 to 10 ft (*4*). This rule of thumb for predicting snowdrift lengths was a useful tool for highway designers desiring to provide a drift-free cross section. Finney's work gained wide acceptance and was not seriously challenged until R. D. Tabler's research in the early 1970s. His studies of snow fences along Wyoming highways led to the observation that the slope of snowdrifts in roadway cut sections did not agree with Finney's research or other derivative literature.

Tabler developed a regression model based solely on topographic data to predict drift formation (*2*). The model was compared to existing drift locations and found to give reliable results. The significant difference between Finney's and Tabler's research was that while Finney found drift length to be directly proportional to embankment height, Tabler's regression model shows that it varies exponentially with height. For example, the length of a drift created by a 4-ft roadway cut will extend about 195 ft beyond the top of the cut—nearly 50 times the depth of cut.

The reason given for the disparity between Tabler's and Finney's findings is that Finney's wind tunnel experiments did not satisfy modeling similitude requirements, with the likely result that embankment heights tested were much higher than intended. This explanation is supported by the fact that the two theories converge for embankments of considerable height. Tabler's work provides a method for dynamic design and analysis of roadway cross-sections with respect to their potential for drifting. If drifting is indicated, the roadway may be redesigned in an iterative fashion until the model indicates that the roadway will remain drift-free.

The Wyoming State Highway Department uses a computer algorithm based on this theory to design drift-free roadways and redesign existing roadways where drifting is a problem (*5*). Redesign options for roadway cut sections include flattening upwind and downwind cut slopes, widening ditches on both sides of the road, and raising the road's profile above the ambient snow cover. Embankments may be made drift-free by providing leeward fill slopes that are equal to or flatter than 4:1 (*6*).

Guiderail often creates drifting onto the roadway at locations that would otherwise be drift-free (*6*). The process results from corrugated-beam guiderail performing as a miniature snow fence, inducing snow deposition downwind. The guiderail also tends to catch snow plowed off the road and prevent it from being thrown farther from the road. This further exacerbates the problem by creating a new snow berm at the guiderail, which may cause blowing snow to cross the road near driver eye level, resulting in reduced visibility. This snow berm may also act as a ramp that can direct a vehicle into the same obstacle from which the guiderail is designed to protect the motorist. The New York State Thruway Authority was found to be negligent in a lawsuit that resulted from an accident caused by this ramping effect (*7*).

For these reasons, at locations where drifting or poor visibility can be attributed to guiderail, the guiderail should be eliminated if possible. If elimination is not possible, then use of cable guiderail is recommended. Use of corrugated-beam guiderail is discouraged. Good road design, although an effective method for preventing drifting onto the roadway, will not obviate the need for other measures if improved visibility is also an objective. Also, because of the significant work involved, road design may not prove to be a cost-effective solution for existing roadways, but road redesign can be evaluated as an alternative to solving drifting problems.

## Snow Fences

The basic function of a snow fence is to produce a reverse airflow area that will cause wind-driven snow to be deposited upwind of the area requiring protection. Although the history of snow fences dates back to their use by railroads in the 1800s, modern engineering criteria for design of snow fence installations have existed for less than 25 years. Tabler was the first to design snow fences for a specific snow storage capacity, on the basis of seasonal snow transport (*8*). He developed a method for estimating snow transport at a given location that depended on seasonal precipitation and unobstructed upwind distance, referred to as the "fetch" (*1*). Another method for estimating snow transport, which Tabler derived from work by Pomeroy, depends on wind speed (*9*).

The underlying premise of the precipitation-based method is that sufficient wind exists to transport all the relocatable snow—i.e., there is "more wind than snow." Conversely, the wind-based method assumes that the amount of snow relocated is limited by the available wind—i.e., "more snow than wind." The equations used to estimate seasonal snow transport are as follows:

### *Precipitation-Dependent Seasonal Snow Transport*

$$Q = 0.5kPT[1 - 0.14^{(F/T)}] \quad (1)$$

where

$Q$ = total snow transport (cubic feet of water per foot of width);
$k$ = transport coefficient, percent of snowfall that is relocatable, expressed as a decimal (0.5 to 0.7);
$P$ = seasonal snowfall (water equivalent, ft);
$T$ = maximum transport distance (usually 10,000 ft); and
$F$ = fetch distance or upwind open distance (ft).

### *Wind-Dependent Seasonal Snow Transport*

$$Q = 0.004895 \sum_i [D_i \sum_j (F_{ij})\ (U_j^{4.04})] \quad (2)$$

where

$Q$ = total snow transport (pounds per foot of width),
$D_i$ = number of now accumulation days in Month i,
$F_{ij}$ = frequency of occurrence for wind speed Group $j$ for Month $i$, and
$U_j$ = composite speed for wind speed Group $j$.

Tabler believes that the wind-based equation is valid for fetch distances of 1,000 ft or more. For this reason, the precipitation-based transport equation should be used exclusively for locations with a fetch distance less than 1,000 ft. In locations where fetch distance is greater than this, total snow transport to be used for design and analysis should be the limiting value from the two equations. The design snow transport is then used to determine size and location of the snow fence required to store this volume of snow.

There are several different types and shapes of snow fences. They have been constructed using materials ranging from steel to paper. The expert system presented here uses four types of snow fences: (a) Wyoming-type wood-slat, (b) synthetic

(plastic), (c) wood-picket, and (d) chain-link. Although chain-link fence is not recommended, it is included to evaluate its placement adjoining the roadway. The other three are the most common types of highway snowfence in use today. Because they have different porosities, they have different storage capacities and different drift profiles (*10*)

### Shelterbelts

Also referred to as "living snow fences," shelterbelts are rows of trees or shrubs planted to provide protection from blowing snow. The known history of shelterbelts in this country dates from the early 1900s when they were used by railroad companies (*11*). Use of living snow fences to protect highways dates from the 1920s. Many states currently have formal living snow fence programs.

Living snow fences have many advantages compared with fabricated ones, including roadside beautification, environmental benefits, little or no maintenance costs after they become established, long service life, and possible lower life-cycle costs. A disadvantage is that they generally require 5 to 10 years before beginning to reach effective heights, although snow fences may be used during the establishment period if immediate protection is desired.

Proper design of shelterbelts depends on many factors, including design snow transport, height of plantings, plant type, number and spacing of rows, and available upwind distance. No quantitative methods now exist for design of living snowfences (*11*). Designs are based on experience and planting schemes that are certain to provide some degree of protection.

An important consideration in shelterbelt design is change in drift pattern with growth and densification of the plantings. A living fence will perform like a porous snow fence during the first few years. As the plantings become more dense with crown closure they will perform more like a solid barrier. Change in drift pattern with growth must be considered during design. Also, an effective seasonal shelterbelt could be achieved in some areas by leaving several rows of cornstalks standing through the winter (*6*).

## PROTOTYPE EXPERT SYSTEM DEVELOPMENT

Nine distinct phases of system development were identified for this project:

1. Identification and acquisition of domain knowledge;
2. Selection of expert system development environment;
3. Development of computer algorithms to compute predicted drift profiles before and after implementation of recommended control measures, and development of other support programs;
4. Acquisition, tabulation, and manipulation of climatological data;
5. Formulation of rules;
6. Development of system reasoning behavior;
7. Organization and formulation of the system;
8. Testing, verification, and fine tuning of the system; and
9. User interface development.

These phases are described in the following sections.

### Identification and Acquisition of Domain Knowledge

#### *Knowledge Sources*

Several data bases were searched for information about passive snow control on highways, including those of the Transportation Research Information Service (TRIS), the U.S. Army Corps of Engineers, and (files of) the New York State Department of Transportation (NYSDOT). The literature survey identified Ronald D. Tabler of Tabler & Associates, Colorado, as a leading expert on passive snow control. He has developed several new methods that have been successfully adopted on Wyoming highways. He agreed to be the domain expert. Several state, county, and local officials within New York State were also identified who are known to have some knowledge of the subject.

The expert system has incorporated knowledge in three basic areas: (a) history of passive snow control methodologies, (b) current practices in western New York, and (c) global knowledge of the domain. Data on evolution of passive snow control techniques was obtained in large part from available literature. Knowledge of current practices in western New York was obtained from several state, county, and local officials who had experience in snow control on roads. They also provided information on why they did not use certain methods, which helped focus the PASCON expert system on addressing problems existing with some of the current techniques.

Much of the domain knowledge for the PASCON expert system was provided by Tabler, who offered general guidance during system development, pointed out several idiosyncrasies, explained unclear principles, advised on proper application of research, and judged the correctness of the expert system's output.

Flow of knowledge transfer is shown in Figure 1.

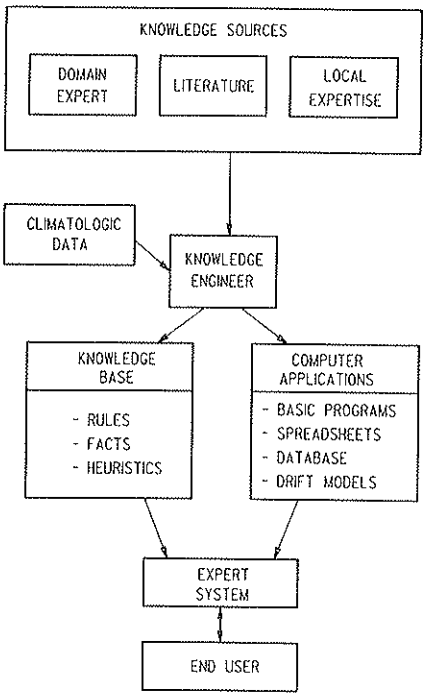



**FIGURE 1 Schematic representation of knowledge transfer.**

*Knowledge Acquisition Methods*

Available literature was examined and relevant domain knowledge was noted. Methods described throughout the literature were evaluated with an eye toward identifying common ideas and basic rules. For example, although there was no consensus regarding proper height, a recurrent characteristic for drift control was that raising the roadway profile a few feet above the surrounding terrain would help alleviate the problem. Experience and knowledge from various highway officials were obtained through informal personal communication. Discussions were conducted in an attempt to collect knowledge on techniques in general use and also to gain insight as to why certain methods are not used.

For example, one common reason for not using 4-ft picket-type snow fences was that they were not deemed cost-effective. This is validated by modern snow control technology, which shows that this type of fence is generally not effective.

Information from Tabler was obtained over several months through spoken and written communication. He was asked questions on applicability of various analytical techniques, research findings, and the current state-of-the-art. Also, he provided completed analyses that were used to validate various system subprograms.

### Selection of Expert System Development Environment

It was decided to develop the expert system using a commercially available microcomputer-based expert system shell. The selected expert system was required to satisfy three major requirements:

1. Ability to execute the necessary external program easily,
2. Graphics capability so that predicted drifting could be displayed or plotted, and
3. Ability to handle advanced mathematical functions necessary to perform required calculations.

GURU® from Micro Data Base Systems, Inc., was selected because it includes spreadsheet, data base, text processing, graphics capabilities, mathematical functions, and simple external interfacing. All of GURU®'s features were used in the PASCON expert system development. The spreadsheet utility was used to store wind speed data and to perform wind-dependent snow transport computations. Precipitation data for 15 gage stations were stored and accessed by means of the data base utility. The text processor was used to provide user instructions and to present results. Graphs of existing and predicted profiles and several help illustrations were prepared through the graphics utility. System architecture is shown in Figure 2.

### Development of Supporting Computer Programs

Computer algorithms were developed using an incremental application of Tabler's regression models to predict snowdrift profiles created by topographic features and by snow fences. Four support programs were written in BASIC language for the various procedures required by the PASCON expert system. A complete listing can be found in Kaminski (*12*).

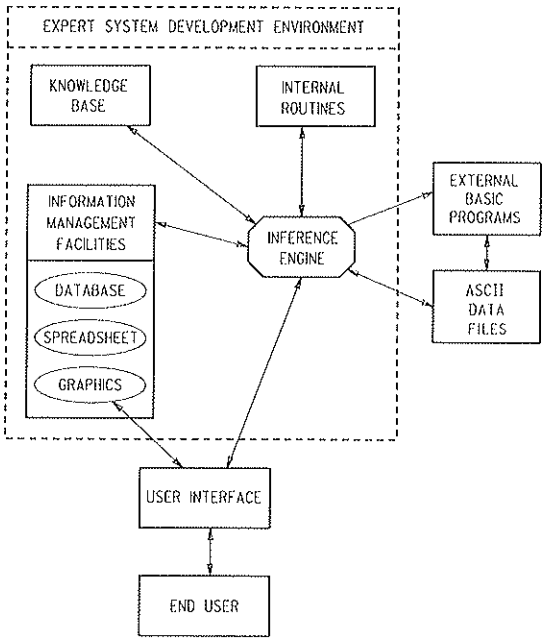



**FIGURE 2 Architecture of expert system for passive snow control.**

1. XSECTION stores cross section points of the road by offset and elevation, for use in plotting and analysis of cross section drift potential.
2. ROUGHXS is used to approximate a cross-section when a detailed survey is not available. It constructs a cross section on the basis of user responses to general questions about the problem site.
3. OGSNOW draws road cross section data from either XSECTION or ROUGHXS and predicts existing snowdrift formation.
4. FNCDRFT computes predicted drift profile as created by a snowfence and determines the fence storage capacity.

### Climatological Data

Surface wind data are used to determine prevailing wind direction with respect to snow transport, and to estimate wind-dependent snow transport. Wind data are available as part of 10-year airport climatological summaries from the National Climatic Data Center in Asheville, North Carolina. The data consist of tables of wind direction versus wind speed (percent frequency of observations). These data are tabulated for 9 wind speed groups and 16 compass points. Because of the tabular format of the wind data, manipulations have been managed by the expert system's spreadsheet mode. A composite wind speed was required for each of the nine speed groups. Also, the wind speed transport equation requires that the wind speed be that at the standard height of 33 ft. Wind speed measurements taken at other heights require adjustment to the standard height. Wind speeds are adjusted by assuming that the velocity provide can be approximated by

$$U = 2.5u_* \ln [(z + h')/h'] \quad (3)$$

where

$U$ = wind speed at height $z$,
$u_*$ = friction velocity, and
$h'$ = aerodynamic roughness height.

(This theory is included in most fluid mechanics textbooks.)

Wind speed at 33 ft was computed by solving for friction velocity at the given wind speed and height. This friction velocity was then used in the equation to determine $U$ at $z$ = 33 ft. An average roughness height of 0.05 in. was assumed because roughness varies depending on vegetation and snow cover. An additional consideration is that the wind-based transport equation is exponential, which means that composite speed is not the median of the wind speed group.

After the conversion of knots to miles per hour and determination of the speed at height of 33 ft, if necessary, the composite speed for a speed group is given by:

$$U = \frac{[u_l^{4.04} + (u_l + 0.5)^{4.04} + \ldots + u_h^{4.04}]^{(1/4.04)}}{[2(u_h - u_l) + 1]} \quad (4)$$

where

$U$ = composite speed for speed group (mph),
$u_l$ = low speed of speed group (mph),
$(u_l + 0.5)$ = intermediate speeds at 0.5-mph increments, and
$u_h$ = high speed of speed group (mph).

Monthly precipitation normals are used to estimate precipitation-dependent snow transport. Precipitation data for cooperative gage stations were obtained from the Northeast Regional Climate Center at Cornell University. The method of polygons was then used to determine domain boundaries for these gage stations. Each town is then assigned to a specific gage station. The system's data base facility managed the precipitation data. This procedure allows the system to select proper precipitation values easily for the problem location. An average snowfall water equivalent of 11 in. was selected as the default value to be used if the town was unknown (or the name misspelled).

Use of wind and precipitation data depends on the duration of the snow accumulation season. Tabler's method for estimating dates of the snow accumulation season on the basis of latitude, longitude, and elevation has been incorporated into this project (*13*).

### Formulation of Rules

The expert system software used in this work permits calling external programs by one-line commands such as RUN "BASICA OGSNOW" or #dskout = "RUNFENCE.ASC." These commands link the support program OGSNOW and the data file RUNFENCE.ASC, respectively, to the expert system. This facility allowed compressing the large knowledge base into 45 rules, because all the design algorithms and computations were executed outside the PASCON expert system. The total rule base is provided by Kaminski (*12*).

A typical rule in GURU©'s syntax includes rule name, comment, priority, IF-THEN clauses, reason, needs, etc., as in the following example:

```
RULE:     RUNFENCE
IF:       TRYFENCE & REQHT < 10
THEN:     e.odsk = true
          #dskout = "RUNFENCE.ASC"
          output designq
          output foffset
          output reght
          e.odsk = false
          e.wfu = false
          run "basica fncdrft"
          e.wfu = true
          HANDLE1 = FOPEN("FDIMENS.ASC," "R")
          FTYPE = FGETL(HANDLE1)
          FH = FGELT(HANDLE1)
          FCLOSE(HANDLE1)
          FHEIGHT = TONUM(FH)
          RUNFENCE = TRUE
NEEDS:    SECTYPE
          ROWWIDTH
          DESIGNQ
          FOFFSET
          TRYFENCE
CHANGES:  FTYPE
          FHEIGHT
```

A simple example of how rules are used to represent knowledge can be illustrated by the following rules to determine the design snow transport:

```
Rule X: IF FETCH <= 1000
        THEN Q = PBASEDQ
Rule Y: IF FETCH > 1000
        THEN Q = MINIMUM(PBASEDQ,WBASEDQ)
```

where

$Q$ = design snow transport,
FETCH = Fetch distance,
WBASEDQ = precipitation-dependent transport, and
WBASEDQ = wind-dependent transport.

These rules represent knowledge that design transport for fetch distances of 1,000 ft or less should be based on precipitation, whereas design transport for fetch distances greater than 1,000 ft should be the limiting value found by the precipitation-based method or the windspeed-based method.

Use of rules to control system logic can be most easily explained by use of an example:

```
Rule a: IF: SURVEY
        THEN: RUN "BASICA XSECTION"
              AND XSECTION = TRUE.
```

The English translation of this rule is "if you have a survey of the site, then input the data points using the BASIC program 'XSECTION'." The "XSECTION = TRUE" clause represents knowledge that the cross section has been stored.

### System Reasoning Behavior

Development of the system logic was completed in stages by directing it to seek specified subtotals, for example, requesting

the system to seek the beginning and end of the snow accumulation season. As individual segments were verified, they were added together to form larger segments. After the system was found to work properly from a mechanical standpoint, the next step was to program it to verify rules and take actions in the desired manner. This procedure was accomplished by assigning a priority order to competing rules that seek values for the same variable. The premise of a rule with higher priority ratings (0 to 100) would be tested before rules with lower priority ratings.

Inferencing was controlled by GURU®'s capability to define the rigor with which a goal or subgoal is sought. This process allows the system either to stop seeking a value for a variable after one is found, or to seek all possible values until all pertinent rules are tested. This process is used, for example, to allow the system to seek multiple solutions to the problem. It can be controlled dynamically within the system by setting the inferencing rigor on the basis of the initial value. Thus, if the system quickly determines that no solutions are feasible, it will not seek additional values for recommended solutions. However, if it determines that road redesign is a possible solution, it will continue to seek other permissible solutions.

The system uses a goal-driven approach, with the goal of providing a recommended solution to the identified snow problem. The system performs three basic tasks: (a) problem identification, (b) problem evaluation, and (c) problem solution, as shown in Figure 3.

Problem identification involves entering cross-section data and providing other site-specific information. The system then uses this information to determine the type of problem that exists. Once identified, the system will then evaluate it to determine its probable causes and estimate its severity. This result is accomplished by a combination of evaluating cross-section elements for their potential to cause problems and estimating drift profiles if needed.

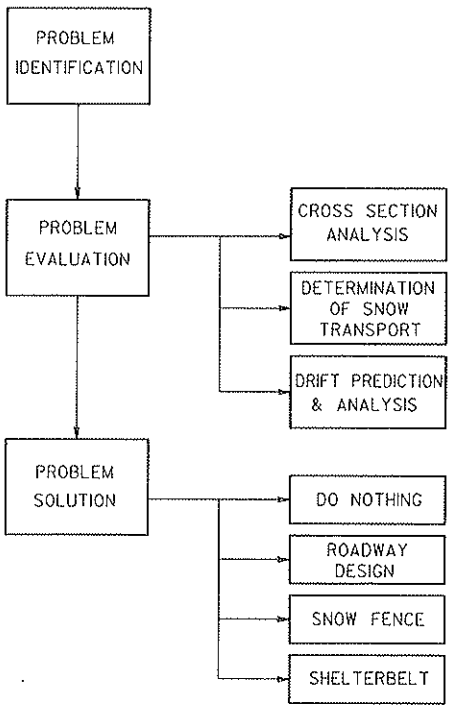



**FIGURE 3 Basic tasks of expert system.**

When evaluation of the problem is completed, the system then tries to find solutions that will mitigate or eliminate it. Possible solutions are (a) do nothing, (b) redesign the roadway, (c) install snow fences, and (d) plant shelterbelts.

Snow fences will be recommended only if they can store at least 50 percent of the total design transport. (This recommendation may be overridden by the user if desired.) This decision is based on the opinion that a fence should store a minimum of half the estimated seasonal snow transport.

## System Organization and Formulation

The system can be divided into three major phases:

1. Initialization,
2. Evaluation, and
3. Completion.

Initialization defines the system goal, introduces the user to the system, initializes all variables, and ensures that the user is prepared to enter the consultation phase.

Evaluation or consultation is the main part of the system. All reasoning and evaluation are performed during this phase. The user is asked for information needed to evaluate the problem, which is then used to infer new information by validating rules and using the information to execute external programs.

As described earlier, historical wind data are stored on a spreadsheet. They can then be accessed and manipulated to determine the wind-dependent snow transport for the determined snow accumulation season. Precipitation data from November through March from 15 precipitation gage stations across western New York are stored in the data base facility. Existing roadway cross section is input using a survey cross section or the user's knowledge of the site to rough in the cross section.

The roadway is then evaluated for its susceptibility to drifting by executing the external roadway drift prediction program. If a significant drifting or whiteout problem is indicated, control measures are investigated. Snow fences are then evaluated by estimating fence height necessary to store the anticipated snow transport and executing the drift prediction model for the external fence. Road redesign options are evaluated within the system and then checked by executing the roadway drift prediction program OGSNOW using the redesigned roadway cross section.

Completion is the end of the consultation in which the recommended snow control measures are delivered to the user. Screen plots of the roadway before and after implementation of the recommendations are displayed to the user. These plots may also be routed to a line printer if desired.

## Testing and Verification

The four external programs were tested for accuracy by comparing results with computations by hand. After debugging and final organization, all the external programs performed

accurately. Spreadsheet and data base manipulations were also compared to hand calculations and were found to give accurate results.

Flow of the PASCON expert system was verified by using the tracing facility of GURU©. This procedure allowed dynamic analysis of the order in which rules were selected for testing. Also, different methods could be quickly and easily checked for their effects on reasoning behavior.

### User Interface

This expert system is being considered by NYSDOT for statewide use. A user interface friendly to NYSDOT engineers is being written. It is planned to include several graphical help screens, as shown in Figure 4, to answer questions that may arise during consultation sessions.

## EVALUATION

### Comparison with Human Expert

An example consultation was performed for a snow problem location previously analyzed using traditional methods. This location is along Route 219 in the town of Boston, New York, where a severe whiteout problem exists. Because this road section is on an embankment and no guiderail is present, drifting is not a problem. It was selected for installation of a demonstration snow fence as part of the Strategic Highway Research Program (SHRP) project to demonstrate the effectiveness of the snow fence on a full-size scale. It was chosen primarily because of the wide right-of-way which allowed the fence to be installed within the state right-of-way. Design of this installation was assisted by Tabler under auspices of SHRP, and was completed in 1989.

The final design was an 8-ft synthetic fence placed near the right-of-way line 160 ft west of the southbound lanes. The total design snow transport would normally require a 10-ft fence for a site on level terrain. This specific site is aided by a small ravine just upwind of the road, which provides additional storage capacity. Also, a fence taller than 8 ft would have to be placed farther from the road to prevent it from potentially casting a drift onto the pavement. The drift predicted by the fence's traditional design indicated that drifting cause by an 8-ft fence would not encroach onto the roadway. The calculations also indicated that it would not be adequate to store the total design snow transport. However, the expected storage indicated that the installation would provide protection for most of the winter months.

The PASCON expert system was used to analyze this location and results were then compared with the original completed design. The system properly determined that road redesign was not a viable solution because the problem was poor visibility. The system found that drifting was not a problem by directly asking the user. This fact was also verified by the system after it determined that the road was on an embankment and that the fill slopes were aerodynamic with respect to snow drifting. The system then analyzed the location for suitability for installation of snow fence. On level terrain, the right-of-way would only be adequate for a 5-ft fence, but that would not provide the declared minimum storage capacity of 50 percent of the total design snow transport. Because this site was on an embankment and not on level terrain, the system did not rule out the placement of fence. The PASCON expert system selected an initial fence height of 8 ft as the minimum allowable that would provide the required minimum 50 percent storage. This fence was then analyzed to determine if the predicted drift would encroach the roadway. It was found that a 9-ft fence would not create drifting onto the road even as the fence approached capacity. The system did not evaluate a taller fence because the predicted drift elevation near the road was found to be within 0.1 ft of the edge of shoulder elevation.

In summary, the system recommendations were very close to the previously completed design. The system recommended that a 9-ft fence offset 160 ft from the road would be a viable long-term solution. This agreed with advice from domain ex-

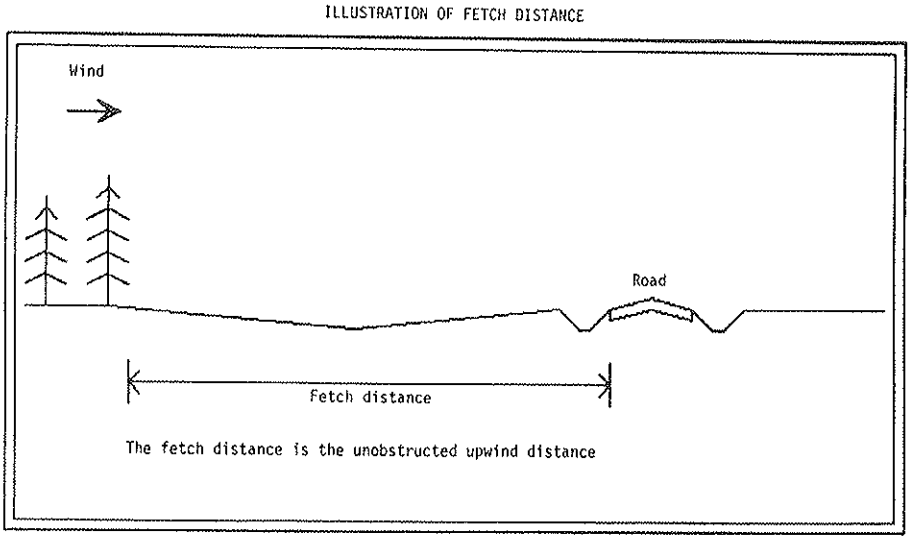



**FIGURE 4 Sample help screen.**

pert Tabler. The consultation for this problem location is shown in Figure 5. The time required to evaluate this problem site with the expert system was considerably less than that spent for the original manual design. The major time savings were in analysis of potential drifting and in fence drift prediction.

Several typical road sections and problems were analyzed to determine accuracy of the system and evaluate its reasoning behavior. The sections chosen were relatively simple problems whose solutions were easily determined beforehand. For example, the solution for a whiteout problem at an at-grade highway location with no right-of-way restrictions would be to install a fence of proper height at the proper distance upwind of the road. Use of several typical examples provided opportunities to observe system logic and determine if the rules were evaluated as intended.

Also, the system followed intended reasoning paths. This behavior was indicated by the observation that the expert system did not ask for information pertaining to impractical solutions, and further, all expected options were evaluated as desired.

GURU® also allows a rule set to be consulted to determine a specific subgoal or variable. This feature was extremely useful because it allowed reasoning behavior and system execution to be checked without going through an entire consultation. An example of this technique was to have the system seek the value for the precipitation-dependent design snow transport whose variable name is PBASEDQ. This operation was easily accomplished by invoking the command "consult direct to seek PBASEDQ." Running this command verified whether proper precipitation values were used and that necessary rules were evaluated. Although the trial problems were designed to be simple to accommodate system evaluation, the system performed as intended and all recommendations were valid.

```
Before we begin it is important that you are prepared to answer
general questions relating to cross-section and general topography
of the problem location.  This information can be from a survey,
record plans, or your personal knowledge of the site.  You should
also
have a USGS map available.
Keep in mind that the accuracy of any recommendations will be a
reflection of the accuracy of the information that you provide.
ARE YOU PREPARED TO BEGIN?
Enter Y or N  ===>  Y

Is drifting snow a problem at this location?
Enter Y or N  ===>  N

Are 'whiteouts' a problem at this location?
Enter Y or N  ===>  Y

The following figure illustrates the various types of cross
sections.  After viewing the figure you will be asked to identify
the type of cross section that approximates your problem site.
     [WIND IS FROM LEFT TO RIGHT]
(Press any key to see figure)

1.  EMBANKMENT SECTION
2.  CUT SECTION
3.  AT-GRADE SECTION
4.  SIDEHILL CUT WITH WINDWARD SIDE ON FILL
5.  SIDEHILL CUT WITH WINDWARD SIDE IN CUT
Enter the NUMBER of the section type that corresponds
to your problem site :  ===>  1

Is there guide rail along the problem section?
Enter Y or N  ===>  N

Have you already entered the cross-section of this site ?
Enter Y or N ===>  Y

The snow accumulation season is partly a function of ELEVATION
(above sea level), LATITUDE, and LONGITUDE.  This information is
easily obtained from USGS maps and some commercial maps.

Enter the ELEVATION of this site (feet)   1560
Enter the LATITUDE of the site, to the nearest .25 degree.   42.5
Enter the LONGITUDE of the site, to the nearest .25 degree.  78.75

LOADING WIND DATA !

Wait ....

DETERMINING PREVAILING WIND DIRECTION
                &
   ESTIMATING TOTAL SNOW TRANSPORT

Please be patient ....
```

**FIGURE 5 Consultation listing.**

```
The upwind ground terrain has a significant effect
on the amount of snow transported.
1.  PAVED SURFACE
2.  FROZEN LAKE
3.  LOW VEGETATION ( 1'' -  6'' )
4.  CULTIVATED FIELD
5.  TALL VEGETATION ( ) 6'' )
6.  NATURAL FIELD OR MEADOW
Enter the NUMBER of the terrain type that most closely corresponds
to your problem site :  ===>  5

The unobstructed distance upwind of a site is referred to as the
'fetch' distance.  The beginning of the fetch is essentially any
barrier across which there will be no blowing snow.  This includes
woods, unfrozen bodies of water, deep ravines, etc.
The following figure illustrates the fetch distance.
(Press any key to see figure)

Enter the fetch distance (feet):  3000

What TOWN is the site located in?  BOSTON

The ROW may not be adequate for passive snow control measures.  Is
an easement at this location possible (Y/N)?   N

Please enter the right-of-way width from the centerline
of the roadway to the windward boundary.  --->  170

Is road redesign an alternative?
Enter Y or N  ===>  N

THE FOLLOWING FIGURE SHOWS THE DRIFT CREATED
BY A  9.0' TALL FENCE

PLOTTING FENCE DRIFT

Please wait ....

PREDICTED SNOW BARRIER DRIFT PROFILE

SUMMARY OF RESULTS:
The snow accumulation season begins on NOVEMBER  30
and ends on MARCH 18
The prevailing wind direction is  WSW
The design snow transport =  939 FT^3/FT

BASED ON THE INFORMATION PROVIDED, I THINK THAT
YOU SHOULD INSTALL A SNOW FENCE
FENCE HEIGHT =            9.0
OFFSET FROM CL =            170
press any key to continue
```

**FIGURE 5 *(continued)***

### Problems Encountered in System Development

Some difficulties surfaced during development of the PASCON expert system, but most obstacles were eventually overcome. Controlling reasoning behavior posed a particular challenge. As with traditional methods, the extent and type of analysis changes as more information is found about the problem. The GURU® expert system environment offers many options to control system reasoning behavior, including forward, backward, or mixed chaining. Also, the order in which rules are selected for evaluation can be controlled by assigning a priority order to different rules that can determine values for the same variable. In addition, many system environment controls may be used to control whether a goal or other variable should continue to be sought. The difficulty was in determining how to use these options properly in a manner that would enable the system to closely emulate a human expert.

## CONCLUSIONS AND RECOMMENDATIONS

Testing the prototype expert system has indicated that it can provide good results, within constraints of the rule set, and

that expert systems for technical applications can and do work. The system is limited only by accuracy of the information used to develop it. Limitations caused by uncertain information will exist whether the problem is analyzed with traditional methods or with an expert system.

The computer algorithms developed for drift prediction are a powerful tool for analysis of current drift potential and for predicting effects of road redesign or snow fence installation. A proposed road design or fence installation can be evaluated in a short time. Analyses of drifting is nearly impossible without a computer program. These programs can also be used as stand-alone options or in an iterative analysis for design of a snow-free roadway, similar to those of the Wyoming State Highway Department. Snowdrifting problems should be addressed at the design stage, where opportunity for alternative designs is at a maximum. This expert system provides a tool for analysis and solution of these problems. It will generally not recommend infeasible solutions. Recommendations it offers will be based largely on information supplied by the user.

This system should not be expected to replace the expert. It offers expert advice to the designer, who must use experience and judgment in accepting and applying each consultation to the problem at hand. It is designed to assist users in solving typical everyday problems, thus freeing the expert to spend time on advancing the technology and solving difficult problems. For this reason, it was not developed to solve all conceivable problems, but is capable of addressing most typical problems.

This project has fulfilled its objective of providing a tool to assist design and maintenance engineers in finding solutions to blowing and drifting snow problems. The system illustrates the effectiveness of expert system technology as a medium for transfer of knowledge to those who can use it. It allows users who do not have knowledge of passive snow control analysis and design to find real solutions to real problems.

## ACKNOWLEDGMENT

The authors wish to express their deep appreciation to Ronald Tabler of Tabler & Associates in Niwot, Colorado, for providing the domain expertise through several long interviews and written communications.

## REFERENCES

1. R. D. Tabler. New Engineering Criteria For Snow Fence Systems. In *Transportation Research Record 506*, TRB, National Research Council, Washington, D.C., 1974, pp. 65–78.
2. R. D. Tabler. Predicting Profiles of Snowdrifts in Topographic Catchments. *Proce., Western Snow Conference*, Coronado, Calif., Vol. 43, April 175, pp. 87–97.
3. R. D. Tabler. *Snow Fence Handbook (Release 1.1)*. Tabler & Associates, Niwot, Colo., 1988.
4. M. Mellor. *Blowing Snow*. CRREL Monograph, Part III, Section A3c. U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, N.H., 1965.
5. S. L. Ring, J. D. Iversen, J. B. Sinatra, and J. D. Benson. *Wind Tunnel Analysis of the Effects of Planting at Highway Grade Separation Structures: Final Report*. HR-202, Iowa Highway Research Board, Ames, 1979.
6. F. W. Cron. Snowdrift Control Through Highway Design. *Public Roads*, Vol. 34, No. 11, 1967, pp. 227–234.
7. J. C. Vance. Supplement to Liability of State and Local Governments for Snow and Ice Control. *NCHRP Report 14: Legal Research Digest*. TRB, National Research Council, Washington, D.C., June 1990, p. 7.
8. Problem of Blowing Snow in Wyoming Under Attack by Highway Department. *Highway Research News*, No. 47, HRB National Research Council, Washington, D.C., 1972, pp. 52–56.
9. R. D. Tabler, C. S. Benson, B. W. Santana, and P. Ganguly. Estimating Snow Transport From Wind Speed Records: Estimates Versus Measurements at Prudhoe Bay, Alaska. Presented at Western Snow Conference, 1990.
10. R. D. Tabler. Geometry and Density of Drifts Formed by Snow Fences. *Journal of Glaciology*, Vol. 26, No. 94, 1980, pp. 405–419.
11. D. L. Shaw. *Living Snow Fences: Protection That Just Keeps Growing*. Colorado Interagency Living Snow Fence Program, Colorado State University, Fort Collins, 1989.
12. D. F. Kaminski. *An Expert System for Passive Snow Control in Western New York* Master of Engineering Project, State University of New York, Buffalo, May 1990, 56 pp.
13. R. D. Tabler. Estimating Dates of the Snow Accumulation Season. *Proce., Western Snow Conference*, Vol. 56, 1988, pp. 35–42.



*Publication of this paper sponsored by Committee in Winter Maintenance.*

# Application of Routing Technologies to Rural Snow and Ice Control

EDWARD HASLAM AND JEFF R. WRIGHT

The design of routes for intrastate highway snow and ice control is perhaps the most difficult and complex of all public-sector routing problems. In addition to the random and usually unevenly distributed effects of a snow event, service must be provided rapidly, equitably, and simultaneously across the network. The task is made more difficult by the presence of multiple and conflicting objectives on the part of maintenance engineers responsible for this service. The design of snow removal routes is addressed from the perspective of multiple objective optimization. The strengths and weaknesses of several mathematical programming approaches are discussed and an efficient heuristic routing methodology is proposed. Experience with the analysis of a portion of the Indiana highway network is described.

Many public services are provided along predetermined routes with the overall effectiveness of those activities determined, at least in large part, by the efficiency of those routes. The design of service routes is characterized by multiple and conflicting objectives, many of which are difficult to quantify. Perhaps the most complex service activity for which careful route planning is important in wintertime snow and ice control on our nation's intrastate highway system.

Snow and ice control during the winter months in northern states is a major operation. The Indiana Department of Highways (INDOT), for example, must routinely maintain some 11,414 mi of roadway throughout the state. Because each traffic lane must receive service, this requirement translates into more than 29,000 lane-mi overall. The resources needed for this operation include nearly 1,500 trained personnel and some 1,088 maintenance vehicles. The cost is enormous. Over $15 million was budgeted by Indiana to support the operation during the 1987–1988 winter season and these costs are expected to increase by 20 percent for the 1988–1989 season.

The management of the operation is complex because of several factors. Unlike other route-oriented public-sector activities, snow removal must be initiated simultaneously throughout the entire event region, which, for Indiana, may mean as many as 1,200 vehicles commencing service at the same time. In addition to the snow removal fleet, a team of radio operators and mobile supervisory units is also required. The operation often extends through more than one worker shift, creating additional personnel management problems as well (*1*).

In addition to the magnitude of the overall operation, uncertainties as to the duration and severity of the snow emergency pose special problems. State roads are categorized into three classifications on the basis of historical average daily traffic (ADT). Class I roads (major traffic arteries including Interstates and their associated ramps and roads with ADT greater than 5,000) receive continuous service including plowing and the application of chemicals and abrasives as needed to keep the road surface wet and bare. Class II roads (routes having ADT between 1,000 and 5,000) receive continuous plowing and sufficient chemicals and abrasives to maintain a bare wet pavement in the center portion of the roadway. Class III roads (ADT less than 1,000) receive enough service to keep the routes passable, with chemical treatment only for hills, curves, and intersections. Any and all routes created should maintain as much route class continuity as possible; that is, each route should be created to service only one class of road. However, because of uncertainties associated with snow episodes, these criteria serve only as general guidelines; actual service levels are often determined while the operation is in progress.

Following a snow episode, extensive clean-up activities must be instituted, consisting of additional plowing and spot application of chemicals to remove all remaining snow from driving surfaces. This operation also includes clearing shoulder areas and special servicing of drains, overpasses, and bridges; drifts; and the maintenance equipment itself. Timing is extremely important in all phases of the operation.

In Indiana, as in most states, snow and ice control is administered at the subdistrict and site level consistent with preestablished snow routes of which there are some 986 in Indiana at present. Though all vehicles begin operation from a specific site location, routes proper do not necessarily begin at a site location. In cases where a truck must travel to the starting point of a route, no service will be conducted during this time. Such travel is referred to as "deadhead" travel, with deadhead miles accounting for nearly 9,000 vehicle-mi per season. Though travel times will fluctuate with overall conditions, design operation calls for plowing intensities dependent on roadway classification. In general, a route must be completely serviced in 2 to 2.5 hr.

Finally, the operation is complicated by the presence of public traffic. Plowing and spreading is most effective within a range of speed for the maintenance vehicle. When vehicle movement is restricted, performance decreases and road conditions will further deteriorate, causing additional problems and potentially dangerous conditions.

In designing an overall management strategy for snow and ice control, a number of difficult questions may be posed: What is the best set of routes for maintenance vehicles so as to minimize overall deadhead miles (minimize excess cost)? What characteristics of individual routes are most important in terms of overall safety of operation? How best should ve-

E. Haslam, Corporate Research and Development, United Airlines, Chicago, Ill. 60666. J. R. Wright, Department of Civil Engineering, Purdue University, West Lafayette, Ind. 47907.

hicles be restocked with sand and chemicals during the operation? What contingencies should be provided to compensate for the uncertainties of storm intensity? What are the best allocations of vehicles to sites and routes to vehicles? Are the current administrative boundaries optimal or should the configuration of subdistricts be modified? Where should new facilities be constructed to best enhance overall system performance? These questions and more should be considered in the design of an effective strategy for conducting winter road maintenance.

The focus of this research is the design of an efficient algorithm for the specification of snow and ice control routes, given complete and precise information about the underlying target network. All road segment lengths and adjacencies are known, and the locations of vehicle and abrasive storage facilities are fixed. Furthermore, a partition of the network for service from those sites is assumed. The problem is one of assigning road segments within the partition to garage sites (vehicles) consistent with a service policy that sets a specific performance level on the basis of time of service to a particular road type. Following a review of the relevant vehicle routing literature, the details of a multiobjective heuristic procedure for solving this problem are described. An actual route design exercise is presented and computational experience is discussed.

## ROUTE DESIGN TECHNOLOGY

There is an extensive body of literature addressing the topic of vehicle routing and scheduling. In this section, a review is offered of that portion of this literature relevant to the winter service vehicle routing problem. Bodin et al. (*2*) provide a textbook synopsis and more complete literature review of the entire field of routing and scheduling of vehicles and crews.

### General Methodologies

The topic of vehicle routing and scheduling covers such activities as retail distribution, school bus routing, mail delivery, street sweeping and snow removal, waste collection, and communications system management. Yet given this diverse spectrum of applications, the entire field of study can be partitioned into two major categories: (a) routing vehicles when the only concern is developing a set of routes to satisfy demand at several points or streets, and (b) scheduling vehicles when the concern is satisfying demand given time windows or precedence relations with regard to the demand points. Some authors (*3*) choose to make the major partition of the topic at demand type. That is, does the demand for service originate from points, making it a node-covering problem, or does it exist along the length of the street, making it an arc-covering problem?

The class of solution methodologies most relevant for this research fall under the classification of the Chinese postman problem (CPP). This classical operations research problem is one of finding the minimum cost cycle that visits every arc in a given network at least once. The basic CPP is well solved (*4*) but is computationally intractable on mixed networks (*5*). When a simple constraint such as a vehicle capacity restriction is imposed on the basic CPP on a directed or undirected network, the CPP becomes known as the "capacitated" CPP and is NP-hard (*6*). A polynomial time algorithm has been developed (*7*) for the CPP on a network with precedence relations on the arcs, which is of particular interest with regard to roadway clearing priorities for the snow removal problem.

### *The Chinese Postman Problem*

In the world of network algorithms, the problem of snow removal is an arc-covering application in that the concern is covering every, or a subset of every, link or road in the network. The CPP is the basic mathematical programming formulation for this application; it is a classic problem statement within the operations research literature (*8*). As mentioned, the objective of the CPP is to find the least-cost path through the network that covers every arc at least once and starts and ends at the same node. For the snow removal problem, least cost is defined as minimum distance and arcs that are traversed more than once are tallied up as deadhead miles. Deadheads miles or arcs are segments traversed by a truck that is not servicing those segments either because they are serviced by another route or are used exclusively as access to a serviceable segment. An initial problem formulation was solved to find a CPP solution using one truck with infinite capacity starting from and returning to a common depot with the objective of covering every road in the network while minimizing total deadhead miles. Although this formulation is not a realistic representation of the actual multiple truck and depot problem, it does provide a lower bound on the solution and a good starting point for more complex models. The CPP mathematical program is a minimum-cost, flow-based formulation (*9*), as follows:

$$\text{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} \tag{1}$$

Subject to

$$\sum_{k=1}^{n} x_{ki} - \sum_{k=1}^{n} x_{ik} = 0 \qquad \text{for all } i, \text{ and} \tag{2}$$

$$x_{ij} \geq 1 \text{ and integer} \qquad \text{for all } (i,j). \tag{3}$$

In this formulation, the decision variables $x_{ij}$ is the number of times the arc from Node $i$ to Node $j$ is covered in the optimal solution. The objective (Equation 1) seeks to minimize the total distance traveled, which is the summation of how many times an arc is covered, $x_{ij}$, times the length of the arc, $d_{ij}$. The first constraint set (Equation 2) enforces flow continuity for all $n$ nodes in the networks by ensuring that every time a truck enters a node it must also leave that node. The final constraint set (Equation 3) states that every arc must be covered at least once and the answers must be integer; an arc being covered a fractional number of times has no meaning. It is a well known characteristic of this formulation that when solved as a continuous linear program using the simplex algorithm, an integer solution will result without the need for use of branching or cutting-plane methods.

*The Rural Chinese Postman Problem*

The first step toward model realism is to realize that INDOT is only responsible for a subset of all the roads in the state. That is, INDOT only has to service state roads, highways, and Interstates; however, its trucks may need to traverse a county road to provide service to a state road. Therefore, city streets and county roads should be included in the data, but with a flag designating them as merely nonrequired arcs that may be used to provide more efficient service or reduce deadheading miles. With the data so marked, a variant of the CPP known as the "rural postman problem (R–CPP) can be used to solve the problem. The formulation is identical to the CPP except for the addition of the following constraint:

$$x_{ij} \geq 0 \text{ and integer} \qquad \text{for all } (i,j) \in (A - R) \tag{4}$$

The modified formulation uses the notion of the set $R$, denoting the set of all arcs that INDOT is required to service, and the set $A$, which is the set of all of the arcs in the network. The additional constraint set (Equation 4) thus allows that all arcs that are not required to be serviced by INDOT do not have to be covered, but may be traversed. This discrete model will also terminate integer when solved as a linear program.

*The Multiple Truck Rural Chinese Postman Problem*

The previous model, R–CPP, used the idealized one truck of infinite capacity. Indiana's current snow control operations use multiple trucks departing and returning to a common depot of which there are several grouped into subdistricts that in turn are part of larger districts. If the resolution is restricted to one depot in any given subdistrict, the R–CPP may be extended to accommodate multiple vehicles and result in a formulation called the "M-rural postman problem" (MR–CPP). This mathematical program is a bit more complex, as follows:

$$\text{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m} d_{ij}\, x_{ij}^{m} \tag{5}$$

Subject to

$$\sum_{k=1}^{n} x_{ki}^{m} - \sum_{k=1}^{n} x_{ik}^{m} = 0 \qquad \text{for all } (i,m) \tag{6}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}\, x_{ij}^{m} \leq \text{MAXMILES} \qquad \text{for all } m \tag{7}$$

$$\sum_{k=1}^{m} x_{ij}^{m} \geq 1 \text{ and integer} \qquad \text{for all } (i,j) \in R \tag{8}$$

$$\sum_{k=1}^{m} x_{ij}^{m} \geq 0 \text{ and integer} \qquad \text{for all } (i,j) \in (A - R) \tag{9}$$

$$x_{\text{depot}}^{m} = 1 \qquad \text{for all } m \tag{10}$$

The MR–CPP has a modified decision variable $x_{ij}^{m}$, that identifies which trucks will cover each arc $(i,j)$, so that there are now $m*a$ integer decision variables in the formulation, where $a$ = number of arcs. The objective function (Equation 5) is the same: minimize total distance traveled by all $m$ trucks. The continuity constraint (Equation 6) is the same flow continuity set, but now they are summed for every truck at each node. Vehicle capacity constraints (Equation 7) enforces vehicle capacity by setting the constant MAXMILES equal to the maximum number of miles each truck may travel. The next two constraint sets (Equation 8) and (Equation 9) are the same as before, but they must also be summed over all the vehicles. The final set of constraints (Equation 10) forces every vehicle to emanate from a designated depot.

*Subtours*

Solutions for the R–CPP and the MR–CPP may contain subtours. Subtour generation is the major stumbling block in the formulation of any type of exact routing algorithm. There are several different ways to address the problem of subtours, but for the current purpose, an exact solution for the R–CPP or MR–CPP, it is sufficient to discuss only one (*10*). The subtours could be eliminated if there existed constraints in the R–CPP or MR–CPP formulation that would not allow them to be formed in the first place. An example of such a set of subtour-breaking constraints is as follows (*6*):

$$\sum_{i \in R} \sum_{j \in R} x_{ij} - n^2\, y_{1r} \leq |R| - 1$$

$$\text{for every nonempty subset } R \text{ of } \{2, 3, \ldots, n\} \tag{11}$$

$$\sum_{i \in R} \sum_{j \in R} x_{ij} + y_{2r} \geq 1 \tag{12}$$

$$y_{1r} + y_{2r} \leq 1 \tag{13}$$

$$y_{1r}, y_{2r} \in \{0,1\} \qquad r = 1, \ldots, (2^{n-1} - 1) \tag{14}$$

The problem with these constraints is that there are an exponential number of possible subtours that can be formed on the basis of the number of nodes; therefore there is an exponential number of possible constraints that could be added to the model to stop their formation. Two problems exist with these additional constraints: (a) for problems of any realistic size, say >100 nodes, the number of subtour breaking constraints grows to an outlandish size; and (b) these constraints force the model off of integer linear programming solutions. With the continuous linear program no longer guaranteed to terminate with an integer solution, the computational burden of iterating to an integer optimal solution can become prohibitive (*11*).

*Conclusions*

With the inclusion of subtour elimination constraints and the fact that there are $m*n$ integer variables in this formulation, the MR–CPP becomes even more hopeless as far as an exact solution by mathematical programming is concerned. But this formulation does provide an idea of the complexity and size of exact routing models as more real-world or problem con-

strains are imposed on the problem. Furthermore, an essential dimension of the problem is missing from the MR–CPP, the fact that there are multiple depots in each subdistrict. At a minimum, any design or analysis of routes and their configurations should be done at a subdistrict resolution. A multidepot MR–CPP formulation would require still more constraints and variables, and given the combinatorial explosion at the R–CPP and MR–CPP level, further listing and explanation are not necessary, given the purpose of this research.

What has been useful in creating and manipulating these exact formulations is the insights into the problem complexity they have unveiled. So far, no efficient method has been discovered to handle the different road classifications or priorities, given that each route should have class continuity. One way to handle this would be to partition the network into three subnetworks, one for each classification, and run a routing algorithm on each one. This procedure assumes that route class continuity is an absolutely binding constraint. (It is not, because the current set of INDOT routes violates route class continuity.) Therefore, this topic is better classified as a minor objective that could possibly be handled with a penalty function. This consideration leads to the point of identifying the true objective of this problem.

The public sector problem of snow and ice control is in reality a multiobjective problem. While minimizing cost, total distance, and number of trucks, INDOT also wishes to maximize the level of service to the public. This could be handled with some multiobjective optimization paradigm such as iterative trials, holding level of service constant and minimizing cost or maximizing efficiency (*12*).

A final point on these exact procedures is that although they are computationally impossible, they do provide the basic groundwork for formal heuristic methods such as linear programming column generation schemes, Lagrangian relaxation procedures, and others (*10,13*). These heuristic methods are appealing because they are based on the exact models and are therefore easy to justify and build the routes up from a trivial initial feasible solution to near optimal in a logical form.

### Arc-Covering Applications and the Snow Removal Problem

The work by Marks and Stricker (*14*) addressed the problem of urban snow removal with the CPP as their base model. In their analysis of the snow plowing problem, they reveal the shortcomings of the CPP model with respect to real-world constraints such as the need for multiple plows and multiple-lane roads. The final and most restricting consideration mentioned is that of road priorities for which they provide several heuristic methods to overcome: (a) a weighting method that would make higher priority roads more attractive, (b) partitioning the network into three class-consistent subnetworks and solving each individually, and (c) ignoring priority and hope for a good solution. They conclude with an example problem in which they present a *cluster first–route second* heuristic that uses a CPP model for routing.

The algorithm provided by Lemieux and Campagna (*15*) is intended to enforce priorities by tracing an Eulerian circuit on a directed graph and picking higher-priority roads first whenever possible. This algorithm would have to be implemented with some heuristic procedure because it assumes that the truck can traverse all the arcs in the network.

Other arc-coving applications that use the CPP model or heuristic procedures include the routing of electric meter readers (*16*), waste collection (*17*), and routing street sweepers (*18–20*).

Finally, system studies (*21*) and simulations (*22,23*) have been conducted for the snow removal problem to help modularize this complex problem so that subsets of the problem may be solved separately and evaluated individually or used to drive another solution procedure. Of particular interest is the use of simulation to help the decision maker, human or machine, evaluate routes produced by some procedure, as described earlier. This evaluation could stimulate recommendations for route modifications or regeneration with additional constraints.

In developing a strategy for winter season snow and ice removal, the goal of the state highway authority is to provide efficient service within the constraints on available resources; plowing and abrasive spreading equipment, sand and salt supplies, and manpower. Although holding down overall cost is a primary consideration (*15*), the safety of the public is the major objective (*24*). Public safety in this context has two distinct but related components: (a) the condition of the road surface (*22*) and (b) the performance of the snow removal fleet during the operation. An effective snow removal operation is one that provides rapid and orderly snow removal and abrasive application without excessive interference with public transportation activity (*22,24,25*).

## A MULTIOBJECTIVE HEURISTIC APPROACH

The CPP-based exact mathematical programming models presented in the previous section proved to be computationally infeasible as the sole solution procedure for generating snow removal routes. This was true for the case of one simple objective, the minimization of distance traveled, but with multiple objectives the exact models become all the more intractable and require alternative solution procedures.

A heuristic or nonexact procedure is an algorithm that will generate a feasible solution with no guarantee on optimality. The interested reader is directed to Chapter 7 of Parker and Rardin (*11*) for an overview and classification of nonexact algorithms with respect to discrete optimization. For the purpose of Indiana's snow route design problem, two possible heuristic approaches are investigated: (a) improving the existing route configuration, and (b) building routes exclusively from a definition of the network.

The snow route design problem may be defined as follows: develop the best set of routes for a given depot, subject to a given minimum level of service and with adherence to route class continuity. Best can be defined as a minimum cost configuration in which cost is measured by the number of trucks needed and total number of miles traveled. Levels of service vary for each class type and are defined in the INDOT policy (*26*). Adherence to route class continuity is not a binding constraint; it may better be defined as a secondary goal. From discussions with INDOT personnel from various levels, the predominate objective seems to be minimization of deadhead miles, because of the negative public reaction resulting from

a snow plow traveling any road while it is not plowing. The public would also like to see its respective streets plowed first, which translates into maximizing the level of service. Therefore, the public sector problem of designing snow routes is now better defined as minimizing cost (taxes) and maximizing the level of service (public satisfaction). Thus, the snow route design problem becomes one of having multiple and conflicting objectives.

There are many different ways to approach a multiple-criterion optimization problem, depending on the way the analyst and the decision maker wish to impose the decision maker's preference or utility on the problem (*27*). The decision maker may specify preferences for each of the objectives with respect to each other a priori, a posteriori, or interactively during the actual solving of the problem (*12*). All of these methods are based on a tractable initial single objective problem formulation and are therefore of little use to the solution procedure being developed in this research.

The multiple criteria theories referenced earlier are used in papers by Henig (*28*) and White (*29*) with respect to the classical network theory problem of finding a shortest path on a given network with the added consideration of multiple objectives. The multiple-objective shortest-path problem is important to this research because some heuristic routing procedures, including the one developed for this research, use or require a set of shortest paths at some point in the algorithm.

Several usable heuristic procedures are described for improving and designing routes.

### Improvement Procedures

One class of heuristic approaches is to randomly or arbitrarily create routes and try to improve them. Because INDOT has a set of routes that have historically evolved and were created by experts, a heuristic procedure that tries to improve these existing routes is probably the best approach. Two such methods are a swap-improvement heuristic and a route elimination heuristic.

The swap heuristic answers the basic question: Can the current routes be improved? Indicators of improvement could be reduction of total deadheading, better enforcement of route class continuity, greater route compactness, and other factors that experienced highway engineering personnel may provide. A heuristic algorithm of this type requires the network data previously mentioned and the current routes to be digitally encoded in a similar format. It then tries to modify the routes on the basis of the objectives listed by swapping arcs between routes or cutting down on deadheading. The multiple objectives may be considered by either an a priori weighting vector that would cause swaps that improved the weighted objective function the most to be done first, or by an interactive swapping procedure that relied on the decision maker to specify preferred improvements. This method results in more efficient and easier-to-drive routes that in turn result in a better level of service.

The elimination heuristic answers the question: Are all the routes needed that are currently used to service the network? This method requires the same digitally encoded data as the improvement method. It ranks and sorts the current routes on the basis of total and deadheading length, total number of arcs it covers, and nodes it shares with a route of the same classification. The algorithm will then try to eliminate a route by breaking it up and distributing its arcs among other routes. In this case, an a posteriori analysis of the reduced set of routes will have to be performed by the decision maker to evaluate the feasibility of the new configuration. If successful, this method results in fewer needed trucks, thus reducing one of the major costs involved in snow and ice control.

### Generation Procedures

Route generation procedures attempt to build routes from nothing more than the defined network. Of particular interest are the procedures outlined by Golden (*18*)—nearest neighbor, Clark, and Wright savings, insertion, nearest merger, and others—and the interactive set partitioning method developed by Cullen (*13*). An ad hoc route generation heuristic was developed and it will be the focus of the remainder of this section.

The seed node–based snow route generator is an ad hoc, nonexact procedure that reads in a definition of a network, prompts the user for a set of seed nodes on the basis of an empirical analysis, and tries to grow snow routes from the seed nodes, subject to the objectives and constraints outlined previously.

The underlying network for this algorithm is assumed to be directed; an arc for each lane with an associated distance and class. The data file has a format that includes an ID number, start node, and end note; length and class are provided for each line; and an ID number, number of arcs directed in and out, list of arc ID's for arcs directed in and list of arc ID's for arcs directed out are provided for each node. The algorithm stores the network data in two dynamically created arrays, one containing a structure or object that has the given data file information, the distance from and to the depot for the node, and two lists of pointers to each arc directed into and out of the node for each node in the network; and the other contains a structure or object that has the given data file information plus pointers to the start and end nodes. A complete discussion of the implementation of this scheme for network research is presented by Haslam (*30*).

The first step in the algorithm is to read in the network data from the data file previously mentioned and dynamically create the arrays of nodes and lines. In this step, all the node-to-arc and arc-to-node pointers found in these arrays are established. If $n$ denotes the number of nodes and $a$ denotes the number of arcs in the network, then computationally this first step is at worst $O(n*a)$. The next step is to calculate the remaining information needed by the algorithm, the all-to-all shortest paths by the Floyd-Warshall algorithm, known to be $O(n^3)$ in the worst case (*9*).

The next step in the algorithm is to do an empirical analysis on the network. This analysis consists of calculating the total number of lane-miles for each class of road, which, when divided by the maximum distance a truck may travel when servicing each class of road, yields the minimum number of trucks needed to service a given network. The user is also prompted for the location of the depot. This theoretical minimum is then used when prompting the user for the number

of trucks the user would like to try to generate routes for. In the worst case, this step is $O(a)$.

At this point, the user is asked to supply a seed node and its corresponding route classification for each truck. The user specifies the seed nodes on the basis of the spatial configuration of the network with regard to arc classifications. That is, seed nodes should be placed among clusters of arcs sharing a common classification. The placement of seed nodes is also driven by the experience of the user as the algorithm evolves through iterative trials. With this information, the array of route headers is initialized and each linked list has its first leg created. Each seed node is tested for feasibility by checking its distance from the depot against the maximum distance allowed for its route classification. In the worst case, this step is $O(n)$ and ends the initialization and user interaction phase of the algorithm. The remaining steps constitute the procedures that attempt to generate or grow routes on the basis of the seed nodes provided.

The grow-to-depot and the grow-from-depot phases of the algorithm each contain two substeps and are essentially the same procedures, with one working in a reverse direction. The purpose of these phases is to try to grow the route back to the depot from a seed node and to try to grow the route from the depot to a seed node. For the grow-to-depot phase, the algorithm starts adding arcs to the solution that are not already covered, are of the same class as the route that is currently being built, do not force the current route over its maximum distance constraint, and push the route closer to the depot. This process continues until the route includes the depot or there are no candidate arcs that satisfy the given conditions. In the second step of this phase, if the route does not include the depot, then the shortest path back to the depot is added to the solution ignoring class and whether the arcs used have been covered before; it is added to the route as deadhead coverage. This procedure is repeated for all of the routes and until they all contain the depot.

The grow-from-depot phase works in the same manner, but moves in the reverse direction. In both phases, the worst case will be $O(n*a)$. These procedures are denoted as Stage 1 and Stage 2 growth by the algorithm.

The next step in the procedure, expand-the-routes, is denoted as Stage 3 growth by the algorithm. At this point in the procedure, each of the routes consists of a shortest path to and from the depot for each seed node containing as many feasible arcs as possible in the solution. The algorithm now tries to add all of the currently uncovered arcs to the solution by expanding the routes so that they cover the desired arcs. In this Stage 3 growth, an uncovered arc will only be added to a route if it is the same class as the route, shares at least one node with the route, and if its inclusion in the route will not cause the route to exceed its length limitation. If only one node of the candidate arc is shared with the route, then all arcs adjacent to the node not in the route are checked to see if they (a) contain a node in the route, and (b) do not violate any of the previously listed constraints. The State 3 generation thus adds uncovered arcs one at a time or in pairs to adjacent routes on the basis of class and length restrictions. In the worst case, this phase could be $O(a^4)$.

The final step of the algorithm, denoted as Stage 4 growth, is a relaxed expansion procedure. This phase is the same as Stage 3 with the exception of relaxing strict class continuity. Any uncovered arcs left at this point become candidates for route inclusion as before, but now they must only have a class that is not greater than the route that is being considered. The idea behind this phase is that it is all right for a higher-class route to provide a better level of service to a lower-class route, but not vice versa. Again, in the worst case this phase could be $O(a^4)$.

At all phases of the algorithm, printouts are provided so that the user may track the progress of the routes as they are generated. The user has to provide different sets of seed nodes until a feasible or acceptable solution is reached. This algorithm often is not able to cover every arc in the network, but by intelligently varying the seed nodes and the number of trucks used, the user should be able to converge on a solution in polynomial time.

### Seed Node–Based Route Generation Example

An all Class I, two-lane network that was used as the base network for an example solution is shown in Figure 1.

On the basis of the empirical analysis, four seed nodes were chosen: Nodes 4, 11, 11, and 7. Node 11 was chosen twice because of its extreme distance from the depot. Figures 2–4 represent each stage of the route generation procedure.

For this example, the maximum route length was set at 45 mi. With 122 lane-mi total, the theoretical minimum number of trucks needed to clear the network is three. With seed nodes of 4, 11, 11, and 7, the four routes covered the entire network with a total of 162 mi traveled, or 40 mi of total deadhead. This simple example was presented to illustrate the proposed routing heuristic. In the following section, some results are described from the application of several different models and analysis procedures on an actual test site in Indiana.

## RESULTS AND DISCUSSION

The methods discussed in the previous sections have been tested using data from the Fowler subdistrict. INDOT, project

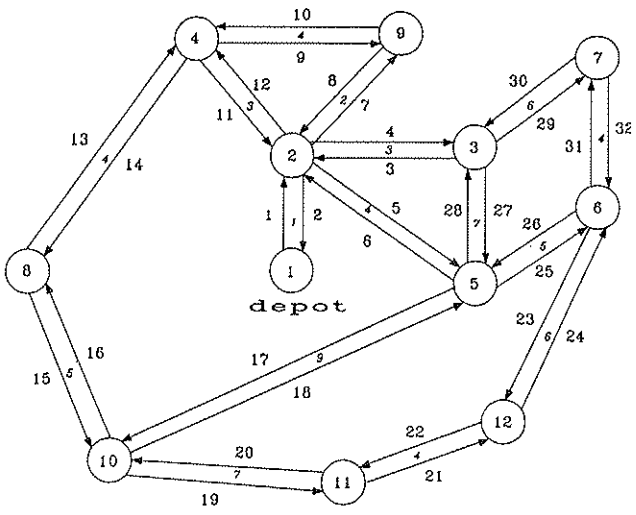



**FIGURE 1 Network representation for example problem.**

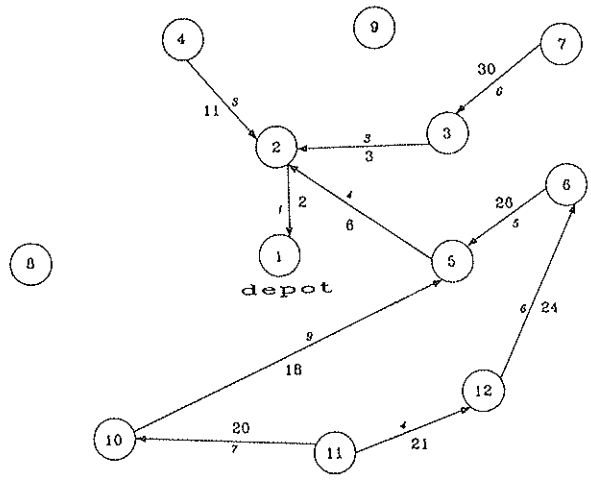



FIGURE 2 Solution after Stage 1 growth.

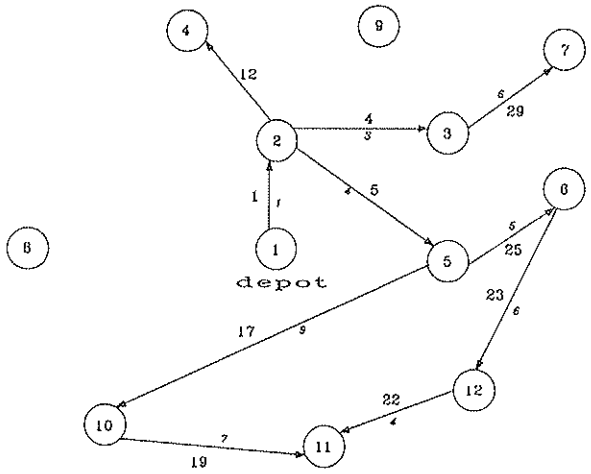



FIGURE 3 Solution after Stage 2 growth.

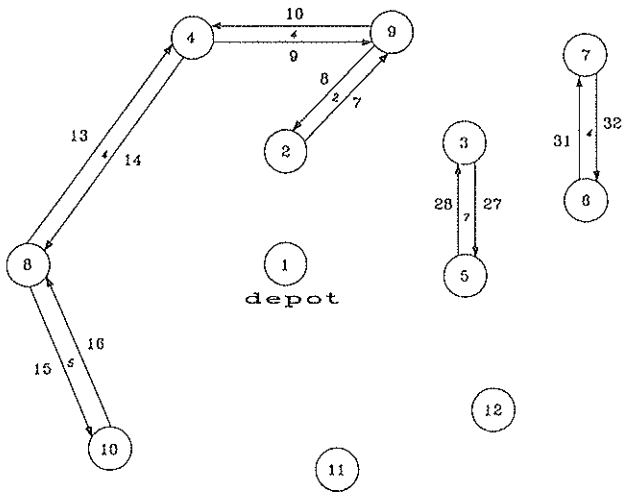



FIGURE 4 Solution after Stage 3 growth.

test site. At this point in the evolution of this project, the U.S. Geological Survey (*31*) digital data are still not fully compatible with the vehicle routing research of the previous sections, yet they are inherently important to any successful design or analysis previously outlined. To overcome this preliminary incompatibility some digital data was created with maps and a ruler. This effort has resulted in a 362-arc, 99-node link-node diagram that was hand encoded from a terminal. There is no precise indicator of the accuracy for these data. One limitation of the data is that only those roads currently used by INDOT in snow servicing were considered.

An empirical analysis of the Fowler subdistrict was undertaken to determine the minimum number of routes required to service the network. There are approximately 457 lane-mi of Class I roads, 271 lane-mi of Class II roads, and 188 lane-mi of Class III roads and Class I, II, and III routes cannot exceed 35, 50 and 65 mi, respectively. The Fowler subdistrict thus optimally (no deadheading) needs 13 Class I routes, 6 Class II routes and 3 Class III routes, or a total of 22 routes. The current INDOT configuration calls for a total of 26 routes or 4 extra routes more than the minimum needed. Given that there must exist some deadheading, a total of 26 routes is not far from a feasible optimal. This comes as no surprise given the fact that each set of routes was designed by a local expert, in terms of local geography and snow plow routing.

## Exact Models: The CPP Formulation

A CPP was formulated and run using the test data. The result was a linear program with 362 variables, 99 flow continuity constraints, and an optimal solution of 905.9 lane-mi required to service the network. Only 2.39 mi were deadheading.

The digital data are currently the hand-encoded network. In the future, digital data will be used as described by Kurmas (*32*) and Wright et al. (*1*). Network data structures are two interlinked linked lists of nodes and arcs and their respective attributes created by a program that reads the raw digital data and generates the lists of structures. These linked lists of structures can be used to generate an XML model or a mathematical program, or to feed a heuristic algorithm. All mathematical programming formulations were solved using the XMP mathematical programming software (*33*).

The results of the CPP model run are presented in Table 1, Column 3. These results shed some light on the underlying structure of INDOT's network of roads and the accuracy of the network data. Out of a total of 903.5 lane-mi of roads that consisted of 362 arcs, only 2.39 mi (nine arcs) were needed to zero the polarity at every node in the network. Therefore, the state highway and road system taken as a completely directed network is a good underlying network for forming Euler or CPP tours. In Table 1, the difference in total required lane-miles between Columns 2 and 3 points out the shortcomings of relying too heavily on hand-scaling the map information and not having a digitally calculated and encoded data base for the network data. The exact mathematical models reviewed in the previous sections may be used as the platform for future route generation schemes in the following contexts: (a) they could be considered as the underlying model for a previously mentioned heuristic enumeration scheme, or (b) given an appropriate graphical interface, they could be

TABLE 1 METHODOLOGY EVALUATION PARAMETERS

| Route Parameters (1) | Current Policy (2) | Theoretical Lower Bound via CPP (3) | Swap Heuristic Improvement (4) | Elimination Heuristic Elimination (5) |
|---|---|---|---|---|
| Total Required Lane Miles | 915.4 | 903.5 | 915.4 | 915.4 |
| Total Deadhead | 175.8 | 2.39 | 135.2 | 160 |
| Total Lane Miles | 1091.2 | 905.9 | 1050.6 | 1075.4 |
| # Routes | 26 | 1 | 26 | 25 |

used without the exponential number of subtour breaking constraints and produce possible routes for a user to evaluate and impose needed constraints.

## Heuristic Methods

The heuristic methods previously described hold the greatest promise for the actual generation of usable snow routes. Heuristics are especially well suited for problems that are a special case of a well-researched topic or algorithm, and snowplow routing with its road priorities and multiple objectives is a special and harder case of the general topic of vehicle routing and scheduling.

### *Improvement and Elimination Heuristics*

The swap-improvement and elimination heuristics were evaluated with hand-encoded data with the routes as currently defined by INDOT as the data source. In the process of imposing the INDOT route definitions on the network definition, many shortcomings or conflicts were discovered between the two. Although many of the conflicts were resolved by editing the hand-encoded data, some still exist that will probably require ad hoc consideration. This is yet another justification for a digital network data base.

The computational effort and routing expertise required for the definition and implementation of the swap-improvement heuristic made its implementation prohibitive in the current project. This method requires the explicit definition of tradeoffs or preferences with respect to the objectives of minimizing cost and maximizing the level of service. These preferences could be defined by actual examples that are best demonstrated in a graphical decision support environment. In this environment, the decision maker could specify or evaluate swaps and then see the implications in the context of how the particular swaps affected the entire set of routes. Column 4 of Table 1, when compared with Column 1, demonstrates the effects of what would be a successful run of a heuristic of this type.

A rudimentary elimination heuristic was hand tested with the Fowler data and routes. This procedure evaluated routes with a relatively large number of deadhead miles, a relatively small number of arcs actually serviced, and at least one neighboring route of the same class. The result was the elimination of one route. In eliminating this route, three others were consequently modified. The removal of the route resulted in a loss of 15.8 deadhead miles and a corresponding surplus truck. The three modified routes all remain within the prescriptions set forth in the documented policy. Column 5 of Table 1 lists the results of this implementation of a elimination heuristic.

### *Seed Node–Based Snow Route Generation Heuristic*

The ad hoc route generation heuristic developed for this research was tested on a subset of the Fowler Subdistrict data. The subset of data represents the arcs currently covered by one of the three depots in Fowler. There are 21 nodes and 54 arcs with all three classes of roads represented. Under the current INDOT route configurations, there are six routes used to cover this network, resulting in a total of 41.8 deadhead miles and 274.3 lane-mi traveled.

First, the heuristic could not cover all of the arcs in the network. However, the six routes currently used by INDOT contain two Class II routes that exceed the maximum allowable Class II length constraint by 10 to 13 mi each. The route generation heuristic currently contains no mechanism for relaxing the maximum distance constraints, but perhaps should, given the weak adherence to policy reflected by this example. For current purposes, a visual inspection of the routes produced by the heuristic and the underlying network indicates that the currently uncovered Arcs 41 and 42 would add 25.6 mi to any route to which they are added. Arcs 41 and 42 are adjacent to Route 4 (Class III, length = 58.9 mi) and Route 5 (Class II, length = 34.3 mi). With a relaxation of 15 mi for a Class II route, its maximum distance is now 60 mi, and the inclusion of Arcs 41 and 42 increases the length of Route 5 to 59.9 mi, but within the relaxed bound. Therefore, Arcs 41 and 42 are now covered by Route 5 with a new length of 59.9 mi.

This configuration of routes generated by the heuristic results in a total of 253.33 lane-mi traveled with 50.53 mi of deadhead travel and only one route violating its distance constraint. Therefore, it might be argued that although the heuristically generated routes contain more deadhead miles, they provide a higher level of service than the current INDOT routes. This demonstrates the tradeoff dilemma characteristic of multiple-objective problems, and is best evaluated by an appropriate decision maker.

The seed node–based snow route generator is an ad hoc heuristic procedure and consequently has several technical drawbacks:

1. Having the user specify the seed nodes used to generate the routes assumes that the user is familiar with the road network and thoroughly understands how the algorithm generates the routes. These requirements could be avoided by having the heuristic enumerate a set of candidate seed nodes and store the best sets of routes generated.
2. In its current form, the heuristic only looks for pairs of arcs to add to the solution when it is expanding the routes. In a rural setting, this is not a serious drawback because most of the network has zero-polarity nodes. In an urban setting with more complex interchanges and intersections this drawback may become a more serious problem. One way to overcome this deficiency would be to run a CPP algorithm on the original network and hand the heuristic the network with arcs added that ensured a CPP tour. The best way to remedy this fault would be to modify this part of the algorithm, a nontrivial task.
3. The only automated relaxation in the algorithm is that of route class continuity. There is a need for a route length relaxation procedure as was previously demonstrated and possibly for more interaction with the user while the problem constraints are being relaxed.
4. The algorithm is greedy. An enhancement for this would be to add an improvement module onto the procedure.
5. Other than finding shortest paths, the heuristic does no optimization. It is uncertain how this procedure might drive some sort of optimization model.

The major point in the heuristic's favor is that it is capable of generating feasible snow routes that, at least in one case, are comparable to existing implemented routes. Furthermore, it is quick, O(No. of arcs$^4$) in the worst case, and simple to understand and trace. Given its simplicity and problem-specific domain, it would be easy to modify to fit a particular snowplow route design problem.

A final point is that this heuristic algorithm is a relatively untested procedure that needs a few of the previously mentioned drawbacks ironed out before it could be compared with any other established vehicle routing solution procedure.

### Future Directions

Throughout this final section, the need for accurate and comprehensive digital network data has been stressed. The reasons are twofold.

Without a common source of reliable data, the evaluation of any route generation scheme can be questioned. Many of these route design procedures may be able to find a better route configuration by traversing a local or county road that is not part of the state network, but will allow a state road or roads to be more efficiently serviced.

The second justification for digital data is perhaps the most important practical point unveiled in this research for INDOT. The Indiana Department of Highways would benefit greatly from establishing some form of a roadway information system (*34*). In the context of this research, such a system could be the backbone of a snowplow routing decision support system. Vehicle routing and scheduling systems have already been reported (*35–37*).

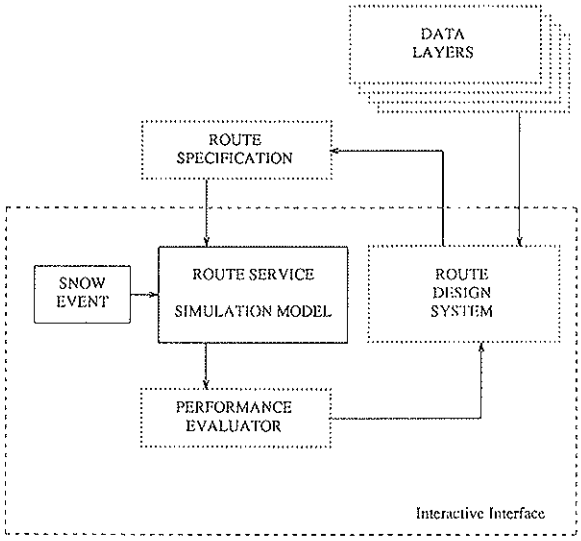



**FIGURE 5 Schematic representation of route design system.**

Set within the context of a decision support system as suggested by Figure 5, the complex and domain-dependent problem of snowplow route design could be simplified by intelligent user interaction and performance evaluation and analysis aids, such as a simulation model or a knowledge base.

## REFERENCES

1. J. R. Wright, E. P. Haslam, P. Kurmas, K. Buehler, and S. Benabdallah. *Development of a Prototype System for Snow Route Design and Management*. Purdue University, West Lafayette, Ind., 1988.
2. L. Bodin, B. L. Golden, A. Assad, and M. Ball. The State of the Art in the Routing and Scheduling of Vehicles and Crews. *Computers & Operations Research*, Vol. 10, No. 2, Pergamon Press Ltd., 1983, p. 211.
3. R. C. Larson and A. R. Odoni. *Urban Operations Research*, Prentice-Hall, Englewood Cliffs, N.J., 1981, pp. 383–427.
4. E. Minieka. The Chinese Postman Problem for Mixed Networks. *Management Science*, Vol. 25, No. 7, July 1979, pp. 643–648.
5. C. H. Papadimitriou. On the Complexity of Edge Traversing. *Journal of the ACM*, Vol. 23, No. 3, July 1976, pp. 544–554.
6. B. L. Golden and R. T. Wong. Capacitated Arc Routing Problems. *Networks*, Vol. 11, 1981, pp. 305–315.
7. M. Dror, H. Stern, and P. Trudeau. Postman Tour on a Graph with Precedence Relation on Arcs. *Networks*, Vol. 17, pp. 283–294, 1987.
8. K. Mei-ko. Graphic Programming Using Odd and Even Points. *Chinese Mathematics*, Vol. 1, 1962, pp. 237–277.
9. R. G. Parker and R. L. Rardin. *IE 537—Class Notes*. Perdue University, West Lafayette, Ind., 1986.
10. B. L. Golden. Introduction To Recent Advances In Vehicle Routing Methods. In *Transportation Planning Models*, M. Florian, ed. Elsevier, North Holland, 1984, pp. 383–418.
11. R. G. Parker and R. L. Rardin. *Discrete Optimization*, Academic Press, San Diego, Calif., 1988.
12. R. E. Steuer. Multiple Criteria Optimization: Theory, Computation, and Application. John Wiley, New York, 1986.
13. F. H. Cullen, Jr. *Set Partitioning Based Heuristics For Interactive Routing*. Ph.D. dissertation, School of Industrial and Systems Engineering. Georgia Institute of Technology, Feb. 1984, p. 216.
14. D. H. Marks and R. Stricker. Routing for Public Service Vehicles. *ASCE Journal of the Urban Planning and Development Division*, Vol. 97, Dec. 1971, pp. 165–178.

15. P. F. Lemieux and L. Campagna. The Snow Ploughing Problem Solved by a Graph Theory Algorithm. *Civil Engineering Systems*, Vol 1, Dec. 1984, pp. 337–341.
16. H. I. Stern, and M. Dror. Routing Electric Meter Readers. *Computers and Operations Research*, Vol. 6, 1979, pp. 209–223.
17. E. J. Beltrami and L. D. Bodin. Networks and Vehicle Routing for Municipal Waste Collection. *Networks*, Vol. 4, pp. 65–94, 1974.
18. C. Mandl. *Applied Network Optimization*. Academic Press, London, 1979.
19. L. D. Bodin and S. J. Kursh. A Detailed Description of a Computer System for the Routing and Scheduling of Street Sweepers. *Computers & Operations Research*, Vol. 6, Pergamon Press Ltd., Great Britain, 1979, pp. 181–198.
20. T. M. Liebling. Routing Problems for Street Cleaning and Snow Removal. *Models for Environmental Pollution Control*, R. Deininger, ed. Ann Arbor Science Publishers, Ann Arbor, Mich., 1973, pp. 363–374.
21. D. L. Minsk. A Systems Study of Snow Removal. In *Special Report 185: Snow Removal and Ice Control Research*, TRB, National Research Council, Washington, D.C., 1979, pp. 220–225.
22. W. B. Tucker and G. M. Clohan. Computer Simulation of Urban Snow Removal. In *Special Report 185: Snow Removal and Ice Control Research*, TRB, National Research Council, Washington, D.C., 1979, pp. 293–302.
23. R. England. Computer Analysis Ensures a Clean Sweep. *Surveyor*, Vol. 6, May 1982, p. 15.
24. G. L. Russell and H. K. Sorenson. A Value Engineering Study of Snow and Ice Control. In *Special Report 185: Snow Removal and Ice Control Research*, TRB, National Research Council, Washington, D.C., 1979, pp. 66–73.
25. T. M. Cook and B. S. Alprin. Snow and Ice Removal in an Urban Environment. *Management Science* Vol. 23, No. 3, 1976, pp. 227–234.
26. *INDOT 1985–86 Snow Packet*. Public Service Office, Indiana Department of Highways, Dec. 1985.
27. G. W. Evans. An Overview of Techniques for Solving Multiobjective Mathematical Programs. *Management Science*, Vol. 30, No. 11, The Institute of Management Sciences, Nov. 1984, pp. 1268–1282.
28. M. I. Henig. The Shortest Path Problem With Two Objective Functions. *European Journal of Operations Research*, Vol. 25, Elsevier, 1985, pp. 281–291.
29. D. J. White. The Set of Efficient Solutions for Multiple Objective Shortest-Path Problems. *Computers and Operations Research*, Vol. 9, No. 2, 1982, pp. 101–107.
30. E. P. Haslam. *The Application of Routing Technologies to the Problem of Snow Removal*. M.S. thesis, School of Civil Engineering, Purdue University, West Lafayette, Ind., 1988.
31. *Digital Line Graphs from 1: 100,000 Scale Maps*. Data Users Guide 2, U.S. Geological Survey, Reston, Va., 1985.
32. P. Kurmas. *Processing of Digital Line Graph Data for Decision Support Applications*. M.S. thesis, Purdue University, West Lafayette, Ind., 1988.
33. R. E. Marsten. The Design of the XMP Linear Programming Library. *Transactions on Mathematical Software*, Vol. 7, No. 4, Dec. 1981.
34. F. L. Mannering and W. P. Kilareski. The Common Structure of Geo-Based Data for Roadway Information Systems. *ITE Journal*, July 1986, pp. 43–49.
35. P. Duchessi, B. Salvatore, and J. P. Seagle. Artificial Intelligence and the Management Science Practitioner: Knowledge Enhancements to a Decision Support System for Vehicle Routing. *Interfaces*, Vol. 18, No. 2, The Institute of Management Sciences, March–April 1988, pp. 85–93.
36. B. L. Golden and A. A. Assad. Perspectives on Vehicle Routing: Exciting New Developments. *Operations Research*, Vol. 34, No. 5, Sept.–Oct. 1986, pp. 803–810.
37. J. Potvin, G. Lapalme, and J. Rousseau. ALTO: A Computer System for the Design and Experimentation of Routing Algorithms. *Publication 525*, Centre de Recherche sur les Transports, University of Montreal, June 1987.

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Integrating GIS and CAD for Transportation Data Base Development

JIN-YUAN WANG AND JEFF R. WRIGHT

An accurate and reliable transportation network data base is a prerequisite of a successful service route design system. The focus of this research is the design and development of an interactive procedure for use in the reclassification and verification of the U.S. Geological Survey's 1:100,000 digital line graph data for use in a system that automatically designs rural snow and ice control routes for the Indiana Department of Transportation. Details of the digital line graph data format and conversion routines are presented.

An important though often neglected factor in the development of strategies for delivery of public services is the design of service routes. The importance of effective route design in the public sector relates not only to the cost of service provided, but also to a variety of intangible benefits such as equity and quality of service. The quality of service provided by public institutions is often difficult to quantify (*1*). In contrast to the private sector, where engineering management objectives are usually specified in terms of economic efficiency, government agencies strive to provide the best level of service possible as measured by public welfare and safety. These performance criteria are generally difficult to quantify for most public service activities. An important factor, however, in the quality of services provided is the planning and management of those services. For an institution such as the Indiana Department of Transportation (InDOT), a key factor in effective planning and management is the ability to design efficient route configurations for the delivery of services. The removal of snow and ice from the intrastate highway system is a good case in point.

In developing a strategy for winter season snow and ice removal, the goal of InDOT is to provide efficient service within the constraints on available resources (plowing and abrasive spreading equipment, sand and salt supplies, and manpower). Although holding down overall cost is a primary consideration, the safety of the public is the major objective. Public safety in this context has two distinct but related components: (a) the condition of the road surface, and (b) the performance of the snow removal fleet during the operation. An effective snow removal operation is one that provides rapid and orderly snow removal and abrasive application without excessive interference with public transportation activity (*2*).

Many public sector engineering management problems may be formulated and solved using network-based models (*3*). Among the most complex problems are those that involve the routing of service vehicles for such items as trash collection, police patrol, road painting and cleaning, roadside weed control, and snow and ice control. Although a rich set of analytical methodologies has evolved to solve these problems, the data required for their applications are often enormous. For the design of snow and ice control routes on rural highways, for example, complete network distance and adjacency data are needed, including precise configuration information for all major intersections. The effort required to collect, verify, and digitize these data on an ad hoc basic generally precludes their use in model development and problem solving.

Although ad hoc efforts to develop data bases to support these models are not cost-effective, many general data collection efforts are being conducted that contain these data and in digital form. For example, the U.S. Geological Survey (USGS) presently provides complete digital line graph (DLG) data at a scale of 1:100,000 for the entire United States. Data layers at this scale exist for such man-made features as hydrography, roads, railroads, and transmission systems. Magnetic tapes containing these data may be purchased for any region for interest.

The focus of this research is the design and implementation of an intelligent data filter for use in conditioning these data in a cost-effective manner so that the resulting spatial data base is suitable for use in the automated design of service routes.

## DLG DATA CONSIDERATION

For the application of vehicle routing over the state highway network, a complete description of the network is required. This description must include all major intersections in each district, adjacency information that indicates which intersections are joined directly by roads, and the lengths of each of these roads. Planning for snow and ice control requires additional information, such as the width in lanes of every roadway, and more complete descriptions of intersections, including features such as turn lanes or traffic control devices. Currently, the most effective means of assembling such data has been to produce a network by reading the information from a highway map, and then to add additional arcs to the graph to account for multiple lanes. On a statewide basis, this process could prove prohibitive, and the accuracy of the resulting maps could be less than desired.

### Digital Map Data Sources

A major effort of this research has been to explore available digital map data sources, and to find the best methods to

School of Civil Engineering, Purdue University, West Lafayette, Ind. 47907.

convert these data to a usable network representation. One class of geographic data considered was a raster, or grid, format. Data gathered by aerial or satellite photography is in this form. A raster format is impractical for a network application because of the large amount of processing required to extract the roadway information from the data grid. Also, to obtain accurate distances along roads, the sampling size of the raster format must be fairly small, or the errors in calculation would be multiplied many times over.

The most promising format for this application is a vector format in which roadways are represented as polylines, that is, a set of short segments connecting two intersections. Many basic functions are simplified, especially those of finding adjacent intersections and road lengths. The source for vector data thus far has been the USGS's DLG format. Roadway information is available from this source for the state of Indiana at 1:100,000 scale. As part of the National Mapping Program, the USGS is creating this data base by digitization of existing 1:24,000 scale archive maps. Current estimates from the agency indicate that the coverage of the United States in 1:24,000 scale digital maps is only 2 percent complete for hydrography and transportation information. However, to complete a data base for the 1990 decennial census, the USGS has produced a set of digital 1:100,000-scale maps for transportation and hydrographic information. Because this research project was primarily concerned with representation of the highway network, the available 1:100,000-scale DLG format data is adequate. The critical information for use of these data in a geographic information system is the availability and expense of the distributed data, their logical format, and their accuracy. These issues will be discussed in the following sections.

The DLG data are a vector representation of the environment. Map information is distributed by the National Cartographic Information Center of the USGS in two formats: standard and optional. The primary difference in the formats is the internal measurement units used. In the standard format, all coordinates are given in pure integer values, which can be mathematically transformed into geographic coordinates. The USGS recommends use of the optional format distribution of the data, in which all coordinates are recorded in Universal Transverse Mercator (UTM). The optional format also includes a more complete set of linkages between node, line, and area information.

The structure of the distributed DLG file can be described as a heading section identifying the geographic location of the cell, followed by one or more categories of data (see Figure 1). The map heading includes the name, scale, creation and modification dates, and corner registration points of the map. This heading also includes edge matching information, indicating the extent to which adjacent maps are compatible. Each category within the DLG file begins with a record stating the category name, and the number of node, line, and area records comprising that category. Although the capability exists to represent 15 categories within map files, currently distributed maps contain one category per file. As a result, the files are much more easily managed.

The node, line, and area records are designed to provide the maximum amount of usable network information in as compact a form as possible. This information includes an internal identifier, an integer code unique to other nodes within the map in as compact a form as possible. Node records currently contain the following information:

```
USGS-NMD  DLG DATA - CHARACTER FORMAT - 06-10-86 VERSION
WATSEKA, IL IN                15 1984,      100000  S15
 .0000000000000000D+00  .0000000000000000D+00
 .1000000000D+01 .0000000000D+00 .0000000000D+00 .0000000000D+00
SW     40.750000 -87.250000      478895.54 4510824.90
NW     40.875000 -87.250000      478935.34 4524698.35
NE     40.875000 -87.125000      489466.18 4524675.85
SE     40.750000 -87.125000      489446.70 4510802.36
ROADS AND TRAILS     0  253  253 010  152  152 010  403  403  1
N   1  478895.54 4510824.90       2       0   0
 -385  386
N   2  478935.34 4524698.35       2       0   0
   -5    6
N   3  489466.18 4524675.85       2       0   0
   -9   10
N   4  489446.70 4510802.36       2       0   0
A   1  484185.94 4517750.36      38   0   1   0   0
 -386 -385  383  381  379  378  375  374  371  372  356  201
  179  156  124  104   72   54   39    9   10   -1    2   -3
   -4   -6   -5  -22  -47  -50   52  -71  -91 -111 -135 -166
 -229 -366
    0    0
A   2  484185.94 4517750.36       4   0   0   0   0
  386 -369 -384  385
A   3  484185.94 4517750.36       4   0   0   0   0
    8    5    6   -7
A   4  484185.94 4517750.36       4   0   0   0   0
   13  -10   -9  -12
L   1    5    6    1    6          2   0   0
  486938.89 4524686.33  488335.89 4524683.35
L   2    5    7    7    1          2   0   0
  486938.89 4524686.33  486161.65 4524687.99
L   3    8    7    1    8          2   0   0
  483621.64 4524690.88  486161.65 4524687.99
L   4    9    8    1    9          2   0   0
  480512.69 4524700.06  483621.64 4524690.88
```

**FIGURE 1 Sample DLG data file structure.**

- Coordinate pair: a pair of real numbers that represent the UTM coordinate that locates the node.
- Length of line segment list: an integer value representing the number of map lines incident to the node.
- Length of attribute list: an integer value stating the number of attribute codes assigned to the node.
- Line segment list: a list of integers that represent the internal identifiers of each line incident to the node.
- Attribute list: a possibly empty list of pairs of integers that describe the purpose of the node.

For an application of a network analysis algorithm, a node record must contain only the incidence list. However, to implement a geographic data display of these data, all coordinate information must also be maintained. Line information is stored in a format similar to that of the node. Each line record contains the following data:

- Internal identifier.
- Starting node: an integer representing the internal identifier of the starting node.
- Ending node: an integer representing the internal identifier of the ending node.
- Left area: an integer representing the internal identifier of the area to the left of the line, when traversed from start to end node.
- Right area: an integer representing the internal identifier of the area to the right of the line, when traversed from start to end node.
- Length of coordinate list: an integer value stating the number of points which constitute the line between the start and end nodes.
- Length of attribute list.

- Coordinate list: a list of pairs of real numbers that represent coordinates through which the line passes.
- Attribute list: a possibly empty list of pairs of integers that describe the purpose of the list.

The bounding area codes can be used to identify whether the line includes part of the neatline. A neatline represents the boundary of a map, and serves to clip all lines so that the map represents a regularly shaped area. In order to implement a network analysis algorithm, the distance between adjacent nodes must be known. Because the coordinates of data points in the optimal distribution format DLG is the UTM system, distance between adjacent nodes can be calculated as the total length of all segments making up a line. The attribute information should also be included in a problem data base, as the algorithm must be able to consider what feature of a map the line represents. Once again, to represent the map information in graphic format, the coordinate information must be maintained.

Area records are stored in the same structure as node records. The primary differences between the meanings of the various fields lie in the meanings of the line segment list. For an area record, the number of line segments making up the boundaries of the area is stored, rather than the length of the incidence list of a node. Also, the number of islands within the area is stored in an area record. This procedure allows the area to contain holes that could not otherwise be consistently represented in the DLG format. Network analysis does not require any of the area information from the map. For other tasks, however, this information could be managed and manipulated exactly as that of the node and line data.

### Accuracy of Distributed Data

The USGS does not quantify the accuracy of the DLG maps that it distributes. Because the maps are produced from original archived maps, few allowances can be made for changes to the map area since the last revision of the source data. Future goals of the National Mapping Project are to standardize all published maps, and to implement a 5- to 10-year cyclic inspection process. Most necessary changes to the map data base would be made after the entire national data base is digitized. Critical changes to the map data base may be made by the Survey on special request. Alternatively, the data user could attempt to modify his copies of the map data. The USGS does, however, make an effort to ensure that the digitized map correctly represents the source map.

The DLG data users guide (*4*) states that manually digitized source maps have a resolution of 0.001 in., with an accuracy of no less than 0.005 in. If the source map was digitized automatically, the resolution is 0.0013 in. The guide further states that the digitized data are plotted and compared to the source material. The plots are checked for positional accuracy and completeness of the graph. Attribute information for the map is entered manually using the original source map as a reference, and then checked by software to guarantee that only valid codes were entered. This process does not, and could not easily check the entered attributes to verify that they provide a necessary and sufficient description of the line. Topological fidelity or planarity of the data is verified by software. This procedure guarantees that no lines cross except at nodes, and that any line entering a node ends at that node. Edge matching of positions and attributes is said to be performed whenever the adjacent cells are available in the data base. Experimental experience with these data will indicate that this edge-matching procedure may not be sufficient.

## DLG DATA AND TRANSPORTATION NETWORK FOR VEHICLE ROUTING

The use of USGS DLG data for route design and analysis is hindered by two distinct factors. First, the DLG data set includes considerably more information than needed (about county roads and walking trails, etc.). Next, the information provided by DLG data may not be adequate for routing purpose. These concerns are discussed in detail in the following subsections.

### Important Characteristics of a Transportation Data Base

In developing a transportation network data base, a number of important characteristics must be considered independent of the positional accuracy and reliability of the data. These data base elements include the following concepts.

#### *Geometric Relationships*

An accurate coordinate system is the fundamental element necessary to represent the topologic relationships among roads and intersections.

#### *Highway Classifications*

In most of the cases, highways are categorized in terms of their importance. Higher-class highways are supposed to receive higher quality of services. For example, InDOT uses average daily traffic (ADT) as the basis for highway classifications (*3*). There are three different classes used by InDOT, plus a special class called "no service" (to specify road segments that can be traveled but that receive no service).

#### *Number of Road Lanes*

One of the most important considerations during the route design procedures is the number of lanes of a road, because each lane should receive a minimum level of service, especially for street sweeping and snow and ice control.

#### *Directions*

The direction of traffic flow is, of course, an important consideration in snow and ice control on the intrastate highway system. Maintenance equipment must generally follow exist-

ing traffic patterns. For example, in Figure 2 suppose that the path indicated by road segments linking (sequentially) Node 7 to Node 6 to Node 5 to Node 0 to Node 8 to Node 9 to Node 10, and path Node 4 to Node 3 to Node 0 to Node 1 to Node 2 represent two Interstate segments with an overpass crossing at Node 0. In this configuration, path Node 2 to Node 6, path Node 5 to Node 4, path Node 4 to Node 8, and path Node 9 to Node 2 are entrance and exit ramps. Although original Survey data make no direction distinctions, it is clear that travel along a path Node 2 to Node 6 to Node 5 would not be possible, and that a vehicle traversing the path Node 2 to Node 6 must necessarily visit Node 7 next.

The over- or underpassing intersection is another concern. For example, Node 0 in Figure 2 represents an over- or underpassing intersection. Therefore, the path Node 5 to Node 0 to Node 3 is not possible. One way to handle this situation is to delete this over- or underpassing node. (However, this solution usually implies loss of the planarity and the geometric relationship of the graph.) The restricted direction method mentioned earlier can also address this situation. For the ramp area shown in Figure 2, the following statements specify the over- or underpassing situation of Node 0.

- Node 5 is the only feasible destination if travel is from Node 8 to Node 0.
- Node 3 is the only feasible destination if travel is from Node 1 to Node 0.
- Node 8 is the only feasible destination if travel if from Node 5 to Node 0.
- Node 1 is the only feasible destination if travel is from Node 3 to Node 0.

### *Turn-Around Point*

The possibilities for making turns at a particular node (location) is important from the standpoint of a route designer. Any intersection of two roads is considered a candidate for making a turn. Also, any median crossovers existing on highways may be other possible candidates. Information regarding the locations and maneuverability at such locations are important for any routing procedure.

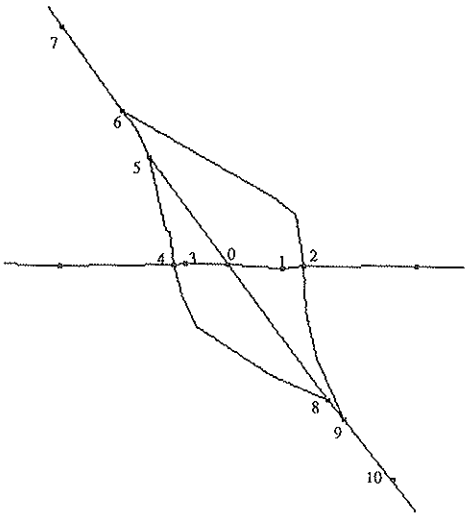



**FIGURE 2 The DLG presentation of a ramp area.**

### *Data Compactness*

Another serious concern of applying network-flow-based models to the real-world routing problems is the computational intractability caused by the excessive data size. It is desirable to represent the service area by using a minimum amount of data to reduce computational complexity.

## Limitations of Survey 1:100,000 DLG Data

Although Survey DLG data provide excellent coordinate systems to represent geometric relationships among roads and nodes (intersections), the USGS DLG data have some shortcomings. The USGS DLG data set includes considerably more data than needed for route design. A portion of the transportation network as represented by USGS DLG data for an area of Indiana is shown in Figure 3. The left panel displays a complete Survey 1:100,000 data set for a representative area in Indiana including all road map neat lines and boundaries. Each road is made up of sets of vectors, with nodes positioned at each vector intersection. Each vector and node is assigned a specified classification according to Survey conventions. The right panel displays only those roads and intersections of interest in the design of snow and ice control routes for the same area. Most of the vector and node information has been removed. In general, approximately 94 percent of the original Survey data are unnecessary for this application.

In addition, USGS DLG data cannot address most of the characteristics listed earlier. For example, the number of lanes, crossover positions, and restricted directions are lacking in the existing USGS data or are not be represented appropriately. In order to overcome shortcomings, an interactive filter program has been developed that is introduced in the next section.

## AN INTERACTIVE FILTER FOR DLG DATA

In order to overcome the limitations of the USGS 1:100,000 DLG data set in support of automated route design, an in-

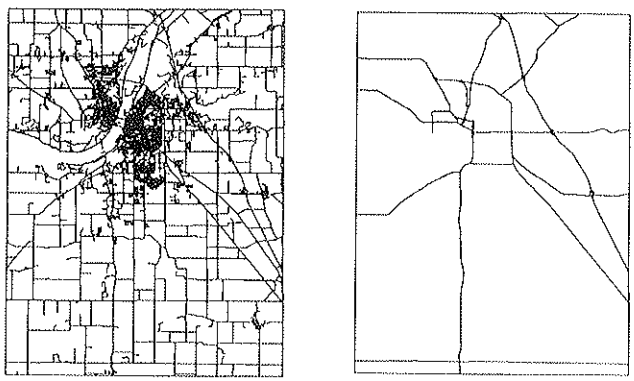


**FIGURE 3 DLG data reduction for route design GIS.**

teractive data filter has been designed and implemented. This system is designed for use by transportation engineers within the maintenance division of InDOT with the goal of developing a comprehensive state transportation data base that can be accessed by computerized route design algorithms. The structure of this interactive data filter is shown schematically in Figure 4. It consists of three separate functions: (a) input of the original DLG data files (b) interactive reclassification of road segments and node (intersection) objects, and (c) output of reclassified data files in DLG format. This filter uses USGS DLG format as its input and output format. That is, the output data of this filter program can be read by any software that can accommodate DLG data format.

### Data Input

The original USGS DLG files are used as the input of the data filter program. There are two steps that must be considered during this stage.

#### *Maps Selection*

USGS divides each 30-minute area into 4 or 16 separated files, and refers to each file by a name. By pointing a mouse cursor at any map panels, the user selects the appropriate files for filtering, and those files will subsequently be displayed as for interactive manipulation. A representation of such a mechanism is shown in Figure 5.

#### *Neat Line and Maps Merge*

Boundary lines (called neat lines) are used by USGS to locate edge nodes and to maintain the topologic continuity among DLG files. Because it is necessary to merge the selected files into a single map, some of the neat lines are redundant and need to be removed. The concept of the neat lines removal and map merge is shown in Figure 6. Notice that some of the neat lines still need to be kept for the future possible merge.

### Network Object Classification

The ultimate goal of this data filter program is to produce a suitable data set that can represent the network properly. Road and node specification play two major roles in this data filter program. An interactive method is presented for the purpose of road and node reclassification.

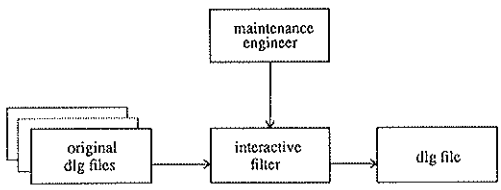



**FIGURE 4 Schematic of the data filtering process.**

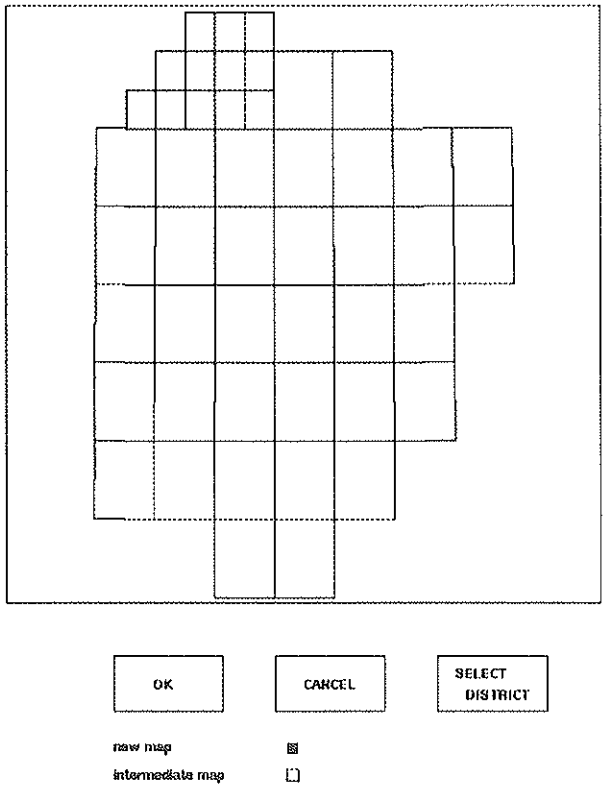



**FIGURE 5 Map selection stage.**

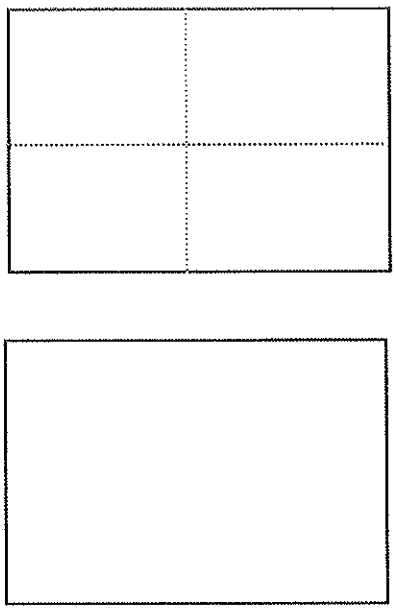


**FIGURE 6 *Top*, before neat line removal, *bottom*, after removal.**

#### *Road Classifications*

The sequence of activities conducted during road reclassification is shown in Figure 7. In this application, five major road classifications are required: Classes 1, 2, and 3, based on ADT data; no-service roads, which do not require service

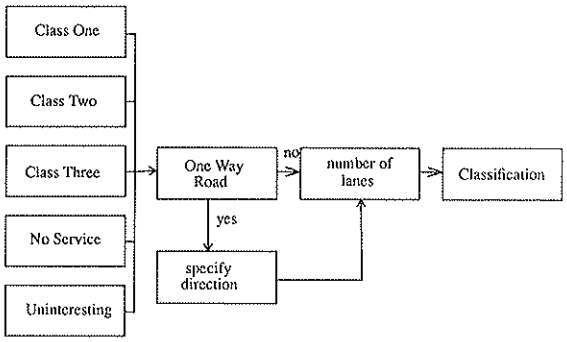



**FIGURE 7 Road classification.**

but can be used to connect two highways; and all other (uninteresting) roads. Separate functions are provided for each group of classifications.

*Restricted Directions*

The restricted direction is specified by selecting three nodes in a row. For example, a restricted direction is declared if we select three nodes, Node 1, Node 2, and Node 3, as a sequence. This means that if travel is initiated from Node 1 to Node 2, continued travel to Node 3 is feasible.

*Node Classifications*

The sequence of activities conducted during node reclassification is shown in Figure 8. Node classification is used to label nodes and to indicate maneuver restrictions at individual nodes such as turn-around restrictions. In addition, it may be necessary to add nodes to the network representation such as to add locations not contained in the USGS data files. Node classification may be important for vehicle maneuverability in snow and ice control operations.

There are five different functions provided for node specification: unit location, no turn around, add new node, label node, and uninteresting. The unit location function is used to designate the location of a storage or maintenance facility. The no-turn node is used to specify locations where a U-turn

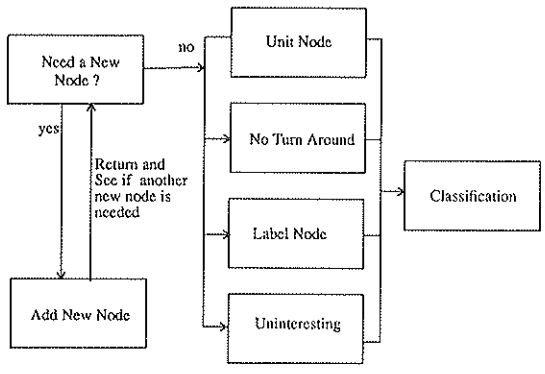



**FIGURE 8 Node classification.**

is not allowed. Label node is used by users to mark any node for future reference.

**Utilities**

In addition to the basic road and node classifications, this data filter also provides several convenient functions to help user conduct the filtering task. These capabilities include:

- Zoom: zoom in or out of a certain area.
- Highlight: highlight two certain road simultaneously in order to help users identify certain roads.
- Information: show information of a certain road or a particular node.
- Showme: show the current status of classification.
- Check direction: verify restricted directions.
- Save: saves current configuration every 20 min.

**Output**

The output stage of the interactive data filter consists of two components: (a) a debugging function, and (b) a routine that prepares the completed data set for subsequent use. The debugging function is designed to verify each road and node segment and to report inconsistencies to the user. The data preparation routine writes the completed map to a standard USGS DLG–3 format. The only difference between the final data representation and the original is that the major and minor attribute codes have been changed, and that the area information is discarded because it is not needed to accomplish route design functions. The files may now be read and manipulated by any program environment that can accommodate this format.

## CONCLUSION

The routing of vehicles for snow and ice control is perhaps the most difficult task of public sector routing problems, yet the data required for solving this problem are not significantly different than those required for other service planning and management problems. Consequently, accurate and complete representation of the physical network suitable for snow control vehicle routing will be of use to maintenance engineers for a wide range of applications such as scheduling and routing of mowing, painting, weed control, facilities and equipment servicing, inspection, and possibly some pavement maintenance activities.

Experience has demonstrated that the USGS 1:100,000 data represent an excellent and cost-effective source of transportation network data for use in applications such as route design and evaluation. The necessary filtering and reclassification of those data can be achieved with interactive computer support routines used by maintenance personnel knowledgeable about the snow and ice control operation and the physical features of the network. In Indiana, field engineers at InDOT district offices are completing this activity, and the quality of the resulting network representation is excellent. In addition to

data conditioning, inaccuracies and deficiencies in the original USGS data are easily identified and repaired.

## REFERENCES

1. D. H. Mark and R. Stricker. *Routing For Public Service Vehicles.* ASCE, Urban Planning and Development Division, Dec. 1971, pp. 165–178.
2. *The Indiana Department of Highways 1985–86 Snow Packet.* Public Service Office, Indiana Department of Transportation, Indianapolis, 1985.
3. B. L. Golden. *Introduction to Recent Advances in Vehicle Routing Methods.* In *Transportation Planning Models*, M. Florian, ed. Elsevier Science Publishing Co., Inc., N.Y., 1984, pp. 383–418.
4. *Digital Line Graphs from 1:100,000 Scale Maps, Data Users Guide 2*, U.S. Geological Survey, Reston, Va., 1985.

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Negatively Buoyant Jet (or Plume) with Applications to Snowplow Exit Flow Behavior

WILLIAM R. LINDBERG AND JOSEPH D. PETERSEN

The initial findings of an ongoing study of negatively buoyant jet (or plume) behavior are described. The motivation for this research is to quantify the dynamics of the exit snow plume from displacement and rotary snow plows. The effects of injection angle and cross flow on the jet (or plume) behavior were of particular interest in this phase of the study. A water tow tank was used for the experiments, in which the exit jet was towed through quiescent water at a constant velocity. Photographic records of the jet (or plume) structure were used to measure the jet (or plume) dimensions. Dimensional reasoning yielded a set of dimensionless parameters that correlated the jet or plume length scales over a wide range of the experimental parameters. Three flow regimes have been identified, which depend on the Froude number, $F_0$, of the cross flow: $F_0 << 1$, negligible cross flow; $0 < F_0 < 1$, weak cross flow; and $F_0 > 1$, strong cross flow; where $F_0 = U_0/(g_0')^{1/2}$, $U_0$ is the cross-flow velocity, $g'$ is the reduced gravity, and $r_0$ is the jet radius. Correlations of the measured length scales with $F_0$ were determined for all three flow regimes.

An important aspect of the performance of snowplows is the behavior of the snow after it exits from the moldboard. The exiting snow experiences a different dynamic environment from the flow conditions along the moldboard itself. At the exit plane, the flow has acquired considerable momentum in the transverse direction (i.e., normal to the plow motion). The flow then persists as a free jet whose trajectory depends on this momentum and the modifying external forces. The forces on this exit fluid include gravity and a complex three-dimensional interaction with the surrounding air, in the forms of entrainment and drag.

Knowledge of the exit plume flow for a variety of operational conditions is viewed as an important element in understanding plow performance, including cast distances, plume trajectories and dilutions, visibility, and the potential for altering these characteristics by plow design and operator control.

Little prior work applicable to exit plume behavior has been reported. In order to begin to address this shortcoming, a laboratory program was initiated to systematically examine the characteristics of negatively buoyant jet (or plumes) subjected to cross flows. A fundamental approach to this study has been adopted, in which the full complexity of the problem is initially reduced to allow focus on the basic flow phenomena. Additional parameters may then be introduced and their effects placed in the context of previous observations.

Department of Mechanical Engineering, University of Wyoming, Laramie, Wyo. 82071.

## BACKGROUND

The wide variety of circumstances occurring for the behavior of a jet (or plume) whose density is different from its surroundings has stimulated research on this topic for decades. The majority of this work has centered on the positively buoyant case, for which the density increment of the jet (or plume) provides a body force in the same general direction as the initial injection direction (i.e., a light fluid injected upward, or a heavy fluid injected downward). Examples of a positively buoyant jet (or plume) include a cumulus cloud, an industrial smokestack, and the rising column of smoke from a campfire.

A negatively buoyant jet (or plume) has a body force in the opposite direction from the vertical component of the injection velocity. In this case, equilibrium heights or flow reversals are always present. The term "vertical component of the injection velocity" is used because these jets can be aligned at an angle to the vertical. Examples of this type of flow other than the behavior of a plume of snow particles after exiting from a snowplow blade include the vertical injection of a dense gas into the atmosphere and the introduction of saline water into a lake or river.

A distinction can be made between a jet and a plume. A jet is a flow caused by a source of concentrated momentum such as the flow induced by the discharge from a pipe or orifice. The driving force for a plume is the downward body force or upward buoyancy of the plume itself. As an example, an initially positively buoyant jet in a stream expands, slows down, and dilutes with increasing downstream distance. Eventually, the momentum of the fluid decreases to the same order as that caused by the buoyancy. As buoyancy begins to dominate the flow dynamics, the jet becomes plume-like and behaves differently. The same processes are present for a negatively buoyant jet (or plume), with the distinction of the reversed body force for the plume, which retards and ultimately reverses the vertical motion.

The presence of an ambient cross flow further complicates this process. This external flow can have the effect of turning the jet (or plume) flow toward the downstream direction and will also alter the internal flow behavior of the jet (or plume) itself.

Laboratory studies of turbulent jets and plumes have been useful in applications in which the prototype and laboratory scales are considerably different. Dimensional and dynamic similarity arguments are insufficient to justify this success without the underlying premise of Reynolds number independence in these turbulent shear flows. It has been observed

that the large-scale turbulent motions are dynamically similar for flows with widely differing Reynolds numbers. The large-scale motions dominate such processes as entrainment, momentum transfer, and mixing. All of these processes are central to the dynamical behavior of jets and plumes. Laboratory studies in which the basic shear flow is turbulent may then be used to stimulate flows whose Reynolds numbers are different, but whose basic behavior is similar.

A dimensional analysis of jets and plumes may be performed to identify the basic parameters of these flows. Figure 1 shows a schematic of the geometry of the negatively buoyant jet (or plume) in an ambient flow. The jet of density $\rho_j$ exits the nozzle of radius $r_0$ at a velocity of $U_j$ into ambient fluid surroundings of density $\rho_a$. The ambient flow velocity is $U_0$. The axis of the jet is inclined an angle $\theta$ from the horizontal in a plane perpendicular to the ambient flow direction.

The two most important dynamical parameters of the initial jet are the jet momentum $M$ and buoyancy $B$ fluxes (normalized with jet exit density):

$$M = \pi r_0^2 U_j^2$$

$$B = \pi r_0^2 U_j[(\rho_j - \rho_a)/\rho_j]g$$

where $g$ is the gravitational acceleration.

The desired observable geometric parameters include $\mathbf{X}_m$, $\mathbf{W}_m$, $\mathbf{X}_t$ and $\boldsymbol{\alpha}$ where $\mathbf{X}_m$ is the $(x,y,z)$ length scale triad when the jet is at its maximum height, $\mathbf{W}_m$ is the jet width $(W_y,W_z)$ at the same location, $\mathbf{X}_t$ is the $x,y)$ length scale for the return of the jet (or plume) centerline to the original height of the jet, and $\boldsymbol{\alpha}$ is the spread-rate parameter ($dW_y/ds$, $dW_z/ds$, where $s$ is directed along the jet or plume) centerline. Dynamic parameters such as velocity scales and dilution rates are also of importance but have not been measured in these studies.

Dimensional analysis, using the defined parameters, yields the following functional set of dimensionless parameters:

$$\mathbf{L}^* = F_j\mathbf{f}(F_0,\mathrm{Re}_j,\theta) \tag{1}$$

where

$\mathbf{L}^* = \mathbf{L}/r_0$ (or $\alpha$),
$\mathbf{L}$ = dimensional length scale (i.e., $\mathbf{X}_m$, $\mathbf{W}_m$, or $\mathbf{X}_t$),
$F_j = U_j/[g'r_0]^{1/2}$ (the jet Froude number),
$F_0 = U_0/[g'r_0]^{1/2}$ (the cross-flow Froude number),
$\mathrm{Re}_j = 2U_jr_0/\nu$ (the jet Reynolds number),
$g' = [(\rho_j - \rho_a)/\rho_j]g$ (the reduced gravity), and
$\nu$ = kinematic viscosity.

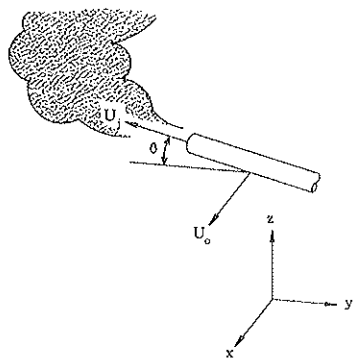



**FIGURE 1 Basic geometry of the negatively buoyant jet (or plume) that is translating relative to the ambient fluid.**

The assumption of Reynolds number independence for sufficiently large $\mathrm{Re}_j(\mathrm{Re} \gtrsim 500)$, reduces the number of independent parameters to three: $F_j$, $F_0$, and $\theta$:

$$\mathbf{L}^* = F_j\mathbf{f}(F_0,\theta) \tag{2}$$

All the experimental data are presented in this form.

Three vertical length scales for this flow may be identified (*1,2*). These length scales are based on the vertical scale where there is a dynamical balance of the various forces acting on the jet or plume:

$$l_b = (M \sin\theta)^{3/4}/B^{1/2}$$

$$l_m = M^{1/2}\sin\theta/U_0$$

$$l_p = B\sin\theta/U_0^3 \tag{3}$$

These length scales represent the vertical transition points between various modes of the jet or plume behavior:

$z \sim l_b$ jet-plume transition (for negligible ambient flow),

$z \sim l_m$ transition point for bent-over jet in an ambient flow, and

$z \sim l_p$ transition point for bent-over plume in an ambient flow.

Not all of these length scales are independent, for example:

$$F_0 \sim (l_b/l_p)^{1/3} \sim (l_b/l_m) \tag{4}$$

The effect of the ambient crossflow may then be characterized in terms of the magnitude of $F_0$, as follows: $F_0 << 1$, negligible cross flow; $F_0 < 1$, weak cross flow; and $F_0 > 1$, strong cross flow.

In addition, the ratio of $l_b$ to $r_0$ is proportional to $F_j$:

$$l_b/r_0 = (\pi\sin\theta)^{3/4}F_j \tag{5}$$

Because there is a transition region of approximately 10 jet diameters for the basic jet shear flow to develop, small values of $F_j$ would indicate that the flow was still developing when the buoyancy forces were beginning to be important. Low values of $F_j$ were avoided.

At this point, it is appropriate to point out those parameters that have not been included in this study. The four most important parameters are the crossflow Reynolds number, $\mathrm{Re}_0$; the fluid density ratio, $S_\rho$; the azimuthal jet angle in a vertical plane relative to the plane normal to the crossflow direction (identical to a spherical coordinate system), $\phi$; and the nozzle geometry.

$$\mathrm{Re}_0 = U_0l_s/\nu$$

$$S_\rho = \rho_j/\rho_a$$

where $l_s$ is the transverse length scale of the jet or plume.

Studies have clearly shown that cross flows directly alter the internal dynamics of jet or plume flows, but the resulting flow behavior is relatively insensitive to variations in $Re_0$, provided $Re_0$ is sufficiently large.

For the present experiments, the flow is approximately Boussinesq [i.e. $S_\rho \sim O(1)$], so these experimental results should be adequately molded without the inclusion of $S_\rho$ as an independent parameter. The density ratio is incorporated in the buoyancy flux parameter, $B$, as it has been defined, which does reflect non-Boussinesq effects.

This study was intended to provide the benchmarks for comparison to other studies incorporating such effects as large density ratios, azimuthal injection angles, and noncircular exit jet geometries. At present, the importance of these effects remains unresolved.

## RELATED STUDIES

A limited amount of study of snowplow plume behavior has been performed. As has been pointed out by Minsk (*3*), the cast distance ($Y_{max}$) for a snowplow has historically been modeled by equations of the form

$$Y^* = \eta_d g(\theta,\phi)\ (\eta F_0)^2 \tag{6}$$

where

$\eta_d$ = dynamic efficiency,
$g(\theta,\phi)$ = sin 2$\theta$ sin $\phi$, and
$\eta$ = moldboard efficiency ($\sim U_j/U_0$).

This functional form of the cast distance has its theoretical origin in simple particle trajectory analysis. The effects of drag, entrainment, turbulence, and density are accounted for in the dynamic efficiency parameter, $\eta_d$. For example, Vinnicombe (*4*) reports $\eta_d$ to range from 0.53 to 0.68. Shalman (*5*) cites the following form for $\eta_d$:

$$\eta_d = \tanh\ [a_0(H\rho_s)^{1/2}] \tag{7}$$

where

$a_0$ = 6.2 (cm$^3$/g-m)$^{1/2}$,
$H$ = undisturbed snow depth (m), and
$\rho_s$ = undisturbed snow density (g/cm$^3$).

Such a formulation is probably most appropriate for wet, cohesive snow, where the plume entrains little, if any, surrounding air, similar to the behavior of a water jet. The appropriateness and limitations of this type of approach are not at present resolved.

In contrast, an exit plume composed of dry, cohesionless snow will behave as a heavy fluid, characterized by strong entrainment, drag, and mixing with the surrounding fluid. It is this situation that is most appropriately modeled by the approach of the present research.

The fluid dynamical modeling of snow avalanche motion using water tanks has been discussed by Hopfinger and Beghin (*6*) and Hopfinger (*7*). They were successful in properly scaling many of the observed features of avalanche behavior in the laboratory. The non-Boussinesq effects of $S_\rho > 1$ were examined and found to be relatively minor up to $S_\rho \sim 10$.

Studies of negatively buoyant jet or plumes are quite limited. In particular, only a few laboratory studies of these flows have been reported (*8–16*).

An experimental study of a vertically injected dense jet in a water tank was reported by Turner (*9*), for the case of 0 cross flow. Unfortunately, Turner incorrectly reported his correlated results, and the error has been propagating in the literature since that time. The correct result for the maximum rise height of the vertically injected dense fluid is

$$Z^*_{max} = 4.17\ F_j \tag{8}$$

This result is consistent with the dimensional arguments of Equation 2.

The work of Hoot et al. (*12,13,17*) is of particular significance, both in its scope and in the quality of the experimental measurements. They report on wind tunnel studies of dense gases injected vertically in a low turbulence flow. Their quiescent flow studies were in complete agreement with Equation 6. Density ratios, $S_\rho$ of between 1.5 and 3 were used. An integral analysis for the case of a cross flow was performed and compared to their experimental results. The functional agreement was excellent, and their data correlate well with the following relationships:

$$Z^*_{\mathbb{C}} F_j = 2.1 F_0^{-1/3} S_\rho^{1/3} \tag{9}$$

$$X^*_{max}/F_j = F_0 \tag{10}$$

where $Z^*_{\mathbb{C}} \equiv Z' = Z_{\mathbb{C}}/r_0$ (at the plume centerline). Other measurements, including concentration distributions and longitudinal touchdown distances, were also reported.

On the basis of a very brief review by Ooms and Duijm (*18*), it would appear that the work of Anderson et al. (*14*) is of direct relevance. Unfortunately, the work was never published and a copy of their report has not been obtained. They used a water flume and varied $\theta$ between 45 and 90 degrees.

The response of a neutrally buoyant jet in a cross flow has been studied extensively. A summary to 1981 has been given by Crabb et al. (*19*). The characteristics of the jet deflection entrainment rates and mean turbulent flow properties have all been measured and documented. The observation of a double vortex structure in the downstream portion of the jet has intrigued investigators for years [see Abramovich (*20*) and Keffer and Baines (*21*)]. The interest lies both in the mechanism of generation and in the subsequent influence on the jet's boundary shape, trajectory, and entrainment rate.

In terms of numerical or analytical models of negatively buoyant jet or plume behavior, the survey by Hanna and Drivas (*22*), the volume edited by Britter and Griffiths (*23*), and the brief review by Ooms and Duijm (*18*) represent most of the currently available models and provide a comparative assessment of these models to the fairly small experimental data base.

## DESCRIPTION OF THE EXPERIMENT

A tow tank was used to simulate the ambient flow. A round tube (the jet source) was towed through the tank at a constant

velocity. The use of a towing facility ensured a uniform, repeatable, and low turbulent mean flow (as seen by the jet or plume). The working fluid in the tow tank was water; the tank dimensions were 394 cm long by 40.6 cm wide and 50.8 cm deep. Salt water of various densities was used for the injection fluid. The water jet was mounted 10 cm above the bottom of the tank so that multiple test runs may be made with the same ambient fluid. The water jet was attached to a tow carriage that could be moved at a constant velocity along the tank's horizontal axis. The speed of the carriage could by varied between 0 and 12.4 cm/sec. The jet flux was controlled by a valve located between a constant head tank and the jet nozzle. The flow rate was measured by a tapered-tube flowmeter (calibrated for the variation in the fluid density). The resulting jet behavior was recorded photographically with two cameras, located at the side and end of the tank. The side camera used a shadowgraph technique to record the projected image of the jet (or plume) density structure on an opaque grid surface. The end camera recorded an illuminated slice of the jet that was produced by directing a sheet of light from the side of the tank and using a small amount of white tempera paint in the jet solution. Examples of the images recorded by this technique are shown in Figures 2–4.

The experimental program consisted of 260 separate tests. The ranges of the experimental parameters used in these studies follow:

| *Parameter* | *Range of Values* |
|---|---|
| $V_j$ | 80–280cm/s |
| $U_0$ | 0–12.4 cm/s |
| $\theta$ | 30°–90° |
| $\Delta\rho/\rho_j$ | 0.01–0.20 |
| $r_0$ | 0.046–0.24 cm |
| $Re_j$ | 870–9220 |
| $Re_0$ | 0–700 |
| $F_j$ | 12–170 |
| $F_0$ | 0–15 |

The values of the Reynolds numbers are based on jet diameter.

A typical experiment began with the installation of a selected nozzle at a predetermined angle setting. The tow tank was then filled and allowed to reach thermal equilibrium with the laboratory overnight. The vertical temperature structure was measured prior to testing. Thermal stratification within the tank never exceeded 2°C/m, so that internal wave motion and stratification within the tank were not important. Approximately 2 L of water in the tank temperature was placed in the constant head tank. Sodium chloride was added to the water in this supply tank until the desired jet density was reached. Density was determined with a hand-held refractometer. A small amount of white tempera paint was then added to the mixture in negligible amounts to affect the density. A series of tests at various towing speeds and jet velocities was then made, without the necessity to drain the tank. The tow tank density was monitored during these tests to ensure that the tank density above the level of the jet remained unstratified.

Photographic enlargements of the side and end views were made for each run. Measurements of the various length scales were then made from the photographs. A single operator

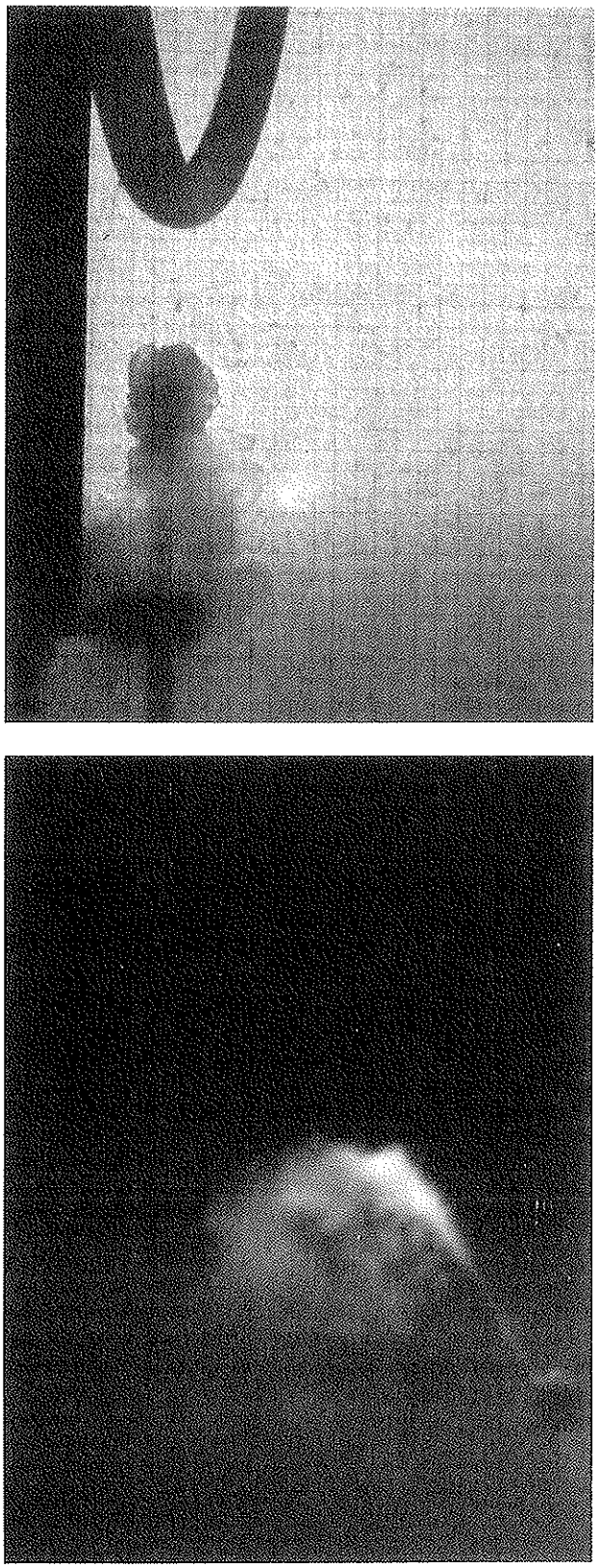


**FIGURE 2 Side (*top*) and end (*bottom*) views of a jet (or plume) for the case of negligible cross flow; $F_j$ = 25.2, $F^0$ = 0, $Y^*$ = 67.2, $Z^*$ = 60.2, $\theta$ = 60 degrees. The grid in the side (shadowgraph) image is 1 cm square.**

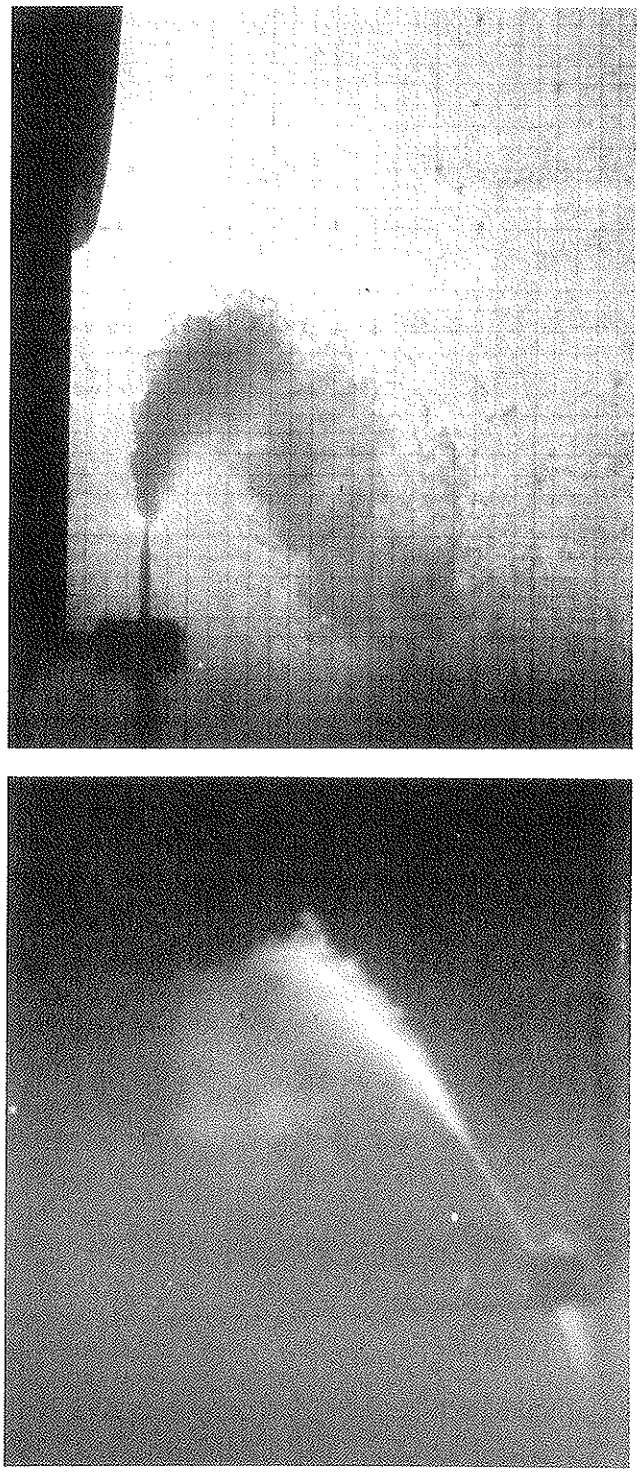


**FIGURE 3 Side (*top*) and end (*bottom*) views of a jet (or plume) for the case of weak cross flow; $F_j$ = 28.9, $F_0$ = 0.51, $X^*$ = 34.9, $Y^*$ = 54.6, $Z^*$ = 81.3, θ = 60 degrees.**

performed all of the measurements. In order to provide a check on any bias this procedure may have imposed, random tests were chosen and independent measurements of the length scales were made by a second observer. There was good agreement between the two estimates of these length scales.

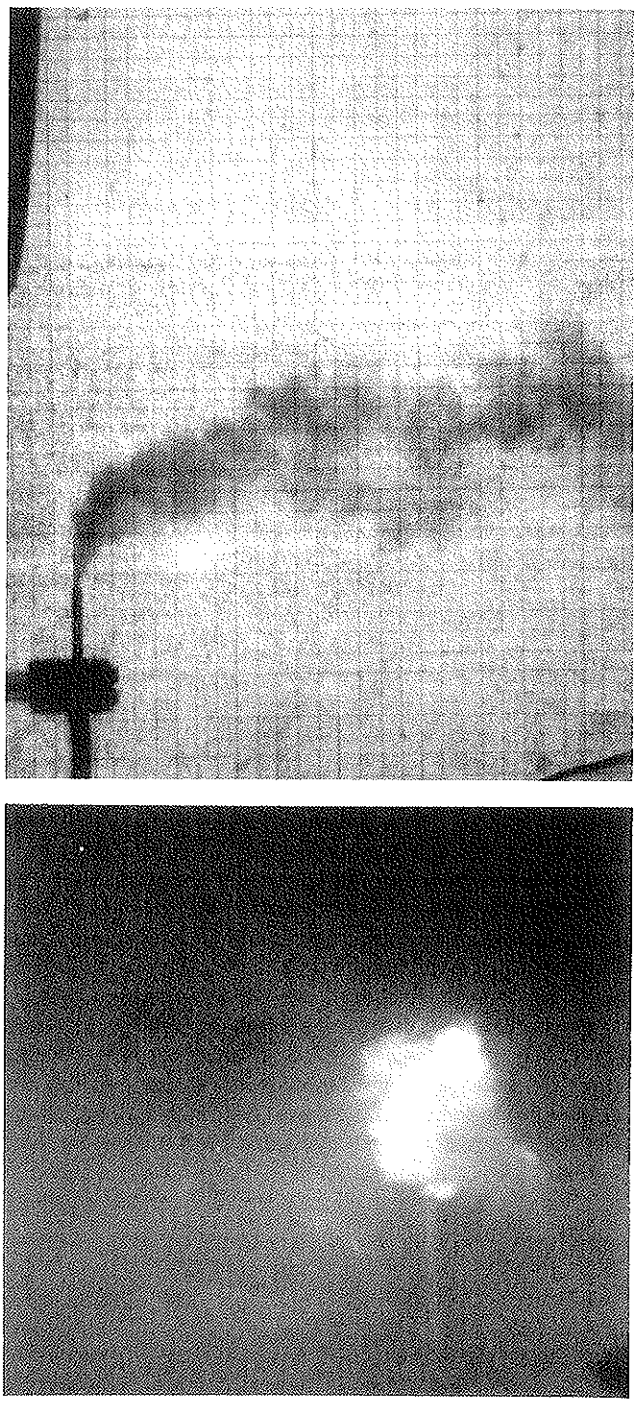


**FIGURE 4 Side (*top*) and end (*bottom*) views of a jet (or plume) for the case of strong cross flow; $F_j$ = 84.5, $F_0$ = 5.5, $X^*$ = 179.0, $Y^*$ = 67.0, $Z^*$ = 95.0, θ = 60 degrees.**

A summary of the experimental uncertainties in these measurements follows:

| *Parameter* | *Uncertainty (%)* |
|---|---|
| $F_j$ | 9.1 |
| $F_0$ | 2.4 |
| $L^*$ | 9.6 |
| $L^*/F_j$ | 13.3 |

However, the major source of scatter in the data is the inherent convoluted shape of the instantaneous boundary of the jet or plume. The instantaneous structure is captured in the photographs, not the average boundary shape, which would require averages over a number of observations.

## EXPERIMENTAL RESULTS AND OBSERVATIONS

### Observations

A complete photographic record of all tests has been obtained. Examples of these photographs are shown in Figures 2–4. These three photographic pairs are for the three ranges of $F_0$ applicable to negligible, weak, and strong cross flows, respectively.

The observed variability in all length scales increased at low $F_0$ and at the higher values of $\theta$. Under these conditions, the plume length scales were observed to fluctuate with time around a mean value. These fluctuations were of the order of $\pm 20$ percent of the mean length scales, while at the same time, the jet flow rates were observed to be constant. These low-frequency fluctuations are then a result of the interaction of the jet or plume with the column of fluid that is falling back around the upward-directed flow. At the lower angles and higher values of $F_0$, this interaction was not as pronounced or did not occur, and the variability in these length scales was much smaller. The highly convoluted entraining interface of these jet or plume flows is also quite apparent in these photographs, where the large-scale turbulent motion imposes a highly irregular boundary between the jet or plume fluid and the surrounding lower-density fluid. The difficulty of determining the mean interfacial length scales from a single photograph is apparent, and the scatter of the measurements is caused predominantly by this convoluted structure.

The double vortex structure mentioned earlier is shown in the photograph in Figure 5. This situation entailed a strong cross flow ($F_0 > 1$). The darker fluid on the inside of the double vortex was unmarked external fluid that was being entrained in a preferential way by this vortical structure. The photograph also shows the elliptical cross-sectional shape of the plume, whose minor axis is aligned in the same general direction as the initial jet angle. The entrainment processes observed are in sharp contrast to the turbulent interfacial entrainment processes of negligible or weak cross flow jets and plumes.

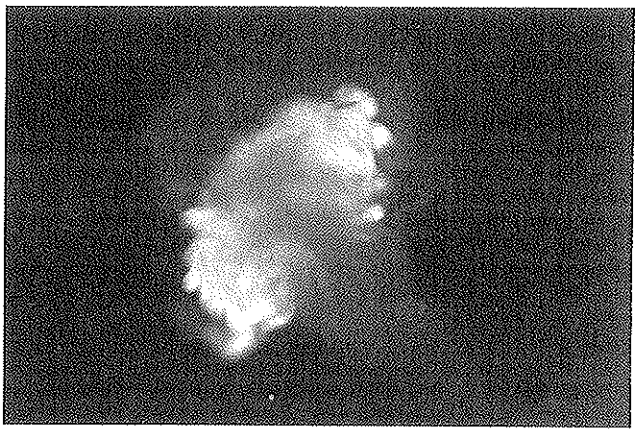


**FIGURE 5 Double vortex structure, as seen from an end view; $F_j = 51$, $F_0 = 4.0$, $\theta = 60$ degrees.**

### Quantitative Results

#### *Comparisons with Previous Work*

The special case $\theta = 90$ degrees has, as has been noted, some limited prior research history. In Turner's (*9*) work, for $F_0 = 0$, a principal result was Equation 6. Figure 6 shows the data for $\theta = 90$ degrees and $F_0 = 0$ for the present study, along with Equation 6. The best fit line for these data (on the basis of eight tests) is

$$Z^* = 3.7F_j \tag{11}$$

which is 11 percent lower than that predicted by Equation 5 or the results of Hoot et al. (*12*). The observed fluctuations of the plume for high $\theta$ and low $F_0$ provide some explanation for this discrepancy. The maximum plume height excursion was not recorded (unless by chance), and a small number of instantaneous photographs were obtained. This procedure resulted in effective plume heights that varied over the range of instantaneous plume heights. Given a sufficiently large set of observations, the data then reflected the average plume height, rather than the maximum observed height. For these observations, a consistently lower value by 10 percent would then be expected. For the case of finite $F_0$, some visual averaging of the photographs was performed, to remove some of the instantaneous large-scale excursions form the data. This procedure is in contrast to the time lapse photographs used by Hoot et al. (*12*). The effect of such a procedure would be to record the maximum excursion over the time duration of the photograph.

A direct comparison with the data of Hoot et al. (*12*) for the case of finite $F_0$ is also possible. This comparison is shown in Figure 7, where the present data are plotted along with Equation 7 over the range of the experimental data used to

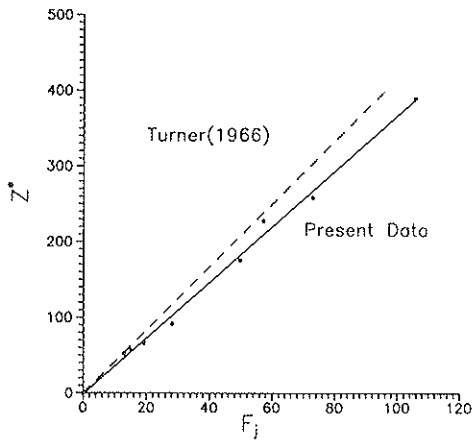



**FIGURE 6 Comparison of vertical jet (or plume) behavior for no cross flow, with the results of Turner (*9*).**

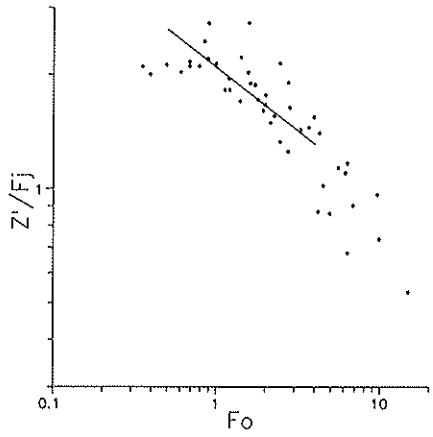



**FIGURE 7 Comparison of vertical jet (or plume) behavior in a cross flow with the results of Hoot et al. (*12*). Equation 7 is shown as a solid line, over the limits of the experimental data.**

obtain the correlation. The low-$F_0$ data of Hoot et al. were consistently below their reported asymptotic curve, in agreement with the present notion of a transition for $F_0 = 1$, as is discussed in the following sections.

*Data Correlations*

The measurements of the maximum vertical excursion of the plume $Z^*$, as a function of $F_j$, $F_0$, and $\theta$ will be discussed in detail. Many of the features of the jet or plume behavior are apparent from these observations alone. The other scales will then be presented and briefly discussed.

**Maximum Plume Rise, $Z^*$** There is an ambiguity in the interpretation of the definition of maximum plume rise when $F_0$ is varied continuously from 0 to values of the order of 15. For weak (or negligible) cross flows ($F_0 < 1$), the only observable is the maximum plume top, as the plume descends in the near vicinity of the initiating jet. This definition of plume rise was used in the previous subsection for the $F_0 = 0$ case study. For strong cross flows (i.e., $F_0 > 1$), the plume is advected in the downstream direction, so that a plume centerline may be considered. The example comparison study for finite $F_0$ in the previous section used such a definition. The two ways to define $Z_{max}$ are not consistent, but may be related as

$$Z' \equiv Z_{max,\mathbb{C}} = Z_{max,top} - W/2 \tag{12}$$

for $F_0 > 0$. For negligible cross flow, the exact interpretation of $Z_{max,\mathbb{C}}$ is not clear, because the internal dynamics of an ascending jet (or plume) and the subsequent annular descending flow are not easily reduced to a single vertical length scale. The choice of the definition of $Z_{max}$ for the data to be presented is

$$Z_{max} = Z_{max,top} \tag{13}$$

This choice was for reasons of consistency, so $Z_{max}$ is defined the same for $F_0 < 1$ and for $F_0 > 1$. Additionally, the experimental techniques used required separate determinations of $Z_{max,top}$ and $W_{max}$. The increased uncertainty of combining these measurements is avoided by this choice. (Such increased experimental uncertainty is in evidence in Figure 7.)

Logarithmic plots of $Z^*/F_j$ versus $F_0$ are shown in Figure 8(a–d). The dimensional arguments leading to Equation 2 are seen to be valid representations of the processes involved, as the data are reasonably well correlated. There is a definite change in the functional behavior of $Z^*/F_j$ in the neighborhood of $F_0 \sim 1$, which marks the transition between weak and strong cross flows.

For the $F_0 = 0$ cases, $Z^*/F_j = f_{z,0}(\theta)$. The values of this angular dependence for $F_0 = 0$ were determined from a regression analysis of $Z^*$ versus $F_j$, as was shown in Figure 6. The linear dependence of $Z^*$ on $F_j$ was consistent for all cases. The weak cross flow ($F_0 < 1$) behavior is essentially $Z^*/F_j \sim f_z(\theta)$, which is consistent with the scaling arguments leading to Equation 2. For $F_0 < 1$, $l_b < l_m$, so the transition to a pure plume would be expected before the effects of a cross flow became important. For $F_0 > 1$, the slope of the data in the form

$$Z^*/F_j = f_z(\theta)F_0^n \tag{14}$$

was also determined from a regression analysis. A summary of the results of these correlations follows:

| *Initial Jet Angle,* $\theta$ *(degrees)* | $f_{z,0}(\theta)$ | $f_z(\theta)$ | $n_z$ |
|---|---|---|---|
| 90 | 3.70 | 3.61 | −0.53 |
| 60 | 2.46 | 3.01 | −0.59 |
| 45 | 1.76 | 2.53 | −0.53 |
| 30 | 1.47 | 2.14 | −0.50 |

A comparison of the values of $f_{z,0}(\theta)$ and $F_z(\theta)$ indicates that as $\theta$ decreases from 90 degrees, the value of $f_{z,0}(\theta)$ decreases much faster than $f_z(\theta)$. A plausible explanation for this behavior is the observed change in the interaction of the descending fluid with the rising central column of fluid. Comparison photographs of this process indicate that the interference of the descending fluid on the jet (or plume) decreases for $F_0 > 0$, where the low-momentum descending fluid is advected by the cross flow. Without this sweeping away of this descending fluid, the rising and falling fluid interaction is stronger, both by increasing the mixing rate within the plume and by an increased momentum exchange between the two flows.

The asymptotic analysis of Lindberg and Petersen (*1*) yielded a scaling for $Z^*/F_j$ for the case of a jet with weak cross flow (i.e., $F_0 < 1$) and $\theta < 90$ degrees. The experimental values of $f_z(\theta)$ agree with the following equation to within 4 percent.

$$Z^*/F_j = f_z(\theta) = 2.3\,[(\sin\theta)(1 + \sin^{1/2}\theta)]^{1/2} \tag{15}$$

As would be expected, the values of $f_{z,0}(\theta)$ do not correlate well with this expression, in which it is noted that the entrainment hypothesis used in the analysis is only consistent with weak crossflow behavior.

**Transverse Distance, $Y^*$** $Y^*$ is defined as $Y_{max}/r_0$, where $Y_{max}$ is the measured maximum transverse distance of the jet

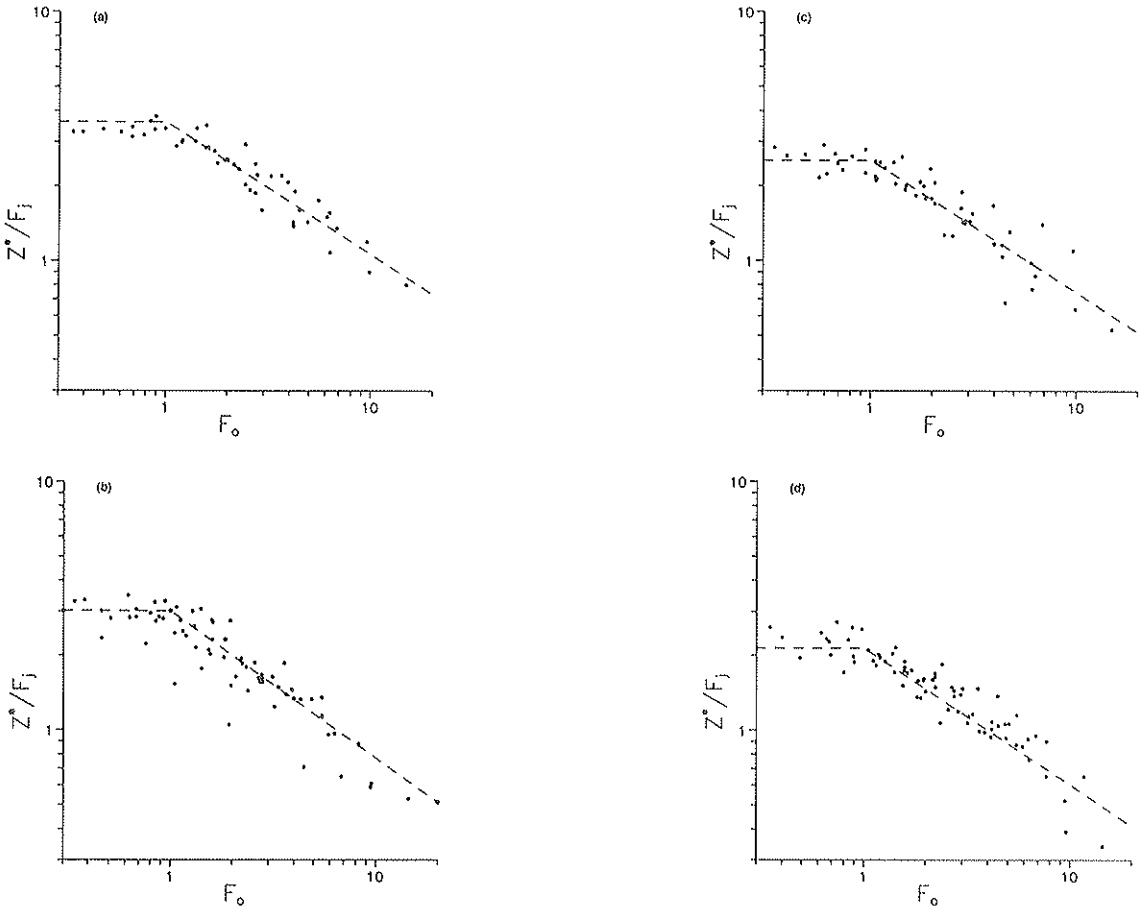



**FIGURE 8 Dimensionless maximum plume rise, $Z^*/F_j$, in a cross flow as a function of cross flow Froude number, $F_0$; (a) θ = 90 degrees, (b) θ = 60 degrees, (c) θ = 45 degrees and (d) θ = 30 degrees.**

(or plume) at the same location as $Z_{max}$. Figures 9(a–c) shows the measured $Y^*/F_j$ values as a function of $F_0$. The asymptotic behavior of $Y^*/F_j$ with $F_0$ is similar to the $Z^*/F_j$ data of Figure 8. Adequate representations of this data are thus:

$$Y^*/F_j = \begin{cases} f_{y,0}(\theta) & \text{for } F_0 = 0 \\ f_y(\theta) & \text{for } 0 < F_0 < 1 \\ f_y(\theta) F_0^n & \text{for } F_0 > 1 \end{cases} \quad (16)$$

A data correlation summary of the linear regression analysis results for $Y^*$ behavior (in which $Y^*$ is zero for θ = 90 degrees) follows:

| *Injection Angle* θ *(deg)* | $f_{y,0}(\theta)$ | $f_y(\theta)$ | $n_y$ |
|---|---|---|---|
| 90 | — | — | — |
| 60 | 1.7 | 1.7 | −0.51 |
| 45 | 1.6 | 2.2 | −0.59 |
| 30 | 3.0 | 2.7 | −0.62 |

With the exception of $f_{y,0}$(45degrees), both $f_{y,0}$ and $f_y$ increase with decreasing θ. The dependence of $Y^*/F_j$ on $F_0$ continues to have an approximate $F_0^{-1/2}$ behavior, although the calculated slopes are even more negative.

Analysis of end view photographs of plume behavior subsequent to the maximum plume rise location indicates that $Y$ does not appreciably increase further downstream. The initial $y$-momentum has become so diffuse through entrainment at this point that further transverse transport is small. For the present, then, $Y^*$ may be assumed to be indicative of the maximum transverse plume distance, at least for $F_0 > 1$.

**Longitudinal Distance, $X^*$** The downstream location where $Z_{max}$ occurs is defined as $X_{max}$. The scatter in the $X^*$ data is significant for $F_0 > 1$ and illustrates the difficulty of determining $X_{max}$ when the irregular plume boundary is essentially horizontal for a finite distance downstream. Within the limits of the data, an asymptotic equation of the form

$$X^*/F_j = f_x(\theta)\, F_0^n x \quad (17)$$

correlates the data for $F_0 > 0$. The following table summarizes the linear regression analysis for these data. Note that $X_{max} = F_0 = 0$.

| *Injection Angle* θ *(degrees)* | $f_x(\theta)$ | $n_x$ |
|---|---|---|
| 90 | 1.20 | 0.92 |
| 60 | 1.61 | 0.55 |
| 45 | 1.43 | 0.48 |
| 30 | 1.58 | 0.46 |

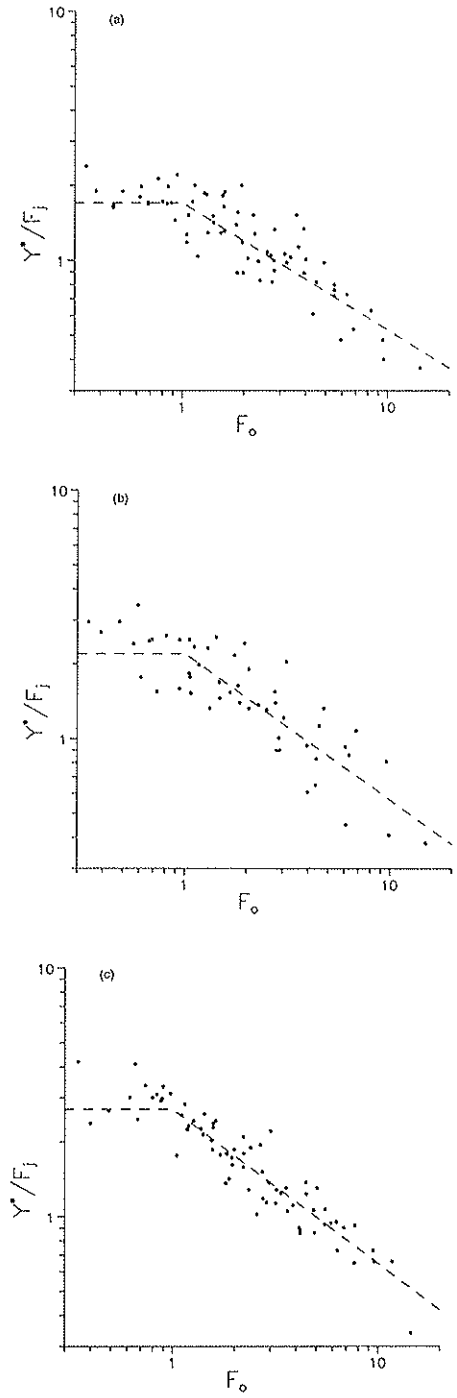



**FIGURE 9 Dimensionless transverse length scale, $Y^*/F_j$, in a cross flow as a function of cross flow Froude number, $F_0$; (a) θ = 60 degrees, (b) θ = 45 degrees, and (c) θ = 30 degrees.**

The variations of $f_x(\theta)$ are not significant; however, the comparison of $n_x$ for θ = 90 degrees and all other angles is to be noted. The asymptotic theory of Hoot et al. (*12*) and Lindberg and Petersen (*1*), for θ = 90 degrees, predicts that $X^*/F_j \sim F_0$ (Equation 8), in agreement with the measurements. For θ < 90 degrees, this exponent, $n_x$, decreases by as much as a factor of two. A more detailed analysis will be necessary to clarify this observation; however, it is clear that the nature of the downstream plume trajectory changes when the initial jet angle is not vertical.

**Plume Width Scale, $W^*$** Observations off the plume width scale, $W^*$, were made a the $Z_{max}$ location. These measurements were made from the side, by shadowgraphic images, so they represented the total vertical width of the plume at that location. Because of the elevation of the injection jet above the bottom of the tow tank, some measured widths exceed $Z_{max}$; however, these data were retained.

The following table summarizes the regression analysis for the four angles for $F > 1$, for the same asymptotic forms as Equation 12. The similar values of $f_w(\theta)$ and $n_w$ for the various injection angles indicate that the vertical plume widths, for all injection angles, are similar. If a conceptual elliptically shaped plume in the downwind direction is used, the lower injection angles have a smaller $Z_{max}$, but a larger projected plume width, because of the tilting of the ellipse.

| *Injection Angle* θ *(degrees)* | $f_w(\theta)$ | $n_w$ |
|---|---|---|
| 90 | 2.5 | −0.54 |
| 60 | 2.2 | −0.61 |
| 45 | 2.2 | −0.55 |
| 30 | 2.2 | −0.55 |

## APPLICATION TO SNOW PLUME BEHAVIOR

In terms of predicting snow plume cast distance or plume height for snowplow applications, Equation 12 can be rewritten as follows:

$$L_i^* = \eta f_i(\theta) F_0^{n_i+1} \tag{18}$$

for values of $F_0 < 1$, which is appropriate for most displacement plowing operations. Note that the length scales depend linearly on the moldboard efficiency, η.

Because nominal values of $n_i$ for $Y^*$ and $Z^*$ are of the order of −0.5, the exponent on $F_0$ in Equation 18 is of the order of +0.5, in sharp contrast to the squared exponent of Equation 5. The characteristic length scale for these laboratory studies has been the nozzle radius $r_0$. An equivalent circular length scale for the case of a snow plume yields the following scale:

$$r_{0,e} = (A_{exit}\pi)^{1/2} = (bH/\pi\eta\alpha)^{1/2} \tag{19}$$

where $b$ and $H$ are the width and depth of the plowed snow, respectively, and α is the ratio of the final to initial snow densities. This equivalent scaling is limited to plow flow exit areas, which are approximately circular (or elliptical).

Example predictions of cast distance, $Y_{max}$, and maximum cast height, $Z_{max}$, using Equation 18, are shown in Figures 10 and 11. The assumed parameters for these calculations are snow depth, $H$ = 0.20 m; plow width, $b$ = 3.3 m; snow density, $\rho_s$ = 100 kg/m$^3$; plow efficiency, η = 1.0; plow angle, φ = 0 degrees, and effective exit radius, $r_{0,e}$ = 0.46 m. Figure 10 also shows the prediction of cast distance using Equation 5 with the estimate of $\eta_d$ as suggested by Shalman (*5*) for the case of θ = 45 degrees.

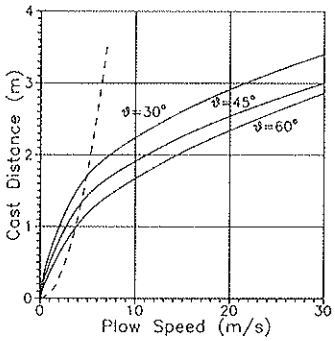



**FIGURE 10 Predicted cast distance as a function of plow speed and initial moldboard exit angle, θ. The example parameters are summarized in the text. The dashed line is Equation 5, for θ = 45 degrees.**

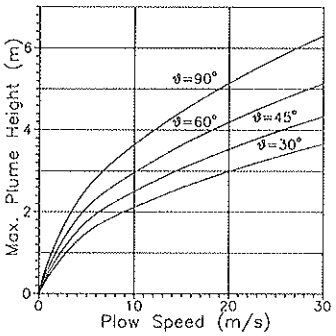



**FIGURE 11 Predicted snow plume case height for the same parameters as in Figure 10.**

The calculations are presented to illustrate the potential benefit of the current research in exit plume behavior. The current limitations on the parameters examined to date should be kept in mind.

## SUMMARY AND CONCLUSIONS

The results of an experimental research program on the behavior of negatively buoyant jets (or plumes) are summarized in this paper. The effects of initial jet angle and cross flow have also been investigated.

Plume trajectory behavior, in the form of coordinate length scales at the location of maximum plume height, was determined from photographs of the visualized plumes. These length scales are functions of jet angle, θ, and the two Froude numbers: $F_j$ and $F_0$. To the accuracy of the observations, the results have been correlated to an asymptotic power law as a function of flow regime in the form

$$\mathbf{L}^*/F_j = \begin{cases} f_{0,i}(\theta) & \text{for } F_0 = 0 \\ f_i(\theta) & \text{for } 0 < F_0 < 1 \\ f_i(\theta)F_0^{n} & \text{for } F_0 > 1 \ (F_0 > 0 \text{ for } X^*) \end{cases} \quad (20)$$

The flow regimes of $F_0 = 0$ (no crossflow), $0 < F_0 < 1$ (weak crossflow), and $F_0 > 1$ (strong cross flow) exhibit significantly different dynamical behavior, as is seen in the experimental observations and in the length scale results.

This study has successfully demonstrated the usefulness of using a water tow tank for the present jet (or plume) research. The range of parameters attainable in this facility has allowed for a clarification of the distinct flow regimes that have been discussed. Observations and photographic records have revealed much of the detail of these jet (or plume) flows from a variety of view angles.

At the outset of this phase of the study, the primary objectives were to

- Examine the effects of injection angle and cross flow on jet (or plume) behavior and relate the observations to previous work.
- Attain a wide range of parameter values to demonstrate the validity of the scaling analysis, and
- Determine the parameter space that may be identified with the various asymptotic flow regimes.

Within the limits of a photographic study, a large range of parameter space, and a relatively small number of tests, the objectives of the study have been met.

Further laboratory study will be directed toward the following objectives:

- Study the effect of changes in the exit jet geometry on jet (or plume) behavior, with emphasis on the strong crossflow regime;
- Include the azimuthal angle to study the effect of both fore and aft injection angles;
- Investigate the non-Boussinesq effect of $S_\rho$ significantly different than 1, by using different solutes and suspensions;
- Examine the interaction between adjacent objects and the downstream plume behavior; including the incorporation of flow control devices attached to the adjacent structures;
- Compare jet (or plume) behavior in a quiescent flow to a turbulent environment; and
- Extend the analytical work to include higher level numerical models, primarily integral-based models.

## ACKNOWLEDGMENTS

The research described herein was supported by the Strategic Highway Research Program (SHRP). SHRP is a unit of the National Research Council that was authorized by Section 128 of the Surface Transportation and Uniform Relocation Assistance Act of 1987. The authors would like to express their appreciation to Wayne Foslien, Jeff Rogers, and Roland Miller for their help with the data reduction and to R. N. Meroney of Colorado State University.

## REFERENCES

1. W. R. Lindberg and J. D. Petersen. *The Negatively Buoyant Jet/Plume in an Ambient Cross Flow*. Fluid Mechanics Technical Report FMTR-90-1, Department of Mechanical Engineering, University of Wyoming, 1990.

2. H. B. Fischer, E. J. List, R. C. Y. Koh, J. Imberger, and N. H. Brooks. *Mixing in Inland and Coastal Waters*, Chapter 9, Academic Press, New York, 1979.
3. L. D. Minsk. Snow Removal Equipment. In *Handbook of Snow*, D. M. Gray and D. H. Male, eds., Pergamon, New York, 1981, pp. 648–670.
4. G. A. Vinnicombe. *Comparative Tests on Model Vee Snow Ploughs*. Report LR 180, Road Research Laboratory, Crowthorne, Berkshire, U.K., 1968.
5. D. A. Shalman. *Snowplows, Construction, Theory and Design*, 2nd ed. (in Russian). Mashinostroenie, Leningrad, U.S.S.R., 1973.
6. E. J. Hopfinger and P. Beghin. Buoyant Clouds Appreciably Heavier Than the Ambient Fluid on Sloping Boundaries. *Proc., 2nd International IAHR Symposium on Stratified Flows*. Troudheim, Norway, 1980, pp. 495–504.
7. E. J. Hopfinger. Snow Avalanche Motion and Related Phenomena. *Annual Reviews of Fluid Mechanics*, Vol. 15, pp. 47–76.
8. F. T. Bodurtha. The Behavior of Dense Stack Gases. *Journal of Air Pollution Control Association*, Vol. 11, 1961, pp. 431–437.
9. J. S. Turner. Jets and Plumes with Negative or Reversing Buoyancy. *Journal of Fluid Mechanics*, Vol. 26, 1966, pp. 779–792.
10. F. M. Holly and J. L. Grace. Model Study of Dense Jets in Flowing Fluid. *Journal of Hydraulics Division, ASCE*, Vol. 98. No. 9365, 1972, pp. 1921–1933.
11. A. B. Pincince and E. J. List. Disposal of Brine into an Estuary. *Journal of the WPCF*, Vol. 45, No. 11, 1973, pp. 2335–2344.
12. T. G. Hoot, R. N. Meroney, and J. A. Peterka. *Wind Tunnel Tests of Negatively Buoyant Plumes*. Report CER73-74TGH-RNM-JAP-13, Fluid Dynamics and Diffusion Laboratory, Colorado State University, 1973.
13. T. G. Hoot and R. N. Meroney. The Behavior of Negatively Buoyant Stack Gases. *Proc., 67th Annual Meeting, APCA*, Denver, Colo., 1974.
14. J. L. Anderson, F. L. Parker, and B. A. Benedict. *Negatively Buoyant Jets in a Cross Flow*. Report 660/2-73-012, U.S. Environmental Protection Agency, 1973.
15. V. H. Chu. Turbulent Dense Plumes in Laminar Cross Flow. *Journal of Hydraulic Research*, Vol. 13, No. 3, 1975, pp. 263–279.
16. A. Badr. Temperature Measurements in a Negatively Buoyant Round Vertical Jet Issued in a Horizontal Cross Flow. In *Atmospheric Dispersion of Heavy Gases and Small Particles*, G. Ooms and H. Tennekes, eds., Springer-Verlag, New York, pp. 167–176.
17. R. N. Meroney. Wind Tunnel Experiments on Dense Gas Dispersion. *Journal of Hazardous Materials*, Vol. 6, Nos. 1–2, 1982, pp. 85–106.
18. G. Ooms and N. J. Duijm. Dispersion of a Stack Plume Heavier Than Air. In *Atmospheric Dispersion of Heavy Gases and Small Particles*. G. Ooms and H. Tenekes, eds., Springer-Verlag, New York, 1984, pp. 1–23.
19. D. Crabb, D. F. G. Duraõ, and J. H. Whitelaw. A Round Jet Normal to a Crossflow. *Journal of Fluids Engineering, Transactions of ASME*, Vol. 103, 1981, pp. 142–153.
20. G. N. Abramovich. *The Theory of Turbulent Jets*. Chapter 12, MIT Press, Cambridge, Mass., 1963.
21. J. F. Keffer and W. D. Baines. The Round Turbulent Jet in a Cross Wind. *Journal of Fluid Mechanics*, Vol. 15, 1963, pp. 481–496.
22. S. R. Hanna and P. J. Drivas. *Guidelines for Use of Vapor Cloud Dispersion Models*. American Institute of Chemical Engineers, 1987.
23. R. E. Britter and R. F. Griffiths. Dense Gas Dispersion. *Journal of Hazardous Materials*, Vol. 6, Parts 1 and 2, Elsevier, 1982.

---

*This paper presents the views of the authors only, not necessarily the views of the National Research Council, of SHRP, or of SHRP's sponsors. The results reported are not necessarily in agreement with the results of other SHRP research activities. They are reported to stimulate review and discussion within the research community.*

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Chemical Undercutting of Ice on Highway Pavement Materials

ROBERT R. BLACKBURN, KARIN M. BAUER, A. D. MCELROY, AND JEAN E. PELKEY

Experiments were conducted to investigate the destruction of the ice-substrate bonds by chemical undercutting. The undercutting experiments determined the undercut area, as a function of time, of three deicers at three temperatures on highway core samples and laboratory-produced specimens. Tests were conducted at temperatures of 25°F (−4°C), 15°F (−9°C), and 5°F (−15°C) using all the substrates in the as-received, but cleaned, condition. The three deicers—NaCl, $CaCl_2$, and ethylene glycol—were used for all temperature-substrate combinations. The substrates included portland cement concrete, dense-graded asphalt, open-graded asphalt, rubber-modified asphalt core samples, and laboratory-produced specimens made of portland cement concrete and dense-graded asphalt. Undercutting action was also investigated on smooth highway core samples and smooth laboratory-produced specimens at selected temperatures. Linear regression models with good predictive power were developed to predict the undercutting behavior of deicer material for given combinations of pavement type, surface condition, and temperature as a function of time. Using these regression models, the following conclusions were drawn: (a) at 25°F, sodium chloride produces larger undercut areas than does calcium chloride when applied to as-received core samples of dense-graded and open-graded asphalt; (b) however, at 25°F, calcium chloride produces more undercutting action than does sodium chloride when applied to as-received core samples of portland cement concrete and rubber-modified asphalt; (c) at 15°F and 5°F, calcium chloride produces more undercutting than does sodium chloride for all four as-received pavement core samples; and (d) at 25°F and 15°F, the as-received laboratory-produced specimens tended to be undercut more extensively than the corresponding core samples at these temperatures.

Deicing salts, primarily sodium chloride, have been applied to highways for control of snow and ice since early in this century. Before 1941, little straight salt was applied to the roadways; most was mixed with sand or other abrasives to freeze-proof stockpiles and to treat locally hazardous highway locations such as curves, hills, and intersections. Experiments were begun in New England in 1941 using NaCl alone as an ice preventive (*1*). The total amount of NaCl applied across the country remained relatively low until the mid-1950s when usage surged dramatically to around 1 million tons annually, partially in response to the public demand for better all-weather roadway conditions. This demand eventually led to the adoption of a bare pavement policy by the highway departments in the snow-belt states (*2*).

Currently, the economic well-being, livelihood, and strategic defense of the United States depend to a large extent on the year-round mobility of trucks, buses, and passenger cars on the nation's highway network. Winter conditions of ice and snow still cause serious disruptions in the economies of nearly all states. The great dependence on highway transportation for the movement of goods, services, and people has resulted in the demand for more rapid and effective clearance of ice and snow from the highways. As a result, there has been an increased use of chemicals and abrasives by highway agencies to assist in providing a clear roadway. The current usage of salt is approaching a rate of 10 million tons per year (*3*).

Sodium chloride has become the chemical of choice because it is effective at subfreezing temperatures (eutectic of −6°F), is substantially less expensive than other deicing materials, and is readily available. At temperatures near the freezing point of water, sodium chloride is an effective agent for the control of ice and light snow by processes that include melting, penetration through layers of ice or snow, and disbondment of ice or packed snow from pavement surfaces. However, concern has steadily been voiced over the last 12 years about the effects of heavy salt use on the roadside environment, water supplies, vehicles, and highway structures.

Removal of ice (and compacted snow) from highway surfaces has been accomplished in part by mechanical means (scraping) and in part by use of deicing chemicals or, in many locations, by a combination of these techniques. Neither the combination of these methods nor their singular use is completely satisfactory. Some chemical methods have potential side effects such as corrosion of vehicle and highway structures, and contamination of the roadside environment and water supplies, while mechanical methods do not provide complete ice removal.

Past work aimed at the development of methods for removal of ice from pavement surfaces has been rather narrow in scope (*4*). There is a small amount of basic data on the adhesive properties of ice, the mechanisms of adhesion, and the effects of different substrates on their adhesion. Proper application of existing technical information and the development of needed basic knowledge covering ice adhesion should provide a sound basis for the development of new, practical measures for ice removal. A much better understanding of the physical and chemical phenomenon observed at the ice-pavement interface is required before bond destruction can be achieved in an economical manner.

In late 1987, the Strategic Highway Research Program (SHRP) of the National Research Council funded a study entitled "Ice-Pavement Bond Disbonding—Fundamental Study." The overall objective of this research was to conduct a fundamental study of the ice-pavement bond structure and

Midwest Research Institute, 425 Volker Boulevard, Kansas City, Mo. 64110.

the mechanics of its formation, to provide a sound basis for the development of techniques and deicers for destroying or disrupting the ice-pavement bond once it has formed. Embedded within this overall objective were several specific objectives including one to characterize the physical and chemical processes that cause deterioration in the bond formed between ice and asphalt or portland cement concretes. Many of the study activities focused on investigations of techniques for the destruction of the bond between ice and highway pavement materials.

One such activity involved conducting laboratory tests to investigate the destruction of ice-substrate bonds by chemical undercutting. The highway pavement materials used during this activity consisted of laboratory-produced substrates and core samples taken from several in-service highway pavements. The results of these chemical undercutting tests are described.

## GENERAL DESCRIPTION OF UNDERCUTTING TESTS

The destruction of the ice-pavement bond by chemical undercutting is a physically complex process that is dependent on a number of variables, including the following: type of pavement material, pavement porosities and irregularities, heat transfer rates, brine concentration and associated density gradients, and chemical species diffusion rates. The experimental procedures used during the undercutting tests were designed to minimize and control as many of these variables as possible in order to generate reproducible data.

The main objective of the chemical undercutting experiments was to determine the influence of temperature and substrate condition on the time-dependent undercut area produced by three deicers at the ice-substrate interface. The three deicers used were ethylene glycol, aqueous sodium chloride (26.3 percent by weight saturated solution at 0°C), and aqueous calcium chloride (37.3 percent by weight saturated solution at 0°C). Six substrates were used in the tests and consisted of core samples taken from four types of highway pavements and two types of laboratory-produced specimens of highway pavement materials. Measurements of the undercutting action of the three deicers were made at three test temperatures of 25°F (−4°C), 15°F (−9°C), and 5°F (−15°C).

### Substrate Configurations and Undercutting Test Specimen Preparation

The substrates used in the undercutting tests included portland cement concrete, dense-graded asphalt, open-graded asphalt, rubber-modified asphalt core samples, and laboratory-produced specimens made of portland cement concrete and dense-graded asphalt. The portland cement concrete and dense-graded asphalt core samples were 4 in. (10.2 cm) in diameter and were obtained from the Connecticut Department of Transportation (DOT). The open-graded asphalt core samples were 6 in. (15.2 cm) in diameter and were obtained from the New York DOT. The rubber-modified asphalt (Plusride) core samples were 4 in. in diameter and were obtained from the Montana Department of Highways and from CALTRANS. All core samples were taken from in-service roads.

The laboratory-produced, dense-graded asphalt specimens were 4 in. in diameter and were supplied by Michigan Technological University through SHRP. These specimens were made in a Marshall test mold using AC20 asphalt and aggregate obtained from the SHRP asphalt-aggregate library in Austin, Texas.

The laboratory-produced portland cement concrete specimens had approximate dimensions of 6.5 × 8.5 × 1 in. deep (16.5 × 21.6 × 2.5 cm deep) and were made by researchers at the South Dakota School of Mines and Technology. These specimens were made in accordance with the mix design used by the South Dakota DOT for highway (nonbridge deck) pavement construction. The core samples were of varying lengths when received. Each core sample was cut to about a 2-in. (5.1-cm) length with a diamond saw in a plane parallel to the exposed, wearing surface. This process produced core samples with one end consisting of the as-received surface and the other end with a relatively smooth, not previously exposed surface. One end of the laboratory-produced dense-graded asphalt specimens was ground smooth, with a final grinding using 400-grit powder. The other end of the specimens was left in the as-compacted state. Three of the laboratory-produced portland cement concrete specimens were ground smooth, with a final grinding using 600-grit powder.

The handling, cleaning, and storage of all substrate surfaces in this study were accomplished in accordance with protocols developed under SHRP (*5*). These protocols were followed to maximize measurement reproducibility. The substrates were cleaned just before ice was grown on the surfaces to minimize any surface contamination that might have taken place during storage. A special diagnostic method was used to characterize the surface condition of the substrates after being subjected to the cleaning procedures. The diagnostic technique, known as the drop-diameter method, was developed under SHRP (*5*) and provided contact angle measurements that were based on the diameter of a drop of test liquid.

Three undercutting test sites were made on the 4-in.-diameter core and laboratory-produced samples. Nine undercutting test sites were made on each of the portland cement concrete slabs.

Ice of 3/16-in. (0.48-cm) thickness was formed on the substrates in a prescribed manner. Before growing the ice, the substrates were placed in a refrigerator maintained at 39°F (4°C) and allowed to equilibrate to that temperature for at least 16 hr. Deionized water was also placed in the same refrigerator and allowed to equilibrate to that temperature. The substrates were then transferred to an ice growing chamber inside a walk-in cold room that was maintained at −9°C. The substrates rested on a thick aluminum plate inside the ice growing chamber. The chilled (4°C) deionized water was then added to the substrate surface whose edges had been dammed with aluminum tape to a height of ¼ in. (0.64 cm) above the substrate surface. Two 100-watt lamps were positioned above the specimens in the chamber and their power was gradually reduced with a Variac®. This procedure allowed the ice to grow slowly and from the substrate upward. Tapered holes were produced in the ice during freezing by use of Teflon® plugs resting on the substrate surface. The configuration of a plug is shown in Figure 1a. The plugs were held in the correct relative position on the substrate with a ¼-in.-thick acrylic sheet that was prebored to accept the pre-

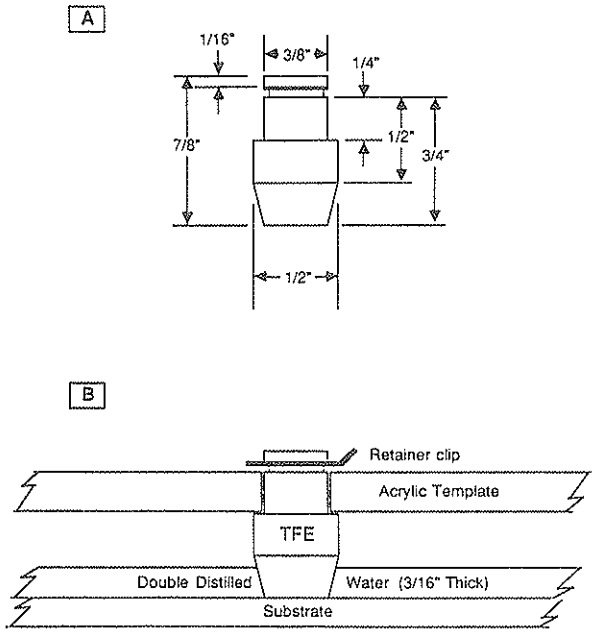



**FIGURE 1 Configuration of undercutting test specimen preparation: (a) teflon plug configuration, and (b) Sketch of experimental configuration.**

determined number of plugs. A sketch of this arrangement is shown in Figure 1b.

After the ice had frozen, the Teflon™ plugs were removed from the ice. The substrate-ice samples were then removed from the ice growing chamber and placed in plastic bags where they remained until the walk-in cold room and the samples equilibrated to the undercutting test temperature.

A small quantity (0.3 to 0.4 mg) of the disodium salt of fluorescein was placed in each cavity before placement of liquid deicer in the cavity. The weight of dye relative to deicer weight was approximately 1 part dye to 1,000 parts deicer. The dye accordingly contributed slightly (less than 0.1 percent) to total ice melting capacities. Photographs were then taken with the aid of a UV light source at eight predetermined time points over a 1-hr period during the undercutting action. The 35-mm negatives used in the photographic process exhibited good contrast between the undercutting areas and the substrate material. The undercutting patterns observed on the highly textured surfaces, such as on the open-graded asphalt cores and laboratory-produced portland cement concrete specimens, were irregular in shape and followed the surface voids around the exposed surface aggregate. The undercutting patterns on the smoother substrates tended to be circular in shape.

Undercutting tests with the three deicers were performed at each of the three temperatures on each of the six substrates in the as-received, but cleaned, condition. Undercutting tests with the three deicers were also performed on the smooth highway core samples and laboratory-produced specimens at 25°F and 5°F. Three replicate tests were performed for each unique combination of substrate, deicer, and temperature at a given time point.

### Data Reduction

The data reduction technique used with the undercutting results followed an approach used successfully in other undercutting experiments (*6*). The 35-mm color slides of the undercutting action were projected onto a screen. The individual undercut areas were then traced onto a vellum paper sheet. An area determination was made by use of a planimeter. A digital clock with a liquid crystal display and a standardized area grid were included in the field of view of the substrate to provide a photographic record of the time lapsed during the tests and a standard reference area for use in determining the undercut area.

All measured undercut areas were adjusted to account for the surface area, at time zero, that the Teflon plug was in contact with the substrate. This area was subtracted from all measured undercut areas obtained from the 35-mm photographic records. The resulting differences were then normalized by dividing them by the appropriate weight of the deicer. These final results are herein referred to as adjusted undercut areas. The replicate results were averaged before construction of the plots.

## RESULTS

Plots of the average adjusted undercut area versus time were developed for each of the three deicer and six substrate combinations. Of these 18 plots, a selection is shown in Figures 2–9. Each figure shows the results for as-received and smooth surfaces for the three temperatures.

The tests conducted with each deicer involved primarily an evaluation of the effects of different types of substrates and their physical surface characteristics (particularly surface roughness) on the undercutting characteristics. All substrate surfaces were cleaned before testing by standardized procedures. Thus, the chemical properties of the substrate surfaces in all cases, including the core samples, do not reflect the chemical contamination that accompanies highway uses. The undercutting results should, however, reflect the chemical properties of the substrate material (i.e., asphalt or portland cement concrete). Potential differences might exist between core samples and laboratory-produced substrates of the same type of material. The core samples were exposed to long-term weathering and use effects; the laboratory-produced specimens were not exposed to these environmental impacts.

### Estimating Undercut Areas as a Function of Time

Linear regression analyses were performed on the data to investigate the relationship between the adjusted undercut area (in square centimeter per gram of deicer in solution) and time (in minutes). These analyses were performed separately for each deicer, substrate material, sample type, surface condition, and temperature combination. Thus, a total of 90 analyses are reported. All three replicate results at each of nine time points were used in these analyses. In a first step, the adjusted undercut area was regressed against time$^{1/2}$, time, and time$^2$. The residuals from each model were examined for

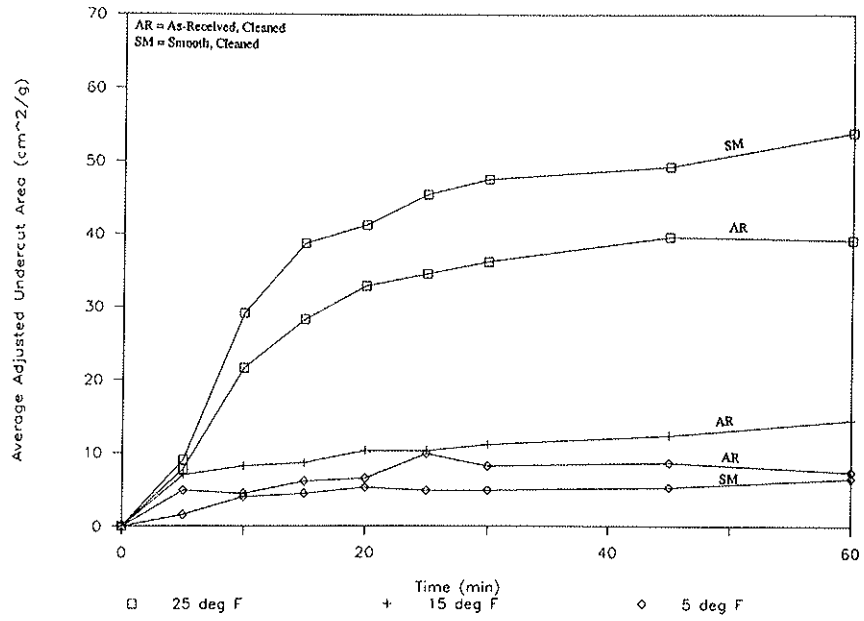



**FIGURE 2 Adjusted undercut area versus time: NaCl on core samples of dense-graded asphalt.**

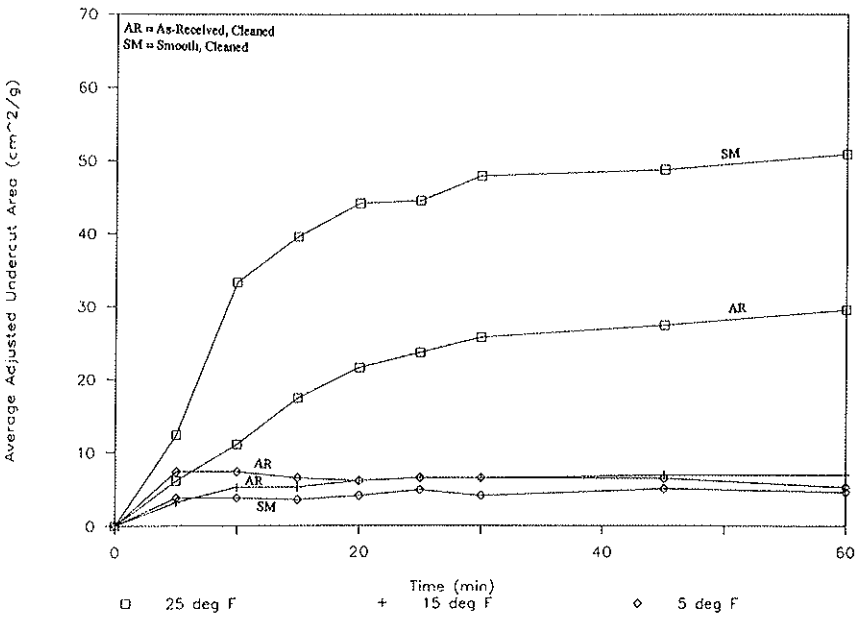



**FIGURE 3 Adjusted undercut area versus time: NaCl on core samples of open-graded asphalt.**

normality and, as a result, the square root of the adjusted undercut area was used. This approach considerably improved the normality of the model residuals.

The square root of the adjusted undercut area was therefore regressed against time$^{1/2}$, time, and time$^2$ for each of the 90 combinations of deicer, substrate type, sample type, and temperature. The regression model is given by the following:

$$(\text{adjusted undercut area})^{1/2} = A * \text{time}^{1/2} + B * \text{time} + C * \text{time}^2 \quad (1)$$

where $A$, $B$ and $C$ denote the regression coefficients. The model was forced through the origin, that is, the intercept was set to zero because at time zero, the undercut area was also zero.

The regression analyses were performed in a stepwise fashion. After considering all three variables, time$^{1/2}$, time, and time$^2$, if any of the coefficients $A$, $B$, or $C$ was not statistically significant at the 10 percent significance level, then the model was rerun without the corresponding variables.

Table 1 presents the final regression results, sorted by deicer, substrate material, sample type, surface condition, and

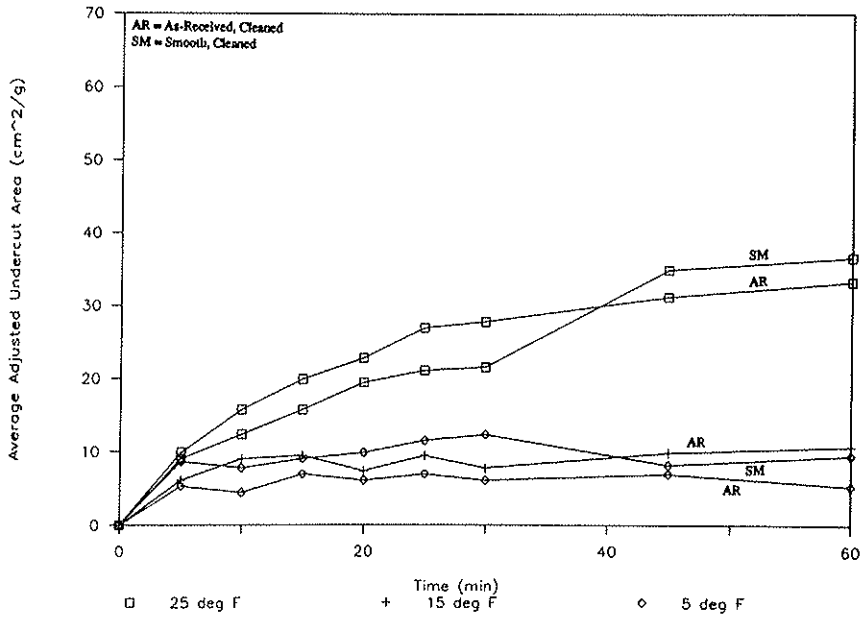



**FIGURE 4 Adjusted undercut area versus time: NaCl on core samples of rubber-modified asphalt.**

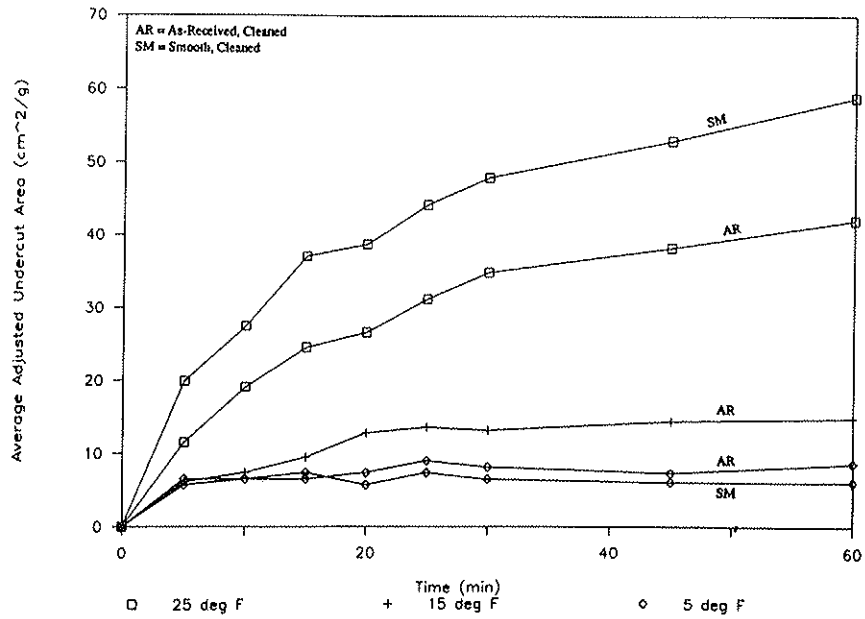



**FIGURE 5 Adjusted undercut area versus time: NaCl on core samples of portland cement concrete.**

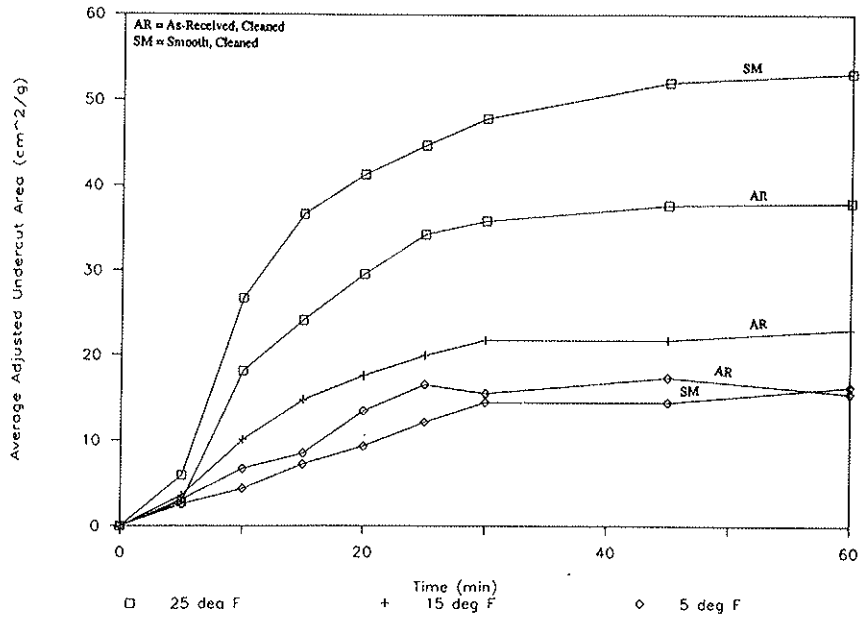



**FIGURE 6 Adjusted undercut area versus time: $CaCl_2$ on core samples of dense-graded asphalt.**

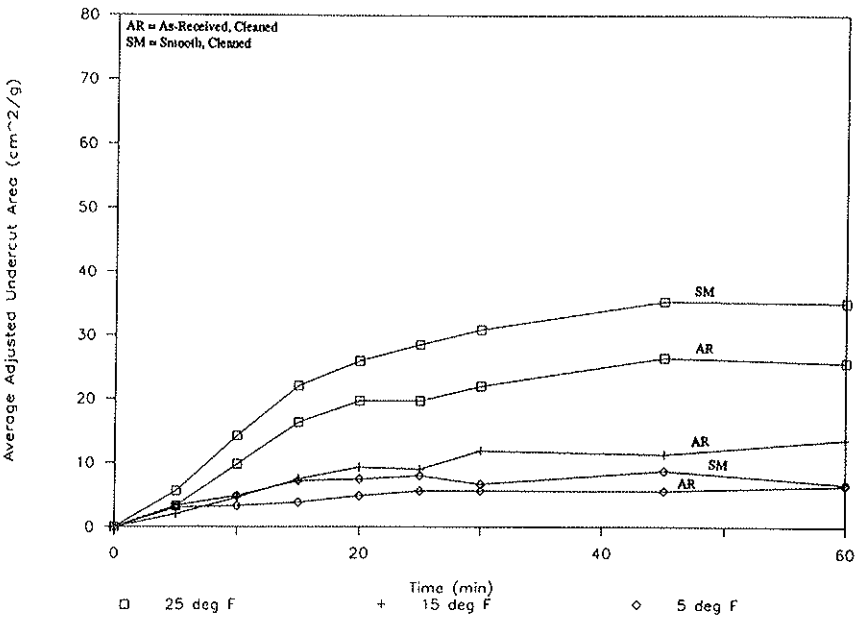



**FIGURE 7 Adjusted undercut area versus time: $CaCl_2$ on core samples of open-graded asphalt.**

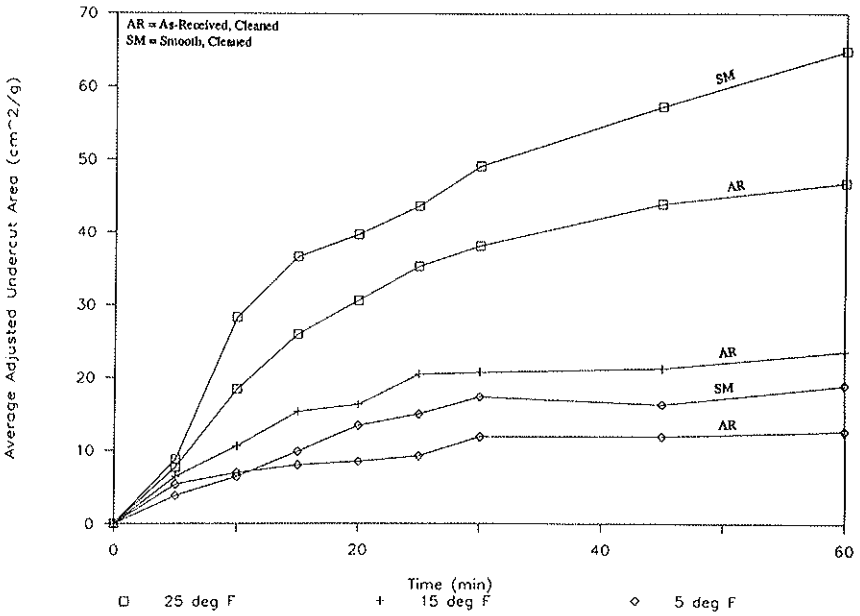



**FIGURE 8 Adjusted undercut area versus time: $CaCl_2$ on core samples of rubber-modified asphalt.**

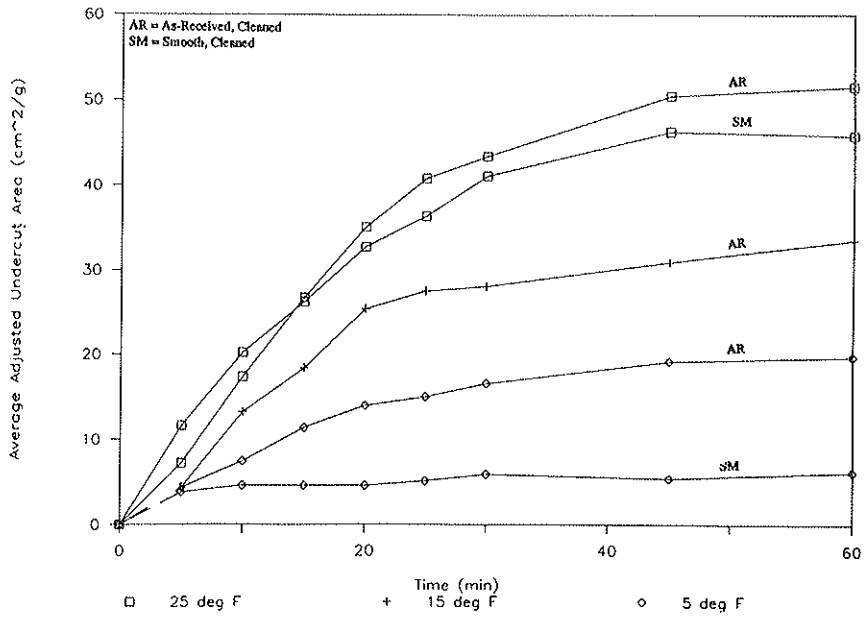



**FIGURE 9 Adjusted undercut area versus time: $CaCl_2$ on core samples of portland cement concrete.**

temperature. All models were statistically significant at the 10 percent level or better. For each model, Table 1 presents the number of test results on which each regression is based; the average values of the square root of the adjusted undercut area over time; the regression coefficients $A$, $B$, and $C$; and two measures of model fit. A missing entry for a regression coefficient indicates that the corresponding variable did not significantly contribute to the model at the 10 percent significance level. This procedure is equivalent to setting that coefficient to zero.

The last two columns are two measures of the fit of the model to the data. The root mean square error (RMSE), in the same units as the dependent variable, the square root of the adjusted undercut area, is an estimate of the standard deviation about the regression. It is a measure of the error with which any observed value of the square root of the adjusted undercut area could be predicted, at a given time interval after deicer application, using the estimated regression equation. Thus the smaller the RMSE, the smaller the error in the prediction.

TABLE 1 LINEAR REGRESSION RESULTS FOR UNDERCUTTING MODELS

## Deicer Material: NaCL

| Substrate Material | Sample Type | Surface Condition | Temperature (deg F) | Number of Tests | Mean Sqrt(Area) | Regression Coefficients Sqrt(Time) (A) | Time (B) | Time*2 (C) | Measures of Model Fit Root Mean Square Error | R-Squared (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 4.76 | 1.046 | | | 1.115 | 70.6 |
| | | | 15 | 27 | 2.83 | 1.546 | -0.211 | 0.0013 | 0.301 | 92.6 |
| | | | 5 | 27 | 2.32 | 0.982 | -0.082 | | 0.358 | 85.7 |
| | | Smooth and Cleaned | 25 | 27 | 5.45 | 2.064 | -0.146 | | 0.404 | 97.0 |
| | | | 5 | 27 | 1.87 | 0.863 | -0.093 | 0.0004 | 0.201 | 93.3 |
| | Lab | As Received, Cleaned | 25 | 27 | 6.06 | 2.349 | -0.173 | | 0.978 | 86.0 |
| | | | 15 | 27 | 3.67 | 1.405 | -0.101 | | 0.292 | 96.4 |
| | | | 5 | 24 | 2.39 | 1.520 | -0.228 | 0.0015 | 0.280 | 91.6 |
| | | Smooth and Cleaned | 25 | 27 | 5.34 | 1.180 | | | 1.604 | 61.8 |
| | | | 5 | 27 | 1.93 | 1.342 | -0.221 | 0.0015 | 0.351 | 79.3 |
| Open-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 3.91 | 0.868 | | | 0.739 | 81.2 |
| | | | 15 | 27 | 2.13 | 1.081 | -0.128 | 0.0005 | 0.164 | 96.0 |
| | | | 5 | 27 | 2.26 | 1.635 | -0.262 | 0.0015 | 0.315 | 86.5 |
| | | Smooth and Cleaned | 25 | 27 | 5.55 | 2.217 | -0.170 | | 0.334 | 97.8 |
| | | | 5 | 27 | 1.82 | 1.099 | -0.158 | 0.0009 | 0.233 | 88.6 |
| Rubber-Modified Asphalt | Core | As Received, Cleaned | 25 | 27 | 4.24 | 1.584 | -0.110 | | 0.326 | 96.5 |
| | | | 15 | 27 | 2.61 | 1.650 | -0.251 | 0.0016 | 0.259 | 93.2 |
| | | | 5 | 27 | 2.17 | 1.255 | -0.165 | 0.0007 | 0.281 | 88.5 |
| | | Smooth and Cleaned | 25 | 27 | 4.00 | 1.292 | -0.067 | | 0.473 | 92.9 |
| | | | 5 | 27 | 2.73 | 1.624 | -0.221 | 0.0010 | 0.448 | 82.6 |
| Portland Cement Concrete | Core | As Received, Cleaned | 25 | 27 | 4.65 | 1.690 | -0.112 | | 0.513 | 93.3 |
| | | | 15 | 27 | 2.95 | 1.160 | -0.088 | | 0.528 | 82.6 |
| | | | 5 | 27 | 2.43 | 1.483 | -0.215 | 0.0012 | 0.249 | 92.5 |
| | | Smooth and Cleaned | 25 | 27 | 5.50 | 2.038 | -0.141 | | 1.252 | 75.2 |
| | | | 5 | 27 | 2.25 | 1.496 | -0.226 | 0.0012 | 0.238 | 92.0 |
| | Lab | As Received, Cleaned | 25 | 27 | 5.07 | 1.122 | | | 1.120 | 74.9 |
| | | | 15 | 27 | 3.50 | 1.417 | -0.111 | | 0.324 | 94.7 |
| | | | 5 | 27 | 2.19 | 1.268 | -0.171 | 0.0008 | 0.230 | 92.1 |
| | | Smooth and Cleaned | 25 | 27 | 6.42 | 2.777 | -0.260 | 0.0006 | 0.289 | 98.7 |
| | | | 5 | 27 | 2.30 | 1.407 | -0.211 | 0.0013 | 0.232 | 92.9 |

## Deicer Material: CaCl2

| Substrate Material | Sample Type | Surface Condition | Temperature (deg F) | Number of Tests | Mean Sqrt(Area) | Regression Coefficients Sqrt(Time) (A) | Time (B) | Time*2 (C) | Measures of Model Fit Root Mean Square Error | R-Squared (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 4.47 | 1.264 | | -0.0010 | 0.646 | 91.1 |
| | | | 15 | 27 | 3.52 | 0.992 | | -0.0008 | 0.231 | 97.8 |
| | | | 5 | 27 | 3.00 | 0.851 | | -0.0007 | 0.234 | 97.0 |
| | | Smooth and Cleaned | 25 | 27 | 5.35 | 1.513 | | -0.0013 | 0.475 | 96.1 |
| | | | 5 | 27 | 2.72 | 0.726 | | -0.0004 | 0.292 | 94.9 |
| | Lab | As Received, Cleaned | 25 | 27 | 5.15 | 1.430 | | -0.0011 | 0.401 | 97.0 |
| | | | 15 | 27 | 4.26 | 0.936 | | | 0.954 | 72.6 |
| | | | 5 | 24 | 3.42 | 1.515 | -0.130 | | 0.238 | 97.2 |
| | | Smooth and Cleaned | 25 | 27 | 6.54 | 1.446 | | | 1.366 | 77.1 |
| | | | 5 | 27 | 1.32 | 0.506 | -0.037 | | 0.071 | 98.2 |
| Open-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 3.64 | 1.008 | | -0.0008 | 0.271 | 97.4 |
| | | | 15 | 27 | 2.52 | 0.564 | | | 0.549 | 78.8 |
| | | | 5 | 27 | 1.91 | 0.757 | -0.058 | | 0.350 | 80.8 |
| | | Smooth and Cleaned | 25 | 27 | 4.13 | 0.917 | | | 1.566 | 52.2 |
| | | | 5 | 27 | 2.25 | 0.973 | -0.083 | | 0.302 | 89.2 |
| Rubber-Modified Asphalt | Core | As Received, Cleaned | 25 | 27 | 4.80 | 1.070 | | | 0.874 | 83.4 |
| | | | 15 | 27 | 3.58 | 1.345 | -0.094 | | 0.255 | 97.1 |
| | | | 5 | 27 | 2.68 | 1.280 | -0.147 | 0.0007 | 0.279 | 93.2 |
| | | Smooth and Cleaned | 25 | 27 | 5.54 | 1.906 | -0.113 | | 0.471 | 96.4 |
| | | | 5 | 27 | 3.07 | 0.849 | | -0.0007 | 0.255 | 96.5 |
| Portland Cement Concrete | Core | As Received, Cleaned | 25 | 27 | 5.02 | 1.369 | | -0.0010 | 0.284 | 98.5 |
| | | | 15 | 27 | 4.07 | 0.911 | | | 0.920 | 77.6 |
| | | | 5 | 27 | 3.18 | 1.127 | -0.071 | | 0.184 | 98.2 |
| | | Smooth and Cleaned | 25 | 27 | 4.89 | 1.079 | | | 1.361 | 64.3 |
| | | | 5 | 27 | 1.98 | 1.174 | -0.168 | 0.0010 | 0.201 | 92.8 |
| | Lab | As Received, Cleaned | 25 | 27 | 4.76 | 1.343 | | -0.0011 | 0.326 | 97.7 |
| | | | 15 | 27 | 4.23 | 1.691 | -0.130 | | 0.304 | 96.9 |
| | | | 5 | 27 | 2.65 | 0.588 | | | 0.630 | 73.1 |
| | | Smooth and Cleaned | 25 | 27 | 5.77 | 1.622 | | -0.0014 | 0.575 | 95.1 |
| | | | 5 | 27 | 3.17 | 0.886 | | -0.0007 | 0.344 | 94.3 |

(*continued on next page*)

TABLE 1 LINEAR REGRESSION RESULTS FOR UNDERCUTTING MODELS (*continued*)

**Deicer Material: Ethylene Glycol**

| Substrate Material | Sample Type | Surface Condition | Temperature (deg F) | Number of Tests | Mean Sqrt(Area) | Regression Coefficients Sqrt(Time) (A) | Time (B) | Time*2 (C) | Measures of Model Fit Root Mean Square Error | R-Squared (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 2.96 | 0.252 | 0.129 | -0.0013 | 0.167 | 99.0 |
| | | | 15 | 27 | 2.51 | 0.573 | | | 0.232 | 96.3 |
| | | | 5 | 27 | 1.84 | 0.633 | -0.056 | 0.0005 | 0.184 | 95.9 |
| | | Smooth and Cleaned | 25 | 27 | 3.59 | 0.323 | 0.169 | -0.0020 | 0.253 | 98.3 |
| | | | 5 | 27 | 2.02 | 0.788 | -0.082 | 0.0006 | 0.165 | 96.9 |
| | Lab | As Received, Cleaned | 25 | 27 | 3.45 | 0.538 | 0.098 | -0.0013 | 0.187 | 98.9 |
| | | | 15 | 27 | 2.62 | 0.593 | | | 0.641 | 77.6 |
| | | | 5 | 27 | 2.04 | 0.520 | | | 0.242 | 94.1 |
| | | Smooth and Cleaned | 25 | 27 | 2.57 | 1.092 | -0.105 | 0.0004 | 0.193 | 96.6 |
| | | | 5 | 27 | 1.94 | 0.786 | -0.093 | 0.0008 | 0.209 | 95.0 |
| Open-Graded Asphalt | Core | As Received, Cleaned | 25 | 27 | 2.43 | 0.548 | | | 0.557 | 78.7 |
| | | | 15 | 27 | 2.24 | 0.558 | | -0.0002 | 0.140 | 98.4 |
| | | | 5 | 27 | 2.59 | 1.504 | -0.223 | 0.0015 | 0.316 | 90.3 |
| | | Smooth and Cleaned | 25 | 27 | 3.09 | 0.693 | | | 0.802 | 72.4 |
| | | | 5 | 27 | 1.92 | 0.772 | -0.075 | 0.0004 | 0.159 | 96.3 |
| Rubber-Modified Asphalt | Core | As Received, Cleaned | 25 | 27 | 3.00 | 0.683 | | | 0.402 | 92.4 |
| | | | 15 | 27 | 2.51 | 0.866 | -0.052 | | 0.126 | 98.6 |
| | | | 5 | 27 | 1.85 | 0.913 | -0.121 | 0.0008 | 0.273 | 88.6 |
| | | Smooth and Cleaned | 25 | 27 | 3.17 | 0.720 | | | 0.906 | 72.4 |
| | | | 5 | 27 | 1.97 | 0.943 | -0.117 | 0.0007 | 0.298 | 87.2 |
| Portland Cement Concrete | Core | As Received, Cleaned | 25 | 27 | 3.42 | 0.611 | 0.074 | -0.0010 | 0.164 | 99.1 |
| | | | 15 | 27 | 2.80 | 0.709 | | -0.0003 | 0.178 | 98.3 |
| | | | 5 | 27 | 1.87 | 0.675 | -0.067 | 0.0006 | 0.241 | 93.4 |
| | | Smooth and Cleaned | 25 | 27 | 3.75 | 0.971 | | -0.0005 | 0.265 | 97.9 |
| | | | 5 | 27 | 1.91 | 0.940 | -0.126 | 0.0009 | 0.391 | 79.4 |
| | Lab | As Received, Cleaned | 25 | 27 | 3.31 | 0.607 | 0.063 | -0.0009 | 0.192 | 98.7 |
| | | | 15 | 27 | 2.88 | 0.747 | | -0.0004 | 0.098 | 99.5 |
| | | | 5 | 27 | 1.91 | 0.610 | -0.042 | 0.0003 | 0.154 | 97.3 |
| | | Smooth and Cleaned | 25 | 27 | 3.99 | 0.779 | 0.072 | -0.0012 | 0.218 | 98.8 |
| | | | 5 | 27 | 1.62 | 0.845 | -0.103 | 0.0004 | 0.150 | 94.1 |

The last column in Table 1 presents the $R^2$ value for each regression model. This figure, in percent, is the proportion of the variation in the data that is explained by the model. That is, the higher the $R^2$ value (maximum is 100 percent), the better the model fits the data. Of the 90 regression models, 82 have an $R^2$ value above 75 percent, that is, they have good predictive power.

Using the results of these regression analyses, the undercutting action of a deicer on a pavement can be predicted under given conditions (i.e., temperature, pavement type, and surface condition) as a function of time. For example, Figure 10 shows the observed and predicted adjusted undercut areas for the three deicers on as-received, cleaned core samples of portland cement concrete at 25°F. The equations used for the three deicers are (from Table 1):

$$\text{NaCl: Area} = (1.690\ \text{Time}^{1/2} - 0.112\ \text{Time})^2 \quad (2)$$

$$\text{CaCl}_2\text{: Area} = (1.369\ \text{Time}^{1/2} - 0.0010\ \text{Time}^2)^2 \quad (3)$$

$$\text{Ethylene Glycol: Area} = (0.611\ \text{Time}^{1/2} + 0.074\ \text{Time} - 0.0010\ \text{Time}^2)^2 \quad (4)$$

The equations from Table 1 were used to compute estimated adjusted undercut areas for all comparisons made in the following sections.

### Effects of Substrate Surface Texture (As-Received Versus Smooth) on Undercutting Results

Comparisons were made of the undercutting results obtained for as-received and smooth substrate pairs of the same material. These comparisons were made for 25°F and 5°F results only because tests were not run on smooth surfaces at 15°F. Ratios of undercutting results for as-received to smooth substrate surfaces were formed at 20 and 60 min for a given substrate-deicer-temperature combination using the coefficients in Table 1. The results at 25°F and 5°F are presented in Table 2, for the six substrates and three deicers.

Table 2 indicates that the extent of undercutting at 25°F was considerably less on as-received substrates than on smooth substrates in a majority [75 percent of the cases (27 out of 36 ratios) are below 1], with ratios ranging from 0.4 to 2.4. At 5°F, no clear-cut conclusions can be drawn from comparisons of undercutting results for as-received and smooth substrate surfaces. For one-third of the results, the ratio is below 1. For the remainder of the cases, the ratios vary widely, ranging from 1.0 to 8.5. The inconsistencies of the ratios are caused, in large part, by the fact that the undercut areas are quite small at 5°F and are subject to more error in measurement than at 25°F. In addition, the small undercutting patterns at 5°F are approximately of the same dimensions as the surface irregularities (surface voids or asperities) of some of the as-received samples. The undercutting patterns at 5°F may thus

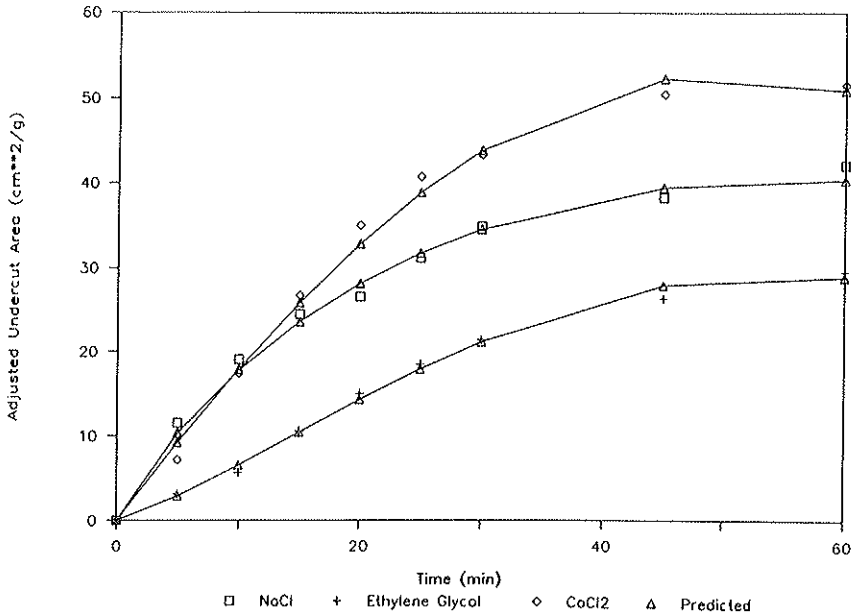



**FIGURE 10 Observed and predicted adjusted undercut areas for three deicers on cleaned, as-received portland cement concrete core samples at 25°F.**

TABLE 2 RATIOS OF UNDERCUTTING RESULTS FOR AS-RECEIVED TO SMOOTH SUBSTRATE SURFACES AT 20 AND 60 min AFTER APPLICATION

**a) Ratios of as-received to smooth results at 25 deg F**

| Substrate Material | Sample Type | Ratio at 20 min. NaCl | CaCl2 | Ethylene Glycol | Ratio at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 0.5 | 0.7 | 0.6 | 1.2 | 0.7 | 0.8 |
| Open-Graded Asphalt | | 0.4 | 1.0 | 0.6 | 0.9 | 0.5 | 0.6 |
| Rubber-Modified Asphalt | | 1.2 | 0.6 | 0.9 | 0.9 | 1.1 | 0.9 |
| Portland Cement Concrete | | 0.7 | 1.4 | 0.8 | 0.7 | 0.7 | 0.8 |
| Dense-Graded Asphalt | Lab | 1.8 | 0.8 | 1.7 | 0.7 | 0.4 | 2.4 |
| Portland Cement Concrete | | 0.5 | 0.7 | 0.7 | 1.1 | 0.7 | 0.8 |

**b) Ratios of as-received to smooth results at 5 deg F**

| Substrate Material | Sample Type | Ratio at 20 min. NaCl | CaCl2 | Ethylene Glycol | Ratio at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 1.6 | 1.3 | 0.8 | 1.2 | 1.0 | 0.9 |
| Open-Graded Asphalt | | 1.8 | 0.7 | 2.0 | 1.2 | 0.9 | 1.6 |
| Rubber-Modified Asphalt | | 0.6 | 0.8 | 0.8 | 0.7 | 0.7 | 0.9 |
| Portland Cement Concrete | | 1.1 | 2.6 | 1.0 | 1.5 | 3.3 | 1.2 |
| Dense-Graded Asphalt | Lab | 1.6 | 8.5 | 1.3 | 1.8 | 5.0 | 1.3 |
| Portland Cement Concrete | | 1.0 | 0.5 | 1.0 | 0.7 | 1.2 | 2.8 |

Note: Undercut areas were estimated using the regression models from Table 1.

reflect behavior over one or two surface irregularities rather than behavior over an average population of irregularities. At 25°F, however, the larger undercutting patterns should more closely reflect the average surface condition.

### Effects of Substrate Material on Undercutting Results

The determination of the effects of substrate material on undercutting characteristics of the three deicers is best made by examining the undercutting results at fixed times after application. These results, for both pavement core samples and laboratory-produced specimens, are presented in Table 3 (Parts a through e) for each of the substrate, deicer, and temperature combinations at 20 and 60 min after deicer application. These results were obtained using the regression coefficients presented in Table 1. From the estimated undercut areas presented in Table 3, the following observations can be made:

- Of the three deicers investigated, ethylene glycol produced the smallest estimated undercut area on both laboratory-produced specimens and pavement core samples at 20 and 60 min after application. This result is true both for smooth and as-received substrates at 25°F.

TABLE 3 ESTIMATED UNDERCUT AREAS (cm² PER 1 g OF DEICER IN SOLUTION) AT 20 AND 60 min AFTER APPLICATION

### a) Smooth and Cleaned Samples at 25 deg F

| Substrate Material | Sample Type | Undercut Area at 20 min. NaCl | CaCl2 | Ethylene Glycol | Undercut Area at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 40 | 39 | 16 | 53 | 50 | 31 |
| Open-Graded Asphalt | | 42 | 17 | 10 | 49 | 50 | 29 |
| Rubber-Modified Asphalt | | 20 | 39 | 10 | 36 | 64 | 31 |
| Portland Cement Concrete | | 40 | 23 | 17 | 54 | 70 | 34 |
| Dense-Graded Asphalt | Lab | 28 | 42 | 9 | 84 | 125 | 12 |
| Portland Cement Concrete | | 56 | 45 | 20 | 67 | 59 | 37 |

### b) Smooth and Cleaned Samples at 5 deg F

| Substrate Material | Sample Type | Undercut Area at 20 min. NaCl | CaCl2 | Ethylene Glycol | Undercut Area at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 5 | 9 | 5 | 6 | 16 | 12 |
| Open-Graded Asphalt | | 4 | 7 | 4 | 5 | 7 | 9 |
| Rubber-Modified Asphalt | | 11 | 12 | 5 | 9 | 18 | 9 |
| Portland Cement Concrete | | 7 | 5 | 4 | 6 | 6 | 9 |
| Dense-Graded Asphalt | Lab | 5 | 2 | 4 | 6 | 3 | 12 |
| Portland Cement Concrete | | 7 | 13 | 4 | 9 | 18 | 4 |

### c) As-received, Cleaned Samples at 25 deg F

| Substrate Material | Sample Type | Undercut Area at 20 min. NaCl | CaCl2 | Ethylene Glycol | Undercut Area at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 22 | 27 | 10 | 66 | 37 | 25 |
| Open-Graded Asphalt | | 15 | 18 | 6 | 45 | 25 | 18 |
| Rubber-Modified Asphalt | | 24 | 23 | 9 | 32 | 69 | 28 |
| Portland Cement Concrete | | 28 | 33 | 14 | 40 | 51 | 29 |
| Dense-Graded Asphalt | Lab | 50 | 35 | 15 | 61 | 50 | 29 |
| Portland Cement Concrete | | 25 | 31 | 13 | 75 | 40 | 29 |

### d) As-received, Cleaned Samples at 15 deg F

| Substrate Material | Sample Type | Undercut Area at 20 min. NaCl | CaCl2 | Ethylene Glycol | Undercut Area at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 10 | 17 | 7 | 15 | 22 | 20 |
| Open-Graded Asphalt | | 6 | 6 | 6 | 7 | 19 | 14 |
| Rubber-Modified Asphalt | | 9 | 17 | 8 | 11 | 23 | 13 |
| Portland Cement Concrete | | 12 | 17 | 9 | 14 | 50 | 20 |
| Dense-Graded Asphalt | Lab | 18 | 18 | 7 | 23 | 53 | 21 |
| Portland Cement Concrete | | 17 | 25 | 10 | 18 | 28 | 20 |

### e) As-received, Cleaned Samples at 5 deg F

| Substrate Material | Sample Type | Undercut Area at 20 min. NaCl | CaCl2 | Ethylene Glycol | Undercut Area at 60 min. NaCl | CaCl2 | Ethylene Glycol |
|---|---|---|---|---|---|---|---|
| Dense-Graded Asphalt | Core | 8 | 12 | 4 | 7 | 16 | 11 |
| Open-Graded Asphalt | | 7 | 5 | 8 | 6 | 6 | 14 |
| Rubber-Modified Asphalt | | 7 | 9 | 4 | 6 | 13 | 8 |
| Portland Cement Concrete | | 8 | 13 | 4 | 9 | 20 | 11 |
| Dense-Graded Asphalt | Lab | 8 | 17 | 5 | 11 | 15 | 16 |
| Portland Cement Concrete | | 7 | 7 | 4 | 6 | 21 | 11 |

Note: Undercut areas were estimated using the regression models from Table 1.

- However, as the substrate temperature decreases, the previous conclusion holds true in most cases at 20 min but not at 60 min after application.

Comparisons involving the smooth substrates (Table 3, Parts a and b) should reflect the differences caused by different chemical properties of the substrate materials (i.e., dense-graded asphalt and portland cement concrete). These comparisons reflect the following:

- No consistent relationship exists between the undercutting results obtained on dense-graded asphalt and those ob-

tained on portland cement concrete for either laboratory-produced specimens or pavement core samples.

- The smooth dense-graded, open-graded, and rubber-modified asphalt core samples cannot be consistently ranked on the undercutting action produced by either of the three chemical deicers.

The results indicate that a complex interaction exists between chemical deicers and the chemical properties of the smooth substrate surfaces.

At 25°F and 15°F, the two laboratory-produced specimens in the as-received condition were generally undercut more than the corresponding core samples at those temperatures (Table 3, Parts c and d). This pattern holds true at 20 min after deicer application; however, at 60 min, this pattern becomes more inconsistent.

Comparisons of the undercutting results for the four as-received core samples should be indicative of undercutting to be expected under field conditions without contamination. The results in Table 3, Parts c, d, and e, indicate the following:

- At 25°F, when applied on dense-graded and open-graded asphalt, sodium chloride produces larger undercut areas than does calcium chloride.
- However, when applied on portland cement concrete and rubber-modified asphalt at 25°F, calcium chloride produces more undercutting action than does sodium chloride.
- At 15°F and 5°F, calcium chloride produces more undercutting than does sodium chloride for all four as-received pavement core samples.

## COMPARISON WITH LITERATURE UNDERCUTTING RESULTS

The principal comparable study of undercutting reported in the literature is that of Trost et al. (*7*). The experimental approach taken by these authors is similar to that described herein. That is, the ice penetration action of the chemical deicer was not part of the chemical undercutting process as it was in another undercutting study (*8*). In the study by Trost et al., small, solid deicer particles (0.5 mm unit in dimension) were placed at the bottom of a cylindrical cavity formed in a larger cylinder of ice. The ice cylinder was prepared separately and then frozen onto various substrates via a thin film of water which was frozen to bond the ice cylinder to the substrates. Substrates tested were glass, brick, portland cement concrete, and asphalt. Solid forms of sodium chloride, calcium chloride, sodium hydroxide, urea, and calcium-magnesium-acetate (CMA) were used in the experiments. The quantities of deicer used ranged from 50 to 500 mg and the test temperatures were −5°C (23°F) and −10°C (14°F). These temperatures are close to two of the three temperatures used in the present study.

Trost et al. observed no significant differences in the extent or rate of undercutting with the different substrates, with the exception that initiation of undercutting was slower on glass.

The undercutting results obtained in the present study with aqueous solutions of sodium chloride and calcium chloride were compared with those obtained by Trost et al. for the solid form of these deicers. Before the comparisons could be made, it was necessary to develop a basis for comparing the aqueous deicers with solid forms of the same deicer. In the present study, adjusted undercut areas are reported in square centimeters undercut by 1 g of deicer in solution. The melting capacity of 1 g of deicer in solution is, of course, less than that of 1 g of an essentially dry form of the deicer.

A substantial portion of the rationale proposed and developed by Trost et al. is that the maximum extent of undercutting is a direct function of ice melting capacity. Consequently, it was decided to normalize the rest results on the basis of ice melting capacities and to compare the resultant data. A comparison of the results obtained for sodium chloride and calcium chloride with those of Trost et al. is presented in Table 4, where the undercutting data for 60 min are normalized on the basis of ice melting capacities. The deicer weights tabulated for the present study were determined by adjusting the aqueous deicer weights to equivalent weights of the solid form of the deicers.

The comparisons presented in Table 4 indicate that the Trost et al. study and the study reported herein yield similar undercutting results when like deicer weights are considered. The present study indicates, however, that the physical characteristics of the substrates have a substantial impact on undercutting.

## CONCLUSIONS

Linear regression models were developed to predict the undercutting behavior of each deicer for given combinations of pavement type, surface condition, and temperature as a function of time. These models have, overall, good predicting power, with $R^2$ values ranging from 52 to 99 percent. Of the 90 estimated regression models, 82 have an $R^2$ value greater than 75 percent. On the basis of these models, the examination of the undercutting data generated with pavement core

TABLE 4 COMPARISON OF UNDERCUTTING RESULTS FOR TWO DEICERS FROM TWO LITERATURE SOURCES

**a) Deicer Material: Sodium Chloride**

| Source of Data | Temperature (deg F) | Deicer, Solid Weight (mg) | Undercut Area (cm2 per gram of Ice Melting Capacity) |
|---|---|---|---|
| Present Results | 25 | 65 | 5.2[a]/4.7[b] |
| Trost et al. | 23 | 50 | 5.9 |
| Present Results | 15 | 45 | 5.8[a] |
| Trost et al. | 14 | 50 | 6.2 |

**b) Deicer Material: Calcium Chloride**

| Source of Data | Temperature (deg F) | Deicer, Solid Weight (mg) | Undercut Area (cm2 per gram of Ice Melting Capacity) |
|---|---|---|---|
| Present Results | 25 | 110.5 | 6.6[a]/4.3[b] |
| Trost et al. | 23 | 100 | 3.6 |
| Present Results | 15 | 95 | 6.5[b] |
| Trost et al. | 14 | 100 | 5.9 |

[a] Average value for smooth laboratory substrates of dense-graded asphalt and portland cement concrete.

[b] Average value for as-received laboratory substrates of dense-graded asphalt and portland cement concrete.

samples and laboratory-produced substrates and three deicers revealed some practical considerations.

A consistent temperature pattern was found during chemical undercutting. For a given deicer and substrate, the undercut area at a fixed time following application of the deicer decreased with decreasing temperature.

At 25°F, sodium chloride is the deicer of choice on in-service dense-graded and open-graded asphalt pavements. However, calcium chloride is the deicer of choice on in-service rubber-modified asphalt and portland cement concrete pavements. At 15°F and 5°F, calcium chloride is the deicer of choice on all in-service pavement types considered in this study. These results would suggest that the rate of chemical application should be varied by pavement type and temperature to achieve a desired undercutting area at a selected time interval after application of a chemical.

The laboratory-produced specimens of portland cement concrete and dense-graded asphalt in the as-received and cleaned condition were undercut more extensively at both 25°F and 15°F than the four pavement core samples in the as-received and cleaned condition. These results would indicate that care should be used when estimating undercutting characteristics on in-service highways from undercutting data obtained from tests performed with laboratory-produced specimens.

In a majority of the cases investigated, the extent of chemical undercutting by NaCl, $CaCl_2$, and ethylene glycol at 25°F on as-received (textured) highway core samples and laboratory-produced substrates was considerably less than that on smooth substrate surfaces of the same materials.

The undercutting results achieved for sodium chloride and calcium chloride generally agreed with similar data reported in the literature when like deicer weights were considered and when the undercutting results were normalized on the basis of ice melting capacities.

## ACKNOWLEDGMENTS

The work reported herein was conducted under SHRP Contract No. H-203. L. David Minsk was the SHRP research manager for the contract.

## REFERENCES

1. L. D. Minsk. A Short History of Man's Attempts To Move Through Snow. In *Special Report 115: Snow Removal and Ice Control Research*, HRB, National Research Council, Washington, D.C., 1970.
2. B. H. Welch, et al. *Economic Impact of Highway Snow and Ice Control—State of the Art Interim Report*. Report FHWA-RD-77-20, FHWA, U.S. Department of Transportation, Sept. 1976.
3. *Research Plans*. Strategic Highway Research Program, National Research Council, Washington, D.C., May 1986.
4. R. R. Blackburn, A. D. St. John, and P. J. Heenan. *Physical Alternatives to Chemicals for Highway Deicing*. Final Report, U.S. Department of Transportation, Dec. 1978.
5. R. R. Blackburn, et al. *Ice-Pavement Bond Disbonding—Fundamental Study*. Report SHRP-H/FR-90-XXX. Strategic Highway Research Program, National Research Council, Washington, D.C., Aug. 1990.
6. A. D. McElroy, R. R. Blackburn, and H. Kirchner. Comparative Study of Chemical Deicers—Undercutting and Disbondment. Paper presented at TRB 69th Annual Meeting, Washington, D.C., Jan. 1990.
7. S. E. Trost, F. J. Heng, and E. L. Cussler. Chemistry of Deicing Roads: Breaking the Bond Between Ice and Road. *Journal of Transportation Engineering*, Vol. 113, No. 1, Jan. 1987.
8. G. C. Sinke and E. H. Mossner. Laboratory Comparison of Calcium Chloride and Rock Salt as Ice Removal Agents. In *Transportation Research Record 598*, TRB, National Research Council, Washington, D.C., 1976, pp. 54–57.

*Publication of this paper sponsored by Committee on Winter Maintenance.*

# Misunderstood Applications of Urban Work Zone Traffic Control

MICHAEL A. OGDEN AND JOHN M. MOUNCE

Traffic control manuals do not sufficiently address many of the problems associated with urban arterial work zones. The traffic control devices applied within these work zones are sometimes misunderstood by the motorist. In order for there to be safe operations within urban roadway work zones, it is important that the traffic control be applicable to this specific environment. A motorist survey was initiated on Abrams Road in Dallas, Texas, to investigate comprehension of construction traffic control devices. Abrams Road is a four-lane undivided major urban arterial roadway. The survey was designed to meet the following objectives: (a) to ascertain knowledge about work zone traffic control, (b) to determine problematic construction traffic control devices with respect to motorist comprehension, (c) to elicit information from motorists concerning overall problems with the Abrams Road project, and (d) to substantiate or negate findings from a similar study in Houston, Texas. Personal interviews were conducted with 345 respondents in the Abrams Road area. These participants were asked to respond to questions regarding work zone signing and other forms of traffic control devices. The response percentages substantiated most of the findings from the earlier study revealing that motorists have some difficulty comprehending selected work zone traffic control applications.

Arterial street systems are being forced to sustain a significant portion of the traffic burden caused by increased congestion on freeways. The Texas State Department of Highways and Public Transportation (SDHPT) implemented a $100 million program [Principal Arterial Street System (PASS)] in September 1987 to upgrade urban arterials. This program was intended to provide additional capacity and improve traffic flow. The initiation and implementation of the PASS program has led to the recognition of shortcomings in the construction traffic management of urban arterials.

Construction traffic control on arterials in highly developed urban areas faces many problems that are not currently addressed in the *Manual on Uniform Traffic Devices for Streets and Highways* (*1*). These problems include increased driver workload associated with limited right-of-way, variable speeds and volumes, excessive turning movements, extensive driveway access points, and construction signing requirements.

There is a lack of documented research relating to urban arterial construction traffic control. In addition to this, there is a large discrepancy between applicable signing for freeway and highway work zones and that recommended for urban arterial work zones (*2,3*). Motorists appear to be confused by selected traffic controls that are applied within urban arterial construction (*4*). The Texas Transportation Institute (TTI) has documented motorist general confusion in understanding roadway construction signing (*5,6*).

Texas Transportation Institute, The Texas A&M University System, College Station, Tex. 77843.

Several problematic work zone traffic control applications were identified in an earlier survey by TTI (*7*) that was performed in Houston, Texas. The results indicated that motorists had problems understanding certain word and message symbols applied in roadway construction work zones. The objective of this study was to substantiate or negate similar findings from the previous study in Houston (*7*).

## INTRODUCTION

Abrams Road is a four-lane undivided major arterial located on the north side of Dallas, Texas (Figure 1). Reconstruction began in July 1989. This 2-mi segment of roadway extends from Forest Lane to Kingsley Boulevard. The reconstructed facility will feature a six-lane divided cross section. The traffic volume in January 1990 was approximately 20,000 vehicles per day (vpd) and included four signalized intersections. The primary land use along the arterial is residential with some retail businesses.

The traffic control plan initiated on the project, for the most part, exceeded the requirements of the *Texas Manual on Uniform Traffic Control Devices* (*8*). This survey was administered in May 1990 to investigate motorists' interpretations of construction traffic control devices and their perception of the urban arterial work zone.

## STUDY DESIGN AND METHODOLOGY

The Abrams Road survey was designed to meet the following objectives:

1. Ascertain knowledge about work zone traffic control,
2. Determine motorist comprehension of applied construction traffic control devices,
3. Elicit information from motorists concerning overall problems with the Abrams Road project, and
4. Confirm findings from a similar study conducted previously in Houston, Texas.

Surveys were conducted with 345 respondents in May 1990 at three locations. Respondents at all locations were approached by the surveyors and asked if they would like to participate voluntarily in the survey. A daily demographic total was kept to address any biases that might develop. The result was that 147 respondents were interviewed at a Texas Department of Public Safety licensing office and a total of 198 respondents were interviewed at two commercial locations.

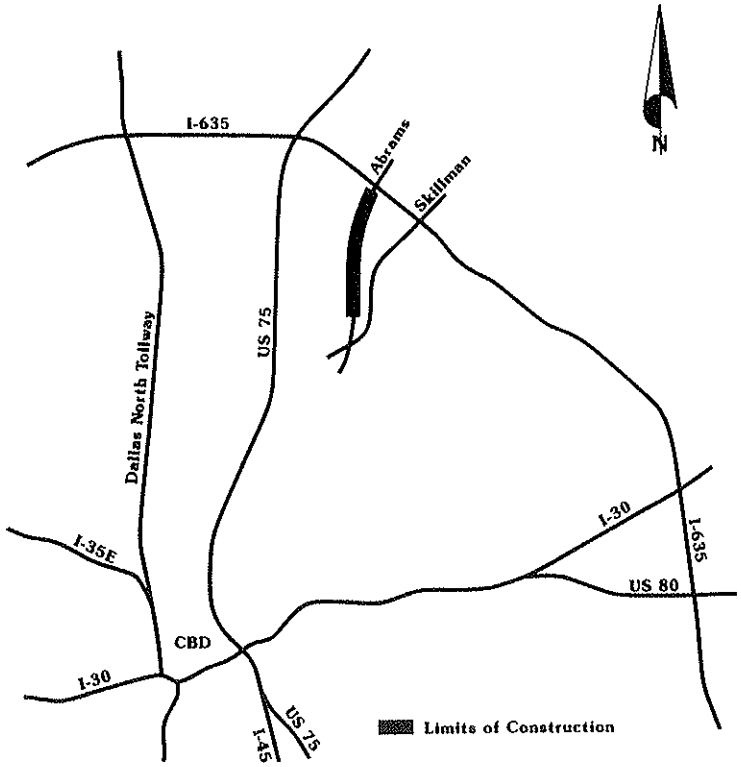



**FIGURE 1 Route map of Dallas, Texas.**

Survey participants were queried on their opinions about various aspects of the reconstruction project. This was followed by a brief set of biographical questions. The participants were then asked to respond to questions regarding work zone signs and other forms of traffic control devices that were presented photographically both in and out of context, as an independent element of the overall traffic control plan.

The Dallas survey results were compared with responses from the Houston survey, which was of a similar type. However, the sample size used in Houston was approximately half of that used in Dallas. It was intended that the surveys support one another as representing similar results from different geographical areas.

## RESULTS

Motorists were first asked about their opinions of the Abrams reconstruction. Figure 2 shows a listing of all the questions with their response frequencies.

"Hazardous road conditions" was the biggest problem as seen by the participants, followed by "The construction is taking too long." Subsequently, drivers were asked, "if they were utilizing alternate routes?" Approximately 65 percent responded that they were using alternative routes.

When asked, "if there are too many, too few, or the right amount of construction signs that give directions to places alongside the construction area?" the response given most often (54 percent) was that there are the right amount of directional signs for the construction area. However, 21 percent said there were too few, 21 percent said they were not sure, and 4 percent said there were too many.

When asked the same question about the number of barrel drums used for channelization, the participants responded most often (50 percent) that there were the right amount. On the other hand, 23 percent said they were not sure, 16 percent said there were too many, and 11 percent said there were too few.

The drivers were also asked, "if they had any trouble getting to specific places because of the construction?" Over 68 percent of the drivers responded that they did not have any problems getting to their destinations. Finally, when asked, "Do the future benefits of this construction outweigh the present inconveniences?", approximately 84 percent responded "yes."

The second part of the survey asked the participants to respond to questions regarding work zone traffic control. This part of the survey revealed that drivers have some difficulty correctly interpreting messages on construction signs. A sample of the work zone control device questions is shown in Figure 3. The response percentages for each device are discussed in the following sections.

### Road Construction 500 ft

Over two-thirds (69 percent) of the participants correctly interpreted the sign in Figure 4. However, approximately 22 percent of the respondents interpreted the sign to mean that the next 500 ft of road are under construction. The Houston

1. Have you travelled on Abrams Road during the current construction?
   84% - Yes  16% - No
2. How often do you travel on Abrams?
   44% - One or more trips each day
   32% - One or more trips each week
   18% - One or more trips each month
   6% - Less than once a month
3. What is the biggest problem in the Abrams construction area?
   25% - Hazardous road conditions
   24% - The construction work is taking too long
   18% - Travel delay caused by construction
   8% - Too much traffic
   8% - Difficulty making turns due to congestion
   7% - Other
   5% - The construction zone is too long
   5% - Inadequate or confusing lane stripping

4a. During what time do you experience delay, if any, due to construction?
   34% - Evening rush hour
   26% - Morning rush hour
   26% - Other time period
   14% - No delay experienced

4b. How much time does the construction delay you?
   34% - Less than 5 minutes
   29% - 6-10 minutes
   19% - No delay experienced
   10% - 11-15 minutes
   8% - More than 15 minutes

5. Is this delay reasonable?
   52% - Yes  16% - No delay experienced
   21% - No  11% - Not Sure

6a. Are you using alternate routes?
   65% - Yes  35% - No

6b. If you are using alternate routes, what are they?
   39% - Greenville  26% - Skillman
   35% - Other

7a. Do the future benefits outweigh the present inconveniences?
   84% - Yes  5% - No
   11% - Not Sure

7b. If no or not sure, why?
   36% - Other
   28% - None Given
   16% - Construction work is taking too long
   12% - Fine before widening
   8% - Cost too high, bad construction planning

8. Do you have trouble getting to specific places because of the construction?
   68% - No  32% - Yes

9. How are the construction signs?
   54% - Right amount  21% - Not Sure
   21% - Too few  4% - Too many

10. How are the construction barrels?
   50% - Right amount  16% - Too many
   23% - Not Sure  11% - Too few

11. How would like to receive roadway project information?
   27% - Radio  12% - Newsletter, flier, etc
   26% - Newspaper  9% - Utility bill stuffer
   23% - Television  3% - Local cable channel

12. Which radio station(s) do you listen to for news?
   36% - Other  5% - KKDA-104.5 FM
   19% - None  5% - KSCS-96.3 FM
   17% - KRLD-1080 AM  5% - KERA-90.1 FM
   8% - KVIL-1150 AM, 103.7 FM
   5% - KLIF-190 AM

13a. Where do you currently get traffic/road closure information?
   54% - Radio  18% - Television
   19% - Newspaper  9% - Other

13b. Other traffic/road closure information?
   66% - No information received
   14% - Road Signs
   8% - Other
   6% - Call traffic information agency
   6% - Homeowners association

14. Do you have concerns about other highway projects in the Dallas area?
   39% - Other
   29% - None given
   24% - N. Central-Miscellaneous
   7% - Skillman-Miscellaneous
   6% - N. Central-Should have been done sooner
   5% - N. Central-Too much traffic, causes delays

**FIGURE 2 Questionnaire response summary.**

1. What does this sign tell you? (Figure 4)
   A. There are 500 feet of construction 500 feet ahead
   B. The next 500 feet of road are under construction
   C. A construction area is located 500 feet ahead
   D. Not sure
2. How would you respond to this sign? (Figure 5)
   A. Turn left
   B. Stop
   C. Change lanes
   D. Not sure
3. Why are these signs different colors? (Figure 6)
   A. Yellow is for school zones, Orange is the standard color for warning signs
   B. Yellow is the standard color for warning signs, Orange is for construction signs
   C. There is no difference between the two
   D. Not sure
4. What does this sign tell you? (Figure 7)
   A. Low shoulder
   B. Uneven pavement
   C. Bumpy road
   D. Not sure
5. What do the orange and black arrows tell you? (Figure 8)
   A. Do not turn left between signs
   B. Shows the direction of the roadway
   C. Sharp turns in the road
   D. Not sure
6. On which side of this sign would you drive? (Figure 9)
   A. Drive to the right of these signs
   B. Drive to the left of these signs
   C. Drive to either side of these signs
   D. Not sure
7. Where would you turn left? (Figure 10)
   A. Before the Crossover sign
   B. After the Crossover sign
   C. Either before or after the Crossover sign
   D. Not sure
8. What do the white posts on the right tell you? (Figure 11)
   A. Shows the driveway locations along the roadway
   B. Shows the right edge of the pavement
   C. Park between these posts
   D. Not sure
9. What does this sign tell you? (Figure 12)
   A. Road construction ahead
   B. Flagger ahead
   C. Guard for school crossing ahead
   D. Not sure
10. What does this sign tell you? (Figure 13)
   A. Median narrows
   B. Right lane ends
   C. Right lane turn marker
   D. Not sure
11. What does this sign tell you? (Figure 14)
   A. Leave room for traffic crossing at intersection
   B. If your car stalls, move it out of the intersection
   C. Avoid driving through the intersection
   D. Not sure

**FIGURE 3 Sign questionnaire.**


FIGURE 4 Advance road construction sign.

survey produced similar results, with only 66 percent correctly identifying the sign.

### Right Lane Ends Sign

Ninety percent of the respondents correctly interpreted the sign in Figure 5 to indicate that a change in lanes was necessary. The correct response to the Houston survey was approximately 93 percent.

### Color Cue Difference

In the past, distinguishing different color cues has not been accomplished well by the motorist. When drivers were shown the Two Way Traffic sign (see Figure 6), one yellow and one orange, and asked the meaning of the two different colors, only 50 percent knew that orange is the color designated for construction. Twenty five percent were not sure of the difference and the other 25 percent gave an incorrect interpretation. In Houston, only 44 percent knew the correct meaning of the two colors and 40 percent said they were not sure of the difference.


FIGURE 5 Right lane ends sign.

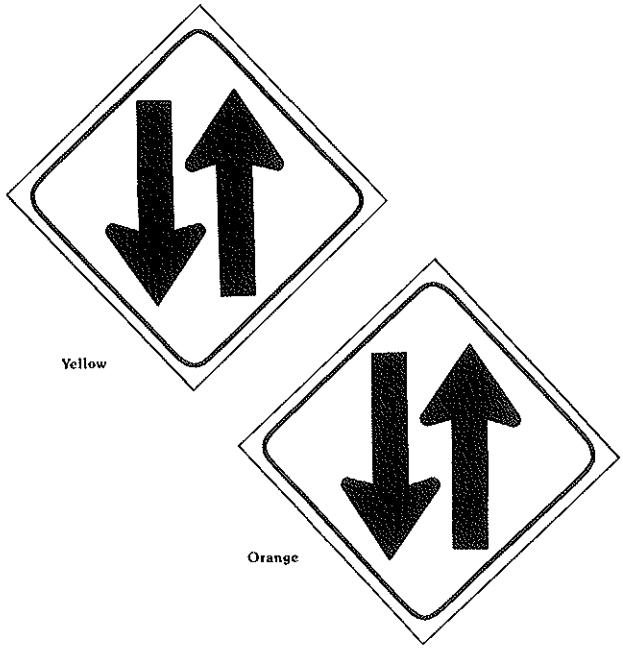



FIGURE 6 Color cue difference of signs.

### Low Shoulder Symbol Sign

This sign (shown in Figure 7) was incorrectly interpreted by a majority of the respondents. Over three-quarters (76 percent) thought this sign meant uneven pavement. The Houston survey had a similar response of 84 percent.

### Chevron Alignment Sign

Eighty five percent of the respondents correctly interpreted the sign in Figure 8 to indicate that the signs show the direction of the roadway. The correct response to the Houston survey was approximately 92 percent.

FIGURE 7 Low shoulder symbol sign.

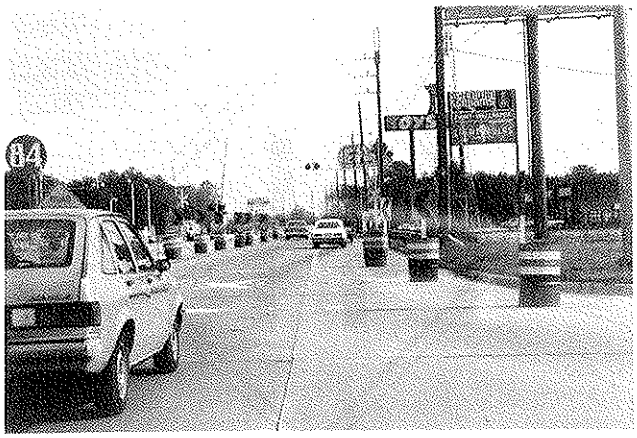


**FIGURE 8 Chevron alignment sign.**

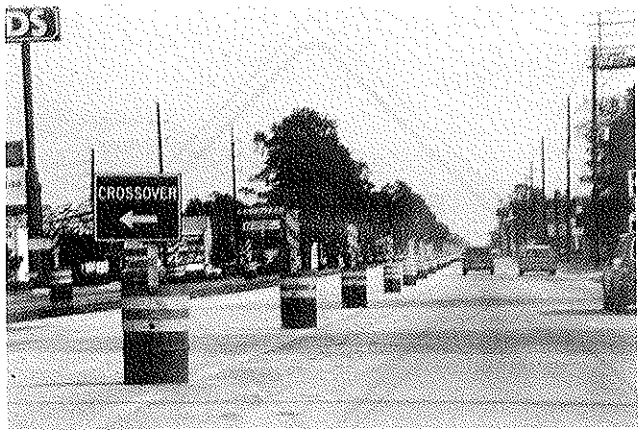



**FIGURE 10 Crossover sign.**

### Vertical Panel

Orange and white hazard markers were shown to the survey respondents (see Figure 9). Drivers were asked, "on which side of these signs they would drive?" Thirty eight percent responded incorrectly, whereas 46 percent were not sure. The Houston survey did not specifically address this sign. However, some indications of the previous survey exhibited incorrect interpretations of the sign.

### Crossover Sign

These signs do not clearly convey where to cross over within the construction area (see Figure 10). Because of the limited spacing requirements of urban arterial construction zones, it is extremely difficult to provide sufficient spacing for crossover situations. There are no current signing alternatives available to delineate a crossover situation within a construction zone. Therefore, the standard green crossover sign was used. When drivers were asked where they would turn left, 53 percent said before the sign, 26 percent said after the sign, and 13 percent were not sure. The Houston survey had similar results in that 55 percent responded that it was permissible to cross over before the sign, 42 percent indicated it was permissible to turn after the crossover sign, and 3 percent were not sure.

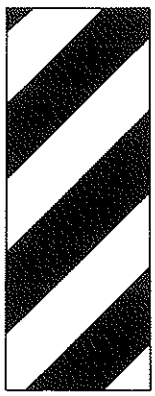


**FIGURE 9 Orange and white hazard marker (vertical panel).**

### White Delineator Posts

These delineators (Figure 11) were used in conjunction with white raised pavement markers within the construction area to delineate clearly the edge of pavement. Seventy five percent of the drivers interpreted these correctly, whereas 9 percent did not correctly interpret the markers, and 16 percent were not sure of their meaning. The Houston survey revealed that 58 percent interpreted the markers correctly, 36 percent misinterpreted them, and 6 percent of the respondents were not sure.

### Advance Flagger Symbol Sign

The Flagger Ahead symbol, shown in Figure 12, was interpreted correctly by 79 percent of the respondents; 21 percent interpreted the symbol incorrectly. The Houston survey re-

**FIGURE 11 White delineator posts.**

FIGURE 12 Advance flagger symbol sign.

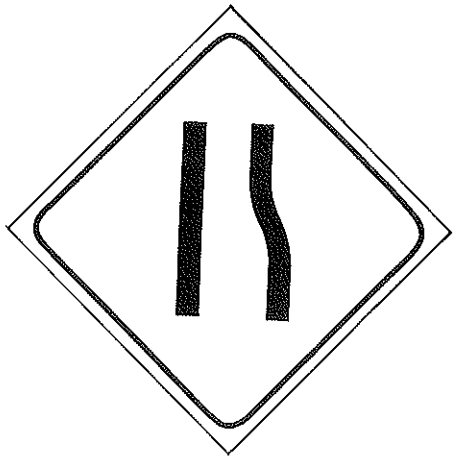


FIGURE 13 Lane reduction transition symbol sign.

vealed similar results with 78 percent correctly interpreting the symbol.

### Lane Reduction Transition Symbol Sign

The Lane Reduction Transition symbol, shown in Figure 13, was interpreted correctly by 74 percent of the survey, whereas 20 percent misinterpreted the symbol and 6 percent were not sure. The Houston survey revealed that 78 percent interpreted the symbol correctly, 19 percent incorrectly interpreted the symbol, and 3 percent were not sure.

### Do Not Block Intersection Sign

Eighty eight percent of the survey participants correctly interpreted the sign in Figure 14 to indicate that the driver must leave room for traffic crossing at the intersection. The Houston survey revealed that 74 percent of the respondents answered correctly.

A brief demographic summary concluded the interview. The results are presented in Table 1 along with the Dallas regional population statistics. The survey sample size was representative of the overall population.

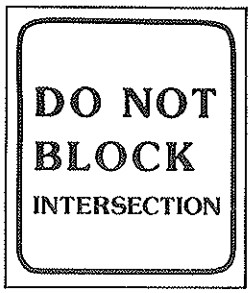



FIGURE 14 Do Not Block Intersection sign.

## CONCLUSIONS

The survey format of in- and out-of-context traffic control device photographs provided as identifying assessment of motorist confusion in understanding work zone arterials. How-

TABLE 1 DEMOGRAPHIC SUMMARY

| Demographic Survey | | Survey Sample | ** Regional Population ** |
|---|---|---|---|
| SEX: | Male | 48% | 48% |
| | Female | 52% | 52% |
| AGE: | < 25 | 17% | 20% |
| | 26 - 55 | 60% | 56% |
| | > 55 | 23% | 24% |
| ETHNICITY: | | | |
| | Anglo | 72% | 76% |
| | Black | 16% | 15% |
| | Hispanic | 6% | 8% |
| | Other | 6% | 1% |
| EDUCATION: | | | |
| | College Graduate | 52% | ---- |
| | Some College | 31% | ---- |
| | High School Graduate | 13% | ---- |
| | < High School | 4% | ---- |

** Source: Bureau of Census (1988)

ever, the interpretations of the survey may not necessarily predict responses within the roadway environment when used as part of a complete traffic control plan. Specific problems identified in the development of the survey included presenting appropriate responses in an unbiased instrument that would serve the researcher's needs and being able to collect a representative sample of the population.

The survey did indicate that the motorists have problems understanding selected work zone signing. There was some confusion experienced by 31 percent of the respondents with the interpretation of the advanced construction sign. The difference between the standard color cues, yellow warning sign and orange construction sign, was incorrectly identified by 25 percent of the respondents, and 25 percent were not sure of the difference. The orange and white hazard markers (vertical panels) were also identified as a problematic sign by 84 percent. The behavioral response of the placement of the standard Crossover sign was interpreted incorrectly by 47 percent. Over three quarters (76 percent) responded incorrectly to the low shoulder symbol sign. The other survey results showed some indications of minor problems.

The Houston survey indicated similar misinterpretations of the advanced construction, difference of color cues, orange and white hazard markers (verticle panels), Crossover, and low shoulder signs. Further education of the motoring public or investigation of alternative signing schemes to improve comprehension was not within the scope of this study. However, further research of motorist understanding and appropriate response for ensuring safe negotiation of roadway work zones seems justified.

Motorist concerns on the Abrams Road project seem to be concentrated around the construction work taking too long and the hazardous road conditions, 24 and 25 percent, respectively. Alternate routes were being used by 65 percent of the respondents. Motorists indicated that there were a sufficient amount of signs and channelizing barrel drums on the project, 54 and 50 percent, respectively. The overall perception of 84 percent of the respondents was that the benefits of widening the roadway would be worth the inconveniences they are experiencing now.

## ACKNOWLEDGMENTS

This study conducted by the Texas Transportation Institute was sponsored by the Texas State Department of Highways and Public Transportation (SDHPT).

The successful completion of this study required the cooperation and assistance of numerous agencies and individuals. The authors would particularly like to thank Laura Moore of the SDHPT, District 18, and the Texas Department of Public Safety for their assistance in this undertaking.

## REFERENCES

1. *Manual on Uniform Traffic Devices for Streets and Highways.* FHWA, U.S. Department of Transportation, 1988.
2. Safety Design and Operational Practices for Streets and Highways. Report 80-228. FHWA, U.S. Department of Transportation, 1980.
3. Effectiveness of City Traffic - Control Programs for Construction and Maintenance Work Zones. In *Transportation Research Record 833*, TRB, National Research Council, Washington, D.C., 1981, pp. 6–9.
4. Work Area Traffic Control: Evaluation and Design. *Journal of Transportation Engineering*, ASCE, 1984.
5. K. N. Womack, P. K. Guseman, and R. D. Williams. *Measuring Effectiveness of Traffic Control Devices: An Assessment of Driver Understanding.* Texas Transportation Institute, College Station, June 1981.
6. M. A. Ogden, D. E. Morris, and B. C. Rymer. *FM 1960 Origin-Destination Study.* Texas Transportation Institute, College Station, Nov. 1988.
7. M. A. Ogden, K. N. Womack, J. M. Mounce. *Motorist Comprehension of Signing Applied in Urban Arterial Work Zones.* Texas Transportation Institute, College Station, Jan. 1990.
8. Texas Manual on Uniform Traffic Control Devices. Texas State Department of Highways and Public Transportation, Austin, 1980, revised to 1988.



*Publication of this paper sponsored by Committee on Traffic Safety in Maintenance and Construction Operations.*

# Need To Stripe No-Passing Zones During Resurfacing of Lower-Volume Rural Roads

MARK R. VIRKLER AND DAVID L. GUELL

The lack of no-passing zone markings during the resurfacing of two-lane highways may produce a hazard to the driving public. The objectives of the research were to evaluate the potential safety problems and to recommend a traffic volume at which there is a significant hazard associated with not having no-passing markings in place during a resurfacing project. Analytic models and simulation models were used to predict the number of passes, the potential for passing conflicts, and the number of delayed passes at various traffic volumes. Traffic volumes were also related to highway level of service and accidents in Missouri involving improper passing. Potential reductions in accident costs were related to the cost of temporary no-passing zones. Recommendations for marking no-passing zones were based on highway classification, average daily traffic, and terrain type.

It is common practice to place centerline markings on two-lane, two-way paved rural highways. The *Manual on Uniform Traffic Control Devices* (*MUTCD*) (*1*) recommends centerlines when the two-lane pavement is 16 ft wide or more and the prevailing speed exceeds 35 mph. If centerline markings are present, the MUTCD requires no-passing zone markings where sight distance is restricted.

During a pavement resurfacing project, a road generally remains in service to traffic. A temporary broken yellow centerline marking (of 4 ft dashes), without the associated no-passing markings, is generally placed during the paving operation. Note that this is clearly contradictory to the MUTCD. The permanent marking system, including no-passing markings, is placed later, often after completion of the entire project. The temporary nonconforming marking system could be in place for as long as 2 weeks.

A road with horizontal and vertical curves would generally have sections lacking adequate passing sight distance. The lack of no-passing zone markings may produce a significant hazard to the driving public. The extent of the hazard would be related to the number of passing maneuvers that typically occur on the road. The number of passing maneuvers is related to the traffic volume.

The objectives of the research were to

1. Evaluate the degree of safety hazard associated with not having no-passing markings in place during resurfacing operations and
2. Recommend a traffic volume at which there is a significant safety hazard associated with not having no-passing markings in place during a resurfacing project.

Department of Civil Engineering, University of Missouri—Columbia, Columbia, Mo. 65211.

Traffic volumes were related to the following indicators of potential hazard:

1. Passes per mile per hour,
2. Passes per mile per hour that would conflict with oncoming traffic if the passing vehicle ignored the presence of an oncoming vehicle,
3. Potential passing vehicles that are delayed because of the presence of oncoming vehicles,
4. Highway level of service (LOS), and
5. Accidents in Missouri involving improper passing.

A benefit-cost analysis was also used to relate potential reductions in accident costs to the cost of temporary no-passing striping.

## PAST STUDIES

### Glennon

Glennon (*2*) used a benefit-cost analysis to determine whether no-passing zones were appropriate for low-volume roads [average daily traffic (ADT) of 400 vpd or lower]. Glennon estimated that on roads with an ADT of 400 vpd the cost of striping would be three times the cost of accidents prevented.

The assumptions used in Glennon's analysis were as follows:

1. One-half of the roadway has restricted sight distance,
2. One-half of all head-on collisions involve passing, and
3. The presence of no-passing stripes reduces head-on accidents involving passing maneuvers in restricted sight distance areas by one-half.

Glennon estimated that the accident rate on roads with ADT values of 400 vpd was 0.367/mi-yr and that 13.7 percent of these accidents were head-on. The cost per accident was $9,500. The cost of striping 50 percent of the roadway was $176/mi-yr. The expected additional accident cost without striping was $60/mi-yr. If all accidents were fatal or involved injury, the additional cost would be $123/mi-yr.

### Josey

The simulation model used by Josey et al. (*3*) predicted passing rates and passing conflicts that would occur if all drivers

initiated passes without regard to the presence of conflicting vehicles. The simulation involved 1 hr of operation on 3 mi of road. The simulation runs indicated no passing conflicts when the volume was 80 veh/hr or lower. The simulation resulted in 1.25 conflicts/mi-hr for a volume of 90 veh/hr and 1.0 conflicts/mi-hr for a volume of 100 veh/hr. The least squares equation calibrated from the simulation runs indicate, for a 50–50 directional split, the following correlation:

| *Hourly Volume* | *Conflicts/mi-hr* |
|---|---|
| 60 | 0.32 |
| 70 | 0.50 |
| 80 | 0.72 |
| 90 | 0.99 |
| 100 | 1.33 |
| 120 | 2.20 |
| 140 | 3.36 |
| 160 | 4.86 |

When the volume on the two-lane highway became high enough, many passes were not completed because of the large number of opposing vehicles. Figure 1 shows the simulation results. At low volumes, passes are roughly proportional to the square of volume. At high volumes, the number of passes decreases because of the small number of gaps available for passing.

The conclusion of the report (*3*) stated, "The probability of passes and emergency indicators (conflicts) approaches zero as traffic volumes decline below a value of 100 vehicles per hour." A conclusion derived from this study was that no-passing zones should be provided when the volume exceeds 1,000 veh/day.

## RESEARCH APPROACH

The number of vehicle overtakings and potential passing conflicts were related to traffic volumes by three methods:

1. Glennon's (*2*) formulation,
2. Wardrop's (*4*) formulation,
3. Simulation by ROADSIM (*5*).

The *Highway Capacity Manual* (*6*) was used to describe LOS for various traffic and roadway conditions. Accident experience involving improper passing was evaluated using 1988 Missouri accident data (*7*).

## OVERTAKING MODELS

### Analytical Models

Glennon (*2*) developed a method to estimate expected number of passes, probability of an oncoming vehicle being in conflict with the passing vehicle, and expected number of passing conflicts. The expected number of passes is based on an assumption of random vehicle headways. At a given point in time, two vehicles with a headway of 1 sec or less are assumed to be engaged in a passing maneuver. The expected number of passes per unit length per unit time are extrapolated from this headway assumption. The number of passes is a function only of flow rate and is independent of the mean speed or the speed distribution.

The probability of an oncoming vehicle's being in conflict with a passing vehicle is based on Poisson (random) arrivals of vehicles at a point. The probability of a conflict increases with the flow rate in the opposing direction and the time period required to complete the pass.

$$P(A) = 1 - P(0) = 1 - e^{-Vt/3,600} \quad (1)$$

where

$P(A)$ = probability of a passing vehicle encountering a conflict,
$P(0)$ = probability of a passing vehicle encountering no opposing vehicles,
$V$ = flow rate in the opposing direction (veh/hr),
$t$ = time period in which passing vehicle is vulnerable to a conflict (sec), and
3,600 = number of seconds per hour.

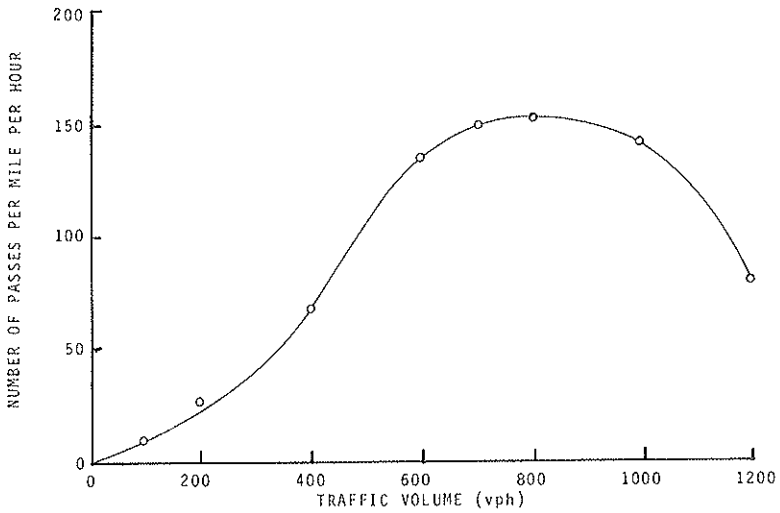



**FIGURE 1 Number of passes versus two-way hourly volume for higher flow rates.**

The expected number of passing conflicts per unit time per unit length would be the expected number of passes multiplied by the probability of a conflict.

Wardrop (*4*) developed an expression to determine the frequency with which vehicles overtake one another. If speed normally distributed and every driver chooses to pass when a slower vehicle is encountered, then the number of overtakings ($N$) per unit length per unit time for vehicles traveling in one direction is as follows:

$$N = \frac{Q^2\sigma}{\bar{v}^2\pi^{1/2}} = \frac{0.56\ Q^2\sigma}{\bar{v}^2} \quad (2)$$

where

$Q$ = one-way flow rate,
$\sigma$ = standard deviation of speed, and
$\bar{v}$ = space mean speed.

If there is interference with overtaking because of oncoming traffic, this expression might be taken to be the number of desired overtakings. For a given distribution of speeds (mean and standard deviation), the number of desired overtakings increases with the square of flow.

Matson et al. (*8*) described a model that was based on similar assumptions. The model involved the summation of overtakings between vehicles within different speed groups. The model yields results identical to the Wardrop formulation.

### Simulation: ROADSIM Model

ROADSIM (*5*) is a two-lane highway simulation model developed by the FHWA during the 1980s. An earlier version of this model (TWOWAF, developed by Midwest Research Institute) was modified and used by the Texas Transportation Institute to develop the two-lane highway procedure for the 1985 edition of the *Highway Capacity Manual (6)*. The version of ROADSIM used for the research reported here has a November 1987 revision date and runs on a microcomputer. It was obtained from the FHWA's Traffic Safety Research Division in October 1989.

## RESULTS OF OVERTAKING MODELS

### Analytical Results

The Glennon (*2*) and Wardrop (*4*) methods were used to predict the number of passes as a function of hourly volume. The probability of a passing vehicle's encountering an opposing vehicle was determined from the Poisson distribution (assuming random vehicle arrivals in the opposing direction).

The following assumptions were used to predict passing demand for the Glennon formulation:

1. Two vehicles within 1 sec of each other are engaged in a passing maneuver, and
2. Nine seconds are required for the passing maneuver.

For the Wardrop formulation, the following assumptions were used to predict passing demand:

1. Space mean speed = 45 mph and
2. Standard deviation of speed = 5 mph.

Figure 2 shows the number of passes per mile per hour derived from the two analytical models.

To predict potential conflicts, the distance traveled by the passing vehicle plus the distance traveled by the opposing vehicle in 5 sec was taken as the conflict distance. If an opposing vehicle was within this 10-sec window, then a potential conflict would occur. The implicit assumption is that all passes are initiated without regard to the presence of opposing vehicles, and the presence of an opposing vehicle within a 10-

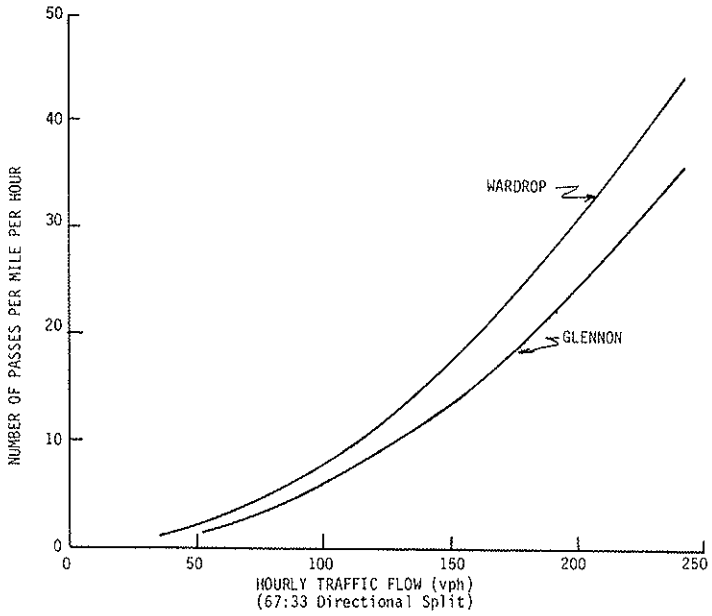



**FIGURE 2 Number of passes versus two-way hourly volume.**

sec window results in a conflict. Figure 3 shows the number of passing conflicts per mile per hour for each of the analytical models.

The Glennon (*2*) formulation depends on the assumption of random vehicle arrivals at a point without regard to speed or the standard deviation of speed. The Wardrop (*4*) formulation explicitly considers speed characteristics. The more variation in speed (the higher the standard deviation of speed), the higher the number of passes that will occur. For that reason, the Wardrop formulation was considered preferable to the Glennon formulation for predicting number of passes.

The Wardrop results for potential conflicts per day as a function of ADT are presented in Table 1. An average day was assumed to consist of 1 hr at 15 percent of ADT, 3 hr at 10 percent of ADT, and 11 hr at 5 percent of ADT. Each hour has a 67–33 directional split.

TABLE 1 POTENTIAL PASSING CONFLICTS PER DAY VERSUS ADT BY WARDROP (*4*) FORMULATION WITH 67-33 DIRECTIONAL SPLITS IN EACH HOUR OF THE DAY

| ADT | Potential Conflicts/day[a] |
|---|---|
| 225 | 0.08 |
| 400 | 0.43 |
| 600 | 1.44 |
| 800 | 3.37 |
| 1,000 | 6.49 |
| 1,200 | 11.1 |
| 1,400 | 17.5 |
| 1,600 | 26.4 |
| 1,800 | 37.2 |

[a]The potential conflicts reported in this table are based on the assumptions that (a) all drivers are willing to pass when adequate sight distance is not available, and (b) passing sight distance is never available.

## Simulation Results

ROADSIM was first used to determine the number of passes that would be initiated in one direction of traffic flow if there were no opposing traffic to restrict the passing maneuvers. Two-way flow simulations were then run to determine how opposing traffic affects the passing pattern.

Simulations were conducted for one-way flow rates ranging from 25 to 175 veh/hr. Two-way simulations were run for flows rates ranging from 50 to 255 veh/hr. The range of hourly flow rates was selected to evaluate a range of ADT volumes from 400 to 1,700 veh/day. The upper value of hourly flow rates (255 veh/hr) is 15 percent of 1,700 veh/hr. Twenty-five vehicles per hour is the lowest nonzero value the simulation model will accept.

Because it was desired to determine the maximum number of passes that would likely occur for a given flow rate, the simulation runs were conducted on an ideal roadway. A 4-mi-section of straight and level roadway was used. At each end of the simulated roadway there was a half-mile section in which no passing was allowed. ROADSIM requires these end sections, over which no data is collected, because of the car-following logic used. For the simulated roadway, data were collected on the number of passes initiated in each 1-mi section. Three miles was selected as the length of roadway over which to collect data because this length was considered to be a typical length of road on which there would be few vehicles entering or leaving a typical rural, low-volume, two-lane Missouri highway. The ROADSIM model puts vehicles into the roadway only at the ends.

Simulation runs were made for 1 hr. A speed distribution with an average of 45 mph and a standard deviation of 5 mph was selected as input to ROADSIM. Only passenger cars were included in the simulated traffic stream. The assumption of no trucks (other than pickup trucks) was considered reasonable for typical low-volume rural roads of the collector and local functional classes in the state of Missouri.

Two direction distributions were used, 67–33 and 50–50. The average number of passes per mile (total in both directions) initiated in the two-way flow simulations (over the range

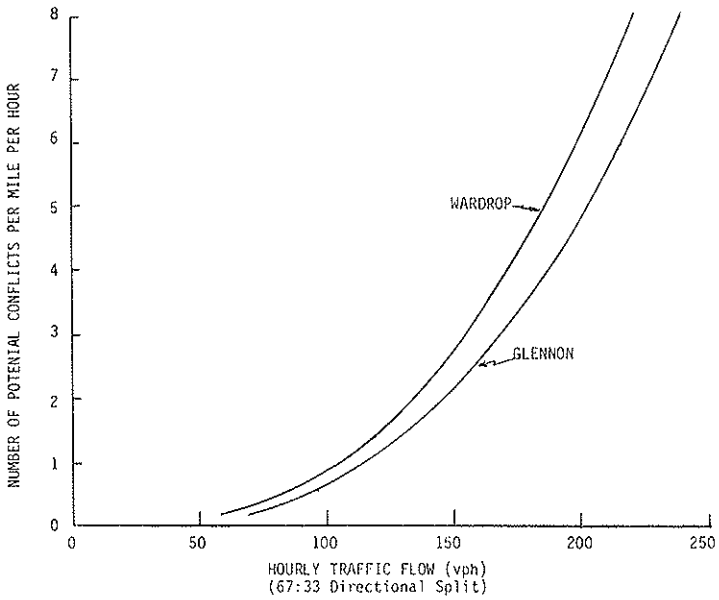



**FIGURE 3 Passing conflicts versus two-way hourly volume.**

of flows from 50 to 255 veh/hr) was found to be independent of the direction distribution of the traffic. The relationship between the two-way flow rate and the average number of passes initiated per mile in the simulated hour is given by

$$PPM = 0.00203Q_2^{1.79} \quad (3)$$

where

PPM = passes per mile initiated, and
$Q_2$ = two-way hourly flow rate.

This equation was obtained by a least squares fit of 16 data points over the range of hourly flow rates described earlier. The equation had a coefficient of determination ($R^2$) value of 0.92. Within the volume range studied over the 3-mi section, the number of passes per mile was found to be approximately the same with and without opposing traffic. Figure 4 shows the ROADSIM results along with the Wardrop (*4*) results.

Conflicts between vehicles desiring to initiate a passing maneuver and opposing vehicles were determined by comparing the number of passes initiated in each direction for the two-way flow and those initiated for one-way flow and hence no opposing traffic. The number of passes in each of the three 1-mi links of the roadway was compared between the two-way and the one-way flow. If the number of passes initiated on a 1-mi link was lower with opposing flow than for the same one-way volume with only one-way flow, then that number of delayed passes was counted for that 1-mi link because those passes were delayed into the next 1-mi link. If then, for example, in the next 1-mi link, the numbers of passes initiated for one-way and two-way flow were the same, an additional number of delayed passes would be counted equal to the same number in the first link.

Within this definition of delayed passes, the relationship between the two-way flow rate and the average number of delayed passes per mile in the hour was found to be

$$DPPM = 0.0000045Q_2^{2.57} \quad (4)$$

where DPPM is the number of delayed passes per mile. This equation was developed by a least squares fit to seven data points with nonzero number of delayed passes. The coefficient of determination was 0.975.

Delayed passes are not identical to potential conflicts as discussed under the Glennon and Wardrop formulations. With ROADSIM analysis, a pass can only be identified as being delayed if conflicting traffic postpones a desired pass from one section to a downstream section. A pass that is delayed but still occurs within the same 1-mi section is not detected in comparing unopposed and opposed simulation runs.

Table 2 presents a tabulation of Equations 3 and 4. Also presented in this table is the percent of the passes that are delayed.

By assuming a distribution of the average daily traffic volume throughout the day, it was possible to determine the number of passes initiated per mile per day and the average number of delayed passes per mile per day. An average day was as assumed previously. Table 3 presents the number of passes and delayed passes, and the ratio over the range of ADT values from 1,700 to 400 veh/day.

TABLE 2 RELATIONSHIP BETWEEN HOURLY FLOW RATE, PASSES, AND DELAYED PASSES

| Flow Rate (veh/hr) | Passes (per mile per hour) | Delayed Passes (per mile per hour) | Delayed Passes per Pass (%) |
|---|---|---|---|
| 250 | 39.8 | 6.55 | 16.4 |
| 225 | 33.0 | 4.99 | 15.2 |
| 200 | 26.7 | 3.69 | 13.8 |
| 175 | 21.0 | 2.62 | 12.4 |
| 150 | 16.0 | 1.76 | 11.0 |
| 125 | 11.5 | 1.10 | 9.6 |
| 100 | 7.7 | 0.62 | 8.0 |
| 75 | 4.6 | 0.29 | 6.4 |
| 50 | 2.2 | 0.10 | 4.7 |

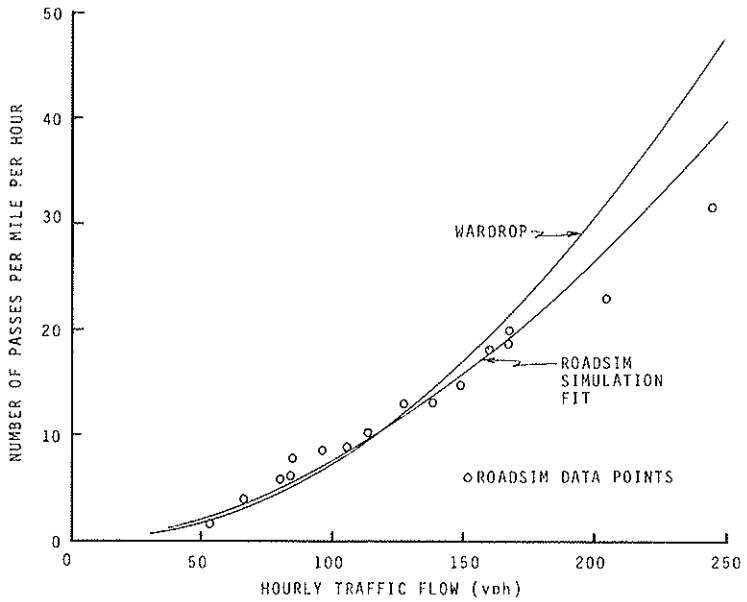



**FIGURE 4 Number of passes versus two-way hourly volume.**

TABLE 3 RELATIONSHIP BETWEEN ADT, PASSES, AND DELAYED PASSES

| ADT (veh/day) | Passes (per mile per day) | Delayed Passes (per mile per day) | Delayed Passes per Pass (%) |
|---|---|---|---|
| 1,700 | 165 | 18.7 | 11.4 |
| 1,400 | 116 | 11.3 | 9.8 |
| 1,200 | 88 | 7.6 | 8.6 |
| 1,000 | 64 | 4.8 | 7.5 |
| 800 | 43 | 2.7 | 6.3 |
| 600 | 26 | 1.3 | 5.0 |
| 400 | 12 | 0.4 | 3.7 |

## LOS

The *Highway Capacity Manual* (*6*) describes the quality of flow associated with each LOS for two-lane highways. In the description of LOS A:

> The highest quality of traffic service occurs when motorists are able to drive at their desired speed. Without strict enforcement, this highest quality, representative of LOS A, would result in average speeds approaching 60 mph on two-lane highways. The passing frequency required to maintain these speeds has not reached a demanding level. Passing demand is well below passing capacity, and almost no platoons of three or more vehicles are observed. Drivers would be delayed no more than 30% of the time by slow moving vehicles. A maximum flow rate of 420 pcph, total in both directions, may be achieved under ideal conditions.

In LOS B, the passing demand becomes more important:

> LOS B characterizes the region of traffic flow wherein speeds of 55 mph or slightly higher are expected on level terrain. Passing demand needed to maintain desired speeds becomes significant and approximately equals passing capacity at the lower boundary of LOS B. . . .

From these descriptions it appears that no-passing pavement markings are desirable for LOS B operations, even if this LOS value is present only in the peak hour. On the other hand, LOS A operations may be acceptable without no-passing pavement markings. The relatively small number of drivers desiring to pass would have little difficulty finding appropriate passing opportunities.

The procedure of the *Highway Capacity Manual* has been adapted for the lower-volume two-lane highways considered in this study. The following assumptions were used:

- General terrain segment operating at LOS A;
- Traffic includes 6 percent trucks, no RVs, and no buses;
- 60/40 directional split;
- 12-ft lanes and 6-ft usable shoulder (with relatively low volumes, narrow lanes and restricted shoulders have minimal effects on flow);
- For level terrain, 20 percent no passing zones;
- For rolling terrain, 40 percent no passing zones; and
- For mountainous terrain, 60 percent no passing zones.

Table 4 presents the ADT values associated with providing LOS A in the peak hour. The *K*-factor is the proportion of ADT in the peak hour. Two-lane rural highways in Missouri are classified as presented in Table 5.

TABLE 4 MAXIMUM ADT VERSUS TERRAIN FOR LOS A IN PEAK HOUR ON TWO-LANE RURAL HIGHWAYS

| *K*-Factor | Level Terrain | Rolling Terrain | Mountainous Terrain |
|---|---|---|---|
| 0.10 | 2,979 | 1,561 | 741 |
| 0.11 | 2,708 | 1,419 | 673 |
| 0.12 | 2,482 | 1,301 | 617 |
| 0.13 | 2,291 | 1,201 | 570 |
| 0.14 | 2,128 | 1,115 | 529 |
| 0.15 | 1,986 | 1,041 | 494 |

In terms of the peak-hour LOS, Table 4, indicates that, for a peak hour equal to 15 percent of ADT, LOS A should not be expected on most arterials. LOS A would be expected on all local roads. For collectors, LOS A would be expected in level terrain but would not be expected in mountainous terrain. In rolling terrain, LOS A would be expected when the ADT is below about 1,041.

## MISSOURI TWO-LANE HIGHWAY ACCIDENT CHARACTERISTICS

The accident rates in Missouri for 1988 were reported by R. Coplen (unpublished correspondence). Table 6 presents 1988 accidents by route marking designation on the state system. The route markings of primary interest in this study are state lettered. Table 7 presents accidents by type and Table 8 presents accidents by contributing circumstances. The contributing circumstances of primary interest to this study is improper passing. Improper passing contributed to 23 fatal accidents (3.0 percent), 571 injury accidents (2.5 percent), and 1,742 property damage accidents (3.3 percent).

The accident rate on Missouri routes with letter designations is 285 accidents per 100 million miles traveled. Using this rate for the roads considered, the expected numbers of accidents per mile per year and accidents per mile per day are presented in Table 9. The accident cost per mile per day is based on Missouri accident patterns (R. Coplen, unpublished correspondence) and the work of Miller et al. (*8*). The assumed average cost per accident is $32,900. Fatal accidents were valued at $2,300,000, injury accidents at $22,000, and property damage accidents at $5,423. The value for fatal accidents was based on the concept of rational investment levels and "is consistent with universal Federal practice in benefit-cost analysis"(*8*). The value is approximately four times the "cost to society."

TABLE 5 FUNCTIONAL CLASSES OF RURAL HIGHWAYS IN MISSOURI (R. Coplen)

| Functional Classification | ADT | Approximate Percentage of Miles |
|---|---|---|
| Arterials | Over 1,700 | 7.0[a] |
| Collectors | 400–1,700 | 22.6 |
| Local roads | Under 400 | 70.4 |

[a]Some arterials would have more than two lanes.

TABLE 6 ACCIDENTS BY ROUTE MARKING DESIGNATION (R. Coplen)

| ROUTE MARKING DESIGNATION | FATAL ACCIDENTS | TOTAL FATALITIES | INJURY ACCIDENTS | TOTAL INJURIES | PROPERTY DAMAGE ACCIDENTS | TOTAL ACCIDENTS |
|---|---|---|---|---|---|---|
| Interstate | 127 | 143 | 4,725 | 7,148 | 11,867 | 16,719 |
| U.S. Numbered | 187 | 222 | 5,258 | 8,646 | 12,934 | 18,379 |
| State Numbered | 232 | 269 | 7,484 | 12,035 | 16,978 | 24,694 |
| State Lettered | 184 | 209 | 4,169 | 6,492 | 7,164 | 11,517 |
| Others | 19 | 22 | 1,240 | 1,961 | 3,495 | 4,754 |
| Totals | 749 | 865 | 22,876 | 36,262 | 52,438 | 76,063 |

TABLE 7 ACCIDENTS BY TYPE (R. Coplen)

| ACCIDENT TYPE | FATAL ACCIDENTS | TOTAL FATALITIES | INJURY ACCIDENTS | TOTAL INJURIES | PROPERTY DAMAGE ACCIDENTS | TOTAL ACCIDENTS |
|---|---|---|---|---|---|---|
| Accident occurred Off Roadway | | | | | | |
| Overturned & Overturning | 53 | 62 | 1,125 | 1,627 | 847 | 2,025 |
| Pedestrian | 0 | 0 | 1 | 1 | 0 | 1 |
| Motor Vehicle in Traffic | 1 | 2 | 39 | 66 | 53 | 93 |
| Parked Motor Vehicle | 2 | 2 | 80 | 108 | 148 | 230 |
| Railroad Train | 1 | 1 | 0 | 1 | 0 | 1 |
| Bicyclist/Pedalcyclist | 0 | 0 | 0 | 0 | 0 | 0 |
| Animal (other than deer) | 0 | 0 | 1 | 1 | 2 | 3 |
| Deer | 0 | 0 | 0 | 0 | 1 | 1 |
| Fixed Object | 277 | 302 | 5,018 | 6,938 | 6,284 | 11,579 |
| Other Object | 0 | 0 | 5 | 5 | 7 | 12 |
| Other, Non-Collision | 2 | 2 | 11 | 14 | 48 | 61 |
| Other | 35 | 38 | 775 | 1,101 | 919 | 1,729 |
| Subtotals | 371 | 409 | 7,055 | 9,862 | 8,309 | 15,735 |
| Accident Occurred on Roadway | | | | | | |
| Overturned & Overturning | 2 | 2 | 199 | 266 | 142 | 343 |
| Pedestrian | 49 | 50 | 316 | 352 | 5 | 370 |
| Motor Vehicle in Traffic | 304 | 379 | 14,251 | 24,401 | 37,814 | 52,369 |
| Parked Motor Vehicle | 1 | 2 | 165 | 247 | 536 | 702 |
| Railroad Train | 2 | 3 | 11 | 21 | 15 | 28 |
| Bicyclist/Dedalcyclist | 1 | 1 | 115 | 117 | 17 | 133 |
| Animal (Other than Deer) | 0 | 0 | 76 | 91 | 531 | 607 |
| Deer | 0 | 0 | 99 | 114 | 2,962 | 3,061 |
| Fixed Object | 14 | 14 | 476 | 638 | 1,078 | 1,568 |
| Other Object | 0 | 0 | 49 | 56 | 667 | 716 |
| Other, Non-Collision | 5 | 5 | 64 | 97 | 362 | 431 |
| Other | 0 | 0 | 0 | 0 | 0 | 0 |
| Subtotals | 378 | 456 | 15,821 | 26,400 | 44,129 | 60,328 |
| TOTALS | 749 | 865 | 22,876 | 36,262 | 52,438 | 76,063 |

Accidents by weather condition are presented in Table 10; accidents by light condition are presented in Table 11. It is expected that recent paving will be obvious in daylight when the pavement is dry and many drivers may not expect pavement markings to be complete.

## ANALYSIS OF POTENTIAL BENEFITS AND COSTS

Marking no-passing zones during resurfacing projects would probably require temporary pavement markings. Permanent markings would then be placed at the completion of the project. The cost of applying preformed removable solid yellow marking tape was approximately $112 per 100 ft in 1989 (*9*). The removal cost was $16 per 100 ft. Temporary pavement striping cost $21.25 per 100 ft of 4 in. solid yellow. For two 4-in. solid lines, the cost would be $1,122 per mile of no-passing zone.

A resurfacing project lasting 14 days was taken as an upper limit of project length. Assume a road with an ADT of 2,000 veh/day, a repaving project lasting 14 days, and that no-passing zones were marked at the end of the project. On average, a given no-passing zone would be unmarked for about 7 days. If no-passing markings could reduce the expected number of all accidents in the no-passing zone by one-half, the value of those savings would be one-half of $188/mi-day times 7 days, or $658 per mile of no-passing zone. Because the cost of the temporary pavement marking would be about $1,122 per mile of no-passing zone, the benefit-cost ratio would be about 0.6. Missouri accident statistics indicate that improper passing is involved in only about 3 percent of accidents. This would seem to imply that the benefit-cost ratio is, at best, only 0.12. It appears obvious that marking no-passing markings during resurfacing projects cannot be justified by benefit-cost analysis unless the ADT is much greater than 2,000 veh/day or benefits other than accident reduction are considered.

TABLE 8 ACCIDENTS BY CONTRIBUTING CIRCUMSTANCES (R. Coplen)

| CONTRIBUTING CIRCUMSTANCES | DRIVERS INVOLVED IN FATAL ACCIDENTS | INJURY ACCIDENTS | PROPERTY DAMAGE ACCIDENTS | TOTAL ACCIDENTS |
|---|---|---|---|---|
| Speed, Exceeded Limit | 110 | 853 | 910 | 1,873 |
| Speed, Too Fast for Conditions | 164 | 4,591 | 7,433 | 12,188 |
| Failure to Yield Right-of-Way | 73 | 4,380 | 9,867 | 14,320 |
| Improper Passing | 23 | 571 | 1,742 | 2,336 |
| Violation, Electrical Signal | 19 | 302 | 460 | 781 |
| Violation, Stop Sign | 7 | 570 | 909 | 1,486 |
| Wrong Side (Not Pasing) | 172 | 1,446 | 1,491 | 3,109 |
| Following Too Closely | 12 | 2,723 | 6,886 | 9,621 |
| Directional Signal, Failed to or Wrong | 1 | 91 | 312 | 404 |
| Improper Backing | 1 | 50 | 639 | 690 |
| Improper Turn | 14 | 616 | 2,036 | 2,666 |
| Wrong Way on One Way | 11 | 49 | 91 | 151 |
| Improper Start, From Park | 0 | 46 | 219 | 265 |
| Improper Parking | 3 | 148 | 247 | 398 |
| Vehicle Defects | 20 | 836 | 2,051 | 2,907 |
| Drinking | 136 | 1,697 | 1,249 | 3,082 |
| Drugs | 6 | 72 | 63 | 141 |
| Other Violation | 0 | 0 | 0 | 0 |
| Inattention | 269 | 11,295 | 23,774 | 35,338 |
| None | 413 | 17,472 | 43,406 | 61,291 |
| Totals | 1,454 | 47,808 | 103,785 | 153,047 |
| For Drivers (Number = ) | 1,149 | 41,053 | 95,168 | 137,370 |
| In Accidents (Number = ) | 749 | 22,876 | 52,438 | 76,063 |

TABLE 9 ACCIDENT RATES AND COSTS

| ADT | Accidents per Mile per Year | Accidents per Mile per Day | Accident Cost per Mile per Day |
|---|---|---|---|
| 400 | 0.42 | 0.00114 | $ 38 |
| 800 | 0.83 | 0.00228 | 75 |
| 1,200 | 1.25 | 0.00342 | 113 |
| 1,600 | 1.66 | 0.00456 | 150 |
| 2,000 | 2.08 | 0.00570 | 188 |

## ADDITIONAL CONSIDERATIONS

### Driver Expectations

Drivers expect different driving characteristics from different types of roads. The principal function of rural local roads is to provide access to adjacent land. Travel distances are short and mobility is not a primary concern. Arterials are expected to provide for high speeds and high volumes. Many of the trips are long and there should be little interference for the through movements. On rural collectors, both land access and mobility are important. Trip lengths are longer than those on local roads but shorter than those on arterials. Speeds are generally higher than on local roads but lower than on arterials (*10*). Drivers expect higher mobility on arterials than on collectors and higher mobility on collectors than on local roads.

### Driver Information

Most resurfacing projects are conducted during the spring, summer, and fall, so most peak volumes will occur in daylight hours. The number of passes is roughly proportional to the square of hourly volume and the number of potential conflicts is roughly proportional to the cube of volume. Recent resurfacing and new 4-ft dashes will be obvious to most drivers during daylight hours. Therefore most drivers in the critical

TABLE 10 ACCIDENTS BY WEATHER CONDITIONS (R. Coplen)

| WEATHER CONDITION | FATAL ACCIDENTS | INJURY ACCIDENTS | PROPERTY DAMAGE ACCIDENTS | TOTAL ACCIDENTS |
|---|---|---|---|---|
| Clear | 509 | 14,573 | 30,329 | 45,411 |
| Cloudy | 166 | 4,933 | 10,402 | 15,501 |
| Rain | 42 | 2,225 | 4,689 | 6,956 |
| Snow | 11 | 421 | 1,180 | 1,612 |
| Sleet | 4 | 82 | 182 | 268 |
| Freezing | 5 | 238 | 508 | 751 |
| Fog/Mist | 12 | 260 | 558 | 830 |
| Other | 0 | 0 | 0 | 0 |
| Not Stated | 0 | 144 | 4,590 | 4,734 |
| Totals | 749 | 22,876 | 52,438 | 76,063 |

TABLE 11 ACCIDENTS BY LIGHT CONDITIONS (R. Coplen)

| LIGHT CONDITION | FATAL ACCIDENTS | TOTAL FATALITIES | INJURY ACCIDENTS | TOTAL INJURIES | PROPERTY DAMAGE ACCIDENTS | TOTAL ACCIDENTS |
|---|---|---|---|---|---|---|
| Daylight | 381 | 437 | 15,023 | 23,977 | 37,151 | 52,555 |
| Dark w/Streetlights on | 59 | 64 | 2,890 | 4,599 | 6,082 | 9,031 |
| Dark w/Streetlights Off | 1 | 1 | 146 | 211 | 329 | 476 |
| Dark - No Streetlights | 307 | 362 | 4,770 | 7,396 | 8,675 | 13,752 |
| Not Stated | 1 | 1 | 47 | 79 | 201 | 249 |
| Totals | 749 | 865 | 22,876 | 36,262 | 52,438 | 76,063 |

time periods will be aware of the recent resurfacing. In addition, the majority of drivers on local and collector roads are nearby residents who will use the road many times during the course of the resurfacing project. Therefore, many drivers may not expect permanent pavement markings to be present.

## CONCLUSIONS AND RECOMMENDATIONS

The primary considerations leading to the recommendations on marking no-passing zones during pavement resurfacing operations are as follows:

1. No-passing zones are most important to drivers with a high expectation of mobility. Drivers on arterial roadways expect provisions that enhance their perceived mobility.
2. For roads in level terrain with relatively few sight distance limitations, the number of passes and potential conflicts do not present a significant safety hazard if the ADT is below 1,700 veh/day.
3. For a road in rolling terrain, sight distance limitations and the reduced speeds of some vehicles increase the hazard caused by passes and potential conflicts.
4. In mountainous terrain, with heavy vehicles operating at crawl speeds and relatively few sections of road with adequate passing sight distance, the hazard caused by the desire to pass and potential passing conflicts becomes significant at relatively low volumes.
5. The monetary value of reduced accidents that might result from marking no-passing zones does not justify the additional cost of no-passing zones on collector roads. However, drivers on collectors that have few sections with inadequate passing sight distance probably expect some positive guidance to support their decisions to pass or not to pass.

The recommendations are presented in Table 12.

If these recommendations are followed, traffic on roads with temporarily unmarked no-passing zones will operate at a high quality of flow. In the peak hour, few platoons of three or more vehicles will develop and the number of passes and potential passing conflicts will be reasonably low. Platooning, passes, and potential conflicts will be much lower in nonpeak hours.

The state rural system in Missouri is 70.4 percent local and 22.6 percent collector. Most state systems include fewer miles of low-volume roads. Decisions on pavement markings for rural local and collector roads in most states are more likely to be made at the county level.

TABLE 12 RECOMMENDATIONS ON MARKING NO-PASSING ZONES DURING RESURFACING PROJECTS

| Rural Road Classification | Recommendation |
|---|---|
| Arterial (ADT > 1,700) | Mark during project |
| Collector (1,700 > ADT > 400) | Mountainous terrain: mark during project<br>Rolling terrain: mark during project when ADT exceeds 1,000 veh/day<br>Level terrain: mark at end of project |
| Local (ADT < 400) | Mark at end of project |

## ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the Missouri Highway and Transportation Department in the conduct of this work.

## REFERENCES

1. Manual on Uniform Traffic Control Devices. FHWA, U.S. Department of Transportation, 1988.
2. J. Glennon. *NCHRP Report 214: Design and Traffic Control Guidelines for Low-Volume Rural Roads.* TRB, National Research Council, Washington, D.C., Oct. 1979, Appendix F.
3. J. L. Josey, F. A. Gerig, Jr., F. Kern, and D. J. Frankenfield. *Simulation of Passing on Rural Two-Lane Highways.* Study 71–8, Missouri State Highway Department, Jefferson City, 1971.
4. J. G. Wardrop. Some Theoretical Aspects of Road Traffic Research. *Proc. Institution of Civil Engineers*, Vol. 1, Part 2, London, June, 1952, pp. 333–334.
5. ROADSIM Microcomputer Version. FHWA-IP-86-15. FHWA, U.S. Department of Transportation, Nov. 1987.
6. *Special Report 209: Highway Capacity Manual*, TRB, National Research Council, Washington, D.C., 1985, Chapter 8.
7. T. M. Matson, W. S. Smith, and F. W. Hurd. *Traffic Engineering.* McGraw-Hill, New York, 1955.
8. T. R. Miller, S. Luchter, and C. P. Brinkman. Crash Costs and Safety Investment. *32nd Annual Proc., Association for the Advancement of Automotive Medicine*, Sept. 1988, pp. 69–88.
9. *Unit Bid Prices—1989.* Missouri Highway and Transportation Department, Jefferson City.
10. *A Policy on Geometric Design of Highways and Streets.* AASHTO, Washington, D.C. 1984, Chapter 1.

*The conclusions and recommendations presented herein are solely those of the authors.*

*Publication of this paper sponsored by Committee on Traffic Safety in Maintenance and Construction Operations.*

# Effect of Radar Transmissions on Traffic Operations at Highway Work Zones

GERALD L. ULLMAN

A series of field studies was conducted at highway work zones where low-output radar transmissions were emitted (by motion detection devices) without the presence of visible law enforcement. Data were collected on vehicle speeds upstream and within the work zones, on speed changes made by vehicles as they approached the work zones, and on vehicle conflicts occurring in the 1,500-ft approach to the work zones. The results indicated that radar signals had only a small effect on average speeds within the work zone and on the change in speeds by motorists as they approached the work zone. However, the radar signals did appear to have a slightly greater effect on vehicles approaching the work zone at speeds greater than 65 mph and on trucks. Such results appear plausible, given the likelihood of greater radar detector use among these types of driver. The vehicle conflict study performed on the approach to the work zones found that severe braking-vehicle conflicts may increase in the presence of radar signals. There was an indication that increases in vehicle conflicts at a given work zone may depend on the amount that the average speed in the work zone (without radar) exceeds the posted work zone speed limit.

Despite impressive improvements in work zone traffic control procedures during the past two decades, work zone safety continues to be a topic of major concern to highway agencies. One of the difficult issues that has not yet been fully resolved is that of speed control within work zones. Although it is generally recommended that work zones be designed so as not to require drivers to reduce their speeds, the unusual and dynamic characteristics of work zones sometimes necessitate slower travel. When the need for reduced speeds is readily perceived by drivers, it is believed that most make appropriate adjustments so as to maintain safe and reasonable travel (*1*). However, if the need for slower speeds is not readily apparent, drivers cannot be expected to reduce their speeds without some active form of speed control.

Research throughout the decade has focused on various techniques available to highway agencies for controlling speeds in work zones (*2–7*). Of those tested, law enforcement has consistently proven to be one of the most effective work zone speed control methods available. Reductions in average speeds of up to 13 mph have been found in some instances (*2*). This result is not surprising; other research has found enforcement to be effective in reducing speeds in special situations such as school zones as well as on normal highway sections (*8*).

Unfortunately, law enforcement in most jurisdictions is a costly speed control method. Perhaps more important, enforcement resources are limited and must be distributed among a number of activities (in addition to traffic control) to preserve public safety. As a result, highway agencies continue to search for methods of work zone speed control that are less costly and easier to implement than law enforcement.

Recently, attention has turned to the possible use of radar transmissions to reduce speeds. Past research indicates that radar has an additional speed-reducing effect when used in conjunction with law enforcement (*8*). More recently, a limited amount of research has been performed evaluating the effect of radar *without* visible enforcement present (*9,10*). These studies, conducted on sections of highway other than work zones, suggest that average speeds can be reduced slightly when radar signals are emitted. These studies also found radar to affect high-speed vehicles more significantly.

Radar transmissions have the potential for reducing speeds at work zones as well. Furthermore, they may also serve as an attention-getting device, increasing the awareness of drivers as they approach the work zone. However, the overall effect that radar signals have on safety at work zones must first be determined. Radar, unlike other forms of work zone speed control (including visible enforcement), does not present a speed-reducing stimulus to each driver approaching a work zone. Rather, only those vehicles using a radar detector will receive any type of signal. Conflicts may develop between vehicles with detectors (that may decelerate suddenly when a radar signal is received) and vehicles without detectors. This study was conducted to evaluate these and other possible effects of radar transmissions at work zones.

## STUDY DESCRIPTION

### Objectives

The objectives of this study were twofold:

1. Determine the effect of radar signals (without the presence of visible law enforcement) on vehicle speeds approaching and passing through work zones without visible enforcement present.
2. Determine what effects radar signals may have on vehicle maneuvers and interactions between vehicles as they approach the work zone.

These objectives were accomplished through field studies at a total of eight work zone locations in Texas.

### Study Approach

Prototype radar transmitters, constructed in a previous study of radar transmissions by the Texas Transportation Institute

Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843.

(*9*), were used during this study. These transmitters consisted of a microwave transmitter and battery installed in a small (2- by 4- by 8-in.) box. In the field, the unit was mounted to a sign, barrel, or railing at the beginning of the work zone. The unit was turned on and off by means of a small switch located on the top of the box. The transmitter itself was a standard motion detector that, when operating, emitted a traffic radar signal approximately 1,500 to 2,000 ft upstream, depending on geometric and environmental conditions.

A radar on–radar off analysis was used at each study site. Data were collected for a 30- to 45-min period without transmitting the radar signal. The transmitter was then turned on, and data were collected for another 30 to 45 min. This cycle was repeated throughout the day. The use of multiple time periods helped negate any effects differences in traffic volumes over the day at a given site may have had on speeds. Data collected while the transmitter was turned on were then compared with data collected with the radar off to determine what effect the presence of a radar signal had on traffic.

### Study Site Section

Vehicle speeds in work zones are affected by a multitude of factors. These factors include the typical geometric, traffic, and environmental elements that affect speeds on normal roadway sections (*8*), as well as the unique and dynamic features of the work zone itself (*11*). The effectiveness of speed control methods may be influenced by these and other factors as well. Principal factors considered in the study design and site selection included the following:

1. Roadway type (Interstate or multilane highway);
2. Traffic volumes (low, moderate, high);
3. Work zone lane closure present (yes, no); and
4. Work zone speed limit (none, 10 mph below normal, more than 10 mph below normal).

The studies were limited to Interstate or multilane highways to ensure that a suitable vehicle sample size was obtained. Also, radar detector use would be highest on these types of roadway. Because the response to a radar signal at a location would be directly related to the percentage of vehicles with radar detectors, focusing the study on these types of roadways provided an indication of the maximum effects to be expected from radar. Testing over a range of traffic volumes was desired to see if undesirable vehicle conflicts increased at higher volume levels because of the radar signals. It was desirable to examine the influence of work lane closures on the effectiveness of radar transmitters. A lane closure reduces the capacity of the roadway dramatically, whereas a work zone without a lane closure may have little or no effect on capacity. Finally, because the premise of a radar signal is the simulation of the presence of enforcement, the effect of radar is expected to depend on the normal and work zone speed limits posted and whether actual speeds are dramatically higher than the posted limit.

Unfortunately, it was not possible to evaluate the radar transmitter at enough sites to fill a complete factorial design. Likewise, the limited number and location of potential study sites precluded the use of an incomplete factorial design. Therefore, sites were selected and categorized according to the factors given earlier, and the data collection effort was designed to maximize the statistical strength of an individual evaluation.

Table 1 presents a summary of the characteristics of the study sites. Sites 1 and 2, located on a section of four-lane

TABLE 1 SUMMARY OF STUDY SITE CHARACTERISTICS

| Site | Road Type | No. of Lanes | Normal Speed Limit, mph | Average Operating Speed, mph | 1987 ADT | Approx. vphpl Observed | Type of Work Zone | Work Zone Speed Limit, mph |
|---|---|---|---|---|---|---|---|---|
| 1 | Suburban Divided Highway | 2 | 55 | 58.5 | 14,300 | 200 | Detour with Lane Closure | 40 (R) |
| 2 | Rural Divided Highway | 2 | 55 | 60.3 | 12,600 | 200 | Detour with Lane Closure | 40 (R) |
| 3 | Suburban Interstate | 2 | 55 | 61.3 | 22,000 | 300 | Temporary Lane Closure | none posted |
| 4 | Suburban Interstate | 2 | 55 | 59.3 | 22,000 | 250 | Temporary Lane Closure | none posted |
| 5 | Suburban Interstate | 2 | 65 | 59.2 | 51,000 | 650 | Work Adjacent to Roadway | 55 (R) |
| 6 | Suburban Interstate | 2 | 65 | 57.3 | 67,000 | 800 | Work Adjacent to Roadway | 55 (R) |
| 7 | Suburban Interstate | 3 | 55 | 53.9 | 163,000 | 1400 | Work Adjacent to Roadway | none posted |
| 8 | Suburban Interstate | 3 | 55 | 56.2 | 163,000 | 1250 | Work Adjacent to Roadway | none posted |

(R) = regulatory speed limits vphpl = vehicles per hour per lane mph = miles per hour ADT = average daily traffic

divided highway with low traffic volumes, involved in long-term work zone lane closure (using barrels) and detour onto adjacent frontage roads. A reduced work zone speed limit of 40 mph was posted at these sites. Sites 3 and 4 were located on a section of a suburban four-lane Interstate with moderate traffic volumes. The work zones at these sites involved the temporary closing of one traffic lane (using cones); however, no work zone speed limits reduced below the normal 55-mph speed limit were posted. Sites 5 through 8 were work zones also located on suburban sections of four- and six-lane Interstate highways. No long- or short-term lane closures were present at these sites, however. In addition, Sites 5 and 6 were posted with a speed limit of 55 mph, reduced from the normal 65-mph limit. The speed limits of 55 mph at Sites 7 and 8 were not reduced in the study section.

## Data Collection and Reduction

### *Vehicle Speeds and Speed Changes*

Researchers collected two types of data during the studies. Figure 1 shows the basic data collection layout at each study site. Vehicle speeds, measured by traffic radar detuned so as to be undetectable by radar detectors, were collected at three stations upstream and within the work zone. The first station, situated approximately 3,000 ft upstream of the work zone and determined to be beyond the influence of the work zone or the radar signal, was used as a control. The second station was located about 750 to 1,250 ft upstream of the beginning of the work zone. The radar transmitter, always installed at the beginning of the work zone, had a range of approximately 1,500 ft. Therefore, speeds measured at Station 2 represented conditions immediately after those vehicles with radar detectors were first able to receive a signal. Because the quality and capabilities of radar detectors vary from model to model, some variation in the exact location individual drivers first received the signal was likely. The third station was positioned within the work zone immediately beyond the radar transmitter location.

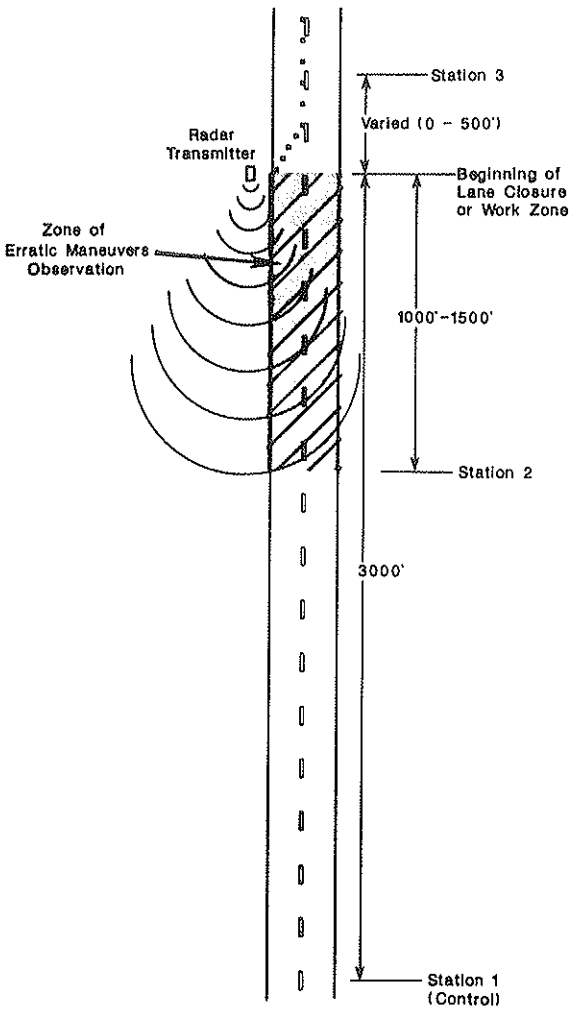



**FIGURE 1 Layout of data collection plan.**

At each station, data collection personnel recorded the speed of vehicles along with a description of the vehicle on a cassette recorder. This procedure allowed vehicles to be tracked through the study section so that changes in speed from Station 1 to Stations 2 and 3 could be examined. This approach provided a strong statistical design for evaluation.

Over 20,000 speed observations were collected at the eight study sites. Consolidated over all sites, approximately 60 percent of the vehicles recorded at Station 1 were tracked to Station 2, and 49 percent of vehicles at Station 1 were tracked to Station 3. On a site-by-site basis, these percentages were much greater for Sites 1 through 4, where traffic volumes were lower.

### *Vehicle Conflicts*

The second type of data collected at each site was vehicle conflicts occurring within the 1,000- to 1,500-ft approach to the work zone. Traffic volumes were collected simultaneously to develop vehicle conflict rates for comparison purposes. Conflicts occurred in isolation (e.g., a single vehicle braking severely) and also because of vehicle interactions (e.g., vehicles behind a hard-braking vehicle were forced to swerve out of the lane or to also brake severely), and an attempt was made by the observer at each site to document the type of conflicts that occurred. However, the frequency of conflicts was not sufficient to maintain this distinction during analysis. Therefore, vehicle conflicts were categorized into four main types: (a) severe braking, identified by a dramatic nosedive or skidding by the vehicle; (b) abrupt last-second lane-changing; (c) accelerating into the work zone at high speeds to get around one or more vehicles before the lane closure or to exit at a downstream ramp; and (d) other vehicle conflicts (stopping on road, running off the road, etc.).

## RESULTS

### Effect of Radar on Vehicle Speeds

Table 2 presents a comparison of the average and standard deviation of speeds measured in the work zone (at Station 3) with and without a radar signal transmitted. In general, the effect of radar was fairly consistent, albeit slight. Average

TABLE 2 EFFECT OF RADAR ON SPEED CHARACTERISTICS WITHIN THE WORK ZONE

| Site | Average Speed, Mph | | | Standard Deviation, Mph | | |
|---|---|---|---|---|---|---|
| | Radar Off | Radar On | Diff. | Radar Off | Radar On | Diff. |
| 1 | [186] 46.3 | [203] 45.7 | -0.6 | 6.04 | 6.91 | +0.87 |
| 2 | [395] 54.7 | [543] 53.1 | -1.6* | 6.56 | 6.55 | -0.01 |
| 3 | [339] 54.5 | [484] 53.7 | -0.8* | 5.49 | 5.50 | +0.01 |
| 4 | [295] 56.8 | [356] 56.5 | -0.3 | 5.08 | 5.33 | +0.25 |
| 5 | [726] 55.4 | [514] 54.9 | -0.5 | 4.50 | 5.22 | +0.72* |
| 6 | [483] 54.1 | [493] 53.9 | -0.2 | 4.69 | 4.49 | -0.20 |
| 7 | [723] 52.7 | [409] 52.7 | 0.0 | 5.70 | 5.89 | +0.19 |
| 8 | [184] 53.1 | [194] 52.8 | -0.3 | 5.26 | 4.85 | -0.41 |

*Statistically significant (0.05 level of significance)

Numbers in brackets [ ] are the sample size at each site

speeds were slightly lower (0.2 to 1.6 mph) at seven of the eight sites when the radar was transmitting, but the change in average speed was statistically significant at only two sites. Statistical significance was measured by a *t*-test of the comparison of means (*12*). Because of the large sample sizes available at each site, however, this test approximated a standard *z*-test comparison of sample means with known variances.

Meanwhile, the standard deviation of speeds increased slightly at seven sites, although only one of these changes was found to be statistically significant. The changes are reported as differences between the radar off and radar on conditions (to illustrate the absolute magnitude of the changes observed). However, an *f*-test of the ratio of the variance estimates was used to detect statistical significance. This statistic was compared with a critical *f*-value with 0.05 level of significance and number of degrees of freedom equal to sample size with radar on and sample size with radar off.

Table 3 presents a comparison of how motorists adjusted their speeds between data collection Stations 1 (control) and 3 (within the work zone) for vehicles that could be tracked through the study site. Such a paired comparison increases the statistical strength of the analysis, providing stronger evidence about the actual effect of the transmitter on vehicle speeds. These changes in speeds between stations were averaged for the radar off and radar on conditions and compared using a paired *t*-test (*12*). These data indicate that the effect of radar was somewhat greater than that suggested in Table 2. The average change in speed between the stations was negative, indicating that speeds decreased as vehicles approached the work zone (as would be expected). The difference in the average speed changes, representing the effect of radar, indicated that an additional 0.2- to 4.5-mph reduction occurred when the radar was transmitting. On the basis of this analysis, the difference in speed changes was statistically significant at four of the eight sites.

The speed changes made by drivers as they approached the work zone appeared to be slightly more variable when the radar was in operation. The standard deviation of the changes in speed between Stations 1 and 3 increased slightly at every site. This increased variability might have been caused by those drivers with radar detectors slowing their vehicles much more dramatically than drivers without detectors. Unfortunately, this hypothesis could not be proven in this study. An *f*-test of the estimates of sample variances was again used to test the statistical significance of the changes between the radar on and radar off conditions.

The analysis of speeds taken 1,000 to 1,500 ft from the work zone (Station 2) did not exhibit as consistent trends as were evident at Station 3. Apparently, motorists with radar detectors had just received the radar signal and had not yet adjusted

TABLE 3 EFFECT OF RADAR ON CHANGES IN SPEED BETWEEN STATIONS 1 AND 3

| Site | Average Change in Speed Between Stations, Mph | | | Standard Deviation of Speed Changes Between Stations, Mph | | |
|---|---|---|---|---|---|---|
| | Radar Off | Radar On | Diff. | Radar Off | Radar On | Diff. |
| 1 | [186] -12.1 | [203] -12.5 | -0.4 | 6.82 | 7.83 | +1.01 |
| 2 | [395] -3.9 | [541] -5.4 | -1.5* | 6.03 | 6.64 | +0.61 |
| 3 | [279] -5.4 | [424] -6.0 | -0.6 | 6.00 | 6.50 | +0.50 |
| 4 | [294] -4.5 | [355] -5.3 | -0.8* | 6.21 | 7.01 | +0.80* |
| 5 | [461] -2.9 | [160] -3.1 | -0.2 | 4.72 | 4.80 | +0.08 |
| 6 | [341] -3.0 | [351] -3.6 | -0.6 | 4.72 | 4.88 | +0.16 |
| 7 | [162] -1.1 | [79] -2.8 | -1.7* | 5.80 | 6.30 | +0.50 |
| 8 | [7] -0.4 | [23] -4.9 | -4.5* | 2.37 | 4.39 | +2.02 |

Statistically Significant (0.05 level of significance)

Numbers in brackets [ ] are the sample size at each site

their speeds. As a result, no clear trends were evident, and so data from that station are not presented here. This information can be found in the study documentation (*13*).

It is generally recognized that the primary use of radar detectors is to avoid ticketing by law enforcement for exceeding the posted speed limit (*9*). Therefore, radar detectors could be assumed to be in more prevalent use on vehicles traveling at higher speeds. Also, the effect of the transmitter on these high-speed vehicles would be expected to be more pronounced. Figure 2 shows a comparison of speed changes between Stations 1 and 3 that supports this hypothesis. The effect of radar on average speed changes between Stations 1 and 3 is shown for the entire speed sample taken at each site and for the portion of the sample that exceeded 65 mph at Station 1 (the control station). As the figure shows, radar generally had a larger speed-reducing effect on those vehicles that were initially exceeding 65 mph as compared to the sample as a whole. At Sites 1 through 4, the effect of radar was from 1 to 3 mph greater for the portion of traffic exceeding 65 mph than it was for the entire sample size overall.

The influence on high-speed vehicles is less pronounced at the other sites, although a small difference is still evident. Because of congestion and data collection problems, no vehicles at Site 8 were observed to have exceeded 65 mph at

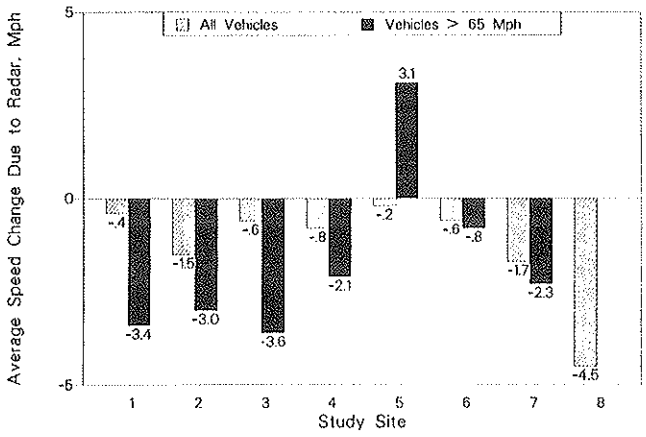



**FIGURE 2 Comparison of radar effect on all vehicles and vehicles exceeding 65 mph.**

Station 1. Also, the data from Site 5 suggest that radar had less effect on high-speed vehicles than on the entire vehicle population. Sporadic police presence was noted by the data collection personnel throughout the day at this site, which may explain these inconsistent findings. Also, the results at

Site 5 may have been caused by some other extraneous factor not accounted for in the analysis.

Pigman et al. (*10*) found that radar detector use is generally greater among trucks than among automobiles. Consequently, the effect of radar would also be expected to have a more pronounced effect on trucks than on automobiles. Figure 3 shows the difference between automobiles and trucks in terms of the effect of the transmitter on the changes in speed between Stations 1 and 3. Generally, the transmitter had a more pronounced effect on trucks than on automobiles, although this result was not the case at Sites 1 and 5. Again, site-specific factors unaccounted for in the analysis likely were the cause of the differing results at these locations.

### Effect of Radar on Vehicle Conflicts

Vehicles conflicts were recorded manually at each site during each study. Traffic volumes were recorded simultaneously so that conflict rates could be computed for comparison purposes (with and without the transmitter in operation). As stated previously, vehicle conflicts were categorized into three main types:

1. Severe braking (evidenced by a dramatic nosedive by a vehicle or by vehicle skidding),
2. Last-second or abrupt lane-changing, and
3. Accelerations into work zone (to pass a vehicle before reaching the lane closure or exit ramp).

A final category was simply labeled "other" to include any other maneuvers considered by data collection personnel to have resulted in conflict.

Results of the conflict analysis are presented in Tables 4 and 5. The category of other was used so infrequently that it was not included in this analysis (only three maneuvers total from all eight sites). A Poisson analysis was used to determine if the changes observed were statistically significant (*14*). Analyses were performed for each type of conflict at each site, for all three types of conflicts combined at each site, and for each type of conflict for all sites combined.

Vehicle conflict rates varied significantly from site to site, presumably because of the differences in volumes, work zone

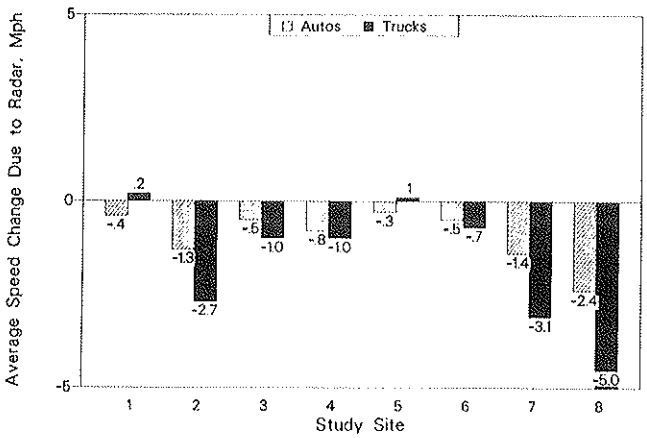



**FIGURE 3 Comparison of radar effect on automobiles and trucks.**

activity and traffic control, roadway geometrics, etc. At seven of the eight sites, severe braking conflict rates were higher when the radar was transmitting. The increases were statistically significant at two sites. Although an attempt was made to determine whether these maneuvers occurred in isolation or because of another vehicle, the small sample sizes at some of the sites and extremely high volumes at the other sites made this impossible to accomplish.

Conversely, last-second lane changes and vehicle accelerations into the work zone did not appear to be significantly affected by the radar transmissions. None of these changes was found to be statistically significant on a site-by-site basis.

For summary purposes, Tables 4 and 5 also present vehicle conflict rates averaged over all sites. Severe braking maneuvers increased from 21.7 to 26.6 conflicts per 1,000 vehicles, a statistically significant increase of 22.6 percent. Overall, last-second lane changing and accelerations into the work zone decreased slightly when the radar was transmitting (4.6 and 5.3 percent, respectively), but neither change was statistically significant. Consolidating all types of vehicle conflicts observed at all sites, the conflict rate increased by 7.9 percent, increasing from 47.1 conflicts per 1,000 vehicles without radar to 50.8 conflicts per 1,000 vehicles with the radar transmitting. The increase in total conflicts was also not found to be statistically significant.

### Discussion

The results just described indicate that a radar signal has some effect on speeds at work zones. However, these effects are small, generally less than 2 to 3 mph. Because of these small changes, there is no way of discerning, either statistically or through engineering judgment, how the site specific factors (traffic volume, work zone type, work zone speed limit) considered in this study influence the effectiveness of radar transmissions at work zones.

Overall, the small reductions in average speeds found in this study are consistent with those obtained in other studies of unmanned radar (at nonwork zone locations). In most situations, radar will not reduce overall speeds in a dramatic way. Radar does affect the behavior of drivers using radar detectors who are exceeding the posted speed limit by large amounts.

Interestingly, the results of this study are not consistent with those of other studies with respect to the effect of radar on speed variability. Although the past studies found speed variance lower when radar was transmitting, this study suggests that the variability of speeds at locations within the work zone may actually increase in some cases. The results of the comparison of speed changes between stations suggest that the variability of these speed changes increases when radar is present. Presumably, this increased variability is caused by drivers with detectors who decelerate dramatically on receiving a signal from their detector. Evidence collected during this study suggests that these drivers tend to be the high-speed motorists, and the transmitter appears to have a more pronounced speed-reducing effect on them.

The vehicle conflict data indicate that the presence of a radar signal increases the frequency of severe braking maneuvers. This increase is expected to be related to the work zone speed limit posted and the actual driving speeds at the

TABLE 4 VEHICLE CONFLICT DATA (RATE PER 1,000 veh)

| Site | Severe Braking | | Last-Second Lane-Changing | | Accelerating into Work Zone | | Total | |
|---|---|---|---|---|---|---|---|---|
| | Radar Off | Radar On | Radar Off | Radar On | Radar Off | Radar On | Radar Off | Radar On |
| 1 | [7] 11.9 | [9] 18.3 | [2] 3.4 | [7] 14.2 | [9] 15.3 | [7] 14.3 | [18] 30.6 | [23] 46.8 |
| 2 | [16] 27.8 | [26] 47.8 | [4] 6.9 | [11] 20.2 | [5] 8.7 | [2] 3.7 | [25] 43.4 | [39] 71.7 |
| 3 | [2] 2.5 | [3] 4.9 | [16] 20.3 | [14] 22.9 | [14] 17.8 | [11] 14.9 | [32] 40.6 | [28] 42.7 |
| 4 | [18] 29.9 | [30] 40.6 | [1] 1.6 | [3] 4.1 | [4] 6.6 | [4] 5.4 | [23] 38.1 | [37] 50.1 |
| 5 | [11] 4.6 | [5] 2.7 | [23] 9.5 | [16] 8.5 | [9] 3.7 | [13] 6.9 | [43] 17.8 | [34] 18.1 |
| 6 | [16] 8.5 | [36] 12.1 | [13] 6.9 | [27] 9.1 | [1] 0.5 | [5] 1.7 | [30] 15.9 | [68] 22.9 |
| 7 | [45] 15.5 | [78] 18.7 | [45] 15.5 | [51] 12.2 | [0] 0.0 | [0] 0.0 | [90] 31.0 | [129] 30.9 |
| 8 | [207] 40.2 | [238] 51.6 | [196] 48.1 | [181] 49.3 | [0] 0.0 | [2] 0.4 | [403] 88.3 | [421] 101.3 |
| Total | 21.7 | 26.6 | 21.6 | 20.6 | 3.8 | 3.6 | 47.1 | 50.8 |

* Statistically significant (0.05 level of significance)

Numbers in brackets [ ] are the sample size at each site

work zone. Specifically, work zone sites at which drivers normally travel much faster than the posted work zone speed limit would be expected to have higher vehicle conflict rates in the presence of radar, as drivers with detectors try to slow down quickly to comply with the posted limit, and any vehicles following are forced to respond in a similar fashion.

To examine this hypothesis, the relationship between the percentage increase in total vehicle conflicts at each site was plotted against the difference in the average speed in the work zone (without radar) and the posted speed limit in the work zone. This relationship is shown in Figure 4. Clearly, a trend towards larger increases in conflicts exists at sites where average speeds are much higher than the speed limit posted in the work zone. Although the actual relationship between accidents and vehicle conflicts at work zones is not known, these data suggest a potential safety problem with the use of radar at sites where the posted speed limit is considerably lower than the normal speed of traffic.

## CONCLUSIONS

This study examined the effect of radar transmissions (without visible enforcement) on vehicle speeds and vehicle conflicts at eight work zone locations on multilane roadways in Texas. The work zones varied with respect to the amount of traffic present, type of work zone (with or without a lane closure), and the reduction in normal speed limits through the work zone. Overall, the effect on speeds was small, as average speeds at the study sites were generally reduced by less than 2 mph. From the analysis of the changes in speeds between data collection stations as vehicles approach the work zone, there may be, in some cases, a greater effect of radar on trucks (in comparison with automobiles) and on high-speed vehicles (in comparison with the entire vehicle sample). Such results correlate well with expectations that radar detector use may be more prevalent among trucks and among high-speed vehicles.

TABLE 5 EFFECT OF RADAR ON VEHICLE CONFLICTS (RATE PER 1,000 veh)

| Site | Change in Severe Braking Conflicts | Change in Last-Second Lane-Changing Conflicts | Change in Accelerating into Work Zone Conflicts | Change in Total Conflicts |
|---|---|---|---|---|
| 1 | +6.4 (+53.8%) | +10.8 (+317.7%) | -1.0 (-6.5%) | +16.2 (+52.9%) |
| 2 | +20.0* (+71.9%) | +13.3 (+192.8%) | -5.0 (-57.5%) | +28.3* (+65.2%) |
| 3 | +2.4 (+96.0%) | +2.6 (+12.8%) | -2.9 (-16.3%) | +2.1 (+5.2%) |
| 4 | +11.7 (+39.1%) | +2.5 (+156.3%) | -1.2 -(18.2%) | +12.0 +(31.5%) |
| 5 | -1.9 (-41.3%) | -1.0 (-10.5%) | +3.2 (+86.5%) | +0.3 (+1.7%) |
| 6 | +3.6 (+42.4%) | +2.2 (+31.9%) | +1.2 (+240.0%) | +7.0* (+44.0%) |
| 7 | +3.2 (+20.7%) | -3.3 (-21.3%) | 0.0 (---) | -0.1 (-0.3%) |
| 8 | +11.4* (+28.4%) | +1.2 (+2.5%) | +0.4 (---) | +13.0 (+14.7%) |
| Total | +4.9* (+22.6%) | -1.0 (-4.6%) | -0.2 (-5.3%) | +3.7 (+7.9%) |

* Statistically significant (0.05 level of significance)

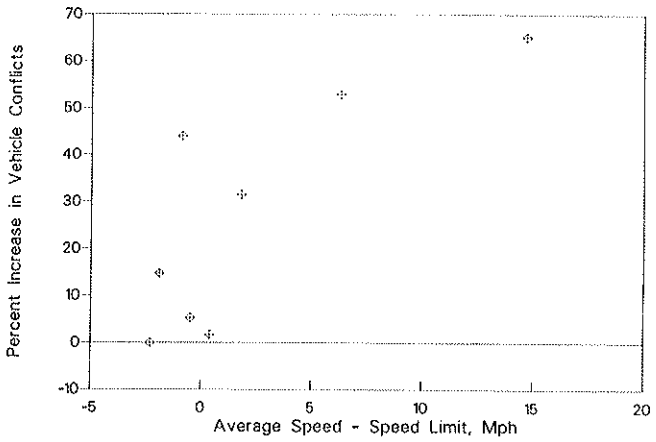



**FIGURE 4 Change in vehicle conflicts versus difference in average speed and posted speed limit.**

Severely braking vehicle conflicts increased significantly at two of the eight sites when radar was operating in comparison to when radar was not in operation. Increases (although not statistically significant) were also observed at the other six sites. The increases were larger for the four sites where lane closures were present and were the highest at the two sites where average speeds were normally much higher than the posted limit. Overall, it appears that the use of radar may result in additional conflicts on the approach to the work zones, particularly if the posted speed limit is considerably lower than that at which drivers wish to travel.

## ACKNOWLEDGMENT

This research was funded as part of an HP&R study sponsored by the Texas State Department of Highways and Public Transportation (SDHPT) and the FHWA.

## REFERENCES

1. *Procedure for Establishing Speed Zones*. Texas State Department of Highways and Public Transportation, Austin, 1985.
2. S. H. Richards, R. C. Wunderlich, and C. L. Dudek. *Controlling Speeds in Highway Work Zones*. Research Report FHWA/TX-84/58+292.2. Texas Transportation Institute, College Station, Feb. 1984.
3. S. H. Richards and C. L. Dudek. Implementation of Work Zone Speed Control Measures. *Transportation Research Record 1086*,

TRB, National Research Council, Washington, D.C., 1986, pp. 36–42.
4. H. W. McGee, D. B. Joost, and E. C. Noel. Speed Control at Work Zones. *ITE Journal*, Vol. 58, No. 1, Jan. 1988, pp. 17–19.
5. Z. A. Nemeth and A. K. Rathi. Potential Impact of Speed Reduction at Freeway Lane Closures: A Simulation Study. In *Transportation Research Record 1035*, TRB, National Research Council, Washington, D.C., 1985, pp. 82–84.
6. J. R. Jarvis. The Effectiveness of Road Work Speed Limit Signs. *Australian Road Research*, Vol. 13, No. 3, Sept. 1983, pp. 185–194.
7. J. Jackels and D. Brannan. *Work Zone Speed Limit Demonstration in District 1A*. Minnesota Department of Transportation, St. Paul, Oct. 1988.
8. D. L. Warren. Chapter 17—Speed Zoning and Control. *Synthesis of Safety Research Related to Traffic Control and Roadway Elements—Volume 2*. Report FHWA-TS-82-233. FHWA, U.S. Department of Transportation, Dec. 1982.
9. V. J. Pezoldt. *The Influence of Radar Detectors on Texas Highway Traffic Speeds*. Texas Transportation Institute, College Station, Sept. 1987.
10. J. G. Pigman, K. R. Agent, J. A. Deacon, and R. J. Kryscio. Evaluation of Unmanned Radar Installations. In *Transportation Research Record 1244*. TRB, National Research Council, Washington, D.C., 1989, pp. 7–16.
11. N. M. Rouphail and G. Tiwari. Flow Characteristics at Freeway Lane Closures. In *Transportation Research Record 1035*, TRB, National Research Council, Washington, D.C., 1985, pp. 50–58.
12. R. E. Walpole and R. H. Myers. *Probability and Statistics for Engineers and Scientists*. MacMillan, New York, 1978.
13. G. L. Ullman and D. R. Riesland. *Catalog of Work Zone Speed Control Methods*. Report FHWA/TX-90/1161-2. Texas Transportation Institute, College Station, May 1990.
14. P. C. Box and J. C. Oppenlander. *Manual of Traffic Engineering Studies*. Institute of Transportation Engineers, Washington, D.C., 1976.



*Publication of this paper sponsored by Committee on Traffic Safety in Maintenance and Construction Operations.*

# Evaluation of Flagger Training Session on Speed Control in Rural Interstate Construction Zones

RAHIM F. BENEKOHAL AND LYNN M. KASTEL

A study was conducted on the evaluation of a flagger training session on speed control in rural Interstate construction zones. Two flaggers were randomly selected to participate in a training session, in which they were taught the *Manual on Uniform Traffic Control Devices* procedures for flagging on rural Interstates. Speeds of traffic outside of the construction zone, inside of the zone but far from the flagger, and inside the zone near the flagger were observed to determine the effectiveness of the flaggers before and after training and to note any increased effectiveness after training. The findings indicate that both cars and trucks have speeds exceeding the speed limit outside the construction zone and far from the flagger, but speeds near the flagger are lower than the speed limit. For both flaggers, the average speeds near the flagger were 4 to 9 mph lower after training than before. For the first flagger, it is not clear whether or not this is a result of an increased effectiveness due to training. However, for the second flagger, it is likely that the reduced speeds are caused by an increased effectiveness after training.

Flaggers are provided to assist moving vehicles and pedestrians safely and efficiently through or around work areas while protecting on-site workers and equipment, either by stopping traffic intermittently or by maintaining continuous traffic past a worksite at reduced speeds to help protect the work crew (*1*). Lack of uniformity in flagging procedures in work zones prompted the question of whether or not contractors' flaggers are flagging according to procedures outlined in the *Manual on Uniform Traffic Control Devices* (MUTCD) (*1*). They may or may not be familiar with the procedure and they may not have learned it through formal training. Training the flaggers to flag according to MUTCD procedures may improve their effectiveness in controlling traffic and increase the safety of workers and motorists. Consequently, a pilot study was performed in which the effectiveness of flaggers in maintaining traffic reducing traffic speeds was evaluated in before and after training conditions. The objectives of evaluating trained flaggers were as follows:

- Determine the effectiveness of the flaggers before training,
- Determine the effectiveness of the flaggers after training, and
- Compare effectiveness in before and after training and suggest possible improvements that may increase the effectiveness of flaggers.

R. F. Benekohal, University of Illinois at Urbana-Champaign, 205 North Mathews Avenue, Urbana, Ill. 61801-2397. L. M. Kastel, Bucher Willis Ratliff Consulting Engineers, 7920 Ward Parkway, Kansas City, Mo. 62114

Previous studies by Richards et al. (*2,3*) reported flagger effectiveness when the flagger was located before the beginning of the taper area. However, because most flaggers are located in the construction zone near the crew, this study evaluated the effectiveness of the flagger when in the zone, rather than in front of the zone. This study considers only the case when the purpose of the flagger is to maintain continuous traffic past the worksite at reduced speeds. Because flaggers are required on this construction project, this study evaluated only the conditions of trained and untrained flaggers, rather than the condition of having no flaggers present. The condition of having no flaggers present would have created substandard construction procedures.

In the study, a contractor was asked to identify and send a few of his regular flaggers to a training session. The contractor identified two male workers who work for him as flaggers. The two flaggers were observed during the study. The flaggers were not told that they were selected to go to the training session until an hour before the session began. The reason that they were not informed about their participation in the program is that they might have performed differently if they knew that they were being trained and that their performance was being monitored.

Flagger effectiveness data were collected for each flagger, both before and after training. The data were collected using a traffic counter and two teams equipped with radars. One measure of effectiveness used to evaluate flaggers' performance was speed of vehicles. Other measures of effectiveness were evaluated (*4*) but because of space limitations, only the evaluation based on vehicle speed is presented.

According to the MUTCD, a flagger should possess the following minimum qualifications: average intelligence, good physical condition, including sight and hearing, mental alertness, courteous but firm manner, neat appearance, and a sense of responsibility for safety of public and crew. The flagger is required to use orange clothing, such as a vest, shirt, or jacket. The flagger must be clearly visible to the approaching traffic for a distance sufficient to permit proper response by speed before entering the worksite. The manual suggests that the flagger stand 200 to 300 ft from the crew. The flagger should stand either on the shoulder adjacent to the traffic being controlled, or in the barricaded lane, and should not stand in the lane being used by moving traffic under any circumstances (*1*).

## BACKGROUND

Flagging is used for traffic control and speed reduction in various types of work zones. Because flaggers are responsible for the safety of the crew and drivers, it is important that flaggers be effective in gaining the attention of drivers and reducing the speed of vehicles. Consequently, numerous studies have been made in order to determine the most appropriate manner of increasing the flaggers' effectiveness.

Presently, proper flagging procedures are given in the MUTCD (*1*). However, alternatives to flagging have been considered for study. For example, Ullman et al (*5,6*) investigated alternatives to using flaggers, such as the use of Yield to Oncoming Traffic signs to self-regulate traffic operations, and the use of portable traffic signals. For one-way operations, Bookers et al. (*7*) evaluated the effectiveness of using a reusable temporary stop bar (stop line) and a freestanding oversized Stop/Slow sign paddle to enhance the safety of the flagger.

Several studies have been performed in which variations on the MUTCD flagging procedures were evaluated. Noel et al. (*8,9*) studied the long-term (2 weeks) and short-term (within 3 days) effectiveness of implementing MUTCD flagging and innovative flagging on speed reduction on multilane freeways. Richards et al. (*2,3*) evaluated the immediate effects of MUTCD flagging and innovative flagging on one or both sides of the traveled lane, both of which were supplemented by advisory or regulatory signs.

It is not known if flaggers are following the procedures outlined in the MUTCD. As stated earlier, flaggers may not be familiar with the procedures, or they may not have learned them through formal training. This study is being performed to determine if there is an increased effectiveness of the flagger in slowing down vehicles after the flagger receives proper training on flagging according to MUTCD.

## TRAINING SESSION

The training session was conducted by the work zone manager of the Illinois Department of Transportation (IDOT), who had years of experience in training and educating workers on work zone traffic control. The training was conducted outside of the IDOT field office of the construction project site, where the trainer met with the flaggers and informally discussed proper flagging techniques for slowing down vehicles on multilane highways when one lane is closed. The purpose of the training session was to demonstrate the proper use of the Stop/Slow paddle and proper flagging procedures used with this kind of lane closure project. The trainer did not discuss other flagging situations. The flaggers were able to ask questions as they arose. The actual demonstration lasted approximately 15 min. The description of the training session is based on the researchers observations, and does not necessarily reflect the views of the research team.

### Key Points of the Training Session

The information taught at the training session was taken from the Illinois Department of Transportation Flaggers' Handbook and was consistent with the flagging techniques described in the MUTCD. The training session reviewed the proper equipment needed, the correct standing position, and the necessary hand gestures needed to slow traffic. Both flaggers were given copies of the Flaggers' Handbook.

#### *Equipment*

The flaggers were told that they should wear a high-visibility vest so that they could obtain the attention of the driver. It was suggested that the flagger wear bright orange gloves to help get the driver's attention and make the hand gestures more visible. The trainer recommended using a high-visibility body suit, such as the type that hunters wear, but said that this is not required. The trainer gave a pair of the bright orange gloves to Flagger 2. A standard Stop/Slow paddle should be used to help direct traffic. In Illinois, the minimum required height of the paddle is 6 ft, and the sign should be 24 in. wide.

#### *Position*

The flaggers were told that they should stand 200 to 300 ft ahead of the construction crew in the closed lane where they would be visible to the oncoming traffic.

#### *Hand Gesture*

The trainer demonstrated the proper hand gestures that should be used when flagging. The Stop/Slow paddle should be held with one hand in an upright position with the Slow sign visible to traffic, and the flagger should use his other hand to direct the traffic. The free hand should be held directly in front of the flagger with the palm outward and continuous sweeping motions downward should be used. When a driver fails to respond to the flagger, the flagger should point to the driver to get his attention and then either use the sweeping motion again or point to the Slow sign.

### Comments Made by the Flaggers After the Training Session

After the training session, the flaggers were asked questions pertaining to their perception of the training session. Their comments are given in later sections. Both flaggers felt that what the trainer had told them about flagging techniques was not anything new or unfamiliar to them. Both flaggers thought that it was necessary for them to stand close to the construction crew, rather than stand 200 to 300 ft away from them, because the traffic would move from the shoulder back into the normal driving position in the lane immediately after passing the flagger and would be too close to the crew for safety. When asked what type of treatment would be the most effective in reducing the speeds of traffic, the flaggers replied that police presence with radar would be most effective. They felt that improving flagging techniques was not effective because drivers were already able to tell when the construction crew was present and that they needed to slow down, but that

they were unwilling to do so. The fear of getting a speeding ticket was more effective in slowing down traffic. When asked if they felt any more, the same, or less conscientious about the importance of flagging and the proper techniques to use, they both said that they felt the same amount of conscientiousness about flagging after training as they had felt before the training session.

## SITE DESCRIPTION

The study site is located on the northbound lanes of a rural section of Interstate 57 just south of Champaign, Illinois. A typical section of the construction zone is shown in Figure 1. The construction zone, which was approximately 5 mi long, began shortly after the interchange of I-57 and US-45 and ended before the interchange of I-57 and I-72. The terrain in the study site is level, and the roadway does not have sharp curves. Along the construction zone, Interstate 57 is four-lane highway, with two lanes in each direction. However, in each direction, one of the lanes was closed because of the construction work. The flagger treatment was studied on the northbound lane, where the type of construction occurring was joint repair and surface preparation for an asphalt overlay. The speed limit was 65 mph for cars and 55 mph for large trucks, unless specified by speed limit signs with flashing

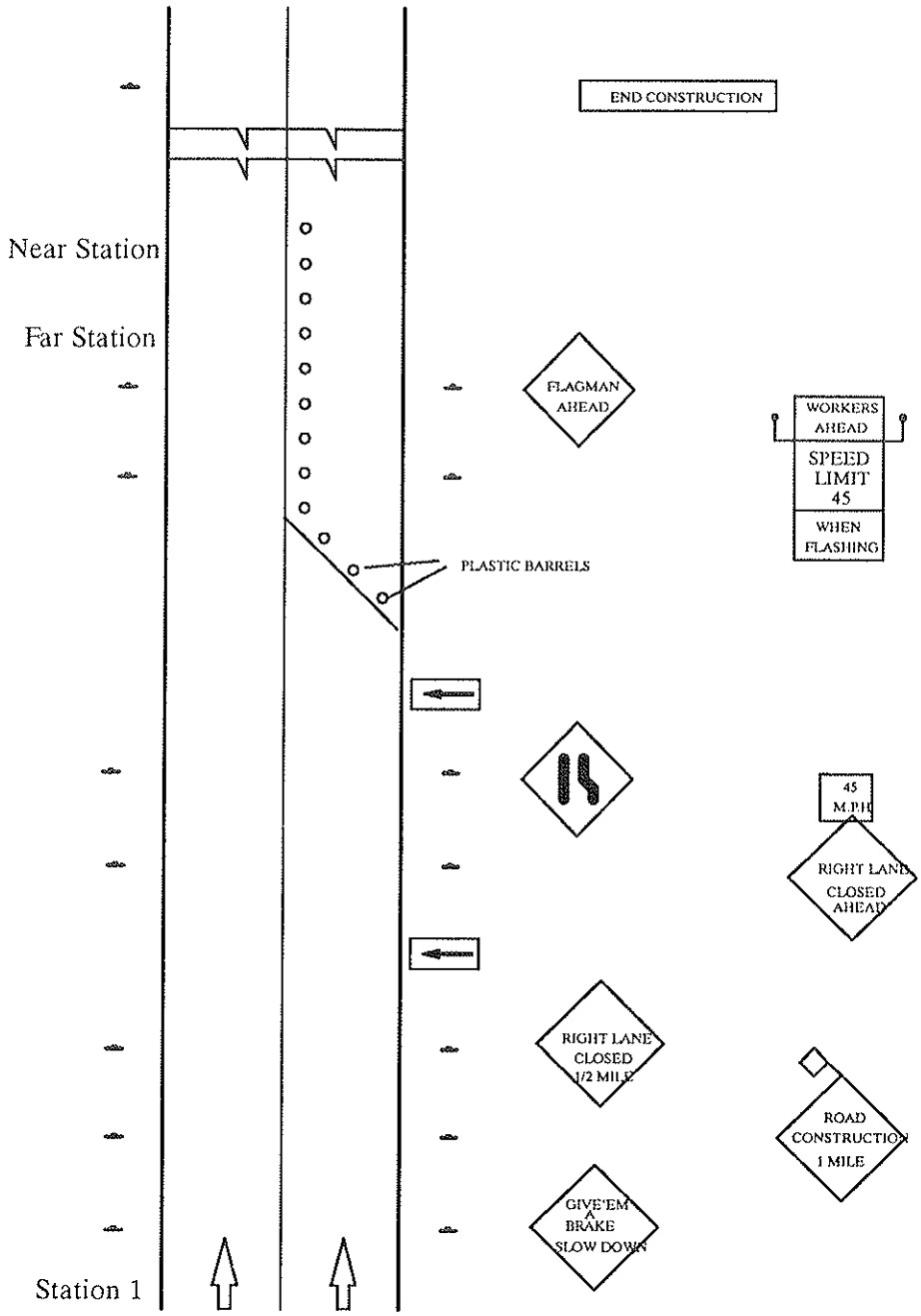



**FIGURE 1 Traffic control plan in the construction zone.**

lights mounted on them. When the flashing lights were turned on, the speed limit was 45 mph for all vehicles.

The traffic volume in the construction zone was light, being approximately 400 veh/hr. In the average daily traffic (ADT) maps published by the Illinois Department of Transportation, the ADT in the zone is 14,500 veh/day with approximately 17 percent large trucks.

## DATA COLLECTION

In order to determine the effectiveness of the flaggers, free flow traffic (not platooned) data were collected at three locations: approximately 1 mi before the construction zone (Station 1), far from the flagger but within the construction zone (far station), and near the flagger (near station). Speed data at Station 1 were collected using a traffic counter and at the far and near stations using radar guns. The speed data at each station are all for the same time period and indicate the speeds for free flow traffic only. Both lanes were open to traffic at Station 1. However, because Station 1 was located upstream of the taper and was not influenced by the construction zone, the inside lane served mostly as a passing lane and was not used heavily. Speeds on the outside lane tended to be lower than those on the inside lane. Consequently, only observations from the outside lane were used as control data, yielding a more conservative analysis. The speeds measured at the far station are a subset of the speeds measured at Station 1, as some vehicles platooned after entering the work zone. By collecting speeds at these points, it was possible to establish a control station and determine how the traffic reacted to the flagger only. Station 1 exhibits travel speeds of vehicles not influenced by the flagger or the construction zone. The far station exhibits travel speeds of vehicles influenced by the construction zone and radar. The near station exhibits travel speeds of vehicles influenced by the construction zone, radar, and the flagger.

All data were collected on weekdays when there were good weather and dry pavement; night conditions and peak traffic periods were avoided. Data were collected either in the early morning or shortly after the lunch break, and at these times, the flaggers did not appear to be tired. Because the flaggers had not been flagging for an extended period of time before data collection, it is assumed that they were not excessively fatigued and that they performed at their usual effectiveness level. During data collection periods, there was no significant change in the layout of the construction zone, the activity taking place, or the type of construction equipment used. Thus, there were not any external factors that influenced the drivers' speeds.

The time, type, color, and speed of vehicles passing each flagger before and after training were recorded by two teams. The speeds of a vehicle where measured near to the flagger, at the near station, and far away from the flagger, at the far station, using two radar guns. The distance between the two stations was about 1,600 to 1,800 ft. A typical data collection setup inside the construction zone is shown in Figure 2. The radar guns had an antidetection bottom that transmitted signals only when it was activated. The radar guns were used because the flaggers did not stay at one location during the entire data collection period. The patching crew worked at

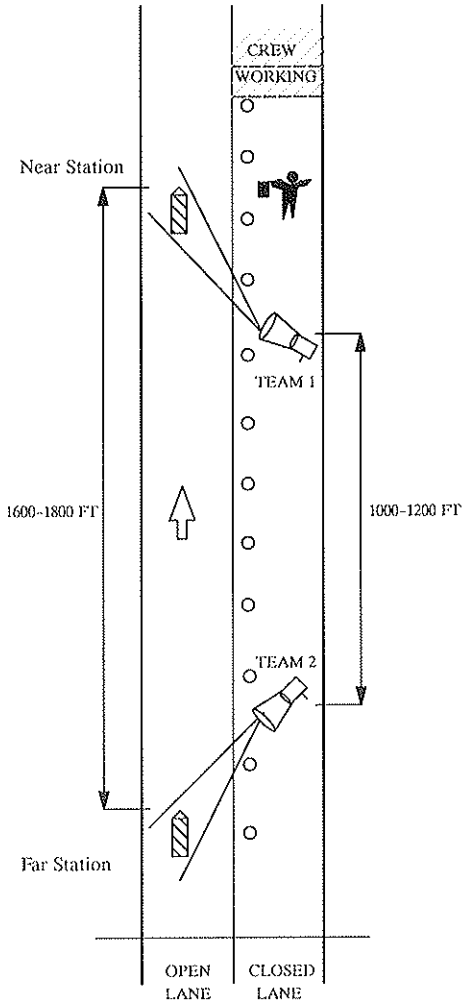



**FIGURE 2 Data collection setup inside the construction zone.**

one location for approximately half an hour and then moved to the next location that needed repair, with the flagger moving with the crew as well. The radar guns were the most appropriate device for data collection under the moving operations because they could be used anywhere that the flaggers were working.

The radar gun at the far station was pointed upstream and had a range of half a mile to a mile. It is assumed that vehicles that had radar detectors would receive the radars signal and that the drivers would slow down by the time they reached the far station. Drivers of vehicles that did not have radar detectors would not be influenced by the radar at either station. Thus, for both cases, any reaction to radar would take place before reaching the far station, and it is assumed that the use of radar guns did not influence the results of the study.

It is possible that the vehicle speeds were affected in the working area as a result of driving through the lane closure and that a speed difference between the far and near stations could be influenced by the lane closure. However, the locations of the far and near stations were a minimum of 1 mi downstream of the beginning of the lane closure, and it is probable that the drivers adjusted their speeds as a result of the lane closure before the far station. In the event that the

vehicle speeds were influenced by lane closure in the data collection area, this influence would be present both before and after the flagger was trained. Consequently, the influence of the lane closure would be the same both before and after training, and the net effect in the comparison of before and after training would be insignificant.

The two teams used their personal cars to collect the data. Each team consisted of two students, who stayed inside the car during the data collection periods. During data collection, one person activated the radar to measure the speed of the vehicles, and the other person wrote a description of the vehicles. The cars used were subcompact cars, which did not resemble police cars. There were no police on the northbound lanes during data collection. The cars were parked on the shoulder or the side of the road in the closed lane section. The vehicles of the construction crew were located close to those of the team. Drivers did not mistake the team cars for police cars.

## DATA REDUCTION

The flagger effectiveness data were reduced in two parts: data from Station 1 were reduced using a Fortran program called TRAFFIC, whereas data from the far and near stations were reduced manually. The data from Station 1 gave the free flow traffic data outside of the construction zone for the desired time period. Free flow vehicles at Station 1 were considered to be those that had a minimum of 6 sec time headway between them and the preceding vehicle on that lane. Traffic data collected at Station 1 were the time a vehicle passed over the tubes, the vehicle speed, the number of axles, and the axle spacings. The TRAFFIC program, written at the University of Illinois specifically for this type of data reduction, separated the traffic data by lane, eliminated data that contained obvious errors, determined if a vehicle was moving at free flow or in a platoon, and then separated the records for the free flow vehicles for a give time period. The reduced data were also examined for any errors that the program may not have eliminated.

Speed reduction between the far and near stations was computed by subtracting speeds of the same vehicle at these two points. The information about each vehicle recorded by the two teams was matched for the near and far station descriptions to ensure that the speeds of the same vehicle were measured at the two locations. The team far from the flagger identified the vehicle whose speed was not influenced by the other vehicles in the work zone. The identified vehicles were either the lead vehicle in a platoon of vehicles of a vehicle traveling alone (not in a platoon). Once the vehicle was identified, the team far from the flagger would measure the speed of the vehicle and send a signal to the team near the flagger to measure speed of the vehicle and record type, color, and other descriptive information about the vehicle. The signal was either visual or audio. For each flagger, over 120 observations were made in before and after conditions. At both stations, the vehicles were not influenced by the vehicle ahead of them.

The data recorded by the two teams about the vehicles matched well. Comparing recorded time was helpful in matching the vehicles. Some data points were not used because the speed of the vehicle near the flagger was influenced by the presence of other vehicles. Every attempt was made to reduce the effect of a leading car on the speed of the following car. In the remaining observations, the speeds were not affected by the presence of other vehicles in the work zone. The descriptions of the vehicles recorded by the two teams were compared and if they did not match, the data were not used. Only the matching data points were used for further analysis.

## DATA ANALYSIS OF THE EFFECTIVENESS OF FLAGGING

The effects of flagging before and after the training session are described for two cases for each flagger: (a) effect on cars only, and (b) effect on trucks only. For each case, the analysis consists of describing speed statistics for the vehicles before training, speed statistics for the vehicles after training, and a comparison of the statistics before and after training.

The speed statistics are calculated for Station 1, the far station, and the near station. The computed statistics are average speed, variance of speed, number of observations, difference between the mean speed and the speed limit, and the percentage of speeding vehicles. The average speeds, variances, and percentages exceeding the speed limit are used to evaluate the effect of the flaggers on vehicle speeds and the effect of the training session on the flaggers' effectiveness.

All statistical analyses were done on a microcomputer using the *Statistical Analysis System for Personal Computers* (PC–SAS). Three tests that were used in the statistical analysis are discussed briefly. The variances of speed of each group of data were compared to determine whether the shapes of the distributions of speeds for two conditions were significantly different from each other. An $F$-test was used for this comparison, with a 95 percent confidence level. The calculated $F$-value was

$$F = \frac{s_1^2}{s_2^2} \tag{1}$$

where

$s_1^2$ = variance of speeds for Case 1, and
$s_2^2$ = variance of speeds for Case 2.

The following distributions were compared using the $F$-test: the far station versus the near station for before-training conditions, the far station versus the near station for after-training conditions, the far station before training versus after training, and the near station before training versus after training. Making these comparisons allowed detecting a change in speed distributions and deciding what type of $t$-test to use for comparing average speeds. In the event that the distributions of speeds were not the same, an approximate value of $t$ was used to reflect the difference in distributions (*10*).

A second test, the paired $t$-test, was used to compare speed reductions between the far and near stations, in which the speed observations were made for the same vehicle at both stations. The test was used in before- and after-training conditions. The $t$-value is computed as

$$t = \frac{\bar{d} - D_0}{s_d/n^{1/2}} \tag{2}$$

where

$\bar{d}$ = sample mean of the $n$ differences,
$s_d$ = standard deviation of the $n$ differences,
$n$ = number of pairs used, and
$D_0$ = expected speed difference.

The following comparisons were made using the paired $t$-test: the far station versus the near station before training, and the far station versus the near station after training.

In order to compare the average speeds before and after training, a third test, the $t$-test was used. The $t$-value was determined as

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]^{1/2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{1/2}} \quad (3)$$

where $\bar{y}_1$ and $\bar{y}_2$ are the sample means, and $n_1$ and $n_2$ are the number of observations in each sample.

The $t$-test was used to make the following comparisons: the far station before training versus after training, the near station before training versus after training, and speed reduction between the far and near stations before training versus the speed reduction after training. The tests determined if the flagger was more effective in decreasing speeds after training and if the traffic was affected by the construction zone and radar differently before and after training at the far station.

The $t$-tests were made on the assumption that the speeds had a normal distribution. However, the $t$-tests were insensitive to the assumption of normal distributions (*11*). Consequently, the use of $t$-tests was acceptable in analyzing the data. The $F$-test was used to determine if the distributions were the same and which $t$-value should be used (*10*). Although the calculated and approximate $t$-values in the $t$-test were nearly identical, the approximated $t$-values were used when the distributions were not identical (*10*).

The speed reductions between station 1 and the near station were compared for before- and after-training conditions. This procedure determined the speed reduction that the flagger attained between a control station located outside the construction and a location close to the flagger. Finally, the percents of speeding vehicles were compared at the far and near stations before and after training.

The variances computed in the analysis were used to determine if the distributions of speeds were the same when comparing two different speed distributions. Garber and Gadirani (*12*) found that an increased variability in speeds is likely to increase accident frequency. The speeds of vehicles may exhibit an increased variance after training in comparison with before training, but the speed variance is irrelevant for this study because the data were collected for free flow vehicles only, and therefore the drivers would not be influenced by other vehicles. It is not known what the speed variance of all traffic is, but it was not the purpose of this study to increase the speed variance of vehicles. The purpose of using a flagger in a construction zone was to maintain speeds at the speed limit or at speeds appropriate for the construction area. If a flagger is effective, the speed of vehicles is uniform and the variance is small.

## Effect on Cars Only

### *Effect of Flagger 1 on Cars Only*

**Flagger 1 Before the Training Session** Flagger 1 was observed flagging from 8:05 to 9:35 a.m. on a Tuesday. At this time, the average speed of cars at the near station, on the basis of 58 free flow observations, was 40.27 mph. The variance of the speed was 36.12. The average speed was 4.73 mph less than the posted speed limit of 45 mph, and 20.69 percent of the drivers had speeds greater than the speed limit. At the far station, the average speed was 53.43 mph, and the speed variance was 32.03. The mean speed at the far station was 8.43 mph greater than the speed limit, with 84.48 percent of the drivers exceeding the speed limit. The mean speed at Station 1 was 73.33 mph, on the basis of 156 free flow car observations. The speed variance was 38.06 mph. The average speed was 8.33 mph greater than the speed limit, and 99.36 percent of the observations showed speeds exceeding the speed limit.

**Flagger 1 After the Training Session** After the training session, Flagger 1 was observed from 1:15 to 2:45 p.m. on a Tuesday. At the near station, the average speed was 35.02 mph and the speed variance was 53.14. The average speed was 9.98 mph less than the speed limit, and 5.38 percent of the cars had speeds exceeding the speed limit. At the far station, the average speed was 48.63 mph and the variance was 34.33. The average speed was 3.63 mph greater than the speed limit, and 64.52 percent of the cars were speeding. At both the near and far stations, there were 93 observations of free flow cars and the speed limit was 45 mph. At Station 1, there were 163 free flow car observations, and the speed limit was 65 mph. The average speed was 75.34 mph and the variance of the speed was 38.68. The average speed was 10.34 mph greater than the speed limit, and 93.25 percent of the cars were speeding. Table 1 presents speed statistics calculated for Flagger 1 before and after training.

**Comparison of Before and After Training** The speed variances at the far and near stations before and after training were compared using $F$-tests with 95 percent confidence level. The speed distributions before training at each station were not significantly different from each other ($F$ = 36.12/32.03 = 1.127). The speed distributions at the far and near stations after training were significantly different from each other ($F$ = 53.14/34.33 = 1.548). This difference indicates that in the after-training condition, the speed of vehicles near the flagger had a distribution with a wider range about the mean than the distribution at the far station.

A comparison of the speed variance at the near station before and after training was made. The $F$-value of 1.4712 (53.14/36.12) indicates that there was a significant difference in the speed variances, with the after-training condition indicating a greater dispersion about its mean than did the before-training condition. A similar test was performed at the far station. The test showed that there was no significant difference in the speed distributions at the far station before and after training ($F$ = 34.33/32.03 = 1.0718). The distribu-

TABLE 1 SPEED STATISTICS FOR CARS ONLY, FLAGGER 1 AND FLAGGER 2 BEFORE AND AFTER THE TRAINING SESSION

| | FLAGGER 1 | | | FLAGGER 2 | | |
|---|---|---|---|---|---|---|
| | Station 1 | Far Station | Near Station | Station 1 | Far Station | Near Station |
| Speed Limit | 55 mph | 45 mph | 45 mph | 55 mph | 45 mph | 45 mph |
| **Before Training** | | | | | | |
| Mean Speed | 73.33 | 53.43 | 40.27 | 75.82 | 51.53 | 41.39 |
| Variance | 38.06 | 32.03 | 36.12 | 42.64 | 25.40 | 43.82 |
| Number of Observations | 156 | 58 | 58 | 245 | 66 | 66 |
| Mean Speed – Speed Limit | 8.33 | 8.43 | -4.73 | 10.82 | 6.53 | -3.61 |
| Percent of Speeding Cars | 99.36 | 84.48 | 20.69 | 93.06 | 87.88 | 25.76 |
| **After Training** | | | | | | |
| Mean Speed | 75.34 | 48.63 | 35.02 | 72.86 | 51.69 | 35.59 |
| Variance | 38.68 | 34.33 | 53.14 | 32.14 | 39.81 | 51.40 |
| Number of Observations | 163 | 93 | 93 | 193 | 92 | 92 |
| Mean Speed – Speed Limit | 10.34 | 3.63 | -9.98 | 7.86 | 6.69 | -9.41 |
| Percent of Speeding Cars | 93.25 | 64.52 | 5.38 | 88.60 | 83.70 | 8.70 |

tions of speeds at the far and near stations before and after training, shown in Figures 3 and 4, indicate that there was a shift toward lower speeds in the after training conditions at both stations.

The speed reductions between far and near stations were compared for before and after training. The speed reduction in the before-training session was 12.23 mph, and in the after-training period it was 13.61 mph. The amount of speed reduction before training was compared to the speed reduction after training, using a $t$-test with a 95 percent confidence level. The $t$-test indicated that there was no significant difference in amount of speed reduction before and after training ($t$ = 0.4052). Thus, the speed reduction before training was not different from the speed reduction after training. The distributions of speed reductions between the far and near stations before and after training are shown in Figure 5.

At the near station and the far station, the average speeds before and after training were compared using a 95 percent confidence level. A $t$-test indicated that there was a significant difference in the speeds at the near station before and after training ($t$ = − 4.5970), and that the average speed after training was 5.25 mph less than the speed before training. A similar test was performed on the data at the far station. The $t$-test showed that there was also a significant difference in the average car speeds at the far station before and after training ($t$ = − 4.9540), with the average speed after training being 4.80 mph less than the average speed before training.

The analysis indicates that the speed reduction between the far and near stations was the same before and after training. The average speeds at each station were about 5 mph less after training than before. The speed reduction at the far station after training is not attributed to the flagger. It is possible that the flagger had an increased effectiveness in decreasing the speeds at the near station after training by an additional 5.25 mph, but the reduced speeds may also be a result of slower-moving traffic in the construction zone.

There was a 33.06-mph speed reduction between Station 1 and the near station before training. The speed reduction between Station 1 and the near station after training was 40.32 mph. The speed reduction between the two stations was 7.26 mph greater after training than it was before training. The average speed of cars at Station 1 in the after-training period was only slightly higher than the before-training period, and thus it appears that the increased amount of speed reduction took place at the near station. However, the average speed at the far station was less in the after-training period than it was in the before-training period and consequently, the increased amount of speed reduction cannot be attributed to the near station only. Figure 6 shows the speed profile of the cars as they pass Station 1, the far station, and the near station.

### *Effect of Flagger 2 on Cars Only*

**Flagger 2 Before the Training Session** Data for Flagger 2 before the training session were collected from 2:15 to 3:45 p.m. on a Monday. At this time, the average speed of free flow cars at the near station was 41.39 mph, and the variance of speed was 43.82. The mean speed was 3.61 mph less than the speed limit, and 25.76 percent of the cars were speeding. At the far station, the average speed was 51.53 mph and the speed variance was 25.40. The average speed exceeded the speed limit by 6.53 mph, and 87.88 percent of the cars had speeds greater than the speed limit. There were 66 observations at both the far and near stations, and the speed limit was 45 mph. At Station 1, the speed limit was 65 mph. There were 245 observations of free flow cars, and the average speed was 75.82 mph. The speed variance was 42.64. The mean speed was 10.82 mph greater than the speed limit, and 93.06 percent of the observations had speeds exceeding the speed limit.

**Flagger 2 After the Training Session** Flagger 2 was observed after the training session from 1:15 to 2:45 p.m. on Wednesday. On the basis of 92 observations at the near station, the average speed was 35.59 mph, and the speed variance was 51.40. The 45-mph speed limit was 9.41 mph greater than

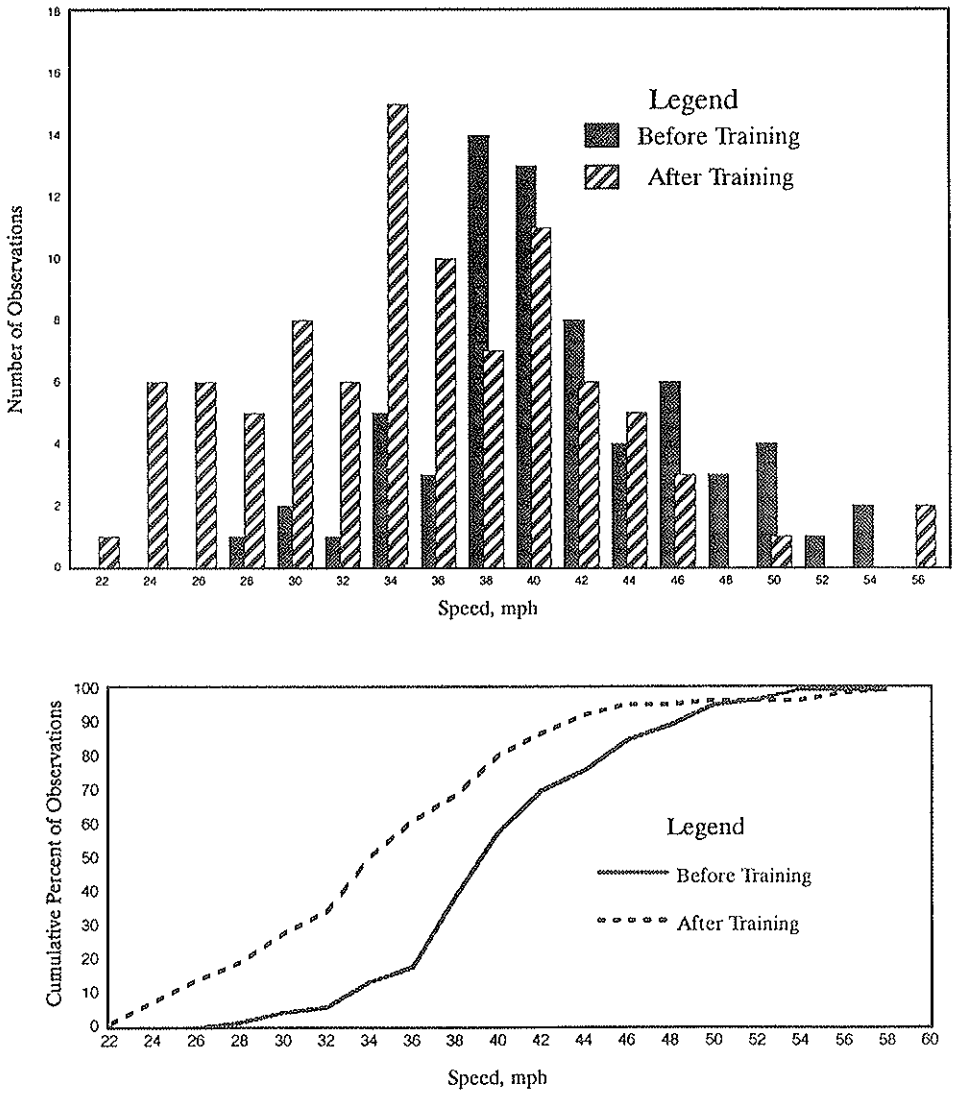
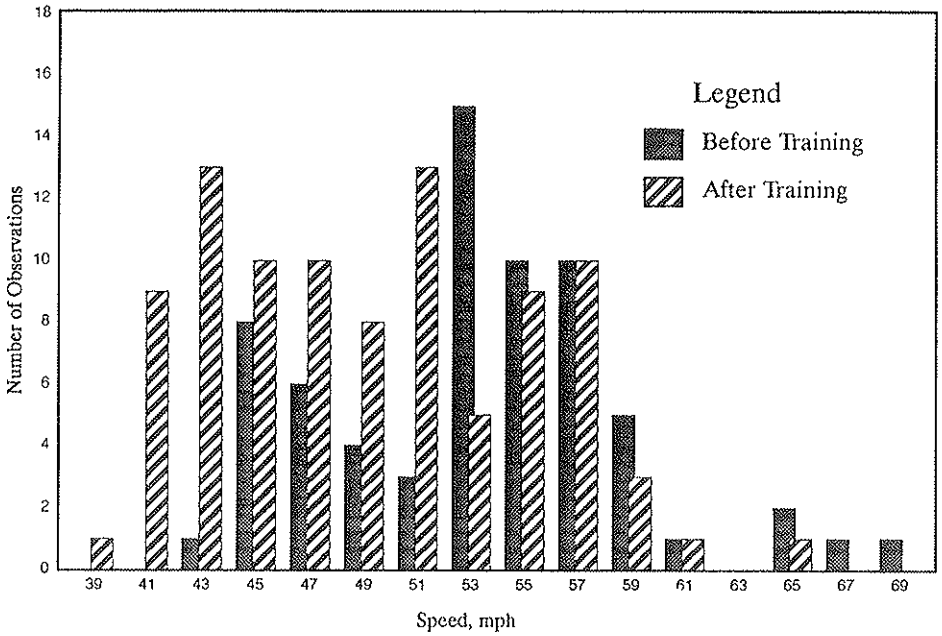



**FIGURE 3 Speed distributions of cars only at the near station before and after training for Flagger 1.**



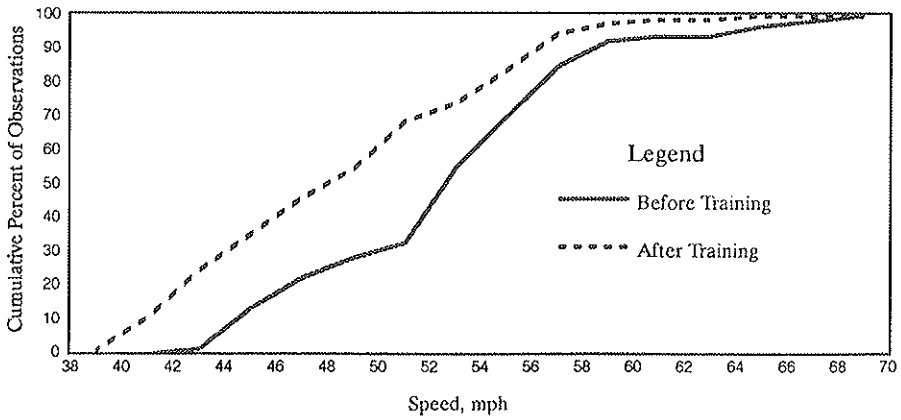




**FIGURE 4 Speed distributions of cars only at the far station before and after training for Flagger 1.**

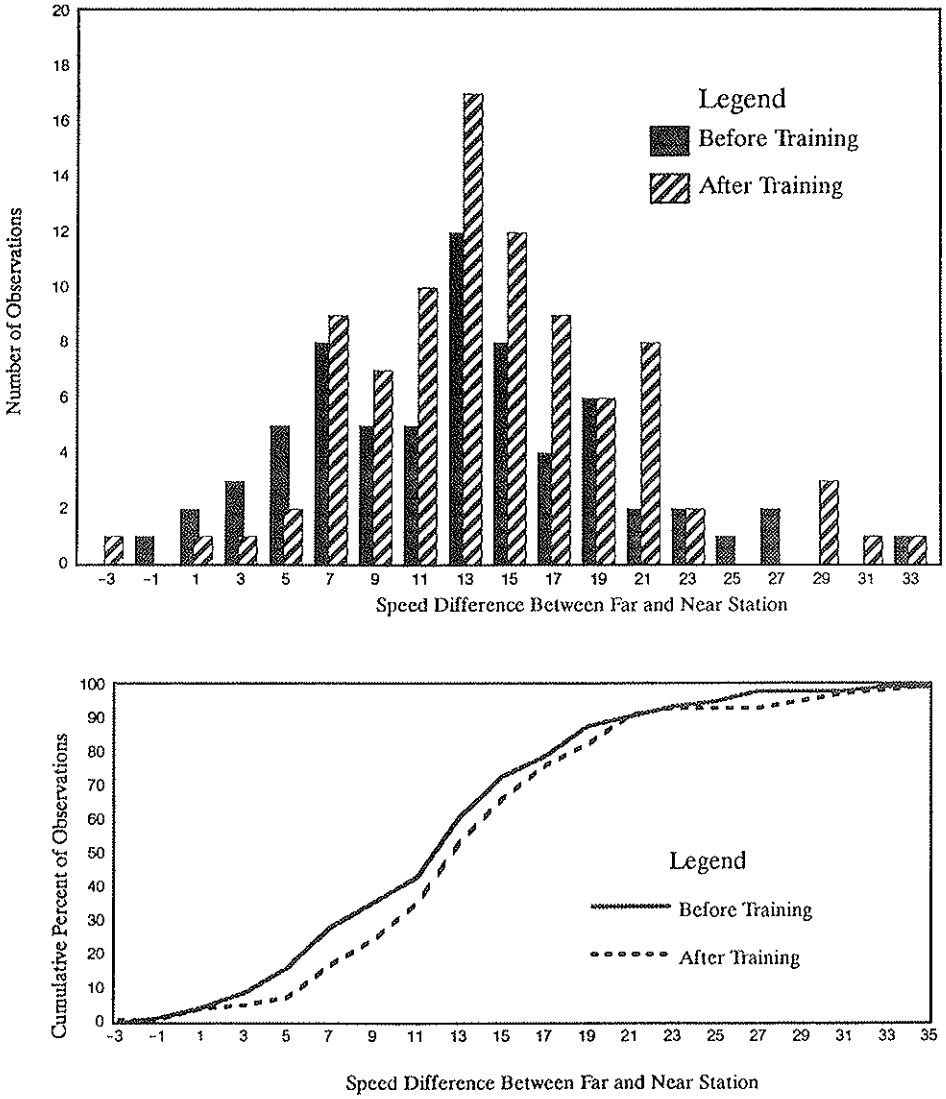



**FIGURE 5 Distributions of speed reductions of cars only between the near and far stations before and after training for Flagger 1.**

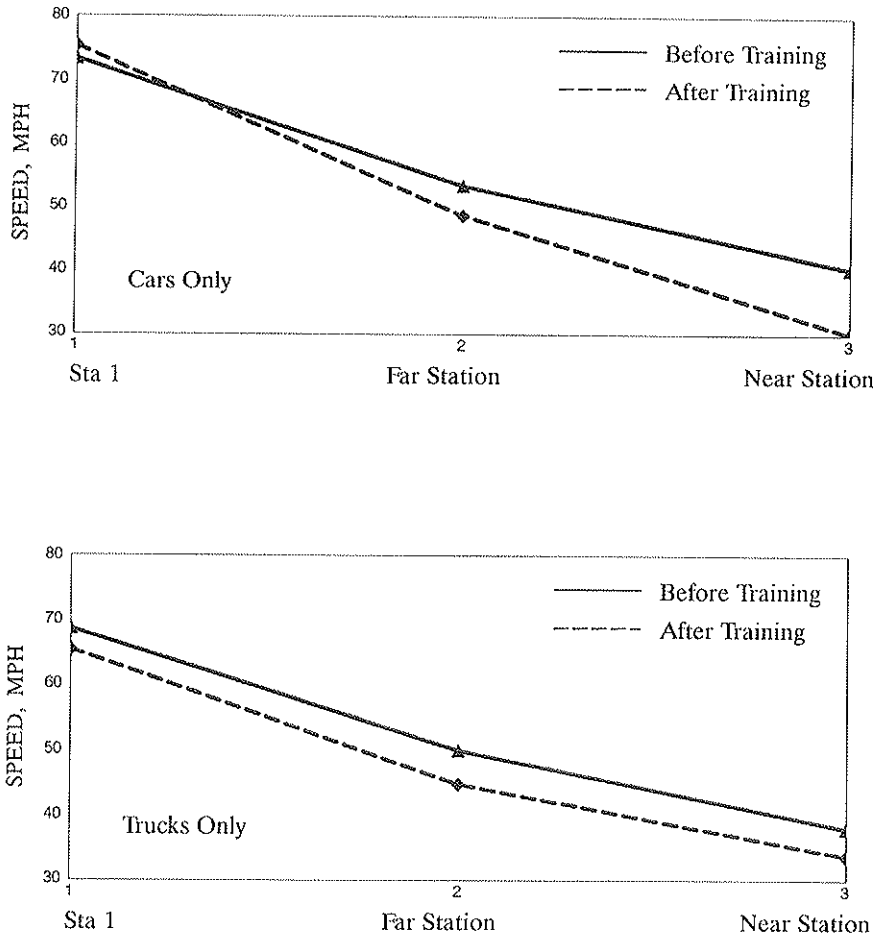



**FIGURE 6 Speed profiles of cars (*top*) and trucks (*bottom*) at Station 1, the far station, and the near station for Flagger 1.**

the average speed. The percentage of speeding cars was 8.70 percent. At the far station, the average speed was 51.69 mph and the speed variance was 39.81. The mean speed was 6.69 mph greater than the speed limit, with 83.70 percent of the observations having speeds greater than the speed limit. There were 92 observations at the far station. The mean speed at Station 1 was 72.86 mph, with a variance of 32.14. This average is based on 193 observations of free flow cars. The average speed was 7.86 greater than the speed limit, and 88.60 percent of the vehicles exceeded the speed limit. Speed statistics for Flagger 2 before and after training are presented in Table 1.

**Comparison of Before and After Training** An *F*-test comparing the speed variance at Station F and the near station before training indicated that the speed distributions were significantly different from each other ($F = 43.82/25.40 = 1.7252$), and that the distribution at the near station had a wider dispersion than the distribution at the far station. However, the distributions of speeds at the far and near stations after training were not significantly different from each other ($F = 51.40/39.81 = 1.2911$).

The variances at the near station were compared for before and after training conditions. There was no significant difference in the distribution of speeds before and after training ($F = 51.40/43.82 = 1.1730$). However, there was a significant difference in the distribution of variances at the far station before and after training ($F = 39.81/25.40 = 1.5673$). The range of the distribution of speeds after training was greater than the distribution before training. The distribution of speeds before and after training at the near and far stations are shown in Figures 7 and 8.

The speed reductions between the two stations were compared for before- and after-training conditions using a paired *t*-test. Before training, the average speed reduction between the far and near stations was 10.13 mph. After training, the speed reduction was 16.09 mph. The results of the test indicate that there was a significant difference in the speed reduction ($t = 5.09$), where the speed reduction after training was 5.96 mph greater than the speed reduction before training. Figure 9 shows the distribution of speed reductions before and after training.

The average speeds of the car observations at the near station were compared for before and after training using a *t*-test with a 95 percent confidence level. There was a significant difference in the speeds ($t = -5.16$), where the average speed of the cars after training was 8.80 mph less than before training. A similar test was used at the far station. There was no significant difference in speeds ($t = 0.1827$). The average

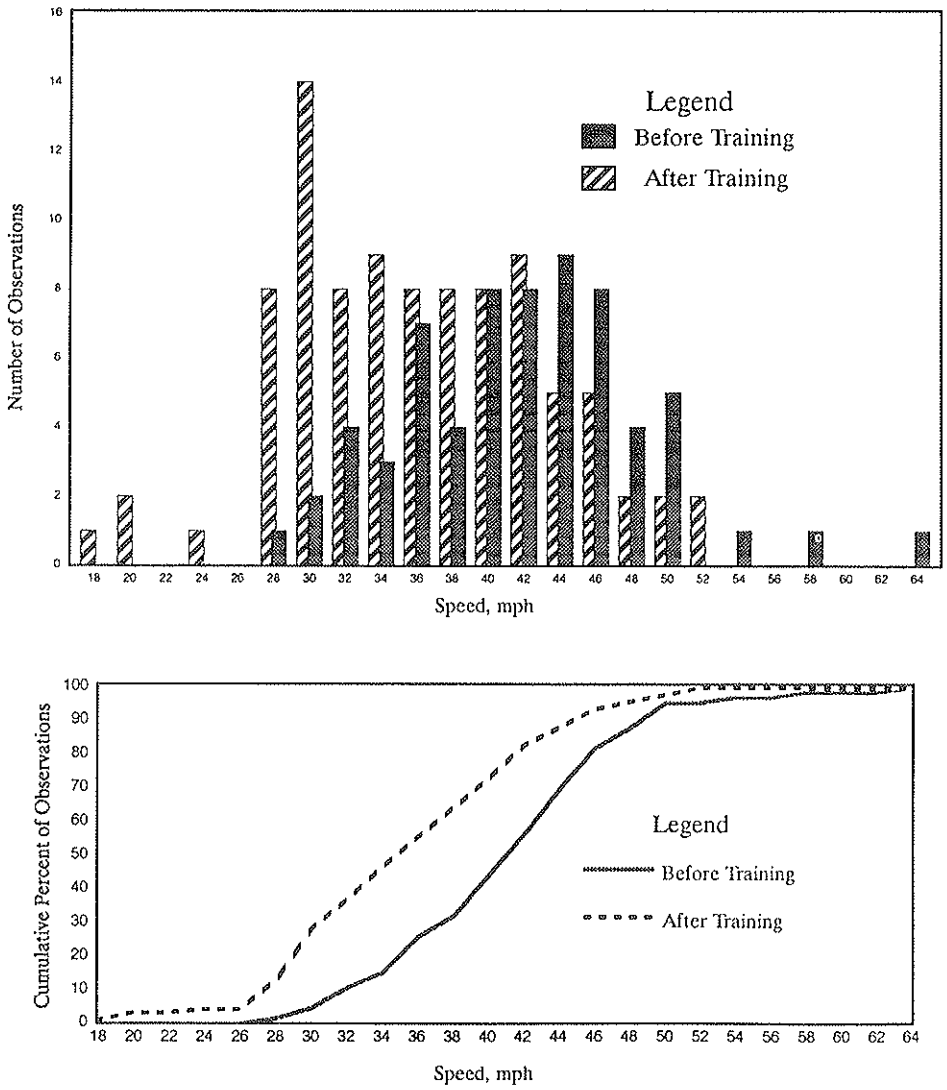
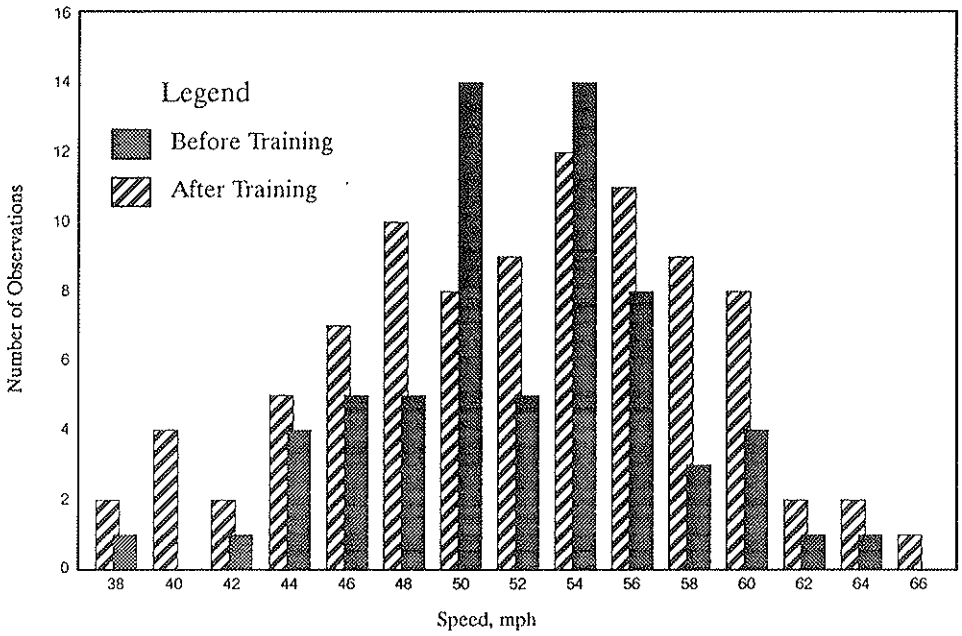



**FIGURE 7 Speed distributions of cars only at the near station before and after training for Flagger 2.**



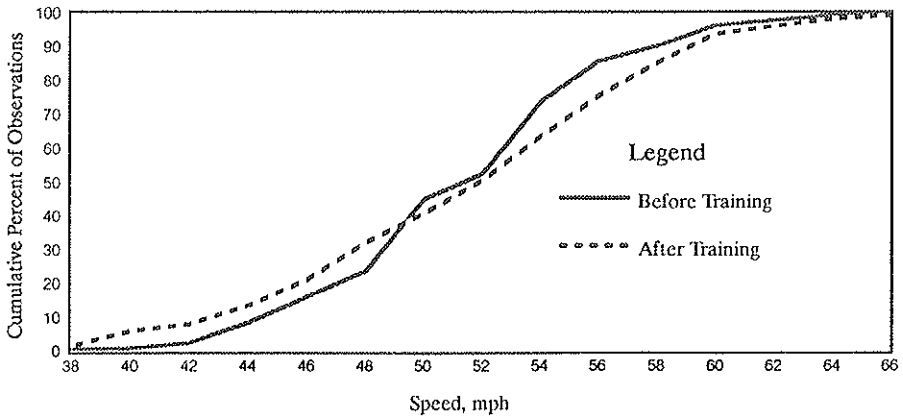




**FIGURE 8 Speed distributions of cars only at the far station before and after training for Flagger 2.**

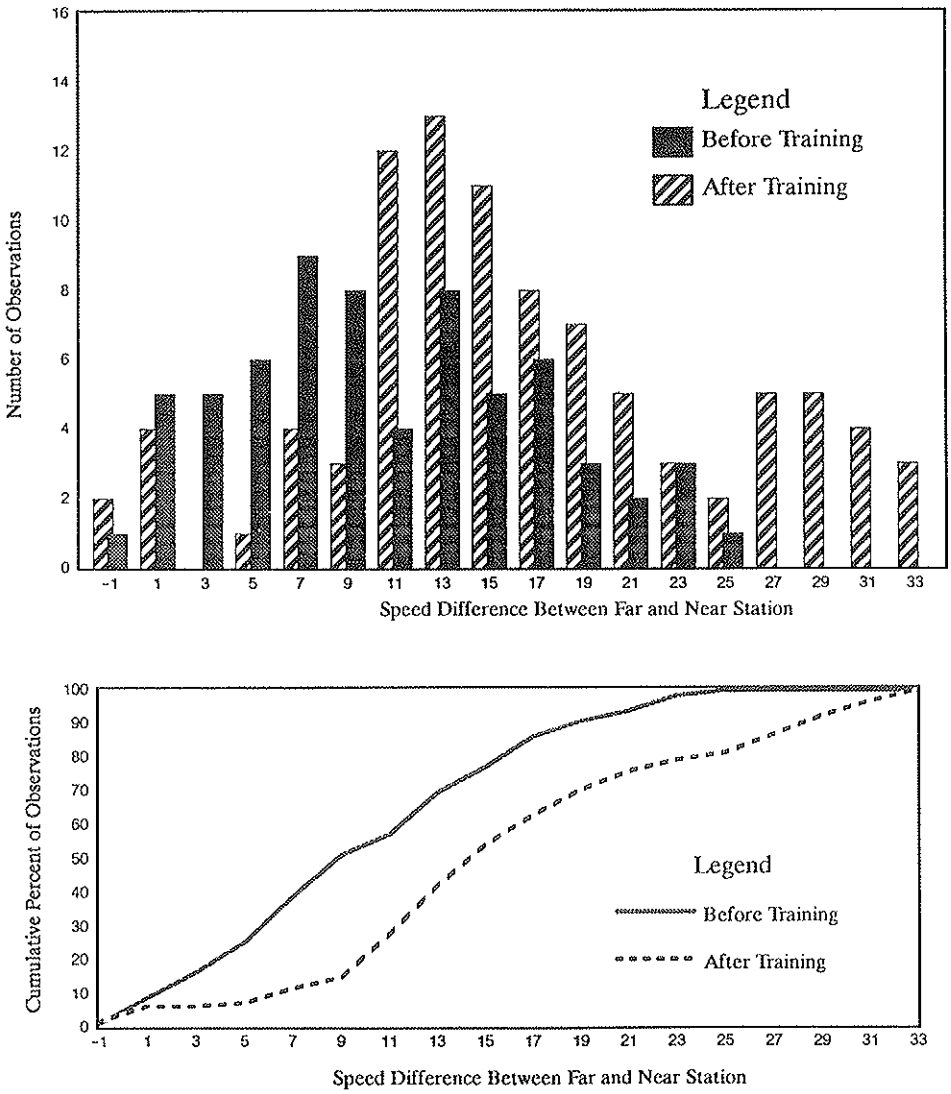



**FIGURE 9 Distributions of speed reductions of cars only between the near and far stations before and after training for Flagger 2.**

speed at the far station before training was not different from the average speed after training.

The average speed of free flow cars at Station 1 was only slightly less in the before-training period than in the after-training period. However, the average speed of cars at the far station was the same for before and after training. Consequently, it appears that the decrease in speed at the near station in the after-training condition, compared to before training, is attributed to an increased effectiveness of Flagger 2.

Before training, there was a 34.43-mph speed reduction between Station 1 and the near station. After training, the speed reduction was 37.27 mph. Hence, the speed reduction after training was 2.84 mph greater than before training. Because the average speed of cars at the far station was the same before and after training, it is possible that the increased speed reduction between the control station and the near station may be attributed to Flagger 2. Speed profiles of cars at Station 1, the far station, and the near station before and after training for Flagger 2 are shown in Figure 10.

## Effect on Trucks Only

### *Effect of Flagger 1 on Trucks Only*

**Flagger 1 Before the Training Session** The average speed at the near station before the training session was 37.93 mph and the variance was 30.47. The average speed was 7.07 mph less than the speed limit, although 6.67 percent of the trucks were traveling faster than the speed limit. At the far station, the average speed was 49.93 mph, and the speed variance was 26.73. The speed limit at the far and near stations was 45 mph, and the number of observations of free flow trucks was 30. The mean speed was 4.93 mph greater than the speed limit, with 83.33 percent of the trucks exceeding the speed limit. At Station 1, there were 45 observations of free flow trucks. The average speed was 68.69 mph, and the variance was 23.52. The average speed was 13.68 mph greater than the speed limit, and 100 percent of the trucks were speeding. The speed limit at Station 1 was 55 mph.

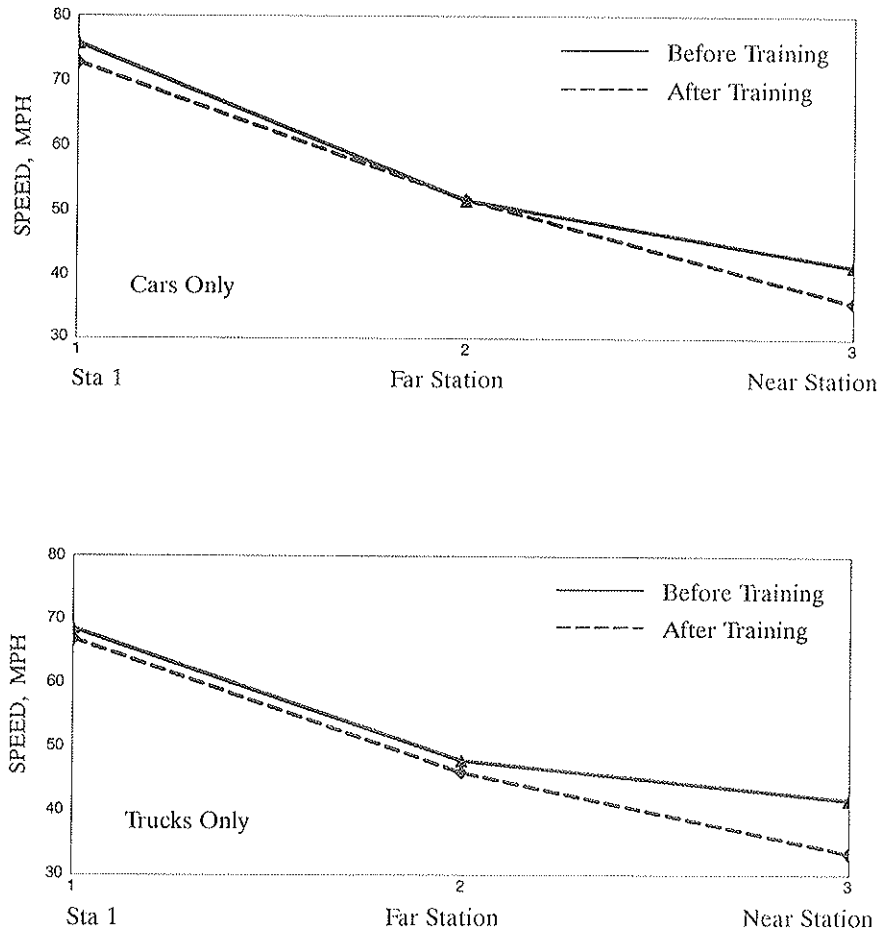



**FIGURE 10 Speed profiles of cars (*top*) and trucks (*bottom*) at Station 1, the far station, and the near station for Flagger 2.**

**Flagger 1 After the Training Session** After the training session, the average speed at the near station, based on 32 observations, was 35.59 mph and the speed variance was 43.82. The mean speed was 11.41 mph less than the speed limit of 45 mph, and 3.13 percent of the trucks were traveling faster than the speed limit. At the far station, the mean speed was 44.75 mph and the variance was 14.59. The average speed was 0.25 mph less than the speed limit, and 40.63 percent of the trucks were speeding. Based on 47 free flow observations, the mean speed at Station 1 was 65.44 mph, and the variance was 45.42. The average speed was 10.44 mph greater than the speed limit of 55 mph, and the percentage of speeding trucks was 95.74 percent. Speed statistics for Flagger 1 for trucks before and after training are presented in Table 2.

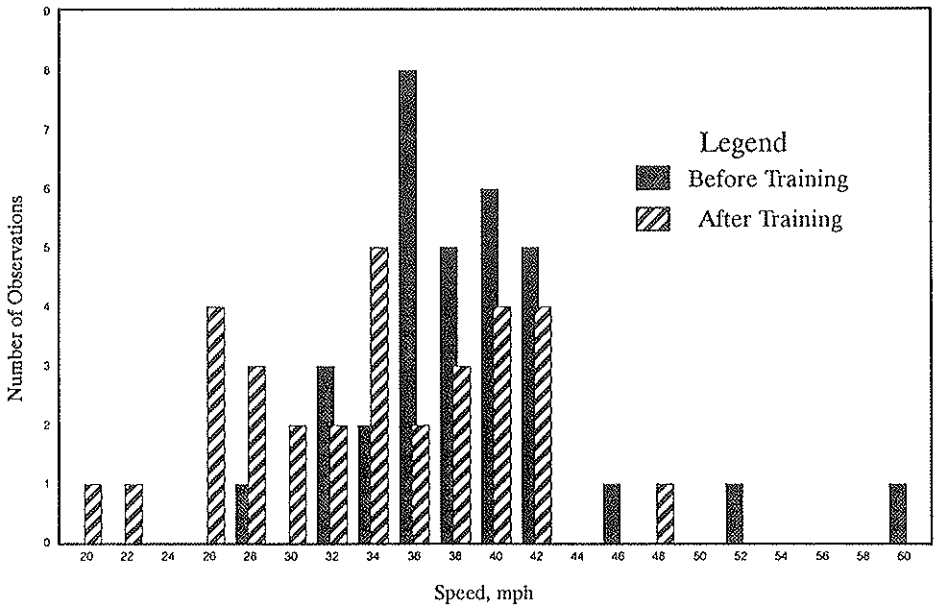
**Comparison of Before and After Training** The variances of speed at the far and near stations were compared using an *F*-test. The results of the test indicate that for before-training conditions, there was no significant difference in the variances at the far station and the near station ($F = 36.47/26.73 = 1.1399$). However, there is a significant difference in the distribution of speeds at both stations after training ($F = 43.82/14.59 = 3.0034$), indicating that the range of the distribution of speeds at the near station after training is greater than the range of the distribution at the far station.

The variances at the near station were compared for before and after training. An *F*-test indicated that there was no significant difference in the distribution of speeds before and after training ($F = 43.82/30.47 = 1.4381$). Likewise, there was no significant difference in the distributions at the far station before and after training ($F = 26.73/14.59 = 1.4381$). The distributions of truck speeds at the near and far stations before and after training are shown in Figures 11 and 12.

The speed reductions between the far and near stations were 12.00 mph before and 11.16 mph after the training session. The speed reductions were compared for before and after training, and there was no significant difference between before and after training conditions ($t = -0.4092$). The speeds at the near station were compared for before and after training, and there was a significant difference in speeds ($t = -2.79$), where the average speed after training was 4.34 mph less than before training. There was also a significant difference in the speeds at the far station before and after training ($t = -4.5050$), with the speed after training being 5.18 mph less than before training. The distribution of speed reduction before and after training is shown in Figure 13.

TABLE 2 SPEED STATISTICS FOR TRUCKS ONLY, FLAGGER 1 AND FLAGGER 2 BEFORE AND AFTER THE TRAINING SESSION

| | FLAGGER 1 | | | FLAGGER 2 | | |
|---|---|---|---|---|---|---|
| | Station 1 | Far Station | Near Station | Station 1 | Far Station | Near Station |
| Speed Limit | 55 mph | 45 mph | 45 mph | 55 mph | 45 mph | 45 mph |
| **Before Training** | | | | | | |
| Mean Speed | 68.69 | 49.93 | 37.93 | 68.43 | 47.79 | 41.61 |
| Variance | 23.52 | 26.73 | 30.47 | 18.74 | 22.46 | 32.03 |
| Number of Observations | 45 | 30 | 30 | 55 | 34 | 34 |
| Mean Speed – Speed Limit | 13.68 | 4.93 | –7.07 | 13.43 | 2.79 | –3.39 |
| Percent of Speeding Trucks | 100 | 83.33 | 6.67 | 100 | 73.53 | 20.59 |
| **After Training** | | | | | | |
| Mean Speed | 65.44 | 44.75 | 33.59 | 66.82 | 45.94 | 33.19 |
| Variance | 45.42 | 14.59 | 43.82 | 18.83 | 25.70 | 27.98 |
| Number of Observations | 47 | 32 | 32 | 72 | 52 | 52 |
| Mean Speed – Speed Limit | 10.44 | –0.25 | –11.41 | 11.82 | 0.94 | –11.81 |
| Percent of Speeding Trucks | 95.74 | 40.63 | 3.13 | 100 | 46.15 | 0.00 |



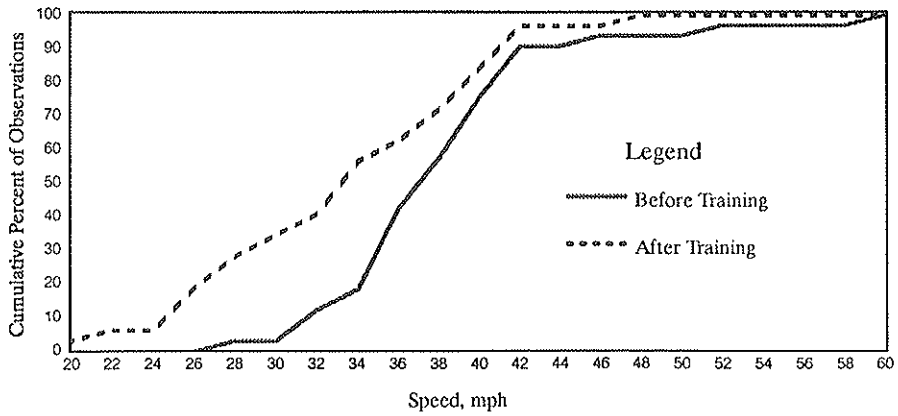




**FIGURE 11 Speed distributions of trucks only at the near station before and after training for Flagger 1.**

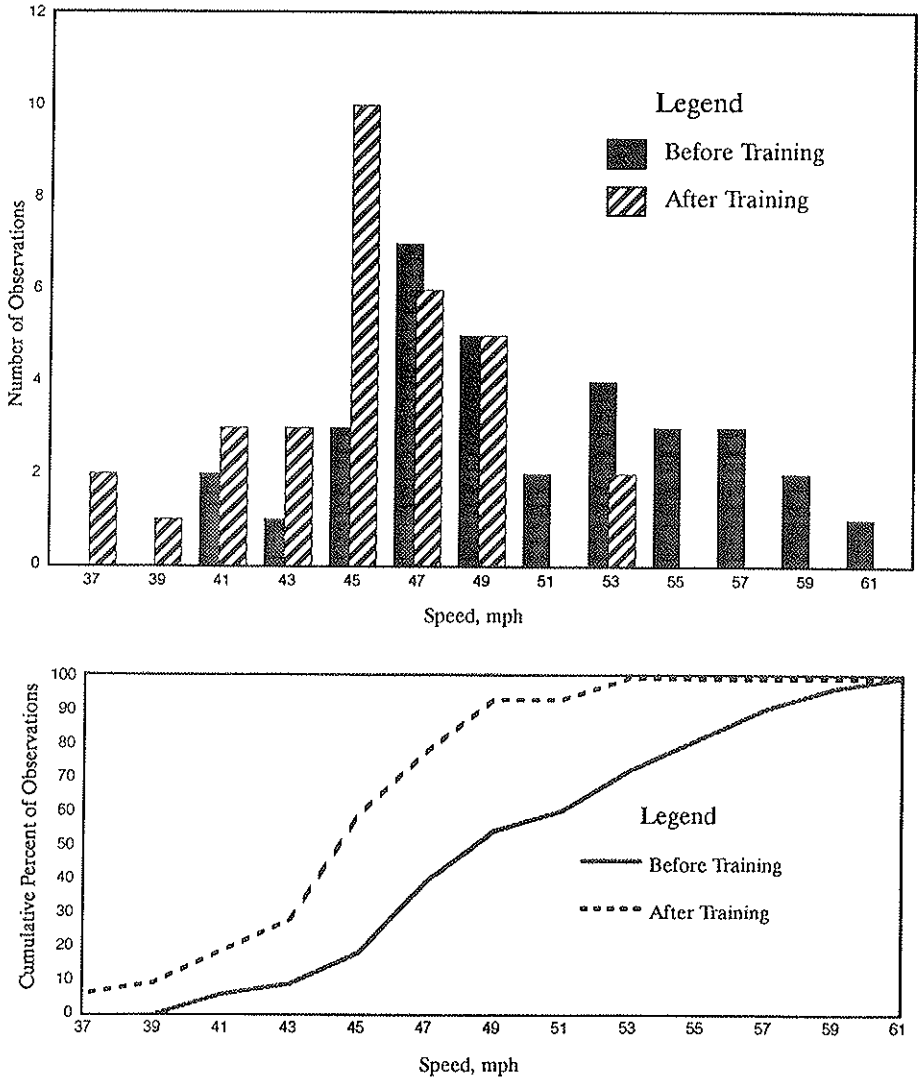



**FIGURE 12 Speed distributions of trucks only at the far station before and after training for Flagger 1.**

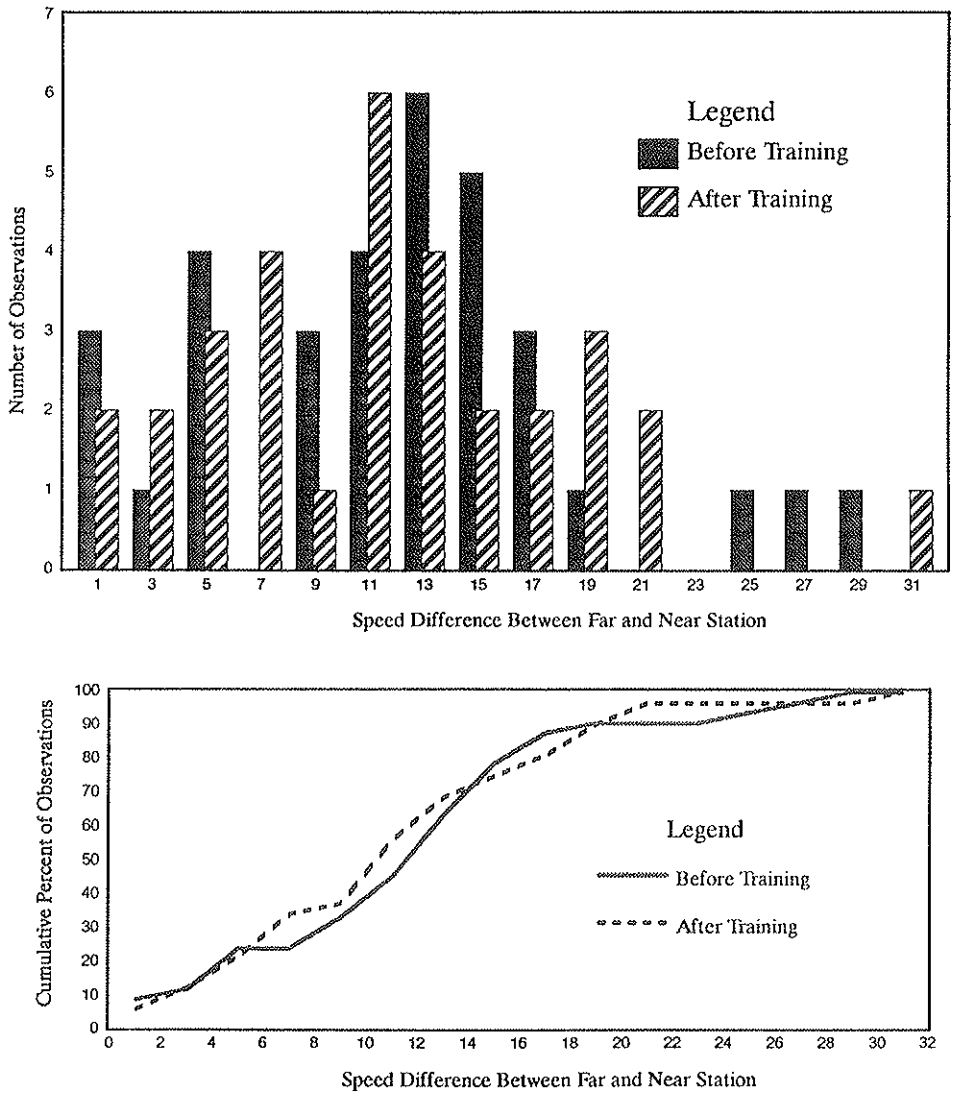



**FIGURE 13 Distributions of speed reductions of trucks only between the far and near stations before and after training for Flagger 1.**

The analysis shows that for after-training conditions, the average speeds of trucks at Station 1, the far station, and the near station, were less than their respective speeds before training. Consequently, although the decreased speed at the near station after training may be a result of an increased effectiveness of Flagger 1 after training, it is also possible that the decreased speed may be a result of a generally slower-moving traffic.

The speed reduction between Station 1 and the near station before training was 30.70 mph. After training, the speed reduction was 31.85 mph. There was a 1.09-mph greater speed reduction between the two stations after training than there was before. Because the speed reduction after training was not much greater than in the before-training period, it appears that the training session had little effect on the speed reduction between the control station and the near station. Speed profiles of trucks at Station 1, the far station, and the near station before and after training for Flagger 1 are shown in Figure 6.

*Effect of Flagger 2 on Trucks Only*

**Flagger 2 Before the Training Session** The average speed at the near station was 41.61 mph, and the speed variance was 32.03. This result is based on 34 free flow truck observations. The average speed was 3.39 mph less than the speed limit of 45 mph, and 20.59 percent of the trucks had speeds exceeding the speed limit. At the far station, the average speed was 47.79 mph and the variance was 22.46. The mean speed was 2.79 mph greater than the speed limit, and 75.03 percent of the trucks were speeding. The average speed at Station 1, on the basis of 55 free flow observations, was 68.43 mph. The speed variance was 18.74. The average speed was 13.43 mph greater than the speed limit, and all of the trucks (100 percent) were speeding.

**Flagger 2 After the Training Session** At the near station, the average speed of trucks after the flagger was trained was

33.19 mph, and the speed variance was 27.98. The average speed was 11.81 mph less than the speed limit, and there were not any trucks speeding (0.00 percent). At the far station, the average speed was 45.94 mph and the variance was 25.70. The average speed was 0.94 mph greater than the speed limit, and 46.15 percent of the trucks were speeding. At both the far and near stations there were 52 truck observations, and the speed limit was 45 mph. At Station 1, the average speed, on the basis of 72 free flow truck observations, was 66.82 mph, and the variance was 18.83. The average speed was 11.82 mph greater than the posted speed limit of 55 mph, and 100 percent of the trucks had speeds exceeding the speed limit. Speed statistics for trucks affected by Flagger 2 are presented in Table 2.

**Comparison of Before and After Training** The variances at the far and near stations were compared for before training conditions using an $F$-test with 95 percent confidence level. The test shows that the distributions at each station were not significantly different from each other ($F$ = 32.03/22.46 = 1.4261). A similar test was used on the speed variances at the far and the near stations after training, and it indicated that there was no significant difference in the speed distributions at both stations ($F$ = 27.98/25.70 = 1.0887). Thus, the distributions of speeds before and after training at the near station were not different from the distributions at the far station.

The speed variance at the near station before training was compared with the variance at the near station after training using an $F$-test with a 95 percent confidence level. The results indicated that the distributions of the speeds at the near station were not significantly different from each other before and after training ($F$ = 32.03/27.98 = 1.1447). Likewise, the same test at the far station indicated that there was no significant difference in the distribution of speeds before and after training ($F$ = 25.70/22.46 = 1.1443). The distributions of speeds at the near station before and after training are shown in Figure 14, and distributions at the far station before and after training are shown in Figure 15.

The speed reduction between the far station and the near station before training was 6.18 mph. After training, the speed

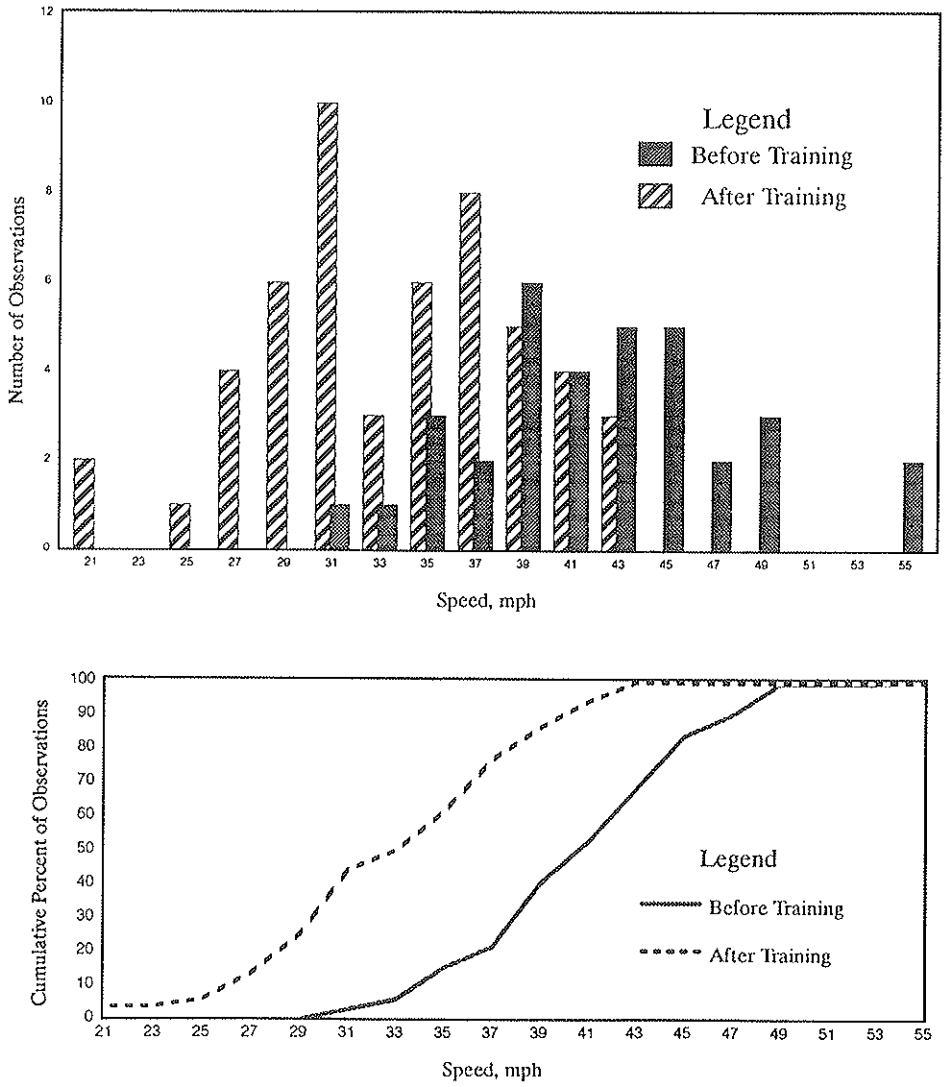



**FIGURE 14 Speed distributions of trucks only at the near station before and after training for Flagger 2.**

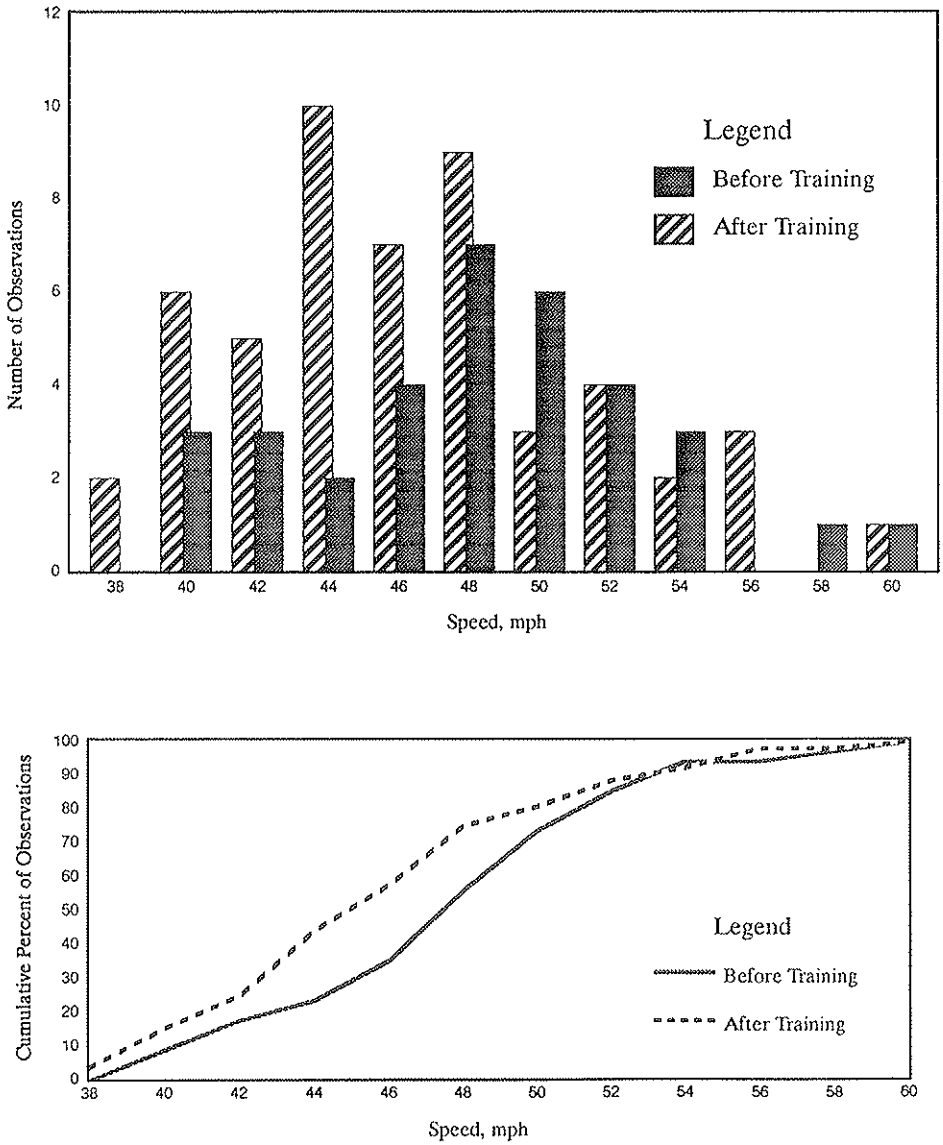



**FIGURE 15 Speed distributions of trucks only at the far station before and after training for Flagger 2.**

reduction between the two stations was 12.75 mph. The speed reductions before and after training were compared using a *t*-test with a 95 percent confidence level. The results indicate that there was a significant difference in the amount of speed reduction ($t = 5.77$). Thus, the speed reduction between the far and near stations was 6.58 mph more after training than it was before training. The distributions of speed reduction of trucks before and after training are shown in Figure 16.

The speeds at the far station were compared for before- and after-training conditions using a *t*-test with a 95 percent confidence level. There was no significant difference in speeds before and after training ($t = 0.0934$). However, a similar test used at the near station indicated that there was a significant difference in speeds before and after training ($t = -7.0191$). Thus, the average speed of free flow trucks at the near station after training was 8.42 mph less than the average speed before training.

The analysis indicates that the average speed at Station 1 after training was slightly less than before training, that the average speed at the far station was the same before and after training, and that the average speed at the near station was significantly lower after training than before. It is possible that because the speeds of trucks were the same at the far station, the speed reduction at the near station is a result of an increased effectiveness of Flagger 2 after the training session.

There was a 26.82-mph speed reduction between Station 1 and the near station before training. After training the speed reduction between the two stations was 33.63 mph. The speed reduction between the two stations was 6.81 mph greater after training than before training. The speeds at Station 1 before and after training were nearly the same (after training was 1.61 mph less than before training), and the speeds at the far station before and after training were the same. Thus, drivers in the after-training period were not considered to be slower-

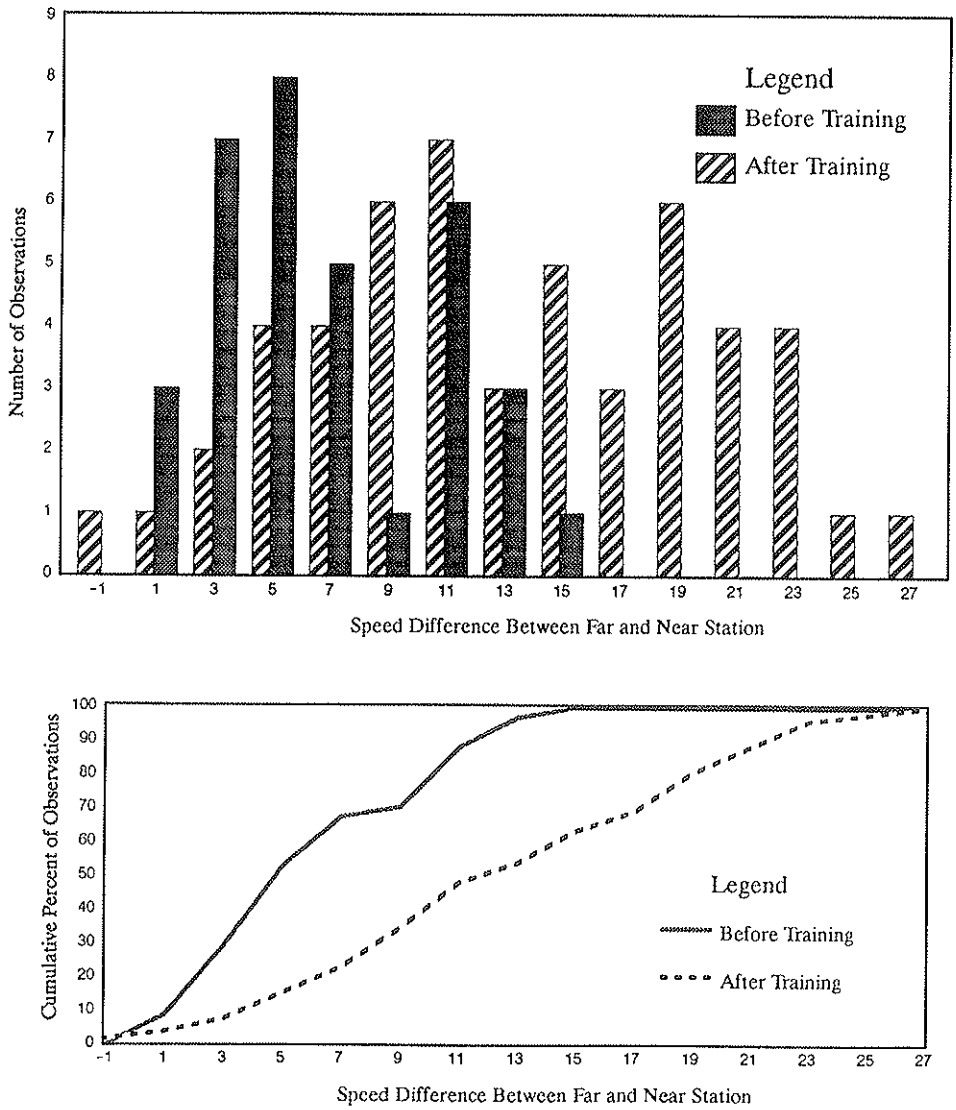



**FIGURE 16 Distributions of speed reductions of trucks only between the far and near stations before and after training for Flagger 2.**

moving traffic, and it considered likely that the greater speed reduction after training was a result of an increased effectiveness of Flagger 2. Figure 10 shows speed profiles of trucks at Station 1, the far station, and the near station for Flagger 2 before and after training.

## CONCLUSIONS AND RECOMMENDATIONS

The results of this study indicate that there was a decrease in the average speed of trucks and cars as they approached the flagger both before and after the flagger received training. Both cars and trucks had speeds exceeding the speed limit at Station 1, outside of the construction zone, and at the far station, inside of the construction zone. However, at the near station, inside of the construction zone, the average speed of free flow vehicles was well below the posted speed limit. At 1600 to 1800 feet upstream of the flagger and inside of the construction zone, the cars and trucks were traveling an average of 7.5 and 3.9 mph faster than the speed limit before training, respectively. After the training session, the cars and trucks were traveling 5.2 and 0.3 mph faster than the speed limit, respectively. Near to the flagger, the average speeds of cars and trucks were 4.2 and 5.2 mph less than the speed limit in before-training conditions, while they were traveling 9.7 and 11.6 mph less than the speed limit in after-training conditions, respectively.

For both cars and trucks, the speed of vehicles near to the flagger is less in the after-training period than in the before-training period. For cars, the speed reduction between before and after training at the near station ranged from 5 to 9 mph, while the speed reduction for trucks was 4 to 9 mph. The speed reduction at the near station cannot necessarily be attributed to an increased effectiveness of the flaggers after having received training. For Flagger 1, the decrease in speeds may be a result of an increased effectiveness of the flagger, or of a slower-moving traffic in the construction zone. For Flagger 2, the decrease in car and truck speeds after the

training session probably was a result of an increased effectiveness because of training. Analysis of field data indicates that training may not have the same effect on all flaggers.

For both flaggers, there was a decrease in the percentage of speeding vehicles as the vehicles approached the flagger, both before and after training.

This study served as a pilot study, evaluating only two flaggers in the same construction zone, with data being collected for a maximum of 6 hr for each flagger. Further studies should be made to evaluate the long-term effectiveness of flaggers after having received training, and should use a large number of flaggers in a variety of study sites. It cannot be said whether training flaggers will have the same effect on all flaggers, or whether or not training will improve the effectiveness of the flagger at all, as the flagger may not follow the instructions given to him in the training session. However, a more comprehensive training program may improve the effectiveness of training and the performance of the flagger.

## ACKNOWLEDGMENT

The authors would like to thank the Illinois Department of Transportation and the FHWA for the financial support to conduct this study and George Houck for training the flaggers. The authors would also like to thank Robin Orloski, Jeff Linkenheld, and Mike Lee for their help in collecting data, and the Project Advisory Committee members for their comments and suggestions during this study.

## REFERENCES

1. *Manual on Uniform Traffic Control Devices for Streets and Highways*. FHWA, U.S. Department of Transportation, 1988.
2. S. H. Richards, R. C. Wunderlich, and C. L. Dudek. Field Evaluation of Work Zone Speed Control Techniques. In *Transportation Research Record 1035*, TRB, National Research Council, Washington, D.C., 1985, pp. 66–78. U.S. DOT 1985.
3. S. H. Richards, R. C. Wunderlich, C. L. Dudek, and R. P. Bracket. *Improvements and New Concepts for Traffic Control in Work Zones, Vol. 4—Speed Control in Work Zones*. FHWA/RD/85/037. FHWA, U.S. Department of Transportation.
4. R. F. Benekohal, and L. Kastel. *A Procedure for Evaluation of Training Flaggers For Traffic Control In Rural Interstate Construction Zones*. Department of Civil Engineering, University of Illinois at Urbana, Jan. 1991.
5. G. L. Ullman, S. Levine, and S. C. Booker. *Flagger Safety and Alternatives to Manual Flagger*. FHWA/TX-87/38 + 406-1F, FHWA, U.S. Department of Transportation, 1987.
6. G. L. Ullman and S. Levine. An Evaluation of Portable Traffic Signals at Work Zones. In *Transportation Research Record 1148*, TRB, National Research Council, Washington, D.C., 1987.
7. S. C. Booker, G. L. Ullman, and S. Z. Levine. Supplemental Devices to Enhance Flagger Safety. In *Transportation Research Record 1148*, TRB, National Research Council, Washington, D.C., pp. 34–38. 1987.
8. E. C. Noel, C. L. Dudek, O. J. Pendleton, H. W. McGee and Z. A. Sabra. *Speed Control Through Work Zones: Techniques Evaluation and Implementation Guidelines*. FHWA/1P/87/4. FHWA, U.S. Department of Transportation, 1987.
9. E. C. Noel, C. L. Dudek, O. J. Pendleton, and Z. A. Sabra. Speed Control Through Freeway Work Zones: Techniques Evaluation. In *Transportation Research Record 1163*, TRB, National Research Council, Washington, D.C., 1998, pp. 31–43.
10. *SAS/STAT User's Guide, Release 6.03 Edition*. SAS Institute Inc., Cary N.C., 1988, pp. 941–946.
11. G. E. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters*. John Wiley, New York, 1978.
12. N. J. Garber and R. Gadiravi. *Speed Variance and its Influence on Accidents*. AAA Foundation for Traffic Safety, Washington D.C., July 1988.

---



*Publication of this paper sponsored by Committee on Traffic Safety in Maintenance and Construction Operations.*

# Guidelines for the Use of Truck-Mounted Attenuators in Work Zones

JACK B. HUMPHREYS AND T. DARCY SULLIVAN

Truck-mounted attenuator (TMA) usage varies for a number of reasons, including apathy, loss of efficiency (real or perceived), fiscal constraints, and lack of information on when and how to use the devices. Although some individual states have adopted policies, there has not been any coordinated effort to develop guidelines for the use of TMAs on a national basis. A literature review was conducted to determine the extent to which guidelines might have been developed but not widely shared. Five states were visited to solicit information regarding support for, and extent of use of, TMAs. There was a wide range in the number of TMAs presently in use. There was more consistency on other issues including the following: (a) initial support for the use of TMAs came principally from administrators; (b) field support is generally good in states using tilt-up versions of the TMA; (c) reported uses included maintenance activities, construction activities, and emergency incident management (use of TMAs on shadow vehicles was, by policy, the most common application); and (d) there seemed to be little factual basis for the existing application policies. A set of recommended guidelines was developed that included priorities for the deployment of shadow vehicles and TMAs. Two limitations on the significance and suggested use of the guidelines are acknowledged. First, the project did not involve collection and analysis of numerical data. Rather, it represented an effort at bringing together appropriate policies and procedures. Second, the guidelines are more appropriately used as a policy formation and budgeting tool.

The hazardous nature of construction and maintenance work zones on and along streets and highways has been recognized for many years. Unfortunately, knowledge all too frequently is not translated into action; when it is, the time required for transition and implementation of newly developed procedures is sometimes lengthy. Only in recent years, for example, has there been implementation of many of the principles set forth in the 1967 AASHTO publication *Highway Design and Operational Practices Related to Highway Safety* (*1*), frequently referred to as the "Yellow Book." Specifically, that document stated that the use of traffic control plans; improvements in signing, channelization and pavement markings; portable barriers; better training of flaggers; arrow panels; changeable message signs; and improved construction scheduling can all combine to produce safer work zones.

During the late 1970s, work zone safety was considered an emphasis area by the FHWA. The impetus for this emphasis largely resulted from a fatal January 1975 work zone accident on the I-495 beltway around Washington, D.C., and subsequent legal action involving the FHWA and other governmental agencies. Research activity into the identification of work zone safety problems, with recommendations for specific research to address those safety problems, was completed in 1979 (*2*). Extensive changes were incorporated into Part VI of the 1978 *Manual on Uniform Traffic Control Devices* (MUTCD) (*3*), many reflecting the principles set forth in the Yellow Book (*1*). Even further changes are noted in the 1989 MUTCD (*4*).

Even with the changes in the 1978 MUTCD and the FHWA emphasis, the number of work zone accidents nationwide has continued to increase. Between 1982 and 1987, the number of construction zone fatalities increased 43 percent nationally. In Illinois alone, there were 23 fatalities in 1988 work zone accidents (*5*). Much of this increase may be attributed to the fact that more and more highway construction and reconstruction is being performed under traffic.

On the basis of a recent six-state survey by Graham-Migletz Enterprises in conjunction with its Strategic Highway Research Program activity (*5*), the five top operations with the largest number of work zone accidents (based on a total of 324 reported accidents) are as follows:

- Snow and ice control,
- Pavement maintenance,
- Flagging,
- Sweeping, and
- Pavement marking.

Adding to the cost of highway accidents nationally is the expense of lawsuits against governmental agencies. It has been estimated that highway agencies paid $120 million in judgments and settlements from tort liability claims in 1986. This amount does not include an additional $20 million required to defend these cases. Because the rate of such suits is increasing at 17 percent per year, engineers and managers are justifiably concerned (*6*).

In order to respond to these work zone accident statistics, both in magnitude and cost, agencies have promoted work zone safety in a variety of ways. Extensive training programs have been undertaken by many states. The authors, for example, have provided 2- and 3-day seminars several times across the State of North Carolina over the last 10 years through the University of Tennessee Transportation Center. Similar seminars have been given in a number of other states and municipalities by the authors. Training in work zone safety is also offered by the American Traffic Safety Services Association (ATSSA), the Institute of Transportation Engineers (ITE), the National Highway Institute (NHI), and others.

In addition to training, the use of more extensive traffic control plans and the upgrading of traffic control devices have

J. B. Humphreys, Department of Civil Engineering, University of Tennessee, Knoxville, Tenn. 37996-2010. T. D. Sullivan, 6956 Riverwood Drive, Knoxville, Tenn. 37920.

both improved and emphasized the need for better work zone traffic controls.

## HISTORY OF TRUCK-MOUNTED ATTENUATORS (TMAs)

Other aspects of highway safety have also been recently addressed. During the 1950s, highway agencies became aware of the large number of fixed roadside hazards that were playing an increasing role in the number of fatalities and injuries. In addition to a realization that such hazards should be removed or relocated, attention was directed to the mitigation of the results of such fixed-object impacts. Crash cushions, or impact attenuators, were considered, and development began.

One of the first such attenuators was the steel drum crash cushion system developed in Texas in the mid-1960s (*7*). Extensive research and development by federal and state governmental agencies and by the highway safety industry has since produced a wide variety of impact attenuators that can be adapted to varying site-specific highway conditions or needs. These facilities include water-filled tubes, sand-filled plastic barrels, and crushable, dry energy-absorbing materials.

Success with these crash cushion designs has stimulated development of mobile systems that are attached to work vehicles. Perhaps the first of these was the Texas crash cushion trailer, developed and tested in 1972 (*8*). Adapted from the fixed-drum attenuators developed and in use in Texas, the design consisted of 55-gal steel drums welded together and mounted on a flat trailer, which was then towed behind a truck. According to the researchers, acceptable collision performance was demonstrated in a head-on impact by a 4,000-lb automobile at 60 mph (*9*).

From this early attenuator, other TMA systems soon following. Designs to date include the following (*8*):

- Energy-absorbing cartridges within a frame [Hex-Foam, by Energy Absorption Systems, Inc. (EASI)] (see Figure 1);
- Aluminum honeycomb with frame (Hexcel by Hexcel, Inc., Alpha 1000 by EASI, Alpha 500 by EASI);
- Water-filled tubular vinyl cells (CushionSafe by TranspoSafety, Inc.); and
- Collapsing (or crushing) steel pipe (developed by University of Connecticut).

The highway safety industry has made extensive improvements to first-generation TMAs. Designs now provide for consistently safe load levels for both light and heavy automobiles over a range of impact speeds, as well as increased maneuverability of TMA trucks because of the tilt-up option with hydraulically activated latching and other improvements (see Figure 2). Overall weights of TMA units have decreased, and the time (and difficulty) of mounting and unmounting the devices from trucks has been greatly reduced. Current TMA designs are thus more effective and easier to use with a vehicle fleet.

## USAGE OF TMAs

With the emphasis on work zone safety exhibited by the FHWA and others, improvements in the level of traffic control provided are quite evident in many states. The use of signing, channelization, markings, etc., has improved vastly in most areas, particularly on larger contract work. The use of temporary concrete barriers, arrow panels, and changeable message signs has also improved motorist and worker safety.

Unfortunately, TMAs have not been so readily and uniformly accepted across the United States. Several factors have apparently contributed to this lack of acceptance, among them the following:

- Negative experience with first-generation TMAs, including mounting procedures, inadequate tilt capabilities, etc.;
- *Perceived* loss of productive work time without significant gain in safety for employee;
- Truck tieup (with dedicated TMA usage);
- Lack of positive local accident experience within the agency;
- Initial cost of TMAs;
- The fact that TMAs are not required by MUTCD; and

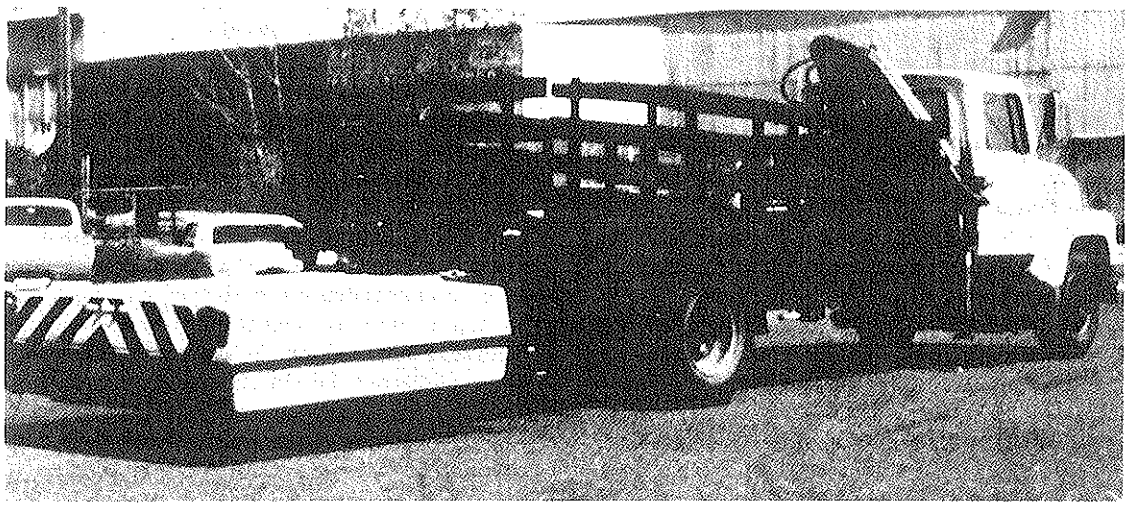


**FIGURE 1 Hex-foam TMA by Energy Absorption Systems, Inc.**

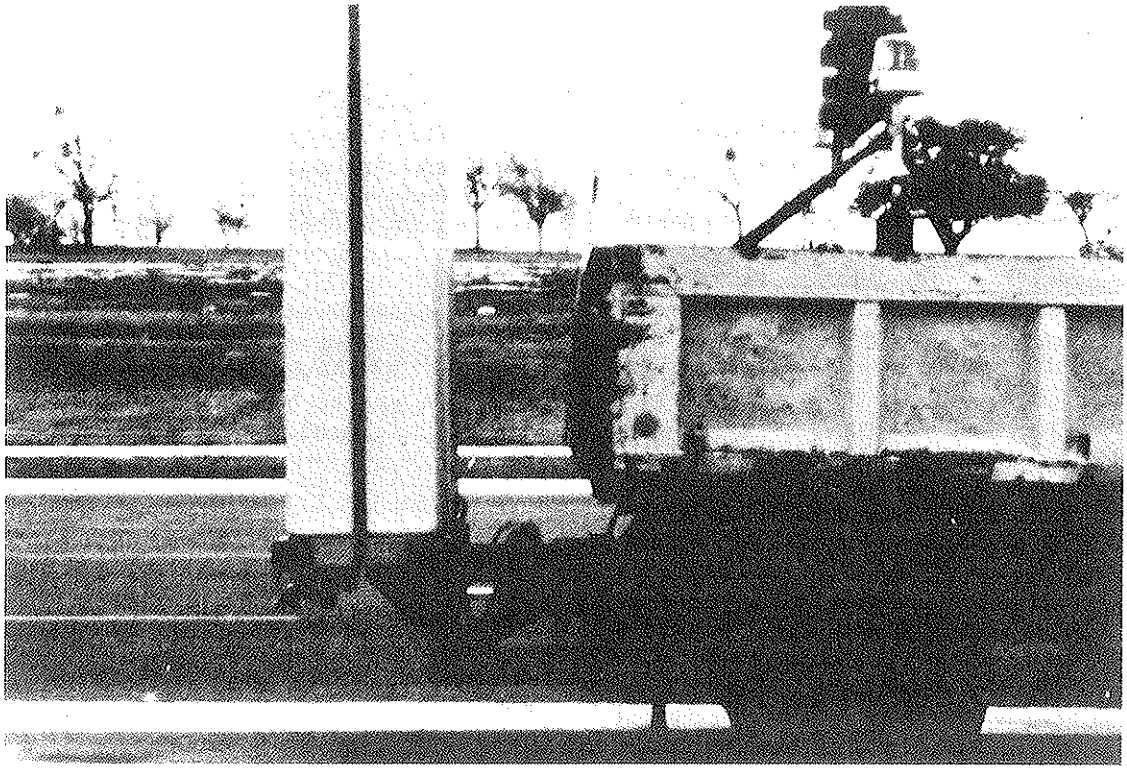


**FIGURE 2 Tilt-up capability of TMA.**

- Lack of widespread (national) policy and procedures for TMA usage (including both *where* and *how* a TMA should be used).

A partial review of the 1980s' TMA experience across the states provides some indication of the lack of uniformity in TMA usage during that period. Perhaps the earliest most specific reference to TMA use was added in July 1981 to the MUTCD in one state, reading as follows:

> At stationary work areas, a shadow vehicle with an attenuator fastened to the rear should be placed upstream of the work area. For moving work areas, the attenuator should be placed on the rear of the work equipment and/or shadow vehicle. (Source intentionally not included.)

Although this text appears to provide sufficient direction and would suggest extensive TMA use, apparently that state, as of early 1990, has only four TMA units within the highway department—certainly not enough to meet the requirements of their MUTCD.

By 1982, the Oakland County, Michigan, Road Commission had one TMA for each of its seven operating districts. Four additional TMAs were purchased in 1985 for use in its more urban districts (*10*).

By 1984, the Texas State Department of Highways and Public Transportation (SDHPT) had several TMAs in use. Each Texas highway district has funds to purchase equipment, with acquisitions to be approved by headquarters personnel. Region 2, headquartered in Forth Worth, was using two units full time in restriping operations alone. They also maintained five TMA cartridges in inventory to meet immediate replacement needs (*11*).

A 1985 report on highway safety devices, prepared for the Texas legislature by the SDHPT, estimated the value of a TMA in such accidents. Savings of $23,000 per accident in injury and damages were estimated for a vehicle hitting a TMA instead of a stationary vehicle, resulting in a favorable benefit-cost ratio (*12*).

Other states moved quickly to use TMAs in their operations. By 1987, California had approximately 500 TMAs in use. By that time, policy required a TMA on the rearmost vehicle in work-in-progress operations. All vehicles moving significantly slower than prevailing traffic, such as in sweeping or painting operations, also had to be equipped with a TMA. CALTRANS agreed that the life-saving benefits to motorists and workers made the crash cushions worthwhile. In addition, savings have been recognized in the repair and replacement of damaged equipment (*13*).

In 1986, a task force was appointed by the North Carolina state highway chief engineer to develop recommendations concerning safer operations for slowly moving maintenance work. A summary of guidelines for maintenance operations was prepared in 1987. Although many of the operations required only rotating beacons on the equipment (such as contour mowers and broom tractors), shadow vehicles with TMAs were recommended for herbicide spraying operations and painting operations using cones, whereas edge line painting (without cones) had the TMA optional on the trailing vehicle. Those guidelines did not address the issues of exposed personnel on foot doing patching, sealing, or other similar work.

A 1988 shadow vehicle policy distributed to all New York regional highway engineers addressed the issue of the required use of shadow vehicles. However, the policy indicated that TMAs were not required on those vehicles, but would be used

"if available and where practical" on both moving and stationary operations on multilane highways. They would be used on two-lane highways "if desirable."

After a St. Louis vehicle struck a TMA involved in a striping operation, with the motorist escaping serious injury, the Missouri Highway Department studied increasing TMA usage. Plans were developed in 1989 to attach TMAs to departmental vehicles performing routine maintenance operations (*14*). Similarly, Florida Department of Transportation officials drafted a set of guidelines for the use of protective equipment, but as of 1989, each district had authority in the decision to require such equipment. In some cases, TMAs are required, such as on contract sweeper operations in Duval County.

Georgia also has developed guidelines for protective equipment, but, as in Florida, those guidelines are not mandatory, and the language is broad. As the assistant state maintenance engineer has stated, TMAs are required "in any instance where there's a high likelihood of impact in an open lane situation" (*15*).

More definitive requirements for TMA usage appeared in the 1987 *Virginia Work Area Protection Manual*, which is a supplement to the Virginia MUTCD; thus, its use is mandatory. Both the 1987 manual and its 1988 revision establish a number of conditions where TMAs are to be used. . . . After July 1, 1988, TMAs were required on all limited access highways," using the following criteria (*16*):

- Pavement marking,
- Stationary lane closures,
- Other mobile maintenance operations, and
- Other situations as warranted.

## PURPOSE OF THIS RESEARCH

As suggested earlier, there is a great variance in usage of TMAs among the states, with some states having virtually none, whereas California has over 500 in use. Even in those states with a number of TMAs, guidelines for usage are in general loosely worded, giving field personnel a great deal of leeway in their application. It would be appropriate to develop some set of nationally accepted guidelines, warrants, or priorities to obtain the usage having the greatest probability of increasing overall safety and reducing total costs. The purpose of this research, then, is to address this issue by suggesting priorities as to how and where available TMAs should be deployed. Then, given the availability of one or more TMAs, supervisory personnel would be able to assign them more effectively on a day-to-day basis. Also, if a priority system can be agreed on within a given agency, the total number of TMAs required to cover a certain level of priority can be better estimated more accurately.

## DEVELOPMENT OF GUIDELINES

Several states were selected as candidate contacts to determine the status of current TMA programs. The states represented a range of attributes with respect to

- Geographic location,
- Apparent interest in the use of TMAs, and
- Number of units in active use.

The states were contacted to determine their willingness to discuss their use of TMAs with the research staff. Initial contacts with the states simply suggested the possibility of a meeting to discuss how TMAs were being used within the agency and what their experiences (good and bad) had been. States ultimately selected for participation in the process were California, Iowa, North Carolina, Tennessee, and Texas.

Discussion sessions were held during July and August of 1989. Agency personnel attending the sessions were selected by the agency and ranged in number from three to seven. Job responsibilities of those in attendance included maintenance foremen, supervisors, and engineers; traffic engineers and technicians; purchasing agents; occupational safety and training officers; garage repair personnel; and construction engineers.

During the discussions agency personnel were invited to comment on the origins of their TMA programs, the general availability of TMAs to field personnel, what were the most common applications, the basis for the assignment of application priorities, and the acceptance of the devices by a broad range of agency personnel. Although there was a wide range of responses on the number of TMAs presently in active use (from fewer than 10 to over 500), there was far more consistency from state to state on other issues discussed. Some of the issues on which there were strong similarities included the following:

1. The initial support for the use of TMAs came principally from the administrative level. In some cases the concern was primarily employee safety, in other cases primarily motorist safety. Most programs dated from the early 1980s.
2. Support for the use of TMAs among field personnel is generally good to very good in states using the tilt-up versions of the TMA. Some field crews are reported to feel so strongly that they virtually refuse to undertake certain assignments unless a TMA-equipped vehicle is available. When available units did not incorporate the more recent technologies including the tilt-up feature and reasonably easy mounting and dismounting of the units, support among field personnel was absent.
3. Reported uses, in order of reported frequency, included maintenance activities, construction activities, and emergency incident management. The use of TMAs on shadow vehicles to moving operations was, by policy, the most common application. However, there was support among the field personnel involved in the discussions for more frequent use of TMAs on barrier vehicles in stationary operations. The safety of exposed personnel was the primary concern of the field forces.
4. There seemed to be little factual basis for any existing application policies. Only one state had comprehensive data available on accidents involving TMA-equipped vehicles, and those data could not be related to exposure in a statistically meaningful way. When TMAs were used regularly, the field personnel often had vivid recollection of specific incidents that did influence usage policies.

On the basis of the information gathered during the agency visits, a draft of suggested TMA use guidelines was prepared.

Those guidelines attempted to reflect the existing practice of the agencies, the expressed concerns of the field personnel who participated, and the experience of the researchers. These draft guidelines were presented to a large group of industry personnel to determine how they thought such information would be received by the various agencies they called on. The draft was modified to reflect comments received, and then was taken back to two of the states originally visited seeking first-hand response. The response was generally favorable, but the guidelines were considered too complicated to be used by field personnel.

The material was again revised to simplify the format and provide more agency flexibility in the application of the suggested guidelines. Draft materials then were distributed to those in attendance at the January 1990 committee meetings of the TRB A2A04 Committee on Roadside Safety Appurtenances and A3C04 Committee on Traffic Safety in Maintenance and Construction Operations. Committee members and others in attendance were asked to review the draft guidelines and were invited to later provide comments on either the content or format of the guidelines.

On the basis of input from the described sources and a number of other informal contacts by the project staff, a final set of guidelines was developed.

## RECOMMENDED GUIDELINES

Before a set of priorities can be established for the uses of TMAs, a system must be available for defining the type of activity taking place. Previously identified factors that affect the type and number of traffic control and protective devices to be used and how they are to be used include the following:

- Speed of traffic;
- Whether the work area is within the roadway, within the shoulder (if one is present), or off the roadway or shoulder;
- Type of activity (moving, intermittent, or stationary);
- Roadway environment: access controlled versus nonaccess controlled and urban versus rural;
- Traffic volumes; and
- Exposure to special hazards.

Although many factors may be important in determining the overall traffic control plan to be implemented at any particular job site, five were selected as particularly relevant to a decision whether or not to use a TMA. Three of those factors are as follows:

- Location of Work Area. Locations of primary concern are those within the traveled lanes and those within all-weather frequently used shoulders. Activities taking place within the traveled lanes are more likely to become involved in an incident than are shoulder activities.
- Type of Activity. Whether the activity is moving, intermittent, or stationary will determine whether or not a standard lane closure or shoulder closure will be implemented. Activities taking place within a formal lane or shoulder closure are less likely to become involved in an incident than are activities fully exposed to approaching traffic.
- Special Hazards. Some activities by their nature expose personnel to greater hazards than do others. Operations involving personnel on foot or located in exposed positions on or within work vehicles (on the platform of a cone pickup truck or in a bucket performing overhead operations, for example) are particularly susceptible to high-severity incidents. Other activities may create conditions that present a significant hazard to vehicles in the passing stream and their occupants.

Table 1 presents a structure for classifying various activities considering the previously discussed lane and shoulder closure and exposure conditions. Examples of typical construction and maintenance activities for each of the closure or exposure conditions also are provided.

Tables 2 and 3 suggest priorities for the assignment of shadow or barrier vehicles and TMAs. Two additional factors that were identified as having an impact on assignment priorities are reflected in these tables.

- Access Control. Access-controlled facilities frequently give drivers a false sense of security with a resulting lower expectation of interruptions to free traffic flow. Therefore, activities on freeways may be more likely to become involved in incidents than are activities on nonaccess controlled facilities in which most drivers are operating at a higher state of alertness.
- Speed Limit. Higher operating speeds leave less time for response, and impacts at higher speeds generally result in more severe injuries and damage. Therefore, activities on facilities with higher speed limits are likely to become involved more frequently and in more severe incidents than are activities on facilities on low-speed facilities.

During the interviews with agency personnel, many of the field personnel felt strongly that the use of a blocking vehicle (generally referred to as a shadow vehicle for moving and intermittent operations and a barrier vehicle for stationary operations) was highly desirable for the protection of exposed personnel *even if* a TMA was not available. Many agencies have a policy regarding the use of blocking vehicles. Those that have a policy may desire to continue to follow that policy. Table 2 suggests priorities that are consistent with the expressed concerns of the field personnel and that may be considered when no policy currently exists.

Table 2 indicates that the suggested priorities for the assignment of blocking vehicles are related directly to protection of agency personnel. In each case in which personnel are exposed, a positive recommendation is provided, with the strength of that recommendation depending on the closure condition, the prevailing speed of traffic, and whether or not the operation is occurring on a freeway.

When exposed personnel are not involved, the use of a blocking vehicle may or may not be justified. That decision will depend on an evaluation of the hazards that exist within the work area and the likely loss if a blocking vehicle is struck. If the evaluation indicates that impact with a blocking vehicle is likely to result in less damage or less serious injury than would impact with a work area hazard or a working vehicle, then a blocking vehicle should be assigned to the operation.

TABLE 1 EXAMPLES OF CLOSURE AND EXPOSURE CONDITIONS

| Closure/Exposure Condition | Examples of Typical Construction/ Maintenance Activities | See Figure |
|---|---|---|
| No Formal Lane Closure | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | Crack pouring, patching, utility work, striping, coning | 3 |
| Shadow Vehicle for Operation Not Involving Exposed Personnel | Sweeping, chemical spraying | 3 |
| No Formal Shoulder Closure | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | Pavement repair, pavement marking, delineator repair | 4 |
| Barrier Vehicle for Operation Not Involving Exposed Personnel | Open excavation, temporarily exposed bridge pier | 4 |
| Formal Lane Closure | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | Pavement repair, pavement marking | 5 |
| Barrier Vehicle for Condition Involving Significant Hazard | Open excavation | 5 |
| Formal Shoulder Closure | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | Pavement repair, pavement marking, guardrail repair | 6 |
| Barrier Vehicle for Condition Involving Significant Hazard | Open excavation | 6 |

Definitions:

- A FORMAL CLOSURE condition (either lane or shoulder) includes a full complement of advance warning devices, a closure taper of channelizing devices, and channelizing devices to define the work area as required.
- A NO FORMAL CLOSURE condition (either lane or shoulder) includes limited (if any) advance warning signs and channelizing devices.
- A SHADOW VEHICLE is a moving vehicle traveling a short distance upstream from a moving operation giving physical protection from approaching traffic.
- A BARRIER VEHICLE is a vehicle parked a short distance upstream from a stationary operation giving protection from approaching traffic.

If the projected damage or injury is greater, then the vehicle should not be assigned. Two examples follow:

- An open excavation several feet deep and several feet across exists on a street in a residential area. A horizontal curve restricts sight distance to the excavation to less than desirable for the 25-mph speed limit. An impact with an appropriate blocking vehicle at 25 mph would probably result in less damage than would driving into a major excavation. Therefore the use of the blocking vehicle would be appropriate.
- A full-depth portland cement concrete patch has been placed and is curing in the right lane of an arterial street with prevailing speeds of >40 mph. An impact with an appropriate blocking vehicle at 40 mph would probably result in greater loss (in both personal and economic terms) than would driving into an uncured patch that might then have to be replaced. Therefore the use of the blocking vehicle would be inappropriate.

Table 3 presents suggested priorities for the assignment of available TMAs. Table 3 indicates that the suggested priorities for the application of TMAs are based primarily on the protection of the approaching motorists. The highest priority is on a freeway where speeds are high and the probability of an impact is greatest. When, because of either the location of the activity or the presence of a formal closure, the probability of an impact is less, a lower priority is assigned.

Figures 3–6 show the use of TMA-equipped vehicles in the closure and exposure conditions identified in Table 1. The relative simplicity of the illustrations compared with illustrations in the MUTCD may be misleading and the following items should be noted:

- In most cases, the use of traffic control devices in the advance warning area and transition area, as defined in the *Traffic Control Devices Handbook* (17), will be appropriate. Because this topic is adequately covered in the MUTCD, in other agency policies, and, where applicable, in the project traffic control plan, those details are not repeated on the figures.
- Figure 3 specifically recommends an arrow panel on the TMA-equipped vehicle. In all of the other figures, it is indicated as an option. In every case, the note indicates that the device is to be operated in accordance with existing agency policy.

TABLE 2 SUGGESTED PRIORITIES FOR THE ASSIGNMENT OF SHADOW AND BARRIER VEHICLES

| Closure/Exposure Condition | Ranking* | | | |
|---|---|---|---|---|
| | | Non-Freeway with Speed Limit | | |
| | Freeway | ≥50 mph | 40-45 mph | ≤35 mph |
| No Formal Lane Closure | | | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | A | A | A | A |
| Shadow Vehicle for Operation Not Involving Exposed Personnel | E | E | E | E |
| No Formal Shoulder Closure | | | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | B | B | C | C |
| Shadow Vehicle for Operation Not Involving Exposed Personnel | E | E | E | E |
| Formal Lane Closure | | | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | B | B | C | D |
| Barrier Vehicle for Condition Involving Significant Hazard | E | E | E | E |
| Formal Shoulder Closure | | | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | C | C | D | D |
| Barrier Vehicle for Condition Involving Significant Hazard | E | E | E | E |

*The ranking letter indicates the priority assigned to the use of a shadow/barrier vehicle. The use of shadow/barrier vehicles:

A is very highly recommended.
B is highly recommended.
C is recommended.
D is desirable.
E may be justified on the basis of special conditions encountered on an individual project when an evaluation of the circumstances indicates that an impact with a shadow/barrier vehicle is likely to result in less serious damage and/or injury than would impact with a working vehicle or the hazard.

- When a formal lane closure or shoulder closure is implemented, a buffer area (or buffer space) as defined in the *Traffic Control Devices Handbook* is typically provided. Because this topic is adequately covered in the handbook, the MUTCD, in other agency policies, and, where applicable, in the project traffic control plan, those distances are not repeated on the figures.
- When a blocking vehicle is hit, it will be moved forward some distance. That distance is commonly referred to as the "roll-ahead distance" and varies depending on the weights and speeds of the two vehicles involved, the extent to which the blocking vehicle is restrained, and certain pavement characteristics. All of the factors except vehicle weights and impacting vehicle speed can be accounted for with a series of assumptions. The likely speed of the impacting vehicle is site specific. The weight of the units used as blocking vehicles and the weight of the impacting vehicle to be accommodated by the system are both policy issues.

Tables 4 and 5 present listings of calculated and rounded roll-ahead distances for various vehicle weight and speed conditions. Calculations were made using the classical conservation of momentum equation and the following assumptions:

- Coefficients of friction between truck tires and pavement surface of 0.50,
- Percent of total vehicle weight on rear axles of shadow or barrier vehicles of 75 percent,
- Engine braking effectiveness of moving shadow vehicle of 80 percent, and
- Values rounded downward as appropriate.

Appropriate values reflecting the agency's policy decisions should be taken from Tables 4 and 5 and inserted in the figures before the figures are distributed for use by field forces.

## CONCLUSIONS

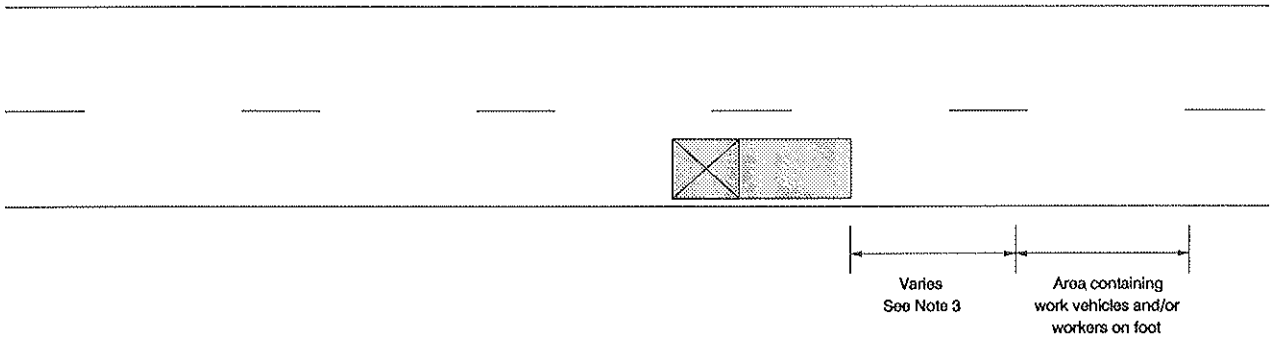
This research effort has resulted in guidelines that may be of assistance in determining the priority of usage of shadow or barrier vehicles and of TMAs. The suggested guidelines

TABLE 3 SUGGESTED PRIORITIES FOR THE APPLICATION OF TMAs

| Closure/Exposure Condition | Priority* | | | |
|---|---|---|---|---|
| | Freeway | Non-Freeway with Speed Limit ≥50 mph | 40-45 mph | ≤35 mph |
| No Formal Lane Closure | | | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | 1 | 2 | 3 | 4 |
| Shadow Vehicle for Operation Not Involving Exposed Personnel | 1 | 2 | 3 | 4 |
| No Formal Shoulder Closure | | | | |
| Shadow Vehicle for Operation Involving Exposed Personnel | 2 | 3 | 3 | 3 |
| Shadow Vehicle for Operation Not Involving Exposed Personnel | 2 | 3 | 4 | 5 |
| Formal Lane Closure | | | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | 2 | 3 | 4 | 5 |
| Barrier Vehicle for Condition Involving Significant Hazard | 2 | 3 | 4 | 5 |
| Formal Shoulder Closure | | | | |
| Barrier Vehicle for Operation Involving Exposed Personnel | 3 | 4 | 5 | 5 |
| Barrier Vehicle for Condition Involving Significant Hazard | 3 | 4 | 5 | 5 |

*The numerical rank indicates the level of priority assigned to the use of a TMA on an assigned shadow/barrier vehicle. The use of a TMA under the defined conditions is:

1 is very highly recommended.
2 is highly recommended.
3 is recommended.
4 is desirable.
5 may be justified on the basis of special conditions encountered on an individual project.




Notes:

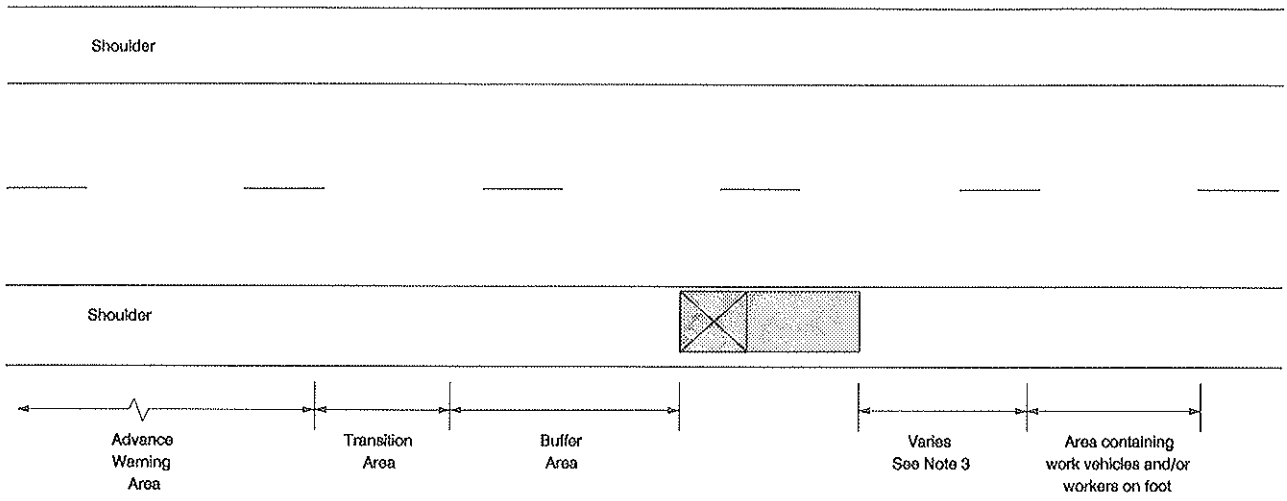
1. Advance warning traffic control devices to be in accordance with MUTCD, TCP, or other agency policies.
2. TMA vechicle should be equipped with an arrow panel operated in accordance with the MUTCD, TCP, or other agency policies.
3. Variable intervening distance to be selected from the following table:

| Prevailing Speed (mph) | TMA Vehicle Stationary (ft.) | TMA Vehicle Moving (ft.) |
|---|---|---|
| 60-65 | Appropriate buffer distance to be obtained from Table 4 or 5 based on policy decision defining shadow/barrier and impacting vehicle weights | |
| 50-55 | | |
| ≤ 45 | | |

Legend:

Vehicle with Attenuator

**FIGURE 3 Work area outside formal lane closure (not to scale).**




Notes:

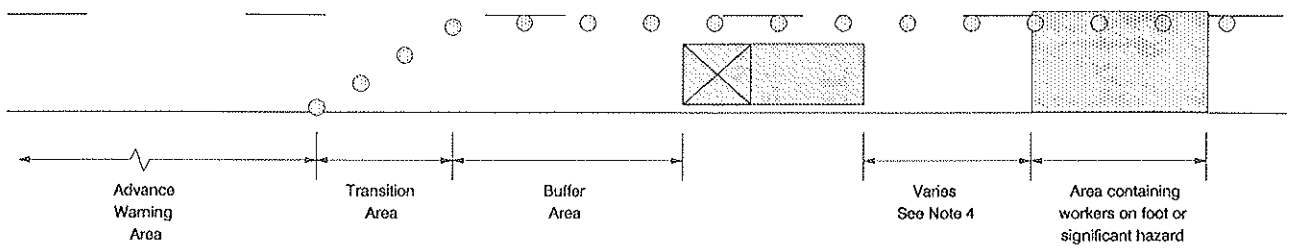
1. Advance warning traffic control devices to be in accordance with MUTCD, TCP, or other agency policies.
2. If the TMA vechicle is equipped with an arrow panel, it is to be operated in accordance with the MUTCD, TCP, or other agency policies.
3. Variable intervening distance to be selected from the following table:

| Prevailing Speed (mph) | TMA Vehicle Stationary (ft.) | TMA Vehicle Moving (ft.) |
|---|---|---|
| 60-65 | Appropriate buffer distance to be obtained from Table 4 or 5 based on policy decision defining shadow/barrier and impacting vehicle weights | |
| 50-55 | | |
| ≤ 45 | | |

Legend:

Vehicle with Attenuator

**FIGURE 4 Work area on shoulder without formal shoulder closure (not to scale).**




Notes:

1. Advance warning and transition area traffic control devices to be in accordance with MUTCD, TCP, or other agency policies.
2. If the TMA vechicle is equipped with an arrow panel, it is to be operated in accordance with the MUTCD, TCP, or other agency policy.
3. Minimum buffer area length to be in accordance with TCP or agency policy.
4. Variable intervening distance to be selected from the following table:

| Prevailing Speed (mph) | TMA Vehicle Stationary (ft.) | TMA Vehicle Moving (ft.) |
|---|---|---|
| 60-65 | Appropriate buffer distance to be obtained from Table 4 or 5 based on policy decision defining shadow/barrier and impacting vehicle weights | |
| 50-55 | | |
| ≤ 45 | | |

Legend:

Channelizing Device

Vehicle with Attenuator

**FIGURE 5 Workers on foot or significant hazard within formal lane closure (not to scale).**

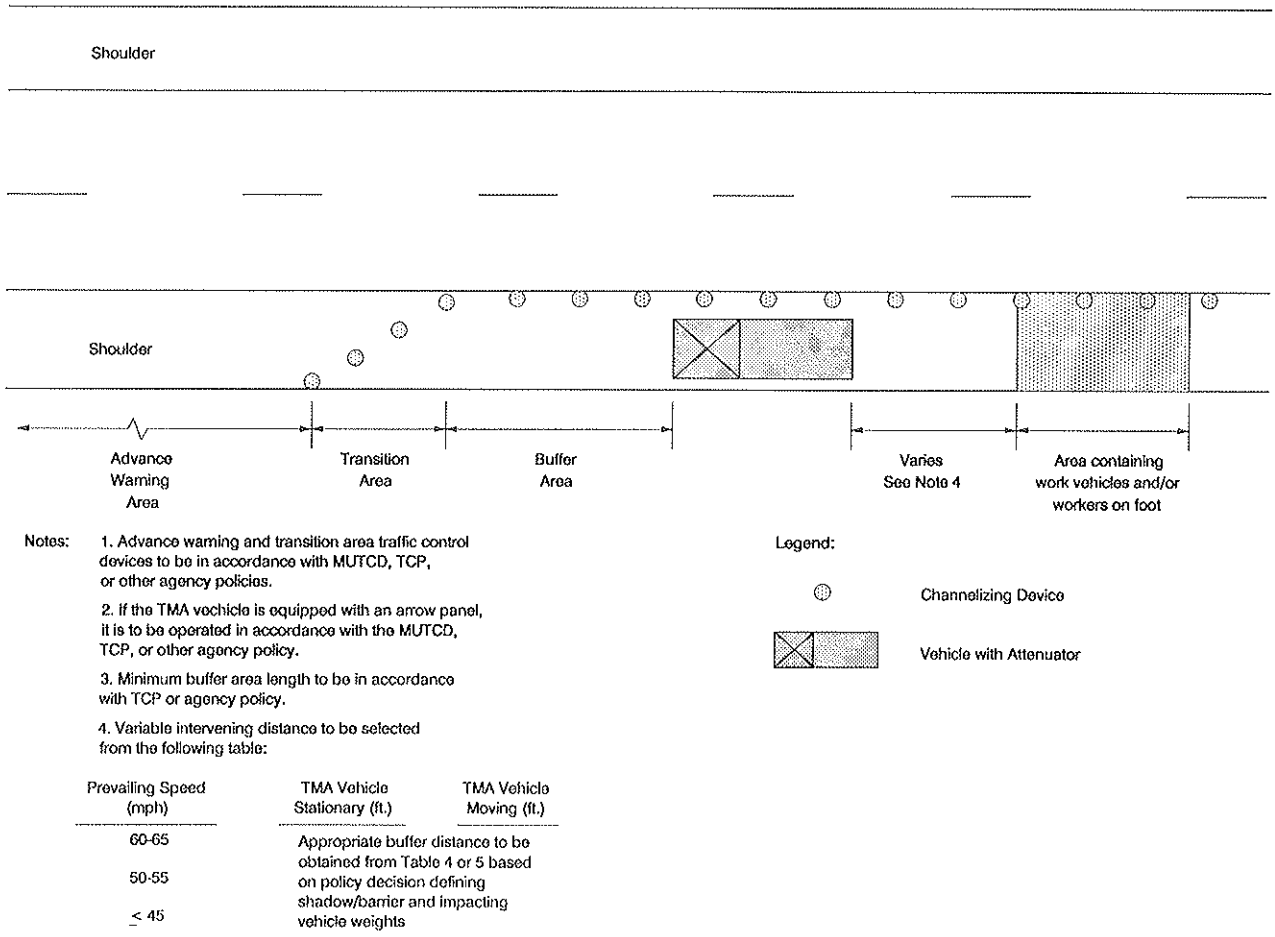



**FIGURE 6 Work area on shoulder with formal shoulder closure (not to scale).**

TABLE 4 ROLL-AHEAD DISTANCE FOR SHADOW VEHICLES

| Weight of Shadow Vehicle (moving)[b] | Prevailing Speed (mph) | Weight of Impacting Vehicle to be Contained[a] | | | |
|---|---|---|---|---|---|
| | | 4,500 lbs | 10,000 lbs | 15,000 lbs | 24,000 lbs |
| 10,000 lbs | 60-65 | 100 ft | 175 ft[c] | 225 ft | 275 ft |
| | 50-55 | 100 ft | 150 ft[c] | 175 ft | 200 ft |
| | ≤45 | 75 ft | 100 ft[c] | 125 ft | 150 ft |
| 15,000 lbs | 60-65 | 75 ft | 150 ft | 175 ft | 225 ft |
| | 50-55 | 75 ft | 125 ft | 150 ft | 175 ft |
| | ≤45 | 50 ft | 100 ft | 100 ft | 100 ft |
| 24,000 lbs | 60-65 | 75 ft | 100 ft | 150 ft | 175 ft |
| | 50-55 | 50 ft | 75 ft | 100 ft | 150 ft |
| | ≤45 | 50 ft | 75 ft | 75 ft | 100 ft |

[a]Weights of typical vehicles:

| | |
|---|---|
| Mid-size automobile | 2,250 lbs |
| Full-size automobile | 3,500 lbs |
| Loaded 3/4-ton pickup truck | 6,000 lbs |
| Loaded 1-ton cargo truck | 10,000 lbs |
| Loaded 4-yard dump truck | 24,000 lbs |

[b]Distances are appropriate for shadow vehicle speeds up to 15 mph

[c]Values suggested for inclusion on Figures 3, 4, 5, and 6.

TABLE 5 ROLL-AHEAD DISTANCE FOR BARRIER VEHICLES

| Weight of Barrier Vehicle (stationary) | Prevailing Speed (mph) | Weight of Impacting Vehicle to be Contained[a] 4,500 lbs | 10,000 lbs | 15,000 lbs | 24,000 lbs |
|---|---|---|---|---|---|
| 10,000 lbs | 60-65 | 50 ft | 100 ft[b] | 150 ft | 200 ft |
| | 50-55 | 25 ft | 75 ft[b] | 100 ft | 150 ft |
| | ≤45 | 25 ft | 50 ft[b] | 75 ft | 100 ft |
| 15,000 lbs | 60-65 | 25 ft | 75 ft | 100 ft | 150 ft |
| | 50-55 | 25 ft | 50 ft | 75 ft | 100 ft |
| | ≤45 | 25 ft | 25 ft | 50 ft | 75 ft |
| 24,000 lbs | 60-65 | 25 ft | 50 ft | 75 ft | 100 ft |
| | 50-55 | 25 ft | 25 ft | 50 ft | 75 ft |
| | ≤45 | 25 ft | 25 ft | 25 ft | 50 ft |

[a]Weights of typical vehicles:

| | |
|---|---|
| Mid-size automobile | 2,250 lbs |
| Full-size automobile | 3,500 lbs |
| Loaded 3/4-ton pickup truck | 6,000 lbs |
| Loaded 1-ton cargo truck | 10,000 lbs |
| Loaded 4-yard dump truck | 24,000 lbs |

[b]Values suggested for inclusion on Figures 3 through 6.

represent the researchers' views of the relative desirability of using a shadow or barrier vehicle (Table 2) or a TMA (Table 3) under a given set of circumstances compared with other circumstances. They should not be used as a basis for evaluating the relative merit of expending resources on providing shadow or barrier vehicles and TMAs compared with the merit of other projects or programs that may be in competition for the same resources.

TMAs have been available for several years, but their use in most states has been limited. As a result, there are no comprehensive guidelines or suggested application priorities. Soon after the study started, the researchers recognized that there was not an existing data base that would support a rigorous scientific analysis and that a comprehensive scientific study would require information derived from TMA use over diverse geographical areas and under a wide range of work zone types. Required data would include the number and severity of accidents (with and without TMAs) by work zone activity and some measure of the frequency of exposure and activities.

Although no scientific work plan was developed, it appeared obvious that developing an adequate data base would require the cooperation of a number of agencies, over an extended period of time, at a cost that would probably be measured in the hundreds of thousands of dollars—far beyond the budget available for this effort. In the meantime, because of the short-term need for a rational basis for assigning available units, this study was conducted.

The guidelines reflect the existing practices of the agencies contacted, the concerns expressed by field personnel who participated in the discussions, and the collective wisdom of the researchers and others (including agency representatives, other researchers, suppliers representatives, etc.) from whom comments were sought and received. Priorities based on scientific research would be desirable and ultimately will be developed. The researchers hope that the present effort will stimulate discussion toward that end, and believe that in the meantime the guidelines in their present form can be used appropriately as a policy formation and budgeting tool.

## REFERENCES


1. *Highway Design and Operational Practices Related to Highway Safety*. AASHTO, Washington, D.C., 1967.
2. J. B. Humphreys et al. *Identification of Traffic Management Problems in Work Zones*. Report FHWA-RD-79-4. FHWA, U.S. Department of Transportation, Dec. 1979.
3. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, 1978.
4. *Manual on Uniform Traffic Control Devices*. FHWA, U.S. Department of Transportation, 1988.
5. Energy Absorption Systems, Inc. *EASI-Reader*, Vol. 9, No. 2, Chicago, Ill., 1989.
6. J. W. Hall, and D. S. Turner. Stopping Sight Distance: Can We See Where We Now Stand? *In Transportation Research Record 1208*. TRB, National Research Council, Washington, D.C., 1989.
7. F. J. Tamanini. *Impact Attenuators: An Overview of Their Characteristics and Effectiveness*. Energy Absorption Systems, Inc., Chicago, Ill., undated.
8. C. E. Buth et al. *Truck Mounted Attenuators*. Report FHWA-TS-88-0018. Texas Transportation Institute, College Station, March 1988.
9. E. L. Marquis and T. J. Hirsch. *Texas Crash Cushion Trailer to Protect Highway Maintenance Vehicles*. Research Report 146-6. Texas Transportation Institute, College Station, Aug. 1972.
10. Energy Absorption Systems, Inc. *EASI Reader*, Vol. 5, No. 2, Chicago, Ill., 1985.
11. Energy Absorption Systems, Inc. *EASI Reader*, Vol. 4, No. 2, Chicago, Ill., 1984.
12. R. W. Anderson, *Transafety Reporter*, Nov. 1988.
13. Energy Absorption Systems, Inc. *EASI-Reader*, Vol. 7, No. 1, Chicago, Ill., 1987.
14. Energy Absorption Systems, Inc. *EASI-Reader*, Vol. 9, No. 1, Chicago, Ill., April 1989.
15. Increasing Work Zone Safety. *Public Works*, March 1989.
16. R. W. Anderson, *Transafety Reporter*, April 1989.
17. *Traffic Control Devices Handbook*. FHWA, U.S. Department of Transportation, 1983.


*Publication of this paper sponsored by Committee on Traffic Safety in Maintenance and Construction Operations.*