TRANSPORTATION RESEARCH

RECORD

No. 1312

Energy and Environment

Energy and Environmental Issues 1991

A peer-reviewed publication of the Transportation Research Board

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

Transportation Research Record 1312
Price: $28.00

Subscriber Category
IB energy and environment

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

Energy and environmental issues, 1991.
p. cm.—(Transportation research record, ISSN 0361-1981; no. 1312)
"A peer-reviewed publication of the Transportation Research Board."
ISBN 0-309-05120-7
1. Transportation engineering—Environmental aspects.
I. National Research Council (U.S.). Transportation Research Board. II. Series: Transportation research record ; 1312.
TE7.H5 no. 1312
[TD195.T7]
388 s—dc20 91-36611
[363.73'1] CIP

Transportation Research Record 1312

Contents

Foreword

Environmental considerations are becoming of increasing concern in the building and operation of transportation facilities. In recognition of the growing environmental problems that are exacerbated by pollutants emanating from transportation systems, new efforts are being directed at significantly improving air and water quality, preserving wetlands, and controlling the handling and disposing of hazardous materials. The papers in this Record deal with a variety of such environmental issues.

Under the general topic of environmental analysis, there are two papers, one dealing with preventing endangered brown pelicans from being struck by vehicular traffic on bridges, and the other suggesting an approach to improving the cooperation between state highway and environmental agencies in dealing with hazardous materials traveling on the right-of-way.

The second group of papers discusses interrelationships between transportation and air quality: (a) the ability of intersection air quality models to simulate carbon monoxide concentrations; (b) a model for environmental-friendly urban traffic flows; (c) highway vehicle emission inventories; (d) effects of trucks on air quality; (e) transportation, land use, and air quality sensitivity analysis; (f) adverse air quality impacts resulting from special events; (g) the effects of modal splits on air quality; (h) transportation-caused air pollution in urbanized areas in developing countries; and (i) a technique for including the costs of air pollution in transportation pricing policy.

Transportation noise continues to be a major problem in urban areas. The papers in this section discuss public attitudes toward transportation noise, the use of sound insulation for aircraft noise abatement, improving the aesthetics of noise barriers, the effect of road texture on traffic and vehicle noise, and the testing of single and parallel highway noise barriers.

The physicochemical phenomena of soils are discussed in two papers. The first paper is a synthesis of available data on electrokinetic soil processing in waste treatment. The second paper recommends site remediation by in situ vitrification. The final paper describes the field experience with electrokinetics at a superfund site.

In the area of emerging technology, there is a paper on altimetric sensing of ocean currents and its potential impacts on ship routings.

Mitigation of Traffic Mortality of Endangered Brown Pelicans on Coastal Bridges

L. KAROLEE OWENS AND RAY W. JAMES

Since the initial 1984 mortality, brown pelican mortalities on the P100 bridge have corresponded to a recovering brown pelican population that has increased markedly. The population in the vicinity of the bridge peaks in the late summer and early fall, but wintering populations in the vicinity of the bridge have also shown a steady increase in recent years. The majority of the brown pelicans forage south of the bridge during the day and cross to the north of the bridge in the later afternoon and early evening. Strong north winds, especially when accompanied by rain or mist, often result in brown pelicans being forced down on the deck where they are struck by cars. On the basis of observed pelican behavior and limited wind tunnel testing, turbulence above the bridge roadway is suspected as a causal factor in the observed mortality, although the actual significance of the turbulence cannot be inferred without somewhat speculative observation of the flight and behavior of brown pelicans in the vicinity of the bridge. Brown pelican mortalities are likely to occur during any strong north winds, and the passage of cold fronts accompanied by rain increases the probability that brown pelicans will be killed on the bridge. Age, experience, and physical condition of the individuals being killed are apparently not related. Few brown pelicans fly under the bridge even without strong north winds. There are still undetermined factors related to turbulence under the bridge, or possibly related to sound, that deter brown pelicans from flying under the bridge. Traffic control measures, including better warning signs and, in particular, decreasing the speed limit during weather conditions are most likely to result in the decrease of brown pelican mortalities; placing telephones at each end of the bridge allows motorists to report birds or accidents on the bridge. This simple and economical approach should result in a mitigation of brown pelican traffic mortalities. Should these measures prove inadequate, further research will be necessary to determine whether brown pelicans can be induced to fly over the bridge to avoid air turbulence close to the deck, to fly higher under the bridge, or to roost south of the bridge. Bird count data indicate that the Texas brown pelican population is increasing. Although an increasing brown pelican population may result in the eventual delisting of this subspecies, the hazard to motorists may increase.

The eastern brown pelican is a large bird weighing about 7.5 lb with average body length of 4 ft and average wingspan of 6.5 ft. Figure 1 shows a sketch of a brown pelican. It is a coastal resident seldom straying inland from its preferred saltwater shores. The brown pelican is capable of flight speeds of 14 to 35 mph and usually flies with slow wing beats close to the water. It forages by diving and, while capable of lifting off from a horizontal surface without a headwind, it commonly takes off into the wind to increase airspeed and to gain lift.

Texas Transportation Institute, Texas A&M University, College Station, Tex. 77843-3136.

The Texas eastern brown pelican population once numbered in the thousands. A recent study (*1*) provides some details about the history of the Texas population. As many as 5,000 pairs have been reported nesting on the Texas coast from the late 1800s until about 1920 (*2*). An early decline in the 1930s was a result of persecution by fishermen (*3–5*). Although legislation was enacted in 1939 to protect brown pelicans from being shot and from the destruction of their nests and eggs, another serious decline became apparent in the early 1950s. By 1962, no brown pelicans were reported in former areas of concentrations of wintering birds, and they had disappeared from former breeding areas. This second serious decline has been attributed to severe weather conditions, disease, and especially to exposure to chlorinated hydrocarbon pesticides (*6*). The Texas subspecies (*Pelicanus occidentalis carolinensis*) was placed on the endangered species list of the U.S. Department of the Interior in 1971.

Historically, the brown pelican has nested along the Texas coast from Galveston Bay to Cameron County, but from 1985 to 1988 brown pelicans nested only on Pelican Island in Corpus Christi Bay. In 1989, breeding colonies expanded to six sites. Brown pelicans winter along the Texas coast from Galveston to Cameron County. It has been estimated that 96 percent of the Texas brown pelican population use the lower Laguna Madre in winter.

The P100 bridge, or Queen Isabella Causeway, is a 2.4-mile-long, four-lane bridge connecting Port Isabel with South Padre Island, Texas. The bridge has a center span rising approximately 84 ft above the Intracoastal Waterway. The bridge was completed in 1974. One study conducted in 1984 and 1985 in the lower Laguna Madre in connection with a proposed transmission line to cross the lower Leguna Madre indicated the area of greatest brown pelican activity was in the vicinity of the causeway with a majority of the observations in the August to October period when the Texas population is supplemented by immature brown pelicans from Mexico. It is thought that these Mexican brown pelicans initiated the recovery of the Texas population.

In September 1984, the first brown pelican mortality on the bridge was recorded. Considerable public concern, already sensitized by the threat to brown pelicans by the proposed transmission line, was expressed after subsequent and increasingly frequent brown pelican mortalities occurred on the bridge.

Several causal factors were initially suggested. In conjunction with the increasing population size, the brown pelicans'

FIGURE 1 Brown pelican (*Pelicanus occidentalis carolinensis*).

foraging and roosting habits in relation to the bridge were implicated. Also, the apparent connection between the passage of cold fronts accompanied by strong north winds possibly resulting in air turbulence around the bridge was proposed as a precipitating factor in the brown pelican mortalities on the P100 bridge.

It is the recovering brown pelican population wintering in the lower Laguna Madre that has come into conflict with the P100 bridge. A growing brown pelican population with an increasing number of nest sites potentially expanding into the lower Laguna Madre will increase the likelihood of fatal encounters on the P100 bridge year round.

OBJECTIVES

The objectives of this study were to (a) identify the factors influencing the presence and resulting deaths of the brown pelican on the P100 bridge, and (b) suggest ways to mitigate those factors.

METHODS

General background information was gathered from a variety of sources, including the following:

1. Letters, memorandums, etc., from correspondence leading up to the funding of this project. These included communications from Texas State Department of Highways and Public Transportation, Texas Parks and Wildlife, U.S. Fish and Wildlife Service, Sierra Club, Bird Rescue, concerned citizens, and newspaper articles.
2. Literature searches, including two computer searches. The initial search explored the literature for information on brown pelicans and their behavior, flight, and aerodynamics, and road kills of birds on bridges and highways. The second search explored the literature with reference to the reaction of birds and wildlife to sound, including auditory perception, hearing, noise, ultrasound, and infrasound.
3. Other information about the brown pelican population was obtained from Audubon Christmas bird counts, breeding bird surveys, ornithological newsletters, bird rescue records, National Oceanographic and Atmospheric Administration weather data, and unpublished reports from Texas Parks and Wildlife.
4. Collection of dead brown pelicans for necropsy.
5. Behavioral observations at the bridge—brown pelican counts were made during four trips to the study site. These visits included an initial survey of the study site (January 12 to 14, 1989) and visits timed with the passage of strong cold fronts (February 2 to 17, 1989, October 19 to 22, 1989, and December 6 to 13, 1989). A brief visit was made to the study site on April 6, 1989, after the Area I Research Committee Meeting in Brownsville to determine if any brown pelicans were present. Another brief trip was made to Port Isabel for a meeting called by Gary Waggerman with the brown pelican volunteers. Observations included several counts of brown pelicans in the general vicinity of the P100 bridge. These observations were made from several vantage points north and south of the bridge on South Padre Island including the state fishing pier and Isla Blanca Park. Observations were also made at Queen's Point Marina, Port Isabel channel, and Long Island. Counts were made of brown pelicans in Laguna Madre, the Brownsville ship channel, and the Gulf of Mexico. Brown pelicans were observed crossing the bridge in late afternoon and early evening. Most observations of brown pelicans crossing the bridge during strong north winds were made from Queen's Point or by driving back and forth across the bridge. One videotaping session was done on the north side of the west end of the bridge.
6. Videotaping of brown pelicans at the bridge.
7. Wind tunnel tests, including videotaping, conducted on two scale models (72:1 and 16:1) of the bridge July 31 and August 1, 1989.
8. Correspondence with personnel in other brown pelican states.

RESULTS AND DISCUSSION

Literature Search

There is little information in the published literature concerning the Texas brown pelican population. Some information was obtained concerning the results of banding studies on brown pelicans in general and some information on flight speeds. There was no information in the literature on road kills of brown pelicans, or even birds in general, on highways or bridges.

The literature search on bird responses to sound yielded little information, as the key words used retrieved papers concerning sound produced by birds and the use of sounds, i.e., bird distress calls and propane cannons, for animal damage control. Only two papers recording bird reactions to sound levels were located. This research involved investigations of the acoustic irritation thresholds of Peking ducks, other domestic and wild fowl (*7*), and ringbilled gulls (*8*). It was found that hungry Peking ducks were discouraged from taking food placed in a low-frequency sound field at 100-db intensity. Additionally, in a report prepared by LGL Limited, environmental research associates for Arctic Gas, it was demonstrated that snow geese were disturbed by sounds made by gas compressors (*9*).

Status of the Brown Pelican Breeding Population

Figure 2 shows the decline and recovery of the brown pelican breeding population, the date of completion of the P100 bridge, and the time of the first recorded brown pelican mortality on the span. The breeding population had risen from nonexistence in 1964 to 230 in summer 1984 just before the first recorded mortality in September 1984. In the 1989 breeding season, six colony sites were used and the population increased markedly as a result of successful nests on five of these sites. Pelican Island in Corpus Christi Bay had 565 nesting pairs that produced about 900 young. This is an increase from the summer 1988 nesting season which had 350 nesting pairs producing 575 young. Other sites included Sundown Island (50 adults and 25 young), Second Chain (10 adults, 10 young), Steamboat Island (12 adults, 6 young), and Dressing Point (25 adults, 22 young). Flooding caused by a hurricane aborted nesting attempts by 14 adults at Cedar Lakes. This new nest site was the result of an attempt to establish a nest colony of brown pelicans at San Bernard National Wildlife Refuge. Thus, a total of 676 nesting pairs produced about 963 young in Texas during the 1989 breeding season.

Although these colony expansions were northward along the coast from Pelican Island, there were indications (Mike Farmer, Audubon Society warden, personal communication) that brown pelicans may have been nesting in Laguna Madre in summer 1990, because adults were seen carrying sticks, a behavior that may be associated with nesting intentions. Two historical sites mentioned by Oberholser (*10*) included a mud dump at Port Isabel and Brazos Santiago Pass, both used in 1927. Steamboat Island used in summer 1989 was last used for nesting by brown pelicans in 1931. Recolonization of historical nesting sites could result in a breeding population in the vicinity of the P100 bridge.

Status of the Brown Pelican Wintering Population

Figure 3 shows the recovery of the brown pelican wintering population along the Texas coast as indicated by the Aubudon Christmas bird counts (CBC) from 1950 to 1988 (*11*). The results of all of the Audubon Christmas bird counts for Texas for Christmas 1989 (90th CBC) will not be accumulated until April or May. However, there was an early report of 104 brown pelicans counted in the coastal tip of Texas count, which includes the area around the study site in a 15-mi-diameter count circle centered (26 deg 02 min N, 97 deg 14 min W) on the Brownsville ship channel. The coastal tip of Texas count was initiated during the Christmas 1986 (87th CBC) count and a total of 12 brown pelicans were seen. Subsequent count totals were 55 in 1987 (88th CBC) and 88 in 1988 (89th CBC), indicating a continuing trend of an increasing winter population.

Thus, there was an increase in the total number of brown pelicans in spite of the severe cold weather in December 1989 when a number of pelicans died. Necropsies were done by the Texas Veterinary Medical Diagnostic Laboratory on 3 of the 18 dead brown pelicans found on Dressing Point. Necropsy results indicated that the pelicans had frozen to death. Another seven brown pelicans carcasses were found on Aransas NWR and at least two others were found along with a dead white pelican along the coast.

Chronology of Brown Pelican Mortalities on the P100 Bridge

Table 1 presents the brown pelican mortalities that were documented in various correspondence and in the course of this study. These deaths have occurred from September through

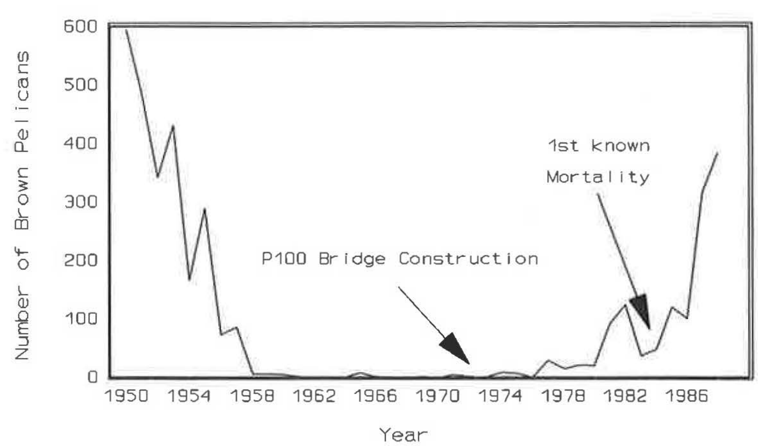


FIGURE 2 Decline and recovery of Texas brown pelican population—Audubon Christmas bird counts, 1950–1988.

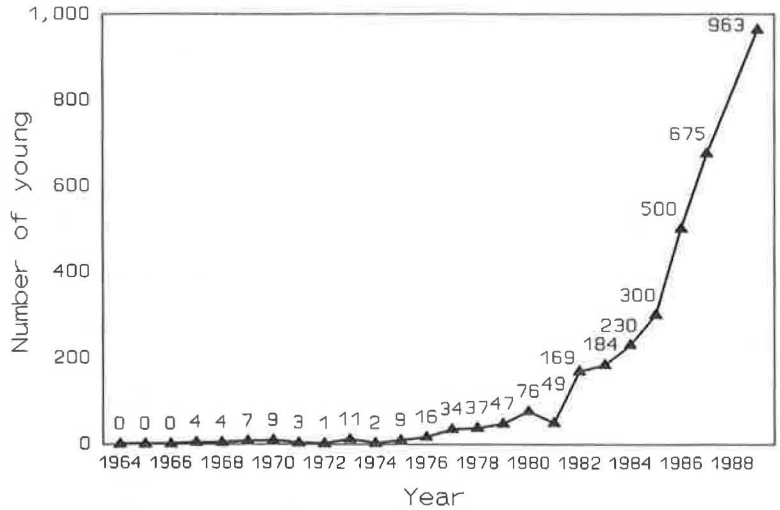


FIGURE 3 Recent historical data for Texas brown pelican wintering population—numbers of young produced, 1964–1989.

TABLE 1 CHRONOLOGY OF KNOWN BROWN PELICAN MORTALITIES ON THE P100 BRIDGE

Date	Number	North Wind (Yes or No)	Wet/Dry	Bridge Lane (North or South)
1984				
19 September[1]	1	Y	W	S
1986				
12 October	3	Y	W	S
13 October	2	Y	W	-
12 November	2	Y	W	S
25 November	1	Y	W	S
1987				
21 January[2]	1	Y	W	N
15 December	2	Y	D	-
1988				
5 February	3	Y	W	-
1989				
6 January[3]	1	N	D	-
18 October[4]	1	Y	D	
16 November	1	Y	D	S
29 November	2			S
2 December	2			S
7 December	2	Y	D	S

1 Tropical Storm Edouard; 19 inches of rain on this date.
2 Upper air disturbance in Northern Mexico.
3 Strong winds but not from the north.
4 An injured Brown Pelican also recovered.

early February. Nine pelicans were killed in the 1986–1987 fall and winter season. Another five pelicans were killed during the 1987–1988 winter season. Only one brown pelican was reported killed during winter 1988–1989 because of the mildness of the winter, which had few fronts with strong north winds passing through the area. Eight brown pelicans were killed from October to December 1989. One recovered alive from the bridge had an irreparably broken wing. After December 1989, no brown pelicans were reported killed on the P100 bridge because the remainder of the winter was mild with few strong fronts.

All but one of the documented deaths occurred during strong north winds, although all of these winds were not associated with the passage of cold fronts. Tropical disturbances, upper air disturbances, or other causes resulted in strong north winds on at least three occasions. Cold fronts accompanied by rain increased the probably of occurrence of brown pelican deaths. Thirteen brown pelican deaths occurred during wet fronts (seven dates) and seven deaths occurred during dry fronts (five dates). All but one of the documented deaths occurred in the eastbound (south) lanes of the bridge.

An examination of the bird rescue reports to the U.S. Fish and Wildlife Service (FWS) from 1983 to April 1988 revealed only four records of brown pelicans hit on the causeway in 1987. There were no other references to brown pelicans that were recovered by bird rescue, dead or alive, from the P100 bridge. These four records included two brown pelicans received from the Coastal Studies Laboratory on August 2, 1987, but which were killed on the causeway in October 1986. Two carcasses were given to Dr. Pauline James in the Biology Department at Pan American University for study skins. The other two brown pelican carcasses were received from Ann Grefke in December 1987, but there was no notation of the disposition of these carcasses. There was also a reference to the brown pelican that broke its wing in a collision with the transmission line and that was later acquired by Bird Rescue from Colley's Fishing Service. This brown pelican was sent to the Victoria Zoo, because it could not be rehabilitated to the wild. Although the Victoria Zoo later lost all its brown pelicans to disease, the injured brown pelican recovered from the P100 bridge on October 18, 1989, initially treated by Bird Rescue, had been transported to the Gladys Porter Zoo in Brownsville where the wing was amputated.

It is virtually impossible to recover carcasses from the bridge in good condition, unless they are picked up immediately after being struck by a car. The birds are struck almost immediately after they land on the bridge and following traffic renders the carcasses damaged beyond usefulness to the study. A coast guardsman reported seeing a live brown pelican and a dead brown pelican in the opposite lane, but by the time he could turn around and get back to the birds, the live pelican had also been killed.

Necropsies of Brown Pelicans Killed on the P100 Bridge

Only one of the brown pelican carcasses removed from the P100 bridge before the initiation of this study was recovered for necropsy. This carcass had been stored in a freezer at the Pan American University Coastal Studies Laboratory on South Padre Island. This bird had been killed on the P100 bridge, probably sometime during winter 1988–1989, but there was no other information on this individual.

Another brown pelican was found dead on the jetty at the Brownsville ship channel on February 6, 1989. This bird had an injured wing that may have been the result of collision with a power line or the result of a gunshot wound. Although this was not a bridge mortality, it was sent for necrospy to get information about the general condition of brown pelicans in the area. Along with the carcass from the Pan American Laboratory, this carcass was sent to Dr. Nancy Thomas of the National Wildlife Health Center Resource Health Team in Wisconsin in February 1989. (We still do not have the results of these necropsies).

Two brown pelicans were retrieved from the south lane at the curve in the causeway on December 7, 1989. One was an adult in winter plumage, and the other was a first-year immature. The carcasses appeared to be fresh, having probably been killed between 7:30 and 10:00 p.m. These two were sent by Continental Airlines to the Texas Veterinary Medical Diagnostic Laboratory at Texas A&M University the next morning. The final necropsy reports indicated that one of these was male and one was female. The birds were in good flesh, and no lesions were noted except those that were the result of trauma. Numerous flukes were found in the small intestine, but there was no indication that these could be a contributing cause of death. Insecticide screens of both livers were negative, and lead levels were less than 1 ppm.

Four other brown pelicans that had been killed on the bridge in earlier cold fronts and that had been stored in the freezer at PAU Coastal Studies Labortory were also necropsied by the Texas Veterinary Medical Diagnostic Lab. These carcasses were badly smashed. Consequently, there was little information gained from these carcasses other than the observation that they had numerous parasitic worms. The insecticide screens were also negative, and lead levels were less than 1 ppm.

Observations of Banded Brown Pelicans

Bird Rescue listed only one of the brown pelican carcasses they disposed of as having an FWS aluminum leg band. Only two brown pelicans having leg bands were observed during this study. On February 3, 1989, there was a winter adult on the breakwater at Queen's Point with a band on its right leg that consisted of a top narrow black stripe followed by a yellow band, another black stripe, and a lower yellow band that was wider than the top yellow band.

On December 7, 1989, leg bands were observed on a winter adult brown pelican perched on the transmission lines south of the state fishing pier. There was an aluminum band on the right leg and a colored band on the left leg. The band appeared greenish with no stripes or other markings. This band may have been one of the red bands put on brown pelicans in Mexico, as these bands tended to fade and could appear greenish. This bird exhibited a bright orange-red color at the base of the pouch on the neck and reddish bill. This coloration was not observed on any of the other pelicans with it on the transmission lines, nor was it noted on any other brown pelicans during the course of this study.

Wind Tunnel Testing

Two series of wind tunnel tests were accomplished in the low-speed wind tunnel at Texas A&M University's Easterwood Airport research facility. The objective of the tests was to document the flow regimes around the roadway to support explanations of the observed pelican behavior in the vicinity of the bridge in times of strong north winds. Some fundamental questions that stimulated the wind tunnel studies were the following:

- Does turbulence below the deck cause pelicans to try to fly over the bridge, rather than under it, in times of strong north winds?
- Does turbulence above the deck affect the pelican's flight above the deck?
- Is turbulence above the deck caused by the railing, the median barrier, or the superstructure and roadway?
- Is there aerodynamic evidence to support a theory that the pelicans might be seeking shelter behind the safety shape median barrier?

Testing of the 72:1 scale model of Spans 36 to 39 took place July 31, 1989, beginning at approximately 8:30 a.m. with installation of the model into the test section and with removal of the model in the afternoon. The larger 16:1 model was installed at approximately 4:00 p.m. Videotape records were made by tunnel staff during the morning and in the afternoon, videotape records were made with the TTI camera, also. Still photographs, including both prints and slides were taken. Smoke tests were conducted at 0 degrees (perpendicular to the centerline of the model) and at various angles simulating N–NE winds at angles up to 45° to visually observe the flow pattern and the presence of regions of turbulence. Dynamic pressure probe measurements were obtained, sweeping the probe along vertical lines just behind the downwind railing (approximately 1.5 in.), which is a continuous trace, and at the median barrier and near the upwind railing, resulting in interrupted traces. These measurements are made in both Spans 36 and 38 at a free field dynamic pressure of 5.0 pounds per square foot (psf). In addition, there is a sweep at 3.0 psf on the record for Span 36 (for checkout only).

Testing of the 16:1 scale model of Span 38 began at approximately 4:00 p.m. on July 31, 1989, and was concluded by approximately 2:00 p.m. on August 1, 1989. Video records were made with both the tunnel camera and the TTI camera. Still photographs were also taken. Smoke tests were conducted at 0 deg and at various angles up to approximately 45 degrees, simulating NE winds. NW winds were not simulated, and because the deck has a downward grade to the east in the region modeled, some difference could be expected between NE and NW winds. NE winds have a negative angle of attack, while NW winds have a positive angle of attack. This effect was not thought to be significant, however. Dynamic pressure probe data were taken along a vertical line just behind the downwind railing, and along a vertical line through the median barrier.

In the afternoon, smoke tests were conducted after removal of a portion of the upstream railing to determine the role played by the upwind railing in the presence of turbulence on the deck. Later a V-shaped leading edge fairing was fabricated to modify the leading edge, and further smoke tests were conducted.

A zone of turbulence and reversed flow was observed above the bridge deck. It is visualized in the smoke tests and may be inferred from the dynamic pressure data. The extent of this zone is estimated best from the dynamic pressure measurements. On the 72:1 model, the deck height is approximately 12.25 in. from the datum (floor), and the zone of turbulence extends up to approximately 14.5, 15.5, and 16 in., respectively, over the upwind railing, the median barrier, and the downwind railing in Span 38. In Span 36, where the deeper steel girders are present, the respective heights are 15.5, 15.75, and 16.75 in. Subtracting the deck height of 12.25 in. from these distances and multiplying by the scale factor of 72 (or 6 ft = 1 in.), it is concluded that height of the turbulent zone above the deck is from 13.5 to 27 ft above the deck on the full-scale bridge. From tests of the 16:1 model, a value of approximately 12 ft is obtained with somewhat higher confidence. This latter number is more consistent with the smoke test observations. Within this region the smoke does not exhibit a static trail but is buffetted significantly. At the level of the deck, the smoke is generally blown upwind in a flow reversal, possibly suggesting horizontally oriented vortices above the deck caused by the bluff leading edge of the bridge—the girders and parapet wall. At 45 degree angles simulating a NE wind, the smoke is blown generally parallel to the traffic flow at the deck level.

This observed zone of turbulent flow extends some considerable distance downstream of the bridge. The horizontal extent of the region of turbulence was not quantified, but it would appear to be at least one deckwidth downstream of the structure, and it probably extends much further than this. It is possible that the pelicans search out this region, as the headwind effect is considerably reduced–a theory which might help to explain why the birds do not fly beneath the bridge during strong winds. However, this theory is at odds with the theory that the pelicans have difficulty flying in the turbulence above the bridge deck and are forced to land on the deck. This second theory appears to be more in line with the observed pelican behavior. Approaching from the south, the pelicans try to climb to an altitude sufficient to clear the bridge and traffic—approximately at the levels of the tops of the light standards. Upon reaching a point of sufficient height above the bridge, the pelicans appear to try to glide or fly across the bridge. It is then theorized, on the basis of reported observations, that the pelicans glide or fly into the region of turbulence above the deck, and disoriented or buffeted by the turbulence, light on the bridge deck rather than flying clear of it. While this theory still cannot be supported by more than observation and knowledgeable interpretation, it still is believed to offer the best explanation of the observed behavior.

The space beneath the spans is observed to be largely free of turbulence. While some turbulence undoubtedly exists near the bottom surface of the deck and near the planes of the piers, there appears to be no obvious aerodynamic explanation for the perceived reluctance of the pelicans to fly beneath the bridge during high winds perpendicular to the bridge. One related question that may not have been resolved is whether winds at angles to the bridge produce such turbulence. The models were tested at beta angles up to approximately 45 degrees, but a comprehensive survey of the air beneath the bridge was not attempted for such configurations.

It is also clear that the size of the region of turbulent flow is only partly influenced by the concrete median barrier. Smoke tests clearly indicate strong reverse flow behind the barrier in the downwind lane. The differences in the flow pattern around the steel girders and concrete girders are observable, but not thought to be significant. In the presence of such a strong flow, it seems unlikely that the pelicans would be attempting to seek shelter behind the barrier.

In an attempt to determine how much of the turbulence was caused by the presence of the parapet wall and railing, a section of the parapet wall and railing was removed from the model. In subsequent smoke tests, it appeared that the size of the turbulent region was visibly reduced, indicating that a significant fraction of the turbulence in the upwind lane is caused by the presence of the parapet wall and railing. However, a region of turbulent flow still remains. It must be concluded that the region of flow reversal and turbulent flow cannot be eliminated even by removal of the median barrier and railing. The flow is, to a large extent, associated with the roadway.

Finally, a modification was made to the shape of the bluff leading edge of the bridge by fabricating a foam fairing, having a 90 deg nose angle and approximately 2.75-in.-long sides, which was then glued to the outer girder of the 16:1 model. It was difficult to assess the effects of the fairing; however, it appeared that the size of the zone of turbulent flow was reduced somewhat in the upwind lane, especially in that portion of the model where the parapet wall and railing had been removed. In the region where the railing and parapet wall had not been removed, the fairing appeared to reduce the height of the turbulent region slightly, but not as noticeably as in that portion of the model where the parapet wall and railing had been removed.

The region of flow reversal, when penetrated by a brown pelican flying upwind, would be perceived by the bird as a wind shear, a suddenly encountered change in airspeed, along with a suddenly decreased angle of attack. The result will be a sudden and rapidly increased rate of descent. Field observations indicate that the birds approaching this region will approach in an orderly formation that degenerates into a confused group at a certain, clearly defined point downwind of the bridge. It may be inferred that the birds are encountering this wind shear and are making large corrections to the right or left to try to avoid the suddenly encountered downdraft. Subsequently, they climb to a higher altitude by a series of parallel traverses downwind of the bridge until they again try to fly above the bridge. Speculatively, the birds that are killed on the deck have probably been forced down onto the deck by this wind shear when they enter the region of flow reversal and turbulence at too low an altitude or when they fail to turn back to gain further altitude. This hypothesis raises the question of whether age of the bird, and in particular flight experience, might correlate with the mortality, with younger, less experienced birds suffering higher mortality rates. Because of the difficulty of collecting physical specimens, this question has not been resolved.

Correspondence With Other Brown Pelican States

A form letter describing the research on the brown pelican problem on the P100 bridge and asking for information on similar problems with birds on bridges was sent to state wildlife agencies and to conservation organizations in the 10 coastal states having brown pelican populations and Puerto Rico. These states included Alabama, California, Georgia, Florida, Louisiana, Mississippi, North Carolina, Oregon, South Carolina, and Washington. No response was received from either state agencies or conservation organizations in California, Washington, Puerto Rico, and South Carolina. Only two responses came directly from the original contact. A number of the original contacts referred our inquiry to 13 other people and 8 of these referees responded. I subsequently wrote three additional contacts suggested by responders, none of whom have responded to my inquiry. A total of 46 contacts have been made directly or indirectly with a total of 10 responses.

The following is a summary of responses:

- Alabama: There is a large nesting colony of brown pelicans in Mobile Bay but they have not noticed any problems with bridges. However, 10 percent of those recovered have been killed as a result of accidents with cars. There was no other information on these traffic mortalities.
- Georgia: They know of no problem with brown pelicans being killed on bridges in Georgia. They have lost some pelicans to collisions with transmission lines.
- Louisiana: They are not aware of any problems with brown pelicans, which are restricted to coastal barrier islands removed from transportation corridors.
- Mississippi: They are not aware of any problems with brown pelicans on bridges. The potential exists, however, because U.S. Highway 90 runs along the coast.
- Oregon: They are not aware of any problems with avian species on highways or bridges, although the South Slough Bridge of Coos Estuary parallels a Pacific Power and Light power transmission line on the southern Oregon coast. A persistent seabird and waterfowl (brown pelicans were not mentioned) mortality is associated with this site, but the mortality is thought to be caused by collision with the power lines after gaining altitude necessary to clear the more visible bridge. This persistent mortality has apparently increased since the power lines were relocated, evidence that the mortality is associated with the power lines rather than the bridges.
- Florida: We received three responses from Florida. (a) There is a problem with royal terns being killed on the causeway bridges between the mainland and Sanibel Island. A year-long study of road kills by the Sanibel-Captiva Conservation Association yielded an estimate of two brown pelicans having been killed along with 102 royal terns, 24 seagulls, 4 anhinga, and 2 cormorants. Brown pelicans commonly fly under the bridge and also feed directly under the spans. (b) A response from the Florida Audubon Society indicated they were not aware of any problems with brown pelican mortalities on bridges. (c) A response from the Florida Game and Fresh Water Fish Commission describes a problem with Least terns and black skimmers being killed on a causeway to St. George Island near Apalachicola. The only breeding colony of brown pelicans in northwest Florida is located 1 km south of the bridge, but it was estimated that no more than 10 pelicans have been killed on the bridge since 1986. However, there have been no consistent surveys for brown pelicans on bridges. They have initiated some traffic control measures; however, results of traffic control measures are difficult to assess because the nesting population increased by 70 percent as a result of habitat manipulations. The percentage of

the adult tern population killed decreased from 18 to 9 percent but the absolute numbers of adults and chicks killed increased. This year, they will reduce the speed limit during the nesting season, install flashing lights and new speed limit signs, and toll booth operators will pass out informational leaflets.

- North Carolina: A response from the North Carolina Wildlife Resources Commission (Charles Fullwood) indicated they were not aware of a problem with brown pelican mortalities on bridges. However, as a result of a response from the State of North Carolina Department of Transportation, it was determined that there have been a number of brown pelican mortalities on the Bonner Bridge across Oregon Inlet between Pea Island and Hatteras Island. The limited information on their problem and the descriptions of the Bonner Bridge indicate that there may be strong similarities to the brown pelican problem on the P100 bridge. This information about the Bonner Bridge problem was conveyed to the North Carolina Wildlife Resources Commission. The Coastal Endangered Species Project leader, Thomas Henson, was instructed to investigate and document brown pelican mortalities on the Bonner Bridge. It would be worthwhile to maintain contact with North Carolina to find out the results of their investigation.

A parallel survey of transportation officials in states having brown pelican populations was also conducted. The following is a summary of the responses:

- Alaska: Karl F. Mielke, Chief Bridge Engineer for Alaska Department of Transportation and Public Facilities, reports that brown pelicans are not thought to live in Alaska. There have been no reports of brown or white pelicans being involved in road kill incidents. They have had some reported road kill problems with bald eagles, but their most serious wildlife-traffic problem is with moose.
- Alabama: No response was received.
- California: James E. Roberts, Chief of Structures Division at California Department of Transportation, responded that no known incidents of pelican-automobile conflicts were known to him. In Southern Califonria, brown pelicans nest on Channel Islands off the coast and feed along the coastline. There are no highway bridges between the coast and the Channel Islands. In Northern California, there is a major population of pelicans along the remote cliffs of Northern Mendocino County where there are no roads. Another population of brown pelicans occurs in the heavily developed Huntington Beach area but does not come into conflict with highway traffic.
- Florida: No response was received.
- Georgia: Mr. Paul V. Liles, Jr., State Bridge Engineer for the Georgia Department of Transportation, responded that their maintenance personnel, bridge personnel, and area engineers have not observed problems with brown pelican traffic mortality.
- Louisiana: Mr. W. L. Haymon, District Administrator in Lake Charles, and Vincent Pizzolato, Public Hearings and Environmental Engineer for Louisiana Department of Transportation and Development, report no known instances of automobile-pelican collisions.
- Mississippi: Mr. W. K. Magee, Environmental Design Engineer, Mississippi State Highway Department, reports that neither he nor the district engineer responsible for coastal counties has any knowledge about incidences of traffic mortality of brown pelicans.
- North Carolina: Mr. L. J. Ward, Manager of the Planning and Research Branch, Department of Transportation, reported that collisions with automobiles do cause some brown pelican deaths on the Hebert C. Bonner Bridge across Oregon Inlet, the NC 12 link between Pea Island and Hatteras Island. Followup discussions with various NC DOT personnel revealed that the extent of the mortality at that site, while not formally documented, may exceed the mortality rate observed at the study site in Texas. The brown pelican population in North Carolina is not endangered, and the mortality, even if more severe than in Texas, is not so significant because the higher population of brown pelicans in North Carolina is not listed as endangered. The Bonner Bridge constructed in 1962 is 2.44 mi long with three 180-ft-long main spans providing a 66-ft vertical navigation clearance. The two-lane roadway is 33.3 ft wide with 31 ft between railings. The superstructure is constructed of prestressed concrete girders in the minor spans and plate girders in the main spans and a 7.25-in.-thick reinforced-concrete deck. In many of these respects, the bridge is similar to the P100 bridge. The bridge is oriented generally north-south, whereas the P100 bridge is oriented east-west. Prevailing winds, however, may generally have similar relationship to the two bridges in spite of their different orientations. In summary, a similar mortality situation with similar causes cannot be ruled out.
- Oregon: No response was received.
- South Carolina: Mr. Ed Frierson, a biologist with South Carolina Highway Department, had no record of traffic mortality of brown pelicans.
- Washington: No response was received.

Chronology of Brown Pelican Presence Near the P100 Bridge

There is some evidence that some brown pelicans remain south of the bridge at night at least some of the time. During an aerial survey on October 26, 1988, Gary Waggerman of TP&WD observed 46 brown pelicans roosting on the CP&L transmission lines next to the west end of the old causeway. Initial observation was at 6:57 p.m. Ten min later, there were 32 brown pelicans on the wires; at 7:20 there were 45 brown pelicans on the wires. Roosting on the wires was later confirmed by storm troopers, a group of volunteer brown pelican observers organized by Waggerman, after dark and before daylight in the same 24-hr period. Similar observations were made October 20–21, 1989. Forty-two brown pelicans were seen perched on the transmission lines on the south side of the west end of the old causeway after dark (7:20 to 7:45 p.m.); 18 brown pelicans were perched on the transmission lines south of the state fishing pier. Before sunrise the next morning, there were still nine brown pelicans roosting on the wires by the state fishing pier, and five were roosting on the wires at the west end of the old causeway. Brown pelicans were not observed roosting on the wires during other trips to the study site.

Brown pelicans began nesting in late February or March. Brown pelicans summering in the vicinity of the P100 bridge are mostly immature or young of the year. On April 6, 1989, only 12 first-year immatures, those hatched in summer 1988, were observed foraging at Isla Blanca Park north of the Brownsville Ship Channel jetty.

Observations of Brown Pelicans at the P100 Bridge

A total of 313 observations of 1,287 individual brown pelicans were made at the bridge. Notation was made when possible of the age class, height of passage over the bridge with relation to the height of the light standards, whether individuals flew over or under the bridge, and the section of the bridge flown over or under. Poor visibility caused by weather conditions and distance limited the amount of information obtained; thus, all information is not available on all observations. This was particularly true of those crossings observed while driving across the bridge. Table 2 presents the information on age class observations and the height at which the brown pelicans flew across the bridge. Seventy-five percent of adults and 67 percent of immature brown pelicans (71 percent of all observations) crossed at or above the height of the light standards.

Table 3 presents the numbers of brown pelicans flying over or under the bridge according to the location along the length of the bridge. The extreme west end of the bridge is denoted Section 1, the west slope denoted Section 2, the main span over the Intercoastal Waterway denoted Section 3, the east slope denoted Section 4, the curve at the bottom of the east slope denoted Section 5, and the east end of the bridge after the curve was denoted Section 6. The majority of the brown pelicans crossed the bridge at the east slope (29 percent), the center span (29 percent), or the west slope (24 percent). All crossings made from the beginning of the west slope to the bottom of the east slope included 84 percent of all observations. Some bias was introduced by observations being made at the west end. The view from Queen's Point allowed maximum visibility of most bridge sections. Consequently, a number of crossings made at Section 5 or Section 6 were probably missed. However, it was apparent that most of the brown pelicans approached the bridge from the south generally heading for the center span. Crossings over other sections were generally a result of the birds turning to fly parallel to the bridge in order to gain altitude after an initial approach from the south.

Only 16 observations (7.5 percent) of brown pelicans flying under the bridge were recorded. All occurred under Sections 1, 2, or 3. Distance and visibility limited observations of brown pelicans crossing under the other sections. An apparent reluctance to fly under the bridge was demonstrated by brown pelicans occasionally flying low toward the bridge as if to fly under the center span, only to turn back, gain altitude, and fly over the bridge. When flying under the bridge during strong north winds, both brown pelicans and cormorants flew just above the water. During southerly winds, birds flew higher

TABLE 2 SUMMARY OF OBSERVATIONS OF BROWN PELICANS CROSSING THE P100 BRIDGE AT HEIGHTS ABOVE OR BELOW THE HEIGHT OF THE LIGHT STANDARDS

Age Class	Number of Observations	Number of Individuals	Height At/or above	Height below
Total adult	49	89	6	2
Adult	5	9	4	1
Winter	36	69	2	1
Breeding	8	11	0	0
Total immature	36	71	2	1
Immature	16	24	2	0
1st year	12	28	0	1
2nd year	8	19	0	0
Unknown	228	1127	155	63
Total	313	1287	163	66

TABLE 3 SUMMARY OF OBSERVATIONS OF BROWN PELICANS CROSSING OVER OR UNDER THE P100 BRIDGE ACCORDING TO SECTION OF THE BRIDGE CROSSED

Bridge section	Number of observations	Number of individuals	Flew over	Flew under	Unknown
1	6	20	18	1	1
1/2	4	5	3	2	0
2	38	73	56	3	14
2/3	1	2	0	2	0
3	27	109	57	8	44
3/4	1	1	1	0	0
4	7	48	47	0	1
4/5	1	1	1	0	1
5	4	6	6	0	0
5/6	1	1	0	0	0
6	3	7	7	0	0
Unknown	220	1014			
Total	313	1287	196	16	61

under the bridge, halfway between the water and the bridge deck or higher. Sharp tilting maneuvers by brown pelicans flying under the bridge were noted, both just as an individual began passage under the bridge from the south side and just as an individual cleared the bridge on the north side. This observation suggests turbulence at the surface of the water on both sides of the bridge, although the wind tunnel studies did not reveal any turbulence below the bridge between piers.

RECOMMENDATIONS

No evidence has been obtained indicating that brown pelicans may be intentionally landing on the bridge to seek shelter or to roost. Consequently, measures to discourage brown pelicans from landing on the bridge such as by flashing lights, propane cannons, or other noise makers are not recommended. Nor would alternative roosting structures, such as platforms or additional railings on the bridge, be useful because it is believed that the birds are not intentionally landing on the bridge roadway. On the basis of the information gathered in this study, the actions most likely to reduce mortality involve using traffic control measures to reduce the possibility of birds being hit once they are on the bridge deck, allowing additional time for the birds to safely depart the bridge. There is no evidence that the existing railings and median barrier present insurmountable obstacles to brown pelicans, but further observation of pelican behavior after adoption of traffic control measures is recommended. Should these traffic measures fail to satisfactorily mitigate the problem, further study of more radical alternatives is recommended.

Traffic Control Measures

Records of brown pelican mortality on the Queen Isabella Causeway indicate that the mortality most frequently occurs during the months of September through February which coincides with both the peak wintering population of brown pelicans and the presence of inclement winter weather conditions. Traffic control measures could be used during this time span to reduce the probability that a pelican would be hit once it is on the causeway, allowing the bird time to egress the roadway. Limited observations have not allowed assessment of the degree to which the existing railing represents an obstacle to a brown pelican on the roadway, but it appears that the birds that are killed, are killed before having much time to negotiate the railing. Birds on the roadway in the downwind lane may be faced with a more difficult problem unless they choose to depart through the downwind railing. The reverse airflow near the deck in the vicinity of the median barrier may serve to confuse birds on the roadway also. With these observations in mind, the following traffic control measures are recommended for consideration by the department:

1. The speed limit on the causeway should be reduced during the months of peak pelican wintering populations when the weather conditions known to be associated with pelican mortality exist. Basically, the significant weather variables include strong northerly winds. The presence of rain or mist makes the weather conditions more critical. Because these conditions only occur a few times a year at the P100 bridge site, special signing would be required; and since these conditions do not occur in a regular, programmable pattern, manually activated signing is recommended. Attention-getting signs, flashing lights, and appropriate enforcement are recommended. It may not be necessary to restrict speeds except on the main spans and the adjacent sloping approaches to the main spans because approximately 85 percent of the pelicans observed crossing the bridge flew over these portions. Studies to determine an optimum or recommended reduced speed have not been accomplished, but the objective should be to allow a driver in poor weather conditions to avoid collisions with birds already on the roadway.

2. The circuits that automatically actuate the causeway lighting should be adjusted so that the lighting is turned on 15 to 30 min earlier in the evenings. Cloudy, rainy, and foggy conditions reduce the likelihood that a motorist can see a brown pelican on the deck soon enough to avoid hitting it under natural lighting at dusk. Furthermore, even though there is no evidence to support a theory that the pelicans cannot see the bridge clearly enough at present, increased lighting might reduce mortality by providing the pelicans better visibility of the structure, especially a better altitude reference.

3. The warning signs that are presently posted at the approaches to the P100 probably have had little effect. The wording of the warning is not sufficiently detailed to properly convey to drivers that there is a potentially dangerous situation. The visibility of those warning signs could be increased by using a more noticeable design. The design using the pelican silhouette, which was originally rejected, would probably be more eyecatching. Wording of the warning should be changed to more accurately reflect the potential danger of hitting the pelicans. Motorists using the bridge daily may become habituated to the signs. Additionally, the high turnover in a temporary winter visitor population results in numerous people crossing the bridge who are unfamiliar with the brown pelican problem and who might miss seeing the warning signs under bad road conditions. The use of flashing lights on the signs connected with a reduced speed warning during periods of severe weather would increase the awareness of both locals and winter visitors. The lights on the sign could be activated remotely via telephone lines. In addition to these recommendations, consideration should be given to the installation of emergency telephones at each end of the bridge to help reduce the risk of motorists who stop on the bridge to aid birds or other motorists. These telephones would be valuable not only for reporting birds on the bridge, but for reporting traffic accidents or disabled vehicles as well. Direct-line emergency telephones would be preferrable to pay phones. Telephones in a highly visible area would reduce the likelihood of vandals' disabling them. Additional signs on the bridge warning motorists not to stop on the bridge should be installed.

Should traffic control allow the birds that land on the deck to survive longer, they may be able to effect safe egress from the deck. The possibility that the strong reverse flow at the deck disorients the birds may explain the mortality rates. Once traffic control measures are enacted, the behavior of the birds should be monitored to determine whether the confusing reverse flow or the railing geometry prevents the grounded birds from safely departing the bridge. In the former case, other measures must be also used. In the latter case, railing modifications, which will allow easier egress, can be considered. The need to consider these two possibilities cannot be determined until traffic control measures are enacted and the subsequent effects on the mortality rates are observed.

Aiding Brown Pelicans in Flying over the Bridge

Because the brown pelicans usually begin to gain altitude several hundred yards before reaching the bridge, it seems evident that they are aware of the bridge and how to cross it. However, they frequently arrive at the bridge at too low an altitude to successfully cross because of strong north winds and possibly the presence of turbulence above the deck. It may be that they are taking aim on the bridge railing. One technique that could be explored involves giving the birds a higher visual reference point so that they arrive at the bridge at a higher altitude. They need to be at or above the height of the light standards when they arrive at the bridge in order to make it across. No other instance is known where something like this has been attempted. One correspondent from Florida mentioned using orange balls on transmission lines to reduce brown pelican strikes. Some modification of this technique, possibly using streamers, could be attempted experimentally. Another possibility is to string a lightweight but visible plastic line between the tops of the light standards. In addition to giving the birds a visual reference, they may be reluctant to fly under the line and thereby avoid the region of turbulence and reverse flow. Careful monitoring of the brown pelican behavior would be necessary to make sure that the solution is not doing more damage than the original problem. The actuation of existing causeway lighting during inclement weather as recommended may give the birds a better visual reference.

Aiding Brown Pelicans in Flying Under the Bridge

Additional research would be required to determine why brown pelicans are apparently reluctant to fly under the bridge. This research could explore the possibilities that sound, including infrasound, or turbulence near the water's surface may be factors. The marked similarities between the structures of Bonner Bridge in North Carolina and the Queen Isabella Causeway suggest the overall configurations of these two bridges may be influencing brown pelican bridge mortalities. Such additional research should be planned if the recommended traffic control procedures do not mitigate the mortality problem.

Alternative Roost Sites South of the Bridge

Roost sites could be provided by the use of floating artificial islands such as Schwimmkampen. These artificial islands were developed by Lothar Bestmann of Bestmann Ingenieur Biologie in Germany and are being distributed by Sven Hoeger of Wetland Habitat by Design in New York. The design of the triangular modules is based on nautical engineering and ship building expertise. They have been used in Germany for 10 years for waterfowl habitat and other purposes and have survived strong winds in service. Artificial islands would also be used as nesting sites by waterbirds and possibly by brown pelicans as well. On the basis of a cursory study, probably the best location to place these would be in the cove formed by the west end of the old causeway and Long Island, south of the old causeway. This area is out of the way of most boat traffic, which could be a problem if the islands were placed between the two causeways. From behavioral observations it appeared that brown pelicans prefer to fly to and from the Gulf of Mexico via the Brownsville Ship Channel rather than flying across South Padre Island. Focusing both winter and summer populations south of the bridge would give them easy access to the Gulf for feeding and might keep them away from the bridge. Obviously, further research would be necessary to ascertain the feasibility of this approach. It might be possible to cosponsor research on artificial islands along with TP&WD and the FWS, as there is some interest in enhancing colonial waterbird habitat as the spoil islands degenerate.

The Texas brown pelican population appears to be increasing at present. Should brown pelican mortality rates increase to the point where a significant threat to the population exists or if the hazards to motorists increase unacceptably, it may be necessary to seriously examine modifications to the bridge structure such as a baffle on the north side of the bridge to deflect the wind currents, changing the design of the railing or changing the design of the center barrier. The findings of this study cannot support major modifications to the bridge structure. Even if much more detailed wind tunnel results were available to allow predictable reductions in the region of turbulence above the deck, the effects of such changes could only be evaluated by field trials.

The findings of this study should be carefully reviewed by designers of other major bridges over waters frequented by the Texas brown pelican. A major new bridge design in areas frequented by the brown pelican should include wind tunnel testing to evaluate the turbulence and potential risk to the brown pelican population.

ACKNOWLEDGMENTS

The work described here was funded by the Texas State Department of Highways and Public Transportation and FHWA, U.S. Department of Transportation, for which the authors are appreciative.

REFERENCES

1. D. W. Mabie. *Brown Pelican Study*. Final Report. Texas Parks & Wildlife Department, Austin, Jan. 20, 1986.
2. T. G. Pearson. Exploring for New Bird Colonies. *Bird Lore*, Vol. 23, 1920, pp. 276–277.
3. R. P. Allen. Notes on Some Bird Colonies on the Gulf Coast. *Auk*, Vol. 52, 1935, pp. 198–200.
4. A. F. Gustafson, C. H. Guise, W. J. Haminton, Jr., and H. Rise. Conservation in the United States, Comstock Publishing Co., Ithaca, N.Y., 1939, 534 pp.
5. C. J. Henny. *An Analysis of the Population Dynamics of Selected Avian Species with Special Reference to Changes During the Mod-*

ern Pesticide Era. U.S. Bureau of Sport Fishing and Wildlife Research Report 1, 1972, 99 pp.
6. R. T. Paul. Pelican Comeback. *Texas Parks and Wildlife*, Vol. 35, No. 3, 1977, pp. 12–15.
7. G. J. Thiessen and E. A. G. Shaw. Acoustic Irritation Threshold of Ringbilled Gulls. *Journal of the Acoustic Society of America*, Vol. 29, 1957, pp. 1307–1309.
8. G. J. Thiessen et al. Acoustic Irritation Threshold of Peking Ducks and Other Domestic and Wild Fowl. *Journal of the Acoustic Society of America*, Vol. 29, 1957, pp. 1301–1306.
9. M. A. Gollop and R. A. Davis. *Gas Compressor Noise Simulator Disturbance to Snow Geese, Komakuk Beach, Yukon Territory, September, 1972.* CAGSL Biological Report Service, Vol. XIV, Chapter VIII, 1974.
10. H. C. Oberholser. *The Bird Life of Texas.* University of Texas Press, Austin, 1974, 1,069 pp.
11. Christmas Bird Counts in American Birds. *Audobon Field Notes*, Vols. 4–24, 1950–1969, Vols. 25–43, 1970–1989.

Publication of this paper sponsored by Committee on Environmental Analysis in Transportation.

Cooperation Between State Highway and Environmental Agencies in Dealing with Hazardous Waste in the Right-of-Way

J. RANDLE SCHICK

The presence of hazardous waste or leaking underground storage tanks in new or existing highway right-of-way challenges state highway agencies. Because state environmental agencies regulate the cleanup of these sites, the relationship between these agencies can become either confrontational or cooperative. Formal cooperative agreements for dealing with hazardous waste in the highway right-of-way, as well as the less structured approaches being taken, are explored. A survey of state highway agencies in the spring of 1990 indicated that few highway agencies had entered a written memorandum of understanding with their environmental counterparts. Many states have worked out informal arrangements in which the state regulatory agency provides consulting or technical services to the state highway agency. In still others, a poor relationship has developed. These survey results, case studies, and written agreements are described.

When state highway departments acquire land to build and maintain roads, they can encounter hazardous waste, petroleum contamination, and asbestos in buildings to be demolished. They are then confronted with the bewildering legal, regulatory, financial, and technical maze that these sites represent. Every state has one or more agencies where the regulatory responsibility reposes to protect the public health and the environment from the damage caused by these wastes and substances. The purpose of this inquiry is to determine whether state regulatory agencies have taken advantage of the expertise, powers, and wherewithal of their respective regulatory agencies in dealing with the challenges of hazardous waste and, if so, whether they have formalized procedures and techniques through interagency agreements.

An interagency agreement of memorandum of understanding in this context refers to a joint agreement that can encompass one or more of the following arrangements:

1. The regulatory agency agrees to perform or provide a service (e.g., identification of contamination); or facility for the highway agency;
2. The regulatory agency and highway agency are to jointly perform a function (e.g., site cleanup);
3. The highway agency is to act on behalf of the regulatory agency (e.g., the highway agency prepares cleanup notices for the regulatory agency); and
4. The agencies merely make sense of the maze (e.g., the agencies delineate their respective functions and responsibilities).

Office of Chief Counsel, Illinois Department of Transportation, Springfield, Ill. 62764.

Interagency agreements are a potentially valuable tool in negotiating the maze and dealing with a variety of problems of concern to both agencies. They can elevate the concerns of the highway department with meeting project deadlines in the eyes of the regulatory agency. They can lead to organization and coordination when otherwise there could be conflict and unpredictability. They can bring the resources of the regulatory agency to bear on the responsible parties and bring to bear the special cleanup funds controlled by the regulatory agency.

However, it must be recognized that interagency agreements or the working arrangements pursuant to them change. They develop in response to particular needs and circumstances and those are changing in the rapidly evolving area of environmental law. They are also subject to changing administrations and personalities.

In addition, it must be recognized that a cooperation agreement can be perceived as "not worth the trouble." Given a choice between self-reliance and having to cooperate, any agency will choose self-reliance. An informal, as-needed relationship can be preferred to a written understanding. In fact, an interagency agreement with an environmental agency can be less a cooperative arrangement than it is a license for the highway agency to conduct cleanups on its own.

The study of cooperation in addressing environmental problems in highway land acquisition was intended to find out whether highway agencies have viewed the benefits of a cooperative agreement as worth the effort. It was desired to find out who was doing what and whether the results had been favorable. Although each state had its own legal and institutional framework, a question was whether any trends were emerging.

SUMMARY OF SURVEY RESULTS

On February 4, 1990, a questionnaire titled "A Survey on the Interaction of State Transportation Agencies with State Regulatory Agencies" was circulated to the transportation agency of each state, Puerto Rico, and Washington, D.C. As of June 1, 1990, 31 responses had been received.

Question 1 stated, "Does your agency have a formal Memorandum of Agreement or Understanding (MOU) with your state regulatory agency (SRA)?" Seven states indicated they did have formal MOUs with their regulatory agencies, while the remaining 24 states said they did not. However, only three of the states' MOUs broadly and specifically involve hazardous waste in the right-of-way. The MOUs of those states—California, Florida, and New Jersey—are outlined in the case studies that follow. The other four states have MOUs that narrowly (or do not yet) deal with hazardous wastes or substances. Vermont's MOU concerns only the reuse of petroleum-contaminated soils. Washington's MOU simply calls for a working agreement, but such an agreement had not been signed to date. Idaho's MOU only concerns emergency situations. Virginia's only deals with solid waste disposal.

Questions 2 and 3 were simply attempts to determine what the states' MOUs dealt with when the respondents did not include copies of their MOUs. Because all respondents included their MOUs, these questions were not needed or answered.

States were then asked what arrangements they have for notification of their SRAs instead of, or in addition to, MOUs. The answer to this question revealed that 68 percent of the responding states have formal procedures for notifying their state regulatory agencies of the discovery of hazardous waste (Figure 1). Such notification is done either in a written form, a verbal form, or through a 24-hr hotline maintained by the state regulatory agencies. In some states, notification is only made after the contamination is determined to be above a certain level. Other states follow regulations that require automatic notification upon the discovery of contamination and in some cases, immediately. Finally, other state departments of transportation (DOTs) notify a central office, either in their own or in another agency, which then notifies all other appropriate authorities.

The survey attempted to uncover at what levels further interagency coordination takes place once this notification has taken place. The largest number of responses (40 percent) indicated coordination primarily takes place at a field or staff level (see Figure 2). This group was followed by 25 percent of the states' indicating primarily middle management involvement and 20 percent indicating coordination of department heads. Only 15 percent of the states indicated they had specific departments or task forces that handled all coordination with state regulatory agencies. There seemed to be a consensus, however, that the level of coordination directly related to the importance of the issue. Although coordination mostly occurs at the staff level, if the issue is either precedent setting or urgent, the level of coordination can rapidly move up the ladder to upper management or department secretaries. Only a few of the states said they had no need for coordination at any level, indicating either that they had been able to handle this issue in house, or that they had not had the occasion to become involved with hazardous waste.

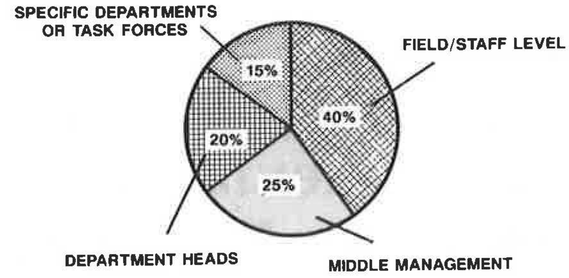


FIGURE 2 Question 5: What administrative level is used by each agency to conduct the coordination on hazardous waste issues?

Question 6 asked how much help state DOTs had received from their companion SRAs. The answers were widely varied (see Figure 3): 28 percent of the states indicated that they had been given access to files, site lists, and hazardous waste reports; 26 percent said that they had only received technical guidance in areas such as sampling, cleanup standards, potentially responsible party (PRP) enforcement, and site inspections; 24 percent of the states said that they had received little or no help from their state regulatory agencies. Some in this latter category either had no hazardous waste dealings at all, or the SRA was strictly reactive in posture. Only 11 percent of the states received consultation as needed or when asked for.

Question 7 generated even more widely varied (yet almost the same) answers as those to question 6. This question asked for examples of past coordination that had taken place between the DOTs and SRAs. The responses ranged from no help at all to as much help as possible. In some cases, the SRAs responded quickly, while in others they did not. The most common encounter was found in planning, discussion, or informational meetings. Help has also been provided in testing and evaluating, reviewing mitigation plans, resolving legal issues, removing leaking underground storage tanks, and disposing of contaminated soils.

Finally, Question 8 asked for a characterization of the cooperation between the two state agencies (see Figure 4). The majority of respondents (60 percent) said cooperation was good; 25 percent said the coordination was good, yet limited

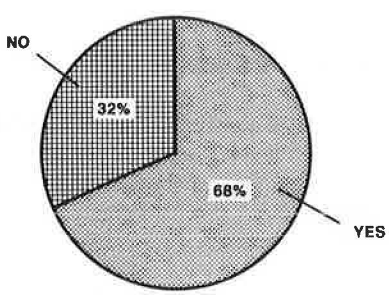


FIGURE 1 Question 4: Does your agency have a procedure to notify the regulatory agency if contamination or a contaminated site has been discovered?

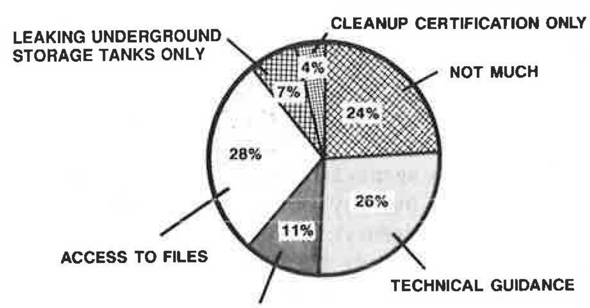


FIGURE 3 Question 6: What assistance has your regulatory agency provided on hazardous waste issues during planning or construction phases?

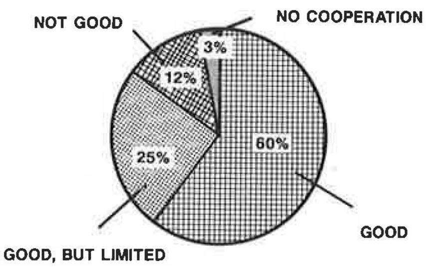


FIGURE 4 Question 8: Characterize the cooperation between your agency and the regulatory agency concerning hazardous waste.

and at times ineffective; finally, 12 percent said it was not good; and 3 percent indicated no cooperation whatsover. However, these numbers may be deceptive because several states that indicated difficulties with SRAs in Questions 6 and 7 said that their cooperation was good, or even excellent. It seems that respondents were reluctant to criticize their SRAs, despite these difficulties.

CASE STUDIES

States with MOUs

New Jersey

The New Jersey Department of Environmental Protection (NJDEP) and the New Jersey Department of Transportation (NJDOT) have signed an MOU that adopts a detailed, step-by-step "Standard Operating Procedure for Managing Soil/ Ground Water Contamination Issues" (SOP). The MOU recognizes that NJDEP is given the responsibility to protect the environment and the public health, safety, and welfare by state law and that NJDOT holds the responsibility for transportation services in New Jersey and has the power to obtain lands by condemnation. The MOU goes on to state that some of these lands acquired by NJDOT trigger the Environmental Cleanup Responsibility Act (ECRA). Because that state law requires an environmental assessment and cleanup, if necessary, at the time NJDOT acquires right-of-way, close cooperation between NJDOT and NJDEP is a necessity. This necessity led to the adoption of the SOP.

The SOP lists an exact 12-step process that details the coordination between NJDEP and NJDOT regarding the assessment and handling of all transportation projects that involve the acquisition of properties with soil or ground water contamination. It seeks the resolution of contamination cases on both an acceptable schedule for NJDOT and in a technically acceptable way for NJDEP. The agremeent includes a flow chart to help the reader follow each step.

During the initial assessment stages of a project, there is close cooperation between NJDEP and NJDOT. As NJDOT screens its transportation projects, it uses NJDEP's Preliminary Assessment format, Field Sampling Procedures Manual, a preliminary assessment of the site's applicability to ECRA. Following NJDOT's initial assessment, NJDEP and NJDOT work together to decide what the next step in the process should be. If the NJDOT determines it is possible to avoid a contaminated property and the PRPs cannot be found, then it may use its own resources to clean up the property in its right-of-way and seek reimbursement through court action. NJDEP will assume responsibility for the non-right-of-way property and seek reimbursement from the PRPs. If, for some reason, NJDOT decides not to spend its own funds for a cleanup, NJDEP will take over the cleanup efforts using its own funds and seek reimbursement from the PRPs at a later date.

This agreement involves a high level of cooperation between NJDEP and NJDOT. The agreement was successfully implemented because of the involvement of the commissioners of both departments.

Each commissioner recognized the need for a mechanism by which NJDOT could become integrated with NJDEP to the extent necessary for speedy site recognition and remediation. Instead of simply giving NJDOT blanket authority to clean up contaminated sites, NJDEP is intricately involved with each step of the cleanup process. NJDEP also has a transportation coordinator who helps enhance the cooperation between the two agencies.

Interestingly enough, NJDEP is apparently willing to not only support NJDOT in its attempt to seek reimbursement, NJDEP will also handle the reimbursement process for NJDOT if circumstances require. In some cases, NJDEP has ordered the PRP to perform a cleanup of a contaminated site ahead of NJDOT's ongoing projects. The only provision for reimbursement under state law is found in a spill fund administered by NJDEP. NJDOT indicated, however, that NJDOT can only receive monies from this fund if the contamination is a direct threat to the public health.

It is also important to note that the MOU merely adopts the SOP, whereas the SOP is the actual working agreement.

Florida

The State of Florida Department of Environmental Regulation (FDER) and the State of Florida Department of Transportation (FDOT) signed an MOU in July of 1989. The purpose of this agreement is to define the role of FDOT in cleaning up contamination sites that accrue to it through right-of-way acquisition and also its fiscal responsibilities for such cleanups. The MOU also reinforces an existing informal agreement between the two agencies. That agreement allowed FDOT to proceed with a cleanup of the necessary right-of-way if the contamination is an immediate threat to public health or the environment. FDOT is also allowed to delay a cleanup if its project will not exacerbate the existing contamination and FDER will have access to the site following the project's completion. However, this informal agreement was not deemed to be truly effective without a formal agreement.

The MOU's definition of contamination encompasses any substance that poses a serious danger to the public health, safety, or welfare that is released into the environment in quantities or concentrations sufficient to cause harm to the public health or the environment. This includes all hazardous wastes (HW) and substances, as well as petroleum and its by-products. There appears to be a strong statutory basis for this agreement. FDOT is protected by the Florida legislature from state-imposed liability that FDOT might incur when it acquires land for transportation purposes that is already contaminated. Florida is unique in that this state law gives FDOT more leverage when dealing with landowners.

The state legislature had also created two trust funds that provide money for cleanup and restoration of contaminated sites. FDER is able to seek reimbursement from these trust funds, and FDOT in turn gains reimbursement for FDER. The MOU details the rights and responsibilities that FDOT will assume as a result of this agreement. Once FDOT discovers some unsuspected contamination, it notifies the FDER, which in turn allows FDOT to examine any and all relevant records pertaining to the contaminated land. The FDOT can then proceed to clean a sufficient amount of the contamination to allow it to advance its project according to schedule.

All of this will be completed by FDOT using its own funds. As long as FDOT receives written approval from FDER to proceed with the cleanup, FDOT can seek reimbursement from the legislatively established trust funds. FDER agrees to fully support FDOT in its quest for reimbursement. FDOT will also assign its own staff to independently oversee and manage cleanups, and the FDER will assist FDOT as much as possible with the cleanup of contamination. Each district of FDOT has an HW coordinator who makes decisions regarding cleanups and acts as a project manager who oversees the mitigation process.

The end result of this agreement is that the FDOT is allowed to clean up any and all contamination sites it encounters in the course of transportation-related projects. Although FDOT has to pay for the cleanup with its own funds, it can receive reimbursement from established state trust funds with the backing of the FDER.

California

The California Department of Transportation (CALTRANS) and the California Department of Health Services (DHS) signed an MOU on July 14, 1989. DHS is given the general responsibility by the California Health and Safety Code to either oversee cleanups of contaminated sites or to perform the actual cleanups itself. In this MOU, contamination refers only to hazardous wastes, and the level of contamination is determined by the federal EPA Hazard Ranking System (HRS). It does not include underground tanks or asbestos as those are regulated by other state agencies. The purpose of this agreement is to stipulate how contaminated sites on CALTRANS right-of-way will be dealt with by these two agencies.

If CALTRANS becomes the owner of a property before abatement, CALTRANS may choose to expend transportation dollars to identify and clean up the hazardous waste contamination if the PRP does not take timely abatement action to meet CALTRANS' construction schedule. It is noteworthy that while the agreement provides that CALTRANS may spend its own funds for a cleanup, there is no mention of any possible reimbursement to CALTRANS for performing the cleanup. Minimal-threat sites, those scoring under 15 on the HRS, can be directly abated by CALTRANS without any DHS involvement. Non-minimal-threat sites, those scoring 15 or over on the HRS, have the option to be listed as California Superfund sites. They will not be listed if the PRP completely funds the site cleanup, and DHS is not asked for formal certification of the site's cleanliness. Under this MOU, however, DHS will still provide a written confirmation of the adequacy of the mitigation action.

If the site is to be listed on the California Superfund list, the DHS takes a much larger role in the site's mitigation. Regardless of whether or not it is Superfund listed, CALTRANS has the option of cleaning up any contaminated site itself. However, CALTRANS' cleanup liability of property acquired by eminent domain only extends to areas directly affected in the right-of-way of a construction project. It does not necessarily have to initiate a cleanup of areas in the right-of-way that are not directly affected, unless there are compelling public health or environmental issues. In the case of a partial cleanup, CALTRANS is required to prevent any increased threat to the public or environment because of the partial abatement.

This MOU outlines the parameters within which CALTRANS will conduct a cleanup of a contaminated site, either minimal or nonminimal, which directly affects its construction project. DHS requires notification of any such cleanups and reserves the right of approval for mitigation projects. Two aspects of this MOU are especially noteworthy. There is no mention of any form of reimbursement for CALTRANS when they undertake a cleanup of lands acquired through eminent domain. Also, although DHS in some cases will not provide a formal certification of a successful cleanup, they will provide a "written confirmation as to the adequacy of the mitigation action."

States Without MOUs

Many of the responding states indicated that the cooperation they had with their state regulatory agencies was less than satisfactory. They would not be expected to have an MOU. For instance, Georgia's Departent of Transportation (GDOT) has had problems getting contaminated sites cleaned up because of the slowness of their SRA. Because of this delay, Georgia has proceeded to perform a cleanup of the needed right-of-way. At least 65 percent of the sites that GDOT tested were found to be contaminated. Because of the SRA's lack of personnel, it may become necessary for GDOT to perform more cleanups. Georgian officials indicated that they had attempted to get an MOU signed, but alleged unwillingness on the part of the SRA has halted the process. Georgia would like to have a memorandum of agreement both to expedite the process and to clarify GDOT's legal position.

Connecticut's Department of Transportation also would like to implement an MOU. The connecticut respondent indicated that "staffing constraints faced by the [Department of Environmental Protection (DEP)] are the greatest impediments to effective coordination." Apparently, it is Connecticut DOT's feeling that an interagency MOU would greatly improve the cooperation between the two agencies. It would expedite the discovery and cleanup process by encouraging the Connecticut DEP to give the Connecticut DOT projects and needs a top priority. As of June 11, 1990, the process of signing an MOU had stalled, but the Connecticut DOT is still pushing to get an agreement signed between the two agencies.

There appears to be a trend among the responses that indicates that primarily urbanized and industrial states see the need for MOUs. States that already have MOUs, such as New

Jersey and California, are highly industrialized and urbanized and are much more likely to encounter hazardous waste problems. States that are more rural, such as New Hamphire or North Dakota, are less likely to encounter hazardous waste contamination and its resulting problems.

North Dakota's respondent indicated that his agency does not have any established procedures for dealing with hazardous waste issues and does not have any experience dealing with such problems. Thus, the agency has neither the need nor the desire for an MOU.

New Hampshire is another example of a rural state that sees no need for an agreement. It is just beginning to experience the problems associated with hazardous waste contamination because of the recent growth in the southern part of the state. Interestingly, this agency handles any necessary discovery and remediation processes itself. The biggest problem NHDOT has encountered is in the structrual mechanism of their SRA. NHDOT has had problems finding the proper people to notify in the SRA. The respondent indicated that while they had considered signing an MOU in the past, the efforts to do so had fallen by the wayside, because NHDOT has only experienced delays caused primarily by consultants and not their SRA.

CONCLUSION

The survey results indicated the following:

1. Far and away, the most frequent impediment to cooperation is a lack of staff in SRAs. Even if staff is available, SRAs complain that that staff is unqualified, incapable, or uncaring.
2. Given that barrier, highway agencies are generally cleaning up hazardous waste themselves and seeking reimbursement from the prior owner of the property or some state fund. These agencies are willing to assume the risk that reimbursement will be forthcoming.
3. SRAs have little formal or informal involvement with requiring right-of-way property owners to clean up their property before the highway agencies acquire their property.
4. Cooperation between state highway agencies and regulatory agencies is generally limited to the following ad hoc arrangements: (a) record searches for environmental problems (e.g., superfund or leaking undergound storage tank list); (b) review of environmental assessments prepared by consultants; and (c) "certification" that the property is cleaned up.
5. Cooperation with regulatory agencies on a case-by-case basis appears to be satisfactory to highway agencies.
6. There is little evidence of teamwork in these surveys. The DOTs and SRAs do not function as a team in solving problems with contamination in the right-of-way.

In a perfect world, highway agencies would build and maintain highways and regulatory agencies would clean up the environment. In that world, highway agencies would be able to contract with regulatory agencies to perform timely environmental assessments in the right-of-way, to oversee the cleanup by the property owner, and to certify that the property is all clear. In the perfect world, highway agencies would not have to have their own hazardous waste expertise, would not have to hire their own hazardous waste consultants and cleanup contractors, and would not have to pursue reimbursement for cleanups from property owners either in eminent domain or cost recovery proceedings.

We do not live in a perfect world. The survey results indicate that highway agencies, with few exceptions, are handling all of these functions related to hazardous waste on their own, with minor assistance from their SRAs. Not surprisingly, the few MOUs that exist are found in states such as California and New Jersey with strong environmental agencies. It is also not surprising that New Jersey has an MOU because its ECRA law requires an environmental assessment and cleanup at the time of sale of property.

There will soon be more formal agreements and more cooperation for the following reasons:

1. State regulatory agencies will grow in size and gain larger staffs,
2. State highway agencies are playing with fire in performing their own cleanups and are going to get burned, and finally
3. What are now informal relationships will become formalized.

Copies of the MOUs from Florida and California and a list of contacts for states that responded to this survey are available from the author.

ACKNOWLEDGMENT

The cooperation of Victor A. Modeer, Jr., and Scott S. Stephens of the Illinois Department of Transportation was necessary for the completion of this paper.

Publication of this paper sponsored by Committee on Environmental Analysis in Transportation.

Comparison of Intersection Air Quality Models' Ability To Simulate Carbon Monoxide Concentrations in an Urban Area

JOHN ZAMURS AND ROBERT CONWAY

In support of an environmental impact statement for a major transportation project in New York City, an air quality study was undertaken to determine model performance of intersection air quality and to ascertain which model should be applied for impact analyses for this project. Instruments were set up at six intersections to measure carbon monoxide concentrations and meteorological data. Traffic data were collected by videotaping. To date, results from two of the six intersections have been analyzed. Model performance was disappointing. Correlation coefficients of observed to predicted concentrations were low (generally less than 0.1), as were the slopes of linear best-fit curves. The models, on average, underpredicted observed concentrations, with only those models that separate composite emissions into their more discrete components indicating a potential for approaching or overpredicting observed carbon monoxide levels.

The New York State Department of Transportation (NYSDOT), with the participation of the city of New York and the FHWA, is proposing a reconstruction of the southernmost 5 mi of Route 9A in New York City, spanning from Battery Place to West 59th Street along the west side of Manhattan, beside the Hudson River.

Because of uncertainties in carbon monoxide mobile source modeling, a major air quality study was undertaken for this project. This study included multiprobe monitoring of carbon monoxide (CO) and meteorological and traffic data collection at six intersections.

The observed CO concentrations were compared to modeled CO values on the basis of observed meteorological and traffic inputs. The performances of eight different methodologies, established by a model selection protocol that was developed before model runs, were compared to determine which model was best able to simulate high observed CO concentrations.

DESCRIPTION OF SITES

For CO monitoring, six sites were selected at intersections likely to have project-induced traffic increases, and, therefore, possible air quality impacts. Spatial coverage of the study area and different traffic levels and geometries were also important considerations. Results for the first two sites analyzed are described.

Site 1, at West and Chambers Streets, is a T-type intersection directly on Route 9A (see Figure 1). Route 9A (also known as West Street in this area) runs north-south; the cross street, Chambers Street, forms the eastern leg of the intersection. Route 9A has two lanes in each direction, with left-turn lanes in the southbound direction. Chambers Street has one lane in each direction at the intersection. At Site 1, Route 9A generally has volume-to-capacity (*v/c*) ratios of about 0.75. Traffic flows fairly well because of the arterial nature of the roadway and the generally well-spaced intersections. The left-turn movements from southbound Route 9A are heavy. At Chambers Street, the left-turn movements are heavier in the a.m. peak hour.

Site 2, at Eighth Avenue and West 34th Street, is not on Route 9A, but in the secondary impact area, where possible traffic diversions could create air quality impacts. At Site 2, Eighth Avenue is one-way northbound (three lanes), and West 34th Street has two lanes in each direction (see Figure 2). This site is near Madison Square Garden, Penn Station, and the General Post Office. Congestion levels are greater than at Site 1 and more representative of Midtown traffic levels. Often during peak hours, *v/c* ratios approach 1.0. Left turns from 34th Street are not allowed at this location during peak hours.

Near Site 1, a background site was selected on an undeveloped portion of the area filled for Battery Park City, approximately 300 ft from West Street, the nearest source of traffic emissions. CO and meteorological parameters were monitored at this location. Near Site 2, a second background site was established on the roof of the nearby General Post Office, approximately 60 ft above street level and 15 ft above the roof line. At this location, only CO was monitored. This rooftop location was chosen because there were no ground-level locations nearby that would not be directly influenced by adjacent roadways.

DESCRIPTION OF EQUIPMENT

At each intersection site and at the Battery Park City background site, the CO monitoring equipment and associated

J. Zamurs, Environmental Analysis Bureau, New York State Department of Transportation, 1220 Washington Avenue, Albany, N.Y. 12232. R. Conway, Allee King Rosen & Fleming, Inc., 117 East 29th Street, New York, N.Y. 10016.

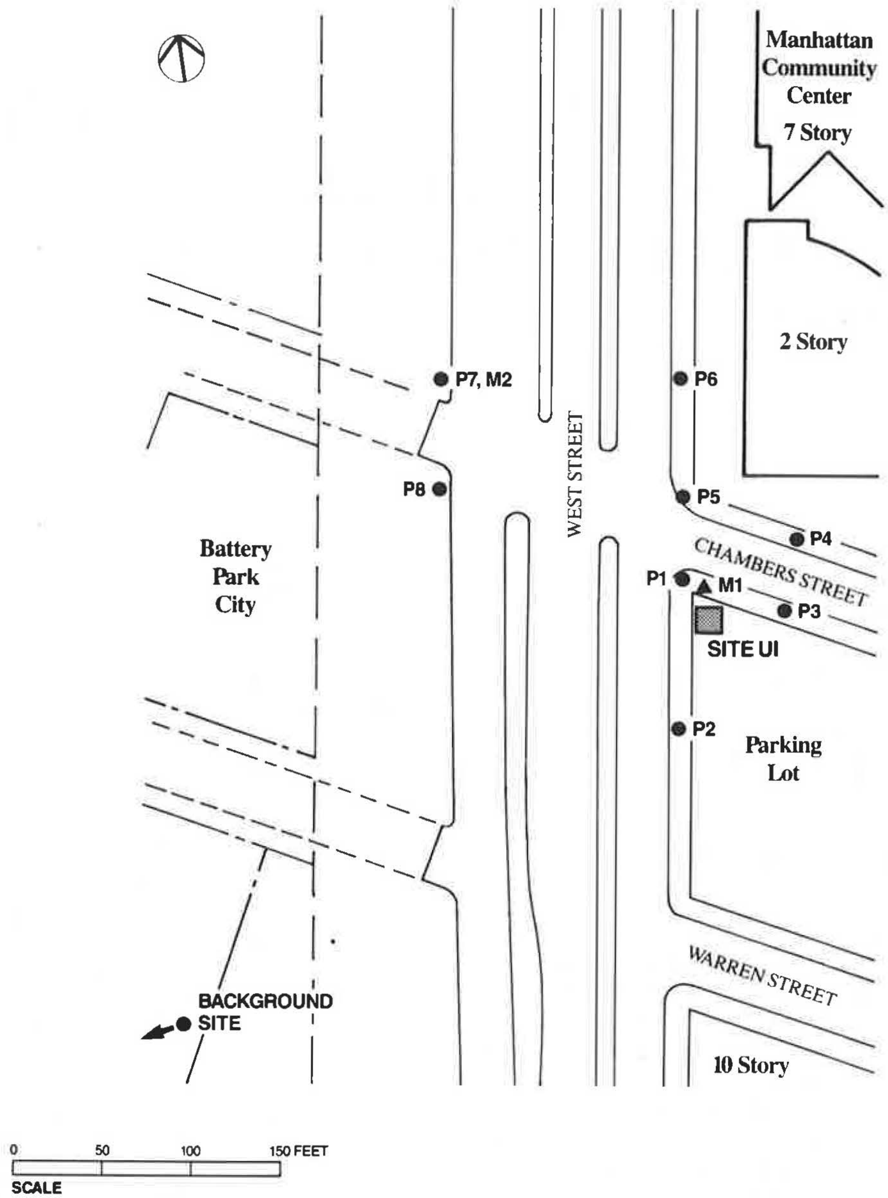


FIGURE 1 Site 1: West and Chambers streets.

instrumentation were placed in environmentally controlled shelters. At the General Post Office background site, the equipment was placed in a small room on the top floor of the building, with only the CO probe exterior to the structure and extending above the roof line.

The shelters at the intersection sites were each equipped with four CO analyzers; one CO analyzer was placed at each background site. Standard CO monitoring equipment was used. Additional details regarding the CO equipment and the probe and intake system design were provided by Sacco and Zamurs (*1*). At each intersection site, eight CO sampling probes were used, requiring that each CO analyzer share two probes. The sample probe and intake systems were designed to meet the specific requirements of each location. The eight probes were placed approximately 9 to 10 ft above street level and approximately 6 ft from the roadway, halfway between the curb and the building line. Figures 1 and 2 show the relative location of the sampling probes and the placement of the meteorological towers at the two intersection sites.

Two meteorological systems were used at Site 1 and three at Site 2. Meteorological monitoring towers of 10-ft height were used. One tower was placed atop the shelters at each of the intersection sites and also at the Battery Park City background site, and one and two additional meteorological monitoring systems were set up on existing light poles at Sites 1 and 2, respectively. The locations of the meteorological equipment are shown in Figures 1 and 2.

At each meteorological monitoring site, wind speed, wind direction, and temperature were measured using standard equipment; standard deviation of the wind direction (σ_θ) was calculated on the basis of an algorithm developed by Nelson (*2*). Initially, the σ_θ values were used to estimate Pasquill Stability Categories using the methodology outlined in EPA guidelines (*3*). However, on the basis of the initial review of

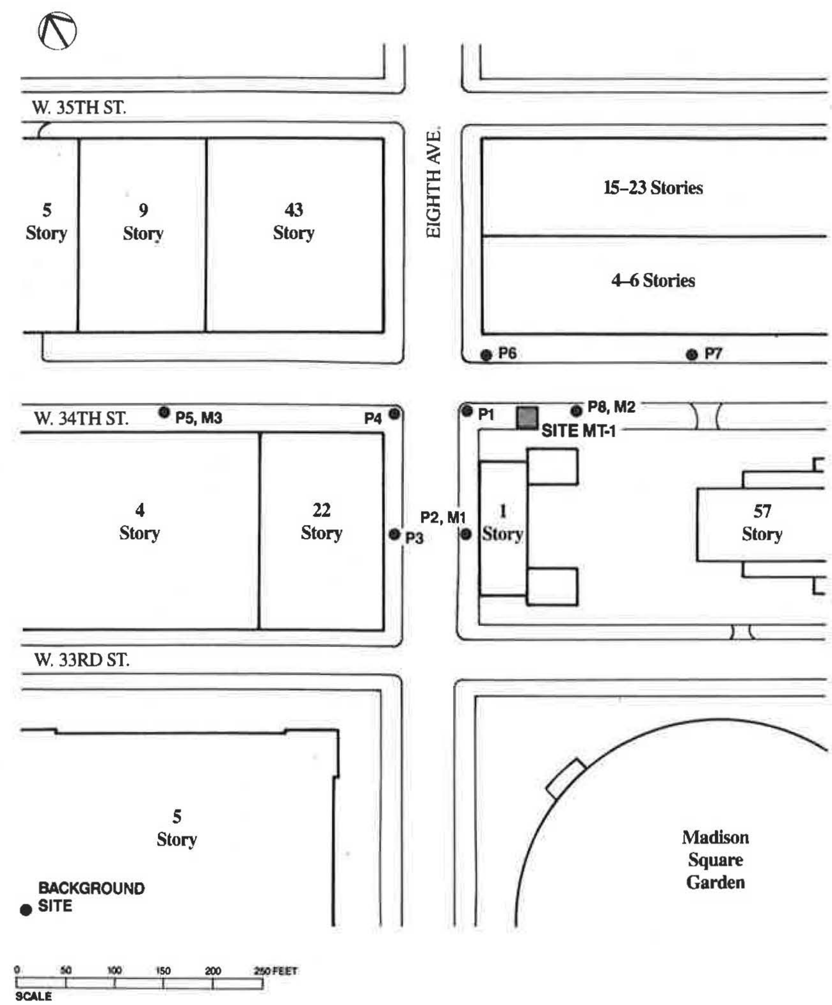


FIGURE 2 Site 2: 34th Street and Eighth Avenue.

σ_θ values and resultant stability determinations (with many excessively high values of σ_θ), and considering the siting of the meteorological sensors (which failed to meet unobstructed distance criteria), the methodology was revised. The stability classes used in the modeling analysis were based on cloud cover data from La Guardia Airport and the on-site wind speed measurements at the intersections (*4*). Except for stability inputs, all meteorological inputs for the modeling were derived from data collected at the intersection sites.

Data acquisition and validation were governed by strict standard operating procedures through a quality assurance program. The operation, data processing, and quality assurance–quality control practices conformed as closely as possible to the provisions of EPA prevention of significant deterioration (PSD) monitoring guidelines (*5*). However, the meteorological measurement systems' locations did not meet EPA siting criteria with respect to the required distance from a structure. On the basis of building heights in Manhattan, it was highly unlikely that a location that met the siting requirements could be obtained. No simple adjustment was readily available to compensate for this. Twice during the monitoring operation, the CO analyzers were audited by representatives of New York State Department of Environmental Conservation. Each of the intersections was monitored until 3 months of valid data were obtained. The background sites were monitored continuously during the intersection site monitoring program.

Traffic data to be used in the air quality prediction modeling scenarios were obtained through the use of videotaping and field surveys. All roadway links within 1,600 ft of the intersection where CO data were being collected were modeled, requiring a substantial traffic data collection effort. Because traffic signals in New York City can be as little as 100 ft apart, and are typically spaced every 250 ft along a north-south avenue, a substantial number of roadway links were required. For Site 1, 17 separate links were used for the analysis; for Site 2, 26 links were used.

The majority of the data collection effort used videotaping to gather real time detailed traffic information. Because an aim of the study was to determine model performance when high levels of CO were observed, videotaping of traffic data was preferred because only some of the hours would be reduced for model input. At both intersection sites, numerous videocameras and recorders were positioned strategically atop buildings and other structures to obtain the various traffic parameters.

The following traffic data were obtained from the videocameras: hourly traffic volumes, average travel speed, average stopped delay, average queue length, average travel speed excluding stopped delay (modified speed), vehicle classification, and traffic signal information. Once the hours of concern were determined from the CO monitoring, the videotapes were replayed to obtain the necessary traffic data. Statistical testing was performed to determine the proper hourly sample size required to obtain accurate hourly average values.

For the first five parameters, a value was obtained for each individual roadway link. Vehicle classification data were obtained on a corridor basis (i.e., northbound, southbound, eastbound, westbound), using data from the intersections where the monitoring was being conducted.

Some of the traffic parameters that were required for one or more of the modeling methodologies could not be obtained from the videotapes. These parameters—acceleration and deceleration rates, cruise speeds, and thermal conditions (i.e., hot or cold start mode)—were obtained from field surveys. The data for each of these parameters were obtained for four different averaging periods—morning peak, midday, evening peak, and the late-night, low-volume period. Each hour of concern that was examined for detailed analysis was assigned to one of these categories and the appropriate thermal state, cruise speed, etc., was used.

DESCRIPTION OF MODELING

Emission Scenarios

The study evaluated the performance of four different emission scenarios using both the HIWAY-2 (*6*) and CALINE3 (*7*) line source dispersion models—a total of eight modeling scenarios, or models. The four emission scenarios, which differed in their spatial and quantitative representation of emissions, were as follows:

- Average speed,
- Modified average speed plus idle,
- Modal emissions, and
- CAL3QHC.

Each of these scenarios is discussed in the following paragraphs, along with a brief description of the line sources used in each application.

Average Speed

The average speed emissions scenario used an average source strength over the entire link, which was based on the total travel time from stop line to stop line for each individual link. A link was defined as a roadway segment (i.e., Eighth Avenue northbound from West 33rd to 34th Street) between two traffic signals. At both sites, each roadway link was modeled using the traffic information from the videotapes and emission factors from the EPA's mobile source emission model MOBILE4 (*8*). This scenario is similar to the FHWAINT model, as described by PEI Associates (*9*).

Modified Average Speed plus Idle

This scenario was used to simulate the effects of idling vehicles at each intersection where a traffic signal was located. The modification refers to the methodology used in calculating the average speed: the vehicle's stop time in a queue (or stopped delay) was subtracted from the total travel time, resulting in a higher average speed. The higher average speed was used with MOBILE4 emission factors to determine the emission source strength over the entire link. This procedure was the same as that discussed for the average speed scenario, but minimized double counting of idle emissions. (The MOBILE4 emission factors were based on a drive cycle in which the proportion of idle emissions increased with decrease in speed.) This scenario is similar to the EPAINT model, as described by PEI Associates (*9*).

The modified average speed for a given roadway segment was still less than the cruise speed, because it included delays caused by slowing, merging, side frictions, etc. The stopped delay was used in this scenario rather than the approach delay, because it was more conducive to quantification from either videotapes or field surveys. It is also a direct output of the *1985 Highway Capacity Manual* (HCM) (*10*), which was being used for the transportation analysis within the overall Route 9A environmental impact statement.

The idle emissions were distributed over a line source that was based on the information obtained from the videotapes. The queue length as measured at the end of the red cycle, averaged over the hour, was used to position this link on the roadway network. It extended back from the stop line of the intersection for the distance computed from the videotapes for each link for each hour of concern. The length was computed assuming a conversion of 8 m per vehicle in queue.

The idle emissions yielded from MOBILE4 were used with the measured stopped delay to estimate the emissions from the queue line source. However, because MOBILE4 only provides idle emission factors for hot-stabilized vehicles at 75°F, the idle emissions were adjusted for the given thermal conditions and measured temperature by developing ratios of MOBILE4 emission factors at 2.5 mph for actual conditions and the hot-stabilized condition.

Modal Emissions

For this scenario, emissions were computed for the four modes of vehicular operation—cruise, deceleration, idle, and acceleration. The idle emissions were obtained from MOBILE4, and the cruise, acceleration, and deceleration emissions were obtained from the modal emission subroutines of the Intersection Midblock Model (IMM) (*11*), which were based on the EPA Modal Analysis Model (*12*).

The modal emissions were distributed along the link in close approximation to where they actually occurred. The idle emissions were computed on the basis of the average vehicle stopped delay, and distributed along the average queue length obtained from the videotapes (the same as discussed for the modified average speed plus idle scenario). The cruise emissions were distributed over the entire roadway link, whereas an additional pseudolink was developed for each segment to simulate acceleration-deceleration emissions. The length of each acceleration-deceleration pseudolink was based on the upstream and downstream cruise speeds and the acceleration and deceleration rates obtained from the field surveys. The method used in the IMM model of placing this pseudolink upstream from the stop point of the link was followed for this scenario.

The subroutines from the IMM model that generate cruise, acceleration, and deceleration emissions for hot-stabilized automobiles at 75°F in 1977 were utilized in the analysis. The 1977 emissions were updated for present-day conditions using correction factors generated by MOBILE4, on the basis of ratios of 1977 conditions to present-day observed conditions. These correction factors also accounted for differences in thermal conditions, temperature, vehicle mix, etc. The MOBILE4 idle emission factors were updated in the same manner as those used in the modified scenario. As indicated, this scenario is similar to IMM. It is also similar to Volume 9 (*13*), because, in many respects IMM is a computerized version of Volume 9. By having acceleration-deceleration emissions included, this scenario is the most similar to TEXIN2 and CALINE4 (*14,15*). These two models, although popular, were not directly evaluated in this study because using videotaped traffic information such as queue lengths and delay would have substantially eliminated some of the significant differences between these two models and among the other models represented by the modal emissions scenario. However, the performance of the modal emissions scenario relative to the other emissions scenarios would be expected to apply to TEXIN2 and CALINE4 as well.

CAL3QHC

A fourth emission scenario was studied that simulated the techniques used in the EPA's CAL3QHC model (*16*). This scenario differed from the modified average speed plus idle scenario in only two respects. First, instead of using the measured stopped delay from the traffic videotapes, the idle time was based on the percent red time at each traffic signal. Second, in the calculation of the queue length, this scenario used six meters per vehicle as opposed to the eight meters used in modified and modal scenarios. This scenario, while similar to the EPA's CAL3QHC, still differed significantly from that model in that the traffic input parameters were based on field-measured data and were not part of the model.

MOBILE4

For all scenarios except the modal scenario, calendar year 1989 vehicular emission factors were solely obtained from EPA's Mobile Source Emission Factor Model MOBILE4. For the modal scenario, the idle emission factors were obtained using MOBILE4, but the cruise and acceleration-deceleration emissions were based on the 1977 Modal Analysis Model. However, MOBILE4 emission factors were used to update and correct the modal emissions to the present-day scenario, as described.

Common parameters for the four emission scenarios used with MOBILE4 included thermal conditions, based on field surveys; ambient temperature; vehicle classification, obtained from the videotapes; inspection-maintenance (I/M) and antitampering credits, obtained from New York State Department of Environmental Conservation (DEC); mileage accrual and vehicle age distribution, obtained from the New York City Department of Environmental Protection (DEP); and fuel volatility (Reid vapor pressure), based on data obtained from DEC. Because of the significant numbers of medallion (yellow) taxicabs on New York City streets, different classes of light-duty gasoline vehicles were used in MOBILE4. Mileage accumulation and vehicle turnover are significantly different for taxis than for privately owned automobiles. Because of the nature of their operation (i.e., cruising), they are predominantly in a hot-stabilized thermal state. Therefore, separate MOBILE4 executions were made to develop the data base of light-duty emission factors at any given speed.

Dispersion Models

The four emission scenarios were used with the HIWAY-2 and CALINE3 dispersion models. Because more than one meteorological data set was obtained at the two intersection sites, the wind speed and direction used in the dispersion model depended on the probe being analyzed. Each of the eight CO probes was assigned to the closest meteorological tower. A mixing height of 1,000 m was used throughout the analysis, as was source heights of 0 m.

For the CALINE3 model, a surface roughness value of 3.2 m was used for both sites. As recommended in the CALINE3 User's Guide (*7*), 3 m was generally added to each side of the line source to account for the width of the mixing zone. However, for all idling links and a left-turn bay at the Chambers Street site, the 3 m was not added to the width of the line source. For CALINE3, six stability classes, A through F, were used in the modeling, whereas HIWAY-2 uses three classifications—unstable, neutral, and stable.

MODELING PROCEDURE

Evaluation studies have typically addressed both the scientific and operational performance of the candidate models (*17,18*). Scientific performance deals with the cause-and-effect relationships within each model and an understanding of how the model components contribute to the various errors. Operational performance deals with how the various models compare within a particular application, such as the prediction of high concentrations. For this study, the performance of the various modeling scenarios at predicting high concentrations, an operational characteristic, was the main concern. However, scientific evaluation of the modeling scenarios added to the understanding of their performance. The predicted versus observed concentrations for the various modeling scenarios under different spatial, meteorological, and traffic conditions were also analyzed.

Data Set Selection

The data set was first developed by subtracting the concentrations measured at the background monitoring locations from concentrations measured at each intersection site probe on an hourly basis. All the hourly CO values for each month were then rank ordered. Each hour appeared only once in the data set, regardless of the number of probes that experienced high CO values. The top 50 hr from each of the 3 months with the highest observed concentrations were then chosen for inclusion into the data set and subsequent traffic reduction.

Performance Measures

The primary goal of the study was to determine which modeling scenario most accurately predicted the highest measured CO concentrations at congested urban intersections. This model would then be used as the predictive air quality model for the Route 9A environmental impact statement's air quality study. Answers were also sought for many other questions, such as: Does one scenario consistently overpredict or underpredict by a constant percentage whereas another has no pattern to its error? Is there a difference in predicting midblock versus intersection concentrations between the various modeling scenarios? Do the scenarios exhibit good correlation but poorly predict the magnitude of the peak levels? Are the mean errors small because of the balancing of large underprediction with large overprediction? etc. With these questions in mind, a variety of statistical measures and graphics were used to evaluate the scenarios (*19–24*).

The modeling scenario predictions were compared with the observed data both in paired and in unpaired form. A paired analysis was used to evaluate each scenario's ability to replicate measured concentrations both temporally and spatially. The unpaired analysis was used to determine each scenario's ability to predict accurately the magnitude of the peak without regard to location or time. These comparisons analyze the scientific and operational performance of each scenario, respectively. The unpaired analysis is often given more weight by regulatory agencies, because it is thought to be more likely to represent the worst-case condition (i.e., the rare event) and, thus, be more protective of the ambient air quality standards.

Unpaired Comparisons

In the unpaired tests, the observed values were compared to the highest predicted concentrations without regard to probe location or time of occurrence. For the top 150 values, the following standard statistical parameters were calculated: (a) bias, (b) absolute gross error, (c) mean square error, (d) variance, and (e) cumulative frequency distribution plots of both the entire data set and the top 25 percent of the data set. (These plots were useful in evaluating the overall, or operational, performance of each modeling scenario. Each modeling scenario's performance in replicating the cumulative frequency distribution was evaluated for goodness-of-fit using a Kolmogorov-Smirnov test.)

Paired in Time and Space

The paired statistical tests were used to assess each scenario's ability to predict concentrations on the basis of the hourly traffic and meteorological conditions and to predict the location and time of distinct concentrations. The following standard statistical analyses were performed:

1. Scatter plots of predictions versus observed measurements, the correlation coefficient, and the slope and intercept of the linear least-squares regression line;
2. The bias and variance of the concentration difference where this difference is paired both in space and in time; and
3. The gross variability, which is a measure of the size of the error produced by each scenario.

Paired in Time Only

The last set of performance measures used in the study involved pairing the observed and predicted data in time, irrespective of probe location. These measures were used to determine if the various scenarios' performances differ when the pairing in space is eliminated from the analysis. The performance measures used for this analysis were correlation coefficient, slope of regression line, bias, and variance. These measures were recomputed using the highest predicted and observed values for each hour.

Results

The results of the statistical analysis for both sites are presented in Tables 1 and 2. In the tables, some of the statistics for the unpaired tests were computed for both the entire data set and the top 25 percent of the data. The latter were used to evaluate each scenario's performance at the highest observed concentrations.

Unpaired Basis

Two sample cumulative frequency plots for the unpaired data are shown in Figures 3 and 4. All the modeling scenarios indicated a negative average bias on an unpaired basis. For Site 1, all the HIWAY-2 scenarios and three of the CALINE3 scenarios significantly underpredicted observed concentrations. At Site 2, two of the CALINE3 scenarios indicated overpredictions on a cumulative unpaired basis (at approximately the 95th percent frequency).

At the highest predicted concentrations, the overpredictions were sometimes quite significant. For example, the CALINE3 modified scenario overpredicted the highest observed concentration by about 55 percent. Finally, at Site 2, at the highest end of the concentration distribution (greater than 75 percent), the HIWAY-2 modified scenario came closer to replicating the observed concentration distribution than the other scenarios.

Paired Basis

The results on a paired basis were even less encouraging. The paired-in-time-and-space regression analyses indicated basically no correlation for any of the modeling scenarios. The slopes of the best-fit linear regression lines were nearly zero in many instances and at times were actually negative. Figure 5 shows a typical scatter plot and regression line.

As expected, the regression analysis on a paired-in-time-only basis indicated some improvement over the analysis paired-in-time-and-space. However, neither the correlation coefficients nor the slopes of the regression lines attained a satisfactory level.

TABLE 1 SUMMARY STATISTICS—WEST AND CHAMBERS STREETS

Statistics	HIWAY-2 Average	HIWAY-2 Modified	HIWAY-2 Modal	HIWAY-2 CAL3QHC	CALINE3 Average	CALINE3 Modified	CALINE3 Modal	CALINE3 CAL3QHC
UNPAIRED								
All Data								
Bias	-4.26	-3.15	-3.19	-1.58	-4.64	-3.71	-3.82	-2.58
Absolute Gross Error	4.25	3.15	3.19	1.58	4.63	3.71	3.82	2.58
Mean Square Error	18.27	10.05	10.31	2.79	21.72	13.82	14.64	6.93
Variance	0.15	0.12	0.13	0.28	0.22	0.06	0.04	0.27
Goodness-of-Fit	6.70	5.10	5.70	4.00	7.50	4.40	5.00	3.70
Top 25 Percent								
Bias	-8.98	-6.35	-6.53	-2.94	-9.90	-7.31	-7.67	-4.64
Absolute Gross Error	8.98	6.34	6.53	2.94	9.90	7.71	7.67	4.66
Mean Square Error	40.77	20.39	21.66	4.66	49.64	26.85	29.50	11.45
Goodness-of-Fit	6.70	5.10	5.70	2.00	7.50	4.00	5.00	2.60
PAIRED IN TIME AND SPACE								
Correlation	0.03	0.04	0.06	0.03	0.10	0.09	0.13	0.06
Slope	0.02	0.05	0.07	0.05	0.06	0.08	0.12	0.09
Intercept	1.04	1.48	1.45	2.48	0.49	0.81	0.61	1.53
Bias	-4.67	-4.06	-3.98	-3.09	-5.00	-4.55	-4.52	-3.79
Variance	1.50	2.21	2.02	3.62	1.20	1.71	1.64	2.78
Gross Variability	23.31	18.71	17.88	13.18	26.18	22.43	22.07	17.14
PAIRED IN TIME ONLY								
Correlation	0.14	0.12	0.18	0.31	0.18	0.05	0.10	0.35
Slope	0.09	0.12	0.17	0.41	0.10	0.05	0.10	0.23
Bias	-4.26	-3.15	-3.19	-1.58	-4.64	-3.71	-3.82	-2.58
Variance	1.22	1.74	1.50	1.83	1.07	2.12	1.71	2.40

TABLE 2 SUMMARY STATISTICS—34TH STREET AND EIGHTH AVENUE

Statistics	HIWAY-2 Averge	HIWAY-2 Modified	HIWAY-2 Modal	HIWAY-2 CAL3QHC	CALINE3 Average	CALINE3 Modified	CALINE3 Modal	CALINE3 CAL3QHC
UNPAIRED								
All Data								
Bias	-3.47	-2.44	-2.61	-2.70	-4.12	-2.76	-3.22	-3.25
Absolute Gross Error	3.47	2.62	2.62	2.70	4.12	2.83	3.24	3.25
Mean Square Error	12.17	6.56	6.56	7.51	17.06	9.11	10.97	10.94
Variance	0.10	0.59	0.36	0.20	0.06	1.51	0.61	0.36
Goodness-of-Fit	4.60	4.30	4.30	4.20	4.70	4.70	4.70	4.40
Top 25 Percent								
Bias	-6.86	-4.00	-4.76	-5.10	-8.09	-3.90	-5.49	-5.71
Absolute Gross Error	6.85	4.00	4.76	5.10	8.09	4.29	5.59	5.71
Mean Square Error	23.68	9.04	11.99	13.40	32.87	11.95	16.68	17.11
Goodness-of-Fit	4.60	2.20	3.10	3.50	4.70	4.10	2.70	2.60
PAIRED IN TIME AND SPACE								
Correlation	0.03	0.04	0.02	-0.01	-0.03	0.01	-0.02	-0.03
Slope	0.03	0.05	0.03	-0.02	-0.03	0.02	-0.03	-0.04
Intercept	2.05	2.20	2.32	2.60	1.89	2.15	2.25	2.40
Bias	-4.33	-4.00	-4.03	-4.05	-4.88	-4.30	-4.49	-4.44
Variance	2.86	3.86	3.55	3.52	2.70	4.90	3.89	4.03
Gross Variability	21.59	19.86	19.8	19.95	26.50	23.37	24.05	23.71
PAIRED IN TIME ONLY								
Correlation	0.06	0.10	0.08	0.05	0.02	0.12	0.09	0.03
Slope	0.07	0.16	0.11	0.06	0.03	0.24	0.15	0.05
Bias	-3.47	-2.62	-2.61	-2.70	-2.70	2.76	-3.22	-3.25
Variance	2.84	4.03	3.65	3.27	3.27	6.32	4.51	4.10

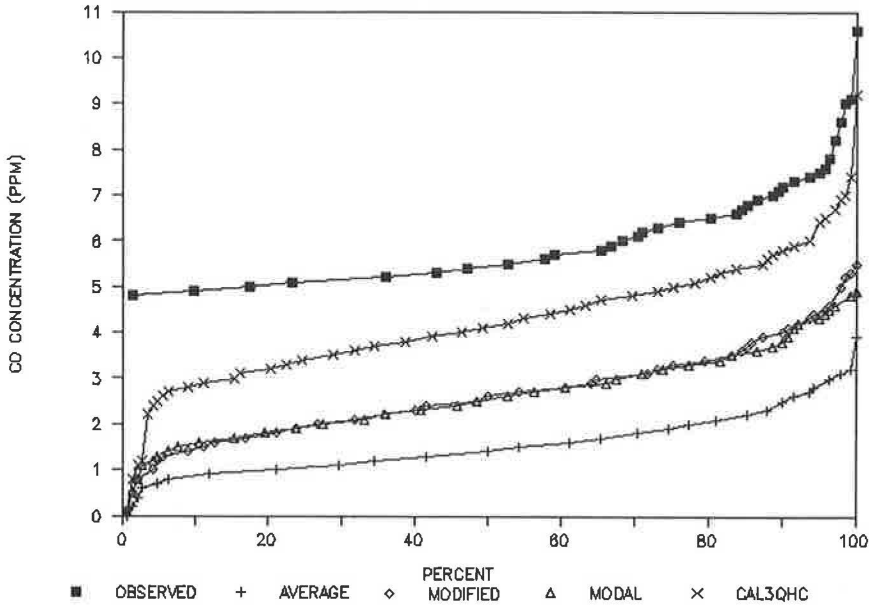


FIGURE 3 Cumulative frequency distribution, Site 1, HIWAY-2.

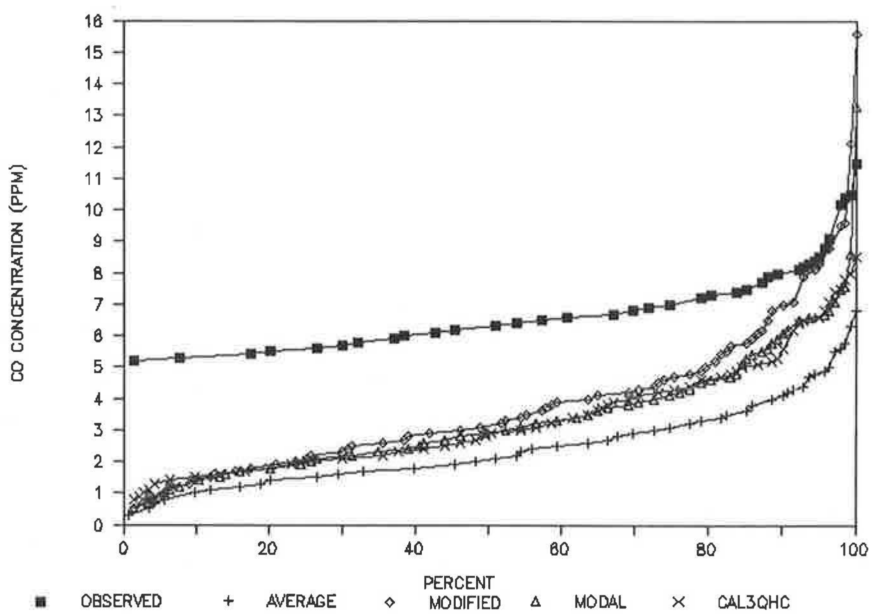


FIGURE 4 Cumulative frequency distribution, Site 2, CALINE3.

Sensitivity Analysis

Various sensitivity analyses were conducted to determine if any systematic error of bias that was affecting model performance could be discovered and corrected.

The stratified data sets for both sites included

- Three wind speed groups (<1.5 m/sec, 1.5 to 3 m/sec, >3 m/sec);
- Three wind direction groups (near parallel, near perpendicular, and other);
- Probe location (midblock versus intersection); and
- Three ranges of standard deviation of wind direction σ_θ (<25 degrees, 25 to 50 degrees, and >50 degrees).

Wind direction appeared to be the most likely source of error that could produce poor correlations for all the modeling scenarios. This fact was corroborated by the actual wind direction data measured at the different meteorological towers. Wind direction was measured at two locations at Site 1 and three locations at Site 2. In many instances, within the data set the average hourly wind direction measured was different from one location to another at the same site. At times, the winds were being channeled along each approach of the intersection toward one another. The variability of the wind direction was also reflected by σ_θ values that exceeded 25 degrees approximately 80 percent of the time, and exceeded 50 degrees approximately 33 percent of the time.

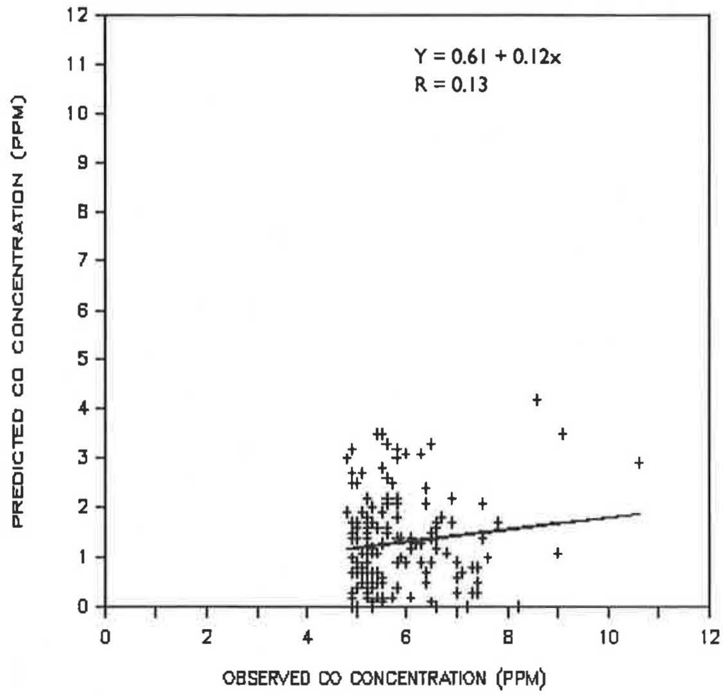


FIGURE 5 Scatter plot and regression line, CALINE3/Modal at Site 1.

It was not possible to determine to what extent this phenomenon was related to the siting of sensors that did not meet PSD monitoring guidelines. It was also possible that the street-level winds were quite different at the various approaches to the intersection of concern. Because both of the dispersion models required a single average wind direction for all the roadway links, this was a possible source of the error.

The results of the regression analysis indicated no improvement when the data were stratified by probe location or by wind direction. However, under low wind speed conditions, some of the modeling scenarios did exhibit better correlation than what occurred under the entire data set. The modeling scenarios also indicated some improvement for one of the σ_θ categories. However, this occurred for the high σ_θ set, which was expected to have the worst performance.

CONCLUSIONS

The performance of the models at the two monitoring sites was disappointing. On average, the models were not able to replicate observed concentrations and tended to underpredict.

At Site 1, the average speed model underpredicted most, typically by about a factor of 5. The performances of modified average speed and modal approaches were similar, with typical underprediction of about a factor of 4. The CAL3QHC method followed, with a typical underprediction of approximately a factor of 2.5. This model was the only one that was able to approach and exceed observed concentrations at the upper end of the concentration distribution. This range is where models need to predict best, so as to be able to predict CO concentrations reasonably, with a degree of conservatism. For this reason, the CAL3QHC model scenarios seemed to come closer to replicating the observed concentration distribution than the other scenarios at this site.

The performance outcome at Site 2 was basically the same in terms of general underprediction. The modal and modified average speed scenarios overpredicted at the upper end of the concentration distribution. Also, the modified average speed and CAL3QHC scenarios were reversed in terms of underprediction. At this site, CAL3QHC underpredicted more than the modified average speed scenario, for the reasons discussed. This approach is probably conservative at intersections that are not seriously congested (such as Site 1), yet is not conservative enough at those intersections where vehicles cannot get through the light in one cycle (such as Site 2).

Although the outcome from the first two intersection sites (work is ongoing at the other four study sites) indicates that the performance of the CAL3QHC, modal, or modified average speed models could yield conservative approximations of air quality, it is clear that the performances of all the models were less than desirable. Clearly, additional research needs to be done to determine how to improve intersection air quality modeling.

There are many uncertainties associated with this type of air quality modeling (*26–27*) and not enough resources and research have been devoted to this area over recent years. In this modeling effort, the uncertainties can be broken down into two basic areas—emissions and dispersion. The uncertainties of emissions modeling at intersections include whether or not modal emissions based on mid-1970s vehicle types are sufficiently accurate to depict today's vehicle fleet. (Indeed, some question whether the level of detail associated with modal emissions is necessary.) Regarding dispersion uncertainties, the proper characterization of dispersion in urban areas with dispersion models that were developed from test track data is a potential shortcoming, as is the proper characterization of meteorological parameters such as stability, wind speed, and direction in a highly turbulent and fluctuating urban environment. In addition, modeling efforts without the benefit of all necessary measured traffic data input have further uncertainties. For example, given the geometry, volume, and speed conditions, are the queue lengths and delay algorithms properly replicating those parameters? One, all, or some combination of the emission and dispersion (as well as traffic) uncertainties may be the cause of the models' apparent difficulty in replicating observed CO concentrations.

Immature CO mobile source modeling has tremendous implications for a variety of concerns. As more and more environmental scrutiny is given to various sources, such as shopping centers, highways, and housing developments, it could become needlessly more difficult to site these sources. In terms of state implementation plans, many unpopular and difficult measures may be implemented to reduce CO concentrations on the basis of modeling that is not up to the task.

Obviously, there are many benefits to be gained from improved CO model performance. These models should be able to perform at the same level of sophistication and confidence as stationary source models.

ACKNOWLEDGMENTS

The air quality meteorological monitoring was performed by ENSR Consulting and Engineering, Acton, Massachusetts, and TRC Environmental Consultants, Inc., East Hartford, Connecticut. Traffic monitoring was performed by Urbitran Associates, New York, New York.

REFERENCES

1. A. M. Sorro and J. Zamurs. The Use of Extremely Long Probes for Monitoring Carbon Monoxide. *Proc., 83rd Annual Meeting and Exhibitions, Air and Waste Management Association*, Pittsburgh, Pa., June 1990.
2. E. W. Nelson. Vehicle Air/Fuel Ratios During Standard Dynamometer Schedules and On-Road Driving: Measurement and Implications. *Journal of the Air Pollution Control Association*, Vol. 34, Nov. 11, No. 1984, p. 1, 139.
3. *On-Site Meteorological Program Guidance for Regulatory Modeling Applications*. Report EPA-450/4-87-013. U.S. Environmental Protection Agency, June 1987.
4. D. B. Turner. A Diffusion Model for an Urban Area. *Journal of Applied Meteorology*, Vol. 3, 1964, pp. 83–91.
5. *Ambient Monitoring Guidelines for Prevention of Significant Deterioration (PSD)*. Report EPA-450/4-87-007. U.S. Environmental Protection Agency, May 1987.
6. W. Petersen. *User's Guide for HIWAY-2, A Highway Air Pollution Model*. Report EPA-600/8-80-018. U.S. Environmental Protection Agency, 1980.
7. P. Benson. *CALINE3—A Versatile Dispersion Model for Predicting Air Pollutant Levels Near Highways and Arterial Streets*. Report FHWA/CA/TL-79/23. Office of Transportation Laboratory, California Department of Transportation, Sacramento, 1979.
8. *User's Guide to MOBILE4 (Mobile Source Emissions Factor Model)*. Report EPA-AA-TEB-89-01. U.S. Environmental Protection Agency, 1989.
9. PEI Associates, Inc. *Evaluation of CO Intersection Modeling Techniques*. U.S. Environmental Protection Agency, July 1989.
10. *TRB Special Report 209: The Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 1985.
11. *Intersection Midblock Model User's Guide*. Environmental Analysis Bureau, New York State Department of Transportation, Albany, 1982.
12. P. Kunselman. *Automobile Exhaust Emission Modal Analysis Model*. Report EPA-460/-3-74-005. U.S. Environmental Protection Agency, 1974.
13. W. Dabberdt and R. Sandys. *Guidelines for Air Quality Maintenance Planning and Analysis, Volume 9 (Revised): Evaluating Indirect Sources*. Report EPA-450/4-78-001. U.S. Environmental Protection Agency, 1978.
14. *User's Guide to the TEXIN2 Model—A Model for Predicting Carbon Monoxide Concentrations near Intersections*. Report FHWA/TX-86/283-2. Texas Transportation Institute, Texas State Department of Highways and Public Transportation, Austin, and FHWA, U.S. Department of Transportation, 1986.
15. *CALINE4—A Dispersion Model for Predicting Air Pollutant Concentrations Near Roadways*. Report FHWA/CA/TL-84/15. Division of Engineering Services, California Department of Transportation, Sacramento, 1984.
16. G. Schattanek, J. Kahng, and J. Soden. CAL3QHC, A Modeling Methodology for Predicting Pollutant Concentrations near Roadway Intersections. Presented at TRB Conference on Transportation and Air Quality, Denver, Colo., July 1990.
17. D. G. Fox. Judging Air Quality Model Performance. *Bulletin of the American Meteorological Society*, Vol. 2. 1981, pp. 599–609.
18. S. T. Rao et al. Resampling and Extreme Value Statistics in Air Quality Model Performance Evaluation. *Atmospheric Environment*, Vol. 19. 1985, pp. 1,503–1,518.
19. S. R. Hanna. Air Quality Model Evaluation and Uncertainty. *Journal of the Air Pollution Control Association*, Vol. 38. 1988, pp. 406–412.
20. W. B. Petersen, S. T. Rao, et al. Evaluation of the Performance of RAM with the Regional Air Pollution Study Data Base. *Atmospheric Environment*, Vol. 19. 1985, pp. 229–245.
21. S. T. Rao et al. Turbulent Diffusion Behind Vehicles: Evaluation of Roadway Models. *Atmospheric Environment*, Vol. 20. 1985, pp. 1,095–1,103.
22. S. T. Rao et al. An Evaluation of Some Commonly Used Highway Dispersion Models. *Journal of the Air Pollution Control Association*, Vol. 30, 1980, pp. 239–246.
23. S. T. Rao and J. R. Visalli. On the Comparative Assessment of the Performance of Air Quality Models. *Journal of the Air Pollution Control Association*, Vol. 31, 1981, pp. 851–860.
24. Stanford Research Institute. *NCHRP Report 245: Methodology for Evaluating Highway Air Pollution Dispersion Models*. TRB, National Research Council, Washington, D.C., 1981.
25. J. Zamurs. Comments on Mobile Source Modeling. Presented at 4th Conference on Air Quality Modeling, Washington, D.C., 1988.
26. R. D. Sculley. Vehicle Emission Rate Analysis for Carbon Monoxide Hot Spot Modeling. *Journal of the Air Pollution Control Association*, Vol. 39. 1989, pp. 1,334–1,343.
27. J. Zamurs. Intersection Carbon Monoxide Modeling. *Journal of the Air and Waste Management Association*, Vol. 40. 1990, pp. 769–771.

Publication of this paper sponsored by Committee on Transportation and Air Quality.

Model Calculation of Environment-Friendly Traffic Flows in Urban Networks

BEN H. IMMERS AND N. G. J. OOSTERBAAN

The environmental impact of traffic flows in urban networks is an increasingly serious problem. The possibilities are investigated of modifying route choice of traffic in urban networks to meet the standards for noise annoyance and the emission of pollutants. Therefore, for each link of the network an environmental capacity is defined, being the minimum capacity of a link resulting from the selected environmental standards. The maximum flow on a link is defined by the minimum of the environmental capacity and the free flow capacity. By assigning the traffic to the network using the equilibrium assignment technique, a desired pattern of flows is obtained that meets the environmental standards as far as possible. The model is applied to the network of the town of Ede-Bennekom, the Netherlands.

One of the drawbacks of increasing use of the automobile is its impact on the environment. The Dutch government has been giving this wide-ranging environmental issue a great deal of consideration, which has resulted in a set of policy measures. First, extensive regulations (e.g., the Noise Act) have been laid down as to the qualification and quantification of environmental annoyance. On the one hand, these regulations specify the limiting values, the standard values, and the preferred values within which particular environmental effects should be kept. On the other hand, standard methods are given by which these effects can be assessed and mapped, e.g., *Technical Aspects of Air Quality Regulations*, and the *Environmental Impact Map*. Second, research is being carried out on the kind of measures that will decrease undesirable and polluting side-effects of various activities (e.g., the introduction of petrol with low lead levels and the establishing of routes for the transport of noxious materials). From the mid-1980s, the Dutch government has adhered to the so-called "stand-still" principle. This principle implies that the levels of the various kinds of environmental pollution must (at least) not be increased.

The main subject is the environmental impact of road traffic. In order to illustrate its scale, in 65 percent of noise annoyance cases road traffic is the chief source; road traffic is responsible for more than half of the emission of carbon monoxide (70 percent), lead (80 percent), and nitrogen oxides (50 percent).

The predicted drastic increase in motor traffic (*1*) will lead to a more extensive use of the existing infrastructure, with all its consequences for the environs. Concern about the environment will result in the phased introduction of stricter standards with respect to environmental effects, making it obligatory for local authorities to issue reports and take particular measures, i.e., reorganizing town and transportation planning in existing situations or else adjusting it to new situations.

A model is especially well suited for determining an assignment (circulation) of traffic in an area in such a way that a favorable situation will arise with respect to meeting the various environmental standards. This approach merits considerable attention for the following reasons:

- The model indicates to what extent desired environmental standards can be met and on which road segments measures to this end will be required;
- By modifying traffic flow circulation, integral measures are taken at the source; by decreasing traffic at an environment-critical site, polluting factors will become less serious.
- Particularly in situations where the greatest problems may be expected to occur (urban agglomerations), alternative approaches are often scarce. Noise barriers cannot be built everywhere.
- By taking into account the future development of traffic, policymakers will be able to deal with the present situation and to anticipate expected developments.

Houtman and Immers (*2,3*) have recently developed a model by means of which an environment-friendly traffic assignment can be established on the basis of permissible noise levels, as specified in the Noise Act. This model was extended with a component for air quality requirements. The most important features on the previously developed model are briefly explained. Then the theoretical background and possibilities for integrating air quality requirements into the existing model are considered. The results of some sensitivity analyses and an application of the model are presented and analyzed, after which some conclusions and recommendations are drawn. An extensive report of the research project was provided by Oosterbaan (*4*).

NOISE ANNOYANCE

Traffic in residential areas may have considerable consequences for human well-being. Many local authorities are faced with the question of which measures to take to meet a desired environmental quality. In order to answer this question efficiently, a model has been developed at Delft University of Technology that optimizes noise annoyance in connection with accessibility.

Central to the model is the environmental capacity (X_a) for Road Segment a, defined as the capacity (veh/hr) resulting from the standards for the emission of noise [see Dutch legal

Department of Transportation Planning and Highway Engineering, Faculty of Civil Engineering, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, the Netherlands.

standard, calculation method I (5)]. Equation 1 was derived by Houtman and Immers (2).

$$X_a = \left(\frac{d}{p_l \cdot 10^{Y_l} + p_m \cdot 10^{Y_m} + p_z \cdot 10^{Y_z}}\right) \cdot 10^{L_{max}/10} \quad \text{(veh/hr)} \qquad (1)$$

where

Y = noise emission parameter;
$Y_l = 5.12 + 0.021u_l - \log_{10} u_l$;
$Y_m = 6.84 + 0.009u_m - \log_{10} u_m$;
$Y_z = 7.62 + 0.003u_z - \log_{10} u_z$;
u_l, u_m, u_z = average speeds [automobiles (l), medium-heavy traffic (m), and heavy traffic (z)], km/hr;
p_l, p_m, p_z = fractions [automobiles (l), medium-heavy traffic (m), and heavy traffic (z)];
d = distance facade to road axis, m; and
L_{max} = maximum noise level, dB(A).

Per road segment (link), the minima are determined of the free flow capacity (C_a, the theoretical capacity, which is environmentally unconstrained) and of X_a.

Figure 1 shows the BPR travel time functions for C_a (solid line) and X_a (dotted line). The minima of both capacities are taken as the capacities determining the travel time when assigning the traffic to the network (the new capacity shown in Figure 2).

If the traffic is assigned to the network according to an equilibrium assignment technique (user-optimal travel time minimization with additional constraints as to the noise level), a desired pattern of traffic flows is obtained. This desired pattern indicates an equilibrium situation as regards travel times and meets the legal standards of noise emission as closely as possible.

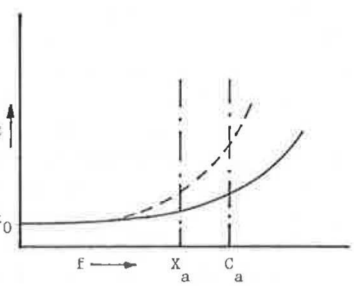


FIGURE 1 Environmental capacity X_a and free flow capacity C_a.

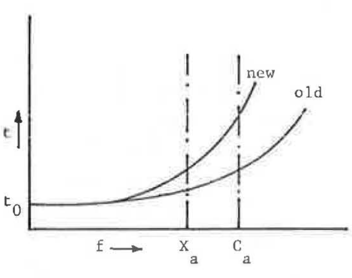


FIGURE 2 Arithmetical capacity, the minimum of the environmental capacity and free flow capacity.

The suitability of the model was tested by applying it to a real site, i.e., Ede-Bennekom (2).

As mentioned, the emission of noxious materials (gases) is another environmental threat. Similarly as for noise annoyance, the extent to which air quality requirements may be incorporated into the environmental capacity (and so into the assignment process) was investigated, so that with respect to this problem, too, the standards can be met as much as possible.

AIR QUALITY REQUIREMENTS

The Dutch government has both set the air quality standards and established the computation methodology of air pollution (6,7). The model calculations were limited to those for nitrogen oxide and carbon monoxide. Lead is left aside because it was assumed that the introduction of fuels with a reduced lead level will considerably decrease the emission of lead.

In *Technical Aspects of Air Quality Regulations* (6,7), standards are set and specifications are given of the model for calculation of air pollution from road traffic (CAR), developed by TNO.

The contribution of traffic to air pollution as a result of carbon monoxide (CO) emission is calculated as follows:

$$[CO] = N * E_s * \Phi * F_{region} * F_b + [CO]_{bg} \quad (\mu g/m^3) \qquad (2)$$

$$E_s = (1 - p_v) * E_p + p_v * E_v \quad (\mu g/m\text{-sec}) \qquad (3)$$

$$[CO]_{bg} = R_{CO} + S_{CO} * F_a \quad (\mu g/m^3) \qquad (4)$$

where

$[CO]$ = resulting concentration of carbon monoxide ($\mu g/m^3$);
N = number of vehicles per 24 hr;
E_s = average emission (μg/m-sec);
p_v = fraction of nonautomobile traffic;
E_p, E_v = emission parameters for private cars and other traffic, respectively, depending on speed (μg/m-sec);
Φ = dilution factor (sec/m^2) depending on the type T of ground cover and the distance s between curb and road axis;
F_{region} = meteorological correction factor with respect to regional differences in wind velocity;
F_b = correction factor with respect to the Type I_b of street in relation to the presence of trees;
$[CO]_{bg}$ = background concentration of CO ($\mu g/m^3$);
F_a = distance between Road Segment a and the edge of the built-up area (km); and
R_{CO}, S_{CO} = constants for the measurement of $[CO]_{bg}$ ($\mu g/m^3$ and ng/m^4, respectively).

The type of trees variable I_b takes on the following values:

1 = No trees or a few trees within 30 m from the road axis.
2 = One or more rows of trees with distance between trees <15 m. There are openings between the crowns of the trees.
3 = Crowns of trees hang over the road.

The ground cover variable T takes on the following values (see Figure 3):

1 = Flat terrain, no or a few buildings or trees within 100 m from the axis of the road.
2 = All types other than 1, 3, 4, or 5.
3 = More or less unbroken buildings on both sides of the road. Distance to axis of road is less than 3 times the height of the buildings.
4 = The same as 3 but distance from buildings to the road axis is less than 1.5 times the height.
5 = Buildings on one side of the road. Distance to the road axis is less than three times the height. On the other side of the road, there are no buildings or they are far away from the road.

In the direct exhaust emission of nitrogen oxides, there is a large quantity of the rather harmless nitric oxide NO (for which there are no air quality requirements) and a comparatively small quantity of the much more toxic nitrogen dioxide (NO_2). However, by a chemical equilibrium reaction with ozone (O_3), NO is fairly soon converted into NO_2 according to the reaction

$$NO + O_3 \rightleftarrows NO_2 + O_2 \qquad (5)$$

On the basis of the emissions of nitrogen oxides (NO_x), the noxious concentration of NO_2 can be calculated as follows:

$$[NO_x] = \gamma * N * E_s * \Phi * F_{region} * F_b \qquad (\mu g/m^3) \qquad (6)$$

$$[NO_2] = \theta * [NO_x] + \frac{\delta * [O_3]_{bg}}{H + [NO_x]} + [NO_2]_{bg} \qquad (\mu g/m^3) \qquad (7)$$

$$E_s = (1 - p_v) * E_p + p_v * E_v \qquad (\mu g/m\text{-sec}) \qquad (8)$$

$$[NO_2]_{bg} = R_{NO_2} + S_{NO_2} * F_a \qquad (\mu g/m^3) \qquad (9)$$

$$[O_3]_{bg} = R_{O_3} + S_{O_3} * F_a \qquad (\mu g/m^3) \qquad (10)$$

where

$[NO_x]$ = concentration of nitrogen oxides from direct emission ($\mu g/m^3$);
$[NO_2]$ = resulting concentration of nitrogen dioxide ($\mu g/m^3$);
θ = coefficient (= 5/6) relating $[NO_2]$ to $[NO_x]$;
γ, δ, H = parameters dependent on the type T of ground cover;
$[NO_2]_{bg}$ = background concentration of NO_2 ($\mu g/m^3$)
$[O_3]_{bg}$ = background concentration of O_3 ($\mu g/m^3$)
R_{NO_2}, S_{NO_2} = constants for calculating $[NO_2]_{bg}$;
R_{O_3}, S_{O_3} = constants for calculating $[O_3]_{bg}$;
E_p, E_v, E_s = emission parameters (differing in numerical value from those for the case of CO); and
Φ, F_b = coefficients (differing in numerical value from those for the case of CO).

Analogous to the way noise annoyance was dealt with, the maximum flow per link is established for which the various air polluting materials (gases CO and NO_2) do not exceed the relevant air quality standard.

From Equation 2, the maximum day flow (Q) can be determined on the basis of the air quality standard $[CO]_{max}$ for CO from the expression

$$Q(CO) = \frac{[CO]_{max} - [CO]_{bg}}{E_s * \Phi * F_{region} * F_b} \qquad (veh/day) \qquad (11)$$

The environmental capacity $X(CO)$ for the emission of CO can then be computed from the expressions

$$X(CO) = \frac{Q(CO)}{10} \qquad \text{(for one-way roads)} \qquad (12)$$

$$X(CO) = \frac{Q(CO)}{20} \qquad \text{(for two-way roads)} \qquad (13)$$

During peak hours, the flow is assumed to be 10 percent of the 24-hr flow, and for two-way roads, 5 percent of the total flow (in both directions).

For the maximum concentration ($[CO]_{max}$), a temporarily increased limiting value of 15 000 $\mu g/m^3$ was established. This value will be decreased in four stages down to 6,000 $\mu g/m^3$ in the year 2000.

The maximum day flow on the basis of the standard for NO_2 can be calculated as follows:

$$Q(NO_2) = \frac{[NO_x]_{max}}{P * E_s * \Phi * F_{region} * F_b} \qquad (veh/day) \qquad (14)$$

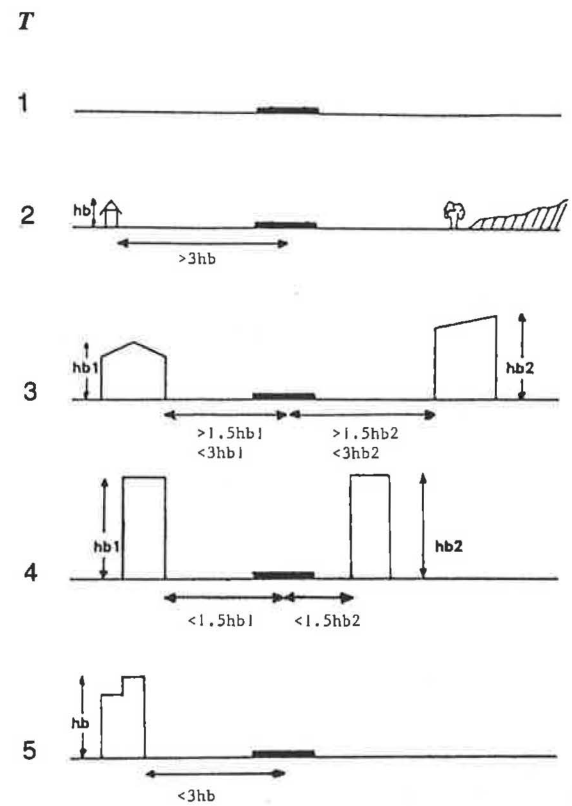


FIGURE 3 Description of ground cover variable *T*.

$[NO_x]_{max}$ can be calculated from Equation 7 as follows:

$$[NO_x]_{max} = [-B + (B^2 - 4AC)^{1/2}]/2A \quad (15)$$

where

$$A = \gamma \quad (16)$$

$$B = \gamma * H + \delta * [O_3]_{bg} + [NO_2]_{bg} - [NO_2]_{max} \quad (17)$$

$$C = H * ([NO_2]_{bg} - [NO_2]_{max}) \quad (18)$$

For $[NO_2]_{max}$, a temporarily increased limiting value of 160 μg/m³ was established. This value will be decreased in two stages down to 135 μg/m³ in the year 2000.

The environmental capacity for the emission of nitrogen oxides can be calculated by dividing the maximum day flow by 10 (for one-way roads) or by 20 (for two-way roads), respectively.

$$X(NO_2) = \frac{Q(NO_2)}{10} \quad \text{(for one-way roads)} \quad (19)$$

$$X(NO_2) = \frac{Q(NO_2)}{20} \quad \text{(for two-way roads)} \quad (20)$$

Incorporation of the air pollution standards in the model means that a great number of additional data must be introduced. Road segment coefficients will have to be introduced relative to

- Type I_b of trees,
- Type T of ground cover,
- Weighting coefficient J indicating the number of dwellings or residents per road segment.

Incorporation of J, the number of dwellings or residents per road segment, is based on the assumption that the effects of air pollutants should be weighted higher for road segments that are densely populated (many dwellings or residents along the road) than for road segments that are sparsely populated.

Furthermore, for every road segment a the distance up to the edge of the built-up area (F_a, see Equations 4, 9, and 10) is to be measured (see Figure 4). A number of additional nodes located at the edge of the built-up area were therefore defined (see Figure 5). The distance of a road segment up to the edge of the built-up area can then be computed by means of the x and y coordinates from the expression

$$F_a = \text{Max}(D_{pI}, D_{pII}) \quad (21)$$

where

$$D_{pI} = \text{Min}(d_n) \quad (22)$$

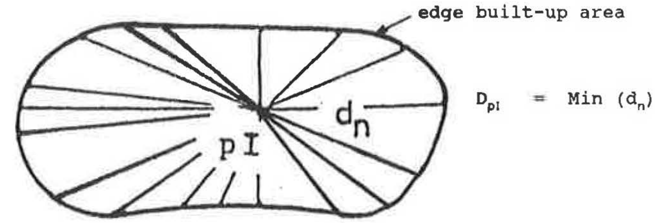


FIGURE 5 Distance from a node to the edge of the built-up area.

As with the noise aspect, it is possible to omit capacity computations with relation to the air quality for particular links, if, for instance, these have a clear traffic flow function. Per link, the capacity may now be calculated, i.e., the minimum value of the free flow capacity (C_a) and the environmental capacities for, respectively, noise, air quality for CO, and air quality for NO_2.

Per environmental aspect, an indication of the capacity reduction of the entire network can be obtained, on the basis of the limits set by the environmental standards, from the expression

$$\text{RESCAP}(I) = \left[\sum_{j=1}^{\text{NLINK}} \frac{\text{Min}(\text{ENVCAP}_j(I), \text{CAP}_j)}{\text{CAP}_j} * 100\right] / \text{NLINK} \quad (23)$$

where

RESCAP(I) = remaining network capacity with respect to environmental Aspect I, percent;
CAP_j = free flow capacity of Link j;
$\text{ENVCAP}_j(I)$ = environmental capacity of Link j with respect to environmental Aspect I; and
NLINK = number of links.

For the total of environmental aspects considered, the remaining network capacity can be calculated as follows:

$$\text{RESCAP(O)} = \left[\sum_{j=1}^{\text{NLINK}} \frac{\text{Min}(\text{INFO2}_j, \text{CAP}_j)}{\text{CAP}_j} * 100\right] / \text{NLINK} \quad (24)$$

where

RESCAP(O) = remaining network capacity with respect to all the environmental aspects, percent; and
INFO2_j = lowest environmental capacity of Link j.

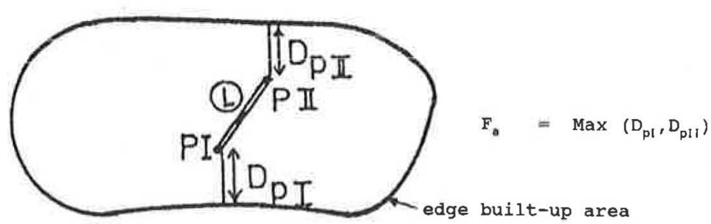


FIGURE 4 Distance from a link to the edge of the built-up area.

ANALYSES

Sensitivity Analysis

By means of a sensitivity analysis, the effects of the following five parameters on the environmental capacity were investigated for CO and NO_2.

1. Distance to edge of built-up area F_a,
2. Average speed u,
3. Percentage P_v of other vehicles (nonautomobiles),
4. Type of trees I_b, and
5. Type of ground cover T.

As to the values of the other parameters, three situations were distinguished in the calculations (with regard to air quality):

- The other parameters have favorable values,
- The other parameters have average values, and
- The other parameters have unfavorable values.

Some of the results of these sensitivity analyses are shown in Figures 6 and 7. The figures indicate the effects of modifications of average speed and percentage of other traffic. For CO, the maximum flow (environmental capacity) increases with average speed and with the percentage of other traffic; as for NO_2, on the other hand, the maximum flow decreases in both cases. In all cases, the limiting value for NO_2 determines the outcome.

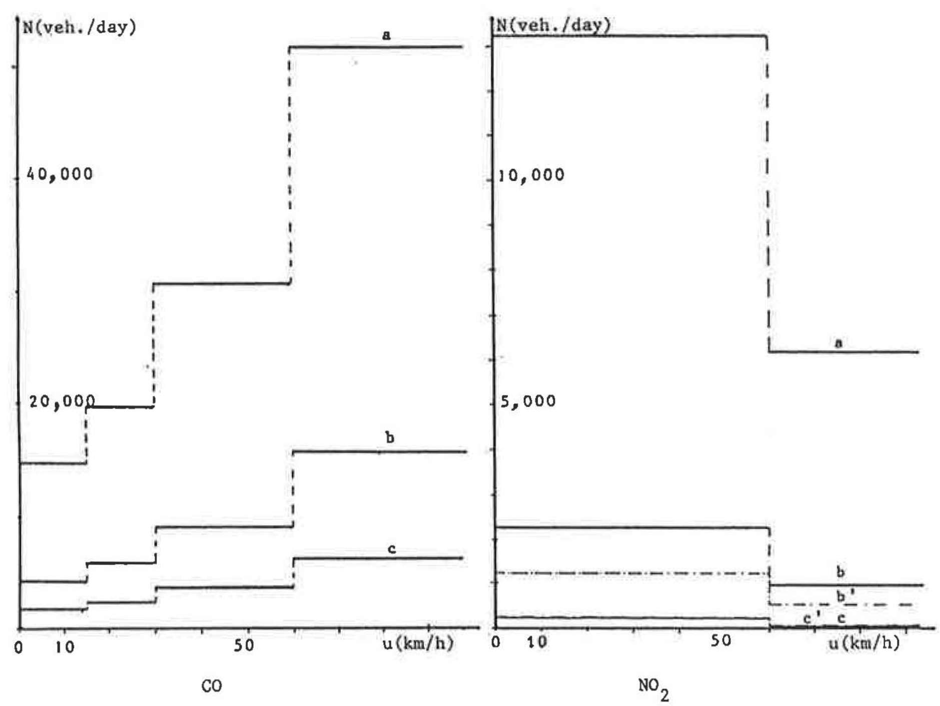
Other important results of the sensitvity analysis follow.

- The sensitivity of the maximum flow for a particular parameter increases with the extent to which the values of the other parameters are favorable,
- The maximum flow decreases with increasing ground cover,
- The maximum flow decreases with increasing distance to the edge of the built-up area, and
- The maximum flow decreases with an increasing number of trees along the road. (This is a short-term effect. In the long term, the presence of trees will have a positive effect on air quality.)

Model Calculations

In order to test the model, the network and the OD table of Ede-Benekom for the year 1995 were used, as drawn up by IWIS/TNO (*8*).

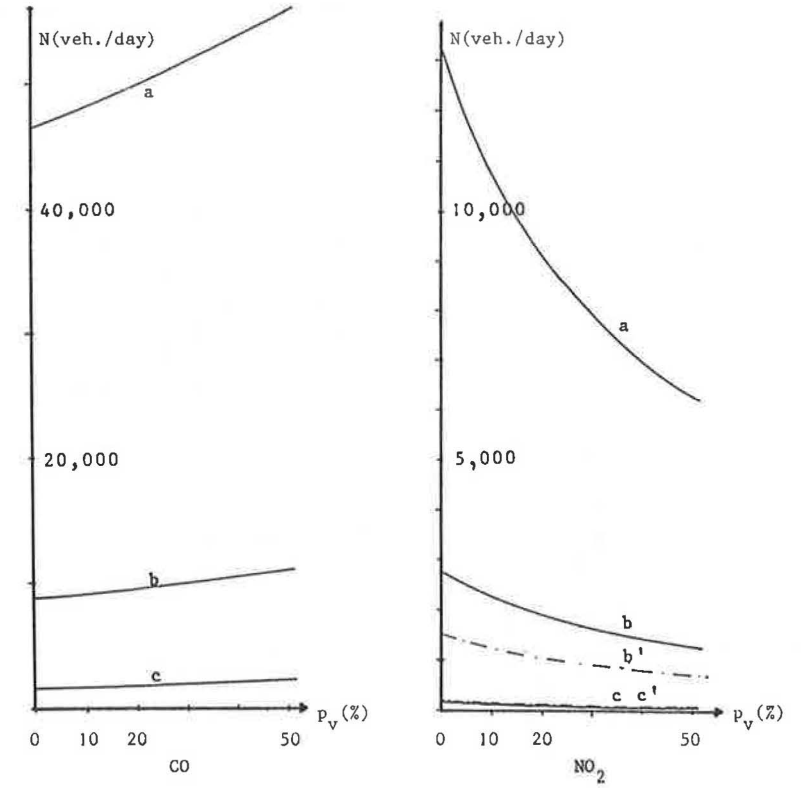
The network contains 264 nodes, comprising 57 centroids, and 859 one-way links (see Figure 8). The categorization in



Standard CO = 15000 $\mu g/m^3$, standard NO_2 = 160 $\mu g/m^3$, F_{region} = 1.05

a	F_a = 0 km,	p_v = 0%,	T = 1,	I_b = 1
b	F_a = 5 km,	p_v = 10%,	T = 3A,	I_b = 2
b'	F_a = 5 km,	p_v = 10%,	T = 4,	I_b = 2
c	F_a = 10 km,	p_v = 30%,	T = 3B,	I_b = 3
c'	F_a = 10 km,	p_v = 30%,	T = 4,	I_b = 3

FIGURE 6 Relation between maximum day flow (*N*) (from the point of view of the standards for CO and NO_2) and the average speed (*u*).



Standard CO = 15000 $\mu g/m^3$, standard NO_2 = 160 $\mu g/m^3$, F_{region} = 1.05

a	: F_a = 0 km,	$u = u_a$,	T = 1,	I_b = 1
b	: F_a = 5 km,	$u = u_b$,	T = 3A,	I_b = 2
b'	: F_a = 5 km,	$u = u_b$,	T = 4,	I_b = 2
c	: F_a = 10 km,	$u = u_d$,	T = 3B,	I_b = 3
c'	: F_a = 10 km,	$u = u_d$,	T = 4,	I_b = 3

FIGURE 7 Relation between maximum day flow (*N*) (from the point of view of the standards for CO and NO_2) and the percentage of other traffic (p_v).

types of road was made according to, among other things, road cross section, average flow, and average speed. Per category of link, a fixed free flow speed, free flow capacity, and percentage of medium-heavy and heavy traffic were determined. The distance facade to road axis, ground cover, type of trees, and distance road segment to the edge of the built-up area vary for each link. The variables *F* region (meteorological correction coefficient for wind velocity), $[CO]_{max}$, and $[NO_2]_{max}$ have the same values for the whole network.

For the area, calculations were carried out for four scenarios:

1. A traditional equilibrium assignment, not taking into account environmental standards.
2. An equilibrium assignment taking into account the noise standard [60 dB(A)].
3. An equilibrium assignment taking into account the noise standard and the temporarily increased limiting values with respect to the air quality standards (CO: 15 000 $\mu g/m^3$; NO_2: 160 $\mu g/m^3$).
4. An equilibrium assignment taking into account the noise standard and very strict requirements regarding the air quality standards (CO: 3 000 $\mu g/m^3$; NO_2: 80 $\mu g/m^3$). These standards are far below the maximum concentrations established for the Netherlands for the year 2000 (CO: 6 000 $\mu g/m^3$; NO_2: 135 $\mu g/m^3$, see (*6,7*)].

The influence of the various quality standards on the traffic pattern is presented in Table 1 in which the remaining capacities [RESCAP(*I*)] for differing scenarios have been calculated.

In Ede-Bennekom, the influence of the air quality standards does not become noticeable until the very strict requirements are applied (Scenario 4).

Table 2 presents the traffic flow effects and the environmental effects resulting from the various assignments. (For explanation of class boundaries, see Table 3.) In order to get an impression of a possible oversaturation all over the network, the average saturation degree for the busiest directions of all roads together and the quietest directions of all roads together is determined. Scenario 2 will always exhibit larger values than Scenario 1 because most saturation degrees depend on the environmental capacity (which usually is smaller than the free flow capacity). Although noise annoyance is a subjective matter, investigations have indicated a remarkable

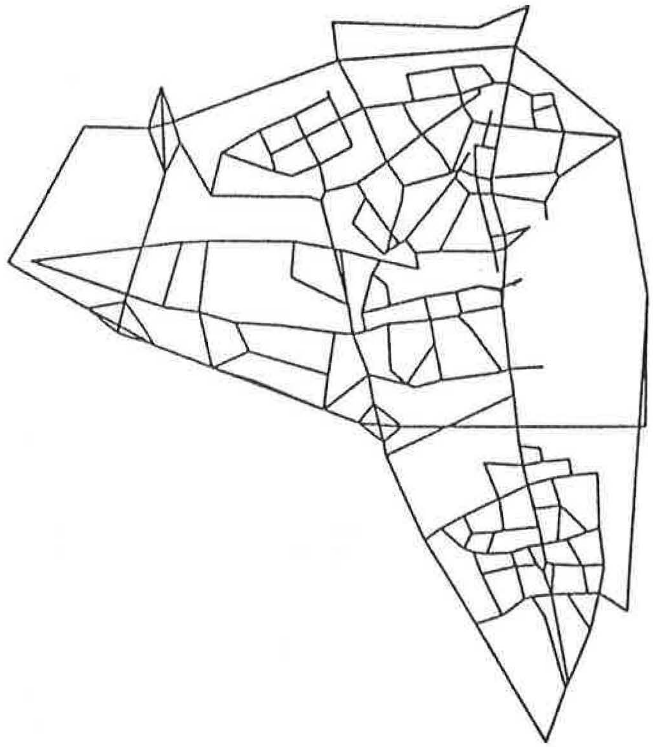


FIGURE 8 Network of Ede-Bennekom (*8*).

TABLE 1 REMAINING CAPACITIES [RESCAP(*I*), *I* = 1–4] IN THE FOUR SCENARIOS (PERCENT)

Criterion	Scenario 1	2	3	4
Noise	100	55.8	55.8	55.8
CO	100	100	100	95.1
NO_2	100	100	99.9	55.5
All criteria together	100	55.8	55.8	50.8

correspondence between the number of people strongly annoyed and the noise level (*2*).

For noise levels beyond the noise standard of 60 dB(A), noise brackets of 3 dB(A) have been defined, because this bracket size corresponds to doubling of the traffic flow.

The class boundaries for CO and NO_2 are more or less arbitrary (6 percent of the temporarily increased standard). No evidence could be obtained about the seriousness of the effects of an excess of the emission standards for CO and NO_2.

The results of the calculations as presented in Table 2 indicate the following:

- Air quality standards are of little or no significance (first column, Table 2, Scenarios 2 and 3). Scenario 3 is identical to Scenario 2. However, introduction of the noise standard does have a distinct influence on the traffic flow (compare Scenarios 1 and 2).
- When applying the very strict air quality standards (second column, Table 2, Scenario 4), a qualitative improvement and a quantitative deterioration occur (compare Scenarios 2, 3, and 4). The total number of road segments and the summed length of the road segments where one or more air quality standards are exceeded increase, but the magnitude of these excesses decreases. These occurrences are caused by a redistribution of the traffic from road segments far above the standard to road segments with a flow smaller than the environmental capacity. Remarkably, this redistribution manifests itself more strongly in the case of the less critical CO.
- When only the noise standard is applied, the air quality is also drastically improved (compare Scenarios 1 and 2, very strict standards). The added incorporation of the air quality standards has a varying effect on the assessments of the acoustic quality (somewhat more annoyance but fewer road segments where standards are strongly exceeded). These findings seem to indicate the need for compromises between the various quality requirements.
- Incorporation of the very strict air quality standards leads to an increase in the total vehicle mileage. In order to prevent environmental standards being exceeded, detours are made. The average number of vehicles (related to the free flow capacity) also increases. This increase is partly caused by the increase in mileage and partly by the assignment of the traffic to alternative routes with a low free flow capacity.

CONCLUSIONS AND RECOMMENDATIONS

On the basis of the tests performed, the following conclusions were drawn:

- Assuming (temporarily) increased limiting values, air quality standards do not play any part in assessing environ-

TABLE 2 RESULTS OF APPLYING STANDARDS

		Temporarily increased standards			Very strict standards		
		Standard Noise = 60 dB(A) Standard CO = 15000 $\mu g/m^3$ Standard NO_2 = 160 $\mu g/m^3$			Standard Noise = 60 dB(A) Standard CO = 3000 $\mu g/m^3$ Standard NO_2 = 80 $\mu g/m^3$		
Scenario		1	2	3	1	2	4
Noise	Class 1	24	55	55	24	55	59
	Class 2	30	24	24	30	24	32
	Class 3	28	13	13	28	13	8
	Class 4	3	0	0	3	0	0
	Class 5	0	0	0	0	0	0
	Index 1[1]	43628	41139	41139	43628	41139	42442
	Index 2[2]	43145	39747	39747	43145	39747	42190
	Index 3[3]	7794	6713	6713	7794	6713	6971
CO	Class 1	0	0	0	20	7	12
	Class 2	0	0	0	1	1	0
	Class 3	0	0	0	0	0	0
	Class 4	0	0	0	0	0	0
	Class 5	0	0	0	0	0	0
	Index 1[1]	0	0	0	4850	1190	3110
	Index 2[2]	0	0	0	1826488	626967	300356
	Index 3[3]	0	0	0	113059	71318	25762
NO_2	Class 1	0	0	0	40	50	55
	Class 2	0	0	0	12	14	19
	Class 3	0	0	0	18	4	4
	Class 4	0	0	0	7	1	1
	Class 5	0	0	0	1	1	0
	Index 1[1]	0	0	0	21310	19050	22400
	Index 2[2]	0	0	0	290734	167672	165858
	Index 3[3]	0	0	0	49130	23326	21191
Kilometrage (km)		89750	108717	108717	89750	108717	113503
Satur. degree (%		20.5	44.8	44.8	20.3	44.8	52.5
of free flow cap)		15.7	34.5	34.5	15.7	34.5	40.7

[1]) Total length of all road sections (a_j) exceeding Noise Standard (L_{norm}), CO standard resp. NO_2 standard.

[2]) For each class the total length of the road sections (see [1]) is multiplied by the excess of the standard (for noise the annoyance factor c_i). For CO: $\sum_j a_j * ([CO]_j - \text{Standard CO})$; for NO_2 : $\sum_j a_j * ([NO_2]_j - \text{Standard } NO_2)$.
The noise standard L_{max} = 60 dB(A) corresponds with an annoyance factor equal to 1; an excess of x dB(A) results in an annoyance factor $c_i = \exp.(0.1143x)$ (see also [2]).

[3])See [2]) but length of road is replaced by number of dwellings along the road.

TABLE 3 CLASS BOUNDARIES FOR ENVIRONMENTAL ASPECTS

Class Boundaries for Noise Annoyance [dB(A)], L_{norm} = 60 dB(A)	
L_{norm}	$< \text{Class } 1 \leq L_{norm} + 3$
$L_{norm} + 3$	$< \text{Class } 2 \leq L_{norm} + 6$
$L_{norm} + 6$	$< \text{Class } 3 \leq L_{norm} + 9$
$L_{norm} + 9$	$< \text{Class } 4 \leq L_{norm} + 12$
	$\text{Class } 5 > L_{norm} + 12$
Class Boundaries for CO Emissions ($\mu g/m^3$), $[CO]_{max}$ = 15 000 $\mu g/m^3$	
$[CO]_{max}$	$< \text{Class } 1 \leq [CO]_{max} + 1000$
$[CO]_{max} + 1000$	$< \text{Class } 2 \leq [CO]_{max} + 2000$
$[CO]_{max} + 2000$	$< \text{Class } 3 \leq [CO]_{max} + 3000$
$[CO]_{max} + 3000$	$< \text{Class } 4 \leq [CO]_{max} + 4000$
	$\text{Class } 5 > [CO]_{max} + 4000$
Class Boundaries for NO_2 Emissions ($\mu g/m^3$), $[NO_2]_{max}$ = 160 $\mu g/m^3$	
$[NO_2]_{max}$	$< \text{Class } 1 \leq [NO_2]_{max} + 10$
$[NO_2]_{max} + 10$	$< \text{Class } 2 \leq [NO_2]_{max} + 20$
$[NO_2]_{max} + 20$	$< \text{Class } 3 \leq [NO_2]_{max} + 30$
$[NO_2]_{max} + 30$	$< \text{Class } 4 \leq [NO_2]_{max} + 40$
	$\text{Class } 5 > [NO_2]_{max} + 40$

mental annoyance in the Ede-Bennekom network. (In a compact town, this may be entirely different).

- Application of strict standards clearly indicate that incorporation of environmental annoyance with respect to air pollution influences the results of the assignment process.
- The effects of the various environmental aspects on the traffic assignment process may be graded as follows: (a) noise, (b) air quality for NO_2, and (c) air quality for CO.
- If different environmental standards are incorporated into the assignment process two effects may be observed: (a) in many cases adjusting the assignment process for one environmental aspect leads to an improvement of other environmental aspects that were not incorporated into the assignment process; and (b) by adding an environmental aspect, the improvements in the traffic circulation calculated for another environmental aspect may be partly nullified.

Besides the recommendations with respect to noise annoyance as stated by Houtman and Immers (*2*), it may, for a greater understanding of the problems of air quality and the applicability of the relevant software, be interesting to investigate the following points:

- Application of the model in a situation more problematic as regards air pollution, e.g., a medium-size, compact town (e.g., Delft).
- Research into the magnitude of the annoyance as a function of the extent to which the air quality standards are exceeded. (In the present model, the excess itself is taken as a measure of the annoyance experienced).
- Incorporation of the increasing travel times in calculations of environmental effects.

The results of the model calculations indicate where a reduction of the free flow capacity should be implemented to meet the environmental standards. Possible measures that can be applied to reduce the capacity of a road section are reducing the number of lanes, reducing the lane width, displacing the road axis, changing the signal settings, and introducing speed ramps (*9*).

REFERENCES

1. R. Hamerslag, B. H. Immers, and W. H. Scheltes. The Influence of Economic Recession and Changing Fuel Prices on Car Mobility in the Netherlands. *Proc., 4th World Conference on Transport Research*, Vol. 2, 1986, Vancouver, British Columbia, pp. 1030–1050.
2. J. W. Houtman and B. H. Immers. A Traffic Assignment Model to Reduce Noise Annoyance in Urban Networks. In *Transportation Research Record 1143*, TRB, National Research Council, Washington, D.C., 1988, pp. 17–21.
3. J. W. Houtman. *Research into the Effects of Speed Reducing Measures on Travel Time in Networks*. Research Report, Department of Transortation Planning and Highway Engineering, Delft University of Technology, Delft, The Netherlands, Feb. 1985.
4. N. G. J. Oosterbaan. *A Model for the Calculation of Environmental-Friendly Traffic Flows in Relation to the Standards for Noise Annoyance and the Emission of Noxious Gasses*. Research Report, Department of Transportation Planning and Highway Engineering, Delft University of Technology, Delft, The Netherlands, June 1988.
5. Ministry for Housing, Physical Planning, and Environmental Protection. *Noise Level Calculation Methods (in Accordance With the Noise Annoyance Act)*. Staatsuitgeverij, Hague, The Netherlands, 1981.
6. Ministry for Housing, Physical Planning, and Environmental Protection. Technical Aspects of Air Quality Regulations, Carbon Monoxide and Lead. *Publication Series Air*, Vol. 52, Staatsuitgeverij, Hague, Oct. 1986.
7. Ministry for Housing, Physical Planning and Environmental Protection. Technical Aspects of Air Quality Regulations, Nitrogen Dioxide. *Publication Series Air*, Vol. 64, Staatsuitgeverij, Hague, The Netherlands, Aug. 1987.
8. IWIS/TNO. *Traffic Study Ede/Bennekom 1979*, Vols. A and B (1980) and Final Report, Phase 2, Vols. I and II, Delft, The Netherlands, 1981.
9. J. W. Houtman. Possibilities for Speed Reduction of Car Traffic in Urban Networks. Research Report, Department of Transportation Planning and Highway Engineering, Delft University of Technology, Delft, The Netherlands, Feb. 1985.

Publication of this paper sponsored by Committee on Transportation and Air Quality.

Preparation of Highway Vehicle Emission Inventories

JOHN H. SUHRBIER, SAMUEL T. LAWTON, AND JOSEPH A. MORIARTY

Passage of the Clean Air Act Amendments of 1990 has initiated a new round of state implementation plans, and their associated current- and future-year mobile source baseline emission inventories. Existing highway vehicle emission inventory practices are assessed in 15 urban areas throughout the country, and these existing capabilities are compared with recommended EPA guidance. Network-based travel demand model approaches are most frequently used and come closest to meeting urban area needs. However, existing approaches often are deficient in their estimation of highway vehicle speeds and also are inconsistent with methodological approaches used for national-level emissions inventories. A variety of institutional problems, though, may present more significant obstacles to the preparation of satisfactory mobile source inventories than any technical limitations. These problems include funding limitations, institutional fragmentation, lack of available technical expertise, and an unfortunately high level of staff turnover.

Passage of the Clean Air Act Amendments of 1990 has occasioned a significant resurgence of interest in the analytical methodologies required to support transportation-air quality related analyses. Issues associated with the preparation of base- and future-year mobile source emissions inventories are explored, particularly the estimation of highway vehicle miles of travel (VMT) and speed.

An important component of the Clean Air Act is the set of planning and analysis activities required by the states and designated nonattainment areas. State implementation plans (SIPs) are comprehensive documents that detail current emissions and air quality conditions, and demonstrate commitments to implement measures that are sufficient to achieve the national ambient air quality standards by a designated date.

As part of an SIP, emissions inventories are developed for all significant mobile, stationary, and area sources of pollutants. Such inventories are developed for both a base year and a projected future year and provide the baseline condition against which the effectiveness of alternative control policies can be measured. Mobile source inventories include all transportation sources of emissions: highway vehicles; off-highway vehicles, aircraft, railroads, and marine. The highway portion of a mobile source emissions inventory classifies vehicles by type (e.g., automobile, light truck, heavy truck, diesel) and estimates both VMT and speed by vehicle and roadway classification.

In addition to urban area and state emission inventories, EPA maintains a variety of national emission inventories covering each of the principal criteria pollutants. Separate trends inventories are maintained as well so as to be able to quickly determine year-to-year variations in emissions.

From a national perspective, it would be desirable if all emission inventories used consistent and identical methodologies, and produced consistent results. Currently, this is not the case and considerable effort is devoted to resolving inconsistencies.

At the state and urban area level, numerous practical problems can be encountered in preparing mobile source emission inventories. For example, it has been customary in past inventory analyses to assume that vehicle travel speeds are the same in future years as they are in the current or base analysis year. This generally has been justified on the basis that highway capacity will expand proportionally with the growth in vehicular travel. Realistically, it is often also based on the lack of information on which to base any other assumption.

Increasingly, however, the assumption of constant vehicular speed over time is being called into question. This change is resulting from an acknowledged cap on highway expenditures, a recognition that congestion is increasing in many urban areas, and results from urban transportation planning analyses. For example, the Southern California Association of Governments is projecting that the average daytime freeway speed in the Los Angeles area will decline from 35 to 20 mph over the next 20 years in the absence of full implementation of an ambitious transportation management and improvement program.

Analyses of the FHWA's Highway Performance Monitoring System (HPMS) traffic data base indicate similar findings. For urban areas, vehicle speeds generally are projected to decrease over time (Figure 1). Speeds decline the most for non-Interstate freeways and expressways, and next most severely for other principal arterials. As expected, the magnitude of the projected speed decrease becomes larger with lower levels of highway funding.

Accurate estimates of speed are particularly important given the nonlinearity with which emissions increase with decreasing vehicle speeds below 20 mph.

RESEARCH AND OBJECTIVES

EPA (*1*) prepared a revised guidance document for the preparation of mobile source emission inventories. This document covers all transportation modes: off-highway vehicles, highway vehicles, aircraft, railroads, and marine vessels. For highways, it provides a comprehensive discussion of factors affecting highway emissions; the use of MOBILE4, vehicle

Cambridge Systematics, Inc., 222 Third Street, Cambridge, Mass. 02142.

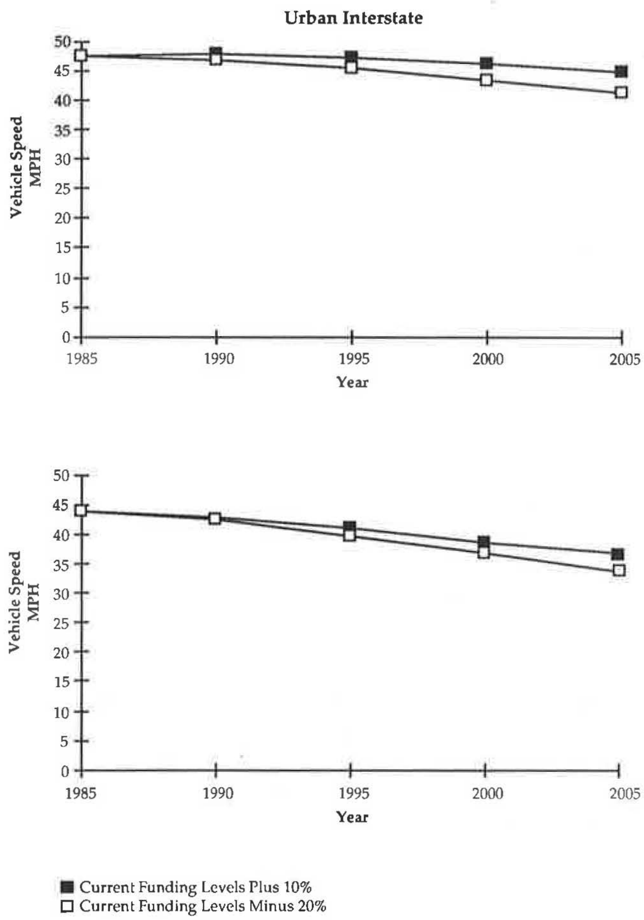


FIGURE 1 Projected changes in automobile operating speeds over time (3).

inspection, maintenance, and antitampering programs; and the urban transportation planning process and associated computerized modeling systems. Particular questions that remained after preparation of this guidance document included the following:

- What is the ability of the states to comply with EPA's emission inventory guidance with respect to the use of urban transportation planning data and model systems?
- How can the interface between transportation models and mobile source emission models be improved?

To gain an improved understanding of specific urban area issues, representatives from environmental and transportation agencies involved in the preparation of mobile source emissions inventories were interviewed to determine the problems being encountered in preparing the traffic-related portion of a mobile source emission inventory. These contacts represent the following urban areas: Detroit, Phoenix, San Francisco, Los Angeles, Atlanta, Tampa, Houston, Pittsburgh, Portland (Oregon), Boston, Denver, Chicago, Nashville, Fresno, and Hartford.

On the basis of the list of nonattainment areas for ozone and carbon monoxide, these cities represent different sizes, levels of sophistication, and technical capability. The objective was to sample the problems and solutions from different tiers of urban areas that could be encountered in preparing mobile source inventories, with the overall goal of providing a reasonably quick assessment of the range of methodologies currently being used, the problems being encountered, and the opportunities available for improvement. The approach centered on open-ended telephone conversations with representatives of urban area agencies involved in preparing the mobile source inventory for their nonattainment area.

AVAILABLE VMT AND SPEED METHODOLOGIES

A diverse range of methodologies is being used to estimate VMT and vehicle speeds, as presented in Table 1 (*2*). These methodologies can result in a wide range of accuracy. The following is an overview of the VMT and speed methodologies that are being used by the interviewed urban areas. In general, the techniques that have the ability to analyze and document disaggregate components offer better precision. Urban areas that rely on more coarse methodologies are less precise.

VMT Estimation Methodologies

Network

Network-based VMT methodologies rely on volumes generated as output from an urban area transportation network model. Those models generally are either FHWA's Urban Transportation Planning System (UTPS) or a variation of UTPS operating on a microcomputer. Volumes are calculated internally by the network modeling process used by many Metropolitan Planning Organizations (MPOs). VMT is calculated by multiplying the volumes by link distances. Depending on the specific urban area, links coded into a network model generally correspond to the following facility types: freeway, expressway, major arterial, minor arterial, and major collector. The amount of links coded into a network and the facility classification depend on the sophistication of the model. VMT can be calculated only for the facility types (links) that are coded.

There are some significant problems associated with network-based techniques. VMT generated by non-network-coded facilities, such as minor collectors or residential streets, are not accounted for in a typical transportation network model. Typically, 10 to 15 percent of low-speed VMT data are generated by these non-network-coded facilities. If the emission inventory area is larger than that covered by the transportation network, VMT generated outside the network also will not be accounted for.

Another important problem is the fact that VMT generated by a network model corresponds to a typical weekday, with no recognition of seasonal variation in volumes. Most urban network models have the capacity to simulate average daily, peak (congested), and off-peak volumes.

Presently, stand-alone urban area transportation network demand models are reasonably accurate at simulating regional transportation demand. However, network models under certain conditions may not accurately simulate the variations in link level traffic operations during peak conditions. An important advantage of the network model approach is the fact

TABLE 1 MATRIX OF SPEED-VMT METHODOLOGIES

SPEED:[1]	Network: V/C HCM	Distance Matrix	Capacity Restraint	Speed Runs	MOBILE4 Defaults
VMT:[1]					
Network	Phoenix Detroit Hartford		San Francisco Nashville Houston Los Angeles		
Network/Hybrid		Portland	Atlanta Denver		
HPMS				Boston	Fresno
Manual					Chicago
Fuel Apportionment			Tampa		

[1] The specific highway vehicle emissions inventory procedures utilized in the Pittsburgh urban area are still to be decided upon by the relevant transportation and environmental agencies

that urban area network models represent a repository of important link level data and information. With refinement, these data can be used to simulate link level traffic operations, an important requisite for an accurate emissions inventory.

Network Hybrid

The network hybrid approach generically identifies urban areas that are attempting to deal with problems of coordinating their inventory over many jurisdictions, or the lack of non-network-coded facility data inside or outside the urban network boundary.

For instance, Portland, Oregon, has a unique problem in which the inventory area encompasses nonmodeled rural areas and modeled urban areas of two states, with three agencies involved in traffic modeling. Because the models of these agencies are not coordinated and rely on different data bases and networks, facility types must be aggregated to reach some level of consistency. This process of aggregating to the metropolitan statistical area (MSA) level significantly complicates the process of a link-by-link network-based analysis VMT estimation method.

HPMS

Some urban areas that were contacted are coordinating their emission inventory with the data collected for the HPMS, a traffic data base developed and maintained by the FHWA (*3,4*). State transportation departments are responsible for collecting the sample roadway travel data for various facility types and reporting it to FHWA. FHWA compiles the information and expands it for analysis.

The problems with HPMS data are as follows: (a) HPMS data are representative at the national, state, and larger metropolitan areas only—samples may be too small to be representative for smaller rural and urban areas; (b) VMT is generated at a regional level; and (c) only higher classification roadways are included in samples (i.e., for highways, arterials, and collectors), and urban and rural local VMT data may be limited or nonexistent.

Manual

The manual method represents the most aggregate or coarse method of developing VMT input data for a mobile source emission inventory. Chicago is the only example of this method. VMT is provided by 11 facility types by county ranging from highway to local roads on the basis of the Illinois Department of Transportation's traffic counting program. No attempt was made to assign VMT to facilities within the county or account for the effect of congestion. Future estimates of VMT are extrapolated on the basis of past trends. This method represents the most consistent coverage from one single data source but lacks the ability to document link level operational characteristics.

Fuel Apportionment

This method uses fuel sales as a factor in determining VMT. Daily average fuel consumption multiplied by miles per gallon represents total VMT. Tampa, Florida, is using a variation of this method to determine VMT, where total VMT minus network-derived VMT equals nonnetwork VMT.

Speed Estimation Methodologies

An underlying premise in performing this work has been that estimates of link-level traffic volumes are regarded as being both reasonably reliable and more accurate than link-level speeds produced by urban transportation model systems such as UTPS or microcomputer-based systems. Speeds developed through the modeling process serve as a means of allocating trips to balance the network. As such, they really are more of an input rather than an output of the model. On the other hand, the vehicle speed data in *Highway Capacity Manual* (HCM) (*5*) are considered to be more reliable than estimates based on network models. In free flow (uncongested) conditions, network and HCM speed estimates appear to be similar. However, under congested conditions in which vehicle speeds may drop below 20 mph, these estimates begin to differ significantly. Estimating speeds under congested conditions is crucial in air quality analysis because large shares of VMT occur during peak conditions and because of the nonlinear relationship between speed and emissions below 20 mph.

Network: Volume to Capacity

This technique uses the link-level volume output from the transportation network model and establishes a volume-to-capacity (*v/c*) ratio for each link that is then used to estimate a speed value for that link. The primary advantages of this method are that the relationship between *v/c* and speed is based on data from actual operating conditions compiled from the HCM, and it closely simulates traffic operations at the link level.

This method appears to offer a higher level of precision because it documents the daily variations in travel speeds due to congestion which significantly influences the quantity of mobile emissions produced. The primary disadvantage to this method is the cost and coordination necessary to customize an existing network model to replicate link level operational characteristics.

Distance Matrix

Portland, Oregon, calculates speed on the basis of the time it takes for a vehicle to travel between the various zone centroids. Speeds from the traffic assignment process are not used. This method had been selected because of criticism of network-derived travel speeds.

A problem with this method is that speed data for urban and rural non-network-coded facilities are limited. For the Portland SIP, factors to disaggregate statewide local road VMT data were used to assign VMT on facilities not coded in the network. This method does not account for variations in link-level operational characteristics that may have an impact on emissions.

Capacity Restraint

The capacity restraint method is a type of traffic assignment algorithm that attempts to model congested speeds during peak conditions for all facility types. The capacity restraint methodology is used as a default formula in many urban areas' traffic assignment models. The capacity constraint function is based on the assumption that speed decreases as congestion increases.

However, the unique manner in which the capacity restraint function manipulates speed for a particular link does not necessarily represent an accurate estimate of speed for that link but rather a value that optimizes traffic assignment over the entire congested network.

The basic problem with the capacity restraint function is that it does not document well the variations in traffic operating conditions, especially on very congested links. It may be unreasonable to assume that a single formula is able to accurately estimate speed for facility types with very different operating characteristics. A more appropriate procedure includes the use of separate methods for estimating speed for each facility class for each condition, i.e., peak versus off-peak conditions. A primary advantage of the capacity restraint method is that it is institutionalized at many MPOs.

Speed Runs

Some urban areas, such as Boston, use manually collected speed runs for various facility types as input to the mobile source emission inventory. This method is based on samples of representative facility types. Speed runs, however, can be both costly and labor intensive. The samples need to be very large to account for daily variation in travel. If sample size is small, an average speed by facility class is used, which may neglect temporal variations in travel.

MOBILE4 Defaults

The Chicago component of the Illinois SIP is using the MOBILE4 internal default speed of 19.6 mph for all facility types, corresponding to the average speed of the Federal Test Procedure (FTP). Using 19.6 mph as a default value may overestimate vehicle speeds on congested freeway and arterial links that are characteristic of larger urban areas during peak-period operations. This, in turn, may underestimate the amount of mobile emissions generated.

SYNTHESIS OF KEY ISSUES

The following is a list of 12 key issues synthesized from the urban area interviews. In conducting these interviews, it was found that institutional arrangements (or lack of) may have more of an impact on the quality of the mobile source inventory than technical methodology. Therefore, the list includes problems of institutional and administrative arrangements as well as technical considerations.

Adequacy of Transportation Network Model for Air Quality Emissions Inventories

A transportation network model establishes the analytical basis for assessing future transportation needs and evaluating projects that will satisfy those needs. Emphasis generally is on planning for major corridor-level projects, and on projecting traffic volumes on major radial and circumferential roadways. Considerably less importance is given to travel speed and to minor or local streets. A stand-alone transportation network model normally needs customization to be an effective tool for the emission inventory process.

Phoenix, Detroit, and Los Angeles are three examples in which customized procedures have been developed to analyze variations in link-level VMT and travel speed, which are important requisites for an accurate inventory. In Phoenix, peak-period spreading of traffic volumes is explicitly considered. Forecast link volumes are then converted to *v/c* ratios for use in a special speed estimation procedure. The Los Angeles DTIM computer program calculates link-level emissions, accounting for VMT and speed on each link rather than aggregated by facility type. These modifications have proven to be both workable and valuable, and need to be undertaken by a larger number of urban areas.

Validity of Network Based Speed Estimates

Significant concern was expressed in a number of urban areas with respect to the validity of the speed estimates produced by network-based traffic assignment procedures. In response, a number of urban areas have developed special speed estimation routines which calculate *v/c* ratios, and then use either HCM or locally derived relationships to convert *v/c* into speed estimates. A direct network-derived speed estimate may overestimate link speeds because the capacity restraint algorithm used is based on the equilibrium adjustment necessary to obtain a reasonable region-wide trip allocation rather than on observed speeds of the roadway link.

San Francisco, Nashville, Atlanta, and Denver are urban areas identified as using network-based speed estimates. The problem is especially crucial during congested conditions. Network-based methodologies are fairly reasonable at simulating freeflow or uncongested conditions. During congested periods, however, speeds associated with links with heavy volumes may be overestimated. In terms of the emission inventory, this condition would underestimate emissions produced by that link.

Coverage of Local Roads

Transportation network models typically do not include minor and local roads, yet an emissions inventory requires that all travel be covered. Available data on nonnetwork modeled local road characteristics may be nonexistent or limited. Procedures used to estimate local road VMT data may be based on judgment and be of questionable accuracy.

To replicate local street travel, a variety of techniques currently are being used. For example, the Hartford inventory assumes the distance of all traffic generated by centroid connectors, a point where all traffic is loaded on the network, to be 0.96 mi long and operating at uncongested level of service C for peak and off-peak conditions. Denver doubles its network-covered local road VMT data to approximate total local road conditions. A consistent and accurate procedure needs to be developed because lower-classification roadways, especially in urban areas, generally operate at low speeds and may be susceptible in certain conditions to congestion during peak periods.

Inconsistency in Accounting for Peak and Off-Peak Travel

There is a lack of a consistent methodology being used to disaggregate VMT and speed data by time of day. The failure to reasonably account for congested (high-volume, low-speed) conditions may underestimate emissions. Using default speeds may not be appropriate considering that most urban area speeds are lowest on facilities that have the highest volume of traffic.

At the lower end of the precision spectrum in terms of accounting for the variation in peak and off-peak travel are urban areas such as Boston, Fresno, and Chicago that use HPMS or aggregate manual methods for estimating VMT and speed. Their use of facility-type VMT data by county and the use of an average daily speed value for that class of facility does not adequately take into account peak-period congestion.

The methods developed by urban areas using a transportation network based technique, such as San Francisco, are able to disaggregate volumes and speeds by a.m.-p.m. peak periods and for daily conditions. This method falls within the middle of the precision spectrum because the methodology to estimate volumes appears to be more reliable than the methodology to estimate speeds.

Of the urban areas interviewed, the Phoenix and Detroit methods represent the highest level of precision. Their methodologies relate speed to volumes on a link-by-link basis, closely simulating real traffic operating conditions for peak as well as off-peak periods.

Lack of Current Transportation Data

The validity of a transportation network model is directly related to how frequently travel behavior data are collected and integrated into the modeling process. Most larger MPOs conducted large-scale household travel surveys in the 1960s and 1970s. Because these efforts require extensive resources, most urban areas cannot replicate these early survey efforts given the current lack of funding. In order to make up for this deficiency, network models increasingly are being validated incrementally with the use of census data and smaller-scale surveys.

The Los Angeles area is a typical example in which their current model was developed using surveys and roadside interviews in 1967, and which subsequently has been updated using information from a 1976 survey and the 1980 census. Representative and up-to-date travel surveys are a critical component of a regional emissions inventory because they provide the base line travel assumptions against which an areawide transportation model can be calibrated.

Problems of County-Wide Reporting

For areas that are not using network model traffic outputs or areas in which the air quality planning area is larger than the transportation network, traffic volume data may be limited. Facility classifications may be inconsistent with the network model. Care must be taken in mixing empirical data with model results. The cost of obtaining detailed information from nonmodeled areas, however, may not be worth the expense because of the small fraction of total emissions these areas produce.

If consistent methodologies are desired, there may be a tendency to aggregate and report VMT and speed data to the least common denominator. HPMS statistics may be used, and inventory results may be reported only at the county level. This approximation will not document the nuances of different urbanized and rural area traffic behavior for a specific facility type.

In the cases of Illinois and Massachusetts, for example, VMT data are inventoried at the county or regional level by facility type with an areawide average travel speed assigned to that class of facility.

Lack of Alternative Approaches

With appropriate adjustment and postprocessing, transportation network demand models may be the best available tool for predicting traffic inputs to the mobile emission inventory process. However, for areas without traffic demand models, alternative acceptable approaches need to be agreed on to assess VMT and speeds. Alternative methodologies can be defined that are consistent with the magnitude of the problem. If congestion is a major component of the transportation system, a detailed link assessment could be required. In areas having relatively little travel, simpler methodologies could be used.

Portland, Oregon, is an example of an inventory area in which both alternative and a mixture of approaches are necessary. Their inventory area encompasses both nonmodeled rural areas and modeled urban areas of two states. Three separate agencies are involved with network modeling, but with none of the models being consistent and in a transferable format for a regional link-level emissions inventory.

Requirements of a Mobile Source Inventory are not Consistent with the Metropolitan Area Transportation Planning Process

The urban transportation planning process, and its associated set of computerized (UTPS) travel demand models, was not explicitly designed to develop mobile source emissions inventories. This inconsistency of objectives has a number of manifestations. The typical day in terms of average traffic used for transportation planning purposes will not correspond to the same time period that should be used for either a CO or VOC emissions inventory.

The urban transportation planning process used by larger MPOs typically results in a regional transportation plan with a single 25-year planning horizon. In contrast, the forecast emission inventory for an SIP is required at 5-year increments. There is no easy mechanism to accurately interpolate transportation data every 5 years. Accurate 5-year forecasts require a separately coded transportation alternative, something that normally is not done.

A related issue is what transportation alternative to use as the basis for the future-year base emissions condition. The preferred alternative may satisfy demand, but is it realistic considering funding restraints? The preferred alternative may also overestimate VMT. The no-build alternative may likewise underestimate VMT.

The base year for transportation planning purposes usually will not be the same as the air quality base year desired by EPA. This means that VMT and other travel projections cannot be directly translated into EPA terms. Ideally, an entirely new set of transportation analyses should be produced, but this is a time-consuming and expensive task. Thus, transportation and emission inventory analyses are frequently out of synchronization in terms of their base and horizon analysis years.

The Tampa inventory process is currently confronted with this issue of what network model to use. The choice is an out-of-date model that would probably underestimate VMT or wait for a calibrated model. The transportation planning and the mobile source emission inventory processes ideally should be developed in tandem with each other.

Available Expertise

Full-scale transportation network modeling requires greater sophistication and capability than is available in many areas. The urban transportation planning process requires extensive expertise in computer programming and operations research. Typically, only the large urban areas have these resources. The computer interface between urban network models and MOBILE4 may require customization that is beyond the resources of smaller MPOs.

Many smaller urban areas are just trying to financially cope with the mechanics of the routine transportation modeling process, i.e., data collection, network coding, calibration, etc., and do not have resources to customize their transportation model for air quality purposes. Tampa is an example of an urban area that is attempting to provide a new model to meet EPA criteria. It is estimated, however, that it will take at least 18 months for this work to be completed.

Staff Turnover

Lower-level staff are usually responsible for running MOBILE4, and may have only a limited understanding of transportation data and complex urban area travel forecasting models. Typically, once they have been trained and experienced, they move on to higher positions. Internal expertise is not institutionalized over the long term.

Problems of staff turnover and inexperience were identified in both the Detroit and Boston interviews. It is assumed that most urban areas are confronted with these issues because of the large amount of time since the last emissions inventory. A related issue is the fact that air quality planning staff may not be trained to appreciate that temporal variations in VMT and speed inputs in the MOBILE4 model can have significant influence on the amount of emissions that are produced.

Funding Limitations

State and local agencies are being asked to assume more responsibility in completing a more detailed mobile source inventory for a larger geographic area, and to do this for less money.

Atlanta and Houston are examples of urban areas that are confronted with the situation of producing a more detailed emissions inventory for a larger geographic area without the benefit of EPA Section 175 funding. During the early 1980s, the Atlanta area produced an emissions inventory for seven counties, which approximated the area covered by the local MPO transportation network model. The current emissions inventory covers a total of 12 counties, in which 8 are included and 4 are not included in the network model.

Institutional Fragmentation

Because there no longer is a formal funding mechanism in place to coordinate transportation and air quality planning, relationships between air quality agencies and the designated transportation MPOs are now largely restricted to informal contact. Development of a mobile source inventory requires a large number of agencies in either a direct or review role. Input from other agencies is frequently accepted on faith, with little understanding of how it was developed and only minimal concern with the consistency or accuracy of underlying assumptions. In some cases, transportation agencies that were responsible for running the emissions model in the 1970s and early 1980s are currently not involved.

According to representatives from the Chicago Area Transportation Study (CATS), the formal institutional relationship that existed in the 1970s and early 1980s between the environmental agency responsible for running the emission models and CATS has ended. Currently, only informal and infrequent data exchange occurs. Consequently, the current Illinois SIP will have only limited urban area expertise to complete their inventory. Reliance instead is being placed on state level data developed by the Illinois Department of Transportation.

RECOMMENDATIONS

These observations imply the following recommendations with respect to the preparation of mobile source emission inventories:

1. The traditional four-stage, urban transportation planning, travel demand model systems are the best available current methodology for urban area level inventories, but generally are not totally satisfactory in their current form for air quality purposes. Refinements and a postprocessor vehicle speed estimation capability need to be added. In addition, improved vehicle fleet information will be required as more emphasis is placed on the production of clean alternative fuels and low-emission vehicles.
2. Emission inventory procedures should not be oriented to urban and rural conditions, but to the categories of nonattainment severity as defined by the new Clean Air Act. Different methodological approaches will be appropriate for different urban areas, and possibly even within an area. In some situations, a hybrid of procedures will be appropriate; for example, in a situation in which the nonattainment area is significantly larger in size than the geographic area covered by the computer-coded transportation network. Standardization is an admirable, but probably neither an obtainable nor even a necessary objective.
3. Mobile source inventory methodologies should support future as well as current-year baseline projections. In addition, whenever possible, the same quantitative methodology as is used for preparation of the inventory also should be used for forecasting the effectiveness of alternative transportation control strategies. Although this is normally routine in transportation analyses, it is not always the standard procedure in stationary and area source analyses.
4. Monitoring or tracking of emission trends will become increasingly important at the urban area level with the new Clean Air Act. This monitoring will have to relate to the overall emissions inventory, but be able to be efficiently performed on an annual basis. Even more important, though, annual monitoring will have to be conducted so as to be able to determine the effectiveness of individual measures relative to a base case condition.
5. The desired level of inventory accuracy and disaggregation should dictate the choice of inventory methodology. If grid-based urban area dispersion modeling is going to be done, then mobile source emission inventory methodologies that are accurate at a zonal level of disaggregation will be necessary. This almost always will imply use of a UTPS-style approach. However, air shed modeling is exceedingly expensive and time consuming, and may not always be needed for analysis purposes. In these cases, non-UTPS approaches may be sufficient.
6. Mobile source inventory procedures that are appropriate at the national level will not be satisfactory at the urban area level. Different methodologies should continue to be used for national than for urban area inventories, with the inevitable but understandable inconsistencies in their results.
7. Institutional and resource considerations are potentially more significant barriers to achieving satisfactory emissions inventories than are any technical problems. Priority, care, and sensitivity need to be devoted to establishing a long-term cooperative working relationship of shared responsibilities between transportation and air quality agencies at the state and local levels of government.

In responding to the Clean Air Act Amendments of 1990, care must be taken not to let the mobile source portion of an SIP become a resource-intensive modeling exercise that loses sight of policy considerations and implementation objectives. The ultimate objective of the SIP analytical process is not a plan or an elegant analytical exercise. It is the implementation of action programs that will contribute to attainment of the National Ambient Air Quality Standards. This implementation-oriented objective must be kept in mind as states and urban areas undertake the development of a new cycle of mobile source emission inventories.

ACKNOWLEDGMENT

The research described in this paper was sponsored by the U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards.

REFERENCES

1. *Procedures for Emission Inventory Preparation—Volume IV: Mobile Sources.* Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, July 1989.
2. Cambridge Systematics, Inc. *Transportation-Related Emission Inventory Preparation.* Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, March 1991.
3. *Highway Performance Monitoring System Field Manual.* Office of Highway Information Management, FHWA, U.S. Department of Transportation, Dec. 1987.
4. *Highway Performance Monitoring System Analytical Process.* Office of Planning, FHWA, U.S. Department of Transportation, Dec. 1987.
5. *Special Report 209: Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 1985.

The findings and recommendations reported in this paper are those of the authors and do not necessarily reflect policy of the sponsoring agencies.

Publication of this paper sponsored by Committee on Transportation and Air Quality.

Managing Trucks for Air Quality: Current Work in Progress

CHRIS NELSON, SARAH SIWEK, RANDALL GUENSLER, AND KELLY MICHELSON

In most areas of California, motor vehicles continue to produce significant amounts of emissions that result in photochemical smog. The smog problem is exacerbated in the South Coast Air Basin (which includes Orange County and major portions of Los Angeles, Riverside, and San Bernardino Counties), a result of local topography and weather conditions. The ozone problem of the South Coast Air Basin is so severe that experts estimate emissions of hydrocarbons must be cut by as much as 82 percent to meet the national ambient ozone standards. Although air pollution levels are not as severe in other areas in California, almost every urban area is currently violating, or close to violating, ambient air quality standards for ozone, nitrogen oxides, or particulate matter. Heavy-duty vehicles are significant contributors to the emission inventory in urban areas. The history of truck-related transportation control measures, current transportation control measures under consideration in California, and uncertain effects of proposed measures are reviewed.

In most areas of California, motor vehicles continue to produce significant amounts of emissions that result in photochemical smog (*1*). Local topography and weather conditions exacerbate the problem in the South Coast Air Basin, which includes Orange County and major portions of Los Angeles, Riverside, and San Bernardino counties. The ozone problem is so severe in the South Coast Air Basin that emissions of hydrocarbons may have to be cut by 82 percent to meet the national ambient ozone standards (*2*). Though air pollution levels are not as severe elsewhere in California, almost every urban area is violating or close to violating ambient air quality standards for ozone, nitrogen oxides, or particulate matter. Heavy-duty vehicles contribute significantly to emissions in urban areas. The history of truck-related transportation control measures, current transportation control measures under consideration in California, and uncertain effects of proposed measures are reviewed.

BACKGROUND

1984 Olympics

The 1984 Summer Olympics, held in Los Angeles, presented a unique challenge to transportation and air quality planners. In anticipation of massive traffic jams and unhealthful air pollution levels, numerous transportation control measures (TCMs) were implemented. These measures included increased ridesharing and moving of commuter traffic to off-peak periods. This effort was marketed as a temporary solution through a massive public relations effort (*3*). Neither congestion nor air quality violations were noted (*4*).

One of the strategies that relieved traffic congestion was a voluntary reduction in truck traffic during anticipated peak periods. The truck trip reduction program had several elements. Restrictions on night deliveries were lifted, aided by union cooperation in accepting regular wages for night work. State laws allowing the night delivery of certain commodities were enacted, and a public information campaign to persuade the trucking industry of the need to adjust routes and activities was used (*3*). Some businesses also increased inventory to reduce the need for deliveries during the Olympics period. Visual counts revealed less truck traffic during peak periods. Peak periods, in this case, were those periods before, during, and after Olympic events and did not correlate exactly with normal peak traffic periods. Evening truck traffic increased, leading to the conclusion that truck trips were shifted out of peak periods, rather than that an outright reduction in trips occurred. Figures 1a and 1b compare traffic by time period, during Olympics and non-Olympics periods.

Traffic moved smoothly and efficiently during the 1984 Los Angeles Olympics and air quality during the Olympics surpassed all expectations (*4*). The direct effect of truck strategies is difficult to isolate, because numerous strategies were used to reduce all traffic. A high level of cooperation and public awareness was achieved during this period, including business using trucks, perhaps partially based on the knowledge that the control measures were temporary.

Trucks as a Source of Motor Vehicle Emissions

There are numerous sources of on-road motor vehicle emissions, and their contribution to the emission inventory in California is much higher than in many other states. Sources generally include automobiles (both commute and recreational trips), trucks, and buses. Overall, trucks constitute a small percentage of total traffic volume; however, they still contribute a significant portion to the mobile source emissions inventory (*5*). Table 1 presents their contribution by air basin for California. Note that the truck contribution to the emission inventory is especially pronounced for NO_x.

C. Nelson and S. Siwek, South Coast Air Quality Management District, 9150 Flair Drive, El Monte, Calif. 91731. R. Guensler, California Air Resources Board, 1102 Q Street, P.O. Box 2815, Sacramento, Calif. 95812. K. Michelson, Lockheed Information Management Services Company, 40 North Central Ave., Suite 2250, Phoenix, Ariz. 85004.

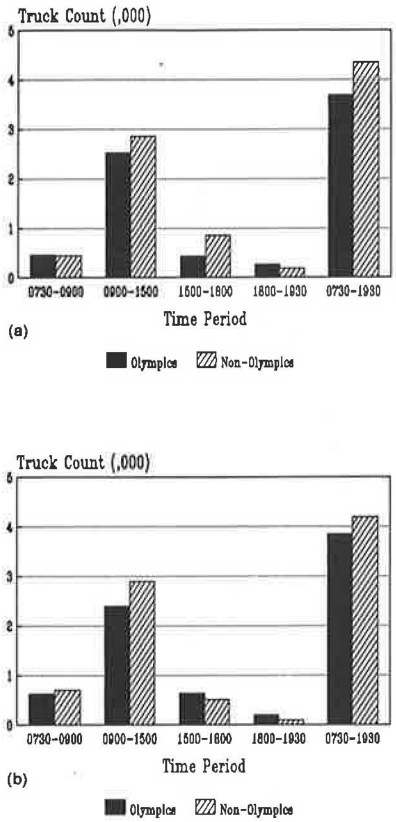


FIGURE 1 Truck traffic on I-10 east of I-110 by day, direction, and time period (*3*): (a) eastbound, and (b) westbound.

TABLE 1 ESTIMATED EMISSION CONTRIBUTIONS OF HEAVY-DUTY TRUCKS (5,500 lb OR HEAVIER) TO TOTAL 1987 EMISSIONS BY AIR BASIN (IN PERCENT)

Air Basin	ROG	NO_x	CO	PM10[a]
South Coast	4	18	12	3
San Diego	4	15	9	2
Sacramento Valley	4	28	7	2
San Francisco Bay Area	3	17	9	3
San Joaquin Valley	3	24	9	2
Statewide	4	21	8	2

[a]Particulate matter less than 10 μm in diameter. Values do not include PM10 associated with resuspended road dust.

What is a Heavy-Duty Truck?

The definition of a heavy-duty truck is not standard. These vehicles can be defined by gross vehicle weight rating (GVWR), gross vehicle weight (GVW), or unladen weight. The GVWR is the weight rating assigned by the vehicle manufacturer and is used by federal and state agencies to determine the applicability of motor carrier safety and air pollution control regulations. GVW is the estimated hauling weight reported by the vehicle operator to regulating agencies for weight-distance tax purposes. It is possible that GVW is either under- or overreported by the operator, depending on the tax structure. Unladen weight is used by the California Department of Motor Vehicle registration purposes. The unladen weight does not account for the load that the vehicle will be hauling. If desired, any weight criteria can be further disaggregated by number of axles, engine size, and fuel type (gasoline, diesel, methanol, etc.).

A legislative working group established a definition of heavy-duty trucks as "any commercial vehicle with a gross vehicle weight rating (GVWR) or 8,501 pounds or greater" (*6*).

The City of Los Angeles originally defined trucks as vehicles with GVW greater than 26,000 lb. Later, the definition was modified to include three or more axles. In their latest version (October 1990), the City has dropped the weight classification completely. This is intended to aid in enforcement by police unfamiliar with weight classifications.

For purposes of this discussion, heavy-duty trucks will be divided into two classes: (a) 8,501 lb to 26,000 lb GVW, and (b) greater than 26,000 lb GVW. Class I generally includes trucks with two axles and gasoline-powered engines. Class II vehicles generally have three or more axles and diesel-powered engines. These categories are used because originally, the City of Los Angeles proposed using GVW for their program and because data are available for GVW categories. The GVW categories were also used in the south coast AQMD's survey to determine the potential regulated population for heavy-duty truck regulations. However, it should be noted that the definitions used by the California Air Resources Board for vehicle emission regulatory purposes are in GVWR. Figure 2 shows the number of vehicles in the South Coast Air Basin and their approximate contribution to the emissions inventory.

Regulating Trucks

The California Air Resources Board has the responsibility for regulating tailpipe emissions from motor vehicles, including heavy-duty trucks. Statewide controls include new-vehicle emission certification standards and in-use inspection and maintenance emission standards. Although local agencies may not, under state law, regulate the emission rates of motor vehicles, they were granted authority to regulate vehicle activity in 1988.

In the South Coast Air Basin, the first vehicle use control strategy (transportation control measure) implemented was an employer-based commute trip-reduction regulation. Regulation XV, adopted by the South Coast Air Quality Management District (SCAQMD) in December 1987, required employers of 100 or more individuals to develop a vehicle trip reduction plan applicable to commute hours and to provide employees with incentives to rideshare.

Regulators are now beginning to examine transportation control measures that will affect heavy-duty trucks. Local air quality districts are developing transportation control strategies designed to change driver behavior and the use patterns of heavy-duty vehicles.

Authority to Regulate Trucks

Two legislative acts have impacted the development of strategies to lower truck emissions, SB151 and the California Clean Air Act (CCAA) AB2595.

SB151 (Presley), passed in 1987, authorized the SCAQMD and gave it the authority to regulate the operation of heavy-

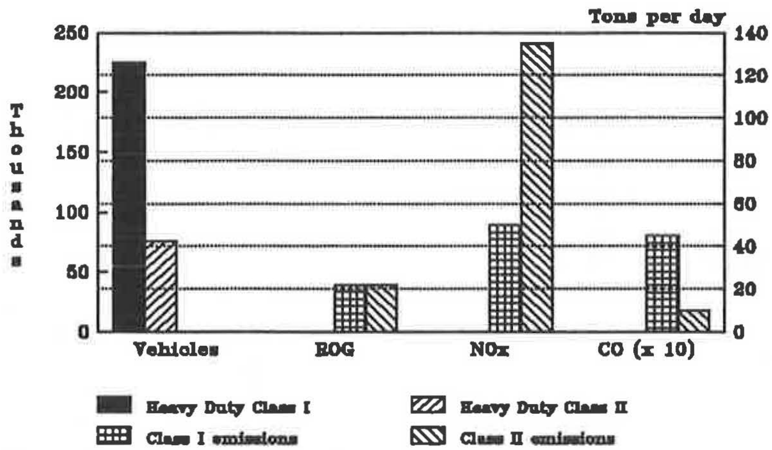


FIGURE 2 Relative contributions of heavy-duty trucks (*1*).

duty trucks during the hours of peak congestion (typically 6 to 9 a.m. and 4 to 7 p.m. in the South Coast Air Basin).

In 1988, the CCAA was passed by the California Legislature and signed by Governor Deukmejian. This comprehensive bill revamped the local agency air quality regulatory authority. The CCAA granted transportation control measure authority, similar to that of the SCAQMD, to all other local air districts in California.

The CCAA also required that a heavy-duty truck technical advisory group, with representatives from industry and regulatory agencies, be established to prepare guidelines for TCMs that would affect heavy-duty truck operations. This group started work in September 1989 and completed its guidelines in August 1990 (*6*). The guidelines, which will be discussed later, present to the districts a discussion of issues surrounding truck operation restrictions and an evaluation of different strategies. All districts other than the SCAQMD must take the AB2595 technical advisory group guidelines into consideration before adopting any TCM that would affect the operations of heavy-duty trucks.

Federal Transportation Law

The 1982 Surface Transportation Assistance Act (STAA) set standards on the size of vehicles using the Interstate highway system and preempted local restrictions on use of or access to the Interstate system by trucks. *New York State Motor Truck Association, Inc. v. City of New York* (S.D.N.Y. 1987, 654 F. Supp.) essentially struck down state laws restricting access to facilities.

In order to restrict truck access, states historically have had to prove that the facility could not safely handle truck traffic, or that the cargoes (hazardous materials, etc.) presented a danger to public health. The federal rulings would seem to limit the scope of truck traffic control measures, because any that were construed to limit access could be challenged in court. Nevertheless, the conflicts between the STAA requirements and local air pollution control regulatory authority have yet to be addressed in the courts, so it is premature to second guess what findings the court will make on these issues.

There are two additional developments that may alter this situation. When the STAA is renewed, the conflict between a cleaner environment and commerce may be resolved by Congress. The 1990 federal Clean Air Act (1990 CAA) appears to have preempted the STAA requirements for nonattainment areas classified as extreme, such that peak-period heavy-duty truck TCMs might be easily implemented. Section 182(e)(4) of the 1990 CAA, Traffic Control Measures During Heavy Traffic Hours, states that transportation control measures applicable during heavy traffic hours to high polluting vehicles or heavy-duty vehicles may be implemented ". . . notwithstanding any other provisions of law. For areas not classified as extreme, the conflict between the STAA requirements and local air pollution control regulatory authority are still to be resolved." The "extreme" nonattainment classification, according to Section 181 of the 1990 CAA, applies to areas with ozone design values of 0.280 ppm and above, with attainment of the primary standard set at 20 years.

Trucks and the Economy

Truck operations affect every part of our economy. Trucks are not only used for goods movement, but also to assist in the provision of needed services (J. Reynolds, unpublished research, City of Los Angeles Transportation Management Program, 1989). A large portion of truck traffic serves the wholesale and retail segments of the economy. Other industries with a high volume of truck vehicle miles traveled (VMT) include construction, chemical and petroleum refining, metal products, lumber, wood products, and furniture (*7*). Any transportation control measure restricting truck movement will have an effect on the economy.

In preparation for the possible implementation of transportation control measures impacting heavy-duty truck operations, the SCAQMD conducted a pilot survey of businesses using large trucks to ship and receive goods. This pilot survey was designed to gain additional baseline information for future survey efforts and to determine the most effective data-gathering methodology. In February 1990, Lockheed Information Management Services was obtained to conduct the pilot survey.

The pilot survey universe consisted of licensed businesses within the city of Los Angeles. Response to the survey was not mandatory. In order to conduct the survey, a stratified sample of businesses was randomly selected from the various business license categories in the city's file (wholesale, resale, transportation, etc.). A 4.5 percent random sample was taken. A resulting 10,125 businesses of the city's 224,000 licensed businesses were mailed survey questionnaires. The survey effort was promoted in the district newsletter and through outreach groups. A telephone support unit was also established to conduct follow-up calls to both encourage response and answer survey-related questions. A total of 1,587 were initially returned. After quality assurance review, 1,563 records (15.83 percent of the original 10,125 surveys) were retained for data analysis, representing only 0.7 percent of the 224,048 licensed businesses in Los Angeles. However, the survey sample was found to be generally representative of the survey population.

The survey questionnaire asked each business to provide general information concerning their business (business type, square footage, number of employees, etc.), and to log and report the shipments they received or sent using heavy-duty trucks over a 1-week period. Businesses were also asked to provide the number of large trucks they owned and leased and how many of these were used to ship goods. The survey instrument identified two types of trucks. This delineation was based on gross vehicle weight (GVW) and included those trucks between 8,501 and 26,000 lb GVW (Class I), and those over 26,000 lb GVW (Class II). These categories were represented on the survey instrument using pictorial examples of trucks falling within each category. Respondents were asked to record shipping activity identifying the truck type (Class I or II), truck ownership, type of delivery (full or partial loads), and time period.

Frequency distributions and simple descriptive statistics of central tendency and dispersion were computed for different variables in the survey. In addition to computing descriptive statistics for each of the survey variables, various statistical tests were conducted to examine interrelationships among the variables. Selected high-level findings from the survey are illustrative of the survey results.

A total of 22.1 percent of the businesses surveyed reported some shipping activity (i.e., sending or receiving shipments) during the 1-week survey period. Businesses reporting shipping activity were found to be larger (in terms of both square footage and number of employees), to be open longer hours, and to be more likely to have a loading dock. Businesses engaged in wholesale or retail trade, or having more than one city business license, most often reported sending or receiving shipments. Conversely, businesses categorized as professional-occupational and hotel-apartment reported shipping activity least often. Only 10.8 percent of the businesses reported sending at least one shipment per week, whereas 21.4 percent reported receiving at least one shipment. Shipments received were more often delivered by the smaller Class I trucks (55.9 percent of deliveries) than by Class II trucks (44.1 percent of deliveries), and were more often partial truck loads (71.1 percent of deliveries) than full truck loads (28.9 percent of deliveries). The percentage of shipments received by the business' own trucks was low (11.2 percent). Overall, businesses classified as wholesale or retail, as well as businesses with more than one business license, reported above-average shipping activity. Conversely, professional-occupational, hotel-apartment, and services (personal-business and health) categories reported below-average shipping activity. Both shipments sent and received varied significantly as a function of time period, with the highest shipping activity generally between 6 a.m. and 4 p.m. A substantial proportion of businesses surveyed did not own or lease trucks (as evidenced by mean values less than 1.0). However, 80 percent of the trucks owned and leased by businesses surveyed were used to ship goods.

Several factors limit conclusions that can be drawn from the survey results. Because survey respondents provided information concerning shipping activity during a 1-week period only, it is not possible to draw any firm conclusions concerning the percentage of businesses in the survey population that send or receive shipments. All that can safely be concluded is that between 20 and 24.1 percent of the businesses in the survey population sent or received, or both, one or more shipments during the 1-week survey period. In order to extrapolate to the full population, it is necessary to obtain shipping activity data for a longer time period, or for perhaps several different 1-week periods during the course of a year (e.g., to attempt to account for seasonal fluctuations). Data are further limited by the large margin of error for many of the population estimates derived from the survey results. These large margins of error were due to two factors: limited sample sizes and significant variation in the survey sample response data. In order to obtain more precise population estimates, a much larger number of businesses should be surveyed. Because of the small size of the survey sample it was necessary to group business types into larger categories. However, some of the businesses that were combined to form each new group may have been so diverse as to make the comparisons for group differences of little value. Although steps were taken to guard against this possibility, data should be collected from a larger sample of businesses to minimize the need to create such groups for data analysis. Finally, there is no other source of information comparable to that collected by the survey against which the survey results can be compared. The survey respondents were asked to provide information concerning shipping activity having full knowledge that the information provided could serve as the basis for future business regulation. Clearly, the potential for bias in the survey data exists in the absence of some method of verification. Future survey efforts should incorporate, at a minimum, an audit procedure along with a legal requirement to report accurately. Note also, only the movements of goods were examined, and not the provision of services.

CURRENT TCM PROPOSALS

The city of Los Angeles (with technical assistance from the SCAQMD) is preparing to implement peak-period truck operation restrictions, based on their staff analyses. The AB2595 Technical Advisory Group, without participation from the Los Angeles City staff, prepared separate guidelines for local air pollution control districts to consider in the development of truck TCMs.

City of Los Angeles Truck Traffic Management Program Background

In 1988, as part of a larger program to reduce congestion on the city's streets, the Los Angeles Mayor's office proposed restricting heavy-duty truck operations during the peak hours of traffic congestion. This program originally had four components: a placard system to restrict heavy-duty truck traffic during peak hours; a rapid-response team to clean up truck accidents quickly; a shipper-receiver program to have loading docks receive goods during off-peak hours; and an advisory hearing panel, to review the program and allow for exemptions on a case-by-case basis.

Several of the components have been altered or omitted as the program has progressed. The rapid response component apparently will be part of a state-wide program, established by legislation (AB480). The AB480 program brings together the California Highway Patrol (CHP), California Department of Transportation (Caltrans), and other participants to develop a more efficient way of cleaning up truck accidents and incidents. In addition, the shipper-receiver component has changed in scope and the advisory panel's role has changed.

The original shipper-receiver component required businesses to stay open at least 4 hr between 8:00 p.m. and 5:00 a.m., if five or more shipments were received during peak traffic hours. The city asked the SCAQMD to implement this portion of the program, by developing a rule limiting hours of shippers and receivers. However, rule-making by the District would require up to 18 months. The city, wanting to move ahead quickly with its program, developed a shipper-receiver component for its ordinance.

The current ordinance language defines a shipper-receiver as any facility that ships or receives commercial goods by heavy-duty truck. Shipper-receivers are limited to five deliveries within the peak traffic hours. If an establishment ships or receives more than eight shipments in the peak, then one-third of the shipments (in excess of five) must be rescheduled out of peak hours.

With respect to financing, the city initially proposed to charge a per unit fee for enrolling trucks in the program. The shipper-receiver portion was to be an SCAQMD rule, with fees set for the implementation and enforcement. In 1989, the state legislature passed AB286, which limited the city's ability to set a fee for the issuance of permits to operate trucks on city streets. One option for the city is to fund the program through the AB2766 program. AB2766, enacted by the legislature in 1990, adds $2 to $4 to vehicle registration. These funds are earmarked for transportation air quality programs. Each city receives 40% of the increase, on the basis of the number of vehicles registered in the city. If this option is pursued, then truck program registrants would not pay directly for the establishment of the program.

The current proposed program (November 15, 1990) restricts heavy-duty truck operations in one or both peak periods. The regulation provides facilities with six complying shipment-delivery options to choose from:

Option	*Percent*	*Peak*	*Conditions*
1	60	a.m.	None
2	60	p.m.	None
3	30	Both	None
4	100	Both	24-hr operation Might violate existing law (noise ordinance) All drivers report out of peak Reduce 50% of SOVs for all employees in peak
5	100	a.m.	Restricted to independent operators
6	100	p.m.	Restricted to independent operators

Independent operators have less than three vehicles.

Exemptions

There are nine general exemption classes proposed by the city of Los Angeles, including emergency vehicles (and tow trucks); military vehicles; mail trucks; trucks licensed to transport household goods while directly en route to or from a point of loading or unloading; alcoholic beverage delivery trucks; trucks transporting hazardous materials; trucks delivering wet concrete, hot asphalt, or structural steel where the workshift begins at 6:00 a.m. (unless prohibited by city ordinance or permit restriction); trucks used to restore electrical power, communications services, and pipelines; and trucks operating in conformance with existing contract conditions, permits, city ordinances, or other regulations that specifically restrict daily starting or ending times or duration of operations.

Companies may also apply for a general exemption on the basis of adverse operational or economic impacts of complying with the program. In addition to permanent exemptions, temporary single-day exemptions would be available through an application process. Route exemptions are also proposed for a number of streets within city limits, including the Los Angeles Harbor area and all identified STAA routes within the city.

AB2595 Working Group Strategies

The AB2595 working group reviewed many possible truck TCM strategies to include in its guidance to local air districts. Some proposed strategies were considered to be out of the group's purview, and they may be addressed by other legislation (accident response) or other agencies (tailpipe emission reduction strategies).

Five strategies for controlling truck emissions were analyzed by the technical advisory group:

- Education and training,
- Reduced idling,
- Freight consolidation centers,
- Shipper-receiver restrictions, and
- a.m./p.m. peak-hour restrictions.

Education and Training

The objective of this strategy is two-fold: (a) to increase heavy-truck driver awareness of the impact trucks have on air quality, encouraging practices that reduce emission; and (b) to increase awareness of truck drivers and the general public on sharing the road. By decreasing accident frequency, accident delays that result in higher emission levels can be reduced. A massive education and public awareness program was used during the 1984 Olympics.

The technical advisory group's guidelines recommend inclusion of good driving techniques in the commercial driver's license handbooks, as well as in the regular driver handbooks used by the public when obtaining or renewing licenses.

An area that is perhaps only cursorily addressed by the AB2595 technical advisory group guidelines is ongoing driver training by companies that operate heavy trucks. Untrained drivers are more likely to be involved in accidents (*8*). The United Parcel Service's ongoing and comprehensive program of driver safety training undoubtedly helped lead to their low accident rate—one-tenth of the industry average (*8*).

The benefits derived from education and training programs are indirect. However, even if a small percentage of truck accidents (perhaps 5 percent) could be eliminated, vehicle hours of delay would be reduced and a corresponding emission reduction benefit would accrue.

Education programs are costly and provide indirect benefits. Thus, it may be difficult for local jurisdictions to justify ongoing expenditures for education programs. Local air pollution control districts are encouraged to make good air-quality-related driving habits part of their overall outreach program.

Reduced Idling

Idling is defined as operating the engine without an engine load. Newer engines require 5 min or less to reach proper operating temperatures. Older engines used to require a considerable period of time to reach proper operating temperature. Thus the AB2595 guidelines recommend the adoption of a statute that restricts idling to 5 min, with some exceptions. It is felt that a uniform state law would provide consistency in enforcement and simplicity for drivers.

Extended idling often occurs as a result of driver habit. Older engines used to require extended time to warm up before operation and to cool down before shut-off. However, much of the extended idling results from facility operating practices that require trucks to inch forward and remain prepared to move at a moment's notice.

Under the proposed regulation, truck drivers as well as the operator of a facility would be liable for excess idling. By also making the facility operator responsible, facility practices that require or encourage idling (slow moving queues, etc.) will be eliminated.

The air quality benefit of this type of strategy could be calculated if truck operating practices were better understood. Excess idling is known to occur within the heavy-duty vehicle fleet; however, the magnitude of excess idling is uncertain. There are approximately 76,000 diesel trucks in the South Coast Air Basin. If the average daily idling were reduced by 10 min per vehicle per day, approximately one-quarter ton of daily hydrocarbon emissions could be saved.

Freight Consolidation Centers

Today's retail shopping center may have hundreds of stores, each restocked by truck. Many stores rely on common carriers rather than company-owned fleets. The objective of this strategy is to reduce trips by for-hire firms by establishing a single freight consolidation area where shipments can be received by the retail center. The retail stores then could pick up their goods and move them into stock rooms with alternatively fueled (e.g., electric) vehicles or with hand trucks.

This strategy works best as a development condition in new construction, although retrofitting could also be effective. The shopping center would be responsible for creating a freight consolidation center and making sure it is available to both merchants and carriers.

The air quality benefit arises through increased delivery efficiency and a reduced number of truck trips to the same address. Common carriers could consolidate orders and ship to multiple stores in a single load. However, data are not available to indicate how many shipments are currently received per store in these facilities. Hence, site-specific delivery issues are a topic for future studies.

There are costs associated with this strategy. The centers would have to not only construct the new facility, but pay personnel to receive and distribute shipments. Common carriers may also incur a slight cost from trip rescheduling. However, there should also be a savings to trucking firms from increased delivery efficiency.

Shipper-Receiver Regulations

The objective of this strategy is to reduce truck VMT during the peak hours of congestion by moving shipping and receiving operations to off-peak hours.

In most cases, the shipper-receiver controls the hours of facility operation and the times when deliveries are allowed to be shipped and received. If the shipper-receiver were open during off-peak hours, deliveries could occur during the off-peak periods, taking trucks out of the peak traffic.

As conceived in the Los Angeles program, those businesses with peak truck trips would have to be open 4 hr between the hours of 8 p.m. and 5 a.m. In many cases, however, this requirement might result in economic hardship if the facility is forced to employ a second operating shift.

There are other ways to move shipping and receiving out of the peak hours. Some retail operations give a key to the supplier, allowing the driver to deliver when the operation is closed. Others provide a safe storage facility with a key to accomplish the same thing.

Many facilities only accept shipments during the morning. If these facilities switched to an afternoon schedule, trucks would not be out in the morning peak. This is applicable to areas like the South Coast, where the morning peak results in the majority of the smog formation. However, this has a drawback in terms of increased traffic congestion during the afternoon peak.

Implementation of this type of strategy can be costly for the regulatory agency as well as industry. With more than 50,000 sources, the SCAQMD would have to provide considerable resources for implementation and enforcement.

It will be difficult to estimate the effect that these strategies will have on traffic volumes and speeds during congested periods. Thus, inherent uncertainties will not allow accurate air quality emissions reduction projections.

Peak-Hour Operating Restrictions

The objective of peak-hour operating restrictions is to reduce emissions from trucks by restricting their operation during peak hours. This type of strategy was discussed as part of the city of Los Angeles' Truck Traffic Management Program. This strategy probably would probably have to be linked to shipper-receiver rules discussed earlier.

There are numerous challenges that may prevent the successful implementation of a.m./p.m. peak-period operating restrictions. The court cases upholding federal transportation policy may limit its application. Most sites within an urban area are within 5 mi of a STAA facility, leaving few roadways covered. The effect of taking trucks off the roadways might also be negated by growth or latent demand. Enforcement may be complicated. Some industries would require at least a partial exemption, which would further complicate implementation and enforcement.

UNCERTAINTY IN THE EMISSION INVENTORY

Emissions from any type of motor vehicle depend on two sets of general parameters: vehicle activity factors (miles traveled, etc.) and activity-specific emission rates (grams of pollutant per mile traveled, etc.). In addition, the activity-specific emission rates are subject to the application of correction factors (to adjust the emission factors for operating conditions).

In general, emission inventories are developed by defining the vehicle activities and multiplying the activity estimates by the appropriate activity-specific emission rates. Unfortunately, there is a fair amount of uncertainty associated with each set of parameters used to calculate the emission inventory. Hence, uncertain activity factors are multiplied by uncertain activity-specific emission rates. Thus, site-specific economic impacts are also a topic for future studies.

Vehicle Activity

In general, vehicle activity includes such factors as number of trips made, vehicle miles traveled, and time spent at idle. Vehicle activity is often estimated through highway counts, survey techniques, and limited reporting requirements. Uncertainty exists in the estimation of heavy-duty vehicle activity through all of the methods currently used.

The uncertainty in the different methodologies can lead to activity estimates that vary widely. For example, the Caltrans cost allocation study (FY 1986–1987 estimates) prepared by Sydec Inc. estimated the annual truck VMT to be approximately 13.9 billion miles. The Caltrans Truck Miles Traveled report estimated approximately 12 billion truck miles were traveled for the same period (*9*). The EMFAC7D-BURDEN7A model, used by the California Air Resources Board to estimate motor vehicle emissions, indicated that 15.5 billion heavy-duty truck miles were traveled in 1986–1987 (the average of 1986 and 1987 figures). It is clear that a large range of uncertainty is likely to exist for any vehicle activity parameter estimated.

Emission Rates

Emission rates, the masses of emissions per unit of activity, are established for each specific activity. For example, specific vehicle emission rates are determined for such activities as engine starts (cold or hot start emissions), engine cool-down (hot soak emissions—gasoline engines only), vehicle miles traveled (running emissions), diurnal evaporation (gasoline engines only), and running evaporative losses. However, because of limited laboratory capabilities, heavy-duty truck emission factors do not yet exist for hot or cold start and hot soak activities.

Vehicle emission rates are determined through laboratory testing, using the methods and procedures established by the U.S. Environmental Protection Agency (EPA) and the California Air Resources Board. Uncertainty exists from the outset, with a range of precision and accuracy associated with the tests.

The set of individual vehicle emission factors are composited into vehicle class emission factors, using vehicle registration data available through the state. In essence, emission rates are developed for an average vehicle in the vehicle class, again introducing some uncertainty.

For heavy-duty trucks, the emission factors were derived by EPA from a limited testing of 40 heavy-duty engines in 1980, and updated in 1984 to account for new vehicle emission standards (Plattey, personal communication 1989). The engines were tested on engine dynamometers (as opposed to chassis dynamometers used for automobiles), so the applicability of the engine-derived emission factors to vehicles in motion are somewhat uncertain. With the variation in heavy-duty vehicle sizes and weights that are in use today, the uncertainty in the heavy-duty vehicle emission inventory warrants further investigation. The California Air Resources Board, in conjunction with the Southern California Regional Transit District, will begin heavy-duty truck chassis dynamometer testing in mid-1991. The new research should provide additional insight into the accuracy of the emission rates used in current models.

Current Correction Factors

In order to prepare the emission inventory, the fleet emission rates are adjusted, through the use of laboratory-determined correction factors, to account for specific operating conditions, such as the operating environment or trip factors.

The operating environment of the vehicle can impact the emission rates and operating efficiency of the engines. Environmental factors such as temperature and altitude are taken into consideration as emission rate corrections. Additional adjustments are also made for assumed effectiveness of inspection and maintenance programs.

The speed of the vehicle trip is an important factor in determining the emissions from motor vehicles. The emissions of hydrocarbons (HC) and carbon monoxide (CO) from motor vehicles (both gasoline and diesel) have been demonstrated to decrease as vehicle speed increases, while the emissions of oxides of nitrogen (NO_x) have been demonstrated to increase as vehicle speed increases (*10*). Recent research by the California Air Resources Board indicates that although automobile NO_x emissions decrease with speed, the emission rates increase rapidly above about 50 mph. In fact, the increase in emissions rates above 50 mph is so significant that improved enforcement of speed limits is currently under investigation as an emission control strategy.

The methodology used to establish the speed correction factors is cycle correction. The fleet emission factor is modified by the results obtained when vehicles are run through test cycles with different average cycle speeds. This process raises the question as to whether standardized cycles, that differ in number of stops and starts, and cruise periods, can be used to establish valid correction factors for vehicle speeds.

Results from field testing during the South Coast Air Quality Study have indicated that a previous version of the CARB's EMFAC model (EMFAC7C) may seriously have underestimated in-use motor vehicle emissions of hydrocarbons, by as much as a factor of 1.4 to 6.9 (*11*). The current version of the EMFAC model (EMFAC7E) now contains additional emission factors associated with automobile running evaporative losses and may better represent emissions from the vehicle fleet. However, additional corrections are probably still required for gasoline trucks.

Future testing by the California Air Resources Board will begin in early 1991 to determine how second-by-second motor vehicle emissions vary during the cycle tests, and additional field data may indicate how well the EMFAC model performs.

Future Correction Factors

Frequent stop-and-go motion, characteristic of peak-hour operation, is also thought to increase emissions from vehicles. This may be especially true of heavy-duty trucks because of their heavy operating loads and power requirements for acceleration. Although the emission contribution has not yet been quantified, the contribution of acceleration-based emission will likely be examined by the California Air Resources Board in the future.

Currently, the emission inventories do not include contributions from engine idling. To some extent, these emissions are accounted for through the use of average vehicle speeds under the cycle correction factors. However, additional research on vehicle idling is warranted.

It is clear that reducing congestion and increasing vehicle operating speeds (below 50 mph), serves as means to reduce vehicle emissions. However, latent demand (the increase in demand that is often noted to occur when additional capacity becomes available on a road segment) may negate congestion reduction benefits. The latent demand phenomena is an issue that deserves further study and analysis.

Driver habits can affect the magnitude of the emission rates for specific activities. For example, the behavior of drivers at stoplights may be important because vehicle emission rates increase significantly when the engine is revved or the vehicle accelerates rapidly.

Uncertainty Conclusions

It is clear that numerous assumptions and generalizations must be made for a bulk emission inventory to be generated. Such is the nature of uncertainty in the calculation of any emission inventory. Public agencies attempt to ensure that the best available data are used, and that research projects designed to improve the methodologies are undertaken.

Overall, there is a large amount of inherent uncertainty in the methodologies used to estimate emission inventories for heavy-duty vehicles. Both the activity factors and emission rates for heavy-duty vehicles can, and will, be improved by the California Air Resources Board through concentrated research efforts. The California Air Resources Board has already identified research needs and proposed a number of projects designed to improve the heavy-duty truck emission inventory during the next few years. However, in developing heavy-duty truck TCMs, it should be recognized that significant uncertainty appears to exist in the estimation of emission reduction benefits.

CONCLUSIONS

Regulations imposed on stationary sources have already exhausted the major emission reductions that can be obtained from most industrial categories. New stationary source regulations are focusing on smaller and smaller portions of the pie. Thus, the regulation of transportation activity is clearly recognized by federal, state, and local agencies as critical for attaining the national ambient air quality standards.

Truck operations are a significant contributor to the mobile source emission inventory, even though truck traffic constitutes a relatively small portion of the total traffic volume. Truck traffic control measures are available and can have a positive impact on air quality. For example, idling restrictions can have a positive impact with little or no economic impact (in nonattainment states, idling restrictions should be explored as a first step).

Some transportation control measures affecting heavy-duty truck operations are certain to be implemented. However, emission reduction effects from truck traffic control measures are difficult to estimate, given the current state of modeling. The impact that truck traffic control measures will have on air quality and the economics of goods distribution (direct and indirect costs) require further study. Future research should also address the impact of latent demand on the emission reductions achieved and the potential indirect impacts of education programs.

The ability to make informed policy decisions concerning the implementation of traffic control measures requires that detailed traffic, motor carrier, and business surveys be conducted. High survey confidence levels and detailed results are necessary. Because a high number of survey responses is usually required, consideration might be given to conducting mandatory surveys of motor carriers and businesses potentially affected by proposed traffic control measures. Minimal one-time fees could be imposed to cover the administrative costs of surveys. Resistance from motor carriers or business communities to mandatory surveys should be outweighed by the need for accurate information, especially before implementation of traffic control measures with potentially significant economic impacts.

REFERENCES

1. *Emission Inventory 1987.* Technical Support Division, California Air Resources Board, Sacramento, March 1990.
2. *Path to Clean Air: A Summary of the 1989 Air Quality Management Plan.* South Coast Air Quality Management District, May 1989.
3. G. Giuliano. *Olympics Transportation System Management Performance Analysis.* Preliminary Report (UCI-ITS-85-3). Institute of Transportation Studies, University of California, Irvine, March 1985.
4. A. Davidson and J. Cassmassi. Ozone Reductions During Olympic Period due to Congestion Reducing Measures. *APCA Journal*, Vol. 35, No. 3, March 1985.

5. Cambridge Systematics, Inc. *Urban Freeway Gridlock Study*. California Department of Transportation, Sacramento, Nov. 1988.
6. *Guidelines for Districts Considering the Implementation of Transportation Control Measures Affecting Heavy-Duty Truck Operations*. AB2595 Technical Advisory Group, California Air Resources Board, Sacramento, July 1990.
7. *Truck Movement Study*. Southern California Association of Governments, Los Angeles, Calif., June 1988.
8. R. Beilock, R. B. Capelle, Jr., and E. B. Page. Speed and Training Factors Associated with Heavy Truck Accidents. *Transportation Quarterly*, Vol. XLIII, No. 4, Oct. 1989.
9. *Truck Miles of Travel on the California State Highway System*, 1974–1987. Division of Transportation Planning, California Department of Transportation, Sacramento, Nov. 1988.
10. *Methodology to Calculate Emission Factors for On-Road Motor Vehicles*. Technical Support Division; California Air Resources Board, Sacramento, Nov. 1986.
11. M. N. Ingalls, L. R. Smith, and R. E. Kirksey. *Measurement of On-Road Vehicle Emission Factors in the California South Coast Air Basin, Volume I, Regulated Emissions*. Southwest Research Institute, San Antonio, Tex., June 1989.

Publication of this paper sponsored by Committee on Transportation and Air Quality.

Sensitivity Analysis for Land Use, Transportation, and Air Quality

JEFF MAY AND GEORGE SCHEUERNSTUHL

The effect on air quality of a higher-density alternative land use scenario that concentrates a high percentage of the employment growth expected between 1989 and 2010 along transit corridors, was compared with that for the expected lower-density suburban development typical in the Denver region. The lower-density suburban growth development pattern was served by a circumferential highway, whereas in the higher-density scenario the circumferential was deleted with the exception of those segments already constructed or reasonably committed. Both scenarios included a rapid transit system. A high-occupancy vehicle lane system was also substituted for other freeway improvements proposed as part of the regional transportation plan. Although transit patronage increased significantly in the higher-density scenario, vehicle miles of travel remained relatively unchanged. Carbon monoxide levels increased under the higher-density land use case, but remained well below the federal standard. Other pollutant levels did not vary significantly between the two scenarios. Concentrating much of the 1989 to 2010 regional employment growth along transit corridors did not improve air quality.

In order to provide insight as to the degree to which changes in land development patterns could affect air quality, a sensitivity analysis of the air quality effects of concentrating development growth along major rapid transit corridors, minimizing new highway construction, and adding a substantial bus and high-occuancy vehicle (HOV) network was conducted for the Denver region. The methodology used in conducting this analysis and the resulting effects on carbon monoxide (CO), O_3, NO_2, and particulate matter less than 10 μm in diameter (PM10) pollutants are described.

A higher-density development pattern alternative was defined. Although the regional control totals on population and employment growth were maintained, more of the new growth was concentrated in the rapid transit corridors. The rapid transit corridors defined in the 2010 regional transportation plan (RTP) are shown in Figure 1. It was estimated that market forces would lead to approximately 40 percent of new employment and 20 percent of new population locating within 1 mi of a rapid transit alignment or $1\frac{1}{2}$ mi of a rapid transit station. For the air quality sensitivity test, it was assumed that growth in the transit corridors would double to include approximately 80 percent of new employment and 40 percent of new population. This higher-density alternative scenario was matched with a transportation system comprising a much reduced highway network, the rapid transit system adopted as part of the 2010 RTP, and a supplemental HOV lane network on major freeways. The major facilities assumed in the two highway systems are shown in Figure 2. The same transportation systems management actions were modeled for the alternative case as was modeled for the 2010 RTP case. Higher parking costs under the alternative scenario in the Denver central business district (CBD) reflected higher employment densities expected. Suburban parking costs were generally not assumed to be in place. No changes in other parameters such as automobile operating costs or trip generation were made. The regional travel model, a traditional four-step Urban Transportation Planning System (UTPS) based model, was run and the resultant highway assignment was used in projecting ambient air quality levels associated with the land use test case for the higher-density sensitivity. Pollutants tested included CO, O_3, NO_2, and PM10.

Denver Regional Council of Governments, 2480 West 26th Avenue, Suite 200B, Denver, Colo. 80211.

DEVELOPMENT SCENARIOS

Lower-Density Suburban Development Pattern

Two alternative development patterns were defined. The base case was the currently adopted market driver development scenario. This land use pattern is typified by

- The majority of new growth occurring outside the City and County of Denver.
- Growth in the City and County of Denver occurring in a few areas such as the CBD, the Platte Valley, and the new airport area, and the Stapleton site.
- Suburban residential development occurring at low densities (typically single-family dwellings at three and four dwelling units per acre).
- Approximately 40 percent of the employment growth and 20 percent of the population growth occurring within the transit corridors.

Higher-Density Transit Corridor Development Pattern

The alternative-scenario land use pattern located approximately 80 percent of the employment growth from 1985 to 2010 within the seven major rapid transit corridors. Population growth within the transit corridors was assumed to about double from 20 percent of the growth to 40 percent. This required that 195,000 future residents and 226,600 future employees be shifted from other locations into the transit corridors.

The relatively short (20-year) time horizon used in this study did not allow for major redevelopment to occur. As such, it was assumed that the existing urban fabric would remain in place and the study dealt only with the growth increment.

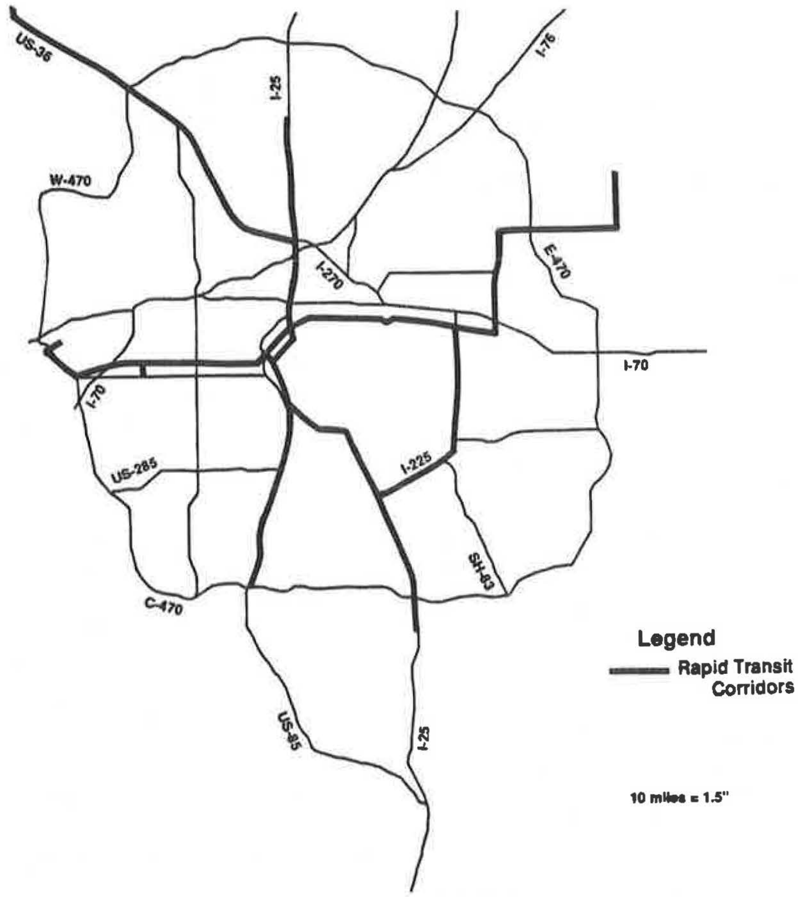


FIGURE 1 Rapid transit corridors.

Not moving existing residences or jobs to the transit corridors had a substantial effect on how much impact the alternative scenario would have, in that 69 percent of the population anticipated in 2010 is already in place, as are 59 percent of the employment locations. Therefore, the percent of 2010 total development actually transferred to the transit corridor zones was only 7 percent of the population and 13 percent of the employment. The majority of development was assumed to be in the same location for both scenarios.

The assumption to place 80 percent of the new employment and only 40 percent of the new residential development in the transit corridors had the effect of concentrating employment much more strongly than residential population in transit corridors. This distribution was based on the assumption that there would be a greater possibility of concentrating employment in higher-density nodes and strips than of concentrating residential development.

Comparison of Development Scenarios

Table 1 presents the net effect on land use distribution by analysis areas between the low density 2010 RTP scenario and the higher density alternative scenario. Development within the rapid transit corridors increased by 20.5 and 22.0 percent for population and employment, respectively. The Denver CBD population was increased by 31.1 percent and employment by 23.3 percent. CBD employment is critical in that it is a major determinant of transit patronage. The percent of population and employment living outside of transit corridors declined by 12 and 32.7 percent, respectively.

In general, future development growth was removed from the 470 beltway corridor and from areas beyond the beltway and in outlying communities such as Longmont and Castle Rock and moved to the transit corridors. Figure 3 shows the concentration of additional growth in the central area. The zonal data sets reveal that the population changes were somewhat more compact, whereas some employment increases were assumed at the end of the major corridors. The population reductions clearly occurred in a ring around the region. The employment reductions were more evenly dispersed and less focused on the ring.

After the redistribution process, a check was made of population and employment densities. The highest population densities do not exceed 50 persons per acre or about 18 housing units per acre. This density is approximately that of one of Denver's oldest and most transit-dependent area: Capitol Hill. Employment densities used were similar to those occurring in the suburban, campus-style, Denver Technological Center. An examination of current land use patterns found many areas along the rapid transit corridors that could be developed at higher densities. Significant development opportunities exist as the planned rapid transit corridors avoid currently developed areas in favor of alignments along existing

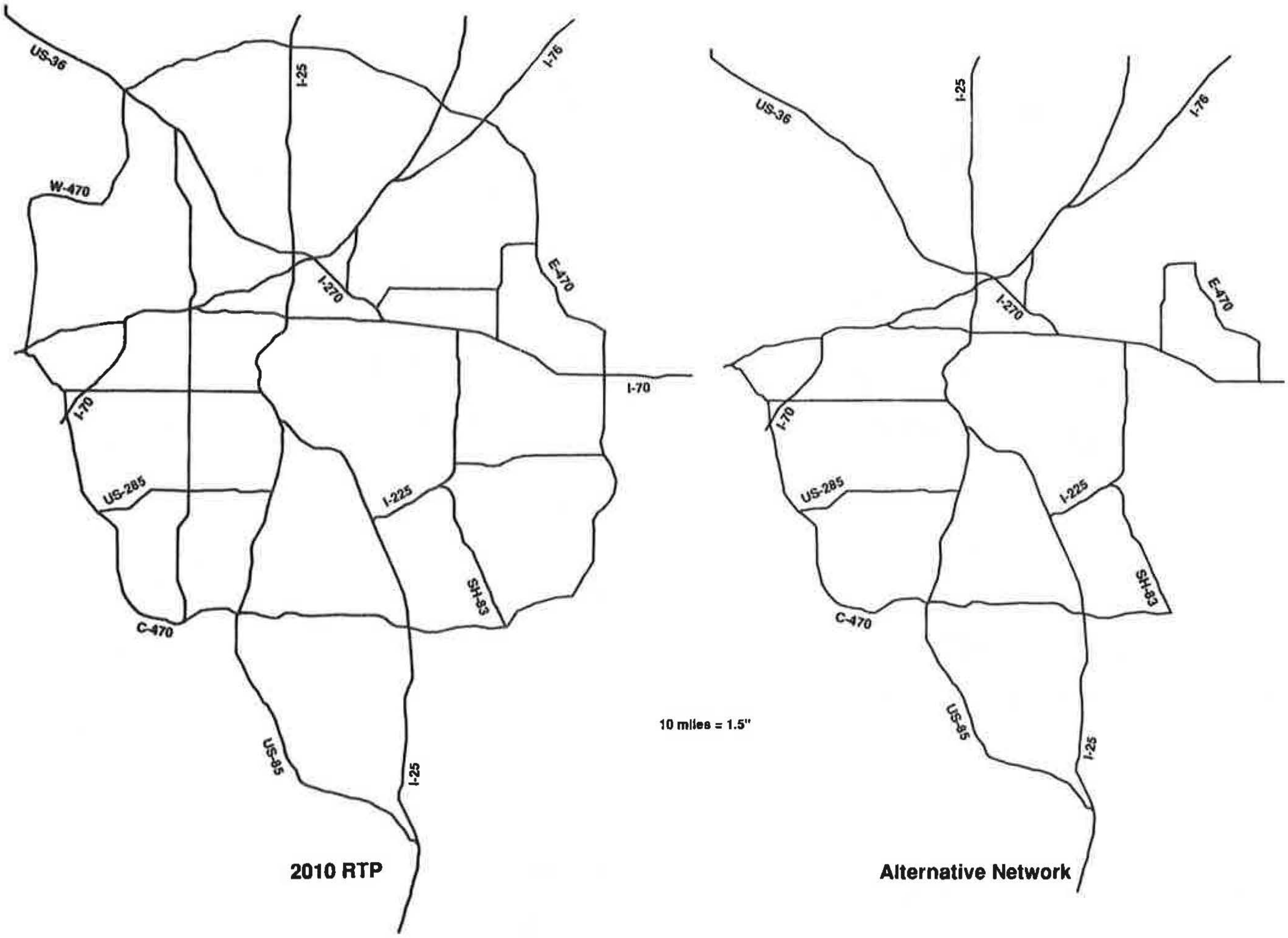


FIGURE 2 Alternative highway systems.

TABLE 1 DEVELOPMENT BY ANALYSIS AREAS OF POPULATION AND EMPLOYMENT

Geographic Area	2010 Plan Pop.	2010 Plan Empl.	2010 LU/AQ Alternative Pop.	2010 LU/AQ Alternative Empl.	% Change Pop.	% Change Empl.
Rapid Transit Corridors	950,890	1,028,007	1,145,741	1,254,617	20.5	22.0
Denver CBD	14,971	164,650	19,628	202,975	31.1	23.3
Remainder of Transit Corridors	935,919	863,357	1,126,113	1,051,642	20.3	21.8
Outside of Rapid Transit Corridors	1,629,010	692,693	1,434,166	466,305	-12.0	-32.7
TOTAL	2,579,900	1,720,700	2,579,907	1,720,922	0.0	0.0

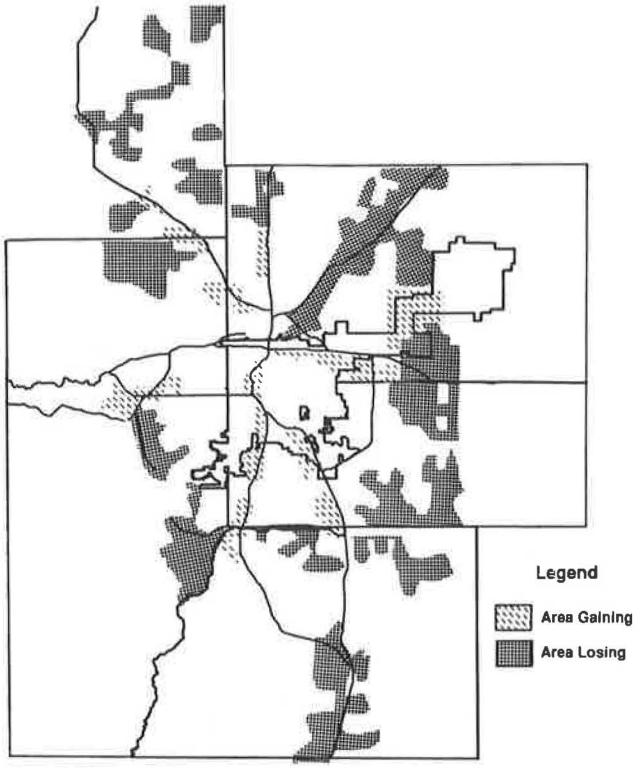


FIGURE 3 Distribution shift from 2010 RDF to alternative scenario.

railroad tracks or freeway facilities where tracts of open land exist.

EFFECT ON TRANSPORTATION DEMAND

Transportation System Supply Characteristics

In order to develop travel demand estimates, it is necessary to make a number of basic assumptions concerning the transportation system supply characteristics. These supply characteristics include the extent and layout of the rapid transit system as well as the background bus system, the extent and layout of the highway system, and similar information concerning the HOV lane system.

The highway system used in testing the higher-density alternative scenario is based on the former year 2000 RTP thoroughfare network and modified on the basis of current knowledge or recent construction, in revised alignments of some roadways and inclusion of new roads to serve the new Denver International Airport. In addition, a minimal number of lane improvements were assumed for the inner freeway system.

The transportation networks assumed for each scenario are shown in Figure 2. The major differences between the networks are (a) the 2010 RTP includes a full circumferential freeway whereas the alternative includes only those portions currently open, under construction, or well along in the planning process; (b) the 2010 RTP includes improvements to the existing freeway network to increase the number of lanes to 8 or 10—in contrast, the alternative network includes minimal improvements to the existing freeway network, and (c) further widening of existing arterials and provision of new arterials in the 2010 RTP are planned—in contrast, widening and new roadways are minimized under the alternative.

The transit system was based on the currently adopted year 2010 RTP transit element. It consisted of approximately 100 mi of light-rail transit (LRT) system supplemented by a large background bus network. Additional stations were identified on the planned rapid transit system to maximize transit development potential. The light-rail system was assumed to have a minimum of 3-min headways and the acceleration, deceleration, and maximum speeds of LRT vehicles.

Transit fare assumptions mimic current-day average transit fares by service type. Rapid transit lines were assumed to have a fare similar to existing express bus fare (i.e., $1.25). Automobile operating costs were those in place during the early 1980s. As such, no additional gaseoline taxes or other vehicle taxes were assumed to be in place under the alternative scenario as compared with the base case. It was also assumed that future increases in gasoline prices will be offset by improvements in fuel efficiency.

Parking costs in the Denver CBD were estimated in direct relation to employment density: the higher the employment density, the higher the parking cost. The average all-day parking costs in the Denver CBD were calculated as 1985, $4.11; 2010 RTP, $6.60; and 2010 alternative, $7.24. The higher CBD employment densities projected in the alternative scenario resulted in a 10 percent increase in projected parking costs over the 2010 RTP scenario. Parking costs were not assumed in suburban areas (except for the new airport) as the employment density was too low to justify parking costs under either scenario.

Transportation Modeling Process

The regional travel model set is a standard four-step model set using the programs and procedures recommended within the UMTA-FHWA-developed UTPS. The model set, originally developed on the basis of a 1971 home interview survey, has been updated extensively using the 1980 Census Urban Transportation Planning Package (UTPP) and information from the 1985 small-sample household survey and 1986 onboard bus survey.

Trip Generation

The total number of internal person-trips in the modeling area increased by roughly ½ percent from 8,453,300 to 8,491,600 between the 2010 plan case and the alternative scenario. This increase occurred as the result of moving development from the small outlying growth centers and rural areas (outside the modeling area) into the Denver urbanized area. Table 2 presents assumed population and employment within the modeling area, as well as the resultant work and total internal person-trips. The increase in trip making occurs mainly in the home-based work trip purpose that is most susceptible to diversion to transit or carpooling. Trip making per person per day decreased slightly (0.7 percent) from 3.83 to 3.80.

The total number of internal person-trips was relatively insensitive to changes in the development pattern. The trip

TABLE 2 REGIONAL TRIP GENERATION

	2010 RTP LU	LU/AQ Alternative	Percent Change
Population	2,206,000	2,232,000	1.2%
Employment	1,527,000	1,575,000	3.1%
Home-Based Work Person Trips	2,227,000	2,300,000	3.3%
Total Internal Person Trips*	8,453,300	8,491,600	0.5%

*Excludes truck trips and internal/external trips.

generation model used assumes that the total number of trips generated per family can be explained by the income and household size of the family. Some individuals have speculated that automobile trips can be converted to walk trips at high enough densities. This relationship is not included in the trip generation model because of the following:

- The quantitative relationship between density and trip making is only speculative; no quantitative relationships have been established.
- Increased density in and of itself may not decrease trip making. In addition to increased density, a diversity of land uses on a microscale would appear necessary. Current land use patterns are centralizing and separating land uses rather than integrating them. For example, in the 1950s the corner grocery store disappeared making way for the supermarket. Recently, some supermarkets are disappearing, being replaced by a few mega-grocery stores. The same trends can be seen in hardware stores and other land uses. The impact on VMT of this change in marketing is unknown. The megastores are often grouped together, providing opportunities to shop for a diversity of items (groceries, drugs, books, hardware) in one general location. This may make trip chaining more likely, offsetting some of the additional distance driven. In order to convert the entire shopping trip for automobile back to walking, it would be necessary to return to the corner grocery store. This implies a change in lifestyle, in which individuals would need to be willing to give up the diversity of goods and low prices associated with mass marketing. It also implies extremely high densities at the residential end to have sufficient numbers of families within easy walking distance of stores.

Trip Distribution

A set of gravity models was used to distribute trips between origins and destinations. Person-trips are distributed for both peak and off-peak periods on the basis of composite impedance measures of spatial separation of zones. The composite impedance is a function of automobile and transit travel times and costs.

The average trip length of 8.5 mi did not change between the two land use scenarios. However, this lack of change on a daily basis masks changes by time of day. The rearrangement of development and congestion in the peak periods caused trip lengths to decrease by 4.1 percent during the a.m. peak period (from 9.4 to 9.0 mi) and 4.8 percent during the p.m. peak period (from 8.4 to 7.9 mi). These decreases indicate that gains were made in reducing the length of the work trip. They were offset by an increase of 3.6 percent (from 8.3 to 8.6 mi) during the off-peak period. The increase during the off-peak may be reflective of a greater separation of activity areas and residents under the alternative scenario than the 2010 RDF/RTP scenario. For example, the ratio of population to employment in the non-rapid-transit corridors increases substantially from 2.3 to 3.0. This increase is not offset by a large corresponding decrease within the transit corridors (the ratio changes only from 0.94 to 0.92). Longer trips may be necessary for those remaining in the areas outside the rapid transit corridors.

The travel model used bases its estimates of travel destinations partially on employment locations. The trip attraction model uses estimates of retail and total employment. In moving employment into rapid transit corridors, the model assumed that service jobs were also being concentrated. Differentiating service jobs and redistributing them separately might have reduced nonwork travel.

Tables 3 and 4 present the trip distribution between analysis areas. Work trips destined to the Denver CBD and the remainder of the transit corridors increase by 29.6 and 20.4 percent, respectively. Work trips destined outside the rapid transit corridors decreased 32.1 percent. Nonwork trips to transit corridors not in the CBD increased 21.3 percent, whereas trips to areas outside rapid transit corridors decreased 22.5 percent. However, the increased work trips are only 3.2 percent for the entire modeling area.

Mode Split

The mode split model was used to divide the trips into mode-specific trips (transit, drive alone, and shared ride—2 and 3 plus carpools) based on the alternative transportation network description. The mode split model is a three-dimensional logit model (drive alone, shared ride, and transit). The mode split model takes into account in- and out-of-vehicle travel time, and monetary costs such as parking fees, transit fares, and automobile operating costs.

Out-of-vehicle travel time is weighted by 2.5. The implied value of travel time between home and work is $5.60 per hour. The mode split model is run for both peak and off-peak periods.

Changing land use assumptions from the 2010 Regional Development Framework (RDF) to the alternative scenario increased the potential for walk access to transit. Households with walk access to either bus or rail transit increased from 58.3 to 63.6 percent, or a gain of 61,900 households having walk access. Employment locations with access increased from 66 to 71.3 percent or a gain of 112,300 employees having

TABLE 3 TRIP DISTRIBUTION (THOUSANDS)

Home-Based Work

Trip Origin	Trip Destination: Denver CBD		Rapid Transit Corridors Outside CBD		Outside Rapid Transit Corridors		Total Modeling Area	
	2010 Plan	LU/AQ	2010 Plan	LU/AQ	2010 Plan	LU/AQ	2010 Plan	Alternative
Denver CBD	3	4	5	7	2	2	10	13
Rapid Transit Corridors Outside CBD	105	154	543	743	231	194	879	1,092
Outside Rapid Transit Corridor	133	153	668	714	540	329	1,340	1,196
Total Modeling Area	240	311	1,216	1,464	773	535	2,229	2,300

Non Work

Trip Origin	Trip Destination: Denver CBD		Rapid Transit Corridors Outside CBD		Outside Rapid Transit Corridors		Total Modeling Area	
	2010 Plan	LU/AQ	2010 Plan	LU/AQ	2010 Plan	LU/AQ	2010 Plan	Alternative
Denver CBD	45	52	86	100	50	43	181	195
Rapid Transit Corridors Outside CBD	141	150	1,718	2,239	661	602	2,520	2,991
Outside Rapid Transit Corridor	130	118	976	1,033	1,947	1,416	3,053	2,567
Total Modeling Area	316	320	2,779	3,372	1,658	2,060	5,753	5,752

TABLE 4 PERCENT CHANGE IN TRIP DISTRIBUTION FOR ALTERNATIVE COMPARED WITH 2010 PLAN

Home-Base Work:

Trip Origin	Trip Destination: Denver CBD	Rapid Transit Corridors Outside CBD	Outside Rapid Transit Corridors	Total Modeling Area
Denver CBD	36.5%	44.0%	-4.8%	31.3%
Rapid Transit Corridors Outside CBD	47.5%	36.8%	-4.8%	31.3%
Outside Rapid Transit Corridor	15.3%	6.9%	-39.0%	-10.8%
Total Modeling Area	29.6%	20.4%	-32.1%	3.2%

Non-Work:

Trip Origin	Trip Destination: Denver CBD	Rapid Transit Corridors Outside CBD	Outside Rapid Transit Corridors	Total Modeling Area
Denver CBD	15.9%	16.1%	-13.3%	8.0%
Rapid Transit Corridors Outside CBD	6.5%	30.4%	-9.0%	18.7%
Outside Rapid Transit Corridor	-9.3%	5.8%	-27.3%	-15.9%
Total Modeling Area	1.3%	21.3%	-22.5%	0.0%

access. In-vehicle travel times were calculated to test how long it would take for area residents to go to work or other destinations using transit. The reported travel times do not include time to walk to a line, wait for a bus, or transfer between transit lines. This test indicated that in the alternative scenario, 34 percent of the work trips and 35 percent of all trips could be completed in under 40 minutes by walking to a bus or rail transit line.

Transit patronage increases 26.9 percent from 316,000 to 402,000 daily between the 2010 RTP case and the 2010 alternative scenario. A significant proportion of the increase is additional work trips taking place on transit, which increases from 149,000 to 216,400, or 44.8 percent. Overall mode share increases from 4.0 to 4.8 percent. Work trip mode share increases from 6.7 to 9.4 percent. The patronage estimates for the 2010 RTP and the 2010 alternative both exhibit significant increases from the current-day transit ridership of approximately 135,000 patrons per day, or roughly 2.6 percent of internal trip making. Ridership on the rapid transit system has increased from 110,300 under the 2010 RTP case to 185,000 under the 2010 land use–air quality (LU–AQ) scenario. Table 5 indicates that most of the increase in transit usage occurred in trips to non-CBD destinations in rapid transit corridors, implying that the rapid transit system is better serving non-CBD trips in the alternative scenario.

Annual transit trips per capita were compared with data from other cities to confirm the reasonableness of the forecasts (see Table 6).

In order to estimate the effects of the HOV system on automobile occupancy, the work mode choice module was supplemented with an HOV choice model. Surveys in other cities having HOV lanes supported this approach, because little use of HOV lanes by nonworkers actually takes place. In addition, nonwork automobile occupancy is most highly affected by family size, because these are mostly families traveling together. The work mode choice model incorporates drive-alone, HOV, and transit travel times and costs into its probability function, and divides work trips into drive alone, shared ride 2, shared ride 3-plus, and transit modes. Time savings using an HOV lane as compared to a general purpose lane are factored into the decision-making process. After incorporating the effects of the HOV system, it was found that the HOV system did not have a significant effect on regional work automobile occupancy; however, the HOV system did increase the number of persons in 3-plus person work-purpose carpools by 11.5 percent over the 2010 RTP, from 152,200 to 169,700 persons per day.

Network Assignment and Analysis

The federally released computer programs were used for building and processing both the highway and transit networks at the traffic zone level. On the highway side, diversion off freeway facilities onto arterials is accomplished using congestion diversion curves to calculate travel times under congested circumstances. Congested speeds from the peak highway assignments are compared to initial estimates of the speeds used in trip distribution and mode split calculations. This typically results in two to three model reruns before highway-side equilibration is reached. Similarly, a check was made to ensure adequate transit capacity for projected demand.

Vehicle miles of travel (VMT) between the two scenarios is effectively unchanged despite the slight increase in population and employment in the alternative scenario. Both scenarios generated approximately 63.5 million VMT daily. VMT in the alternative was 0.2 percent higher than in the 2010 RTP scenario. This can be compared with the increase of population of 1.3 percent and increase in employment of 3.1 percent in the modeling area in the alternative. Table 7 presents the percent of the system operating over capacity. The percent of VMT experiencing congestion increases from 32 to 45.1 percent. The average speeds decrease by 9 to 10 percent during the peak periods. This result increases daily vehicle hours of travel (VHT) by approximately 8 percent.

TABLE 6 ANNUAL TRANSIT TRIPS PER CAPITA

Houston (1985)	21
Denver (1985)	37
Miami (1985)	39
Portland (1985)	49
Seattle (1985)	49
Atlanta (1985)	65
San Francisco (CMSA)(1985)	84
Chicago (1985)	105
San Francisco (Central)(1985)	232
New York (1985)	309
Denver (2010 RTP)	65
Denver (2010 LU/AQ)	83

EFFECT ON AIR QUALITY

Carbon Monoxide

The levels of carbon monoxide associated with each scenario are indicated below:

2010 RTP (ppm)	*Alternative Scenario (ppm)*	*Standard (ppm)*
4.8	6.5	9

TABLE 7 HIGHWAY OPERATING CHARACTERISTICS

Percent of System Operating over Capacity[1]			
% of Roadway Mileage	14.6%	17.2%	
% of VMT	32.0%	45.1%	
AM Peak Avg. Speed	23.8 mph	21.4 mph	-10.1%
PM Peak Avg. Speed	22.3 mph	20.3 mph	- 9.0%
All Day Avg. Speed	26.7 mph	24.7 mph	- 7.3%
Daily VHT	2,382,000	2,575,500	+ 8.1%

(1) Level of Service E and E during the peak period.

TABLE 5 TRANSIT PATRONAGE BY ANALYSIS AREA

Area	2010 RTP	2010 LU/AQ Alternative	Change
Rapid Transit Corridors			
Denver CBD	130,300	164,400	+ 26%
Rapid Transit Corridors Outside CBD	75,600	151,900	+109%
Outside Rapid Transit Corridor	110,100	85,700	- 22%
TOTAL	316,000	402,000	+ 27%

The Denver CBD remains the site with the highest concentrations. The alternative LU–AQ scenario results in significantly higher levels of CO because there is more congestion in the Denver CBD under this scenario than with the 2010 RTP scenario. However, neither scenario violates the federal standard.

Ozone

The currently available O_3 model does not allow prediction of specific levels of O_3; however, it does predict the hydrocarbon (HC) emissions that must be reduced to achieve the standard, or the amount of HC emissions that can be tolerated until a violation is expected in a specific corridor between the CBD and an outlying area. These corridors are called trajectories.

Three trajectories were modeled in this exercise; from the CBD to Boulder, from the CBD to the new airport, and from the CBD to Highlands Ranch. Below is a comparison of the likelihood of a violation, if any, in each trajectory for the 2010 RTP and the alternative LU–AQ scenario. In both cases where a violation is possible, the expected violation would be marginal.

	Possibility of O_3 Violation CBD to:		
Scenario	*Boulder*	*Airport*	*Highlands*
2010 RTP	No	Yes	No
Alternate Scenario	No	No	Yes

PM10

Model results are often best used to compare one scenario with another scenario instead of predicting absolute values. The PM10 (small particulate) model calibration indicates that the model significantly overpredicts the annual average concentration by approximately 50 percent. The model calibration also indicates that the 24-hr concentrations are about 10 percent higher than actual 1986 monitored concentrations. The results should thus be viewed as indications of a problem needing further exploration rather than as an identification of specific predicted results.

The PM10 (small particulate) modeling completed for this exercise resulted in no significant difference in concentrations between the alternatives. The alternative scenario resulted in a 2 percent higher maximum concentration than the 2010 RTP; however, such a small difference is less than the range of accuracy of the model. The modeled 24-hr and average annual concentrations are presented below. Both the 2010 RTP and the alternative scenario result in significant violations of the 24-hr standard; however, the area that experiences violations is slightly larger in the 2010 RTP than in the alternative scenario.

	2010 RTP (mg/m^3)	*Alternative Scenario (mg/m^3)*	*Standard (mg/m^3)*
24-hr average	272	278	150
Annual average	90	91	50

Nitrogen Dioxide

Using a conservative rollback modeling technique, the projected NO_2 concentrations for the two transportation data sets at the CBD monitoring station were modeled. The annual average results are summarized below:

2010 RTP (ppm)	*Alternate Scenario (ppm)*	*Standard (ppm)*
0.058	0.059	0.053

The rollback technique is not sensitive to differences in spatial distribution; therefore, the effects of the spatial difference between the two transportation data sets are not adequately analyzed. Also, because the two data sets produce concentrations that are only slightly higher than the standard, a more detailed modeling technique would be required to determine if either of these two scenarios would cause an actual exceeding of the standard.

CONCLUSIONS

- Changing development patterns by concentrating employment growth along transit corridors to the degree used in this study, restructuring highway capacity improvements, and providing an extensive bus and HOV system did not improve air quality.
- The relatively short (20-year) time horizon used in this study did not allow for major redevelopment to occur. The existing urban fabric was assumed to remain in place and the study dealt only with the growth increment. Not moving existing residences and jobs minimized the air quality difference measured between the two scenarios. Because 69 percent of residential locations and 59 percent of the job locations in 2010 are already in place, the relative impact of even large changes in the location of projected growth is minimized.
- Even though 40 percent of the expected population growth was concentrated in the transit corridors under the alternative scenario, the larger concentration of 80 percent of employment growth in the transit corridors led to an imbalance of population-to-employment ratios. It appears that both population and employment need to be concentrated in the transit corridors. In the alternative higher-density development pattern tested, because substantial employment growth was allocated to the transit corridors without equivalent redistributed population, a greater spatial separation between population and employment resulted than in the base 2010 RDF/RTP planned suburbanization case. This spatial separation led to increased miles of travel (i.e., longer trip lengths) for non-peak-period, nonwork trips, which eliminated much of the benefit gained in transferring peak-period, drive-alone work trips to transit and carpooling.
- Transit patronage increased significantly—from 316,000 to 402,000 riders per day between the 2010 RDF/RTP and 2010 LU–AQ alternative. Overall, 4.8 percent of trips in the LU–AQ alternative use transit. Factors that mitigated against a larger transit mode share included

  —Increased congestion resulting from the assumed constrained capacity roadway network and higher land use densities in the LU–AQ alternative was of limited benefit in increasing transit patronage. In both scenarios, the bulk of transit service is provided by buses operating on-street. These buses operating in mixed traffic flow were slowed proportionately because of increased arterial traffic congestion. Bus services affected by traffic congestion included on-street feeder services to the rail facility. Only for the rail line-haul portion of the transit trip was traffic congestion a benefit in inducing travelers to use transit.

—Placing development within 1 mi of a rapid transit line did not ensure good transit access. The resulting densities were too low to ensure walk access to the rapid transit system. Most of the development in the corridors was still beyond walking distance to a station. Many of the residents and employees within the rapid transit corridor were dependent on feeder buses or park-and-ride lots to access the rail system. Moving more employment to the immediate vicinity of a transit station (i.e., within $\frac{1}{3}$ mi, the assumed maximum walking distance) or provision of small-area circulator buses feeding the transit station might increase transit ridership.

—The rapid transit and HOV systems were radial in nature, mainly serving trips to and from the central area.

—Provision of transit improvements is a necessary but not sufficient condition to cause a large increase in transit ridership. Incentives to use transit, such as employer subsidized bus passes and bus priority lanes intersecting a rapid transit corridor, and disincentives to use the private automobile, such as increased parking costs and higher fuel costs, appear necessary if a large shift to alternative modes is to occur.

ACKNOWLEDGMENT

Preparation of this report has been financed in part through grants from UMTA and FHWA of the U.S. Department of Transportation.

Publication of this paper sponsored by Task Force on Transportation Planning Applications.

Special Events and Carbon Monoxide Violations: TSM, Crowd Control, Economics, and Solutions to Adverse Air Quality Impacts

SUSANNE PELLY SPITZER

Issues involved with the staging of short-term special events, including regulatory concerns, transportation systems management (TSM), crowd control, carbon monoxide (CO) incidents, and economics, are examined. Five of these events in the Twin Cities Metropolitan Area of Minneapolis and St. Paul are then investigated. Although TSM efforts can improve traffic flow during special events, the unplanned nature of some of these events, lack of integrated interjurisdictional coordination, and perceived financial benefits lead to CO incidents. Impacts of these incidents may be felt areawide and have economic and policy costs not usually considered. Therefore, special efforts need to be made to control the impacts of these events on air quality levels.

Special events require special transportation system management (TSM) actions to control their impact on the transportation system. Although literature exists on incident and special event control for improved TSM actions, readily available information does not examine the link between special events and carbon monoxide (CO) incidents. [For purposes of this paper, an incident is an occurrence in which the Minnesota standard of 9.0 parts per million (ppm) or the U.S. standard of 9 ppm is exceeded.] Three types of special events are characterized: short-term repeating, anticipated unique events, and unplanned events and their impacts in the Twin Cities metropolitan area. Each presents different problems for TSM and for CO control.

THREE TYPES OF SPECIAL EVENTS

Short-Term Repeating Special Events

Short-term repeating special events, such as festivals, fairs, and athletic meets are considered beneficial to the economy of the area in which they are held, because they bring new visitors. They also bring unwanted side effects. These include additional costs for traffic and crowd control, in the form of special signage, added police, and overtime duty for traffic engineering and other public works personnel. A less apparent impact they bring is that they can result in adverse air quality impacts.

Air Quality Division, Minnesota Pollution Control Agency, 520 Lafayette Road North, St. Paul, Minn. 55155.

Anticipated Unique Events

These considerations also enter into other types of special events, such as narrowing of capacity because of temporary construction, official visits of foreign dignitaries, and one-time sports events. Although these types of events can considerably disrupt the transportation system, they are easier to plan for, and their impacts tend to be more controllable.

Unplanned Events

The third category of special events is ones whose impacts are not necessarily anticipated, nor the impacts fully known, because they occur without consent of the government of the affected area. Examples of this type of special events are demonstrations and warm-weather cruising by automobile drivers. Although authorities do have published strategies to deal with anticipated crowds of varying sizes, minimizing the TSM and CO problems of the situation is possible only if these events repeat in the same location under similar circumstances.

PREVIOUS INFORMATION ON SPECIAL EVENTS AND TSM

The traffic management of incidents and special events has been a subject of considerable research interest in the last 10 to 20 years, with an emphasis on management of freeway emergencies (*1*). As noted by Dudek (*2*), special events, even with the best of planning, may still result in traffic congestion because drivers are unaware of the extent of congestion along particular roadways leading to the events, or the events themselves are one-of-a-kind. One type of special event, which included a system of off-site ticket sales with guaranteed remote lot parking spots, bus transportation to the event, and admission all in one ticket, alleviated congestion (*3*). Recent work on special events indicates that interagency or intercity cooperation may be the key to managing the events, an approach that requires surrendering some independence for the sake of areawide TSM (*4,5*).

SPECIAL EVENTS IN TWIN CITIES AREA

Short-Term Repeating Special Events

Short-term repeating special events have become common in the Twin Cities area. These events only last one or a few days. Typical yearly street festivals in 1985 to 1990 and their respective attendance levels include the Cinco de Mayo festival (8,000) across the river from downtown, the Grand Olde Days festival (300,000 on 1 day) along Grand Avenue in the Crocus Hill and Macalester-Groveland neighborhoods, and the Taste of Minnesota (35,000 average, 100,000 on July 4) by the Capitol near downtown, all in St. Paul. In Minneapolis, the Uptown Art Festival is held each year at the high-volume Hennepin Avenue and Lake Street intersection, with a total attendance of 250,000. All of these events close off major streets to traffic. The Taste of Minnesota also leads to interruptions in access onto Interstate 94 freeway entrance ramps.

Annual short-term repeating special events that last a week or longer but have varying locations for events and parades include the Winter Carnival (15,000 average, 100,000 for parades) in St. Paul in February, and the Aquatennial Festival in July in Minneapolis. Most of the traffic for these events originates in their respective cities. The State Fair (207,000 on 1 day), which is not located officially in any city or county, but is on a special tract of land that borders the city of St. Paul on the south and east, the Village of Falcon Heights on the north and east, and the University of Minnesota on the west, is a yearly 12-day event. Its traffic is dispersed to St. Paul and surrounding suburbs. There are also short-term repeating special events that occur every few years rather than annually. These events include ones that last a few days, such as the meeting of the National Street Rod Association at the State Fairgrounds. This event also attracted mobile source activity near the intersection of Snelling and University Avenues. As at the Hennepin-Lake intersection, a monitor there has periodically recorded CO violations.

Anticipated Unique Events

Temporary events in recent months include street closures because of sewer construction, the visit of Mikhail Gorbachev to the Twin Cities, and the Olympic Festival. Sewer construction has lasted for several years. Temporary freeway construction lasting more than a year is coordinated with the State of Minnesota Pollution Control Agency (MPCA) to ensure that CO standards are not violated. Gorbachev's visit will probably remain unique to Twin Cities history. Security considerations ensured that every aspect of the visit included TSM planning. Except for the 1992 Superbowl, events such as the Olympic Festival (which does not draw major crowds to any one event) will not likely occur in the near future.

Unplanned Events

Spontaneous demonstrations include marches and protests on political issues, such as U.S. involvement in Central America, and environmental ones, such as air pollution, both at the Hennepin-Lake intersection. Others, mostly political, are held in downtown Minneapolis. Celebrations also fall into this category, such as the one held adjacent to the Metrodome on the edge of downtown Minneapolis when the Minnesota Twins won the World Series. Warm-weather cruising near the Hennepin-Lake intersection occurs in the spring and early summer, when cars cruise from the intersection to Lakes Calhoun and Harriet and Lake of the Isles, some on regular city streets and some on parkways.

REGULATORY HISTORY

The Clean Air Act (CAA) Amendments of 1977 required that areas that would not meet CO standards by December 31, 1982, adopt an inspection and maintenance (I/M) program for automobile emission control systems. They would then be required to meet the standards by December 31, 1987. By mid-1990, however, the U.S. Environmental Protection Agency (EPA) listed over 29 areas that still did not meet CO standards and would have to have strengthened I/M programs.

Failure to Meet 1982 Deadline

Some 1982 areas failed to meet their deadlines because they relied on the Federal Motor Vehicle Emissions Control Program (FMVECP), which tightened emission standards. Implementation of the FMVECP was delayed a year by the EPA. Other 1982 deadline areas failed to meet the deadline because they did not implement all TSM measures required by the CAA [called "transportation control measures" (TCMs)]. Some 1982 areas that did not meet the deadlines relied on modeling that assumed unrealistically high ambient air temperatures for vehicle operation, on the basis of EPA procedures. (These procedures test emissions of cars at higher temperatures than they are actually driven in cold-weather states under worst-case CO conditions.)

Areas Subject to Sanctions

Areas that failed to submit an acceptable state implementation plan (SIP) to meet the 1982 or 1987 deadlines or failed to make good faith efforts to implement the SIP were subject to potential and actual sanctions. These sanctions were left somewhat to the discretion of the EPA under CAA provisions. Sanctions included bans on stationary source construction; halts of federally funded transportation projects except for those of mass transit, safety, or TCMs; cutoffs of sewer construction funds; and cutoffs of program grants to air quality agencies. Lesser-known sanctions included refusal to allow new water or sewer hookups and refusal to give consent to bonds needed for construction of inner city housing. The EPA and FHWA have often disagreed in the past as to whether the transportation sanctions are enforceable without FHWA consent.

CO NONATTAINMENT IN THE TWIN CITIES AREA

The Twin Cities metropolitan area surrounding the cities of Minneapolis and St. Paul has a population exceeding 2,240,850.

Modeling for the SIP predicted that the FMVECP and TCMs (one-way pairs, traffic signal timing changes, and improved transit service) would enable the area to meet the 1982 deadline. Except for the Snelling-University intersection in the St. Paul Midway area, located between the two downtowns, all monitors indicated the expected modeled compliance by the deadline.

NONATTAINMENT STATUS OF THE INTERSECTION

The MPCA with the Metropolitan Council (the metropolitan planning organization), submitted a new SIP amendment that included intersection signal timing changes for the intersection and surrounding ones. The changes would become operative when CO reached a certain level. Attached was a request to redesignate the entire seven-county Twin Cities metropolitan area to attainment for CO standards, except for the Snelling-University intersection, because in previous years the whole area was designated as the CO nonattainment area to watch in the Metropolitan Council's jurisdiction.

The EPA modified the request. It approved instead a cross-shaped area in St. Paul considerably larger than the intersection. In response, the MPCA, in cooperation with the Minnesota Department of Transportation (Mn/DOT), posted additional CO monitors in the cross-shaped area, and at the Hennepin-Lake intersection in Minneapolis to demonstrate that the EPA was in error. Instead of compliance, a number of the monitors (including the Hennepin-Lake one) indicated violations of the CO standard. At the EPA's urging and under the threat of possible EPA sewer and air program sanctions, the MPCA, Mn/DOT, and Metropolitan Council won legislative approval of an I/M program. It will begin in the summer of 1991.

INVESTIGATION OF CO INCIDENTS

In Spring 1990, five CO incidents related to special events were investigated to determine whether such incidents could be prevented in the future through improved TSM. Controlling special events at the source appeared to be a more equitable and less costly way to control CO than tightening the whole future I/M program. If so, these types of events would not tighten the I/M program's cutpoint levels to achieve a reduction in emissions. Governmental officials, peace officers in charge of traffic management and those heading police reserve activities, and traffic engineering staff were interviewed to obtain a complete picture of causes and possible preventions of the incidents.

The events, occurring in 1985 and 1989, included two incidents affecting the Snelling-University intersection monitor south of the State Fairgrounds in St. Paul, one affecting the Snelling-Larpenteur intersection monitor in Falcon Heights north of the Fairgrounds, and one at the Hennepin-Lake intersection. All other CO incidents in 1985 to 1990 (a total of 77 in the Twin Cities area), did not appear to be related to pin-pointable special events, but to general traffic, preholiday shopping, or traffic and inversions.

STREET RODS AND CRUISING

Alteration of Pollution Equipment

Two CO incidents were clearly related to a short-term repeating special event, the meeting of the National Street Rod Association in July of 1985, and again in 1989, at the state fairgrounds. Each meeting drew over 100,000 participants and spectators. Members drove cars whose bodies were designed to look like older classic models. The vehicles had newer engines that were altered to provide added power or other features designed to increase particular performance aspects. Pollution control equipment was sometimes removed in this process. The exact amount of excessive CO produced is not known, but removal of a catalytic converter can increase CO production by 400 percent. The street rods were tuned to extremely high performance standards, perhaps reducing that percentage.

Mobile Source Activity by Street Rod Owners and Traffic Control

Street rods participating in the event were based over the entire Twin Cities area, but gathered for events by cruising north on Snelling Avenue to the fairgrounds, and west and east on University Avenue. Traffic control problems were compounded by owners of the street rods on weekday mornings, who cruised north to the fairgrounds in advance of the gate opening. Those that arrived early at the gates assured themselves of tree-shaded positions on the grounds. The traffic resulting from maneuvering for position backed southwards to the Snelling-University intersection.

Spectators and Traffic Control

The cruising attracted spectators, who cruised in regular cars next to the street rods. Sidewalk-sitters parked near the Snelling-University intersection, walked to it, lawn chairs in tow, sitting as close to the curb as possible to see the maximum number of street rods. Spectators also drove to the fairgrounds to view the street rods after paying an admission price of $8.00. The free viewing at the intersection was therefore more appealing for many.

Efforts of the St. Paul Police, Police Reserves, and Neighborhood Assistance Officers to control traffic by preplanning (through use of public events control guideplans) and spontaneous problem solving (banning left-turns without official prepermission to do so) were somewhat thwarted by spectators. Some tried to pour bleach under the tires of street rods. The bleach caused tires to warm up rapidly, producing clouds of white smoke and extremely rapid acceleration. This led to some loss of control in steering, and danger to persons on the sidewalk. Failure to respond to the requests to do "bleach burns" resulted in threats to "key" a car—ruining the finish by scratching it with a key. In 1989, the cruising street rods and spectators were finally dispersed by Public Works department trucks, which were mobilized to water the streets. The water caused spots, temporarily marring the highly polished wax finish of the street rods. The National Street

Rod Association has declined to return to the Twin Cities in the immediate future, probably because of the Public Works department actions.

STATE FAIR OPERATIONS

The Minnesota State Fair is the fourth largest in the nation, with the largest paid attendance of all state fairs, recording over 153,000 on a weekday in 1989, and the 207,000 noted before on a weekend day. It is owned and operated by the Minnesota State Agricultural Society, a quasi-governmental organization governed by elected representations from county fairs. It passes its own ordinances, and contracts for police patrols. With the exception of the traffic it generates, it is an extremely popular institution with almost all Minnesotans, and with visitors from surrounding states.

Traffic Control at the Fair

During the 12 days of the Fair, traffic is directed from the fairgrounds to the outside by a team of off-duty peace officers (licensed police) from communities around the state, headed by an off-duty Ramsey County sheriff's officer also hired by the Fair. The rest of the year, the St. Paul Police Department patrols the grounds and tries to handle traffic from events, such as the street rods or visitors to the annual July 4th fireworks show. Falcon Heights has in the past contracted with the Ramsey County sheriff for traffic control during the 12 days. For 1990 and future years, its contract is with Roseville, a city whose southern border is several blocks north of the fairgrounds. Roseville is also in Ramsey County.

Most of the traffic leaving the fairgrounds heads to the south, to St. Paul. It is headed directly at the Snelling-University intersection. This is partly because of the general solution used by the Ramsey County and Roseville police forces of sending the traffic elsewhere. It is also caused by the large crowds that find Fair ingress easiest from the Snelling Avenue exit of Interstate 94.

Parking Management at the Fair

Parking space on the fairgrounds is inadequate, with room for at most 15,000 cars. Spaces in lots probably turn over at least once a day, but parking has been free and lots often fill by 10:30 a.m. on weekends. One-third of all fairgoers are from rural areas located many hours from the Fair. They feel uncomfortable trying to find a parking space elsewhere. Like other fairgoers faced with filled on-site lots, they prefer to cruise the fairgrounds and the surrounding neighborhood. Many make use of a St. Paul policy that permits paid parking on lawns of fair neighborhood residences during the 12 days. The off-street parking, which includes officially tolerated curb jumping to access the lawns, slows traffic and may contribute to pedestrian safety problems.

The Fair has tried a number of approaches to manage the parking, from agreements with the University of Minnesota, to leasing of remote lots, to rideshare arrangements. Agreements with the University of Minnesota involve use of the parking lots adjacent to the fairgrounds that are part of the St. Paul portion of the Twin Cities campus. (The University is similar to the Fair in that with the exception of some legislative funding approvals, it is largely independent of any jurisdiction except that of the state of Minnesota.) Spaces in its lots, normally in great demand, are available because the University is not in session at the time of the Fair.

Remote lots are leased by the Fair in 14 locations, serving 10,000 vehicles. Parking in the lots is free to the public, but there was a charge for bus service to the fairgrounds. In 1991, bus fare from the lots will become free. Although buses travel between the lots and the fairgrounds on a regular basis, arrival and departure times are not definite because the buses themselves become enmeshed in Fair traffic. The Fair is attempting to find additional remote lots. An attempt by the Fair to display a portable changeable message sign directing persons to remote lots was withdrawn after St. Paul received a complaint about the noise the generator for it made from a nearby homeowner. The sign was not effectively diverting much traffic. Additional changeable signs remain in place. More remote lots are needed, especially south of the Snelling-University intersection. Although parking remains free in the shuttle lots, a new plan in 1991 will charge vehicles with less than three persons for parking on the fairgrounds. The revenues from parking will pay for free bus rides from the remote lots, and discounts on admissions for remote lot parkers.

Ridesharing at the Fair

The Fair has contacted Minnesota Rideshare, part of the Metropolitan Transit Commission (MTC), and the Regional Transit Board, which governs the MTC, to improve ridesharing among 2,200 permanent, but mostly seasonal, and 3,800 temporary employees during the 12 days. Providing incentives for carpooling by giving close-in spots to employees who rideshare, and working with the MTC to improve already increased bus service to the Fair are strategies it is actively pursuing. The seasonal characteristics and variety of shifts have resulted in few matches.

TRAFFIC BY THE LAKES

Determining the Impact of the Aquatennial

The CO incident at the Hennepin-Lake monitor in Minneapolis was apparently caused by an event of the Aquatennial, the Minneapolis festival centered around its lakes. After talking to the Minneapolis Police Reserves and to Minneapolis Traffic Engineering, it was apparent that the event did not cause the incident. The event was held downtown, and not near the intersection. It was unlikely that the parade sent traffic far south enough to affect the intersection.

Lake Cruising

Further discussions revealed that the day in question was a perfect July day with cooler-than-usual temperatures in the 70s Fahrenheit. Much of the lake traffic was caused by cruising, an unplanned special event that occurs frequently near the lakes. The traffic on the lake parkways was managed by the Minneapolis Park Police. (Except at upper levels of command, they are governed by the Minneapolis Park Board, a separate public entity.) Parks surrounding the lakes were closed in late evening, leading to cruising of the lake parkways. If the Minneapolis Park Police then restricted parkway traffic to stop cruising, the cars moved west to the Hennepin-Lake intersection, where they became a problem for the Minneapolis Police. Given the two-way average daily traffic of 32,000 through the intersection, additional traffic from the lakes tended to increase CO levels there, especially at night. Most CO incidents occurred during late-night hours at this intersection. A one-way pair was implemented for air quality reasons through this intersection in 1990. No violations have been recorded since.

SEWER CONSTRUCTION AND CO VIOLATIONS

One special event became a CO incident by surprise. In late October 1989, many monitors around Minnesota recorded CO incidents as record-breaking warm air combined with a statewide inversion to trap CO. Even background non-hot-spot monitors recorded incidents caused by traffic and the unusual meteorology. None of the incidents could be explained in any other way, with one exception.

The one explainable CO incident occurred in downtown St. Paul. The monitor, located on a one-way street, rarely recorded CO incidents. After investigation, the MPCA staff concluded that the incident was caused by rerouting of traffic from an Interstate freeway exit to avoid temporary construction near the exit. More than 16,000 additional vehicles, many driven by persons unfamiliar with the area, at slower than normal speeds, were routed past the monitor by Mn/DOT. The construction was for mandated separation of storm overflow and regular sewers, required by the EPA.

METEOROLOGY AND REPEATING AND UNPLANNED SPECIAL EVENTS

All CO incidents for monitors in the Twin Cities, St. Cloud, and Rochester, Minnesota, were examined for those years in which they were in operation from 1972, the year monitoring started, through 1990, the last year of verified data. The aim of the examination was to determine what the possible role of meteorology was in creating patterns of incidents, compared to that played by the special events. The sheer volume of incidents (over 1,500) makes it difficult to analyze the meteorology in detail. Instead, the patterns of occurrence were examined.

In the examination of the dates of occurrence of the incidents, a recurring special event that especially affected monitors was Christmas shopping. Although five incidents did occur on the day after Thanksgiving, the busiest shopping day of the year, this is not a high number. Monitors were likely to indicate violations in the period 2 weeks before Christmas (18 in the period December 10 to 17), and in the period right before Christmas (22 in the period December 18 to 24). Hot-spot monitors usually indicated several violations in the latter period if they had any in the former.

Traditional explanations based on meteorology would note that although Thanksgiving can sometimes have mild weather, December is almost always very cold in Minnesota. The atmosphere is also very stable. Table 1 (National Climatic Data Center, U.S. National Oceanic and Atmospheric Administration) indicates that the December mean temperature for 1959 to 1988 is 19.3°F, the January temperature 12.9°F. The atmosphere is probably not more stable. Although there are occasional inversions (often statewide in nature) that are correlated with CO incidents in January, the same type of pattern does not emerge for any particular weeks. More detailed examination of temperatures and wind direction and speed would be needed to come to a definitive conclusion about the effect of Christmas shopping or any other special recurring event during the period of November through January.

All CO incidents were also examined to determine their correlation with cold weather and atmospheric stability, given the popular assumption that incidents are related to winter weather. Although most do occur in colder temperatures, many of the 1,500 incidents occur in temperatures above 60°F. As indicated in Tables 2 and 3 (Air Quality Division, Minnesota Pollution Control Agency), in the years for which summary temperature data are available, 1987 to 1990, there were 10 out of 53 CO incidents for which the temperature was above 60°F. Of these 10, 8 occurred at the Lake-Hennepin intersection monitor, lending credence to the possibility that the unplanned special event of cruising contributed more than meteorology.

RECOMMENDATIONS: COORDINATION FOR SPECIAL EVENTS

Handling Repeating Special Events

The previously mentioned CO incident to the north of the Fair at the Snelling-Larpenteur intersection monitor in Falcon Heights occurred the day the Fair attendance recorded a record 207,000 paid attendees. The cruising from the street rods led to violations at the Snelling-University monitor in 1985 and 1989. (A monitor malfunction made it impossible to determine if a similar incident occurred in 1982.) Thus, repeating special events need special attention to avoid having them cause CO incidents.

A combination of police reserve officers, traffic control agents, regular police, or off-duty officers can be mobilized to handle large special events such as parades or festivals. This procedure is now being implemented in some areas. In order to avoid conflicts on union issues and to ensure that integrated work is done by all forces, these personnel need to be explicitly under one command that can choose which ones to use and how to mobilize them. This policy is probably not feasible for smaller unplanned events, because city budgets do not allow routine traffic direction by police, given more pressing needs of increased crimes and more violent crimes, which are of greater importance to the victims.

Sending the Traffic Elsewhere

Whether an event is repeating or not, long-term or 1 day, and planned or unplanned, traffic control techniques can be

TABLE 1 AVERAGE TEMPERATURES (DEGREES F) FOR MINNEAPOLIS-ST. PAUL, MINNESOTA

YEAR	JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUGUST	SEPT	OCT	NOV	DEC	ANNUAL
1959	10.6	17.2	34.0	47.1	60.6	70.6	74.5	75.2	63.1	44.4	25.4	30.1	46.1
1960	17.5	17.8	19.8	45.9	57.0	63.8	71.7	72.1	61.8	48.7	33.3	16.8	43.8
1961	12.0	22.5	32.0	38.5	54.7	68.1	70.8	71.3	59.1	52.2	33.5	15.3	44.2
1962	7.1	11.7	24.5	42.2	60.6	66.2	67.5	68.3	56.4	50.2	35.0	19.0	42.4
1963	2.9	2.1	34.2	47.3	55.4	69.8	73.5	68.9	62.2	58.1	38.3	10.0	44.4
1964	20.0	23.9	25.8	46.8	61.5	68.7	76.0	68.5	58.9	48.2	35.0	14.8	45.7
1965	10.0	11.8	19.5	41.8	58.7	66.5	70.5	68.6	52.8	50.7	33.1	28.0	42.7
1966	3.3	16.3	35.8	42.2	53.6	68.4	76.8	68.2	60.3	47.5	30.1	18.1	43.4
1967	14.6	8.7	29.8	44.7	52.3	66.9	68.8	66.2	60.3	46.3	30.7	21.8	42.6
1968	14.3	15.2	38.8	48.5	53.4	67.2	71.1	70.7	61.1	50.7	34.0	16.9	45.2
1969	9.4	19.3	24.1	49.3	60.6	61.8	73.6	74.4	63.0	46.5	33.6	20.3	44.7
1970	5.6	15.4	26.0	46.1	58.5	71.2	75.2	71.9	61.2	49.6	32.7	18.2	44.3
1971	6.5	17.0	28.0	47.0	55.4	71.5	68.8	69.6	62.8	51.4	32.7	18.4	44.1
1972	5.5	10.5	26.5	41.9	61.3	66.0	68.5	69.8	57.9	43.7	32.2	11.3	41.3
1973	17.4	21.6	40.2	44.4	55.2	69.5	73.8	73.4	60.1	53.8	34.3	16.7	46.7
1974	11.9	16.9	29.5	47.1	54.4	65.5	76.6	67.3	55.3	49.8	33.7	24.4	44.4
1975	14.5	15.5	22.1	38.9	60.9	68.8	76.3	71.7	57.7	52.8	37.5	21.3	44.8
1976	11.6	27.8	31.4	51.8	58.9	71.7	76.1	73.3	61.8	44.6	28.3	13.6	45.9
1977	0.3	22.7	37.5	53.0	66.9	68.4	74.8	66.1	60.5	47.1	30.8	14.4	45.2
1978	5.5	11.6	30.0	45.2	61.8	67.8	71.1	72.2	67.3	49.8	32.5	15.2	44.2
1979	3.2	10.0	28.9	44.0	55.5	67.3	73.6	69.9	63.4	46.6	31.7	26.0	43.3
1980	15.3	15.3	27.3	49.2	61.5	67.6	75.2	70.7	59.5	45.1	36.6	19.8	45.2
1981	18.0	23.4	37.7	49.1	57.1	67.0	70.9	69.3	60.0	46.7	38.0	17.5	46.2
1982	2.3	15.8	29.0	43.8	62.5	63.7	75.6	71.8	60.9	50.3	31.5	25.7	44.4
1983	19.6	26.9	34.2	42.3	54.6	68.0	77.2	76.8	62.6	48.4	34.0	3.7	45.7
1984	12.0	27.5	24.8	47.1	56.0	69.7	72.2	73.5	57.2	50.7	33.3	17.9	45.2
1985	10.1	16.5	35.6	52.1	62.2	63.9	73.9	67.6	59.9	47.5	24.8	7.7	43.5
1986	17.5	15.7	33.9	49.6	59.4	68.6	73.9	67.1	59.8	49.2	28.2	24.7	45.8
1987	21.2	37.6	38.7	53.5	63.5	72.8	76.0	69.0	62.5	44.6	37.9	25.0	49.7
1988	10.4	13.9	33.8	47.4	65.4	74.4	78.1	73.9	62.4	44.0	32.7	20.5	46.4
RECORD													
MEAN	12.9	17.1	30.0	46.0	58.2	67.9	73.2	70.7	61.5	49.7	32.9	19.3	45.0
MAX	21.5	25.7	38.3	55.8	68.5	77.8	83.3	80.6	71.3	59.0	40.5	26.8	54.1
MIN	4.3	8.5	21.7	36.3	47.9	58.0	63.1	60.7	51.6	40.3	25.3	11.7	35.6

improved to deal with it provided political will and funding are available. Sending the traffic elsewhere can work provided it is evenly dispersed so as not to cause or aggravate an existing hotspot condition. (This technique, for example, was used by the Minneapolis police reserves to send cruising traffic celebrating the Twins victory to sites away from downtown. They forced celebrants involuntarily in four directions onto routes leading directly to Interstate highways.) It is not a good planning tool for repeating special events, which need integrated advance planning to handle unexpectedly large crowds, such as occurred with the record attendance day at the Fair in 1989.

Restricting Certain Events or Actions

Banning certain events or actions is bound to be politically unpopular. Nevertheless, if the events are virtually uncontrollable, or lead to traffic control actions, such as lawn parking, which have possible safety or other adverse traffic management impacts, the temptation is strong to regulate the events. Many of these events (such as the State Fair) are immensely popular with the public. As indicated below, the financial impact of doing so may also outweigh the desire for control of CO incidents. Therefore, it is necessary to carefully investigate other TSM methods, coordination of forces, and methods of neighborhood cooperation to control the CO implications, because it is probably not desirable or possible to eliminate the events. For some events, such as the street rods, increased CO levels may not be controllable except by prohibiting the event.

ECONOMICS OF EVENTS

Most special events are categorized by promoters as contributing heavily to the local economy. The Olympic Festival, for example, was expected to contribute $23 million to the local economy. Its cost, however, is harder to determine. Expenditures by state and regional agencies and the University of Minnesota to speed up construction of new facilities and provide traffic management for the Olympic Festival are expected to exceed that figure. In the case of the National Street Rod Association meet, 90 percent occupancy for Twin City hotels and motels was not uncommon, so its economic impact was great. The Fair, while receiving no legislative appropriations,

TABLE 2 CO EXCEEDANCES FOR 1987 AND 1988

1987 CO EXCEEDANCES

DATE	LEVEL	TEMPERATURE	LOCATION
01-12-87	12.8 PPM	38 degrees F.	Duluth
02-01-87	11.6 PPM	37 degrees F.	University & Lexington
02-06-87	16.0 PPM	43 degrees F.	University & Lexington
02-06-87	10.2 PPM	43 degrees F.	Larpenteur & Snelling
02-06-87	13.7 PPM	43 degrees F.	University & Snelling
02-07-87	10.6 PPM	43 degrees F.	University & Rice
02-07-87	9.7 PPM	43 degrees F.	Nalpak Bldg.
02-09-87	10.1 PPM	41 degrees F.	University & Lexington
02-09-87	9.8 PPM	43 degrees F.	Larpenteur & Snelling
02-09-87	9.3 PPM	41 degrees F.	University & Snelling
02-19-87	9.9 PPM	38 degrees F.	University & Lexington
10-13-87	9.2 PPM	49 degrees F.	University & Lexington
11-02-87	9.9 PPM	61 degrees F.	Lake & Hennepin
11-13-87	12.0 PPM	41 degrees F.	Lake & Hennepin
11-13-87	11.4 PPM	43 degrees F.	St. Cloud
11-13-87	9.0 PPM	48 degrees F.	University & Lexington
11-14-87	9.6 PPM	38 degrees F.	University & Lexington
11-30-87	9.2 PPM	31 degrees F.	University & Snelling
12-19-87	9.0 PPM	____	Rochester
12-23-87	11.6 PPM	____	Rochester

1988 CO EXCEEDANCES

DATE	LEVEL	TEMPERATURE	LOCATION
01-11-88	14.3 PPM	14 degrees F.	University & Lexington
01-11-88	14.2 PPM	14 degrees F.	Lake & Hennepin
01-15-88	9.4 PPM	17 degrees F.	Lake & Hennepin
01-15-88	9.1 PPM	29 degrees F.	Lake & Hennepin
01-17-88	9.8 PPM	20 degrees F.	Lake & Hennepin
01-18-88	9.7 PPM	29 degrees F.	Lake & Hennepin
01-28-88	9.1 PPM	13 degrees F.	University & Lexington
01-28-88	9.4 PPM	8 degrees F.	Lake & Hennepin
05-20-88	10.7 PPM	78 degrees F.	Lake & Hennepin
06-05-88	9.2 PPM	78 degrees F.	Lake & Hennepin
08-15-88	9.6 PPM	94 degrees F.	Lake & Hennepin
09-17-88	9.3 PPM	81 degrees F.	Lake & Hennepin
10-14-88	9.7 PPM	58 degrees F.	Lake & Hennepin
11-03-88	9.9 PPM	48 degrees F.	Lake & Hennepin
12-19-88	9.6 PPM	34 degrees F.	Lake & Hennepin
12-30-88	9.1 PPM	14 degrees F.	Lake & Hennepin

*The Levels reflect the highest 8 hour average exceedance for that date.

charged admission and space rental fees for the Fair and for use of the fairgrounds. It ended 1989 on a strong financial note. Although governmental and quasi-governmental organizations do contribute financially to traffic management efforts, the special events run by them and their financial quasi-independence may make it more difficult to reach agreements with them or require financial commitments than if they were subject to the normal fiscal and legislative controls imposed on municipalities in the state.

The calculation for the cost of Christmas shopping cannot be easily made. Spinoff effects, such as revenue for cities in terms of meter parking fees and taxes on parking ramps would have to be considered. So would the secondary impacts of increased restaurant patronage, gasoline taxes, and other unusual expenditures that affect the local and national economies.

CONCLUSIONS

The management of special events to control CO incidents varies with the type of event, its proximity to a violating monitor, and the institutional arrangements for handling the traffic. Some events appear to have traffic and crowd control impacts that are virtually uncontrollable. If held near monitors, they require special institutional arrangements to avoid CO incidents. For Christmas shopping, as done abroad, these could include shuttle buses and other ridesharing arrangements to ferry shoppers and parcels to and from the stores.

These arrangements are especially critical when coordinating TSM from areas surrounded by multiple jurisdictions. Agreed-on coordinators, a traffic management plan, and integrated police enforcement are the most important elements

TABLE 3 CO EXCEEDANCES FOR 1989 AND 1990

1989 CO EXCEEDANCES

DATE	LEVEL	TEMPERATURE	LOCATION
01-04-89	11.4 PPM	23 degrees F.	University & Lexington
01-04-89	9.9 PPM	28 degrees F.	Lake & Hennepin
03-10-89	9.9 PPM	38 degrees F.	Lake & Hennepin
03-27-89	10.1 PPM	53 degrees F.	Duluth
07-08-89	9.6 PPM	79 degrees F.	Lake & Hennepin
07-20-89	10.8 PPM	76 degrees F.	University & Snelling
07-23-89	9.7 PPM	73 degrees F.	Lake & Hennepin
07-23-89	11.3 PPM	73 degrees F.	Lake & Henneping
09-02-89	10.3 PPM	57 degrees F.	Larpenteur & Snelling
10-23-89	9.9 PPM	55 degrees F.	Duluth
10-23-89	10.3 PPM	59 degrees F.	University & Snelling
10-23-89	9.2 PPM	72 degrees F.	Nalpak Bldg.
10-23-89	12.1 PPM	59 degrees F.	University & Lexington
10-24-89	10.7 PPM	58 degrees F.	Nalpak Bldg.
10-24-89	12.0 PPM	50 degrees F.	1829 Portland Avenue S.
11-24-89	9.4 PPM	30 degrees F.	Lake & Hennepin

1990 CO EXCEEDANCES

DATE	LEVEL	TEMPERATURE	LOCATION
02-08-90	9.9 PPM	29 degrees F.	University & Lexington

*The Levels reflect the highest 8-hour average exceedance for that date.

in preventing violations. Regular meetings of interjurisdictional task forces do much to establish good relations that can then be used when handling these events. Jurisdictions involved need to exercise some self-discipline, both in scheduling events and dealing with their impacts on surrounding areas. They need to refrain from sending the problem elsewhere in hopes that it will disappear. Under certain conditions, it will reappear as a CO air quality incident. If it does, it will affect not only the area where the traffic was sent, but may result in sanctions or an overly tightened I/M program affecting jurisdictions far from the site.

Most special events will not result in a CO incident. Analysis of special events in the Twin Cities indicates local traffic flow problems on streets near monitors that extend onto the freeway entrance ramps, but the problems do not cause shock waves in traffic flow on the freeways. When the problems involve both the freeways and the local streets, involve vehicles that are without operating pollution control equipment, or are during the worst-case meteorological season, the potential for incidents must increase of necessity.

ACKNOWLEDGMENTS

The author wishes to thank the following individuals for information used in the production of this paper: Lts. Caroline Bailey and Lester Benson, Officer Steven Golden and Capt. Richard Eckwall of the St. Paul Police Department; Micheal Klassen, Donald Sobania, and Donald Tufte of the Traffic Engineering portion of the St. Paul Public Works Department; Lt. Robert Beckers of the Minneapolis Police Department; Larry Erb and Glen Carlson of Mn/DOT; Michael Monahan and Thomas Becker of the Traffic Engineering portion of the Minneapolis Public Works Department; Mike Heffron of the Minnesota State Fair; and Nancy Ebert of the MPCA.

REFERENCES

1. C. L. Dudek et al. Field Evaluation of Messages for Real-Time Diversion of Freeway Traffic for Special Events. *Transportation Research Record 682*, TRB, National Research Council, Washington, D.C., 1978, pp. 37–45.
2. *Transportation Research Circular 326: Freeway Incidents and Special Events: Scope of the Problem.* TRB, National Research Council, Washington, D.C., 1987, p. 5.
3. J. E. Thomas. A Special Event Parking and Transit Pass System Using Ticketron: The Rochester, New York, Tall Ships Experience. In *Transportation Research Record 1039*, TRB, National Research Council, Washington, D.C., 1985, pp. 16–21.
4. G. C. Price. Traffic Management Teams in Florida. In *Transportation Research Circular 344*, TRB, National Research Council, Washington, D.C., 1989, pp. 21–24.
5. *Transportation Research Circular 344: Traffic Management of Special Events: The 1986 U.S. Open Golf Tournament.* TRB, National Research Council, Washington, D.C., 1989, pp. 25–34.

Neither the agencies nor the individuals who provided information for this paper are responsible for its contents or conclusions.

Publication of this paper sponsored by Committee on Transportation System Management.

Mode Split at Large Special Events and Effects on Air Quality

CHARLES P. GREEN

In the past few years, the scope and number of large special events, such as concerts and sporting events, have increased. Large special events are being viewed as a means of economic development, that is, as an attempt to bring in spectators from outside the region to obtain tourism dollars. Some cities are planning large special events that may attract as many as 100,000 spectators. Little published material is available applicable to the planning of large special events, especially in the areas of mode split (the mode of travel to the event) and the air quality impacts of these large events. A study was recently performed in nine large metropolitan areas nationwide to gain further knowledge of the impact of these events. The data collected were analyzed to derive conclusions and recommendations for assisting officials in planning of large special events. Analysis of the data revealed that average automobile occupancy for weekend events is much higher than on weekdays; automobile occupancy for higher-priced events is significantly lower than for lower-priced events; automobile occupancy at Western U.S. sites is much lower than at Eastern and Midwestern U.S. sites; transit usage where regularly scheduled special transit service is available is much higher than where charter or nonscheduled special transit service is provided; and transit fares have a significant impact on the transit mode share. Vehicles attending large special events can have large impacts on air quality. If resources become available, a model to be used as a tool for planning special events should be developed.

In the past few years, the scope and number of large special events, such as concerts and sporting events have grown. Many concerts are now being held in large outdoor stadia, whereas 10 years ago they were held mainly in medium- and small-sized indoor arenas. These special events have necessitated using the large stadium or large outdoor setting, to accommodate more spectators in an attempt to recoup the costs. Some cities, including Denver, have sponsored or are sponsoring large special events, such as automobile races, that are expected to attract as many as 100,000 spectators. In addition, many cities are vying for major league sports teams, with the promise of new stadium sites to accommodate the specific sport.

Little material has been published analyzing large special events. In fact, a computer literature search using keywords such as "special events" and "air quality" did not turn up a single study or report on this issue (*1*). Little, if any, work has been done to analyze large special events in terms of mode split or the mode of transportation people take to get to the event. Even less analysis has been performed on the air quality impacts of these events. This study, which was performed in nine large metropolitan areas nationwide, is aimed at collecting and analyzing data pertaining to the mode of transportation people take to get to large special events, evaluating the impact of the large event on air quality, and deriving conclusions and recommendations aimed at improving traffic flow to and from large events.

Denver Regional Council of Governments, 2480 West 26th Avenue, Suite 200B, Denver, Colo. 80211.

Various problems arise from lack of information to plan for large special events. A large sudden influx of automobiles can grind the surrounding roadway network to a halt. This traffic congestion also is directly correlated to air quality; not only is the air quality in the immediate vicinity of the special event site impacted, but air quality at points away from the site can also be impactd by the effects of congestion as well. For instance, a Denver Broncos' football game, which attracts more than 75,000 fans on a given event day (usually Sunday), can generate over 4 tons of carbon monoxide (CO) and over 800 lb of hydrocarbons alone in the area within 1 mi of the stadium. Total pollutant emissions from this crowd regionwide can be as much as 40 tons of CO and 4 tons of hydrocarbons (HC). These impacts can be exacerbated if the event is held on a weekday.

Because of the fiscal constraints of this study, the number of sites sampled is relatively small and should not be considered as a representative cross section of the entire United States. Where possible, both weekday and weekend samples were taken at the same site. This was possible at the Denver and Cleveland sites. Sufficient samples were taken to enable grouping of the data, such as by weekend and weekday, by automobile and transit, etc., to attempt to derive general conclusions regarding large special events. Further study is needed to reduce any large variance in the data. Where possible, local information should replace the averaged information presented here.

DATA COLLECTION

Baseball and football games and large concerts were surveyed to obtain the necessary data. Nine metropolitan areas were chosen for this analysis. Figure 1 shows the metropolitan areas sampled for this study. The metropolitan areas were chosen to attempt to obtain a nationwide distribution of sample sites, and to obtain equivalent samples from each of the three United States regions (East, Midwest, West). Two sites, Denver and Cleveland, were sampled for both weekday and weekend events. Automobile occupancy information was collected on-site; transit information was obtained from the transit agencies serving the event site. The data collection phase started in July 1987 and was completed in July 1990.

Information gathered includes the attendance at the event, the day of the week of the event, automobile occupancy, and

- Baltimore
- Chicago
- Cleveland
- Denver
- Detroit
- Kansas City
- Los Angeles
- New York
- St. Louis

FIGURE 1 Metropolitan areas sampled.

transit patronage, as well as factors that could influence mode choice, such as parking costs, ticket prices, availability of transit, and transit fares.

In order to obtain average automobile occupancy, at least 100 automobiles entering the parking areas at the event site were surveyed. This sample was determined to be a 1 percent sample, which results in an error of ±5 percent. Transit agencies were surveyed to obtain information on transit ridership and level of service to the events. Event sponsors were surveyed for additional data, such as charter bus ridership to the event and ticket and parking price information. Arrival times were tallied to determine the distribution of arriving vehicles.

DATA SYNTHESIS AND RESULTS

The data collected have been synthesized into an electronic data base and are being analyzed. Two-sample *t*-tests and multiple sample correlation tests were performed on the data to determine correlation among certain data elements. Information has been categorized into weekends versus weekdays, and various levels of parking and ticket costs as well as type of transit service (public versus charter) were tested for correlation with automobile occupancy and transit usage.

Distribution of Arrivals to the Event

Figure 2 shows the distribution of automobile arrivals to the event site. The reference lines on the figure show the cumulative percent of arrivals at the interval within 1 hr of the start of the event. In some cases, this can mean as many as 10,000 automobiles arriving in 1 hr. Translated into highway capacity equivalents, this is the same as 5 lanes of freeway, or 15 lanes of arterial roadway. This second equivalency will serve to form a basis of this analysis, because most stadium facilities use the arterial roadway system to access the site, rather than direct freeway access. Because 15 roadway lanes is a large amount of capacity to be supplied, the access roads to the stadium can easily become overloaded before the start of the event and after the event, as well.

In the case of the Denver Broncos and their event site, Mile High Stadium, Figure 3 shows the roadway network accessing the stadium and the available number of arterial traffic lanes. There are 10 arterial traffic lanes provided. In the hour preceding the start of the football game, 10,000 automobiles (plus 200 buses) create a volume-to-capacity ratio

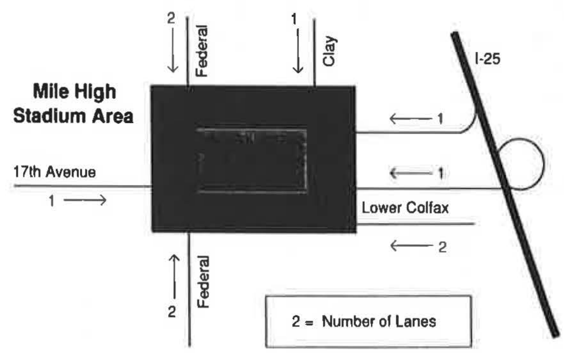


FIGURE 3 Access roads to Mile High Stadium.

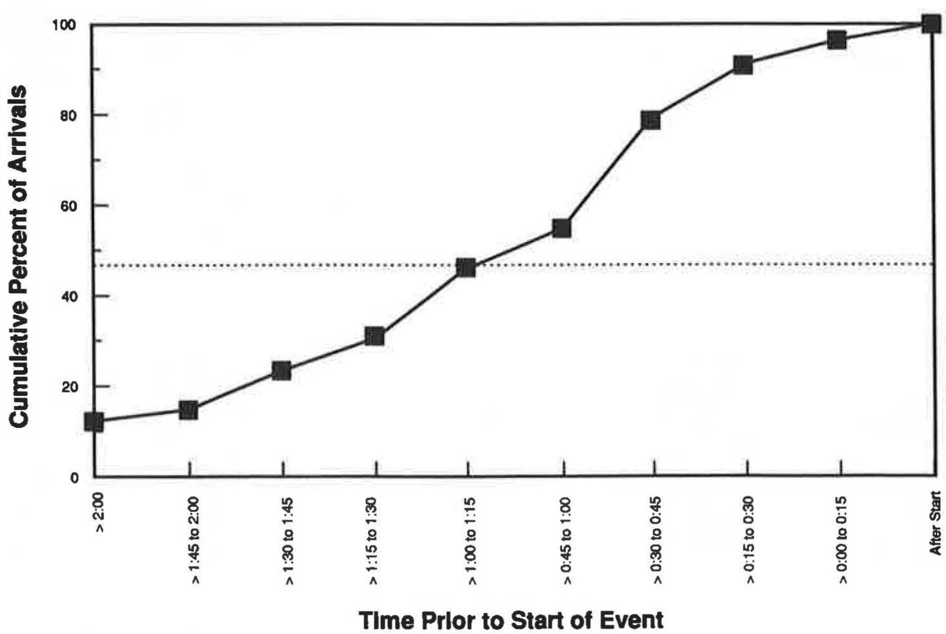


FIGURE 2 Cumulative percent of arrivals by time of arrival.

of 1.40, resulting in speeds in the area within a 1-mi radius of the stadium averaging between 10 and 15 mph. Using automobile emission factors supplied by the Colorado Air Pollution Control Division (from EPA's MOBILE 4 computer program), this translates to 4 tons of CO and 800 lb of HC in the stadium vicinity alone. Assuming the average event-goer travels 10 mi at an average of 30 mph to attend the event (a liberal assumption), the total number of vehicles attending a Broncos' football game can produce as much as 40 tons of CO (based on an average CO emission rate of 15 grams per mile per vehicle) and 4 tons of HC (based on an average emission rate of 1.5 grams per mile per vehicle). These figures can be doubled if the game is held on a week night, when the event traffic is combined with the regular 4:30 to 6:30 p.m. peak period (week night football game starting times are usually either 6:00 p.m. or 7:00 p.m. Mountain Time). Before a recent Monday night football game, gridlock conditions were observed as far away as 5 mi from the stadium. The Broncos experience a rather high 15 percent transit share to their games, so the driving situation could be a lot worse.

The departure from the event is worse. Data collected from this study indicate that as many as 95 percent of the event-goers leave the event within 30 min of the end of the event (this time can vary if the special event is a sporting event and one team has a comfortable margin over the other team as the event nears completion).

Automobile Trips

Figure 4 shows the automobile occupancy for weekday and weekend events. Average automobile occupancy for weekend events is some 11.8 percent higher than for weekdays (2.93 versus 2.62). Although this difference may not seem significant, for an event attended by 75,000 fans, assuming 5 percent transit usage, this 11.8 percent difference could account for as many as 3,000 additional automobiles being driven to the event. These 3,000 automobiles contribute over 900 lb of CO and over 90 lb of HC to the air, just within 1 mi of the event site. The automobile occupancy difference is probably caused by the limited ability for leaving work at 5:00 p.m. on a weekday, returning home, forming a carpool, and arriving at the event on time. (Starting times for large evening events are usually 7:30 p.m.)

In many mode choice models nationwide, parking cost is a major determinant in the decision to drive alone, carpool, or take transit. However, as Figure 5 shows, there is no significant difference in average automobile occupancy between sites with parking costs of less than $4.00 and sites with parking costs of $4.00 or more. For events with parking costs of less than $4.00, the average automobile occupancy was 2.78, compared to 2.88 for sites with $4.00 or more parking costs. This difference is probably because of sharing of parking costs among automobile occupancy; the resultant cost per person is a little over $1.00. (The $4.00 parking cost breakpoint was selected to enable equivalent sample sizes.)

Although parking cost does not have any significant impact on automobile occupancy, there is a strong correlation between ticket price and automobile occupancy. Figure 6 shows that for average ticket prices under $10.00, the average automobile occupany is 15.6 percent higher than for ticket prices over $10.00 (2.96 versus 2.56). (The $10.00 breakpoint was used to maintain equivalent sample sizes.) Ticket price is related to demand for and availability of tickets; higher ticket prices usually suggest that demand for the tickets is higher, which many times is accompanied by limits on the number of tickets one can purchase (or that one can afford). Because carpooling to events is usually family or friend oriented, the ability to form large carpools to high-priced events is limited. Figure 7 shows average automobile occupancies by region of the United States. There is no difference in automobile occupancy between Eastern and Midwestern region sites. However, automobile occupancy at Western region sites was significantly lower. This can be attributable to much lower population densities in the West, as indicated in Table 1. Population density has been demonstrated to be a major influence on automobile occupancy; lower densities result in

FIGURE 5 Average automobile occupancy by parking cost category.

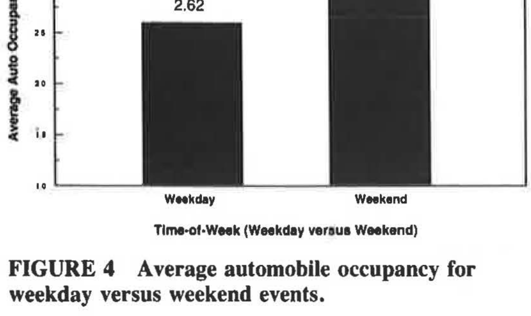


FIGURE 4 Average automobile occupancy for weekday versus weekend events.

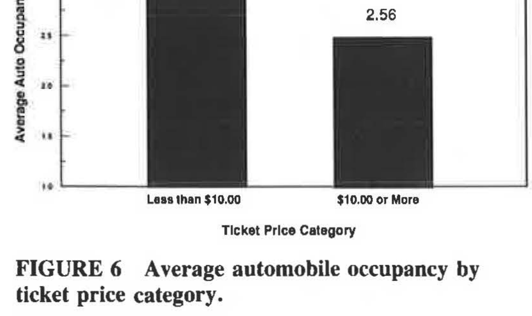


FIGURE 6 Average automobile occupancy by ticket price category.

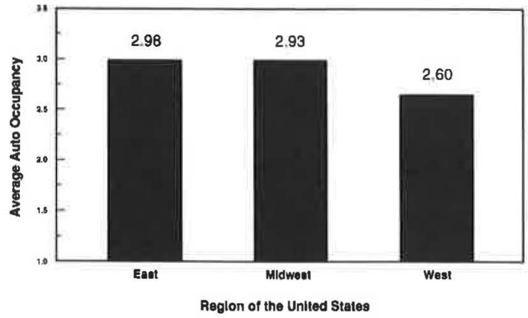


FIGURE 7 Average automobile occupancy by region of the United States.

TABLE 1 POPULATION DENSITY BY REGION, 1986 (*2*)

Region	Density (Pop./Sq. Mi.)
East	1,938
Midwest	931
West	359

longer travel times and distances to form a carpool, thus making carpooling a less attractive mode.

Transit Trips

Transit usage fluctuated widely between sample sites, varying from less than 1 percent to over 20 percent. Several factors seemed to influence the transit mode share. The strongest factor was the availability of public transit to the event site. As Figure 8 shows, sites with public transit available averaged 8.9 percent of total event-goers using transit, whereas the average for charter-only service was around 1 percent. This difference is apparently caused by the familiarity with the public transit operator (one who is visible on a daily basis), as well as the level of service. Also, public transit operators are publicly subsidized, in most cases, allowing them to charge a lower fare than private charter services. Also, places such as Denver, New York, and Kansas City (with the highest transit mode share of the sites sampled) give some preferential treatment to transit, such as close-in parking and exclusive

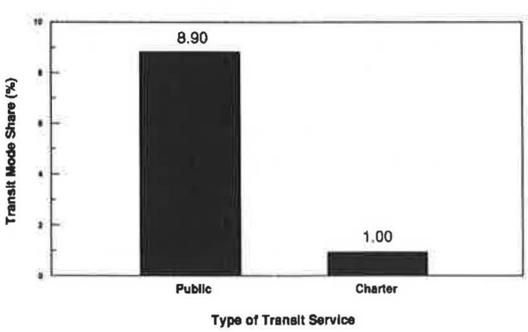


FIGURE 8 Transit mode share by type of transit service.

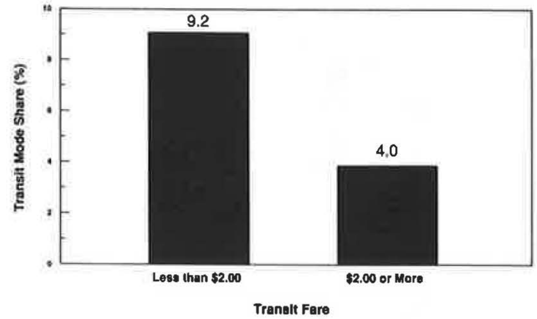


FIGURE 9 Transit mode share by transit fare.

gates for direct access onto entrance and exit roadways at the event site. Figure 9 compares transit fares with transit usage. Also, public transit operators are able to advertise their service to special events, many times on a daily basis (advertisements on board the transit vehicles, brochures or flyers, schedules, etc.).

This difference in transit mode share can have a large impact on air quality. For the example of an event attended by 75,000 spectators, the difference between 8.9 and 1.0 percent mode share can mean an additional 2,000 automobiles being driven to the events site, resulting in an additional 600 lb of CO and an additional 60 lb of HC being emitted into the air.

More people ride transit away from the event than ride it to the event (letter from R. S. Page, Southern California Rapid Transit District). They attributed this phenomenon to people receiving rides to the event (such as friends carpooling from work) and then taking transit to get home. Another possibility is that some people may take a regularly scheduled route to get to the event site, and use the special service to return home. This may be more evident in Denver, where the Denver Regional Transportation District charges users on the trip to the event, and does not charge for the return trip.

The number of attendees from outside the region (defined as the census urbanized area) has an impact on transit share. As Figure 10 shows, as the percentage of attendees from outside the region increases, the transit share declines. This is because of the outsider's unfamiliarity with the local transit system.

The size of the event apparently has no impact on transit usage. For those events with attendance of less than 35,000, the transit share averaged 6.2 percent, whereas for events attended by 35,000 or more, the transit share was 6.8 percent. Statistical tests indicated that there was no significant difference between the transit shares.

RECOMMENDATIONS

After analyzing the data and determining the operational goals of those sponsoring large events, the following recommendations were developed.

When at all possible, regularly scheduled transit service should be provided to the event site.

As shown in Figure 8, those event sites with public, regularly scheduled transit service experienced significantly higher tran-

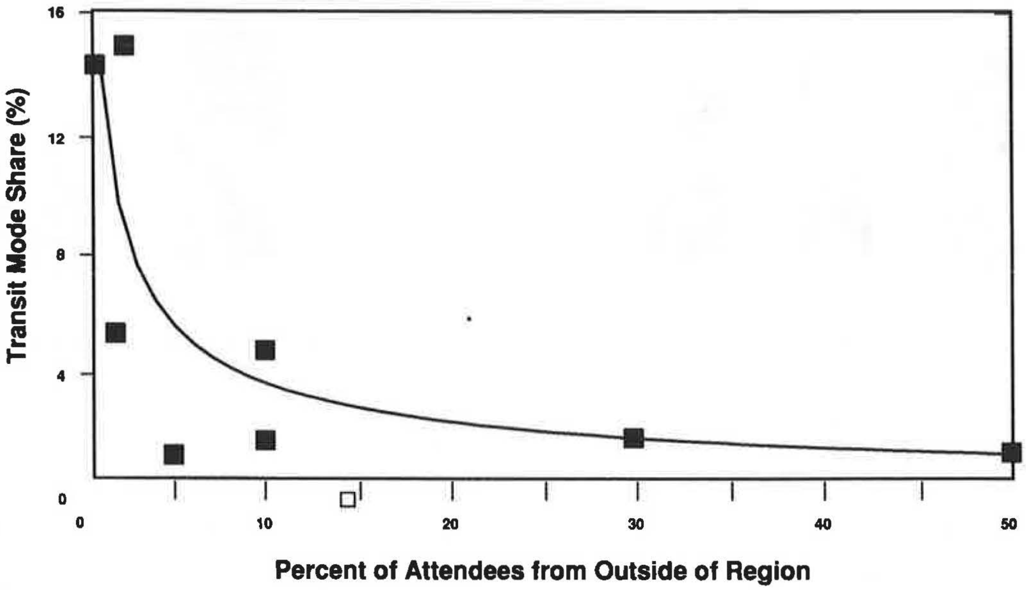


FIGURE 10 Transit share by percent of attendees from outside the region.

sit patronage than those sites served with nonscheduled, charter-type operations. This can be significant, especially in terms of traffic flow and air quality. For the example of an event attended by 75,000 people, if only 5 percent used transit, this would result in some 1,300 fewer automobiles traveling to the site. Using the assumption that the average fan travels 10 mi to the event (this average would be greater for a stadium located in a noncentralized area, such as an outlying suburb), and averages 25 mph on the drive to the event, this would result in a reduction of over 400 lb of CO and over 40 lb of HC for the area within 1 mi of the event site. Regionwide reductions in air pollutants would be higher.

Special events should be held on weekends, or the starting times on weekdays should be delayed at least one-half hour to allow event attendees more time to form carpools or take transit.

Any increase in automobile occupancy or transit share can only be beneficial to air quality. For example, for the event with 75,000 attendees, the difference between weekday and weekend automobile occupany alone results in 3,000 fewer vehicles attending an event of the same size on weekends, compared to weekdays. This means a savings of over 900 lb of CO and over 90 lb of HC in the stadium vicinity alone. Again, region-wide savings would be much higher.

Shuttle services from outlying areas should be provided if there are a large number of event-goers from outside of the region.

As shown in Figure 10, as the percent of event-goers from outside the region increases, their transit share decreases. It is obvious that this is because of a lack of knowledge of the transit service to the event site on the part of those from outside the region. Thus, if a large percentage of event-goers from outside the region are attending an event—such as a sporting event between rival teams in relatively close geographic proximity, a large event that is held a few times per year, or a major stadium concert—to reduce traffic congestion near the event site, a shuttle system should be set up. Obviously, because of the unfamiliarity with the local transit system, shuttle lots should be located on major approaches to the event site (such as Interstate freeways), should be located in outlying areas well removed from the vicinity of the event site, and should be clearly marked on the highway. Park-and-ride lots usually are located along major routes, and would make excellent shuttle lot locations, especially for a centrally located event site.

Transit vehicles should be given priority at events, to make transit usage more attractive.

Many studies of transit systems have indicated that drivers need to be given incentives to shift from automobiles to transit. Transit service alone to an event site cannot be relied on as a major effort to reduce congestion at the event site. Some transit priority measures already in place include close-in parking or transit stations and preferential entrances and exits. These transit priority measures and others, such as bus-only lanes on approaches to the event site, should be considered at all special event sites. Mode split studies performed nationwide have shown that passengers will switch to an alternative mode if that mode can save 1 min of travel time per mile traveled. In the case of a large special event, the time savings will occur in the area within 1 mi of the event site. For an average travel distance of 10 mi, the travel time savings should be approximately 10 min to switch automobile riders to transit. With these transit priority measures in place, it is indeed possible to save 10 min of travel time compared with the automobile mode.

Incentives to arrive early and leave late should be considered to reduce the magnitude of congestion before and after the event.

Figure 2 shows that more than 50 percent of event-goers arrive within 1 hr of the start of the event; other information collected shows that as many as 95 percent of the event-goers leave within 30 min of the end of the event. Incentives should be considered that will allow people to arrive at an event before the peak hour before the start of the event, and to leave the event site well after the end of the event, to reduce the travel "spike" that causes extreme congestion. Incentives to arrive early at events could include pregame (tailgate) parties or preevent festivities. Some sporting events hold Old Timer games or some type of skill contests well in advance of the main event. Incentives to delay departure from the event site could include showing highlights of the event on large video screens, a post-game party (which could be sponsored by a local radio or television station), or coupons for food at nearby restaurants within walking distance of the event facility. (Incentives for the purchase of alcoholic beverages, such as free-drink coupons, should be strongly discouraged.)

CONTINUING STUDY

The data analysis portion of this study should be considered as ongoing, as more information will be included into the inventory in the future. A computer model will be developed and made available for assistance in planning large special events. The model will use as input the following parameters:

- Number of tickets available or expected attendance,
- Type and level of service of transit available,
- Region of the United States where the event is being held,
- Parking costs,
- Transit fares,
- Ticket cost,
- Day of the week of the event, and
- Other parameters affecting mode split that prove statistically significant.

The model will have as output the following:

- Expected average automobile occupancy and number of automobiles,
- Expected transit mode share,
- Estimated air quality impacts,
- Recommendations for traffic congestion alleviation, and
- Other output parameters determined to be of use in planning major events.

If resources become available, work on model calibration and validation could begin in 1991.

SUGGESTIONS FOR FURTHER STUDY

During the data analysis process, several data anomalies appeared that suggested that further study was necessary. These were the following:

- A general decrease in automobile occupancy as the attendance increases, as shown in Figure 11.
- The West experiences the highest transit mode share to large special events, as shown in Figure 12.
- The transit share for event sites with automobile parking costs of less than \$4.00 is more than double the transit share at event sites with parking costs of \$4.00 or more. This is contrary to mode split studies performed for regional transportation planning purposes, however, there may be some correlation between the indifference in automobile occupancy relating to parking costs.
- The transit mode share on weekdays is significantly higher than on weekends. This may be caused by improved transit service levels during the week (i.e., more frequent service).
- There is general decline in transit mode share as the automobile occupancy increases. This decline may indicate that

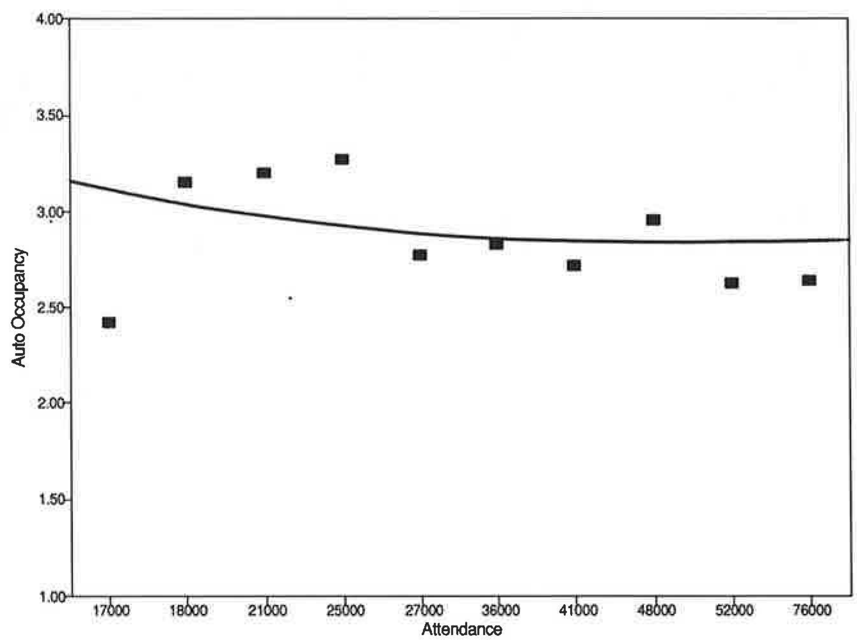


FIGURE 11 Automobile occupancy by attendance.

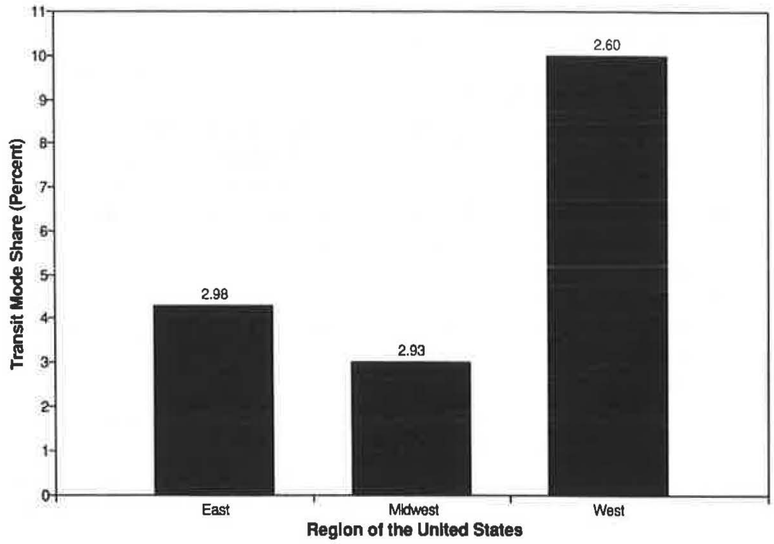


FIGURE 12 Transit mode share by region of the United States.

there is a trade-off between transit riders and automobile passengers.

- There may be some correlation between large special events and air quality monitoring stations reporting air pollution violations. Data collected from the Colorado Air Pollution Control Division and a report from the City and County of Denver (*3*) indicate some correlation between Denver's annual Parade of Lights and high pollution days.

It is also recommended that the study be continued to include more sample data. Because the data included in this report were grouped for testing, in most instances regional travel behavior differences may not appear in these results. Also, because of the relatively small number of samples taken, it is suggested tht continued data collection is needed. If sufficient resources become available, it is also suggested that the above anomalies be studied further. It is possible that a more intensive data collection effort, if not able to explain these anomalies, would serve to reduce or eliminate them.

ACKNOWLEDGMENT

The author acknowledges the assistance of the following agencies in obtaining transit and attendance information: Southern California Rapid Transit District, Chicago Transit Authority, Kansas City Area Transportation Authority, St. Louis Cardinals Baseball Club, New York City Transit Authority, Southeast Michigan Area Rapid Transit Authority, Bi-State Development Agency (St. Louis), Denver Regional Transportation District, Colorado Department of Health—Air Pollution Control Division, and City and County of Denver—Department of Health and Hospitals.

REFERENCES

1. S. Spitzer. Special Events and CO Violations. Presented to the Transportation Research Board summer meeting, Denver, Colo., July 1990.
2. *State and Metropolitan Area Data Book.* U.S. Bureau of the Census, 1986.
3. S. Foute. Air Pollution Prevention for Special Events in Downtown Denver. Presented to the Transportation Research Board summer meeting, Denver, Colo., July 1990.

Publication of this paper sponsored by Committee on Transportation System Management.

Pricing of Air Pollution in the Swedish Transport Policy

LARS HANSSON

Swedish transportation policy has radically changed during the last decade. In 1979, the principle of a social marginal cost responsibility for road and rail traffic was introduced. The consideration of social costs for traffic accidents was the main innovation. Air pollution was only mentioned as an external effect that could not be calculated at that time. In 1988, when the Swedish parliament adopted the new Transport Policy Act, an essential part of the new transportation policy was the principle of internalization of some of the traffic emissions. They were accordingly considered explicitly in the infrastructure charges for road and rail traffic, as well as for domestic aviation. The explicit evaluations of external effects in Sweden imply road user charges (gasoline and kilometer taxes) that amount to a cost recovery almost 3.5 times higher than the budgetary costs for highways and roads. The same cost responsibility for rail traffic would only increase its corresponding budget about 10 percent.

The ultimate goal of Swedish transportation policy (*1,2*) is to help maintain and develop prosperity. The general goal of the policy is "to provide the population and industry throughout the country with adequate, safe, and environmentally acceptable transport services at the lowest possible social costs." The policy is specified in five objectives:

- To increase efficiency of the transport system,
- To reduce the environmental damage,
- To increase road safety,
- To ensure adequate transport services in all parts of the country, and
- To promote regional balance.

These objectives define the orientation of a transportation policy that intends to pursue two different aspects of efficiency:

- To do things right, and
- To do the right things.

To do things right is to increase the efficiency of the transportation system as a whole. A measure of particular importance in current transportation policy is thus a deregulation of the transportation market.

However, to do things right is only one aspect of the market economy and the resource allocation. An equally important aspect, which is emphasized in the Swedish transportation policy, is to do the right things.

In order to provide services in areas where the population base is limited, public transportation is considered a necessity. The infrastructure and communications have played a historical role as an instrument of regional policy. The transport system should continue to play a part in achieving a balanced population trend and providing places of work and essential services in all parts of the country.

Another vital issue is the traffic impact on the environment. The transportation policy must be instrumental in promoting transportation that is environmentally acceptable and in reducing the effects of traffic on the environment. Being mainly a question of road transportation, the traffic safety is another issue of vital interest.

The principles of the Swedish transportation policy are shown in Figure 1. Transport modes and transport carriers shall operate and compete in a free transport market. This maxim is expected to increase efficiency in the transport system. It implies, e.g., that the Swedish State Railways (SJ) should operate traffic in a strictly commercial manner. If traffic is unprofitable for SJ, then it should cease.

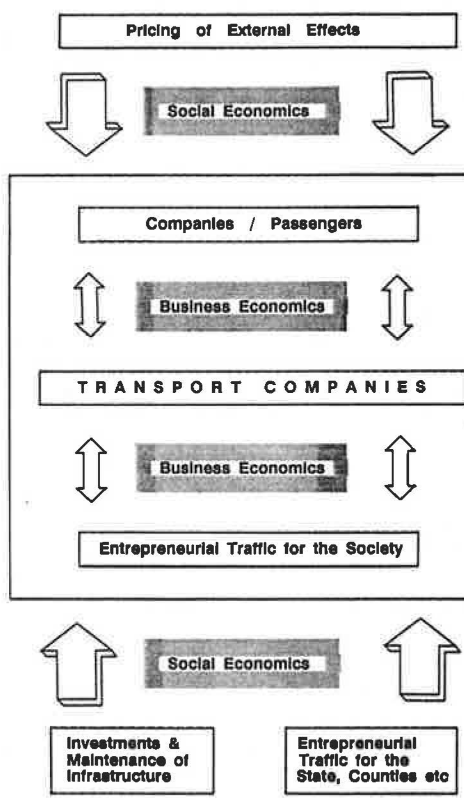


FIGURE 1 Swedish transport policy.

Swedish State Railways, Central Station Building, Stockholm, S-105 50, Sweden.

The transport market is affected by transportation policy that contains "rules of the game" to do right things. The three most important of these activities are

- Charges for traffic accidents and air pollution, the two most important external costs on the transport market;
- Investment appraisals of rail infrastructure, which are based on cost-benefit analysis in the same way as it is used for road investments;
- The entrepreneurial role of transport operators. If the Swedish Transport Board (which can purchase interregional passenger traffic) or the 24-County Passenger Transport Authorities (responsible for regional public transportation) want more traffic than is undertaken by the commercial operators, then they will purchase such traffic.

ROAD AND RAIL USER CHARGES

On July 1, 1988, the SJ was divided into two separate bodies: SJ, still owned by the State, and the National Rail Administration. The former is required to act as a strictly commercial transport enterprise, while the latter is responsible for railway infrastructure, acting as a governmental authority.

The National Rail Administration will act on the same formal inducement as the National Road Administration, i.e., accomplish a railway investment policy on the basis of cost-benefit analyses. SJ (and other prospective railway companies) is charged for the use of the railway network. The principle for charging is a two-part tariff, which also is applied to road traffic.

An important market principle is the internalization of external effects, i.e., pricing of the external costs for society.

The negative external effects are essentially of four different types: traffic accidents, air pollution, noise disturbance, and congestion. If these effects are not considered properly, it will cause an allocation of resources that

- Produce and consume too much transportation, and
- Favor transport modes that create many negative effects at the expense of those transport modes that cause few negative effects.

Traffic safety effects, traffic emissions, and congestion (only for rural road traffic) are considered explicitly in the infrastructure charges. Noise effects are explicitly evaluated in the investment appraisals used both by the National Road Administration and the National Rail Administration. They are hitherto not included in the pricing of external effects.

The pricing of external effects of accidents is applicable only for the transport sector. The air pollution charges are applicable both for the industrial and the transport sector. In the transport sector, international aviation and all maritime traffic are excluded.

These evaluations provide new figures in the analyses of cost responsibilities in transportation. The cost responsibility results should not be confused with traffic charges actually paid. Today, only private cars with catalytic converters, rail traffic on the main line system, and domestic aviation pay according to the current evaluation of external costs. Private cars without catalytic converters (65 percent of passenger car kilometers in 1990), trucks, buses and coaches, together with rail traffic on the county line system, pay only part of their external costs. International aviation and maritime traffic pay no charges at all.

Despite insufficient charges for most of the traffic, the pricing of external effects has affected the user charges to a large degree. One example is the increase of kilometer taxes, weight-distance taxes used for diesel vehicles. In 1989, the calculations accounted for the transportation policy act led to a 45 percent increase in the kilometer taxes for trucks and a 100 percent increase for buses and coaches.

The total social marginal costs for trucks were calculated as follows: 43 percent wear and tear, 26 percent external costs from traffic accidents, and 31 percent air pollution charges. The trucks paid for 58 percent of these costs, i.e., they paid for all their budgetary costs consisting of wear and tear. They paid only for about one-quarter of the external costs for traffic accidents and air pollution. The 45 to 100 percent increases in the kilometer taxes were insufficient, but political difficulties kept them to these levels.

PRICING OF AIR POLLUTION

In economic theory, the basis for the evaluation of costs and benefits is willingness to pay. If someone suffers from an external effect, there will be a certain amount of money that can compensate for the disutility. In cost-benefit analysis, that amount is used as a social cost. If the sufferers are compensated with at least this amount, it will be equivalent to a business transaction in the market, i.e., a price label on the external effect.

External effects such as noise, air pollution, and visual intrusion are normally considered to be incommensurable effects. In spite of this, explicit choices about these effects are sometimes made, i.e., they are traded off. Knowledge about these trade-offs provide the same information about the evaluation of the external effects as market prices do for commensurable goods and services.

It is important to realize that the willingness-to-pay concept is a compensation for those currently being affected. Also the concept is based on existing preferences. Both these prerequisites are to be considered when air pollution is evaluated, which requires a more extensive point of view.

How can the emissions of CO_2 and its effects on the environment be evaluated? There are no financial costs for the feared green-house effect. It's not fruitful to ask people about their willingness to pay for reduced CO_2, or to look for their revealed preferences. It can only be concluded that there are no ways to deduce a proper value or a cost for the emission of an external effect such as CO_2.

Baumol and Oates (*3*) approach this problem by drawing attention to the fact there is an inability to measure marginal social costs:

> If there is little hope of estimating the damage that is currently generated, how much less likely it is that we can evaluate the damage that would occur in an optimal world which we have never experienced or even described in quantitative terms.

The problem is thus that the environment impacts can neither be estimated nor evaluated, i.e., a cost responsibility determined. Even if an optimal charge can't be defined, society is compelled to establish restrictions, regulations, economic incentives, etc., to reduce the environmental impacts. These measures are undertaken in the light of the fact that nature can't manage current pollution.

For the emissions of NO_x, HC, and SO_2, there are some measureable costs. But these costs only cover some of the expected effects. Environmental quality aspects (outdoor life, the historic remnants, etc.) are not included. Of what value is it that people, if it was possible, express a willingness to pay for something they don't know about or can't imagine. Even if some preferences could be observed, the preferences for the environmental qualities change over time as people meet with alterations. Yesterday's evaluations are not the same as today's and tomorrow the effects of air pollution are more severe than they are today.

Finally, there is the inheritance to future generations. Historic remnants are a gift not only to the current generation, but also to the next generations. Most important of all, future ecological catastrophe is equivalent to a consumption, where the bill has to be paid by future generations.

Even if there is no way to make acceptable cost-benefit calculations, it's important to undertake some measures to reduce the emissions. The relevant issue is to find out what nature can tolerate. Most ecologists and environmental researchers (in Europe) agree that NO_x must be reduced by 70 to 80 percent to attain a sustainable ecological balance for environment. The Swedish parliament has adopted the objectives presented in Table 1 for reduction of emissions.

With these objectives as a background, the pricing of external effects may be motivated by two different reasons; on the one hand, a cost responsibility can be considered as a compensation for costs or disutilities; on the other hand, it can be used as an economic instrument to reach an aim or end as cost efficiently as possible. These two aspects concur as an incentive, which results in a better resource allocation. The current policy is based on the second approach.

Once an environmental objective is formulated, e.g., as a specified limitation for an emission, the question is how to reach this objective. Legal restrictions and preventive measures can be used to some extent. The question is how to reach the reduction with a minimum of costs for the variety of restrictions and measures possible.

Cost minima for any specified limitation of emissions are reached if externality charges are used. The principle incentives given by an externality charge for emissions are the choice between the three following alternatives:

- Continue the emissions, and pay the charge;
- Take measures to reduce the emissions, and thus pay a reduced charge;
- Stop production (e.g., stop making the journey) and pay no charge at all.

TABLE 1 ENVIRONMENTAL OBJECTIVES FOR EMISSIONS

Compound	Reduction (%)	Period
NO_x	30	1980–1995
SO_2	80	1980–2000
CO_2	0	1988–
HC	50	–

The level of the charge decides whether it is favorable or not to take measures to reduce the activity creating the emissions. Those activities with the lowest costs for reducing the emissions (to a specified level) are automatically chosen by the market mechanism.

The conclusion is thus that a cost responsibility should be implemented for the incommensurable external effects, even if there are no costs for them.

AIR POLLUTION FEES AND CHARGES IN SWEDISH TRANSPORTATION POLICY

Swedish transportation policy has changed radically during the last decade. In 1979, the principle of a social marginal cost responsibility for road and rail traffic was introduced. The consideration of social costs for traffic accidents was the main innovation. Air pollution was only mentioned as an external effect that could not be calculated at that time. In 1988, when the Swedish parliament adopted the new Transport Policy Act, an essential part of the new transportation policy was the principle of internalization of some of the traffic emissions. They were accordingly considered explicitly in the infrastructure charges for road and rail traffic, as well as for domestic aviation.

The air pollution fees were developed in three steps. The first step was taken in 1982 when a commission appointed by the Department of Transportation tried to calculate the costs for air pollution. In 1985, the commission presented some cost calculations, e.g., loss in productivity in forestry, corrosion from SO_2, and health effects. However, the calculations were rudimentary. They were therefore questioned for being both underestimated and overrated.

The second step was taken when parliament adopted the 1988 Transport Policy Act. The main philosophy was that environmental measures undertaken should be consistent. This means that the same evaluations should be used for reducing the same kind of emissions with different measures.

Implicit evaluations of some current measures implemented, e.g., catalytic converters (compulsory for new cars from 1989); stricter emission standards for diesel engines (compulsory for heavy vehicles from 1994); regulations restricting the emissions from combustion plants; etc., were calculated as a support for the explicit evaluations. Implicit values of NO_x reductions ranged from 10 to 80 SEK per kilogram, where HC and S were weighed together with NO_x to an NO_x equivalent (1 SEK = 0.18 U.S. dollar).

In the Transport Policy Act it was suggested that some caution should be applicable to the incorporation of environmental effects in the social marginal cost calculations. Despite this, the Act nevertheless ascertained that it was necessary to consider these costs when variable charges for the use of transport infrastructure (gas taxes, kilometer taxes, rail charges, and landing charges) are to be determined. The basis for evaluation of air pollution was that they corresponded to a cost responsibility of 15 SEK per kilogram of discharge of NO_x.

The final step was taken in 1990, when the Commission on Economic Instruments in Environmental Policy (*4,5*) presented analyses of the scope for using economic instruments in environmental policy, and submitted proposals to the parliament for the framing of such instruments. The charge proposals applicable to the transport market are presented in Table 2.

Part of these charges have already been adopted by parliament. In 1988, the emissions of NO_x and HC were included in the cost responsibility for road, rail, and domestic aviation. The explicit charges were 15 and 7.50 SEK per kilogram, respectively. In 1990, the charge for CO_2 was implemented for road traffic and domestic aviation. On January 1, 1991, the sulfur charge was imposed on coal, peat, and oil. On January 1, 1992, the charge of 40 SEK per kilogram of NO_x will be imposed on large combustion plants.

In addition to these charges, there is a tax difference between leaded and unleaded petrol amounting to 0.30 SEK/L.

In February 1991, a government bill will be presented to parliament, based on the charge proposals suggested by the Commission on Economic Instruments in Environmental Policy.

EFFECTS OF THE PRICING PRINCIPLE

What do these monetary values stand for? The first calculations for the transport sector in 1985 were mainly based on financial costs. The 1988 evaluations were explicit values based on implicit values. The proposed values laid the main stress on pricing to attain objectives. The reasons for this change in emphasis of evaluation principles are the inherent shortcomings of the traditional basis for social costs.

Even if it is futile to make cost-benefit calculations, etc., in order to find the optimal economics, it is crucial to undertake measures to reduce the emissions. However, the relevant issue is to find out what nature can withstand, i.e., the required reduction in emissions in order to attain a sustainable ecological balance. Once an environmental objective is formulated, e.g., as a specified limitation for an emission, the question is how to reach this goal.

Any specified limitation of emissions will be reached more efficiently, i.e., to decrease total costs, if externality charges are used. Those activities with the lowest costs for reducing the emissions (per kilogram, etc.) are automatically chosen by support of the market mechanism.

Without a charge, there are only costs for measures reducing the emissions. With emission charges, there are counteracting reductions in the total charges paid for the emissions. The levels of different charges thus determine whether it is favorable or not in different branches to take measures to reduce emissions.

When the NO_x and HC charges were introduced in 1988, LIN, the major domestic airline in Sweden, started to replace the combustion chambers on their Fokker F28. The emissions of HC were reduced by 90 percent, NO_x by 15 percent. The pollution charges for an average flight (380 km) were consequently reduced from $100 to $26.

The cost for the replacement program was $4.4 million plus $0.5 million per year in variable costs. On a yearly basis, LIN will save about $3.6 million. If the replacements are written off, e.g., during 5 years and with a 5 percent real discount rate, the net return is 200 percent.

TABLE 2 AIR POLLUTION CHARGES IN SWEDISH TRANSPORTATION POLICY

	Charge	
Substance	(SEK/kg)	($/kg)
Sulfur	30.00	5.25
Nitrogen oxides	40.00	7.00
Hydrocarbons	20.00	3.50
Carbon dioxide	0.25	0.04

SOME EXAMPLES OF THE RELATIVE EFFECTS OF CHARGES FOR EMISSIONS

Highway and Road Traffic

The cost responsibility based on external effects will only affect the variable road user charges. Here the external costs for road accidents and air pollution are much more important than the financial costs for road maintenance and traffic surveillance (*6*).

Concerning nonurban traffic, the social marginal costs (including a minor external cost for congestion) for cars are almost 10 times higher than the budgetary costs for road maintenance and traffic surveillance, compared with the ratio of about 15 for cars without catalytic converters. The marginal external effects (i.e., the average cost responsibilities, which are based on the average of marginal effects for different roads and highways), are presented in Table 3 for vehicles with fuel consumptions of 0.8 and 1.0 L per 10 km.

An external cost of 10 cents per 10 km corresponds to a traffic charge of 10 cents per liter of gas for a vehicle with fuel consumption of 10 km/L. For a car with a fuel consumption of 12.5 km/L, the traffic charge would be 12.5 cents per liter of gas. For cars with and without catalytic converters, the traffic charges imputed as a tax on gas per liter are presented in Table 4.

The variable traffic charge for cars in Sweden is a flat tax on gas, differentiated between leaded and unleaded gas: $0.57 and $0.53, respectively. A value-added tax of 25 percent is charged, based on the gas cost ($0.51) including the traffic charge. The gas price thus is $1.35 and $1.30, respectively (as of December 1990).

The reasons why the traffic charges correspond to the social marginal costs for cars with catalytic converters are, on the one hand, that the Swedish car fleet is rapidly converging to

TABLE 3 EXTERNAL EFFECTS FOR PRIVATE CARS WITH DIFFERENT FUEL CONSUMPTION (CENTS PER 10 km, 1990)

	Catalytic Converter			
	Yes		No	
	Fuel consumption in liters per 10 km			
External Effect	1.0	0.8	1.0	0.8
Road maintenance, traffic surveillance, and congestion	5.4	5.4	5.4	5.4
Traffic accidents	25.9	25.9	25.9	25.9
Pollution	14.1	11.3	44.1	35.4
Total	45.4	42.6	75.4	66.7

TABLE 4 TRAFFIC CHARGES IMPUTED AS A TAX ON GAS

Fuel Consumption (liters per 10 km)	Tax per Liter
With Catalytic Converter	
1.0	$0.45
0.8	$0.53
Without Catalytic Converter	
1.0	$0.75
0.8	$0.83

catalytic converters (2 years after catalytic converters were compulsory, more than 30 percent of all vehicle-kilometers are on the account of cars with catalytic converters); on the other hand, a great number of cars without catalytic converters can use unleaded gas.

The possibilities for differentiating the traffic charges for diesel vehicles are better, as these vehicles pay kilometer (weight-distance) taxes.

Rail Traffic

The National Railway Administration charges SJ and other railway companies according to a principle of two-part tariffs. The variable charges are based on short-run marginal costs for maintenance (wear and tear) of the track, plus external costs. The charges are different for various types of vehicles, locomotives, and tracks. The fixed charges are based on wheel axles for different vehicle-litteras, and driving axles for locomotives. The total payment responsibility for SJ during 1989 is presented in Table 5 in relative figures.

The Swedish rail network is to a large degree electrified. The charges for air pollution emanate from diesel locomotives, for which currently a charge of 4.8 cents per liter of diesel fuel is paid. When the proposed emission charges are adopted, the new charge should be 9.3 cents per liter. This charge excludes charges for the emissions of CO_2. A reason for this is that the competing road traffic (trucks as well as coaches) do not pay for their total external costs at present.

Air Traffic

In addition to the emission charges proposed for air pollution, a charge will also be implemented for noise. For a typical flight, charges are suggested as presented in Table 6.

The environmental charges, determined on the basis of international certification data for civil aircraft engines, account for 35 to 40 percent of the total charges for older aircraft, whereas more modern planes have a share of about 20 percent. [A DC–9–41 will pay $470 out of $1,320 (36 percent), whereas a B737–500 will pay $195 out of a total charge of $1,085 (18 percent).]

As international flights are concerned, the conditions are somewhat different. Swedish authorities are not permitted to charge international traffic for other costs than the management of airports, i.e., external effects are not allowed to be included in the landing charges.

TABLE 5 THE PAYMENT RESPONSIBILITY FOR RAIL TRAFFIC, 1989

Charges	Percent
Variable	
• Traffic safety	20
• Air pollution	2
• Maintenance of tracks	27
Fixed	51
	100

NOTE: The fixed charges cover about 20 percent of the total fixed costs. The basis of payment responsibility for fixed costs (i.e., nonmarginal costs) is formulated ex ante.

TABLE 6 CHARGES PROPOSED FOR DOMESTIC FLIGHTS

Charge	DC–9–41	MD–82
Landing, etc.	$850	$1,030
HC + NO_x	110	100
CO_2	215	160
Noise	150	25

NOTE: (Dollars per flight at Arlanda Airport; distance of 380 km; the cabin factor is 65 percent.)

The Commission on Economic Instruments in Environmental Policy therefore suggests environment-related landing charges, which don't increase the revenues for airports or the state. The revenues will instead be transferred back to the airports to reduce the conventional charges. Before the proposal is implemented, further analyses of the limitations imposed by international agreements, especially within ICAO, may prove necessary.

The principle of the environment-related landing charges is illustrated with two planes, where one is a winner and the other a loser in the proposed system (Table 7).

A DC–9 thus will have to pay an additional landing charge of $115, an increase of 7.7 percent, whereas the charge for an MD-80 will be reduced by $155, a decrease of 7.8 percent. The environmental charge per seat is $6.50 for the DC–9 and $2.80 for the MD-80.

The environmental charge for a DC–10 is net 7.5 percent (the charge per seat is $7.90), whereas the environmental subsidy for a Boeing 676 is net 11.5 percent (the charge per seat is $2.60).

TABLE 7 CHARGES PROPOSED FOR INTERNATIONAL FLIGHTS

Charge	DC–9 ($)	MD–80 ($)
Conventional Charges		
Landing, etc.	1,530	1,970
Proposed Charges		
Landing, etc.	1,530	1,970
HC + NO_x	160	150
CO_2	320	240
Noise	225	40
Reimbursement	–590	–590
Net	1,645	1,815

Navigation

Sweden has almost no inland waterway traffic. The coastal traffic is open to international competition, which limits the possibilities for using environmental charges. However, the Commission on Economic Instruments in Environmental Policy proposes

> a system of charges for shipping, aimed at reducing discharges of sulfur in proximity of Sweden. The proposal is based on the assumption that it is impossible, in the near future, to achieve international agreements to impose direct regulations to reduce sulfur discharges from shipping.

The proposed sulfur charge will affect all shipping using oil with more than 0.5 percent sulfur. The charge is based on the extra cost of using low-sulfur oil, with an upper limit based on a distance of 350 km. It corresponds to a charge of about 10 SEK per kilogram, which can be compared with the general charge of 30 SEK per kilogram.

The proposed system includes both domestic and foreign shipping. Ships with frequent calls at Swedish ports are expected to choose low-sulfur oil, while other ships are expected to pay the charge. A merchant vessel of 20,000 tons will have to pay $1,230 at a maximum.

COST RESPONSIBILITIES FOR AIR POLLUTION

The emissions in Sweden of NO_x, SO_2, HC, and CO_2, weighed together with the proposed price tags, emanate to 45 percent from the transport sector. The proposed charges correspond to the following cost responsibilities for the Swedish transport sector, which are a challenge for the transport policy concerning traffic and price levels of 1990 (Table 8).

The emission costs, being more than 270 times higher for road traffic than for rail traffic, can be compared to passenger traffic being 16 times higher for road traffic (96 400 versus 6 120 million passenger-kilometers for rail) and freight traffic being ⅓ higher for road traffic (25 500 versus 19 100 million ton-kilometers for rail).

The evaluation of emissions has changed drastically during the last decade. This is illustrated for road traffic by the following data (Table 9).

From 1982 to 1990, the cost responsibility for the environmental effects from road traffic has increased by 700 percent, whereas road traffic (in vehicle-kilometers) has increased by only 25 percent.

A more thorough comparison between road and rail indicates how the infrastructure charges are drastically changed when external effects are considered. With the budgetary costs for the state and local communities made comparable by an index of 100, the external effects according to the Transport Policy Act of 1988 increased the costs for road traffic by 150 percent, whereas the costs for rail traffic were increased by only 5 percent. With the proposals presented earlier, the cost difference was even more striking. Road traffic costs increased by 240 percent, whereas rail traffic costs increased by only 8 percent (Table 10).

TABLE 8 COST RESPONSIBILITIES FOR AIR POLLUTION IN THE TRANSPORT SECTOR

Mode	Amount (million)
Road	$2,860
Navigation	460
Aviation	160
Rail	10

TABLE 9 EVALUATION OF AIR POLLUTION FROM ROAD TRAFFIC AT THE PRICE LEVEL OF 1990

Year	Amount (billion)
1979	+
1982 (costs)	$0.4
1987 (explicit)	1.8
1990 (charges)	2.9

SOME COMPLEMENTARY REMARKS

The evaluations and cost calculations of air pollution are important to support measures in obtaining the environmental and ecological objectives of the Swedish transportation policy. They are used in the cost-benefit analyses and investment appraisals of infrastructure measures, as well as in the formal cost responsibilities (excluding maritime traffic) to give incentives for better resource allocation.

However, the cost calculations of air pollution must be interpreted with some care. They should be applied to their full extent for a particular transport mode, only if all external effects, within as well as outside the transport market, are considered appropriately.

So far, the maritime sector is, as an example, excluded from the pricing of air pollution. This can be considered in the cost-benefit analysis for measures where maritime traffic is affected. However, the cost responsibilities for air pollution for the competing transport modes must consider this in a second-best pricing. So far, this has not been especially analyzed.

When more and more of the nonpecuniary external effects are included in the pricing, the matter of how to use the revenues must be considered. From a theoretical point of view, the revenues should be used in the same way as the general taxes, i.e., where the yields and benefits are most needed. Earmarked charges and taxes will suboptimize the use of resources.

In practice, pollution charges are questioned. The Swedish Road Federation, The Hauliers Association, etc., do not ac-

TABLE 10 COMPARISON OF FINANCIAL COSTS AND EXTERNALITY COSTS FOR ROAD AND RAIL TRAFFIC

	Road traffic (%)		Rail traffic (%)	
Cost	1988	1990	1988	1990
Infrastructure	100	100	100	100
Externalities:				
Traffic accidents	91	117	5	7
Air pollution	65	123	0	1

cept nonpecuniary effects to be included in the cost responsibilities. (However, they accept these effects to be included in the cost-benefit analysis of road measures, where the reductions of external effects increase the benefits). Naturally, the politicians take these facts into their considerations.

In practice, there is a risk if charges are used as general revenues. Charges for external effects might be adopted because of the revenue effects.

In the future, there must be a sharper distinction between the incentive effects and the revenue effects. Also there must be a stronger connection between the environmental and ecological objectives and the use of the revenues from pollution charges. If this is done, by financing ecology funds, etc., society is avoiding dependency on the revenues mentioned.

The purpose of the charge for external effects must be the incentive effects, not the revenue effects (except for pecuniary costs, of course). Reduced emissions imply welfare gains, but also a reduction of revenue. If the charging is extremely successful, there will be no emissions. Accordingly, there will be no revenue. This is, of course, not a problem if the revenues have been used for measures undertaken to obtain the objectives for which the charges were formed. It would be regarded as a problem if the revenues have been used for general purposes. The need of money remains, but the source where it came from is empty.

The use of pollution charges proposed for international flights is a good example in which the incentive effects are implemented without creating more revenues for the state. Of course, air pollution charges increase the costs for those transport modes that pollute the most. However, with tax neutrality, they will pay only for those measures undertaken, not for the other nonpecuniary effects. This can be expected to increase the acceptance for pollution charges.

REFERENCES

1. L. Hansson. The Swedish Approach to Multi-Modal Transportation Planning. AAR 1990 Intermodal Policy Conference, Key Largo, Fla., April 6–9, 1990.
2. Trafikpolitiken inför 90-talet; Proposition 1987/88:50 (Government Bill on Transport Policy for the 1990s).
3. W. J. Baumol and W. E. Oates. The Use of Standards and Prices for Protection of the Environment. *The Swedish Journal of Economics*, Vol. 73, No. 1, 1971, pp. 42–54.
4. Sweden's Environmental Charge Commission. Ekonomiska styrmedel i miljopolitiken (Economic Instruments in Environmental Policy). SOU 1989:83.
5. Sweden's Environmental Charge Commission. Sätt värde på miljön. Miljöavgifter och andra ekonomiska styrmedel (Setting a Value on the Environment. Pollution Charges and Other Economic Instruments). SOU 1990:59.
6. L. Hansson and I. Leksell. Vägtrafikens kostnadsansvar (*The Cost Responsibility for Road Traffic*). The Swedish Society for Nature Conservation, Stockholm, Sweden, 1990.

Publication of this paper sponsored by Committee on International Activities.

Transportation and Urban Air Pollution Policies for Developed and Developing Countries

ALAN J. KRUPNICK

Improvements in urban air quality remain elusive in large cities throughout the world, including those in the United States where efforts have continued over 20 years to reduce emissions from vehicles and other sources. Although technological advances in gasoline and diesel-fueled vehicles and in their emission control systems and fuel have resulted in impressive emissions reductions, the combination of more vehicle-miles traveled and other factors offsets these improvements. Tightening further vehicle emissions standards in developed countries would be costly and—judging by the continuing presence of ambient ozone in urban areas—ineffective policy. Rather, efforts to target high-emissions vehicles and to impose fees on fuels, vehicles, or emissions may prove to be more cost-effective. For developing countries, the removal of lead from motor fuels and imposition of economic incentives to control transportation are potentially cost-effective strategies.

The further tightening of vehicle emissions standards in the U.S. Clean Air Act Amendments of 1990, the worldwide effort to develop alternate-fueled vehicles, and the rapid growth of vehicle use in developing countries provide ample evidence of the pivotal role to be played by vehicles in strategies for improving urban air quality, enhancing energy security, and addressing global warming concerns. Yet, technological, economic, and environmental trade-offs complicate the fashioning of strategies to meet any one of these goals, let alone all of them at once.

Of all these issues, that of urban air quality is particularly frustrating. Technological improvements in gasoline- and diesel-fueled vehicles, their emissions control systems, and the fuel they use have resulted in impressive emissions reductions. For instance, emissions of volatile organic compounds (VOCs) (precursors of ambient ozone) fell 33 percent over the 1980 to 1989 period (*1*). However, these reductions have not improved ambient ozone concentrations in U.S. cities, with some 66 million people still living in 96 urban areas that violate the U.S. ambient ozone standard (*2*). The growth in vehicle-miles traveled (VMTs) by 39 percent over the same period (*2*) is one of the major culprits.

Outside of the United States the story is similar (where data permit a story to be told). Many of the largest urban areas in OECD (Organization for Economic Cooperation and Development) countries and South America likewise record concentrations of CO and NO_x (another ozone precursor) above World Health Organization (WHO) air quality guidelines (*3*). Some of the worst air pollution in the world is in major cities in developing countries, such as Mexico City, and, with growing incomes, vehicle ownership and VMTs will skyrocket, leading to more severe air quality problems.

These problems are being approached on a variety of fronts, but with less regard to economic principles than is warranted. In any country, but particularly developing countries, scarce resources should be allocated to maximizing net social benefits. By designing policies that are sensitive to both the costs of obtaining emissions reductions and the benefits derived, this goal can be approached. As these costs and benefits may vary with level of development, among other factors, the efficient set of policies may differ according to stage of development as well.

Some of the benefits, costs, and policy issues associated with reducing urban vehicular air pollution in developed and developing countries are addressed. The benefits of controlling each type of emissions are examined, as well as the relative contributions to urban emissions made by vehicles. Then, a variety of vehicle emissions control policies are examined for their ability to deliver cost-effective emissions reductions. These policies include developing alternate fuels and vehicles, targeting high-polluting vehicles, using transportation controls, and introducing broader economic incentives, such as emissions fees. Focus on benefits and costs is used to guide and organize the discussion, but a full-blown benefit-cost analysis is not presented. The role of vehicles in producing greenhouse gases and the complex interplay between policies addressing fuel economy and those addressing air pollutants are ignored.

BENEFITS

All fossil-fuel burning vehicles emit the conventional pollutants: sulfur dioxide (SO_2), particulates, VOCs, nitrogen oxides (NO_x), carbon dioxide (CO_2) (and small quantities of other greenhouse gases), and carbon monoxide (CO). The VOCs and NO_2 are precursors to ambient ozone (O_3). Some VOCs are carcinogenic, such as benzene; some particulates are also carcinogenic, such as benzo(a)pyrene and other polycyclic aromatic hydrocarbons (PAHs). Significant quantities of lead are also present as additives in leaded gasoline.

Each type of emission from a vehicle has its own dose-response functions that relate exposure to a pollutant to the amount of injury. Because of the different proportion of emissions of various types in gasoline and diesel exhausts, policies that favor one fuel over another will affect the mix of urban

Quality of the Environment Division, Resources for the Future, 1616 P Street, N.W., Washington, D.C. 20036.

emissions and the damages caused, even if VMTs remain constant. To be informed about these trade-offs, information on the health and other effects of the various types of pollutants is needed. Because a policy can cause certain types of health and other effects to increase while others diminish, a comparison of effects requires using a common metric, such as dollar values. Several available estimates of the effects and benefits of controlling various air pollutants emitted by vehicles are described in the following paragraphs. For comparability, they are converted to a per-mile basis.

The effects are limited to those on human health. It is the health damages associated with ozone formed from its precursors—VOCs and NO_x—as well as health effects related to carbon monoxide emissions and, in developing countries, particulate emissions that are fueling concern over vehicular emissions.

O_3

The effects of ozone on health are the most well understood of any conventional pollutant-health interaction. The short-term effects of ozone [respiratory symptoms after 2-hr exposure, daily symptoms, restrictive activity days (RADs), and asthma attacks] are firmly established and quantifiable (within fairly narrow uncertainty bounds) using both clinical and epidemiological studies (*4*). There is currently much debate over long-term effects of ozone on lung tissue and the probability of developing chronic respiratory disease.

The acute health benefits of ozone reductions are calculated as the acute health effects reductions associated with reduced ozone exposure multiplied by the value people place on avoiding such effects. Acute health benefits for a 35 percent reduction in VOC emissions in the northeast United States have been estimated to range between \$100 and \$2,000 per ton with a best estimate of around \$500 per ton (*5*). This contrasts with cost benchmarks of \$10,000 per ton and more (*6*). On a per-mile basis, the benefits per ton translate [at 1.75 grams per mile (g/mi), EPA's estimate of VOC emissions from current gasoline vehicles (*7*)] to benefits per 10,000 mi of \$0.40 to \$8.00, with a best estimate of \$2.00. Said another way, with 73 million people living in this area, benefits range from about \$2 to \$33 annually per person.

CO

CO at ambient levels may increase the probability of experiencing angina for an estimated 5 to 7 million people in the United States who are at risk (a prevalence rate of 3 percent). However, in spite of much effort expended to identify health effects of this pollutant at ambient levels, too little information is available to derive exposure-response functions (*8*).

NO_x

The direct effects of NO_x emissions appear to be of little health significance in the aggregate. The acute health effects from ambient NO_2 exposure appear to be minor (*9*). Further, EPA says that there is presently no reliable scientific evidence of adverse effects in humans for long-term NO_2 exposure at ambient levels (*10*). The fact that NO_2 is 17 times less reactive an oxidant than ozone may account for the difficulties in finding effects.

Particulates and SO_2

In contrast, there is more concern about, but not a good understanding of, the effects of SO_2 and particulates on health; but concern is muted by the fact that few areas, at least in the United States, are violating current air quality standards. The direct effects of SO_2 on health in the United States have been strongly linked to exercising asthmatics exposed for brief periods (1 hr) to SO_2 concentration spikes that are occasionally experienced in the United States. In general, SO_2 and particulates, and their combination, have been linked to increased risks of acute and chronic morbidity and mortality (*11*).

On the basis of a recent epidemiological study (*12*), Portney et al. (*13*) found mortality risk reductions in Los Angeles (with 12 million people) ranging from 0 to 4,000 statistical lives saved annually as a result of the Los Angeles air quality plan (*10*) to reduce sulfate concentrations by about 50 percent (in part with controls on vehicles and diesel fuel). Multiplying this figure by an estimate for the value of a statistical life (from the economics literature)—\$1 million—yields benefits ranging from \$0 to \$4 billion. However, as only about 4,000 people died of respiratory causes in the Los Angeles area in 1988 (personal communication, Los Angeles Department of Public Health), the upper range estimate cannot be viewed as credible. In any event, benefits range from \$0 to \$300 per person, with a best estimate of about \$150. Assuming that the benefits are realized entirely through particulate reductions (projected to be 1,331 tons per day by the South Coast Authorities) and using the diesel particulate estimates in grams per mile presented in Table 1, a best estimate of mortality benefits is about \$29 per 10,000 heavy-duty diesel (HDD) miles displaced [or \$5 per 10,000 light-duty diesel (LDD) miles displaced]. Adding morbidity benefits of \$700 million estimated by the South Coast (*14*) yields estimates of benefits ranging from \$0 to \$78 per 10,000 HDD miles displaced, with a best estimate of \$39. For LDDs, the range is \$0 to \$14 with a best estimate of \$7. These benefit estimates are substantially larger than for ozone.

Lead (Pb)

Finally, consider one of the most important unconventional vehicular pollutants—Pb. The relationship of changes in Pb in gasoline to changes in blood lead levels in the U.S. population is remarkably close (*8*). Thus, unlike other pollutants, there is little uncertainty about the effects of reducing Pb in gasoline on Pb exposure. There is also general agreement that children with high-typical blood Pb levels suffer learning disabilities and recent studies link ambient Pb exposure to high blood pressure. With this link, Pb becomes a risk factor for hypertension, heart attacks, and stroke, particularly in men.

EPA (*8*) estimated the physical and monetary benefits of the Pb phase-down regulations [which reduced Pb in gasoline

TABLE 1 AVERAGE TAILPIPE EMISSIONS (g/mi) (FEDERAL TEST PROCEDURE CYCLE)

	POLLUTANTS								
				HC		Particulates			
VEHICLE TYPE	CO	NO_x	SO_2[f]	THC	Benzene	Total	BaP	SO_4[c]	Pb
Diesel								0.037	
Heavy-Duty	10	28	1.6			3.2[c] (1.4)[d]	54		
Light-Duty	3	1	0.53	0.23[a,b]	0.02	0.6[c] (0.4)[d]	13		0.0[i]
Gasoline								0.006	
No Catalyst	15	4	0.1	5.4[a,b]	0.31	0.1	20		0.020
Catalyst	5	2	0.07	1.8[b]	0.06	0.02	0.4		
U.S. Late Model	8[h]	1.39[g]		.72[g]					
U.S. VEHICLE TAILPIPE EMISSIONS STANDARDS									
Light-Duty Diesel Trucks	10.0	1.2		0.8		0.26			
Light-Duty Gasoline Passenger Vehicles & Trucks	3.4	1.0		0.41		0.6(<1987) 0.2(≥1987)			

a. 1975-82 models.
b. Diesels emit heavier alkanes, which have greater ozone-forming potential.
c. Mostly fine particulates (≤ 1 um).
d. Field experiments in a tunnel.
e. Accounts for nearly 20% of particulate matter mass, mostly as H_2SO_4.
f. Diesel sulfur content of 0.3% (U.S.).
g. EPA (1989c). 50,000 mile in-use.
h. CARB (1989). 50,000 mile in-use.
i. The source mistakenly had 0.019 g/m. Lead is not added to diesel fuel. Unlike for gasoline, the octane enhancement provided by lead makes diesel engines perform worse.

Source: EPA (1990).

from 1.1 to 0.1 grams per leaded gallon (gplg)] for the above endpoints, concluding that the costs of reducing lead (far less than 1 cent per gallon) were far outweighed by a lower bound estimate of the benefits. EPA predicted benefits in avoided medical costs and special education classes of $600 million annually for children with Pb levels brought below 25 μg/dl and benefits of $5.9 billion for men between 40 and 59 years old in reduced risk of premature death from heart attack, reduced incidence of hypertension, and reduced incidence of stroke and nonfatal heart attack. With overall benefits of $6.5 billion, 28.8 billion miles traveled on leaded gas in 1986, and assuming average passenger vehicle fuel economy of 18 miles per gallon in 1986, benefits average about $12 per 10,000 mi, again higher than for controlling ambient ozone, and also exceeding that for LDDs.

EMISSIONS SHARES

In fashioning efficient emissions control policies, knowledge of the relative share of total emissions created by particular types of vehicles (gasoline versus diesel, automobile versus scooters, etc.) is important. If large reductions in emissions are desired, emissions from the vehicle types contributing the largest shares need to be controlled.

The contribution of vehicles to air pollution in an urban area depends on the use of vehicles relative to activity levels of other sources of emissions, on the mix of vehicles, the presence and operation of pollution controls on the vehicles, and the quality of the fuel being burned.

In the United States and most developed countries, there are credible estimates of the contribution of vehicles to total emissions by pollutant. These data for the United States (*15*) indicate that vehicles contribute 45 percent of the VOCs in ozone nonattainment areas, although this share varies across U.S. urban areas from a low of 30 percent to a high, in Los Angeles, of about 66 percent. Vehicular NO_x emissions were only about 30 percent of the total nationwide in 1989 (*1*), but are higher in urban areas. CO emissions are dominated by vehicles. Comparing gasoline and diesel vehicle emissions, gasoline vehicles dominate in all categories of emissions except SO_2.

Yet, controlling diesel emissions can still be productive on a per-mile basis. Table 1 presents average U.S. tailpipe emissions coefficients for HDD and LDD trucks and for passenger and light-duty gasoline vehicles with and without catalytic converters. The most comparable figures are between gasoline vehicles with catalytic converters and LDDs. Of the conventional pollutants, sulfur oxides are over five times larger, and total particulates are 30 times higher for the LDDs, with sulfates about six times larger (mostly in the form of sulfuric acid), whereas CO, NO_x, and HCs (a slightly broader measure of hydrocarbons than VOCs) are substantially lower for LDDs (*16*).

The comparison for particulates is even more unfavorable than it looks, because nearly all of the diesel particulates are fine (and therefore can penetrate deeply into the lung) and some carcinogens such as benzo(a)pyrene, are also far more heavily represented in diesel emissions than in gasoline emissions. According to WHO (*3*), diesel engines generate 10 times more respirable particulates than gasoline engines per kilometer traveled.

The comparison between gasoline vehicles (in this case automobiles) and diesel vehicles (i.e., buses) can be further sharpened by comparing emissions per passenger-mile. Assuming that bus emissions are the same as those for heavy-duty diesel in Table 1, an average of 20 passengers per mile on a bus and one person per mile in an automobile would equate NO_x emissions per passenger-mile between a bus and a late model automobile. For SO_2, 23 passengers, and for particulates, 160 passengers. Thus, in the United States, where transit buses carry only an average of 10 people per mile (*17*), buses would still be more polluting.

In developing countries, data limitations mean that the emissions share of vehicles needs to be inferred. Certainly, transportation is a major activity in developing countries, given that it claims 25 percent of their energy use (except India and China) (*18*). Dependence on trucks and buses is far greater than in developed countries—more than 50 percent of total energy demand comes from trucks and a large additional share from buses (*19*). In terms of passenger trips in motorized vehicles, about 50 percent occur in buses (*20*). Two- and three-wheeled two-stroke vehicles, which can produce high levels of pollution in spite of their fuel economy, also make up a larger share of urban road transport in developing than in developed countries. They are particularly prevalent in cities such as Singapore, which has one motorcycle for every 3.5 automobiles. Also, vehicles are responsible for much of the CO and ozone problems experienced in the largest urban areas of developing countries.

Whatever estimates are available on emissions from vehicles in developing countries appear to support the emphasis on diesel emissions and two- and three-wheel vehicles. In Indian cities, two- and three-wheeled vehicles were estimated to be responsible for 85 percent of the CO and 35 to 65 percent of the hydrocarbons (*21*). Diesel trucks and buses were estimated to have created 90 percent of the NO_x. A recent study (*20*) estimates that for nine developing countries buses and trucks (as separate categories) are each a larger source of total emissions of all types than automobiles, except for CO in Thailand and Tunisia. In three Latin American countries, total emissions of CO and HC were estimated to be larger for automobiles, whereas buses and trucks emit the largest share of SO_x. Particulates and NO_x were not found to exhibit any clear trend.

In many urban areas, air pollution from vehicles appears to be of lesser current concern because vehicle ownership and VMTs are so low. According to the United Nations (*22*), automobiles per 1,000 urban dwellers are nearly always less than 100 and generally less than 50 in developing countries, while ranging from 200 to 500 per 1,000 in developed countries. Yet, of six developing countries reporting VMTs in 1977 and 1987, all experienced increases, ranging from 23 to 250 percent (versus 33 percent in the United States). In Indonesia, vehicle ownership tripled from 1970 to 1981, in Brazil and Lagos it more than doubled, in Nigeria it increased five times. With a doubling of cities with over 4 million people by 2000 (*3*), vehicle use is clearly going to continue its rapid increase.

Emissions, at least on a per-mile basis, are already likely to be higher for the average vehicle in developing countries than in developed countries. This is so because the former tends to be older and less well maintained and less likely to have any emissions control equipment. For example, in Mexico City, 70 percent of gasoline vehicles and 85 percent of diesel vehicles participating in a voluntary inspection and maintenance program fail the emissions tests. Less than half the vehicles in Mexico have pollution control devices of any kind and none have state-of-the-art systems. However, on a passenger-mile basis, bus emissions (as well as automobile emissions) would be far lower than their U.S. counterparts because of the high load factors of buses (and automobiles) in developing country cities. More data are needed to determine which mode has the net advantage in emissions per passenger-mile and what the cost effectiveness would be of reducing emissions for each mode.

There is little ambiguity about the large share of urban Pb emissions from vehicles and the tight relationship between reducing this source of Pb and blood Pb levels. Furthermore, the benefits of avoiding exposure to Pb emissions have been estimated to be large and relatively certain compared to some other pollutants and relative to the costs of reducing or eliminating Pb additives. Thus, phasing out of leaded gasoline deserves a high priority. Beyond this, firm conclusions are harder to make.

In particular, difficult choices loom for developed countries. There, passenger vehicles contribute a large proportion of urban VOCs, NO_x and CO, thereby contributing importantly to violations of ambient O_3 and CO standards. Diesel vehicles, which are only a small part of the vehicle fleet in developed countries, contribute (at least in the United States) most of the vehicular particulates and SO_2 on a per-mile basis, but not in the aggregate. Most urban areas of developed countries are not violating particulate and SO_2 standards. At the same time, the health effects are not well understood for CO and, for O_3, those that are well understood are relatively unimportant and not highly valued in the aggregate when compared to heath effects potentially caused by particulates and SO_2. Should the dominant vehicle type—the gasoline-fueled automobile—come in for most of the control even though its emissions are perhaps more benign than those of diesels?

For all but the largest urban areas of developing countries, vehicular emissions are probably less of a problem and, at least for now, the choices are less clouded. Vehicles in these areas are likely to emit larger amounts of all types of pollutants per mile traveled but, with so many fewer miles traveled and such poor controls on other urban sources of pollutants, particularly those emitting particulates, control of vehicle emissions may not be a major issue. Future growth in urban incomes will change this situation in the future, however, bringing the air quality of many more urban areas down to that of the large primary cities in developing countries. The modal split favoring buses, trucks, and two-stroke scooters over automobiles in developing countries (with the reverse true in developed countries), the poor maintenance and advancing age of the bus fleet, and its associated high levels of emissions per mile, the use of high-sulfur fuel by buses and trucks [in the United States, sulfur content is 0.3 percent, to be reduced to 0.05 percent, whereas in the Philippines it is 1 percent (*20*)], and the potentially high value of benefits from reductions in fine particulates implies that developing countries might be better off focusing on control of these diesel sources rather than on control of automobiles.

POLICY IMPLICATIONS

As long as national air quality standards are generally set on the basis of protecting health rather than balancing costs and benefits, urban air quality policies will be inefficient. Nevertheless, there is still a broad scope for the application of economic concepts in the design of cost-effective pollution control policies involving vehicular emissions. In this section, four types of policies are reviewed for their likely cost effectiveness with respect to one another and in developed versus developing countries.

Vehicle-Fuel Technologies

The primary strategy for reducing vehicular emissions in developed countries has been that of setting tailpipe emissions standards on gasoline vehicles. This strategy has resulted in lower tailpipe emissions primarily through the technological advance of the catalytic converter. Indeed, some more recent advances have enabled some late model year vehicles to meet the very stringent 50,000-mi standards set in the 1990 U.S. Clean Air Act Amendents for 1994 (*23*).

Germany, the Netherlands, Norway, and Sweden have, until recently, taken another route. Each taxed clean cars less than others (or reduced annual vehicle fees) but clean was defined mainly as having a catalytic converter, not in terms of specific types of emissions. These policies were slated for removal in 1989 as United States–style EC-wide vehicle emissions standards will make such differentiation unnecessary (*24*).

Developing countries generally do not have emissions standards, with vehicles emitting at levels consistent with the technology for emissions reductions embodied in them or, if the technology is not working, at uncontrolled levels. Pre-1969 vehicles purchased from any manufacturer would emit at uncontrolled levels.

Cost-effectiveness analyses of the progressive tightening of the U.S. emissions standards suggest that some reductions in vehicle emissions are cost effective (i.e., in comparison to the costs of other emissions reduction options), but that these costs increase rapidly with more stringent standards (*25*). Absent technological breakthroughs in emissions controls or reformulated gasoline, the newest round of emissions reductions in the United States is likely to be even less cost effective. However, for developing countries that have never set requirements on vehicular emissions, relatively lenient emissions standards may prove cost effective.

The mounting frustration with reducing the external effects (environmental and energy security) of gasoline vehicles has led to increasing attention being paid to a set of technologies involving new fuels and vehicles. The new U.S. Clean Air Act, for instance, requires gasoline to be reformulated to have lower ozone-forming potential, more oxygen, and fewer carcinogens, and diesel fuel to have no more than 0.05 percent sulfur content. Although no other alternate-fueled vehicles are mentioned in the Act, very stringent standards in areas with serious air quality problems open the door to them. Alternate-fueled vehicles being considered include methanol, ethanol, compressed natural gas (CNG) (and LNG), and electric vehicles (EVs) (as well as hybrids), with hydrogen and solar vehicles further off in time.

Numerous reviews of the advantages and disadvantages of these options (*26*) have failed to identify a clear frontrunner on grounds of costs and emissions reductions. One fuel-vehicle mix being particularly touted in the U.S.—methanol—has been estimated (*6*) to be a costly option in terms of the reductions in ozone-forming potential that it delivers relative to improved and advanced gasoline vehicles—with a best estimate of costs for M100 vehicles (vehicles that burn 100 percent methanol) of $51,000 per ton of VOCs displaced.

The main drawback of relying on advances in technology for emissions reductions—and this applies at least as much to alternate fuel vehicles as to gasoline vehicles—is that its effectiveness is dependent on turnover in the vehicle stock. This has two implications. First, this strategy is for the long run, as such turnover takes much time; for developing countries, the wait will be even longer because vehicles are held for longer periods than in developed countries. Second, if the new technologies are more expensive than the old and have characteristics that make the vehicles less attractive to consumers, the rate of turnover is likely to slow. Crandall et al. (*25*) found (using an earlier version of EPA's MOBILE4 model) that higher new vehicle prices as a result of the 1981 tightening of emissions standards had the effect of increasing VOCs and CO emissions over that of a scenario where standards were not tightened, and that this perverse effect lasted for 5 years. Overall, they found that in 1982 the aging of the vehicle stock from its 1967 to 1978 average resulted in VOCs, CO, and NO_x emissions being 26, 23, and 11 percent larger, respectively, than they would have been had no aging occurred.

Given the dependence of the success of most of the new fuel-vehicle strategies on consumer acceptance of new vehicle technologies, strategies involving reformulated gasoline (both leaded and unleaded) and reduced sulfur in diesel fuel have more appeal because they can be applied at once to the entire vehicle stock. They also may be reasonably low cost. Formulation changes to reduce gasoline volatility are the most cost-effective option identified by the Office of Technology Assessment ($500 per ton of VOCs reduced) (*15*); ARCO's EC–1 formulation for precatalytic vehicles has a cost effectiveness in reducing hydrocarbons of about $4300/ton, if a 20 percent emissions reduction to a 1979 model year can be obtained for $0.02 per gallon (*27*). Another cost advantage is that reformulated gasoline can be used selectively in areas that have the worst air quality.

Taking the perspective of developing countries, development of alternate-fueled vehicles is generally outside of their ability to influence (the potential market being too small). In addition, advanced emissions control technologies generally require advanced equipment and specialized knowledge to service. Given the already poor state of maintenance of vehicles in developing countries, only specially designed low-maintenance, reliable, and low-cost technologies are likely to make inroads into developing country vehicle markets. Until developing countries switch to unleaded fuels (leaded fuel poisons the catalytic converter), even current catalyst technologies will be beyond their reach.

As seen for developed countries, reformulated gasoline and low-sulfur diesel fuel strategies are promising innovations for developing countries. Given the continued use of leaded fuel in the latter countries, ARCO's EC–1 leaded gasoline is particularly promising if the 2 cents/gal cost differential in the United States would apply to these countries as well. Given the heavy dependence of Asian countries on highly polluting two-stroke motor scooters, successful penetration of gasoline-powered two- and three-wheeled motorized vehicles would improve the air.

Targeting High-Polluting Vehicles

Even in the United States, there has been little effort to identify and mitigate emissions from high-polluting vehicles—the 10 percent of the U.S. vehicle fleet that, according to some sources, is responsible for half the vehicle emissions (*28*). Until recently, the consensus was that the oldest vehicles were the grossest polluters. In this case, fleet turnover would eventually take care of the problem. Older vehicles, because they embody older emissions control technology, do produce more emissions per mile than newer vehicles, on average. A pre-1981 passenger vehicle produces three times the HCs, twice the NO_x, and eight times the CO as a new vehicle (*27*). In the United States, pre-1981 passenger vehicles account for 71 percent of all vehicular HC emissions.

Unfortunately, matters are more complicated than this. Recent research indicates that there is wide variation in emissions performance of vehicles within any given model year. ARCO found that the most polluting of 16 pre-1981 vehicles produced nearly 10 times the HCs of the least polluting. General Motors has also found wide vehicle-to-vehicle variation in emissions, with the distribution of emissions roughly log-normal: at 50,000 mi, 10 percent of 1986 model year vehicles violate the HC standard, 27 perΣN1i =0 ΣN2j =0ijz⁻i1z⁻j2j cent violate the CO standard, and 7 percent violate the NO_x standard (*29*).

Inspection and maintenance (I&M) programs are the primary approach used to address this problem in the United States and other developed countries, but are rarely used in developing countries. However, these programs (at least in the United States) have some major drawbacks: (a) they test cars while they are idle and warm, rather than running or cold, when most of the emissions are emitted; (b) some do not have strict enough penalties for failure to correct violations or even have the car inspected; (c) pass rates are very high, meaning that administrative costs are high to catch the few violators; and (d) those that fail are not required to spend more than a set limit on repairs. This waiver feature leaves vehicles with serious pollution problems still on the road. In addition, (e) very old vehicles are exempted from the program.

In order to better target high-emitting vehicles, Stedman (*28*) recommends using infrared sensors placed on highways to read CO emissions (which are highly correlated with HC emissions) of passing vehicles, take photographs of heavy polluters, and therefore permit officials to require that these vehicles be brought to inspection stations. Unfortunately, the technology is still in the experimental stage.

Programs to identify and prematurely retire, maintain, or relocate high-polluting vehicles, can be potentially well-targeted, cost-effective emissions reduction strategies. The American Petroleum Institute (API) (*27*) considered a variety of alternatives to address the high-polluting vehicle problem, including differential registration fees and vehicle retirement programs. Most developed countries charge registration fees that are constant for all vehicles. Japan, for instance, charges from $250 to $500 per vehicle, whereas the U.S. fees range from $15 to $40. Higher fees, even if undifferentiated by model year or emissions, would tend to encourage scrapping lower-valued vehicles, although the extent of this effect is unknown. Higher fees for older vehicles, while better targeted for pollution reductions, would likely be highly regressive, i.e., politically unpalatable. Charging higher fees for more-polluting vehicles, irrespective of age, is really an emissions tax, discussed in the following paragraphs.

Vehicle retirement programs have begun in California. The most interesting involves the purchase of pre-1971 vehicles by Unocal Corporation for $700 each, for what it hopes will count as emissions credit against emissions from its industrial operations. This privatization of pollution reductions could deliver large cost-savings to the company, estimated to be about $161,000 per ton of VOCs (*30*). The issue for society is whether the vehicles purchased would be scrapped anyway, i.e., whether the emissions reductions are real. Also as currently structured the actual emissions of the vehicle are not taken into account. By basing vehicle purchase prices on emission tests results, emissions reductions could be obtained more cost effectively. Finally, some further screening would help improve efficiency by ensuring that the purchase price did not exceed the price of repairs. API suggests that only vehicles failing the I&M tests and that cannot be repaired within the waiver limits be eligible for purchase.

While I&M programs are sorely needed in developing countries, some observers are quite pessimistic that they can be implemented and run efficiently (*20*). At the same time, with a much higher percentage of the vehicle stock made up of poorly maintained and older vehicles, there is little reason to be concerned about targeting. As noted earlier, cleaner gasoline and low-maintenance vehicles constitute a more reliable, but not necessarily less expensive, option for emissions control.

Transport Controls

Controls on the flow of traffic, e.g., alternative drive-days, no-vehicle zones in city centers, high parking fees, HOV lanes, etc., can have a direct effect on the environment if total VMTs fall or trips are switched to less polluting modes of transport, and an indirect effect if congestion is reduced.

At lower speeds, vehicle emissions per unit distance are higher for most (though not all) pollutants. Hydrocarbon (HC) and carbon monoxide (CO) emissions are twice as large at 15 mph as they are at 30 mph (although NO_x emissions are halved) for automobiles in the United States (*20*). These effects could be particularly important for pollutants that do not disperse far, such as CO.

Both developed and developing countries are increasingly experimenting with transportation controls as a means of improving traffic flow and reducing emissions. European countries have been in the forefront of applying controls, such as restricting traffic in central business districts (CBDs).

Transportation controls can reduce emissions efficiently in general and in developing countries. First, transportation controls have the advantage of being local in nature, and can therefore be fine-tuned, at least in theory, to local conditions. Second, such controls generally have low expenditures associated with them, although nonmonetary costs, in terms of (say) inconvenience and longer commutes, may be inferred to be large, given the resistance to efforts at changing driving and overall commuting behavior.

Third, transportation controls may be more effective in and attractive to developing countries than developed countries. Transportation controls have not been particularly successful in developed countries, in part because preferences for automobile travel, urban design and infrastructure, and transport choices are largely fixed. These features are far less fixed in fast-growing areas of developing countries. If the value of time and comfort is lower in developing countries, some of the nonmonetary costs of transport controls may be lower as well.

Some developing countries have already instituted transport control programs using economic incentives. Congestion tolls have been tried. For example, a Mexico City policy whereby admittance to the CBD on a given weekday depends on one's license plate number, appears to have had some effect in reducing traffic and smog levels in the city. Singapore's Area License Scheme, which required purchase of a sticker to enter the city during the morning rush-hour (current price = \$2.60/day) and was supplemented by parking fees and a park-and-ride service has also been judged a qualified success: a 75 percent decrease in vehicles entering the area during the morning rush hour, a 20 percent increase in bus commuters, a doubling of car pools (*31*), and substantial reductions in downtown air pollution (*32*).

In effect, the sticker policy confers a property right to car owners, or more precisely it creates a property right in what was previously a common property resource. However, the right is not transferable because it is determined by the license plate. An improvement would be to issue certificates or stickers for travel on a given day, perhaps on the basis of license plate number. A wrinkle would be to allow these certificates to be traded or sold; the ensuing market would help ensure that stickers would flow toward those with the greatest willingness to pay to enter the city.

An experiment in Hong Kong (*33*) exhibited the potential for moving towards congestion tolls of a more optimal nature, although this is more realistic for developed countries. By fitting cars with license plates capable of being scanned by computers set up at key arteries within the city, vehicles were to be charged roughly by the amount of driving they did. The plan was ultimately rejected because of concern over what would happen to the collected revenues. However, advances in such technologies and their merger with technologies being tried out in California to measure CO emissions from vehicles waiting to enter Los Angeles freeways could make it possible to charge vehicles for their contribution to congestion and pollution.

Fourth, although some types of transportation controls may be attractive to developing countries, expanding public transit—a long-standing favorite urban transport strategy—is not likely to be one of them. Because of the large share of commuting by bus and the chronic undercapacity of this system in many cities, scope for increasing use of buses is limited, unless capacity is increased, an expensive option. One approach might be to encourage privatization of bus service, an approach judged to have been successful in Calcutta, Mexico City, and Bogotá (*34*). Another approach is to make bus companies, whether private or public, more financially sound by permitting their subsidized fares to increase. An ancillary effect of higher fares might be to improve maintenance of bus engines. The effect of improved maintenance on reducing emissions might be more than enough to offset emissions from any increase in bus capacity.

Fifth, beyond public transportation subsidies, existing perverse incentives on use of vehicles could be reduced. Both developed and developing countries subsidize automobiles and truck use rather than tax or otherwise restrict their use on environmental grounds. For instance, in many cities, parking costs only about 50 cents per day; trucks pay road taxes that are far less than proportionate to the damage they do to roads (*35*); and diesel fuel is often subsidized (in India it is about half the gasoline price) as a way of promoting economic development. Such subsidies may distort transportation choices. In the case of diesel fuel, the subsidy may artificially disadvantage barge and rail carriers.

Finally, the earlier analysis suggests that any policies that encourage modal shifts away from gasoline vehicles to diesel vehicles may involve complex health trade-offs—increasing (highly uncertain) mortality effects from fine particulates present in diesel emissions while reducing (much more certain but less severe) respiratory distress from automobile-related ambient ozone exposures. Analysis of the relative costs and health benefits of reducing pollution from diesel versus gasoline engines is needed to better focus these policymaking efforts.

Emissions Fees

In contrast to the prominent place occupied by what might be called the "mainline" economic incentive approaches—emissions fees, energy taxes, and tradable permits—in the debate over how best to reduce CO_2 emissions (*36*) and acid rain precursors (resolved in the new U.S. Clean Air Act in favor of tradable permits), little discussion of these tools can be found in the debate over policy design for improving urban air. This is unfortunate as such tools, at least in developed countries with the administrative and legal infrastructure to use them, can help bring about low-cost attainment of air quality goals.

For instance, a set of emissions fees that varies by fuel-vehicle type depending on the extent of the environmental externalities created by use would go a long way to encourage (a) reduction of VMTs, (b) mode switching to more environmentally benign forms of transport, (c) research and development of fuels and vehicles that result in low emissions, and (d) more rapid turnover of high-polluting or, at least, older vehicles (depending on how the fee system was structured). Fees levied on any specific fuel, vehicle, or characteristic (such as fuel economy or carbon content) cannot deliver all of these benefits and may distort choices and outcomes away from being socially beneficial. Taxes on oil only, for instance, may reduce the attractiveness of gasoline vehicles but, by not accounting for the polluting characteristics of all fuels, may not give enough incentive to develop and market the fuels that have the lowest social cost (private cost plus external cost).

One further benefit of emissions taxes can be obtained if they are used to replace value-added, excise, or other types of product-based taxes. These taxes distort market prices and associated consumer choices (i.e., consumption versus leisure choices) and, therefore, result in social losses, what economists call "deadweight losses of taxation" (*37*). By replacing such taxes with emissions fees, the full social costs of these products is reflected in price and the deadweight losses are removed. Sweden's plans for replacing its value-added taxes on energy with revenue-neutral environmental taxes are in this spirit (*38*).

CONCLUSION

Tailpipe emissions standards on passenger vehicles have resulted in impressive reductions in VOC, CO, and NO_x emissions per mile in the United States and will doubtlessly do so as other developed countries follow the U.S. approach. Nevertheless, these reductions have come at a high cost and, more important, have failed to significantly improve urban air quality, in part because of increases in VMT. At the same time, by raising new vehicle prices, the mandated standards have slowed vehicle turnover, keeping dirtier vehicles on the road longer, and delaying penetration of the newer, far cleaner models. However, an undeniably wise national-level strategy, based on benefit-cost considerations, has been to phase out leaded gasoline and to reduce sulfur in diesel fuel.

What the latest round of policymaking in the United States suggests is that the future holds more of the same, in the sense of maintaining the focus on vehicle emissions standards without a reasonably certain prospect that air quality (at least in terms of ambient O_3) will improve. By setting these standards below a limit that automobile makers and oil companies say can be reached by gasoline vehicles with or without reformulated gasoline, a major mandate has been placed on alternate fuels and vehicles, though none as yet have the combination of cost effectiveness and consumer acceptance to make them an efficient alternative to gasoline vehicles. Meanwhile, promising options, such as targeting high-emitting vehicles, go begging at the federal level.

Virtually unexplored but much debated are emissions fees and energy taxes that take into account the environmental and other external benefits of reducing various types of emissions. European governments appear more willing to consider and implement this approach than the United States (*24*), but the movement in the United States and elsewhere to internalize environmental costs in electricity supply decisions and prices (*39*) may do much to legitimize and further the application of emissions fees and energy taxes in the transportation sector.

The promising policies for developing countries are different from those for the developed countries. A targeting strategy may be less effective for developing countries, because most vehicles probably have high emissions. Whether inspection and maintenance programs can be effectively operated is also an open question. Penetration of alternate-fueled vehicles will be limited by lower baseline vehicle turnover rates, greater sensitivity of vehicle demand to price, and the need for less sophisticated, low-maintenance technologies. However, the clear health benefits from reducing Pb in gasoline and sulfur in diesel fuel make these fuel improvement strategies a priority. Transportation controls (including, for instance, removal of subsidies to parking, bus fares, and diesel fuel, as well as sticker programs as implemented in Singapore) offer some promise of cost-effectiveness, particularly if they can be instituted before commuting and freight transport patterns become hardened. Because of the possibly large health effects associated with diesel bus emissions relative to those of gasoline passenger automobiles, expansion of the latter at the expense of the former may not be all bad. In any event, reducing subsidies to public bus systems or encouraging their privatization are two ideas for placing bus service on firmer financial footing, with associated reductions in emissions resulting from poor maintenance and an aged capital stock.

REFERENCES

1. *National Air Pollutant Emission Estimates 1940–1989*. EPA-450/4-91-004. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C., March 1991.
2. *National Air Quality and Emissions Trends Report 1988*. EPA-450/4-91-003. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C., Feb. 1991.
3. World Health Organization and United Nations Environment Programme. *Assessment of Urban Air Quality, Global Environment Monitoring System*. Monitoring and Assessment Research Centre, London, 1988.
4. A. J. Krupnick and R. J. Kopp. *Catching Our Breath: Next Step for Reducing Urban Ozone. The Health and Agricultural Benefits of Reductions in Ambient Ozone in the United States*, Appendix to Office of Technology Assessment, U.S. Congress, Washington, D.C., 1988.
5. A. J. Krupnick. *An Analysis of Selected Health Benefits from Reductions in Photochemical Oxidants in the Northeastern United States*. Brian J. Morton, ed., Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C., and Resources for the Future, Washington, D.C., Sept. 1988.
6. A. J. Krupnick and M. A. Walls. The Cost-Effectiveness of Methanol Vehicles. Presented at the Association for Public Policy and Management Conference, San Francisco, Calif., Oct. 1990.
7. *Analysis of the Economic and Environmental Effects of Methanol as an Automotive Fuel*. Special Report, Office of Mobile Sources, U.S. Environmental Protection Agency, Sept. 1989.
8. *Cost and Benefits of Reducing Lead in Gasoline Final Regulatory Impact Analysis*. EPA-230-05-85-006. Office of Policy Analysis, U.S. Environmental Protection Agency, Washington, D.C., Feb. 1985.
9. B. Brunekreef, D. Houthuijs, L. Dijkstra, and J. S. M. Boleij. Indoor Nitrogen Dioxide Exposure and Children's Pulmonary

Function. *Journal of the Air Waste Management Association*, Vol. 40, 1990.

10. *Draft 1988 Air Quality Management Plan*. South Coast Air Quality Management District and Southern California Association of Governments, Los Angeles, 1989.
11. *Review of the National Ambient Air Quality Standards for Particulate Matter: Updated Assessment of Scientific and Technical Information. Addendum to the 1982 OAQPS Staff Paper*. EPA-450/05-86-012. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, N.C., Dec. 1986.
12. H. Ozkaynak and G. D. Thurston. Associations Between 1980 U.S. Mortality Rates and Alternative Measures of Airborne Particle Concentrations. *Risk Analysis*, Vol. 7, No. 4, Dec. 1987.
13. P. R. Portney, D. Harrison, Jr., A. J. Krupnick, and H. Dowlatabadi. To Live and Breathe in L.A. In *Issues in Science and Technology*, Vol. V, No. 4, 1989.
14. *Benefits of the 1989 Air Quality Management Plan for the South Coast Air Basin: A Reassessment*. National Economic Research Associates, Inc., Cambridge, Mass., March 1990.
15. *Catching Our Breath: Next Steps for Reducing Urban Ozone*. Office of Technology Assessment, U.S. Congress, Washington, D.C., 1989.
16. *Health Assessment Document for Diesel Emissions Workshop Review Draft*. EPA/600/8-90/057A. Office of Health and Environmental Assessment, U.S. Environmental Protection Agency, Research Triangle Park, N.C., July 1990.
17. *Transportation Energy Data Book: Edition 10*. ORNL-6565. Oak Ridge National Laboratory, Oak Ridge, Tenn., Sept. 1989.
18. International Energy Agency. *World Energy Statistics and Balances 1971–1987*. Organization for Economic Cooperation and Development, Paris, France, 1989.
19. C. G. Harral. Meeting the Transportation Aspirations of Developing Countries: Energy and Environmental Effects. In *Proceedings of the Energy and Environment in the 21st Century Conference*, Massachusetts Institute of Technology, Cambridge, March 1990.
20. A. Faiz, K. Sinha, M. Walsh, and A. Varma. *Automotive Air Pollution: Issues and Options for Developing Countries*. Infrastructure and Urban Development Department, Policy, Planning and Research Staff, The World Bank, Washington, D.C., Aug. 1990.
21. International Development and Energy Associates, Inc. *Clean Coal Technologies for Developing Countries*. Office of Technology Assessment, U.S. Congress, May 1990.
22. *Environmental Data Report*. GEMS Monitoring and Assessment Research Centre, London; World Resources Institute, Washington, D.C., and United Kingdom Department of the Environment, London; 1989.
23. Sierra Research, Inc. *Reactive Organic Gas Emission Factors and Fuel Economy Estimates for Advanced Technology Gasoline—and Methanol—Fueled Vehicles*. Report SR89-10-01. Western States Petroleum Association, Sacramento, Calif., Oct. 30, 1989.
24. J. B. Opschoor and H. B. Vos. *The Application of Economic Instruments for Environmental Protection*. Organization for Economic Cooperation and Development, Paris, France, 1989.
25. R. W. Crandall, H. K. Gruenspecht, T. E. Keeler, and L. B. Lave. *Regulating the Automobile*. The Brookings Institution, Washington, D.C., 1986.
26. M. A. DeLuchi, D. Sperling, and R. A. Johnson. *A Comparative Analysis of Future Transportation Fuels*. Research Report UCB-ITS-RR-87-13. Institute of Transportation Studies, University of California, Berkeley, Oct. 1987.
27. American Petroleum Institute. *Issue Summary, A Program for Reducing Emissions From Older Vehicles*. Public Policy Committee, Washington, D.C., April 1990.
28. D. Stedman. Automobile Carbon Monoxide Emissions. *Environmental Science and Technology*, Vol. 23, No. 2, 1989.
29. H. M. Haskew and J. J. Gumbleton. *GM's In-Use Emission Performance: Past, Present, Future*. SAE Technical Paper Series 881682, Society of Automotive Engineers, Inc., Warrendale, Pa., Oct. 1988.
30. *Air and Water Pollution Report*. Business Publishers, Inc., Silver Spring, Md., Nov. 5, 1990.
31. P. W. Wilson. Welfare Effects of Congestion Pricing in Singapore. *Transportation*, Vol. 15, 1988.
32. *Urban Transport Development with Particular Reference to Developing Countries*. Department of International Economic and Social Affairs, United Nations, New York, 1989.
33. *The Economist*. Feb. 18, 1989.
34. T. Rallis. *City Transport in Developed and Developing Countries*. Macmillan, Basingstoke, Hampshire, England, 1988.
35. C. Winston. Efficient Transportation Infrastructure Policy. *Journal of Economic Perspectives*, Vol. 5, No. 1, Winter 1991.
36. D. Gaskins and B. Stram. *A Meta Plan: A Policy Response to Global Warming, for Vienna Workshop for Cooperative Research on Climate Change*. Environment Division, Electric Power Research Institute, Palo Alto, Calif., 1990.
37. G. S. Eskeland and E. Jimenez. *Choosing Among Policy Instruments in Pollution Control: A Review*. Country Economics Department, Public Economics Division, The World Bank, Washington, D.C., June 1990.
38. Setting a Value on the Environment: Pollution Charges and Other Economic Instruments. Environmental Change Commission, Sweden, 1990, p. 59.
39. Pace University. *Environmental Costs of Electricity*. Oceana Publications, New York, 1990.

Publication of this paper sponsored by Committee on International Activities.

Internal Consistency and Stability of Measurements of Community Reaction to Noise

R. F. S. Job

Correlations between transportation noise exposure and community reaction indicate that on average only 18 to 22 percent (depending on the type of transportation noise) of the variation in reaction is attributable to noise exposure. Given the importance of these studies in determining not only the acceptable levels of noise but also how the noise is measured (equal energy units or not; with or without time of day weightings, etc.), the impact of reliability of measurement of community reaction and how reliability might be improved are considered. It is often suggested that error in measurement of reaction is a major reason for the relatively low noise-reaction correlations observed. However, few data exist on the internal consistency of composite scales of reaction. Data are presented on the internal consistency of various reaction scales, indicating that the typical general reaction scale and annoyance scales are reliable (in terms of internal consistency) and superior to the typical disturbance and complaint disposition scales. Further, reliability is increased by the use of several questions in a scale, rather than a single question. Thus, the data suggest that the best measure of community reaction is a composite general reaction scale (based on questions such as, "How much are you personally affected . . . "; and "How disturbed are you by . . . ") with a number of contributing questions. Noise-reaction correlations would be increased moderately with the use of an extremely reliable measure of reaction. However, when the limited reliability of the measurement of reaction and noise exposure, and modifying variables (especially attitude and sensitivity) are taken into account, the proportion of variation in reaction left unaccounted for is only around 20 percent.

Since the 1950s, community reaction to noise has been studied through the collection of noise exposure data and self-reported reaction data from residents around the noise source of interest. Such studies serve important purposes by attempting to identify the factors influencing community reaction to noise, most obviously including the noise itself. Such information has a number of consequences: it is relevant to basic theory of psychophysics, of annoyance in general, and of annoyance caused by noise in particular (*1,2*). Because the ultimate aim of these studies and subsequent countermeasures is not reduction of noise but reduction of reaction (annoyance, disturbance, etc.), the precise noise-reaction relationship is central to issues of acceptable noise level, compensation for residents exposed to noise, land-use planning, noise insulation, and adjustments to noise at the source.

As a simple, perhaps extreme, example, one of the theoretical and practical issues of concern is the value of the decibel equivalent number (k), which still receives considerable research attention (*3,4*). Studies have suggested values of k that best predict reaction, ranging from below 0 to 25 or 35 (*4*). High values of k would suggest that adding low-decibel noises would increase community reaction, whereas low values of k would suggest that such additions of low-decibel noise would decrease reaction. If this were true, one way of reducing reaction would be to add noise, not decrease it. If the number factor is irrelevant or trivial, many quieter noises are better, but if the number factor is large, fewer louder noises may have less impact on reaction, and if equal energy units provide the best prediction of reaction [again, a point of contention (*4–6*)], then overall energy reduction is the most likely beneficial option. Clearly then, the practical and theoretical implications of k are profound.

Similar arguments may be made for many features of noise (frequency range, time of day, duration, etc.). Consequently, considerable care with noise-reaction studies is justified. However, such studies have indicated correlations between the noise exposure and the reaction of individuals to the noise, ranging from 0.19 (*7*) to 0.64 (*8*), with a recently reviewed mean of 0.42 (*9*). This suggests that only 4 to 41 percent of the variation in individual reaction is accounted for by noise exposure. Although the average correlations are slightly higher for transportation noise than for impulsive noise [mean r = 0.42, 0.46, 0.46, 0.28 for road, aircraft, rail, and impulsive noise, respectively (*9*)], the percentages of variance in community reaction accounted for by noise still only average 17.6, 21.2, and 21.2 percent for road, rail, and aircraft noise exposure, respectively. Such low correlations in these critical studies deserve explanation. There are three principal explanations for these results:

1. The relationship between noise exposure and reaction is, in reality, not strong;
2. The noise measurement techniques or indices used have been inadequate; and
3. The reaction measurement techniques or indices have been inadequate.

Most socioacoustic surveys have emphasized the search for noise indices to more accurately predict human reaction, rather than examining many possible measures of reaction itself. For example, Fields and Walker (*10*) examined 44 indices, Bradley and Johan (*11*) examined 25 indices, and Bullen et al. (*12*) examined 88 indices. The examination of such large numbers of noise indices is justified by the common aim of testing the indices against one another, and the need for such indices for

Department of Psychology, University of Sydney, Sydney, New South Wales 2006, Australia.

legislation and land-use planning. However, with some exceptions (*13–15*), this concentration has resulted in a lack of research directed at obtaining more accurate measurement of reaction. In most surveys, it is likely that reaction measurements contribute a significant amount of error—more than is contributed by noise measurements as identified by computations of reliability. Therefore, the issue of reliability of measurement of community reaction to noise is addressed in the following actions.

TEST-RETEST RELIABILITY

Following the classic literature, Hall and Taylor (*15*) have distinguished two meanings subsumed under the general label of reliability: internal consistency and stability. Stability refers to the change in the true magnitude of the characteristic being measured, whereas internal consistency refers to the consistency of the items used to construct a composite scale. Of course, stability in the sense of change in true magnitude is not directly measurable. Rather, differences in repeated measurements at different times are contributed to by changes in the true score and changes in the error components of the observed score. Ultimately, in terms of measurement, the difference between stability and internal consistency becomes this: stability is measured by repeating exactly the same questions to the same respondents with a significant time interval—usually several months to 1 year in studies of noise reaction stability. Internal consistency is measured by asking nonidentical questions, designed to measure the same characteristic of the same respondent over a very short interval—within the one interview of typically 30 min to 1 hr.

However, measures of stability may be confounded by error components. For example, the supposedly identical questions may be put slightly differently by the interviewer on the two occasions: voice intonation, eye contact, and nonverbal cues from the interviewer are known to influence survey outcomes (*16–19*). If these effects are conceived of as constituting slight changes in the question, then part of the variation observed as stability is actually a matter of internal consistency. Furthermore, it is possible that some of the variation observed in measurements of internal consistency really reflects changes of stability over the brief interval of the interview. In fact, the conduct of the interview itself may point out many possible effects of the noise not previously identified all at once. For example, for questions on disturbances to sleep, watching television, conversation, telephone calls, symptoms such as restlessness, nerves, and headaches may occasion a change in the true component of reaction. Accordingly, the theoretical distinction between stability and internal consistency may be somewhat blurred at the point of measurement. Nonetheless, the distinction is recognized as being of theoretical and practical value.

Four studies have been reported on the stability of social survey measures of reaction to noise across time (*11,15,20,21*). Despite substantial differences in sample sizes and the scales of reaction examined, the results are similar. Scales based on single items achieved the following reliabilities: (a) seven-point, semantic differential scale of dissatisfaction with traffic noise, 0.642 (*22*), 0.61 (*20*); (b) four-point, fully labeled unipolar scale of the extent to which the person is bothered by traffic noise, 0.63 (*21*); (c) nine-point, bipolar fully labeled scale from extremely agreeable to extremely disturbing, 0.53 (aircraft noise), 0.58 (traffic noise), 0.30 (overall noise) (*15*); and (d) eleven-point (0 to 10), unipolar scale with endpoints labeled "Not at all Disturbed" and "Unbearably Disturbed", 0.50 (aircraft noise), 0.58 (traffic noise), and 0.26 (overall noise) (*15*). Hall and Taylor (*15*) noted that the lower correlations for aircraft and overall noise compared with road noise may be because of changes in the noise environment. Therefore, the results for traffic may be taken as the least ambiguous measure of test-retest reliability in this study. Thus, over the different measurement scales, reliability varied over a very narrow range (0.58 to 0.64).

Bradley and Johan (*11*) reported data on the test-retest reliability of composite scales of reaction to traffic noise. The annoyance scale achieved a time-separated test-retest reliability (stability) of 0.75, while for the interference scale the stability was 0.81. Consistent with Hall and Taylor's (*15*) claim that little data existed on the internal consistency of reaction scales, Bradley and Johan's study is the only one that bears on internal consistency, albeit with a comparatively small sample size. The reaction measurement was based on analysis of 80 response items. The alpha coefficient, a measure of internal consistency of a multi-item scale (*22*), was 0.95 for the annoyance scale and 0.94 for the interference scale.

In consideration of this dearth of data, the first aim is to present analyses of data available from existing surveys to provide estimates of internal consistency of various reaction scales. The impact of the use of composite scales and single item scales of reaction on test-retest reliability is also examined. Finally, the effect of increases in reliability on the amount of variation in reaction that is accounted for (in terms of noise exposure and modifying variables) is examined.

The last purpose is relevant to consideration of the scientific and practical consequences of the small percentage of variation in reaction accounted for by noise exposure. From a practical point of view, the low correlations have led to criticism of the populations studied (*23*) and doubts about the environmental noise standards derived from these studies (*24*). Furthermore, the selection of the indices of noise most predictive of community reaction is hampered by the inaccuracy of reaction measures, which reduces the opportunity to distinguish between noise indices (*25*). From the viewpoint of basic science as well as practical application, research may be motivated by the need to account for the variation in reaction left unexplained by noise exposure. Thus, numerous other factors such as personality, various measures of socioeconomic status, sex, age, marital status, noise sensitivity, and attitude toward the noise source have all been examined. Langdon (*25*) has suggested that this may not be a problem. That is, the failure of noise exposure to account for more than a small percentage of the variation in reaction may reflect unreliability of the measures used. The present paper tests this possibility.

ANALYSIS OF EXISTING SURVEYS

Several studies by the socioacoustic research team of the Australian National Acoustic Laboratories have been conducted along similar lines, although a range of noise sources has been investigated: artillery noise, rifle range noise, and noise from civilian and military aviation. In all cases, respondents have been asked a large number of questions to assess their reactions. In all cases, a composite scale of overall reaction has been calculated using factor analyses and regression. The studies shared several scales of reaction. Thus, not only are data on internal consistency available for these scales, but also the internal consistencies can be compared across studies, noise sources, and populations (civilian or military). The following scales of reaction were used in the studies *(12,14,26–29)* and can be assessed for internal consistency.

1. General scale based on how much the person is affected by, and dissatisfied with, the noise;
2. Annoyance scale;
3. Complaint disposition scale;
4. Disturbance scale; and
5. Overall reaction based on all of the previous scales.

Table 1 presents the relevant results for each scale in terms of a measure of internal consistency, the maximum likelihood estimate of reliability known as the α coefficient *(22,30)*. These data indicate that all the scales used have high internal consistencies regardless of the noise type or the population surveyed (civilian or military). The internal consistencies are themselves extremely consistent, with only a small range of values represented for each scale. The means of the α coefficients for each scale indicate that the scales do not differ greatly in terms of their internal consistency (the range of mean α values is 0.853 to 0.923). However, some of the scales required considerably more items to obtain the observed reliabilities than others. As a guide to the internal consistency of the scales—independent of the number of items used to compose the scale—the average interitem correlations have also been presented. These data suggest that the general scale and the annoyance scale are superior to the disturbance and complaint disposition scales. The general scale is also significantly superior to the annoyance scale ($p < 0.01$). However, in practical terms this statistically significant difference is not of great importance. As the last row of Table 1 indicates, both scales require three items contributing to them in order to achieve an α value of greater than 0.9. This row also indicates the practical consequence of the reduced internal consistency of the disturbance and complaint disposition scales that required large numbers of items to achieve an α value of 0.9.

TABLE 1 ALPHA COEFFICIENTS FOR THE VARIOUS SCALES OF REACTION

Study	Sample Size	Noise	General	Annoyance	Disturbance	Complaint Disposition	Overall Reaction
Bullen, et al. 1986 (28)	3580	Aircraft	2 .90**	3 .92	8 .85	8 .88	6 .93*
Hede & Bullen 1982 (27)	201	Rifle	2 .92**	4 .94	6 .76	6 .90	4 .92 *
Bullen, et al. 1985 (12)	624	Aircraft	2 .88**	5 .94**	6 .90*	- -	3 .90*
O'Loughlin, et al., 1986 [24]	318	Rifle Range	2 .89*	- -	- -	- -	5 .92*
Bullen, et al. submitted (29)	1,626	Artillery Range	2 .88**	3 .89	9 .90	8 .84	5 .877*
Job & Bullen 1987 (14)	45 (interview sample)	Aircraft	- -	- -	- -	- -	5 .841
	41 (group sample)		- -	- -	- -	- -	5 .939
Mean No. Items			2.00	3.75	7.25	7.33	
Mean alpha for each scale			.894	.923	.853	.873	not computed, since only a small sample size study (14)
Mean inter-item correlation			.808	.762†	.445†	.484†	has appropriate data
Number of items required for alpha to exceed .9			3	3	12	10	

* For overall reaction the figures reported are results of multiple regression of the various scales against the general scale.

** Computed from mean inter-item correlations.

† Computed from the alpha, via equation 1 in text.

Effects of the Number of Items in a Scale

As pointed out earlier, with the brief interval test-retest procedure the result is interpreted as internal consistency. However, this measure may be affected by factors other than the differences in question wording aimed at measuring the same underlying feature. Reliability in social surveys may be reduced by unwanted influences such as voice intonation, nonverbal cues from the interviewer, and idiosyncratic interpretations of the wording of the question (*16–19*). Consistent with the possibility that the effects of the interviewer's voice intonation and nonverbal cues may add to the error, a recent study found superior internal consistency of a reaction scale when the questionnaire was self-administered rather than given by interview (*14*). The error variance attributable to such influences is reduced by user of several questions to measure the same variable. Thus, the use of several items not only allows examination of internal consistency, but is also likely to increase the observed stability. If a number of questions are used to form a scale rather than a single question, the maximum likelihood estimate of reliability is given by

$$\alpha = jr/[1 + (j - 1)r] \qquad (1)$$

where r is the average correlation between the questions and j is the number of questions (*31*).

As a guide to the increase in reliability which occurs with the use of more questions, this function is graphed for various values of r in Figure 1.

By the way of confirmation of this increase in reliability in relevant surveys, it is possible to calculate the brief interval test-retest reliability (or internal consistency) of various questions used to measure annoyance in surveys using more than one question. Bullen et al. (*25*) used four questions as measures of annoyance caused by aircraft noise. The average interitem correlation for the four items used by Bullen et al. was calculated, as was the average correlation between scales created by combining any two of the items. In the calculation of the latter correlation, only correlations between scales with no items in common were used; only three such comparisons exist: 1+2 versus 3+4; 2+3 versus 1+4; and 2+4 versus 1+3. The average correlation between the single items was 0.783, which yields a predicted correlation (or brief interval test-retest reliability) of 0.878 for the two-item scales. The average correlation between the two-item scales, calculated from the three comparisons above, was 0.879. This value is close to the predicted value and significantly different from the average single-item interitem correlation of 0.783 ($p <$ 0.05). Similar analyses were conducted for other studies where the relevant data were available. The results are presented in Table 2.

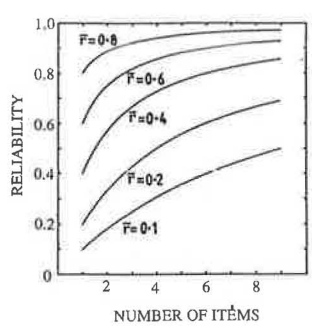


FIGURE 1 Composite scale reliability as a function of the number of items in the scale and the interitem reliability.

The examples in Table 2 indicate that Equation 1 yields a good prediction of the results of socioacoustic studies and that the reliability of self-reported noise reaction measurements may be significantly increased by the use of more than one item to measure reaction. Despite this increase in reliability, numerous surveys have used a single question as the basic measure of reaction even though many relevant questions existed in the questionnaires (*6,32–34*).

By taking into account the reliability of the reaction measure, it is possible to calculate the extent to which the true variance of reaction is predicted by the observed noise exposure. The equation is (*31*)

$$r_{nR} = r_{nr}/(r_{rr})^{1/2} \qquad (2)$$

where

r_{nR} = correlation between noise exposure and the true component of reaction,
r_{nr} = obtained correlation between noise exposure and reaction, and
r_{rr} = reliability coefficient for the reaction measurement.

This equation may be applied to the changes in reliability arising from the use of several items in measurement of reaction by Bullen et al. (*12,27*). Bullen, Job, and Burgess (*12*) found that the reliability of a single-item scale (i.e., the mean interitem correlation) was 0.740, whereas the measures of reaction constructed from several items achieved a reliability of r_{rr} = 0.902, with a noise-reaction correlation of r_{nr} = 0.580. Substituting the last two figures in Equation 2 results in correlation between noise exposure and the true component of reaction, of r_{nr} = 0.611. Substituting this figure in Equation 2 with the single-item reliability of 0.740 reveals that the observed correlation would have been r_{nr} = 0.526 with the use of a single-item reaction scale. Thus the drop in reliability of reaction measurement from r_{rr} = 0.902 to 0.740 would have resulted in a drop in the observed noise-reaction correlation from r_{nr} = 0.580 to 0.526. Similarly, the use of a single-item measure (r_{rr} = 0.783) rather than a composite measure (r_{rr} = 0.931) by Bullen, Hede, and Kyriacos (*27*) would have resulted in the observed noise-reaction correlation being reduced from r_{nr} = 0.361 to 0.331. Thus, in the two studies under consideration, the use of composite scales of reaction resulted in increases of 6.0 and 2.1 in the percentage of variation in reaction accounted for by noise exposure.

A more general case yields a slightly larger effect because the studies analyzed have unusually high reliabilities of reaction measurement. For example, assuming an r_{nr} value of 0.42 [the average correlation in a recent review (*6*)], and a single-item reaction scale with reliability of r_{rr} = 0.630 (*21*), the expected correlation with a reaction scale based on five items with reliability of 0.63 is r_{nr} = 0.500. This represents an increase of 7.4 percent in the variation in reaction accounted for by noise exposure. It would appear that although

TABLE 2 COMPARISON OF PREDICTED AND OBSERVED RELIABILITY FOR COMPOSITE SCALES OF REACTION, AND EFFECTS OF RELIABILITY ON THE NOISE-REACTION CORRELATION

Study	Mean inter-item Correlation ($\bar{r}$)	Number of Items (k)	Reliability of composite scale Predicted	Observed	noise/reaction correlation Predicted	Observed
Bullen *et al.*, 1985 (12)	.740	2	.851	.852		
	.740	4	.919	.902		
Griffith *et al.*, 1980 (22)	.642	2	.782	-	.397	.37
	.642	3	.843	-	.413	.43
	.642	4	.878	-	.421	.44
Hall & Taylor, 1982 (15) aircraft	.733	3	.890	.86		
Hall & Taylor, 1982 [9] road	.687	3	.868	.83		
Hall & Taylor, 1982 [9] overall noise	.653	3	.850	.84		
Bullen, et al. 1986 (28)	.783	2	.878	.879		
	.783	4	.941	.931		

increases in observed noise-reaction correlations will result from the use of multi-item scales of reaction, the increases will not be large.

Changes in the reliability of noise exposure measures will have similar effects on the observed r_{nr} value. However, given the already high reliability in noise measurement (*9*), little effect of improvements in noise measurement reliability can be expected. The correlation between the true component of reaction and the true component of noise exposure can be calculated (*31*) from the equation

$$r_{NR} = r_{nr}/(r_{nn} \cdot r_{rr})^{1/2} \qquad (3)$$

where

r_{NR} = correlation between the true component of reaction and the true component of noise exposure, and
r_{nn} = reliability coefficient of the noise measurement.

Thus, this correlation can also be calculated from the equation

$$r_{NR} = r_{nr}/\{jr_{rr}/[1 + (j - 1) \cdot r_{rr}] \cdot r_{nn}\}^{1/2} \qquad (4)$$

where

j = number of items contributing to the reaction scale, and
r_{rr} = average interitem correlation of the j items.

The results of the application of Equations 3 and 4 to studies in which the relevant details were available are presented in Table 3. Three features of the data in Table 3 are noteworthy. First, the long and brief interval measures of reaction scale reliability differ (means = 0.724 and 0.916, respectively) indicating greater variability with more time between the measurement points. These data suggest that the internal consistency measure is an underestimate of the error component of the scales, or that true reaction changes significantly over time. Second, this difference allows a more powerful correction to the computed r_{NR} correlation in the case of the extended delay reliability test. Thus the originally very similar r_{nr} correlations for the studies using long and brief interval reliability tests (means = 0.382 and 0.386, respectively) diverge after correcting for reaction scale reliability [means = 0.450 and 0.402, respectively (see Table 3)]. Third, estimates of noise measurement reliability were only available in studies in which the brief delay reliability (internal consistency) was assessed. Accordingly, the values of r_{nR} computed would have been larger if those studies had assessed extended delay reliability. For example, substituting the average extended delay reliability (r_{rr} = 0.724) for the r_{rr} observed in each of these studies results in r_{NR} increasing to 0.696, 0.293, and 0.431 in the three studies (means = 0.473 compared with the mean of 0.423 in Table 3). Overall, these data indicate that estimates of the amount of variation in reaction accounted for by noise exposure increase slightly when the reliabilities of the measurements of noise and reaction are taken into account: the coefficient of determination (r_) expressed as a percentage increased from an average of 15.0 to 22.4 percent for the three studies mentioned.

Computing the Influence of Modifying Variables

In an attempt to identify the amount of variation in reaction left unexplained, it is worthwhile to consider the role of modifying variables. (Modifying variables are variables in the respondent or environment that may modify reaction to noise.) Although many have been shown to be significantly related to reaction, only two are commonly assessed and consistently account for more than a small percentage of the variation in reaction: noise sensitivity and attitude to the noise source (*9*).

Three problems arise in further consideration of the impact of these variables. Firstly, a correlation between a modifying variable and reaction does not guarantee that the modifying variable influences reaction. Some third variable may influence both the supposed modifying variable and reaction, or the modifying variable may be influenced by reaction, not vice versa. Secondly, the calculation of the coefficient of determination for exposure and reaction on the basis of (the-

TABLE 3 COMPUTED EFFECTS OF RELIABILITY ON OBSERVED NOISE-REACTION CORRELATIONS

r_{nr} and reference source		r_{nn} and reference source (unless as for r_{nr})	r_{rr} and reference source (unless as for r_{nr})†		Computed r_{nR}	Computed r_{NR}
Long delay test-retest reliability						
.477	Bradley & Jonah Jonah, 1979 (11)	-	.75		.551	-
.36	Griffiths et al. 1980 (22)	-	.642		.449	-
.37		-	.782**		.418	-
.43		-	.834**		.471	-
.44		-	.878**		.470	-
.21	Langdon 1976 (41)	-	.626	(average of .6 (21) & .642 (22)	.265	-
.325*	Langdon 1978 (26)	-	.61		.416	-
.37	Langdon 1978 (37)	-	.76**		.425	-
.46	McKennell 1963, 1973 (38,39)	-	.63	(22)	.580	-
Means	(Long Delay) .382	-	.736	(excluding .626)	.450	-
Brief delay reliability (internal consistency)						
.477	Bradley & Jonah, 1979 (11)	-	.95		.489	-
.29	Bullen & Hede, 1982 (40)	-	.922	Hede & Bullen 1982 (27)	.302	-
.58	Bullen et al. 1985 (12)	.96	.902		.611	.623
.220	Bullen et al., submitted (29)	.78	.877		.235	.266
.361	Bullen et al. 1986 (28)	.969	.931		.374	.380
Means	(brief delay) .386	.903	.916		.402	.423

* Average of .33 and .32 reported.

** Computed by Equation 1 from information reported (k and $\bar{r}$)

† Where r_{rr} was not obtained from the same study as r_{nr}, the same reaction scale, same type of noise source and same country were involved.

oretically existing) completely reliable measures of both factors has already taken into account some variation in reaction possibly attributable to the modifying variables. This is confounded because the variation in reaction attributable to the modifying variables is treated as error variance that is eliminated in calculating for a completely reliable (zero error variance) reaction scale. Third, the modifying variables may themselves correlate.

The first problem is not easily handled. The data are conflicting as to whether attitude, for example, is a genuine modifying variable or not (*9,26,35*). In the ensuing analyses, it is assumed, with some justification (*35*), that both sensitivity to noise and attitude towards the noise source are genuine modifying variables. The second problem may be avoided by identifying the amount of variation in reaction that is reliable (true) and calculating how much of this component is accounted for by the other factors. The third problem presents particular difficulties when attempting to correct for the reliability of the scale used to measure the modifying variable. That is, a multiple regression could be used to compute, for example, the amount of variation in reaction accounted for by sensitivity after attitude is taken into account. However, this partial correlation of sensitivity and reaction would be affected by the increased correlation of reaction and the true component of attitude over and above the correlation of attitude (as measured) and reaction.

Nonetheless, examination of the data where such multiple regression has been carried out (*26,27*) may yield a guide as to the minimum amount of variation in reaction that is accounted for in the analysis. Bullen et al. (*27*) found that in a multiple regression of modifying variables against reaction, the modifying variables account for 59.4 percent of the variation in reaction. Noise exposure accounted for 13 percent of the variation in reaction. However, some small correlations between the modifying variable and noise exposure mean that these two percentages cannot be added together, as an estimate of the amount of variation in reaction accounted for by noise and modifying variables. This problem arises because the partial correlation for noise exposure and reaction could be expected to be less than 0.361 (r_ = 13.0 percent). However, given the reliability of the reaction scale (r_{rr} = 0.931), only 86.7 percent of the variation in reaction is true variation to be explained. This amount still leaves somewhere between 14.3 and 27.3 percent of the variation in reaction unaccounted for. A significant but uncalculated proportion of this would be taken up by consideration of the lack of reliability of the modifying variables and the exposure measure. Hede and Bullen (*26*) also computed the multiple regression of the modifying variables against reaction, but on this occasion included noise exposure. They calculated that 65.5 percent of the variation in reaction is accounted for by sensitivity, attitude, and noise exposure in multiple regression. The reliability of the reaction scale in this study (r_{rr} = 0.92) indicates that 84.6 percent of the variance in reaction is true variance. This leaves less than 20 percent unaccounted for. More variation would be accounted for in consideration of the reliability of the modifying variables and noise exposure. However, these conclusions are based on the assumption, mentioned earlier, that the modifying variables influence reaction (rather than the correlations being based on some other connection). Although there is some justification for this assumption, conclusions based on it must remain tentative.

CONCLUSIONS

1. The data presented correct, at least partially, the dearth of data on the internal consistency of scales of reaction to noise. The analysis indicates that two scales (the general scale and the annoyance scale) can achieve reliabilities (internal consistencies) in excess of 0.9, with only three items. Complaint disposition and disturbance scales are not as reliable.

2. Analyses of the effects of the use of several items in a given scale of reaction indicate that the increase in reliability is worthwhile, in that the increased reliability may allow greater sensitivity in determining which measures and features of noise are the most important predictors of community reaction.

3. When noise exposure, modifying variables, and the percentage of error variance in the reaction measure are taken into account, the variation in reaction left to be explained is substantially reduced (to around 20 percent). This provides partial support for the suggestion that the variation in reaction left unexplained is not a problem (*25*).

REFERENCES

1. S. S. Stevens. *Psychophysics: Introduction to its Perceptual Neural and Social Aspects*. John Wiley, New York, 1975.
2. S. Fidell, T. Shultz, and D. M. Green. A Theoretical Interpretation of the Prevalence Rate of Noise-Induced Annoyance in Residential Populations. *Journal of the Acoustical Society of America*, Vol. 84, 1988, 2109–2113.
3. J. M. Fields. The Effect of Numbers of Noise Events on People's Reactions to Noise: An Analysis of Existing Survey Data. *Journal of the Acoustical Society of America*, Vol. 75, 1984, pp. 447–467.
4. T. Yoshida and S. Nakamura. Effect of Noise Events on Inhabitants' Reactions to Railway Noise. *Journal of the Acoustical Society of Japan (English)*, Vol. 10, 1989, pp. 339–348.
5. R. B. Bullen and R. F. S. Job. Re-analysis of data presented in "Community Response to Blasting," by Fidell et al. [J. Acoust. Soc. Am., 74, 888–893 (1983)]. *Journal of The Acoustical Society of America*, Vol. 78, 1985, pp. 799–800.
6. S. Fidell, R. Horonjeff, T. Shultz, and S. Teffeteller. Community Response to Blasting. *Journal of the Acoustical Society of America*, Vol. 74, 1983, pp. 888–893.
7. R. F. S. Job and A. J. Hede. Community Reaction to Noise from Power Stations. *Proc., Inter-Noise 1989*, Newport Beach, California. Noise Control Foundation, Poughkeepsie, N.Y., 1989, pp. 865–893.
8. J. Lambert, F. Simonnet and M. Vallet. Patterns of Behavior in Dwellings Exposed to Road Traffic Noise. *Journal of Sound and Vibration*, Vol. 92, 1984, pp. 159–172.
9. R. F. S. Job. Community Response to Noise: A Review of Factors Influencing the Relationship Between Noise Exposure and Reaction. *Journal of the Acoustical Society of America*, Vol. 83, 1988, pp. 991–1001.
10. J. M. Fields and J. G. Walker. The Response to Railway Noise in Residential Areas in Great Britain. *Journal of Sound and Vibration*, Vol. 85, 1982, pp. 177–255.
11. J. S. Bradley and B. A. Jonah. The Effects of Site Selection Variables on Human Responses to Traffic Noise, Part 1: Type of Housing by Traffic Noise Level. *Journal of Sound and Vibration*, Vol. 66, 1979, pp. 589–604.
12. R. B. Bullen, R. F. S. Job, and D. H. Burgess. *Reaction to Aircraft Noise on R.A.A.F. Bases*. National Acoustic Laboratories Commissioned Report No. 7. Australian Government Publishing Service, Canberra, 1985.
13. R. M. Edwards. A Social Survey to Examine the Variance of Aircraft Noise Annoyance. *Journal of Sound and Vibration*, Vol. 41, 1985, pp. 41–50.
14. R. F. S. Job and R. B. Bullen. The Effects of a Face to Face Interview Versus a Group Administered Questionnaire in Determining Reaction to Noise in the Workplace. *Journal of Sound and Vibration*, Vol. 116, 1987, pp.161–168.
15. F. L. Hall and S. M. Taylor. Reliability of Social Survey Data on Noise Effects. *Journal of the Acoustical Society of America*, Vol. 72, 1982, pp. 1212–1221.
16. A. Barah and C. F. Cannell. The Effects of Interviewer Voice Intonation on Reporting. In *Experiments in Interviewing Techniques: Field Experiments in Health Reporting, 1971–1977*, C. F. Cannell, L. Oksenberg and J. M. Convere, eds., Survey Research Center, University of Michigan, Ann Arbor, 1979.
17. R. L. Kahn and C. F. Cannell. *The Dynamics of Interviewing*, John Wiley, New York, 1957.
18. G. Kalton, M. Collins and L. Brook. Experiments in Wording Opinion Questions. *Applied Statistics*, Vol. 27, 1978, pp. 149–161.
19. H. Shuman and S. Presser. *Questions and Answers in Attitude Surveys*. Academic Press, New York, 1981.
20. I. D. Griffiths and F. R. Delauzun. Individual Differences in Sensitivity to Traffic Noise: An Empirical Study. *Journal of Sound and Vibration*, Vol. 55, 1977, pp. 93–107.
21. I. D. Griffiths, F. J. Langdon, and M. A. Swan. Subjective Effects of Traffic Noise Exposure: Reliability and Seasonal Effects. *Journal of Sound and Vibration*, Vol. 71, 1980, pp. 227–240.
22. J. Nunnally. *Psychometric Theory*. McGraw-Hill, New York, 1978.
23. M. N. Moreira and M. E. Bryan. Noise Annoyance Susceptibility. *Journal of Sound and Vibration*, Vol. 52, 1972, pp. 449–462.
24. M. E. Bryan and W. Tempest. Are Our Noise Laws Adequate? *Applied Acoustics*, Vol. 6, 1973, pp. 219–232.
25. F. J. Langdon. Reliability of Estimates of Annoyance With Road Traffic Noise. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978.
26. A. J. Hede and R. B. Bullen. Community Reaction to Noise from a Suburban Rifle Range. *Journal of Sound and Vibration*, Vol. 82, 1982, pp. 39–49.
27. R. B. Bullen, A. J. Hede, and E. Kyriacos. Reaction to Aircraft Noise in Residential Areas Around Australian Airports. *Journal of Sound and Vibration*, Vol. 108, 1986, pp. 199–225.
28. R. B. Bullen, A. J. Hede, and R. F. S. Job. Community Reaction to Noise from an Artillery Range (in preparation).
29. B. J. O'Loughlin, R. B. Bullen, A. J. Hede, and D. H. Burgess. *Community Reaction to Noise from Wiliamtown Rifle Range*. Commissioned Report No. 9. National Acoustic Laboratories, Australian Government Publishing Service, Canberra, 1986.
30. C. H. Hull and N. H. Nie. *SPSS Update*. McGraw-Hill, New York, 1979.
31. P. Guilford. *Psychometric Methods*. Tata McGraw-Hill, Bombay, India, 1954.

32. B. J. Murray and B. Avery. *Survey of Community Reaction to Overpressure from Blasting*. Wilkinson-Murray Report W4244, Linfield, Australia, 1984.
33. R. Rylander, M. Bjorkman, V. Ahrlin, S. Sorensen, and K. Berglund. Aircraft Noise Annoyance Contours: Importance of Overflight Frequency and Nose Level. *Journal of Sound and Vibration*, Vol. 69, 1980, pp. 583–595.
34. S. Sorenson and J. Magnusson. Annoyance Caused by Noise from Shooting Ranges. *Journal of Sound and Vibration*, Vol. 82, 1979, pp. 437–442.
35. A. C. McKennell. Annoyance from Concorde Flight around Heathrow. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 562–566.
36. F. J. Langdon. Reliability of Estimates of Annoyance with Road Traffiç Noise. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 567–570.
37. A. C. McKennell. *Aircraft Noise Annoyance Around London (Heathrow) Airport*. U.K. Government Social Survey Report SS337, 1963.
38. A. C. McKennell. Psycho-Social Factors in Aircraft Noise Annoyance. *Proc., International Congress on Noise as a Public Health Problem*, Dubrovnik, 1973, pp. 627–644.
39. R. B. Bullen and A. J. Hede. Assessment of Community Noise Exposure from Rifle Shooting. *Journal of Sound and Vibration*, Vol. 82, 1982, pp. 29–37.
40. F. J. Langdon. Noise Nuisance Caused by Road Traffic in Residential Areas: Part 1. *Journal of Sound and Vibration*, Vol. 47, 1976, pp. 243–263.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Impact and Potential Use of Attitude and Other Modifying Variables in Reducing Community Reaction to Noise

R. F. S. Job

A large number of variables in addition to noise correlate with community reaction, yet to date virtually the only method used to reduce reaction is to reduce noise exposure. It is suggested that this obvious and effective countermeasure may be supplemented with additional measures. Modifying variables may be manipulated to reduce reaction if they meet three criteria: the variable must have a sufficiently strong relationship with reaction to be of practical value; the variable must be a causal factor in reaction (or turn out to be part of reaction); and the variable must be potentially manipulable. Available data suggest that a number of modifying variables may meet these criteria (attitude, reduction in rates, direct sight of the noise source, and fear of crashes). It is suggested that more research is needed to establish with greater confidence the direction of causality and the impact of modifying variables on reaction. In particular, the manipulation of modifying variables in order to reduce reaction (either alone or in combination with noise reduction) deserves direct testing.

Socioacoustic investigations provide valuable data on the correspondence (correlation) between reaction and noise. It has been accepted universally that these correlations represented a causal sequence from noise to reaction such that a reduction in noise would lead to a reduction in reaction. This assumption is supported by data on community reaction associated with changes in noise exposure (*1–4*). However, the notion of an untainted correspondence between noise level and reaction level is a substantial oversimplification of events (*5*). Indeed, the fact that only around 18 percent of the variance in reaction is explicable in terms of noise exposure (*6*) indicates that noise is not remotely the whole story.

Many other factors (attitude to the noise source, sensitivity to noise, age of the hearer, etc.) influence reaction to noise. However, these factors are often seen as inconveniences that obscure the real noise-reaction relationship. The point is to suggest that some of these other factors themselves may be just as real as noise in accounting for reaction, and may also be useable in terms of reducing community reaction. That is, the general argument may be made that if any variable (say x)—not just noise—correlates with reaction, it is possible that changes in x will cause changes in reaction. (Of course, a correlation does not indicate causation, but may only be suggestive of a causal link.)

There is a good reason to expect that a number of features of the noise, the hearer, and other physical factors (e.g., time of day, sight of the noise source) will influence the degree of reaction occasioned by the noise. That some of these factors have a genuine possibility of being causal agents in reaction may be seen in the following examples. Consider hearing a neighbor playing a synthesizer. All the following factors could be expected to influence the level of reaction.

1. The loudness and duration of the synthesizer music may influence reaction (i.e., basic features of the noise as measured by L_{eq}).
2. Changes in the noise (making it a random sequency of notes instead of a melody) may increase reaction (disturbance, annoyance) because it is more likely to be heard as noise rather than music, yet such changes would not be reflected by measures like L_{eq} or any typical measure of noise exposure.
3. Knowledge of and attitude to the noise may have an impact. The synthesized music may be perceived as noise or music depending on whether the hearer knows and likes synthesized music or not, as well as the hearer's opinion of the neighbor (i.e., attitudes to the noise source).
4. The reason for the noise may have an impact—is the noise perceived as useful or frivolous? This factor can be seen more clearly in say, reaction to military aircraft: are overflights seen as joyrides that waste the taxpayers money or as a necessary part of defense (i.e., the value of the noise source, as a part of attitude).
5. Is the noise loud because it is unavoidable, or is the producer simply being inconsiderate? Again, this factor is important in relation to aircraft, road, and rail noise: do the authorities in charge care about the noise (i.e., misfeasance as another part of attitude).
6. Time of day—at 1:00 p.m. it may be music but at 1:00 a.m. it is noise.
7. Features of the hearer that vary over time may have an additional impact—at the time of the noise the hearer may be mowing the lawn or reading a book in the quiet. The influence on reaction of such factors as level of concentration is supported by laboratory studies (*5*).
8. In some cases, reaction may vary with whether or not the hearer can see the source of the noise.
9. Degree of control over the noise: can the hearer move to a less exposed room, or close the windows?
10. How distracted by noise is that particular hearer (i.e., sensitivity to noise).
11. Demographic variables (age, sex, socioeconomic status, length of residence) may influence reaction either indepen-

Department of Psychology, University of Sydney, Sydney, New South Wales 2006, Australia.

dently or as a consequence of the above factors (e.g., a 25-year-old may be more likely to hear a synthesizer as music than a 70-year-old). Home ownership may have an influence through perception of loss of property value.

12. Finally, these factors may interact: for example, loss of melody (Factor 2) may interact with value of the noise source (Factor 4). If the melody is poor because the neighbor is drunk, this may have a worse effect than if the melody is poor because the neighbor's 9-year-old son is learning to play; or more obviously, sensitivity to noise (Factor 9) may interact with loudness (Factor 1).

In these examples, it seems quite plausible to suggest that a variety of factors have an important role in the causation of reaction to noise.

MODIFYING VARIABLES

The previous factors (called modifying variables because they modify reaction to the noise) are often measured in the more sophisticated studies of community reaction to noise. Social surveys may include questions that assess attitudes to the noise source, sensitivity to noise, and a variety of demographic variables. Additional physical information may be collected on windows, visibility of noise source, and age of the dwelling.

A large number of variables, other than noise, have been shown to have a significant relationship with reaction. These include the following: sex (*7*); age (*8–11*); marital status (*7,12*); education (*12*); socioeconomic status or income (*8,13*); personality (*14*); psychiatric status (*15*); home ownership (*9,16*); whether local groups are seen to have done anything about the noise (*12*); whether the neighborhood has complained about the noise (*12*); reductions in rates (taxes) because of the noise (*12*); duration of residence (*7,12,13,16,17*); satisfaction with the neighborhood (*12,13*); age of the dwelling (*12*); satisfaction with life (*12*); fear of aircraft or railway crashes (*10,12,13*); open and closeable windows and control over the noise (*18*); number of people in the household (*12*); being able to see the noise source (*19,20*); annoyance with dirt, air pollution, lights, and loss of privacy (*12*); discussion of noise with friends and neighbors (*21*); belief that noise can harm health (*12*); attitude (*6*); and sensitivity (*6*).

Another variable that has received research attention is complaint behavior (*12,22*). However, although the variable of complaint by the neighborhood (*12*) will be considered as a modifying variable, individual complaints will not be so considered. Individual complaints have not been considered because such complaint behavior is influenced by many factors extraneous to present considerations, such as level of verbal or written skill, or knowledge of to whom the complaint should be directed. Complaints are also known to be elicited more by changes in the noise (i.e., unusual noise) rather than usual but excessive noise (*22*). Complaints also indicate no clear relationship to noise exposure except for the possibility that complaints are more likely in lower noise exposure areas (*22*).

This review has presented a prima facie case for possible causal connection between a large number of factors and reaction to noise. However, much more than a significant statistical relationship such as a correlation is required to provide argument for a causal connection. The next section considers the criteria necessary for a modifying variable to be potentially usable in reducing community reaction to noise.

UTILITY OF MODIFYING VARIABLES

For a modifying variable to be useful in reducing reaction to noise, that variable must satisfy three criteria.

1. The modifying variable must have an impact on reaction that is of sufficient size to be of practical value.

Many of the given variables do not appear to have sufficiently powerful statistical relationships with reaction. In the cases of sex, education, local groups being seen to have done something, socioeconomic status, home ownership, and marital status, the sizes of effect are typically able to account for only 0.2 to 0.5 percent of the variation in reaction. Indeed, studies have often failed to find statistically significant effects for some of these modifying variables: education (*10,16,19*); sex (*10,16*); and home ownership (*10,16*).

In a few cases, the results diverge considerably. For example, duration of residence has sometimes not had a significant effect despite a large sample size (N = 3,575) (*10*), whereas Fields and Walker (*12*) calculated the effect of duration of residence as equivalent to a 12 dB variation in noise [with noise controlled by its inclusion in the regression analysis (*12*)]. This may reflect differences between the studies such as culture (Australia versus Great Britain) or noise type (aircraft versus railway). However, duration of residence does not seem to have a consistently large effect. The effect of age varied similarly across the same two studies—the correlation with reaction was 0.50 with noise controlled (*12*) versus 0.10 (*10*). However, a number of studies have reported small effects of age (*8,11*), or even no significant effect (*16,19*). Finally, number of people in the household has produced results ranging from a correlation of 0.45 (*12*) to a nonsignificant effect (*16*).

The sizes of effect for personality, psychiatric status, and the effect of more openable and closable windows are either small and difficult to determine precisely because few studies have examined these variables. In the latter case, the study was of work rather than residential buildings (*18*). However, in principle any control over the level of noise actually heard in the house should bring substantial psychological benefits, judging from the extensive literature on learned helplessness, uncontrollability, and its stress effects (*23,24*). This variable deserves further investigation.

A total of 11 variables appear to have correspondences with community reaction that are of practically significant size. These variables, along with evidence of their sizes of effect, are presented in Table 1.

2. The modifying variable must be a causal factor in community reaction, or eventually be considered to be part of reaction.

In either of these cases, the variable is usable because a decrease in the modifying variable will either cause or constitute a decrease in reaction. However, there are two other ways in which a correlation between an apparent modifying variable and reaction could arise. First, the correlation could occur as an inadvertent consequence of mutual correlations with a third variable. That is, the modifying variable and

TABLE 1 THE 11 MODIFYING VARIABLES HAVING CORRESPONDENCES WITH REACTION

Variable	dB Equivalent Size of Effect	Correlations with Reaction
Attitude to the noise source	15 (*2*)	0.50[a] (*12*) and 0.41 (*6*; review of 12 studies)
	8 (*12*)	0.46[a] (*12*); 0.43 (*19*); 0.34[a] (*25*); and 0.33 (*19*)
Sensitivity to noise	9 (*12*)	0.45[a] (*12*); 0.30 (*6*; review of 11 studies) 0.27 (*21*); and 0.18 (*19*)
Reduction in rates	15 (*12*)	0.46[a] (*12*); 0.43 (*19*); 0.34[a] (*25*); and 0.33 (*19*)
Satisfaction with neighborhood	13 (*12*)	0.51 (*26*)
Fear of crashes	6 (*12*)	0.52/0.45[a] (*25*); 0.61 (*10*); and 0.46 (*12*).
Whether the neighborhood has complained	9 (*12*)	
Satisfaction with life	10 (*12*)	0.46[a] (*12*)
Sight of the noise source	10 (*20*)	
Dissatisfaction with dust, air pollution, lights, loss of privacy	26 (*12*)	0.6[a] (*12*)
Belief that noise harms health	14 (*12*)	0.38/0.33[a] (*25*) and 0.48 (*12*)

[a]Noise controlled for.
NOTE: Corresponding sources of data are given in parenthesis.

reaction could correlate because both are independently caused, at least in part, by some third variable. Second, reaction may be a causal agent to the modifying variable. In both these cases, the modifying variable is of no value in reducing reaction.

Correlations in no way identify the chain of causality. So, the size or the direction of the correlation is not informative on this matter. In fact, with most of the variables in Table 1 there are few data that indicate the likely causal sequence. However, a few studies presented data that bear on this issue. Bullen et al. (*9*) found an initially significant correlation between duration of residence in the area and reaction. However, when the effects of age and home ownership were statistically controlled, the reaction–duration of residence correlation disappeared, suggesting that its impact only arose from the age and home ownership factors. Similarly, the correlation between reaction and education disappeared (*9*). Finally, in a comprehensive study, Fields and Walker (*12*) assessed the modifying variables independently and simultaneously, so as to take into account the complex interrelationships. They found that the impacts of the modifying variables were generally greatly reduced. These results suggest that many of the modifying variables intercorrelate in a manner that allows that their apparent impacts on reaction arise in part from the impacts of other modifying variables with which correlations are shared. Because the required statistical analysis (e.g., regression) to establish correlations independent of the other identified potential factors is rarely done, it is possible that quite a number of the relevant factors do not have independent correlations with reaction.

However, it is possible to speculate that a number of correlations exist between the modifying variables. People no longer studying or further advanced in their careers are likely to earn more (so, age is an income correlate); middle-aged people are more likely to be married than young adults say 18 to 25 years (so, age and marital status will be related); older people have had more time to save money, and earn more money [so, home ownership is related to age as supported by the data (*9*)]; age is also known to influence attitude; older people have had more time to live in one place, and are more likely to own their home so they will move around less often (so, age and length of residence will correlate); home ownership influences duration of residence [again supported by the available data (*9*)]; a person who likes the neighborhood is more likely to stay (so, satisfaction with the neighborhood and duration of residence will correlate); people are more likely to buy a house in an area they like (so, satisfaction with the neighborhood and home ownership will correspond); dissatisfaction with dust, air pollution, and intruding lights could be expected to influence satisfaction with the neighborhood, so these factors should correlate; age of dwelling could be expected to correlate with age of the occupants and length of residence. Further, unmeasured factors may cause some of the observed relationships. For example, it may be speculated that a general predisposition to feel dissatisfied or to express dissatisfaction could lead to correlations between reaction, level of satisfaction with the neighborhood, and level of satisfaction with life (*12*). Finally, additional correlations could arise from the previous correlations: if factor A correlates with factor B and factor B correlates with factor C, a correlation between A and C may result. A number of possibilities of this form exist in the given instances.

The lack of data on intercorrelations of modifying variables and the lack of comprehensive regression analysis have necessitated these speculative arguments. This problem would be reduced if future socioacoustic studies reported regression analysis or the correlations between modifying variables. Given the absence of complete data, the best procedure may be to proceed with only those variables for which there is no obvious alternative explanation, or for which the size of the effect is so large as to be inexplicable in terms of the possible alternative variables through which the impact may occur. The variable of age of dwelling may be eliminated on these grounds. It also seems likely that a substantial part of the impact of satisfaction with the neighborhood and satisfaction with life (and perhaps even some parts of attitude) may be explained in terms of general predisposition to be dissatisfied or satisfied, and the general predisposition to express that feeling. However, these variables remain in contention because there is no evidence on the size of this hypothetical effect and the variables all have strong relationships with reaction (see Table 1).

Although other cases may be unclear and future data and analysis may eliminate more variables on these grounds, overinclusion at this point is likely to be a less costly error than missing potentially usable variables in the battle to reduce community reaction to noise.

The second half of the present criterion involves elimination if it is likely that reaction affects the modifying variable rather than vice versa. Of course it is possible that a complex interaction is occurring: for example higher reaction to the noise may cause more neighborhood complaints and more political pressure, and lead to reduced taxes; then reduced taxes may lead to greater satisfaction and acceptance of the noise [although cognitive dissonance (*27*) could cause an effect in the opposite direction]. Thus, in the case of reduced taxes it seems likely that the variable is influenced by and influences reaction. However, in the case of neighborhood complaints it seems likely that they are caused by reaction rather than vice versa.

It also seems plausible to suggest that dissatisfaction with noise would cause dissatisfaction with the neighborhood and to some extent with life in general. Therefore, the latter two variables are eliminated because they may be caused in some part by reaction itself as well as by dissatisfaction with other aspects of the area—dust, air pollution, etc.—covered in another variable that has a stronger relationship with reaction (*12*). Additionally, as mentioned earlier, dissatisfaction with neighborhood and life may correlate with reaction through a general predisposition to be dissatisified or to express dissatisfaction.

Some relevant data are available on the variables in Table 1 not yet considered in relation to direction of causality. First, as indicated in Table 1, the variables not yet considered have statistically significant correlations (generally 0.3 to 0.5) with reaction even when the effect of noise exposure is statistically controlled by inclusion in the regression analysis. If these variables were themselves influenced by reaction (which correlates with noise exposure), the correlation between the modifying variable and reaction might be expected to drop substantially when the variance attributable to noise is removed. Thus, the significant correlations (independent of the effects of noise) are consistent with the possibility that these variables modify reaction, rather than are caused by it. Nonetheless, these data are suggestive rather than compelling. More detailed analysis would be useful. The necessary data are often unavailable and in the absence of data to the contrary it will be assumed that reduction in rates and the belief that noise harms health may be genuine modifiers of reaction. However, additional analysis is possible for several variables.

- Attitude. The data on attitude are complicated by the fact that a number of different types of attitude have been tapped in the various studies and it is possible that these different variables behave differently. Furthermore, two (and sometimes three) distinct attitude variables have been identified in single studies using factor analysis (*9,10,19*). Thus, the different attitude scales do not appear to reflect a simple underlying general attitude. Nonetheless, across the studies the low mean correlation between attitude and noise suggests that attitude is not simply part of or caused by noise-induced reaction. However, the slight but positive noise-attitude correlations [mean = 0.15 (*6*)] suggest that a small proportion of attitude may be part of, or caused by, reaction.

Although the possibility that reaction affects attitude rather than vice versa cannot be eliminated directly by correlational data, McKennell's data (*28*) are informative. McKennell's analysis of regression slopes indicated that patriotism regarding the Concord was associated with reduced annoyance with Concord overflights to an equal extent, regardless of noise exposure. Thus, the patriotism effect does not appear to be caused by greater noise exposure (through greater reaction, which could then cause less patriotism). However, it is still possible that more noise-sensitive respondents were more annoyed to begin with and, therefore, became less patriotic regarding the Concord. However, McKennell's data also tell against this possibility because the effect of sensitivity, unlike patriotism, changed with noise exposure. On the other hand, in a study of artillery noise (*9*), attitude and sensitivity were related in a manner that suggested that reaction may affect attitude, although other explanations are possible. Overall, it would appear that attitude can be a genuine factor influencing reaction, but on occasion attitude may also be influenced by reaction. In recognition of the potentially powerful impact of attitude on reaction, attitude has been suggested as a cause of the different levels of reaction to the same equal-energy noise exposure levels (*29*) and as an account of the larger-than-predicted changes in reaction that occur with changes in noise exposure (*30*).

- Sensitivity. If sensitivity to noise is caused by reaction, a positive correlation between reaction and sensitivity combined with a positive correlation between reaction and noise could be expected to lead to a positive correlation between sensitivity and noise. On the other hand, if sensitivity is a genuine underlying factor of the respondent, it should bear no relationship to noise exposure, or its only relationship to noise exposure should arise from self-selection. That is, more noise-sensitive people are less likely to live in high-noise areas, leading to a negative noise-sensitivity correlation. A recent review (*6*) revealed a near-zero average noise-sensitivity correlation (−0.01). Similar results have been reported since, with the two sensitivity factors having insignificant correlations of −0.05 and −0.003 with noise exposure (*19*). The conclusion that sensitivity is a genuine modifier of reaction is not new (*10,31*).

- Fear of Crashes. Fear of crashes generally exhibits a high correlation with reaction even with effects of noise exposure statistically removed (0.46 and 0.45) (*12,25*). In another study of aircraft noise, the fear-reaction correlation was 0.61 compared with a fear-noise correlation of 0.22 (*10*). However, an additional positive fear-sensitivity correlation complicated the issue. If fear was a feature of respondents, a zero-to-negative fear-noise correlation might be expected because of self-selection, as with sensitivity. However, this is not the case. Further, the fear-sensitivity correlation suggests that fear arises at least in part from reaction (or is part of reaction), which is then exacerbated by sensitivity to the noise. Fear would seem, in principle, to be an unpleasant state for a respondent to live in as a result of hearing aircraft or railway noise, and therefore it is suggested that fear be considered part of reaction.

• Seeing the Noise Source. The study by Fidel et al. (*20*) found that respondents who could see the source of the noise from their home were more affected by the noise. Because this study was conducted with noise exposure deliberately held constant across respondents, this finding cannot be explained in terms of noise-induced reaction.

The finding that dissatisfaction with air pollution, intruding lights, and loss of privacy has a large impact on reaction (equal to 26 dB) may reflect largely the same variable. That is, only in line of sight would loss of privacy and intruding lights occur. Therefore, these factors are treated as one variable.

3. The modifying variable must be potentially manipulable. If a modifying variable is to be used to reduce reaction, it must be possible to manipulate the level of that variable in an individual. The remaining candidates after the second criterion are now considered from this point of view.

• Attitude. It is clear that attitude can be manipulated. A great deal of successful propaganda and advertising in the commercial sphere and in health promotion (*32,33*) attests to the manipulation of attitude and behavior change. This could also be achieved in relation to noise sources, but would be best achieved with research into the target population and the underpinnings of their current attitudes. However, examination of the known attitude variables in socioacoustic investigations yields some possible attitude altering messages.

McKennel (*28*) found that degree of patriotism influenced reaction to Concorde overflights in Great Britain. In such cases, patriotism may be increased by pro-British advertising, and the direct connection to the Concorde as a British product may also help.

The most commonly measured attitude variables appear to be the following:

—Misfeasance. The belief that those in a position of authority regarding the noise source do not care about the noise they cause.

—Value. The extent to which the noise source is seen as providing a valuable function. For example, is the airport seen as an unnecessary luxury for the wealthy, or a valuable transport link for holiday makers, goods, and the influx of tourist dollars? Is an artillery range seen as pointless and outdated in modern warfare or a critical link in national defense?

—The extent to which the noise souce is seen as having costly disadvantages to the community (intrusion into privacy, air pollution, etc.) is also sometimes measured as an attitude variable. (This variable will be covered separately in a later section.)

A media or mailbox campaign could be designed to counter these attitudes. For example, in the case of an airport, misfeasance may be reduced directly by having letters addressed to noise-exposed residents from the relevant authorities. These letters could state the awareness and concern of the authorities with the noise, and the measures being taken to alleviate noise (noise reduction measures in aircraft, night-time curfews if they exist, etc.). Further, the letter could explain noise events that are seen as pointless: the safety value of engine spool-up, the determination of runway use by prevailing weather conditions, etc.

The perceived value to the community of an airport may also be addressed in terms of the community's common use of air transport, air mail, goods transported in or out, employment, and tourists carried with the consequent flow of money to that community. A similar case could be made for railway use in terms of employment, goods transported, commuter travel, saving energy, and air pollution. Although some of these points apply to road noise, it is probably more difficult to defend in terms of community value because excessive noise is often seen as the responsibility of particular drivers. This viewpoint may in fact, explain the higher level of community reaction to road noise (*29*).

• Sensitivity. It would seem unlikely that noise sensitivity can be readily manipulated.

• Reduction in Rates (Taxes). Clearly, a reduction in rates is possible for those in high-noise areas. However, it may be better to supply a rebate that goes to the resident rather than to the landlord in the case of rented property, even though home owners are more affected by noise than tenants.

A better mechanism may be to relate the monitary benefit more directly to the noise source. This procedure would mean that the noise maker pays for the measure, and the compensation for the noise is more saliently associated with the noise source. The compensation may really be small in monetary terms, but potentially effective in terms of reduced reaction according to the available data (*12*). Such a psychological effect may arise in part because of reduced misfeasance, as discussed earlier. As examples of more directly connected compensations, noise-exposed residents may be offered reduced airfares or departure taxes, reduced train ticket prices, or reduced vehicle registration depending on the noise source of concern.

• Fear of Crashes. Fear of crashes may be reduced by appropriate education or propaganda. A detailed analysis would indicate that the chance of a train or aircraft crashing into any given house is exceedingly remote. Driving home may be far more dangerous than living in the house. However, because the use and reduction of fear are often difficult to achieve in mass media campaigns (*33*), this suggestion would require further research.

• Seeing the Noise Source or Intruding Lights and Loss of Privacy. Because it appears that these factors are relevant to community reaction, visual barriers such as trees around power plants, factories, or along roads and rail lines may be effective in reducing reaction at much less expense than alternative measures. The effectiveness of such a measure may be increased by messages (mass media campaigns or letters to residents) to the community that the measure is being implemented for them.

• Noise Harms Health. The belief that noise harms health is a difficult issue to deal with, because categorical statements to the contrary cannot be substantiated. Although some data point to stress-related effects of noise, the issue of pathology as a consequence of noise has not been settled, largely because of the lack of rigorous large-scale studies (*34*). Therefore, explicit reassurance of the lack of health effects of noise is not justified at this time.

COST EFFECTIVENESS

An apparent advantage of attempting to reduce community reaction by the means suggested rather than by noise reduction or insulation has the potential for real gains at considerably reduced cost. It is much less costly to send letters and information to a residence than to sound-attenuate that residence. It is likely to be less expensive to offer reduced fares or relevant taxes than to attenuate sound. Planting trees for a line of sight barrier (with very slight sound attenuation) is likely to be less expensive and more ecologically sound than constructing noise barriers.

The potential disadvantage of attitude manipulation is that the manipulation may have to be repeated from time to time if the effects dissipate or when new residents enter the area. The level of benefit achieved by the means suggested, and, in the case of attitude, the longevity of the effect require direct future investigation.

Two counterarguments may be anticipated. First, it may be suggested that these effects are not real, and indeed there is no absolute evidence for the causal link suggested. Although it is true that the evidence is not absolute, there is no reason to believe that psychological effects are any less real than acoustic effects. The sizes of effect for attitude and other variables indicate that they have the potential to produce substantial reductions in reaction. The potentially high cost effectiveness justifies trials on this issue. At the least, emphasis on measures to change attitude and line of sight to the source should, where possible, accompany any noise reduction measures. This tactic is likely to increase the size of the reaction reduction achieved.

Second, it may be suggested that the onus to reduce reaction should be on the noise makers; they have an obligation in principle to reduce the noise. The measures suggested here would not remove whatever onus exists on the noise maker. However, the ultimate aim (in the absence of evidence for harmful effects of community noise other than through reaction) is to reduce reaction. Any available means should be explored to achieve that end. Although changing the hearer (in the case of attitude) may seem an odd way to achieve this end, it may be better than doing nothing because of the extreme expense involved in noise reduction. Furthermore, environmental considerations (energy and resource conservation) may be better served by the means suggested here than by engineering sound attenuation construction additions. Finally, countermeasures such as insulation against house noise may actually have a significant part of their impact through attitude (misfeasance) change rather than noise change. This possibility is supported by the obvious concern about noise indicated by noise-insulating the residence, which may therefore substantially alter misfeasance, and the not uncommon finding that outside noise levels are better predictors of reaction than inside noise levels.

CONCLUSIONS

The strong relationships between a number of modifying variables and community reaction suggest that these variables may be useful in reducing reaction. Although more research is needed to identify with greater confidence the direction of causality, the data suggest that some modifying variables do modify reaction. Therefore, the manipulation of these variables is likely to reduce reaction. This possibility warrants direct testing.

REFERENCES

1. F. J. Langdon and I. D. Griffith. Subjective Effects of Traffic Noise Exposure, II: Comparisons of Noise Indices, Response Scales, and the Effects of Changes in Noise Exposure. *Journal of Sound and Vibration*, Vol. 83, 1982, pp. 171–182.
2. A. L. Brown, A. Hall, and J. Kyle-Little. Response to a Reduction in Traffic Noise Exposure. *Journal of Sound and Vibration*, Vol. 98, 1985, pp. 235–246.
3. G. J. Raw and I. D. Griffiths. The Effect of Changes in Aircraft Noise Exposure. *Journal of Sound and Vibration*. Vol. 101, 1985, pp. 273–275.
4. I. D. Griffiths and G. J. Raw. Community and Individual Response to Changes in Traffic Noise Exposure. *Journal of Sound and Vibration*, Vol. III, 1986, pp. 209–217.
5. S. Fidell and L. Silvati. Laboratory Tests of Hypotheses Derived from a Decision-Theoretical Model of Noise Induced Annoyance. *Proc., Inter-Noise 1989*, Newport Beach, Calif., Noise Control Foundation, Poughkeepsie, N.Y., 1989, pp. 887–900.
6. R. F. S. Job. Community Response to Noise: A Review of Factors Influencing the Relationship Between Noise Exposure and Reaction. *Journal of the Acoustical Society of America*, Vol. 83, 1988, pp. 991–1001.
7. E. Grandjean, P. Graf, A. Lauber, H. P. Meier and R. Muller. A Survey of Aircraft Noise in Switzerland. *Proc., International Congress on Noise as a Public Health Problem*, Dubrovnik, Yugoslavia, 1973, pp. 645–660.
8. J. S. Bradley. Field Study of Adverse Effects of Traffic Noise. *Proc., 3rd International Congress of Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 571–579.
9. R. B. Bullen, A. J. Hede and R. F. S. Job. Community Response to Artillery Noise (in preparation).
10. R. B. Bullen, A. J. Hede, and E. Kyriakos. Reaction to Aircraft Noise in Residential Area Around Australian Airports. *Journal of Sound and Vibration*, Vol. 108, 1986, pp. 199–225.
11. R. Rylander, S. Sorensen and A. Kajland. Annoyance Reactions for Aircraft Noise Exposure. *Journal of Sound and Vibration*, Vol. 47, 1972, pp. 237–242.
12. J. M. Fields and J. G. Walker. The Response to Railway Noise in Residential Areas in Great Britain. *Journal of Sound and Vibration*, Vol. 85, 1982, pp. 177–255.
13. P. N. Borsky. Integration of Multiple Aircraft Noise Exposures Over Time by Residents Living Near U.S. Air Force Bases. *Proc., 4th International Congress on Noise as a Public Health Problem*, Turin, Italy, 1983, pp. 1049–1060.
14. J. Francois. Aircraft Noise, Annoyance, and Personal Characteristics. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 571–579.
15. A. Tarnopolsky, D. J. Hand, S. M. Barker, and L. M. Jenkins. Aircraft Noise, Annoyance and Mental Health: A Psychiatric Viewpoint. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 588–593.
16. A. Lawrence and A. Putra. The Combined Effect of Road Traffic and Aircraft Noise on People. *Proc., Inter-Noise '89*, Newport Beach, Calif. Noise Control Foundation, Poughkeepsie, N.Y., 1989, pp. 891–894.
17. M. Vallet, M. Marin, M. A. Page, B. Favre, and G. Pachiaudi. Annoyance from and Habituation to Road Traffic Noise from

Urban Expressways. *Journal of Sound and Vibration*, Vol. 60, 1978, pp. 423–440.
18. R. B. Bullen, R. F. S. Job, and D. H. Burgess. *Reaction to Aircraft Noise on RAAF Bases*. National Acoustic Laboratories Commissioned Report, No. 7. Australian Government Publishing Service, Canberra, 1985.
19. R. F. S. Job and A. J. Hede. Community Reaction to Noise From Power Stations. *Proc., Inter-Noise 1989*, Newport Beach, Calif., Noise Control Foundation, Poughkeepsie, N.Y., 1989, pp. 865–868.
20. S. A. Fidell, S. R. Teffeteller, and K. S. Pearsons. *Initial Study on the Effects of Transformer and Transmission Line Noise on People: Volume 3: Community Reaction*. Electric Power Research Institute, Palo Alto, Calif., 1979.
21. Tracor, Inc. *Community Reaction to Aircraft Noise*. Research Report CR1761. U.S. National Aeronautics and Space Administration, 1971.
22. G. A. Luz, R. Raspet, and P. D. Schomer. Analysis of Community Complaints to Noise. *Journal of the Acoustical Society of America*, Vol. 73, 1983, pp. 1229–1235.
23. S. F. Maier and M. E. P. Seligman. Learned Helplessness: Theory and Evidence. *Journal of Experimental Psychology: General*, Vol. 105, 1976, pp. 3–46.
24. T. R. Minor, N. K. Dess, J. B. Overmier. Inverting the Traditional View of the "Learned Helplessness" Paradigm: Reinterpretation in Terms of Anxiety and Modulation Operations. In M. R. Denny, (ed.), *Aversive Events and Behavior*, Lawrence Erlbaum, Hillsdale, N.J., in press.
25. A. C. McKennell. Psycho-Social Factors in Aircraft Noise Annoyance. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Dubrovnik, Yugoslavia, 1973, pp. 562–566.
26. B. Rohrmann, R. Schumer, A. Schumer-Kohrs, R. Guski, H. I. Finke. An Interdisciplinary Study on the Effects of Aircraft Noise on Man. *Proc., International Congress on Noise as a Public Health Problem*, Dubrovnik, Yugoslavia, 1973, pp. 765–776.
27. R. A. Baron and D. Byrne. *Social Psychology* (4th ed.). Allyn & Bacon, Boston, Mass., 1984.
28. A. C. McKennell. Annoyance from Concorde Overflights Around Heathrow. *Proc., 3rd International Congress on Noise as a Public Health Problem*, Freiburg, West Germany, 1978, pp. 562–566.
29. J. G. Walker. Railway Noise Exposure: A Possible Method of Establishing Criteria for Acceptability. *Journal of Sound and Vibration*, Vol. 120, 1988, pp. 347–349.
30. R. F. S. Job. Over-Reaction to Changes in Noise Exposure: The Possible Effect of Attitude. *Journal of Sound and Vibration*, Vol. 126, 1988, pp. 550–552.
31. F. J. Langdon. Noise Nuisance Caused by Road Traffic Noise in Residential Areas. *Journal of Sound and Vibration*, Vol. 49, 1976, pp. 241–246.
32. G. W. Ward. How to Sell Health. *World Health Forum*, Vol. 7, 1986, pp. 169–177.
33. R. F. S. Job. Effective and Ineffective Use of Fear in Health Promotion Campaigns. *American Journal of Public Health*, Vol. 78, 1988, pp. 163–167.
34. D. M. DeJoy. A Report on the Status of Research on the Cardiovascular Effects of Noise. *Noise Control Engineering Journal*, Vol. 23, 1984, pp. 32–39.
35. A. K. Dixit and J. O. Reburn. Community Reaction to Railway Yard Noise. *Proc., Inter-Noise 1980*. Noise Control Foundation, Poughkeepsie, N.Y. 1980.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Criteria for the Design of Sound Insulation in Homes Around Commercial Airports

ERIC STUSNICK AND JOHN E. WESLER

Current FAA regulations provide for FAR Part 150 funding for the noise insulation of residences, schools, hospitals, churches, and other approved noncompatible structures located within the 65-dB yearly day-night average sound level (DNL) contour. For residences located where the exterior DNL is 65 dB, a noise level reduction (NLR) of at least 20 dB is required in major habitable rooms. The requisite NLR is increased commensurate with any increase in exterior DNL above 65 dB. This is mathematically equivalent to achieving an interior DNL of 45 dB in major habitable rooms. Although the use of the DNL metric may be appropriate in determining eligibility for funding because it has been found to be correlated to the community reaction to environmental noise, experience has indicated that an individual homeowner's annoyance with the noise from aircraft overflights is more closely related to the average sound exposure level (SEL) of overflights. The relationship between DNL and SEL is examined as a function of the number of aircraft operations. It is demonstrated that, for a given value of DNL, the average value of the allowed SEL increases as the number of operations decreases. Thus the use of an interior DNL metric to determine the NLR criterion for homes around airports results in higher average interior SEL values in homes around smaller airports than in homes around large airports. An alternate SEL criterion is proposed.

Three different acoustic metrics are commonly used to describe the noise exposure from aircraft overflights:

- Day-night average sound level (DNL),
- Sound exposure level (SEL), and
- Maximum A-weighted sound level (ALM).

Each of these metrics is a sound pressure level measured in units of decibels (dB) and each deals with the A-weighted sound level. An A-weighted measurement of an acoustic signal adjusts the level to take into account the fact that the human ear is not equally sensitive to all frequencies of sound—it being most sensitive to frequencies between 1,000 and 6,000 Hz and progressively less sensitive at frequencies farther from that range.

The DNL value is a sound level corresponding to the average A-weighted sound energy that is received during an entire 24-hr period, giving a 10-dB penalty to noise occurring at night (10:00 p.m. to 7:00 a.m.). SEL and ALM are single-event metrics. The former is a level corresponding to the total A-weighted sound energy that occurs during an overflight; the latter is the maximum A-weighted sound level that occurs during an overflight.

Wyle Laboratories, 2001 Jefferson Davis Highway, Suite 701, Arlington, Va. 22202.

Because SEL is the sound level that would contain, in 1 sec, the same A-weighted sound energy as did the actual overflight (which is always longer than 1 sec), the value of SEL for a given event is always greater than the value of ALM for the same event. On the other hand, because DNL contains both quiet and noisy periods in its average, the value of DNL is lower than typical SEL or ALM values for individual overflights that occur during the 24-hr period.

The physical difference between these three metrics should be clearly understood. The public is often confused by discussions of DNL values between 65 and 75 dB when they have seen exterior ALM measurements at their homes that are typically 20 dB higher and occasionally as much as 40 dB higher. The reason for the difference, of course, is that each of the metrics describes a different aspect of the noise.

DAY-NIGHT AVERAGE SOUND LEVEL

DNL is the acoustic metric that has been chosen by the FAA to determine a structure's eligibility for inclusion in a federally supported sound insulation program. The October 24, 1989, edition of the *Airport Improvement Program (AIP) Handbook* states: "Unless specifically justified by the airport sponsor in its NCP and approved by the FAA, the structure must be located within a DNL 65 dB contour." The NCP is the FAA-approved noise compatibility program for the airport.

The AIP Handbook further states: "Normally, unless extenuating circumstances dictate, noise insulation should not be considered for structures within a DNL 75 dB or greater noise contour since it is preferable to change the land use."

The FAA also uses DNL to define the minimum acoustic insulation that should be provided to a residential structure taking part in a federally supported sound insulation program. The AIP Handbook states:

> For residences located in areas where exterior noise exposure is DNL 65 dB, the requisite noise level reduction (NLR) provided by the structure should be at least 20 dB in major habitable rooms. The requisite NLR should be increased commensurate with any increase in exterior DNL above 65 dB.

This condition is mathematically equivalent to requiring an NLR that will produce an interior DNL value of 45 dB or less.

Although the normal program eligibility requirements are fixed (an exterior DNL value between 65 and 75 dB), the NLR requirement is a lower limit. Acoustic insulation that produces a greater NLR value can be provided without violating the conditions in the AIP Handbook. In fact, the Handbook goes on to state: "Since it takes an improvement of at least 5 dB in NLR to be perceptible to the average person, any residential noise insulation project will be designed to provide at least that increase in NLR."

Of course, providing a higher NLR value than the minimum required by the AIP Handbook should be justified.

The DNL metric has been used for two different purposes: (a) to define a structure's eligibility for the program, and (b) to define the minimum noise level reduction that must be provided.

The DNL value is an appropriate metric to use for these purposes because it has been shown to correlate fairly well with community reaction to environmental noise (*1,2*) and it is related to the total acoustic energy received throughout a 24-hr period. As a result, the DNL value takes into account both the sound level of typical overflight events and the number of events that occur. A structure experiencing a large number of overflights in a 24-hr period will have a higher DNL value than a structure experiencing a small number of similar events in that period.

SINGLE-EVENT METRICS

Although the DNL value correlates with community reaction to noise, the relation is statistical in nature, not causal (*1,2*). That is, an individual's annoyance is caused by the interference of specific noise events with some human activity, such as speech or sleep. Questionnaires administered by Wyle Laboratories to residents of homes before their acoustic insulation have indicated the most disturbing aspects of aircraft overflight noise to be interference with radio and television listening and with telephone and general conversation. These activities are followed by interference with general relaxation and concentration and by sleep disturbance.

Whether or not activity interference will occur is best determined by a single-event metric. Because the DNL value does depend on both the sound level of a typical overflight event and the number of events, it is not a good estimator of single-event sound levels.

In order to clarify this statement, note that the relation between the DNL at a given point and the daily mean value of SEL for aircraft overflights at that point is

$$DNL = \langle SEL \rangle + 10 \log (N_{eff}) - 49.4$$

where

$\langle SEL \rangle$ = mean value (on an energy basis) of SEL, and
N_{eff} = number of effective daily operations = $N_d + 10N_n$.

Here N_d is the number of daytime operations and N_n is the number of nighttime operations. (Multiplying the number of nighttime operations by a factor of 10 is equivalent to penalizing nighttime sound levels by adding 10 dB.)

For a given value of exterior DNL, the corresponding value of exterior <SEL> will increase as the number of effective daily operations decreases, as presented in Table 1. Thus, while a structure located on the DNL 65 dB contour around a large airport might experience a mean exterior SEL value of 87.5 dB, a structure located on the same DNL contour around a smaller airport might experience a mean exterior SEL value that is 5 to 10 dB higher. Providing the minimum required NLR in each of these situations results in mean interior SEL values from 67.5 dB at large airports to 77.5 dB at small airports.

TABLE 1 AVERAGE SEL AS A FUNCTION OF DNL AND N_{eff}

Exterior DNL (dB)	N_{eff}	Exterior <SEL> (dB)	Minimum NLR (dB)	Interior <SEL> (dB)
65	500	87.5	20	67.5
	100	94.5	20	74.5
	50	97.5	20	77.5
70	500	92.5	25	67.5
	100	99.5	25	74.5
	50	102.5	25	77.5
75	500	97.5	30	67.5
	100	104.5	30	74.5
	50	107.5	30	77.5

The physical interpretation of this is as follows. Because of the larger number of operations on a given runway at a large airport than at a smaller airport, a given DNL contour is farther from that runway at the large airport than it is at the smaller airport. Thus, for an average overflight at the larger airport, the aircraft is farther from a point on that DNL contour than it would be for a similar point on the corresponding DNL contour at a smaller airport. As a result, the mean value of SEL for the overflight at the larger airport is less than the mean value of the SEL for the overflight at the smaller airport.

Clearly, if the goal of the sound insulation program is to prevent activity interference by maintaining an interior value for a single-event metric that is independent of the size of the airport, the structure near the smaller airport should receive more acoustic insulation than does the structure near the larger airport. This would occur if a single-event metric is used for the criterion for acoustic insulation modification designs. It would not occur if the DNL metric is used for this purpose.

This argument also applies to homes near lesser-used runways at large airports. A home situated on a given DNL contour adjacent to such a runway will be exposed to single-event noise levels that are higher than those experienced at a home on the same DNL contour adjacent to the main runway. Thus, to achieve the same interior value for a single-event noise metric in both homes, more acoustic insulation would be required in the home near the lesser-used runway than would be required in the home near the main runway.

A criterion based on a single-event metric does not conflict with the present AIP Handbook as long as it leads to acoustic insulation designs that provide at least the minimum NLR value discussed earlier.

Which, then, is the best single-event metric for the purpose, SEL or ALM, and what is the appropriate value for that metric? As will be discussed in the following paragraphs, both metrics have advantages and disadvantages.

ALM is the maximum sound level during an overflight. Thus, in general, the interior ALM value is easy to measure above the background sound level inside the home, even when the structure has been well insulated and the overflight does not occur directly overhead. However, the ALM value corresponds to the level at a single instant of time. As a result, it is subject to wide variations from overflight to overflight of the same type of aircraft at the same distance because of the varying effects on sound propagation of atmospheric turbulence.

Use of SEL minimizes some of these problems because it measures the energy in the entire overflight (defined as the interval between the time at which the sound level is initially 10 dB below the ALM value for the overflight and the time at which the sound level is again 10 dB below that ALM value). Thus, rapid temporal variations in instantaneous sound level caused by atmospheric turbulence are averaged out. In addition, an NLR value defined as the difference between measured exterior and interior SEL values is more closely representative of an average for the entire structure than is the difference between the exterior and interior ALM values.

However, practically speaking, if the house has been well insulated and if the overflight does not occur directly overhead, the two 10-dB-down points for the overflight may be below the background sound level inside the house. Thus it may be difficult to measure accurately the interior SEL after the structure has been acoustically insulated.

On balance, although it may require more effort to measure the interior SEL value of an acoustically insulated structure than to measure the interior ALM value, the SEL value is superior for use in a criterion for the design of acoustic insulation modifications.

DESIGN CRITERION

What, then, should be the interior SEL goal for the design of acoustic insulation modifications? Studies (*3*) have shown that essentially 100 percent sentence intelligibility is achieved indoors with an A-weighted background level of 52 dB. Thus, if the interior ALM for an overflight is below this level, speech interference will be minimal.

In order to relate this ALM value to an interior SEL value, the FAA's Integrated Noise Model (INM), Version 3.9, was used in a grid-point analysis to obtain SEL values for departures and arrivals of several Stage 2 and Stage 3 aircraft. At the same time, the data base within INM was examined to determine corresponding ALM values for these flight operations.

Figure 1 shows the results of this analysis for a narrowbody Stage 2 aircraft (727), a narrowbody Stage 3 aircraft (737-300), and a widebody Stage 3 aircraft (L-1011). Data have been included for thrusts corresponding to takeoff and to approach. It is seen that all of the data points lie within a 5-dB band.

Figure 1 also shows the least squares linear fit to these data for which ALM is given as a function of SEL by the expression

$$\text{ALM} = 1.17\ \text{SEL} - 24.5 \quad \text{dB}$$

Alternately, SEL is given as a function of ALM by the expression

$$\text{SEL} = 0.85\text{ALM} + 20.9 \quad \text{dB}$$

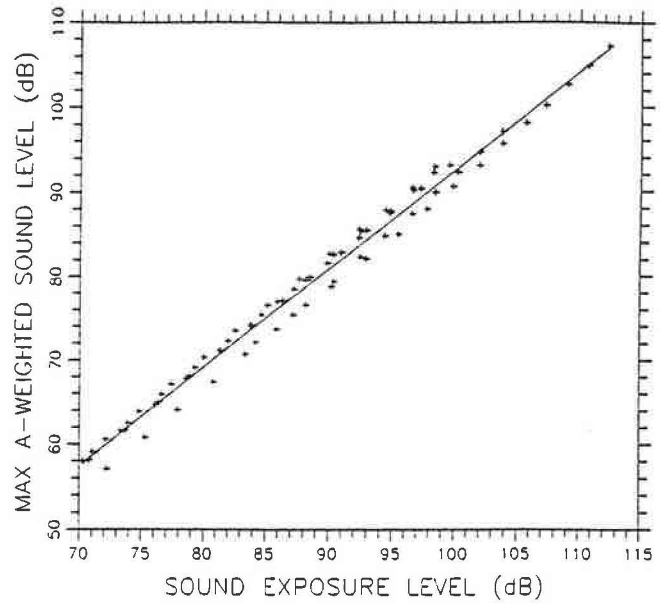


FIGURE 1 Relation between maximum A-weighted sound level (ALM) and sound exposure level (SEL).

Using this relationship, an ALM value of 52 dB corresponds roughly to an SEL value of 65 dB.

Thus, the criteria proposed to be used in developing acoustic insulation modification designs are as follows:

- In all major habitable rooms, provide sufficient acoustic insulation so that the more stringent of the two following interior levels is achieved: DNL = 45 dB and <SEL> = 65 dB, where <SEL> is the mean value, on an energy basis, of the sound exposure levels of all overflights during a 24-hr period.
- In all rooms for which the above criterion indicates that sound insulation modifications are required, increase the NLR by at least 5 dB.

Because the mean value of the daily sound exposure levels is used in these criteria rather than the maximum value, there will still be occasional overflights during which some speech interference occurs. However, past experience has indicated that use of the maximum SEL value, rather than the mean SEL value, as a design criterion leads to NLR goals that are impossible to achieve with commercially available sound insulation materials.

REFERENCES

1. T. J. Schultz. Synthesis of Social Surveys on Noise Annoyance. *Journal of the Acoustical Society of America*, Vol. 64, No. 2, Aug. 1978, pp. 377–405.
2. *Updating a Dosage-Effect Relationship for the Prevalence of Annoyance Due to General Transportation Noise.* HSD-TR-89-009. Human Systems Division, Air Force Systems Command, Brooks Air Force Base, Tex., Dec. 1989.
3. *Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety.* 550/9-74-004. Environmental Protection Agency, Washington, D.C., March 1974.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Techniques for Aesthetic Design of Freeway Noise Barriers

JULIE FARNHAM AND EDWARD BEIMBORN

Guidelines and suggestions for the improved design of freeway noise barriers in urban areas are provided. The guidelines have been developed to encourage an attractive and efficient system of noise control and to achieve design continuity throughout the area. How inventories of the technical, cultural and natural landscape can be used in design is illustrated. A variety of techniques of wall shape and texture and landscaping application that can be adapted to specific environments is provided. The desired approach views the wall and landscape as an integrated system that should be designed to reflect and enhance the surrounding community. Diversity in design of both the wall and landscaping should be sought and systems should be designed to reflect the historic and ethnic heritage of the neighborhoods. A broad range of options should be provided that can creatively enhance the environment that is seen as well as heard along freeways. The concepts developed are illustrated through several prototype designs.

Noise barriers are increasingly used to solve problems of noise pollution from freeways in urban areas. Concern for the negative impacts of noise on the environmental quality of adjacent properties prompted many highway agencies to set maximum allowable noise levels for areas adjacent to freeways. These noise levels are often exceeded in nearly all segments of urban freeway with normal levels of freeway traffic. Consequently noise barriers could be warranted for large portions of the freeway network and there could be a need to build extensive systems of noise barriers over the next decade. As a result of these concerns, many cities and states are in the process of retrofitting their urban freeways with noise barriers. These projects are often implemented with relatively little attention given to their aesthetic impact on the environment.

Noise barriers can have a substantial effect on the visual environment of a highway and its surrounding environment. They are long continuous structures, often more than 10 ft high, made of various materials. They can significantly change the view from the road by blocking the view of the roadside and creating a monolithic uniformity of walls instead of changing urban scenery. They also can change the view towards the road for the surrounding community. There can be an increased sense of the noise barrier as a community barrier that separates an area from other places. Nonetheless, noise barriers may be highly sought after by neighborhoods as a means for reducing noise levels related to freeways. What is needed is a way to provide the benefits of the sound reduction while at the same time creating a positive visual image for road users and the surrounding communities. Ways in which this can be done are suggested.

A larger report (*1*) was prepared in response to a legislative mandate of the State of Wisconsin to examine aesthetic issues related to noise barriers. The report was prepared in cooperation with the Wisconsin Department of Transportation, the city of Milwaukee, and the Southeastern Wisconsin Regional Planning Commission. In addition, previous work in the U.S. (*2–5*) and overseas (*6*) provided a basis for this work.

GOALS AND OBJECTIVES

The goal of this project was to develop a process for planning of noise barriers that enhance and improve the aesthetic quality of the freeway environment. This goal has two primary objectives:

1. To analyze the system-wide area to identify prominent physical and cultural patterns that describe and express the character of urban areas. This objective is based on the fact that barrier locations are not always continuous but must fit into a continuous freeway environment. It is intended that analysis of the system-wide environment will serve as the basis for integrating barriers into the overall urban context.
2. To explore alternative design solutions for barriers in site-specific contexts. Specifically, these designs should accomplish the following:
   a. Integrating the design of new barriers and landscape,
   b. Integrating the new barriers and landscape with existing barriers and landscape.
   c. Designing the new walls as attractively as possible to minimize the need for screening them with landscaping, and
   d. Designing the barrier and landscape to reflect and enhance the specific cultural and physical characteristics of the neighborhoods.

BASIC DESIGN PRINCIPLES

Issues that should be addressed in barrier and landscape design include distance, motion, line, form, scale, balance, rhythm, sequence, and orientation. Each of these factors must be carefully considered to create a design that is compatible with its surroundings.

Distance and motion affect what motorists see when driving through the freeway corridor. Speed alters the peripheral cone of vision and the distance to the motorist's point of focus. In

J. Farnham, Village of Mt. Prospect, 100 S. Emerson St., Mt. Prospect, Ill. 60056. E. Beimborn, Center for Urban Transportation Studies, University of Wisconsin, P.O. Box 784, Milwaukee, Wis. 53211.

general, as speed increases (Figure 1), the cone of vision narrows and the focal distance increases. Likewise, at slower speeds, peripheral vision is expanded and the focal distance is closer to the observer. Vision cones delineate the area within which objects are generally in focus. Objects outside these cones become blurred.

Lines are created by joining two points. Lines are fundamental to three-dimensional forms. Smooth, flowing, horizontal lines may suggest calm and serenity, while bold, vertical, angular lines may suggest strength and tension. The most prominent line created by a noise barrier is the top profile (Figure 2). Lines can be created on the surface of the barrier wall using various materials and texture patterns. Lines may also be implied. For example, a row of street trees may appear to form a solid line.

Form depicts volume in three dimensions—length, width, and depth. A barrier wall itself is a three-dimensional form. This is commonly referred to as positive form. The horizontal wall configuration can be designed to create spaces that are commonly referred to as negative spaces or forms (Figure 2).

Scale establishes a frame of reference. It is a relative measurement. People commonly relate to their environment using the human body or other elements of familiar dimension as the reference measurement. Noise barrier walls can be 15 ft tall or more. Such a wall would appear massive and overpowering located adjacent to a back lot line because residental fences are more commonly 4 to 8 ft tall (Figure 2). The perceived scale of barrier walls can be reduced through the use of landscaping and in the design of the wall itself (e.g., in materials and configuration).

Balance creates order and unity by suggesting a sense of equilibrium. Two basic types of balance are symmetry and

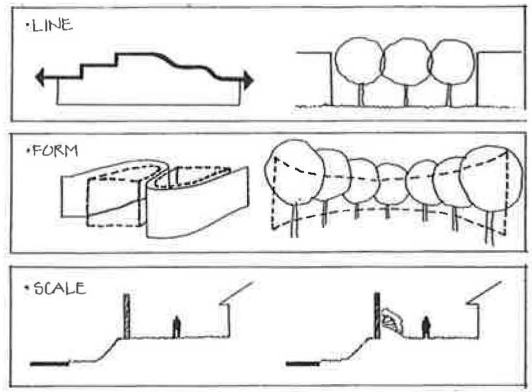


FIGURE 2 Fundamental elements of design.

asymmetry. Symmetry results when elements are arranged equally around a central axis, for example, creating a mirror-image effect. Symmetry is considered to be a formal type of balance. Asymmetry is more informal, lacking a central axis. Elements are juxtaposed in such a way that they counterbalance each other without creating a mirror image.

Rhythm and sequence establish consistent, recognizable patterns. Repeated patterns create a sense of familiarity and comfort. They also provide a sense of progression, unless continued indefinitely. Rhythm and sequence can be created using either the barrier wall or landscaping, or both.

Orientation refers to the dominant visual direction established through design (Figure 3). Horizontal orientation is associated with relatively flat and expansive landforms. This type of orientation tends to direct the eye forward, reducing the apparent height of a barrier wall by emphasizing its re-

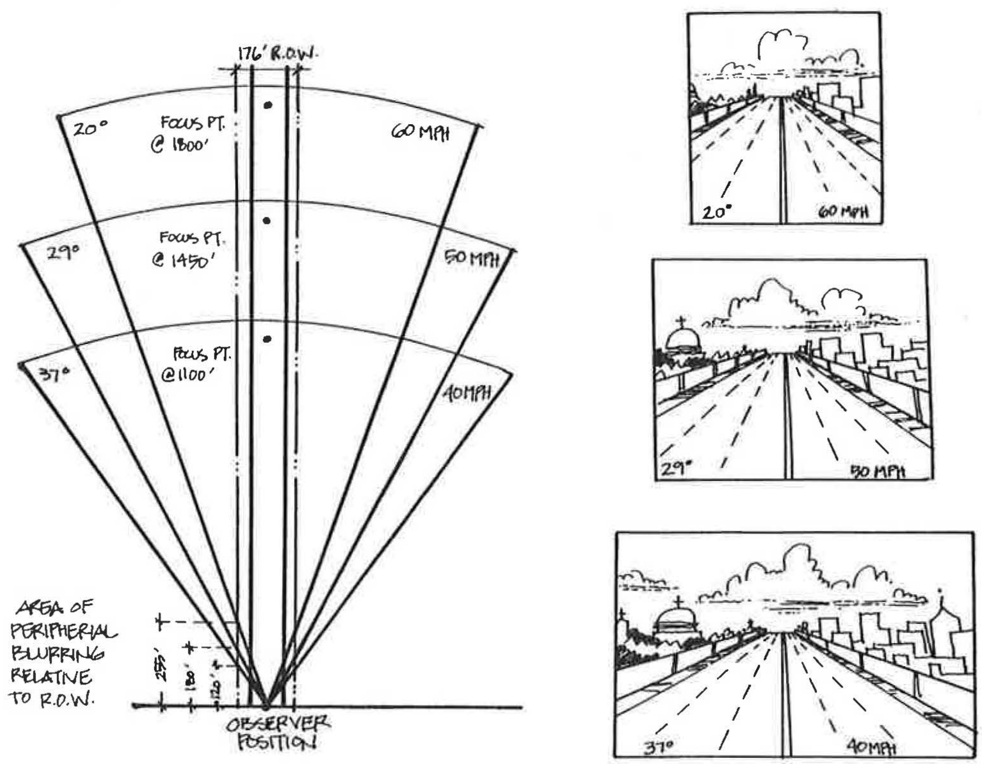


FIGURE 1 Distance and motion.

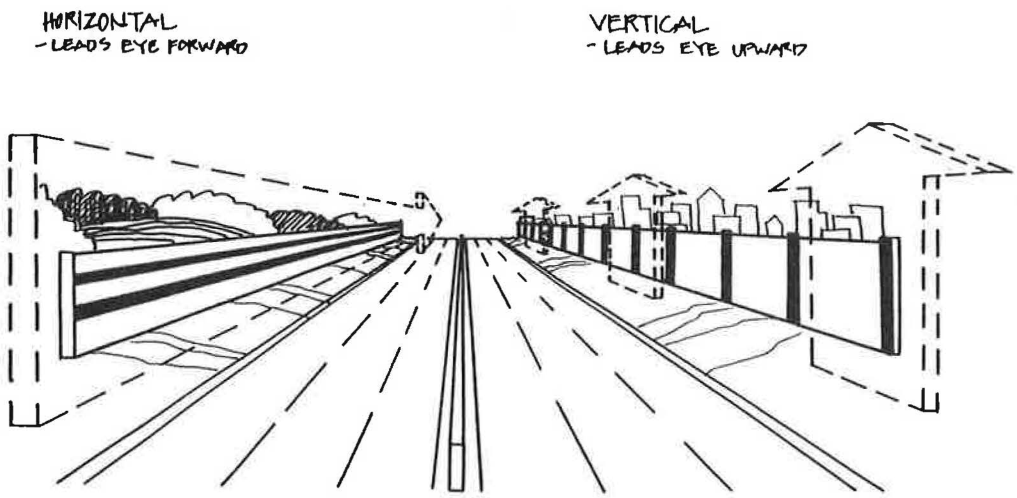


FIGURE 3 Visual orientation.

lationship to the horizon. Vertical orientation is associated with upward reaching forms such as skyscrapers. This type of orientation tends to direct the eye upward, increasing the apparent height of a barrier wall.

INTEGRATION OF WALL, LANDFORM, AND PLANTS

A primary goal in designing noise barriers is to integrate them into the landscape. Problems develop when barrier walls are placed on the landscape with little attempt to integrate them with the surrounding landforms or existing built elements such as bridge abutments, endwalls, and guardrails. As a result, barrier walls can appear as obtrusive objects in the environment. Barrier walls can be integrated with the landscape in two basic ways. The structure itself can be designed to appear to grow out of the landscape, or the landscape (plants and earth) can become part of the barrier structure.

Plants and wall structures can be integrated in a variety of ways. The wall itself can be designed using earth and plants as the primary construction materials. Living barriers, which are used in Western Europe (7), are an example (Figure 4). These essentially vertical earth walls function as the growth medium for willow plants. The earth is contained in a frame constructed of white willow posts interwoven with basket willow twigs. Living barriers provide an attractive alternative to the common barrier constructed of hard materials, both in terms of appearance and maintenance. A similar approach would be to design planter troughs into the wall structure (Figure 5). Irrigation systems could be incorporated into the troughs. The level of planters could be varied to produce a cascading effect. Annuals are commonly used in planters in urban areas and would be appropriate and attractive in wall planters. However, they must be planted annually and are therefore too maintenance intensive for extensive use. Plants can also be integrated with walls by attaching them to the surface or by providing holes in the wall through which they can grow. These approaches are especially appropriate in situations where planting space is limited. Vines planted on the residential side of the noise barrier will eventually cascade over the top. In addition, small holes can be drilled through the wall surface to allow the vines to grow through and spread on the freeway side.

DESIGN IN RELATION TO LOCAL CONDITIONS

These basic principles were used along with additional information on plant materials and barrier wall texture to develop a series of prototype designs for sample locations along the freeways of Milwaukee County. The first step in the devel-

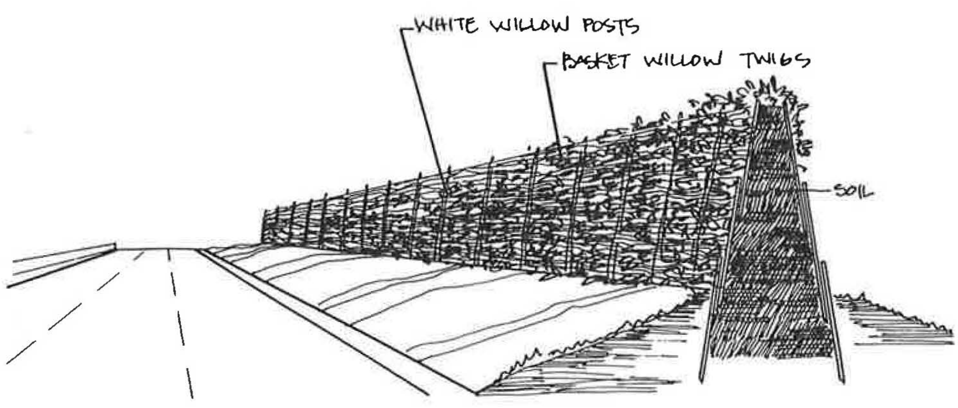


FIGURE 4 Living barrier.

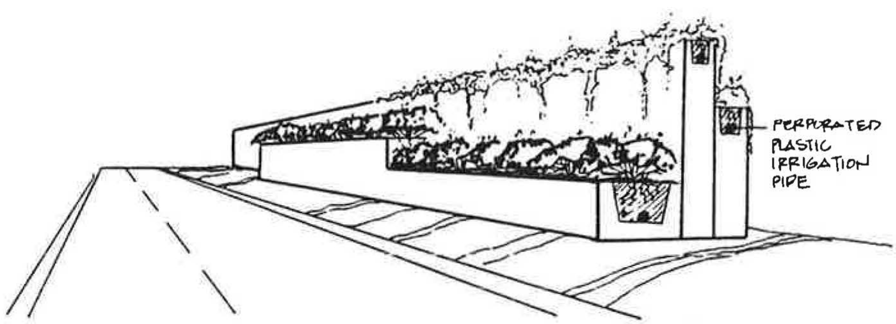


FIGURE 5 Planting troughs nested in the wall.

opment of prototype designs was to identify the opportunities and constraints presented in the community by physical and cultural features. The purpose of the inventory and analysis was to (a) identify regional patterns and features that could potentially impact the visual quality of noise barriers and the freeway corridor; (b) to identify unique cultural, historical, or natural features that could be used to set a theme for the design of a noise barrier–landscaping system, and (c) to identify prototypical situations relative to the location and function of noise barriers.

Both research and a windshield visual analysis were conducted to gather information on the following patterns and features on a county-wide scale:

1. Physical, e.g., topography, vegetation, unique features, and views;
2. Cultural, e.g., ethnic enclaves, landmarks, former uses (historical "ghosts"); and
3. Land use, e.g., urban, suburban, and rural.

Physical Patterns

Analysis was limited to those locations with existing noise barriers and those proposed as candidates for barriers. The most prominent physical features were the existence of mature vegetation in some locations and the visually dominant presence of power lines. Specific locations along the freeway system were identified where existing vegetation provides a significant asset. In many of those areas the vegetation is within the existing right-of-way or near the fence line. A decision must be made as to where (residential or freeway side) a line should be cleared for construction access. Although the most appropriate approach varies by location, a high priority should be given to preserving and using as much of the existing mature vegetation as possible (*8*).

Significant Views

Three locations were identified at which future noise barriers could provide visual obstruction of desirable views:

1. The Milwaukee River crossing. Although the current elevation of the freeway bridge spanning the river precludes expansive views, further minimization of this view opportunity should be avoided.
2. The gate to the city from the south created by the twin spires of a historic church on the west side of the freeway and an industrial clock tower on the east side. Both of these structures are visual landmarks. Driving north, they provide a gateway to the downtown. Driving south, they mark the transition between the industrial heart of Milwaukee and its traditional neighborhoods.
3. The visually prominent dome of St. Josaphat Basilica. The scale and elegant design of this structure signify its importance as a cultural symbol and landmark.
4. The Milwaukee downtown area is a prominent feature when approached from all directions. The view is especially dramatic from the south on the high-rise bridge over the Menomonee Valley.

Cultural Patterns

Cultural patterns and features refer to the location of traditional ethnic neighborhoods along the freeway corridor. This information can be used to extract location-specific design characteristics that could be incorporated into the barrier design. One of the objectives of this study is to provide ideas on how to design noise barriers that better reflect the character of the neighborhoods in which they are located. Substantial effort was made to identify the historic neighborhoods located along the freeway system. Descriptions of the cultural and historic background of neighborhoods were developed (*9*) to provide insight into significant design qualities that could be incorporated into noise barrier design.

Land Use Patterns

The analysis also revealed patterns of land use of adjacent development. The density and age of development contribute significantly to the character of specific locations. The freeway system was separated into zones on the basis of the dominance of particular types of residential development as viewed from the roadway. Three distinct types of development were identified.

- Urban character refers to older residential areas of the city. The freeway cuts through these established neighborhoods, with dense residential development immediately adjacent to the corridor. The design character of these areas is associated with man-made geometric forms, hard edges, hard surfaces, vertical lines, and contrast in forms. An appropriate barrier could be constructed of hard materials, using geometric configurations, with a varied top profile and vertical orientation.

• Suburban character refers to areas of lower density and newer housing. The design character is associated with a mixture of geometric, man-made forms, horizontal lines, and amorphic, naturalistic forms. An appropriate barrier could be constructed of a mix of materials from concrete to wood, using a mix of geometric and curvilinear configurations with a varied top profile. Plants are used primarily to blend, soften, and reduce the scale of the wall.

• Rural character refers to areas that appear to be undeveloped open space or low densities when viewed from the highway. It is unlikely that areas exhibiting actual rural character will be potential candidates for noise barriers. However, some suburban areas have maintained pockets of areas that appear to be rural when viewed from the freeway. The design character is associated with soft, flowing forms. An appropriate barrier could be constructed of natural materials (wood, stone), using curvilinear configurations and a definite horizontal emphasis. Naturalistic planting designs using informal layouts, indigenous plants, and muted color schemes would be most appropriate. Table 1 presents the relationship between the land use type and design considerations.

ECONOMIC CONSIDERATIONS

Retrofitting urban freeway corridors with noise barriers can be an expensive endeavor. For example, costs can be over $2 million per mile for a single side of a freeway. Several factors affect the cost of barriers and landscaping. Materials are probably the most fundamental cost factor. Type of material used affects cost as does the quantity of material used. To provide an estimate of the order of magnitude of landscaping costs as compared to other barrier costs, some typical values were developed as presented in Table 2. Experience in Wisconsin thus far has indicated that metal barriers have been the least expensive. Wood barrier use has had a unit cost about 50 percent higher than metal, while concrete has had a cost 65 percent higher than metal. Landscaping will add cost to a barrier system, but is not as critical as the selection of a material. A limited landscaping may add 5 percent to the cost, an average landscaping 8 to 10 percent, while an elaborate system could add 15 to 20 percent. Actual costs depend on the extent of work, site conditions, and maintenance considerations.

Design trade-offs are inevitable. Decisions must be made as to how much should be spent on the barrier versus the landscape treatment. An elaborate landscaping scheme coupled with a lower-cost barrier material may be cheaper overall than using simple landscaping with more expensive material. Such trade-offs should be made clear to local citizens and elected officials so they can select the option that best fits with their community. In some contexts, dense landscaping may be required to blend the barrier with its environment. In such a situation, a simple barrier could be used, retaining more of the funding allotment for landscaping. In other, especially more urban, locations, the form of the wall may be the most important design feature, with a simpler landscape scheme used to accentuate the wall form.

DESIGN PROTOTYPES

Prototypical designs were generated to illustrate the application of the design guidelines and ideas presented in previous sections. Five sites along the stretch of I-94 were selected to represent prototypical situations. In addition, a short segment of I-43 was selected to represent an example of a "historical ghost." A map of the location of these sites is shown in Figure 6.

Before selecting the specific sites, a windshield analysis was completed to determine unique characteristics of this corridor section. The southern end of the corridor appears to be rural in character when viewed from the freeway. It gives way to newer suburban development as movement proceeds north toward the airport interchange. North of the interchange, houses are close to the freeway and development becomes urban in character. North of these areas, the freeway cuts through older, established urban neighborhoods. Local landmarks, particularly church spires, are prominent on the skyline. At the intersection of I-94 and Mitchell Street, a gateway is formed by the twin spires of St. Stanislaus Church and the Allen-Bradley clock tower. The visual image of a gateway to downtown should be reinforced because just beyond this location, the freeway crosses the Menomonee Valley, which opens an expansive view to downtown Milwaukee.

A series of before and after sketches were made of several sites along this corridor to demonstrate how noise barriers and landscaping can be used to accentuate the characteristics of specific areas to create a more aesthetically pleasing and coherent freeway corridor environment. The intent of each design solution is briefly described in the following paragraphs. They are presented in a sequence from the south to represent views a motorist would have while traveling north to the city center.

Site 1—Rural Character

This site, located just south of the Rawson Avenue interchange, is currently relatively undeveloped. The area is characterized by patches of development interspersed with farm fields, meadows, and woodlands. Historically, this area sup-

TABLE 1 RELATIONSHIP OF LAND USE TYPE TO DESIGN CONSIDERATIONS

Design Component	Urban Land Use	Suburban Land	Rural Land Use
Orientation	vertical	horizontal	horizontal
Lines	angular	mixed	curvilinear
Balance	symmetry	mixed	asymmetrical
Rhythm	regular	mixed	irregular

TABLE 2 RELATIVE COST COMPARISON FOR LANDSCAPE TREATMENT COMBINED WITH BARRIER MATERIAL

Landscape Treatment**	Barrier Material*					
	Metal		Wood		Concrete	
Limited (1)	Barrier	1.00	Barrier	1.50	Barrier	1.66
	Landscape	0.05	Landscape	0.05	Landscape	0.05
	Total	1.05	Total	1.55	Total	1.71
Average (2)	Barrier	1.00	Barrier	1.50	Barrier	1.66
	Landscape	0.08	Landscape	0.08	Landscape	0.08
	Total	1.08	Total	1.58	Total	1.74
Extensive (3)	Barrier	1.00	Barrier	1.50	Barrier	1.66
	Landscape	0.22	Landscape	0.22	Landscape	0.22
	Total	1.22	Total	1.72	Total	1.88

Notes:

Multipliers are given using least expensive noise abatement, metal wall alone, as base. Base = 1.0. These estimates are approximate. Actual costs will depend upon the exact nature of the design, site conditions and maintenance concerns.

* Costs for barrier materials based on average unit cost ($/linear foot).

** Landscape treatment cost based on landscaping along 100 linear feet of barrier and extending 15' into R.O.W. from barrier.

(1) Limited landscaping includes a continuous, single row hedge planted along barrier wall with remainder of landscape area (15' x 100') seeded with wild flower/grass mixture.

(2) Average landscaping includes a multi-row/mass planting of shrubs of various heights and occasional overstory trees (i.e. 3 per 100 linear feet). Remainder of landscape area (10' x 100') seeded with wild flower/grass mixture.

(3) Extensive landscaping includes 3' retaining wall extending partial length of barrier segment (i.e. 40'-50'); multi-row/mass planting of shrubs of various heights and occasional ornamental trees (i.e. 3 per 100 linear feet). Remainder of landscape area (10' x 100') seeded with wild flower/grass mixture.

ported several truck farming and greenhouse operations. Much of its original rural farm character remains intact.

Noise barriers in this area should be designed to reflect its rural character (Figures 7 and 8). Design characteristics of rustic, residential, and utilitarian fences could be adapted to the barrier wall design. The barrier wall could be constructed of rough sawn wood, placed in a somewhat loose pattern to resemble grape stakes. Landscaping should consist of naturalistic drifts of native grasses and wild flowers. Piles of field stones placed at irregular intervals at the base of the wall would add contrast and visual interest.

Site 2—Suburban

This site is located near the airport. The area is characterized by low- to medium-density residential development. This stretch of the freeway is in a slight cut, with gently undulating side slopes along its edge. A significant amount of vegetation has been retained along the top of the side slopes, possibly in the backyards of abutting residences. In this situation, a primary objective would be to design the barrier to resemble a residential fence (Figures 9 and 10). A strong horizontal texture or pattern on the surface of the panels would help to visually reduce the height of the wall by reinforcing the horizontal line of the landform. A wood barrier would blend nicely into the landscape; however, concrete, brick, and even metal could also be used to create the same basic result. Landscaping should consist of informal mass plantings of trees and shrubs on both sides of the barrier. The barrier should appear to meander in and out of the woods. Consideration should also be given to adding berms so that the height of the wall could be reduced.

Site 3—Urban Focal Point

This site is located along the east side of the freeway at a curve just south of Howard Avenue. The sharp curve presents a situation where motorists' views are naturally focussed straight ahead to a potential barrier location. The opportunity exists to capitalize on this situation and create focal points (Figures 11 and 12). This site is located in the Bayview neighborhood, which began as an iron and steel company town. Barriers in this area could be constructed of metal to acknowledge the local historical roots. The design solution proposed suggests attaching decorative metal tracery screens to smooth metal wall panels. Some space should be left between the tracery screen and the wall panel to create depth and shadow patterns. In addition, the wall panel should be light in color to ensure that shadow patterns are visible. To retain their focal impact and contextual uniqueness, decorative panels of this type should be used sparingly. Adjacent wall sections should be rather simple to avoid visual distraction and confusion. No landscaping is required along the decorative panels, because the barrier itself is visually attractive. Landscaping along the ad-

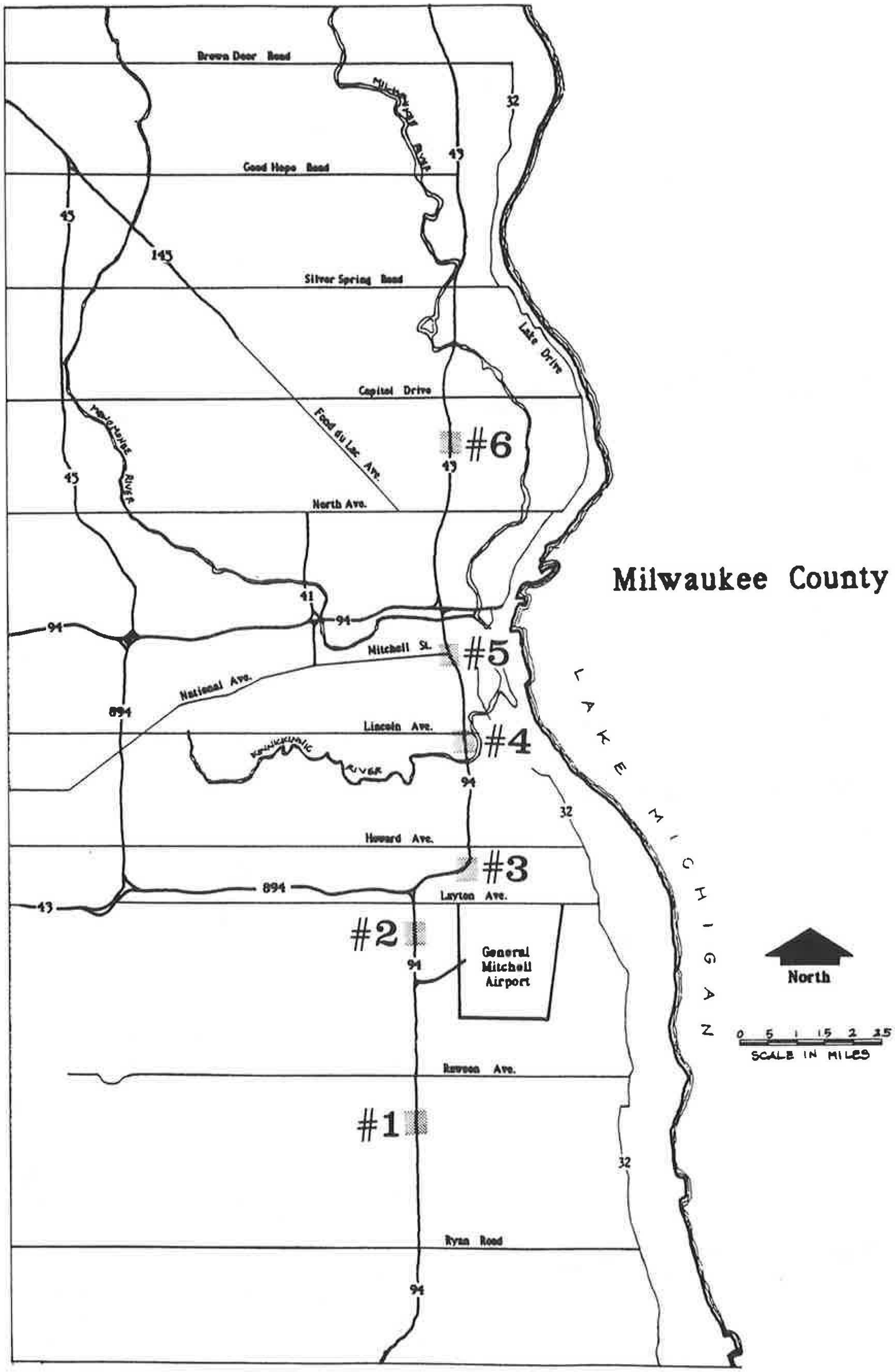


FIGURE 6 Location of sites for prototype designs.

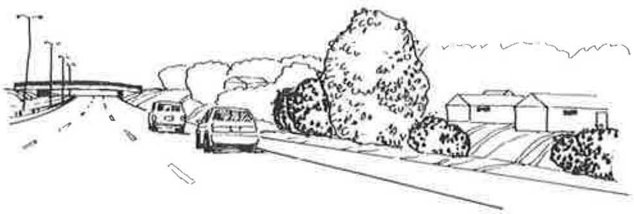


FIGURE 7 Site 1, rural character, before.

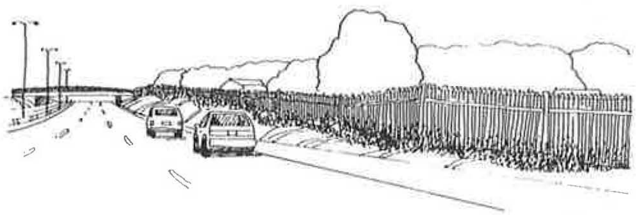


FIGURE 8 Site 1, rural character, after.

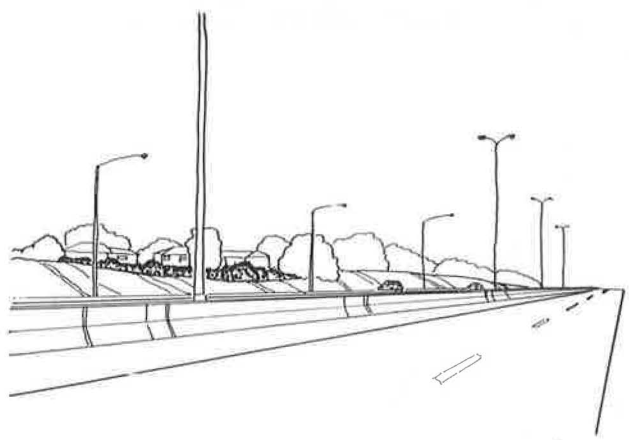


FIGURE 9 Site 2, suburban character, before.

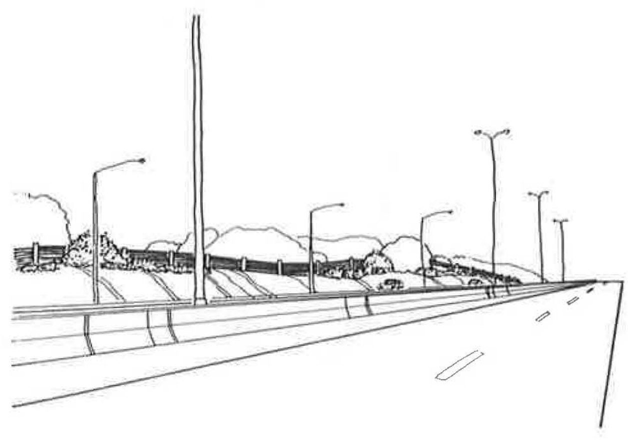


FIGURE 10 Site 2, suburban character, after.

FIGURE 11 Site 3, urban focal point, before.

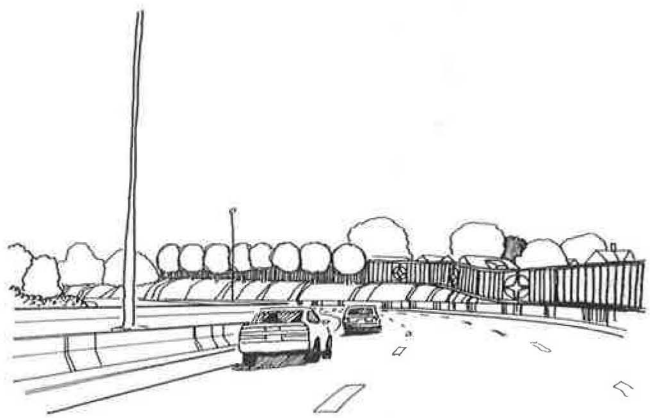


FIGURE 12 Site 3, urban focal point, after.

jacent wall sections should also be simple. In this example, an existing row of trees was retained on the freeway side of the barrier to resemble a residential streetscape.

Site 4—Urban Neighborhood

This site is located in a densely developed, old urban neighborhood. The site contains two significant attributes: a dense row of healthy, mature vegetation, and a prominent local landmark, the dome of St. Josaphat Basilica, a historic church, as a backdrop. The barrier should be designed to accommodate both of these assets. The solution proposed is relatively simple in design (Figures 13 and 14). The wall should be designed to have a finished, more refined, urban character. Pillars could be placed at regular intervals between spans of wall panels. The wall panels should be relatively dark in color, so they visually recede into the landscape plantings. A contrasting (light-colored) cap provides a finishing touch to the top of the wall. Special care should be used when siting and constructing the barrier to save and use as much of the existing vegetation as possible. Some additional landscaping might be necessary to blend the overall design into the existing landscape.

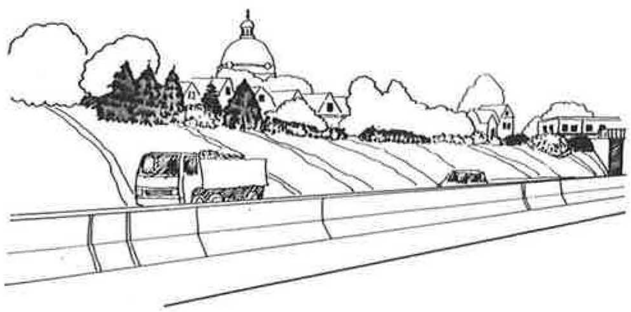


FIGURE 13 Site 4, urban neighborhood, before.

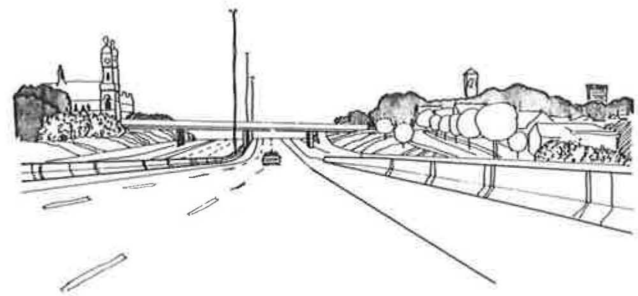


FIGURE 15 Site 5, urban gateway, before.

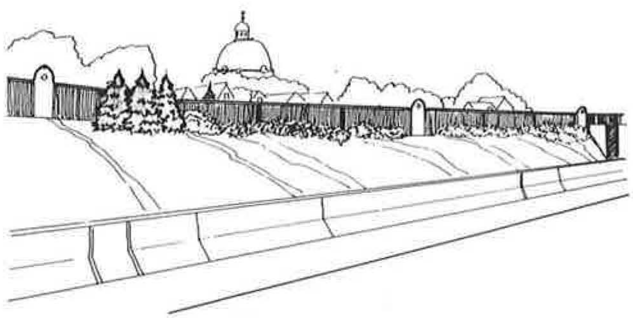


FIGURE 14 Site 4, urban neighborhood, after.

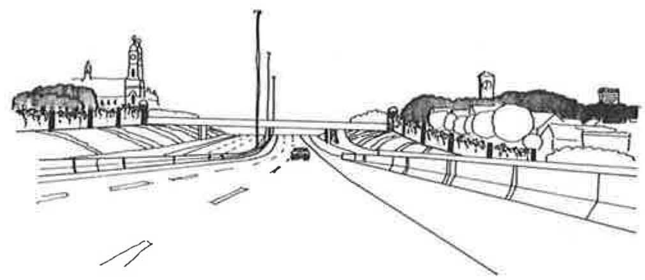


FIGURE 16 Site 5, urban gateway, after.

Site 5—Urban Gateway

This site is located near the downtown along the Mitchell Street interchange. A sense of gateway is created by two visually prominent vertical structures located on either side of the roadway. The twin spires of St. Stanislaus Church and the Allen-Bradley clock tower are prominent Milwaukee landmarks located along the seam where the industrial heart of the city meets its oldest residential neighborhoods. Because of their visual prominence and juxtaposition to each other and to downtown, they create a gateway signaling either the entrance to or exit from the heart of the city (downtown).

A noise barrier on this site could be designed to reinforce the gateway image (Figures 15 and 16). As proposed, the top profile of the walls are stepped to emphasize an upward movement. The walls are tied to the existing bridge structure by vertical pillars. These pillars create a more literal gateway, making reference to the implied gateway created by the two prominent buildings in the background. Vertical posts between wall panels accentuate the upward movement toward the pillars and the surrounding architecture. The vertical elements, end pillars and panel posts, should be darker in color than the wall panels to make them visually dominant.

Site 6—Historical Ghost

This site is located where I-94 crosses Chambers Street. It is the former site of Borchert Field, which served as the major athletic field for Milwaukee baseball teams during the first half of the century until the opening of Milwaukee County Stadium in 1953. In the same year, Borchert Field was demolished to make way for construction of I-43. To reflect this cultural relic, the barrier walls could be designed using a baseball theme.

To avoid being distractive, the design should be simple and uncluttered (Figures 17 and 18). The solution proposed involves the integration of a backstop into the barrier on the west side of the freeway. Home base would be constructed of an appropriately shaped slab of white concrete or marble. The dimensions of the base would have to be exaggerated somewhat to increase its visibility. The base would be placed within a bed of reddish colored gravel to provide heightened contrast. The existing slope of the cut bank could be used to make the base more visible by tilting it toward the roadway.

Borchert Field stadium was a wooden structure. Similarly, the barrier walls should be constructed of wood boards, preferably with a weathered finish, placed vertically, side by side, alluding to the baseball stadium. Numbers stenciled on the east side barrier signify field length. Vertical poles symbolize foul line markers. A change of materials would clearly signal that this area is distinct. Concrete would be an appropriate choice of material for the adjacent barrier because it would provide color contrast and could be given a vertical surface texture to complement the wood barrier design.

Landscaping on the freeway side would be minimal. Vines could be planted to grow on or over the wall. Conceptually, the cut slope should be planted with turf grass. At the point where the barrier material changes, the landscape treatment should also change. The area immediately adjacent to the turf should be planted with a taller ground cover to clearly delineate the edge of the historic ball field.

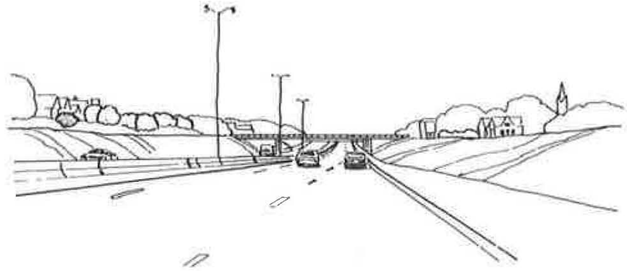


FIGURE 17 Site 6, historical ghost, before.

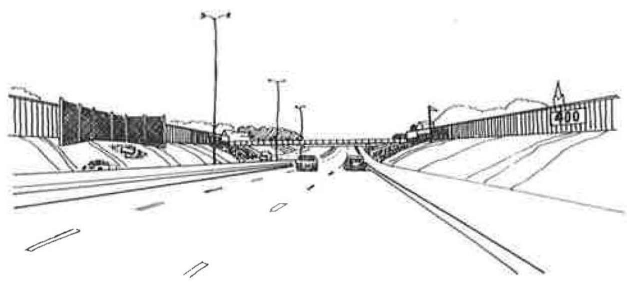


FIGURE 18 Site 6, historical ghost, after.

CONCLUSIONS

Information and tools that can be used to enhance the visual and aesthetic quality of freeway noise barriers have been provided. Noise barriers can be effective in reducing unwanted sound from highway sources in residential neighborhoods. However, they also can have a substantial effect on the visual environment of a highway and surrounding neighborhoods. They can significantly change the view from the road creating monolithic tunnels of walls and by blocking views of changing urban scenery. They can also change the view towards the road for the surrounding community by creating barriers to other areas and developing a sense of isolation. What is needed is a way to provide the benefits of sound reduction from noise barriers while at the same time creating a positive visual image for road users and the surrounding communities.

An attractive and efficient system for freeway noise control can be developed if the following principles are followed.

Diversity

There should be variety in the placement, materials, texture, and landscaping to enhance the aesthetic characteristics of noise barriers. Urban landscapes are diverse mixtures of buildings, plants, streets, and sidewalks; there should be no hesitancy to incorporate this diversity into noise barriers.

Integrate Barriers and Landscaping

Noise barriers and landscaping should be viewed as an integrated, complementary system. Choices of materials, textures, profile, and location should be done in such a way that the various elements fit together into an integrated whole. A balance should be struck between wall decoration and landscaping so that they do not conflict or compete with each other.

Reflect Neighborhood Characteristics

The physical, cultural, and historical characteristics of urban neighborhoods should be incorporated into the design of noise barrier and landscaping systems. Noise barriers should reflect the basic land uses and the heritage of the areas through which the highways pass. Noise barrier design should include efforts to understand urban areas through careful inventories of physical, visual, cultural, ethnic, historical, and land use characteristics of different neighborhoods. These factors should be used to shape design themes as well as details. Specific views should be reinforced and enhanced. Gateways should be identified and accentuated.

Options

Decisions about barriers and landscaping should be made in consultation with neighborhood groups, elected officials, and others. A broad range of options should be provided including materials, profile, and configuration of barriers, and type, location, and configuration of landscaping. Trade-offs should be provided between wall materials and landscaping so that an acceptable balance can be reached.

Creative Process

Finally, a process should be used that maximizes the opportunity for creative design of noise barrier and landscaping systems. This process would include careful inventories of neighborhood characteristics, selection of design themes, preparation of alternative designs, community involvement, and selection of designs that balance the various trade-offs of cost, aesthetics, and noise reduction. Through such a process, systems can be developed that enhance the quality of the environment that is seen as well as heard along freeways.

ACKNOWLEDGMENTS

This report was prepared in response to a legislative request included in the 1989–1990 state budget to look at ways to improve aesthetic characteristics of freeway noise barriers.

The authors would like to thank Rep. Timothy Carpenter for his support in this project and for the help provided through his office by Paula Doyle and Curt Finkelmeyer in arranging meetings. A task force established to guide this project included representatives of the Wisconsin Department of Transportation, the Southeastern Wisconsin Regional Planning Commission, and the City of Milwaukee. This task force provided useful insight and input on the project. In addition, we would like to thank members of the TRB Committee on Transportation Noise and Vibration who supplied information for this study through correspondence about noise barrier

design in Michigan, New York, New Jersey, Minnesota, Maryland, California, and other states.

REFERENCES

1. J. Farnham and E. Beimborn. *Noise Barrier Design Guidelines.* Center for Urban Transportation Studies, University of Wisconson, Milwaukee, July 1990.
2. R. F. Blum. *A Guide to Visual Quality in Noise Barrier Design.* Implementation Package 77-12, FHWA, U.S. Department of Transportation, Dec. 1976.
3. *Metro Noise Abatement Program.* Minnesota Department of Highways, Office of Environmental Services, Feb. 1976.
4. *Roadside Design Guide, Chapter 3: Roadside Topography and Drainage Features.* AASHTO, Washington, D.C., 1989.
5. K. M. Keegan. *The Integration of the Highway and Landform.* Center for Architectural and Urban Planning Research, School of Architecture and Urban Planning, University of Wisconsin, Milwaukee, 1989.
6. *Fichier d'Écrams Acoustiques.* Center d'études des Transports Urbains, March 1987.
7. M. Springer. Vegetative Noise Barriers from Willowtwigs (in German). *Neue Landschaft, Fachzeitschrift für Garten-, Landschafts-, Spiel- und Sportplatzbau,* Patzer Verlag GmbH u. Co. KG, Hannover, Berlin, West Germany, June 1987, pp. 287–389.
8. J. Ritzer. The What and Why of Natural Roadsides. *Public Works,* March 1990, pp. 49–53.
9. J. Gurda. *Discover Milwaukee Catalog.* Department of City Development, City of Milwaukee, Wis., 1986, reprinted 1988.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Fort Hood Noise Study

RICHARD M. LETTY

At the request of the U.S. Army Corps of Engineers, Fort Worth District Office, an installation compatible-use zone (ICUZ) noise study was prepared for Fort Hood, Texas. The purpose of this study was to address the noise impact from military training activity conducted at Fort Hood. The major component of this Fort Hood ICUZ noise study was a comprehensive long-term noise monitoring program and the use of computer modeling to develop noise contours to identify noise-impacted areas. Noise measurements were obtained at a total of 17 noise measurement locations: 9 airfield noise monitoring sites, and 8 weapon-firing blast noise monitoring sites. Because of the day-to-day variations in military training activity, it was determined that 60 days of noise data at each of the 17 noise-monitoring sites would be useful in understanding long-term airfield and weapon-firing blast noise levels. Airfield noise contours were developed using the NOISEMAP computer model, whereas the weapon-firing blast noise contours were developed using the BNOISE computer model. An overview of the Fort Hood ICUZ noise study is presented, and the measured and predicted airfield and weapon-firing blast noise levels are compared.

At the request of the U.S. Army Corps of Engineers, Fort Worth District Office, an installation compatible-use zone (ICUZ) noise study was prepared for Fort Hood. This ICUZ study is required by Army Regulation (AR) 210–20 (*1*) as part of the Department of Army Installation Master Planning Program. The purpose of this study is to address the noise impact issue from military training exercises conducted at military facilities at Fort Hood. The ICUZ program offers a method for analyzing exposure to noise and provides land use guidelines for achieving compatibility between the needs of the army and civilian communities.

One of the major components of this ICUZ study was a comprehensive long-term noise monitoring program and the use of computer modeling to develop noise contours to identify Zone I, Zone II, and Zone III noise-impacted areas. This study deals almost exclusively with the issue of noise at Fort Hood. Noise measurements were obtained at a total of 17 noise measurement locations: 9 airfield noise monitoring sites around Hood Army Airfield (HAAF) and Robert Gray Army Airfields (RGAAF), and 8 blast noise monitoring sites at various Fort Hood weapon-firing range locations. Because of the day-to-day variations in military training activity, it was determined that 60 days of noise data at each of the 17 noise monitoring sites would be useful in understanding long-term airfield and weapon-firing blast noise levels.

PROGRAM DESIGN

Because of the limitations in the available number of noise monitoring units, and the different noise descriptors used in measuring airfield noise and impulsive weapon-firing blast noise, it was decided that the noise monitoring would be performed in two phases. The Phase I noise measurements focused on airfield noise. A total of nine noise measurement locations were selected: two in the vicinity of RGAAF, four around HAAF, and three along the various flight corridors leading to and from the Fort Hood Army Installation. The Phase II noise measurements focused on the blast noise from artillery and weapon-firing activity on the various ranges at Fort Hood. The purpose of the noise monitoring program was not only to define the actual noise levels from airfield and weapon-firing activity at Fort Hood, but also for comparison with the computer modeling results.

The primary noise descriptor for the measurement and evaluation of airfield noise is an A-weighted day-night noise level (ALDN) measured over a 24-hr period with a 10-dB penalty assessed to nighttime (10:00 p.m. to 7:00 a.m.) hourly Leq noise levels. The hourly Leq level is the steady A-weighted sound level over a 1-hr period that contains the same acoustic energy as the fluctuating noise measured during that same 1-hr period of time. Because weapon-firing blast noise is an impulsive noise source, the primary noise descriptor is the C-weighted day-night noise level (CLDN) measured over a 24-hr period with a 10-dB penalty assessed to the nighttime hourly Leq noise levels.

The ALDN and CLDN noise descriptors are consistent with the annual average day noise contours generated by the NOISEMAP (airfield noise contours) and BNOISE (blast noise contours) computer models used in this analysis. According to Army Regulation AR200–1 (*2*), noise impact areas should be divided into three zones where residential housing, schools, and other noise sensitive land uses will be considered as follows:

Zone	*ALDN*	*CLDN*	*Quality*
I	<65	<62	Acceptable
II	65–75	62–70	Normally Unacceptable
III	>75	>70	Unacceptable

INSTRUMENTATION AND COMMUNICATION NETWORK SYSTEM

To perform the noise monitoring portion of the Fort Hood ICUZ noise study, instrumentation was required that could measure both A-weighted (for airfield noise) and C-weighted (for weapon-firing blast noise) noise levels. In addition, the instrumentation also had to have a fast rise time and sampling rate to measure the full amplitude and short duration time of the impulsive weapon-firing blast noise. In order to satisfy these constraints, the U.S. Army Environmental Hygiene Agency (AEHA) at Aberdeen Proving Grounds in Maryland provided their Metrosonics dB–604 noise monitoring units

Acentech Inc., 125 Cambridge Park Drive, Cambridge, Mass. 02140.

with real time detectors (RTDs) for use during the noise monitoring portion of this study.

The Metrosonics dB–604 units have the ability to measure A-weighted or C-weighted noise levels and store the data internally for downloading at a later time. The primary format selected for data storage on these Metrosonics units were hourly multiple intervals that included hourly Leq levels, five hourly statistical noise descriptors (L_{99}, L_{90}, L_{50}, L_{10}, and L_1), L_{max}, single-event noise levels (for analyzing aircraft and helicopter operations and weapon-firing blast noise), and daily LDN levels.

Although these various noise descriptors were useful in evaluating the data on a day-to-day basis, the primary noise descriptor used in evaluating the noise impact from activity at Fort Hood was the LDN day-night sound level. LDN is the specified noise descriptor recommended by the Department of Defense and other federal agencies to be used in assessing community noise impacts in ICUZ noise and land use compatibility studies.

Because of the vast amount of noise data obtained during the noise monitoring portion of the Fort Hood ICUZ noise study, it was necessary to devise a process where the data could be checked and evaluated on a daily basis to minimize instrumentation downtime or questionable noise measurements that could result in loss of data. As a result, each of the Metrosonics dB–604 units were equipped with a telephone modem and a dedicated telephone line. In addition, the units were also connected to AC power to keep the internal gel-cell batteries fully charged. Without AC power, the Metrosonics dB–604 RTD units would require frequent recharging. Because this would result in unacceptable downtime, it was necessary to select noise measurement locations with access to AC power and telephone service. All Metrosonics dB–604 units were also equipped with wind anemometers to inhibit data logging during high wind (>15 mph) conditions.

With the Metrosonics dB–604 units connected to a modem and telephone line, it was possible to communicate directly with each of the noise monitoring sites from the Acentech office in Cambridge. This set-up allowed downloading of the data from the Metrosonics units in Texas directly onto a personal computer in Cambridge for final storage and further analysis and evaluation. Figure 1 shows a schematic of the instrumentation and communication network system. This process allowed verifying that the units were operating properly, downloading and storing the data onto a personal computer, and checking the data on a daily basis. A local technician was also hired to check and calibrate the units on a regular basis, and to provide field maintenance when necessary. Each noise monitoring site was contacted, checked, and downloaded each morning by telephone, and if a problem was detected, the site technician was called and instructed as to which sites required immediate attention. This process allowed minimizing system downtime, as well as reducing the amount of lost or questionable data.

NOISE MONITORING SITES

A total of 17 noise measurement locations were selected to monitor both airfield noise (9 sites) and weapon-firing blast noise (8 sites).

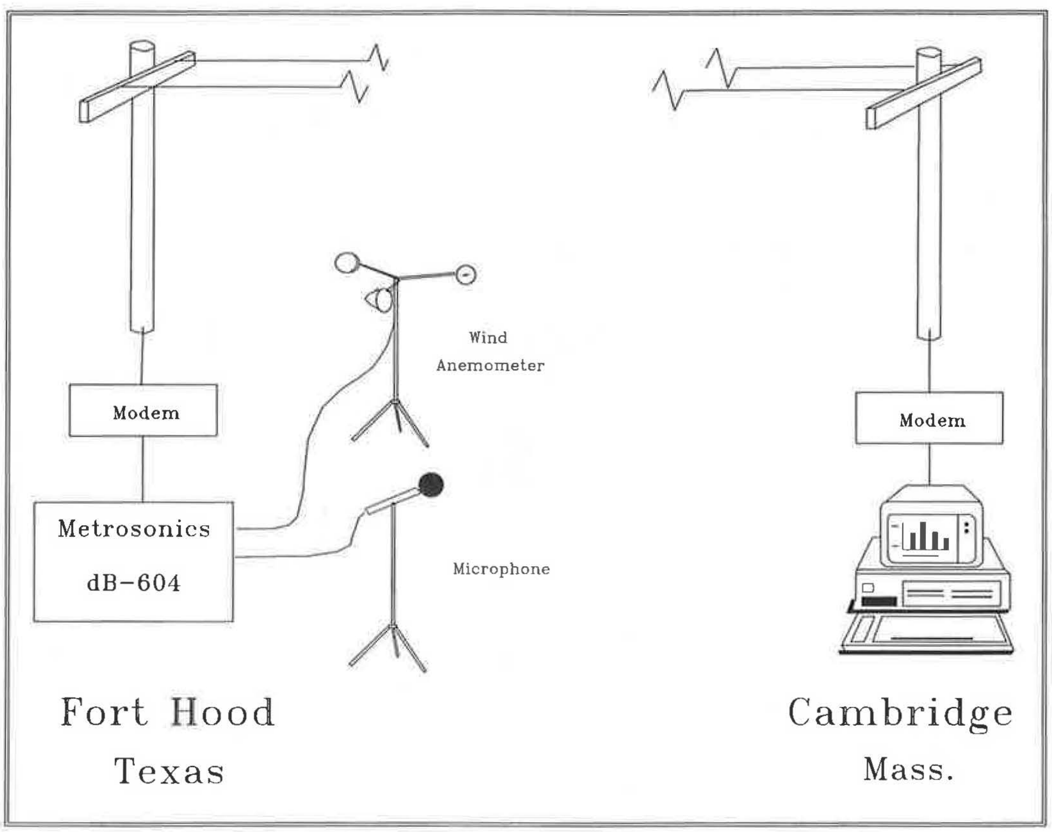


FIGURE 1 Instrumentation-communication network.

Airfield noise measurements were obtained at both HAAF and RGAAF. The location of the airfield noise monitoring sites were as follows: RGAAF (two sites), HAAF (four sites), and Fort Hood flight corridors (three sites). Figures 2–4 show the various airfield measurement locations in the vicinity of HAAF and RGAAF. All nine airfield noise monitoring sites were located at residential homes where access to ac power and telephone service was available.

RGAFF is located at the extreme southwest corner of the Fort Hood Military Reservation. The primary noise sources at RGAAF are helicopter operations and C-5 touch-and-go training exercises conducted by aircraft flying in from other airbases. Because of the reduced level of aircraft and helicopter activity at RGAAF in comparison with HAAF, only two noise monitoring sites were selected in this area.

HAAF is located along the south edge of the military reservation property line adjacent to the Fort Hood cantonment area. Helicopter operations are the primary noise source at HAAF. Although most of the helicopter flight operations are required to remain over Fort Hood property, the proximity of residential housing to the airfield makes this a noise sensitive area.

For noise monitoring along the flight corridors leading to and from the Fort Hood Military Reservation, three sites were selected as representative of this type of activity.

A total of eight noise monitoring sites were selected to measure the weapon-firing blast noise levels at Fort Hood. These measurement locations are shown in Figure 5. Because of the requirements regarding access to AC power and telephone communication with each of the noise monitoring units, and security, all eight of the noise monitoring sites were located at the observation towers at the various weapon-firing ranges. These eight locations were selected because they essentially encircle the entire Fort Hood weapon-firing area and were easily accessible.

AIRFIELD NOISE MEASUREMENTS

The airfield noise measurements were obtained at the nine airfield noise monitoring sites over a 3-month period from January through the end of March. Because of the large amount of noise data obtained during the noise monitoring portion of the Fort Hood ICUZ noise study, only the daily LDN noise levels are summarized and presented in this report.

Figures 6 and 7 show typical daily LDN noise levels measured in the vicinity of RGAAF and HAAF. These figures also show the 65 LDN level that defines Zone I (<65 LDN) where residential housing, schools, and other noise sensitive land uses are considered acceptable.

Although the data in these figures indicate that there are some days when the daily LDN level is above 65 dBA, these days generally correlate with high levels of activity at the airfields. For example, Site 5, which is located close to HAAF, clearly indicates this trend. During the week, when activity at the airfield is generally high, the daily LDN levels tend to be at or above 65 LDN. However, on weekends, when there is substantially reduced levels of activity at the airfield, the daily LDN levels tend to be much lower. This same trend, although to a lesser degree, is also consistent with the daily LDN levels measured at Sites 4 and 6, which are also close to HAAF.

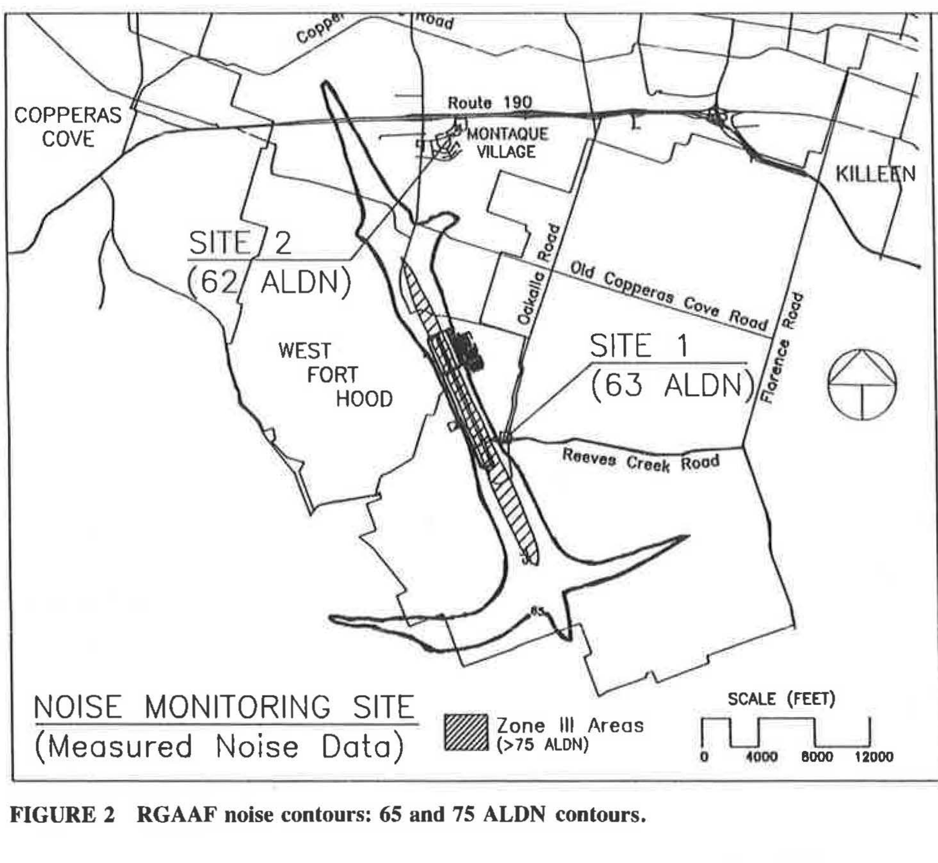


FIGURE 2 RGAAF noise contours: 65 and 75 ALDN contours.

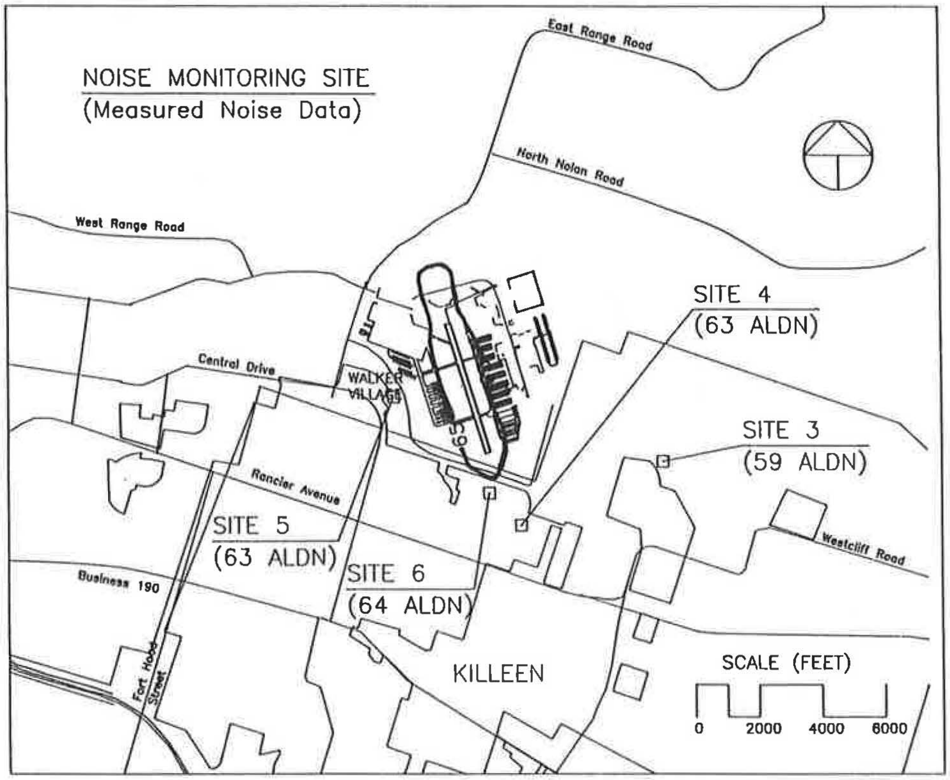


FIGURE 3 HAAF noise contours: 65 ALDN contour.

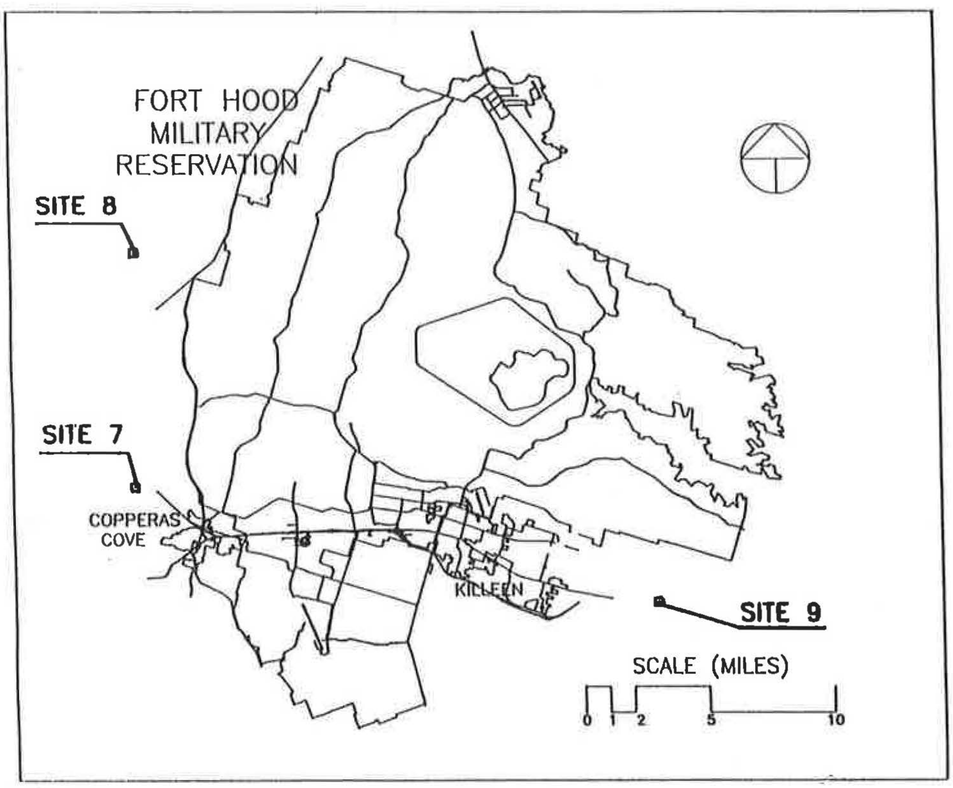


FIGURE 4 Airfield noise monitoring sites: flight corridors.

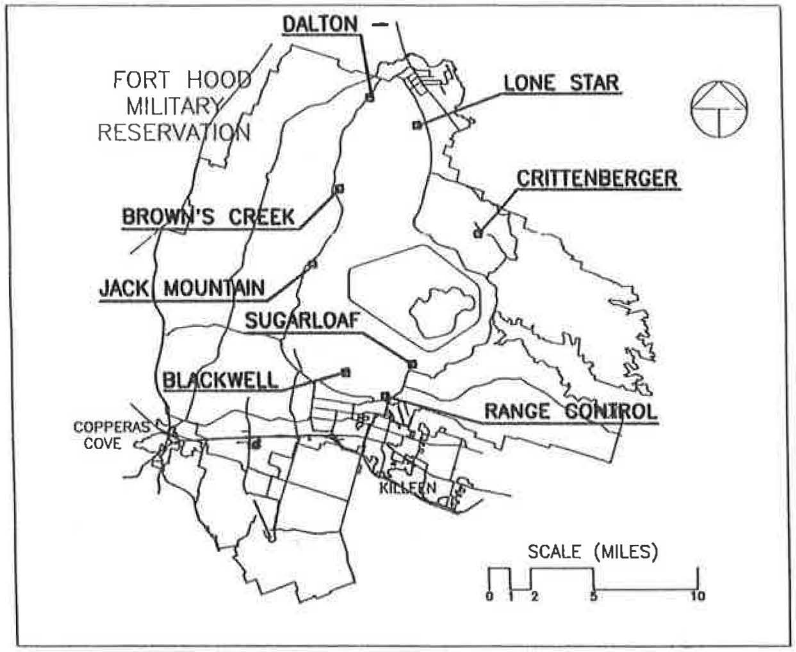


FIGURE 5 Blast noise monitoring sites.

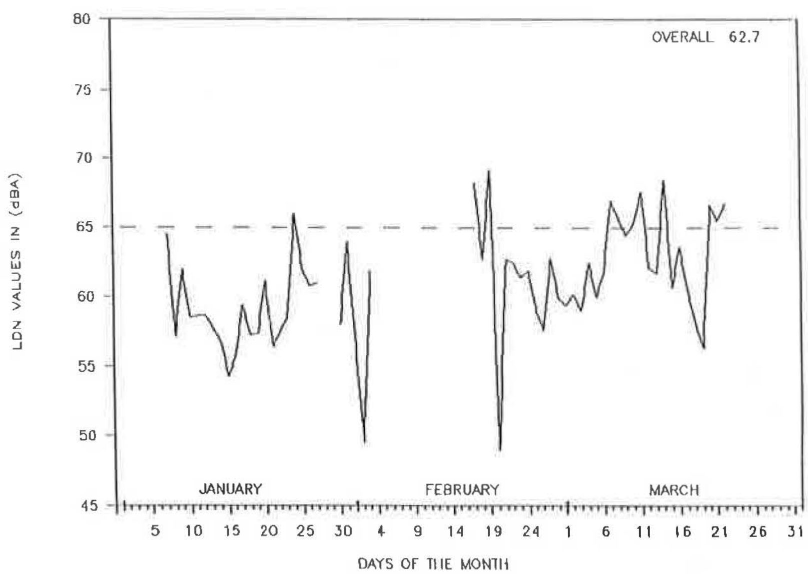


FIGURE 6 Site 1 RGAAF daily LDN noise levels.

Table 1 presents the LDN levels on a monthly basis and also presents the overall LDN level over the entire noise monitoring period. Although there are days when the daily LDN level at some of the sites exceeds 65 dBA, this table clearly indicates that on a monthly and overall basis, the levels at all nine airfield noise monitoring sites are below 65 LDN.

The overall LDN noise levels at Sites 1 and 2 that are near RGAAF are 63 and 62 dBA, respectively. Data obtained at Site 1, which is adjacent to the south end of the runway at RGAAF, indicate that the daily LDN levels usually range between 56 and 62 dBA. However, on days when C-5 aircraft are practicing touch-and-go operations, the daily LDN levels increase to 65 to 68 dBA. Site 2, which is located approximately 3 mi north of RGAAF, is along one of the helicopter flight tracks between Fort Hood and the airfield. In addition, Site 2 is also near Route 190 where traffic noise also contributes to the overall noise levels measured at this location. In any event, the overall levels at Site 2 are still well below 65 LDN.

Sites 4, 5, and 6, which are adjacent to HAAF, all have roughly the same overall LDN level, 63, 63, and 64 dBA, respectively. Because of the flight tracks flown by the helicopters approaching or departing HAAF from the south end of the runway, it could be expected that the daily and overall LDN levels measured at these three noise monitoring sites would be almost identical. Figure 8 shows the daily LDN

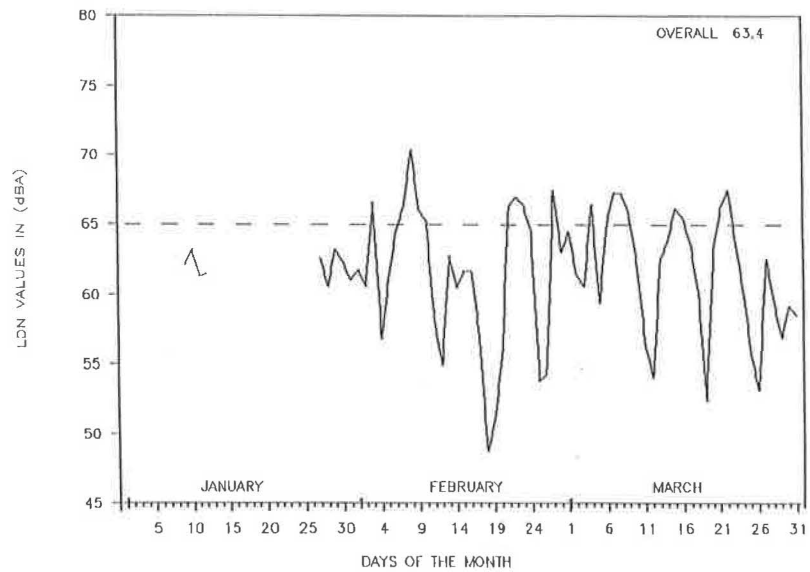


FIGURE 7 Site 5 HAAF daily LDN noise levels.

TABLE 1 HAAF NOISE MEASUREMENTS: SUMMARY A-WEIGHTED LDN LEVELS

	January	February	March	Overall
Gray Army Airfield				
Site 1	60	63	64	63
Site 2	61	63	62	62
Hood Army Airfield				
Site 3	55	56	62	59
Site 4	61	63	64	63
Site 5	62	64	63	63
Site 6	62	64	65	64
Flight Corridors				
Site 7	54	52	56	54
Site 8	-	58	58	59*
Site 9	56	59	58	58

*Overall LDN Level also includes April -61 LDN and May-50 LDN

levels at these three sites for the month of February 1989. Site 3, which is located somewhat farther from HAAF, has an overall LDN level of 59 dBA, which is 4 dBA lower than the overall levels measured adjacent to the airfield.

Noise level measurements at Sites 7, 8, and 9, which were located along flight corridors used by helicopters approaching or departing the Fort Hood Military Reservation, resulted in overall LDN noise levels of 54, 59, and 58 dBA, respectively.

BLAST NOISE MEASUREMENTS

Noise monitoring began at the blast noise monitoring sites on April 1 and continued through the middle of July. Because of the impulsive nature and the relative low frequency of weapon-firing blast noise, noise measurements were made on a C-weighted scale. Figures 9, 10, and 11 show typical daily C-weighted LDN noise levels measured at three of the weapon-firing blast monitoring sites. Sections of missing data in these figures are caused by instrumentation problems and loss of electrical power at several of the range location observation towers. Because there were only eight Metrosonics dB–604 units with real time detectors (RTDs) capable of accurately measuring the peak noise level associated with the short-duration impulsive nature of the weapon-firing blast noise, significant monitoring time was lost when units had to be returned for repairs.

Because of the wide range of training activity that occurs at Fort Hood, evaluation of the measured noise data was difficult. In addition to the weapon-firing activity at the ranges, there were also helicopter and aircraft operations that affected the noise measurements. All of the weapon-firing ranges are located along East, West, and South Range Roads. However, these roads also serve as visual flight track corridors for hel-

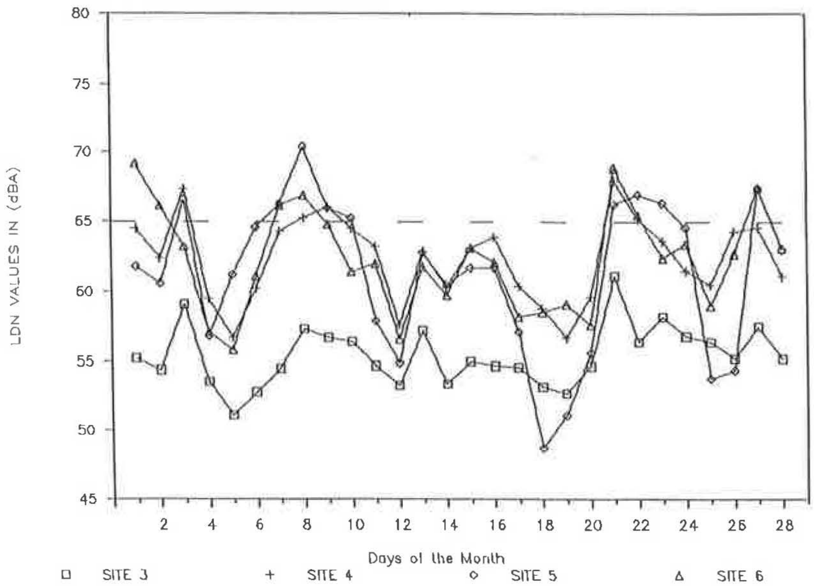


FIGURE 8 Comparison of daily LDN levels for Sites 3, 4, 5, and 6 (Feb. 1989).

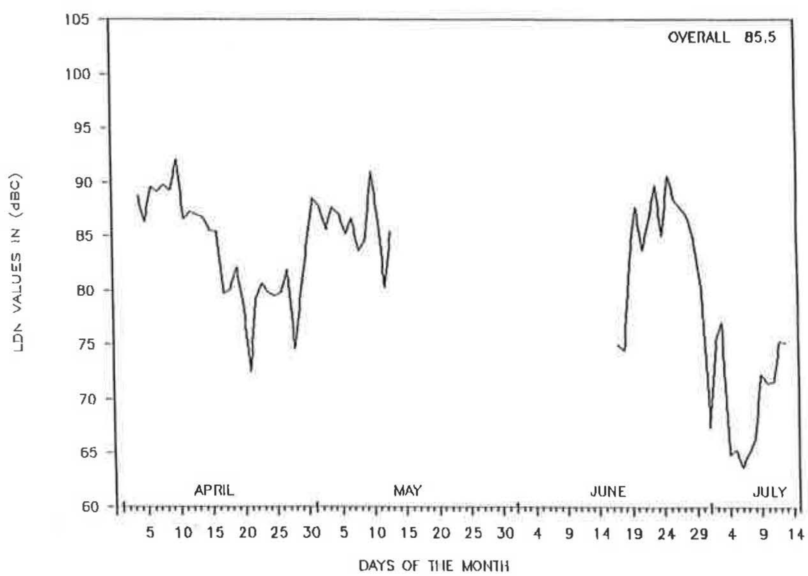


FIGURE 9 Site A: Sugarloaf—daily LDN noise levels.

icopters traveling to and from the various helicopter weapon-firing ranges. Also, aircraft from other military installations use Fort Hood to drop practice bombs at the Shoal Creek bombing range in the north, as well as live ordnances at the Smith Mountain artillery impact area. Therefore, both helicopter and aircraft activity also contributed to the overall noise level measurements obtained at the eight blast noise monitoring sites.

The C-weighted daily LDN noise levels displayed in Figures 9 and 11 clearly indicate the wide range of measured noise levels obtained at the various weapon-firing range locations.

Site A (Sugarloaf) is primarily used by M1 and M60 tanks firing 105-mm cannon rounds. With moderate levels of tank firing activity occurring at this range, daily LDN levels ranged between 85 and 90 dBC. Site C (Crittenberger) is used for multiple tank training exercises. This range is used by tank squads of up to four M1 or M60 tanks firing 105-mm cannon rounds, as well as M2/3 Bradley armored personnel carriers firing 25-mm chain-guns. With Bradley vehicles using the range, daily LDN levels range between 74 and 77 dBC. When the M1 and M60 tanks are on the range, the daily LDN levels increase to 85 to 90 dBC. During the noise measurement program, there were several days when 8-in. and 155-mm artillery firing from locations close to Site C (within 2 to 3 mi) along with actual range activity generated substantially higher daily LDN noise levels (95 to 100 dBC). However, when artillery firing occurred at more distant firing locations (>5 mi), the additional noise impact was not discernible above the actual range activity.

Site D (Lone Star) and Site E (Dalton) are located at the extreme northern section of the Fort Hood Military Reservation. Although Site D is primarily used by Army reserve

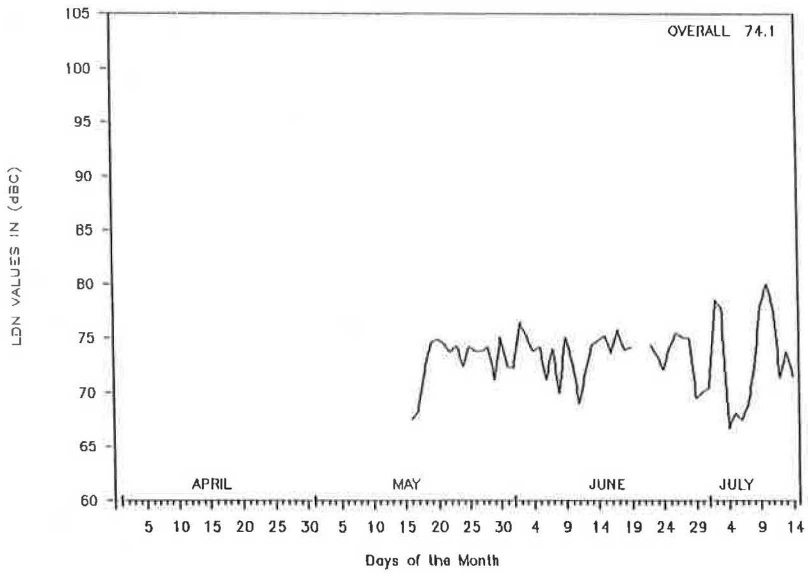


FIGURE 10 Site B: Range Control—daily LDN noise levels.

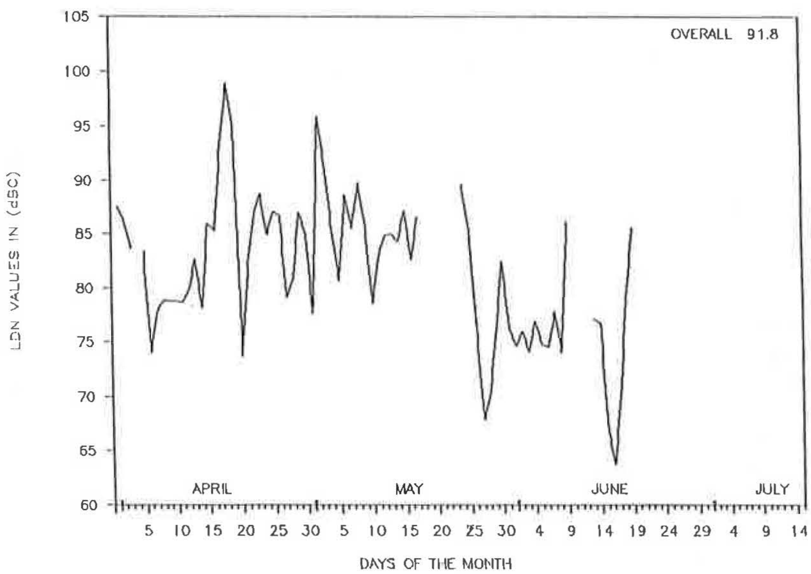


FIGURE 11 Site C: Crittenberger—daily LDN noise levels.

units on weekend training exercises, it is also used during the week as a forward area refuel and rearm point (FAARP) for helicopters awaiting their turn to fire their weapons at the Dalton–Henson Mountain firing ranges. In addition, Sites D and E are impacted by military jet aircraft using the Shoal Creek bombing range to practice simulated bombing runs. These aircraft flyovers contribute to the overall noise levels measured at these two locations. As a result, the overall C-weighted LDN noise levels measured at Sites D and E were higher than might have been expected given the actual activity contained in the range control firing logs.

Site E (Dalton) is primarily a helicopter firing range. The Dalton–Henson Mountain multiuse range complex is the primary firing range for helicopters firing their 2.75-in. rockets. In addition to the rocket-firing noise levels measured at this site, the helicopters themselves also contribute to the overall measured noise levels.

Site F (Brown's Creek) and Site G (Jack Mountain) are primarily firing ranges for the M2/3 Bradley armored personnel carrier firing its 25-mm chain-gun.

Site H (Blackwell) and Site B (Range Control) are not located on active firing ranges. These two sites were selected to define the measured noise levels along the southern boundary of the ranges without being impacted by actual on-site range activity. The higher daily LDN noise levels measured at Site B are caused by the impact of helicopter activity from HAAF. Helicopters departing or approaching HAAF from the west use South Range Road as their visual flight track corridor.

Table 2 presents the data measured at the eight blast noise monitoring sites on both a monthly and overall C-weighted LDN basis. Although these data accurately reflect the noise levels measured on the firing ranges, it was necessary to rely on the noise contours generated by the BNOISE computer

TABLE 2 FORT HOOD WEAPON-FIRING BLAST NOISE MEASUREMENTS: SUMMARY C-WEIGHTED LDN LEVELS

		April	May	June	July	Overall
Site A:	Sugarloaf	86	87	86	72	86
Site B:	Range Control	-	73	74	75	74
Site C:	Crittenberger	88	93	93	-	92
Site D:	Lone Star	82	89	88	-	87
Site E:	Dalton	85	79	80	73	82
Site F:	Brown's Creek	79	81	77	75	78
Site G:	Jack Mountain	81	84	95	78	89
Site H:	Blackwell	-	-	68	67	68*

*Overall LDN Level also includes August - 68 LDN.

model to determine the location of the acceptable Zone I (<62 LDN) noise area.

COMPUTER MODELING—NOISE CONTOURS

The second part of the Fort Hood ICUZ noise study consisted of using computer modeling to generate noise contours for both the airfield and the weapon-firing blast noise impact. In fact, one of the primary reasons for the extensive noise monitoring program at Fort Hood was to compare actual measured noise data with the noise contours generated by computer modeling.

The NOISEMAP computer model was used to generate the noise contours for the two primary Fort Hood Airfields: RGAAF and HAAF. This computer model was developed for the U.S. Air Force and is the model recommended by the Department of Defense for use on all Military AICUZ/ICUZ noise studies with airfield installations. Using airfield data consisting of aircraft or helicopter type, approach and departure flight tracks, runway utilization, and number of daytime and nighttime (10:00 p.m. to 7:00 a.m.) operations, A-weighted LDN noise contours were generated to define the Zone I (<65 LDN), Zone II (65 to 75 LDN), and Zone III (>75 LDN) impact areas. For this study, the most current PC Version 6.0 of the NOISEMAP computer model was used (*3*).

For assessing the noise impact from weapon-firing impulsive blast noise from U.S. military installations, the U.S. Army Construction Engineering Research Laboratory's (CERL) BNOISE computer model was used (*4*). This computer model supports the U.S. Army's ICUZ noise program and is used to generate C-weighted LDN noise contours to evaluate weapon-firing noise impacts at military installations. Using weapon range firing data that consist of weapon type (tank, artillery, mortar, rocket, etc.), size of round fired (105-mm, 155-mm, 8-in., etc.), type of round fired (TPT, HEAT, HE, etc.), firing point location, impact location, number of daytime rounds fired, and number of nighttime (10:00 p.m. to 7:00 a.m.) rounds fired, C-weighted LDN noise contours were generated to define the Zone I (<62 LDN), Zone II (62 to 70), and Zone III (>70 LDN) noise impact areas.

AIRFIELD COMPUTER MODELING—NOISEMAP

HAAF is located on South Fort Hood and borders the Killeen city limits. It is a fully operational airfield designed primarily to handle rotary-wing Army aircraft. The majority of operations at HAAF are conducted by six types of helicopters, the AH-64, AH-1, OH-58, UH-60, UH-1, and CH-47. Annual flight operations were derived from air traffic activity reports and information gathered from air traffic control personnel. The activity modeled was 100 percent rotary wing. The yearly averaged daily operations were 341.6 day flights (between 7:00 a.m. and 10:00 p.m.) and 34.4 night flights (between 10:00 p.m. and 7:00 a.m.).

The LDN noise contour for HAAF is shown in Figure 3. The 65 LDN noise contour, which defines Zones I (<65 LDN), encompasses an area surrounding the immediate vicinity of the runway.

RGAAF, located on West Fort Hood, is a fully operational airfield designed to handle all types of fixed-wing and rotary-wing aircraft. Air Force C-5 aircraft also conduct touch-and-go pattern operations at RGAAF.

Annual flight operations were derived from air traffic activity reports and information gathered from personnel. The activity modeled includes fixed- and rotary-wing flight operations. Fixed-wing aircraft operations at RGAAF average 61.7 daytime and 4.9 nighttime operations. In addition, five types of helicopters make up the majority of rotary wing operations at RGAAF. The AH-64, OH-58, UH-60, UH-1, and CH-47 conduct 80 percent of the total local military operations. Helicopter operations at RGAAF average 170.9 daytime and 16.7 nighttime operations.

The LDN noise contours for RGAAF are shown in Figure 2. This figure shows a 75 LDN noise contour surrounding the immediate vicinity of the runway and extending several thousand feet under the take-off and departure flight tracks.

A small portion of the 65 LDN noise contour is located outside the military reservation boundaries. This portion of the contour lies below the radar pattern and the west VFR pattern, with the primary contribution to the contour being the C-5 touch-and-go operations. The computer modeling analysis of RGAAF exhibits no noise intrusions on residential land uses.

BLAST NOISE COMPUTER MODELING—BNOISE

Fort Hood has over 60 different firing range complexes. These firing ranges essentially encircle the artillery impact area and include tanks, infantry fighting vehicles, helicopters, mortars, rockets, machine guns, and small arms weapon firing activity. In addition to the weapon firing that occurs on the ranges, there is also artillery firing (105-mm, 155-mm, and 8-in. howitzers) that occurs at the various artillery firing points that surround the ranges and impact within the designated artillery impact area at Fort Hood.

The BNOISE computer model consists of a weapon data base for only the large-caliber weapons that are the primary noise source at most military installations. The primary weapons of concern that are fired at Fort Hood include the M1 and M60 main battle tanks, the M2/3 bradley armored personnel carrier, TOW missiles, 2.75-in. rockets, hellfire missiles, dragon missiles, LAW (light anti-tank weapon) rockets, MLRS (multiple launch rocket system) rockets, mortars (60-mm, 81-mm, and 4.2-in.), and artillery (105-mm, 155-mm, and 8-in. howitzers). Most of this heavy-weapon firing activity occurs on approximately 20 of the total Fort Hood ranges and are mostly located along East and West Range Roads.

In order to generate the input data for the BNOISE computer model, it was necessary to summarize total weapon firing activity at the ranges over a 1-year period. Using the range control weapon firing logs, it was possible to generate an annual summary of the total rounds fired by the various weapon types used at each of the primary weapon firing ranges at Fort Hood. The same summary analysis was performed for the artillery firing data using information contained in the same range control weapon firing logs. These summary firing data, including firing point locations and impact locations, were used as input into the BNOISE computer model to generate C-weighted LDN noise contours. Using the computer-aided drafting (CAD) system, these noise contours were then superimposed onto a digital base map of the Fort Hood area.

Figure 12 contains the C-weighted LDN weapon-firing blast noise contours for both the artillery and the weapon-firing activity at the ranges. The larger 70 LDN noise contour defines the noise levels around the artillery impact area, whereas the smaller 70 LDN noise contours are associated with the weapon firing activity on the ranges as well as from the various artillery firing locations. The 62 LDN noise contour is almost totally contained within the boundary of the Fort Hood Military Reservation. Only a small area along the northeast boundary of Fort Hood (near Site C: Crittenberger) is contained within a Zone II area. Except for this location, there are no other off-base Zone II (62-70 LDN) or Zone III (>70 LDN) noise impact areas resulting from weapon-firing activity at Fort Hood.

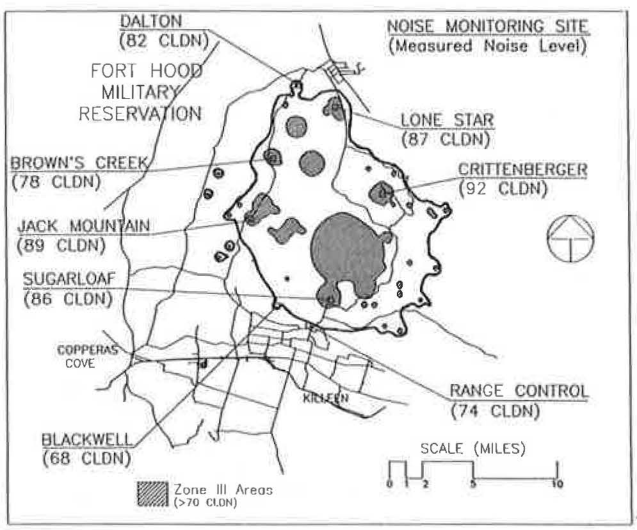


FIGURE 12 Blast noise: combined range firing and artillery firing (62 and 70 CLDN contours).

It should be noted that although the artillery firing noise contours do not extend beyond the Fort Hood Military Reservation, it does not mean that the artillery firing is inaudible in the residential areas surrounding Fort Hood. Because of the impulsive nature of artillery noise, the larger, noisier 8-in. artillery firing can be audible at distances of 10 mi (depending on wind direction and other atmospheric conditions) from the artillery firing point locations and impact area.

COMPARISON OF MEASURED AND PREDICTED NOISE LEVELS

One of the primary reasons for the extensive noise monitoring program at Fort Hood was to compare the actual measured noise data with the noise contours generated by computer modeling.

At Site 1, which is located adjacent to the south end of RGAAF, daily A-weighted LDN levels generally ranged between 58 and 62 dBA. Although there were days when the daily LDN levels exceeded 65 LDN, the overall (3-month logarithmic average) level at Site 1 was 62.7 dBA. Results of the computer modeling noise contours indicated that the LDN level at this location was 62.3 LDN.

At HAAF, Sites 4, 5, and 6 are located at residential noise monitoring sites near the airfield. The overall (3-month logarithmic average) LDN levels at all three sites are approximately identical at 63 dBA. Results of the computer modeling noise contours indicate that the LDN noise levels in this area around the perimeter of HAAF were approximately 62 dBA. Table 3 presents a comparison of the measured and predicted ALDN noise levels around the airfields. This comparison indicates that there is relatively good agreement between the measured data and the predicted LDN noise levels.

Comparison of the overall C-weighted LDN weapon-firing blast noise data from the Fort Hood firing ranges with the BNOISE computer modeling results is presented in Table 4.

TABLE 3 HAAF NOISE MEASUREMENTS: COMPARISON OF MEASURED AND PREDICTED ALDN NOISE LEVELS

	Measured Overall ALDN	Computer Modeling Results	Comparison
Robert Gray Army Airfield			
Site 1	62.7	62.3	+0.4
Site 2	61.8	58.5	+3.3
Hood Army Airfield			
Site 3	59.4	57.5	+1.9
Site 4	63.1	62.7	+0.4
Site 5	63.4	62.9	+0.5
Site 6	63.6	63.3	+0.3
Air Corridors			
Site 7	54.1	-	-
Site 8	59.1	-	-
Site 9	58.0	-	-

TABLE 4 FORT HOOD BLAST NOISE MEASUREMENTS: COMPARISON OF MEASURED AND PREDICTED CLDN NOISE LEVELS

		Measured Overall CLDN	Computer Modeling Results	Comparison
Site A:	Sugarloaf	85.5	88.1	- 2.6
Site B:	Range Control	74.1	61.6	+12.5
Site C:	Crittenberger	91.8	82.2	+ 9.6
Site D:	Lone Star	87.4	81.0	+ 6.4
Site E:	Dalton	81.7	73.1	+ 8.6
Site F:	Brown's Creek	78.3	84.2	- 5.9
Site G:	Jack Mountain	89.4	81.7	+ 7.7
Site H:	Blackwell	67.5	61.7	+ 5.8

Differences between these measured and predicted LDN levels are caused by a number of factors. For example, although the BNOISE computer model includes weapon-firing noise from helicopter rocket-firing activity on the ranges, it does not include the noise generated by the helicopter itself. In addition, military aircraft using the Shoal Creek bombing range generate flyover noise impacts along the northern ranges at Fort Hood that do not appear in the blast noise contours. As a result, measured noise levels at Site E (Dalton), which is a primary helicopter rocket-firing range, were 8 to 9 LDN higher than the results obtained from the computer model. In addition, Site D (Lone Star), which is primarily used by U.S. Army reserve units during weekend training exercises, had a higher-than-expected measured noise level because it is used as a FAARP for helicopters waiting to use the primary rocket-firing ranges.

Although Site B (Range Control) and Site H (Blackwell) were not located on active firing ranges, they were selected as blast noise monitoring sites so that artillery-firing noise levels could be evaluated without direct impact from range activity. However, noise impact from helicopter operations at HAAF, which uses South Range Road as a visual flight track corridor, resulted in higher measured noise levels than those predicted by the BNOISE computer model for weapon-firing activity only.

SUMMARY

Noise measurements at all of the airfield noise monitoring sites indicate that the overall average (logarithmic) LDN levels are below 65 dBA. Although there are days when LDN levels are above 65 dBA, these days generally correlate with high levels of helicopter activity at HAAF, or C–5 touch-and-go flight training activity at RGAAF. Noise contours developed for HAAF and RGAAF exhibit excellent agreement with the measured noise levels. This agreement is primarily because the airfield activity follows more well-defined flight tracks, and the level of activity is relatively constant over the entire year.

The results of the weapon-firing noise measurements and computer modeling analysis exhibit less agreement because of the complexity of military training activity at Fort Hood.

The results of this extensive airfield and blast noise monitoring program conducted at Fort Hood, together with the computer modeling used to develop noise contours for the airfields and the weapon-firing ranges, indicate that specific noise mitigation measures are not required.

REFERENCES

1. *Master Planning for Army Installations.* U.S. Army Regulation 210-20. June 1987.
2. Environmental Protection and Enhancement. U.S. Army Regulation 200-1. June 1982.
3. *Air Force Procedures for Predicting Aircraft Noise Around Airbases: Noise Exposure Model (NOISEMAP) User's Manual.* Report AAMRL-TR-90-011. Wyle Laboratories, Arlington, Va., Feb. 1990.
4. *MicroBNOISE: A User's Manual.* Technical Report N-86/12. U.S. Army Corps of Engineers, Construction Engineering Research Laboratory, Champaign, Ill., June 1986.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Effects of Road Surface Texture on Traffic and Vehicle Noise

GEORGE GLAZIER AND STEPHEN SAMUELS

Seven different pavement types, ranging from asphaltic concrete, open-graded and dense-graded, to cement concrete, with various surface textures, were required to be tested to obtain the relative noise from the interaction of tires on those pavements. Three methods were chosen, and these are discussed, together with an assessment of their respective merits and the overall results.

This study, conducted by the Australian Road Research Board (ARRB), grew out of a previous recommendation of one of the authors "that the use of 'quieter' pavement surfaces such as open graded asphaltic concrete (OGAC) be adopted through residential or other noise sensitive areas, in lieu of concrete (grooved, brushed, or smooth)—on new constructions."

The original intention of that recommendation was to involve testing of two, three, or at most four surface types by a comparative method in which two of the surface types could be found in suitable adjoining locations. These were to be OGAC, dense-graded asphaltic concrete (DGAC), and a few of the portland cement concrete (PCC) pavements (e.g., deep grooved, shallow grooved, hessian dragged, etc.).

Subsequently, the New South Wales Roads and Traffic Authority (RTA) decided to expand the study to cover a total of 10 pavement types, including examples of a cold overlay slurry seal, a 14-mm chip seal, and two examples of polymer-bound OGACs and normal OGACs to two different specifications (1975 and 1987).

TEST METHODS

Traffic Noise

At each site, four 15-min samples of traffic noise were collected using Bruel and Kjaer Type 4426 noise analyzers. These samples provided the conventional indices L10, L50, L90, and Leq as well as histograms of the traffic noise level distributions. At some sites, the L1 index was also monitored. During two of the four samples, tape recordings of the traffic noise were also made. Traffic volumes, composition, and speeds were monitored during all sample periods by means of the ARRB-developed video analysis data acquisition system. It comprises a video camera mounted atop a vertical mast attached to a small trailer. After suitably positioning the trailer adjacent to the road, the mast is extended pneumatically up to an optimum viewing height (not exceeding 10 m). Signals from the camera are recorded on a video cassette recorder mounted inside the trailer. On return to the laboratory, the tapes are replayed into an analysis system to determine the required traffic characteristics. Manual checks were also kept to complement these data.

Vehicle Noise

Sets of individual vehicle noise data were also collected at 10 sites. Peak roadside noise levels were measured for individual, isolated vehicles in the traffic as they passed the monitoring station. These data were obtained using a Bruel and Kjaer precision sound level meter (Type 2203). Vehicles were classified as either cars or heavies as per U.K. Department of Environment (1975) and the carriageway in which they were traveling was also noted. The passby speed of each vehicle was measured by means of a radar speed gun.

Comparative Noise

In several cases, two sites of differing surface type were within 2 km of one another along the one road. Under these conditions, data were monitored simultaneously at both sites. Such simultaneous monitoring was conducted at the following pairs of sites:

- OGAC (1987) versus DGAC;
- OGAC (1987) versus PCC, shallow-grooved;
- PCC, shallow-grooved versus PCC, hessian-dragged;
- PCC, hessian-dragged versus cold overlay slurry seal; and
- DGAC versus PCC, deep-grooved.

General

All data were collected in accordance with the provisions of Australian Standard 2702 under still dry condition—the latter despite an unusually wet period in and around Sydney.

Further, the physical dimensions of each site were recorded, including the distance of the microphone from the center of the nearest travel lane—generally 4 to 7 m.

RESULTS

Traffic Noise

L10 (1-hr) levels were predicted for each site and for each sampling period within each site, by the U.K. Department of

Wilkinson Murray Griffiths Pty Ltd., 246A Willoughby Road, Crows Nest, New South Wales 2065, Australia.

Environment (1975) model, using the measured traffic characteristics and site dimensions—each carriageway separately, then combined. The resultant site sound pressure levels (SPLs) were then related to a chosen reference condition to give a correction factor, *b*, for each site and sample period. Overall, this factor varied from −3.0 to +2.1 dB(A), which is a considerable range.

This correction factor, *b*, quantifies the between-site and between-sample variability. By applying it to the measured levels, all variable factors, except the different pavement textures, were eliminated. The resultant L10 (1-hr) levels are presented in Table 1, together with the L10 (1-hr) values related both to the quietest pavement type (open-graded asphaltic concrete, 1987 specification) and the more commonly used (to date) dense-graded asphaltic concrete pavement.

The deep-grooved PCC pavement was clearly the one from which emanated the most noise from traffic. This conforms to the decree of public disapproval expressed. That texture of pavement is unlikely to be used again in a residential area. The shallow-grooved PCC was shown to be the next loudest, of the order of 3.4 dB(A) quieter than the deep-grooved PCC. The third PCC pavement tested was the hessian dragged PCC, which was found to have an advantage of the order of 3 dB(A) over the shallow-grooved PCC.

In regard to the asphaltic concrete pavements, the 1987 specification for OGAC appeared to have an acoustic advantage of the order of 2 dB(A) over the 1975 specification OGAC and of the order of 6 dB(A) over dense-graded AC, with the cold overlay slurry seal performing slightly better than DGAC [of the order of 1 dB(A) better].

Vehicle Noise

At each site, the vehicle noise data set contained the following information for each vehicle:

- Peak passby noise level,
- Vehicle type,
- Vehicle speed, and
- Lane of travel.

The road surface effects were investigated by comparing the vehicle noise levels having due regard for the (between- and within-site) propagation distances and the vehicle speed. The former was achieved by applying the inverse square law, whereas the latter involved the 40 log (speed) relationship. With speed, all data were corrected to 80 km/hr, whereas with distance a reference value of 4.4 m was adopted. The reference distance was the nearside lane propagation distance for one of the dense-graded asphaltic concrete sites. Including the speed corrections, the vehicle noise corrections spanned a range of around 10 dB(A) in total sites. Aggregated, corrected vehicle noise levels are plotted in Figure 1, and show the same trends as the traffic noise data (Figure 2).

TABLE 1 RANKING OF SURFACE TYPES BASED ON TRAFFIC NOISE

Road Surface Type	L10 (1h) dB(A)	δ L10 (1h)	δ L10 (1h)
PCC "deep grooved"	87.2	3.8	9.7
PCC, "shallow grooved"	83.8	0.4	6.3
Dense graded asphaltic concrete	83.4	0	5.9
Cold overlay slurry seal	82.6	-0.8	5.1
PCC, "hessian dragged"	80.8	-2.6	3.3
Open graded AC (1975 spec)	79.3	-4.1	1.8
Open graded AC (1987 spec)	77.5	-5.9	0

Table 2 indicates similar trends for the different pavement types (for vehicle noise) to those trends for traffic noise in Table 1, given the variability in both vehicle and traffic noise levels at each site. In Table 2, there are three surface types not included in Table 1. The 14-mm chip seal (CS) and the two polymer-based OGAC pavements were tested only for vehicle noise, as described earlier, and not for total traffic noise. The rankings in Table 2 are given as in Table 1, i.e., related both to the quietest pavement type (OGAC) and to dense GAC, and in the case of Table 2 for both cars and heavy vehicles (those weighing in excess of 1,545 kg).

(Heavy vehicles were not measured in the case of the two polymer-based OGAC pavements because this was in the outside lane only, near the end of a long uphill run, where the noise from heavy vehicles was dominated by engine and exhaust components.)

Comparative Noise Measurements

With pairs of sites on the one road (same traffic, same mix of light and heavy vehicles, and same speeds) separated by up to 2 km, the only variables remain the site dimensions (in some cases) and the textures of two different pavement types.

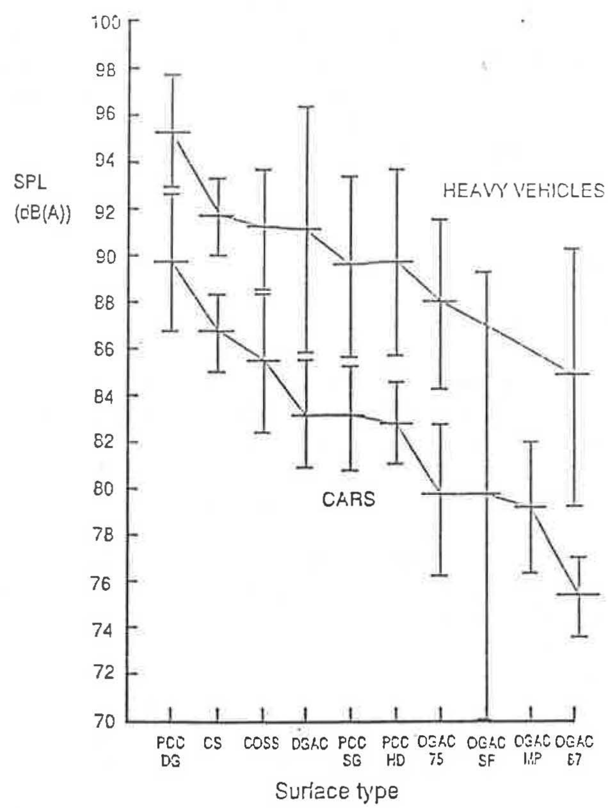


FIGURE 1 Vehicle noise data shown as a mean level plus or minus one standard deviation for each road surface type.

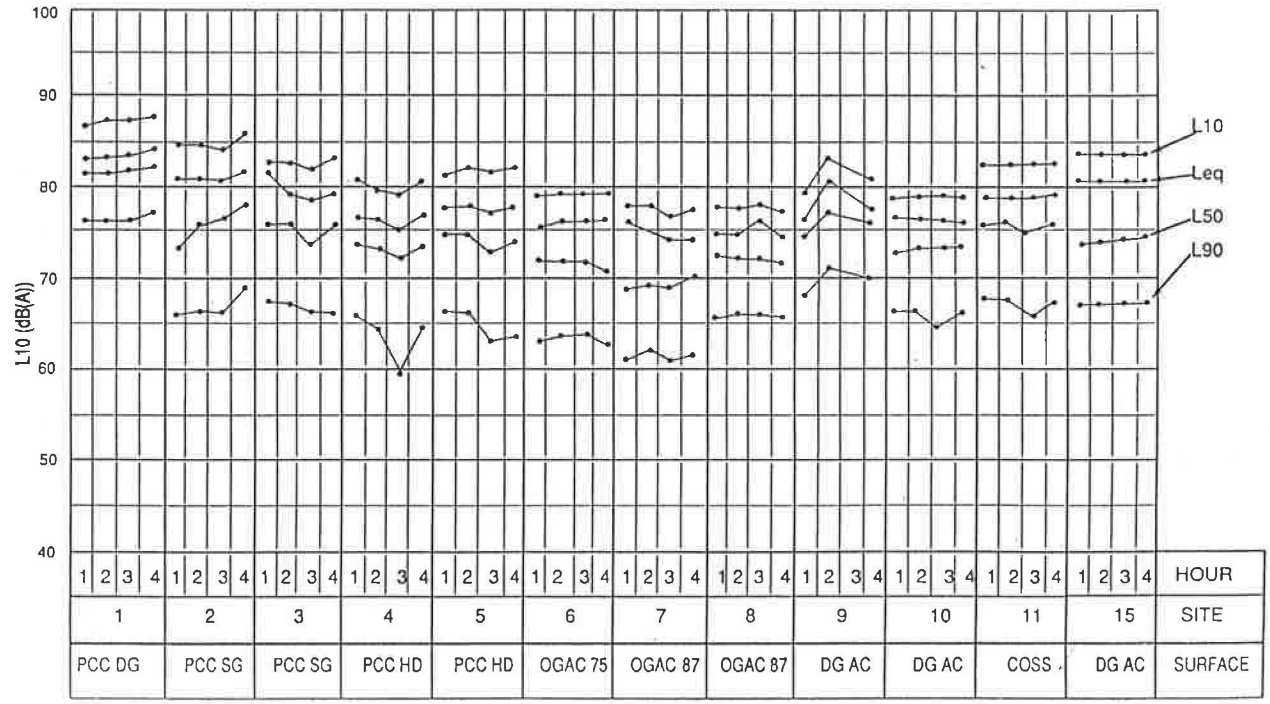


FIGURE 2 Distribution of measured traffic noise levels.

TABLE 2 RANKING OF SURFACE TYPES BASED ON INDIVIDUAL VEHICLE NOISE

Road Surface Type	CARS			HEAVY VEHICLES		
	SPL dB(A)	δ SPL dB(A)	δ SPL dB(A)	SPL dB(A)	δ SPL dB(A)	δ SPL dB(A)
PCC, DG	89.8	6.7	14.4	95.4	4.7	10.7
Chip seal (14mm)	86.7	3.6	11.3	91.7	1.0	7.0
COSS	85.5	2.4	10.1	91.1	0.4	6.4
DGAC	83.1	0	7.7	90.7	0	6.0
PCC, SG	83.0	-0.1	7.6	89.3	-1.4	4.6
PCC, HD	82.8	-0.3	7.4	89.7	-1.0	5.0
OGAC (1975 spec)	79.5	-3.6	4.1	88.1	-2.6	3.4
OGAC - Sealflex	79.5	-3.6	4.1	--	--	--
OGAC - Mobilplas	79.1	-4.0	3.7	--	--	--
OGAC - (1987 spec)	75.4	-7.7	0	84.7	-6.0	0

After any necessary corrections for site dimensions were made, two sites at a time were compared—those pairs of sites listed earlier.

Thus, it was expected to be able to produce a ranking of pavement textures from loudest to quietest. However, the first ranking that resulted was vastly different from the previously determined rankings and it was realized that this was obviously caused by differences in textures even between the two OGAC (1987 specification) pavements (for example), and between the two dense GAC pavements, because in each case these were on different roads, as also were the two PCC hessian-dragged pavements, etc.

Knowing this fact, the two pavements of each type were averaged, then the average values were compared, which gave answers similar to those given earlier and in Table 1, except for the dense-graded AC. The resultant rankings based on those average L10 (1-hr) levels are presented in Table 3.

The difference in ranking of the dense-graded AC pavement as presented in Table 1 and in Table 3 is believed to be because one particular site was excluded in the Table 1 preparation but included in the Table 3 preparation. At that site,

TABLE 3 RANKING OF SURFACE TYPES BASED ON COMPARING PAIRS OF PAVEMENT TYPES AND TRAFFIC NOISE

Road Surface Type	"Average" L10 (1h) dB(A)	δ L10 (1h) dB(A)	δ L10 (1h) dB(A)
PCC "deep grooved"	87.2	7.4	9.4
PCC, "shallow grooved"	83.6	3.8	5.8
Cold overlay slurry seal	82.6	2.8	4.8
PCC, "hessian dragged"	81.0	1.2	3.2
Dense graded AC	79.8	0	2.0
OGAC (1987 spec)	77.8	-2.0	0

the traffic was relatively close to an interchange, with some variations in speed of traffic as it entered the freeway or as it slowed to exit.

CONCLUSIONS

On the basis of the results and analyses, the following conclusions may be drawn:

- Road traffic noise levels are affected by road surface treatment. For the surfaces investigated, the total range of L10 (1-hr) levels was almost 10 dB(A).
- Vehicle noise observations supported those of the traffic noise. Road surface treatment had similar effects on both car and heavy-vehicle noise levels.
- By far the loudest surface was the dense-graded PCC. For the remaining surfaces, the noise levels ranged through approximately 6 dB(A).
- The quietest surface was clearly the OGAC 87, which was approximately 3 and 6 dB(A) down on the hessian-dragged and surface-grooved surfaces, respectively.

- Mobilplas- and Sealflex-type polymer-bound OGAC surfaces produce approximately the same vehicle noise levels. In turn, these levels are approximately equivalent to those of the 1975 specification OGAC.
- A 14-mm chip seal could reasonably be expected to generate traffic noise levels in the order of 2 dB(A) higher than a DGAC and 8 dB(A) higher than the 1987 OGAC.

The following observations may be added:

- The tape recordings taken, as back-up for the field measurements, were given to the client. Frequency analyses from those tapes would have been interesting as further comparisons, but were not part of the brief so were not conducted by the consultants. If the tapes are available, even now those frequency analyses could be of interest.
- The PCC pavements have been described as deep grooved, shallow grooved, and hessian dragged. The RTA concrete pavement group have since advised that "grooved" is not an acceptable term, but that two methods of making the grooves have been used—brushing and tyning. Brushing is the drawing of a broom across the wet concrete surface transversely. Some of these, including the test pavement described as PCC, deep-grooved, have much deeper grooves than is actually required for skid resistance purposes. It is understood that this method has now been discontinued. Tyning is the forming of the grooves by drawing a special rake transversely over the wet concrete. This method produced grooves not as deep as by brushing, and more recently grooves even shallower to what is now quite common and is referred to as PCC, shallow-grooved.
- Some of the earlier tyned surfaces produced nonrandom grooves, but generally a random pattern has more recently resulted from this process. The PCC shallow-grooved pavement tested had a random pattern.
- The dates of laying the OGAC, both in the 1975 and 1987 specifications, were not determined and recorded. This is unfortunate, because it may have shown up different results as being caused by degradation of the voids with age. It is suspected particularly that this may have accounted for the difference in levels measured in the two different specifications of OGAC. It is considered that further testing of the degradation of OGAC with age, from an acoustic point of view, is warranted.
- When choosing sites, some OGAC pavements, even some laid down within 2 years, appeared to have texture characteristics almost of a dense-graded AC pavement. This pointed to the need for careful laying to be carried out if the acoustic (and probably other) benefits of OGAC are to be fully realized.
- Concurrently, a separate monitoring program was being carried out on a new section of freeway. Some sections of PCC, shallow-grooved pavement, and adjacent, hessian-dragged PCC pavement were covered with OGAC. Levels taken at houses adjacent to the freeway corridor, in the vicinity of the junction of the two PCC pavement types, before and after the laying of the OGAC (some houses in an elevated topography and others in a situation level with the freeway pavement), indicated an average improvement of 4.5 dB(A) in the L10 (18-hr) levels. This is close to the average value of the 6.3 and 3.3 values for these pavements indicated in Table 1, which tends to confirm the findings therein.

ACKNOWLEDGMENT

This study was conducted by the Australian Road Research Board under contract to the New South Wales Roads and Traffic Authority.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Full-Scale Testing of Single and Parallel Highway Noise Barriers

LLOYD HERMAN

The results of research conducted by North Central Technical College and cosponsored by Ohio's Thomas Edison Program and Empire-Detroit Steel Division are reported. Laboratory and full-scale outdoor testing of both reflective and absorptive highway noise barriers are discussed. The phenomena of single-barrier insertion loss degradation caused by the addition of a second parallel barrier was confirmed. Further, absorptive barriers evidenced performance benefits over reflective barriers in parallel tests. Single-barrier insertion loss tests supported barrier attenuation theory for the configurations tested.

An important part of comprehensive noise abatement programs devised to alleviate adverse noise conditions involves the use of noise barriers. Highway noise abatement efforts, in particular, have been in effect in the United States for more than 15 years. To date, over 700 mi of highway noise barriers have been constructed along roadways, and another 700 mi are planned for construction over the next 10 to 15 years.

In 1988, Ohio's Thomas Edison Program accepted a proposal submitted by Cyclops Corporation, Empire-Detroit Steel Division, and North Central Technical College to provide funding for noise barrier research.

This paper is a report of the second phase of a three-part program to study highway noise barriers. The three parts are

1. Laboratory testing of acoustical panels.
2. Full-scale barrier tests at an isolated roadway.
3. Field tests at highway sites under traffic operations.

A summary of the first part of the program is included as background information. Part 2 of the program, the full-scale barrier tests at an isolated location, was completed in 1990. The goal of this phase was to produce a data base for both the third phase of the program and future research.

RESEARCH OBJECTIVES

The first objective for this phase of the project was to study in detail a variety of single-barrier configurations. Both a comparison of their performance with established barrier attenuation theory (which considers the path length increase for diffracted sound waves over a barrier), and observation of any differences between reflective and absorptive barriers, were to be made.

North Central Technical College, 2441 Kenwood Circle, Mansfield, Ohio 44901. Current affiliation: Vanderbilt University, Box 1625, Station B, Nashville, Tenn. 37235.

A second objective focused on the degradation effect occurring when a second barrier is installed parallel to the first barrier on the opposite side of a roadway. This phenomenon of barrier performance degradation, attributed to multiple reflections between barriers, has been studied by a number of researchers in recent years. Mathematical models have been developed to predict the effect of these multiple reflections in highway noise barrier applications. Test data, for the most part, have been the result of either acoustical scale model measurements or field data collected under operating conditions at highway sites. As Bowlby et al. (*1*) have observed:

> Full scale field data are limited and generally inconclusive because of the difficulty in isolating the phenomenon. There is a need for more carefully controlled, full-scale data collection to complement the scale modeling results and to confirm the insertion loss degradation phenomenon.

On the basis of this need of a larger data base, the objective to perform full-scale parallel barrier tests was chosen.

Absorptive barriers have been used in lieu of reflective barriers in applications where multiple reflections were predicted to be a problem. As a third objective, absorptive barriers were to be compared with reflective barriers to determine their effect on multiple reflections.

The primary project objectives are summarized as follows:

1. To compare single-barrier configurations with barrier attenuation theory, and to observe any performance benefits of single absorptive barriers compared to single reflective barriers;
2. To obtain full-scale test data to document any degradation of barrier performance resulting from parallel barrier installations; and
3. To determine the effect of absorptive barriers in parallel configurations compared to reflective barriers.

The overall goal was to provide a data base from a variety of barrier configurations with a minimum of outside variables. Highway test sites serving traffic operations were ruled out, because they contain many uncontrollable variables. This decision resulted in a trade-off. The data itself would not have direct application to a similar highway situation. However, the data would be indirectly applicable to a much broader range of highway situations. That is, the data, being foundational in nature, could serve both as a building block for future research, including Part 3 of this research program, and as a basis for formulating conclusions regarding specific highway applications.

NOISE BARRIER RESEARCH

Acoustical Panel Design

The absorptive panels tested were 2 in. thick and contained 1.5 in. of mineral rock wool with a solid steel back and a perforated steel front cover. The mineral rock wool had a density of 6 lb/ft^3. The perforated steel front cover contained a minimum perforation area of 15 percent. This amount of perforation area produced a cover that was essentially transparent to the sound striking the surface of the panels.

Laboratory Testing

Acoustical test panels were evaluated at Riverbank Acoustical Laboratories. The sound transmission properties of the test panels were evaluated using the ASTM standard test methods for sound transmission loss: E90-87 and E413-87. The sound transmission classification (STC) for the test panels was 33 when the seams between panels were sealed; however, in the unsealed condition, the STC rating was reduced to 26. Because an STC of 26 met the design goals for the test panels, it was decided that initial testing in the outdoor full-scale tests would not include sealed seams. Absorption testing was done according to the ASTM *Standard Test Method for Sound Absorption and Sound Absorption Coefficients by the Reverberation Room Method*: ASTM C423-87 and E795-83. The test panels with 1½-in.-thick mineral rock wool (6 lb/ft^3) produced a noise reduction coefficient of 0.95.

Full-Scale Outdoor Testing

Site Configuration

Test objectives called for an isolated outdoor test site where multiple barrier configurations could be evaluated under controlled conditions. The site chosen was an unused roadway in an open area at the Mansfield Lahm Airport in Mansfield, Ohio. A 16-ft-wide north-south roadway was located in a large grass-covered field. A wood support structure was erected on both sides of the roadway 50 ft from the centerline. The structure was designed to test barriers up to 250 ft long and 16 ft high in single or parallel arrangements.

Sound Source

The sound source used for the full-scale testing was a stationary, controlled, artificial point source. An artificial source was chosen to produce a high level of energy in all the frequency bands of interest. A point source was chosen because most of the data base from previous scale model research used a stationary point source. The type of tests scheduled were comparison tests requiring a high level of control and precision. There was no need to produce highway traffic conditions for these tests, but there was a need for repeatability with a minimum of variables. A moving point source was tested initially, but it proved more difficult to monitor random changes in background levels to determine the validity of a given sample. Also, sample time was limited with the moving source at any given position. Therefore, most of the tests were done with the point source held stationary. A Tracoustics Model NS-100 producing pink noise was selected. This source with its pink noise capability provided enough energy in a broad range of frequency bands to prevent masking caused by background levels in most cases. Dispersion testing was conducted to characterize the polar response of the speaker. Sound propagation from the speaker proved uniform and symmetrical. Barrier tests were made with the sound source in various positions; however, the most effective position proved to be with the speaker facing up. This position eliminated directionality effects that were especially important for the parallel barrier tests. The source heights used in the tests were 4, 8, and 12 ft.

Receivers

The receiver microphone was located, in all cases, on the opposite side of the barrier from the source, on the normal line with the source at 25, 50, 75, and 100 ft from the near barrier, at a height of 5 ft. A Larson-Davis Laboratories free-field ½-in. condenser microphone and preamp were used at the receiver positions.

Instrumentation

A Larson-Davis Laboratories dual-channel real time analyzer (Model 3100) was used to record ⅓-octave unweighted levels for a frequency range up to 10,000 Hz. In addition, all tests were recorded on a TEAC TASCAM dual-channel tape recorder.

Test Method

The test method used was taken from Methods for Determination of Insertion Loss of Outdoor Noise Barriers (*2*).

Test Environment

Testing began in August 1988 and ended in November 1989, a period much longer than had been originally anticipated. During this time, tests were run with few exceptions on all suitable days. However, moderate wind conditions prevail at this site, severely limiting the number of suitable days for testing. It was found that wind velocities greater than 7 mph distorted measurements and generally hindered data collection. In addition, adherence to test standards eliminated many days when temperatures, cloud cover, and humidity levels exceeded allowable variations. High levels of background noise because of the airport location further complicated the testing process. Nighttime testing was considered as a solution to the airport noise problem. It was rejected, however, out of concern for the effect on test microphones of the nighttime humidity levels (generally 100 percent at this location).

Test Procedures

All test equipment was set up for an individual test and taken down each day. Once equipment was in place, a checklist was used to review all settings for the instrumentation used. After settings were verified, a calibration procedure was begun using a Larson-Davis calibrator. Serial numbers of all equipment, including the calibrator, were recorded for each test and included in the test data file. All tests used the same receiver microphone. The microphone was calibrated as a first step, then the real time analyzer and microphone were used to record the output of the sound source. This procedure was a verification to ensure that the sound source would not vary in its output over time. This check was made at the beginning and the end of each test. In addition, the output voltage of the sound source was monitored with a voltage meter. This voltage was recorded at each change of the sound source height during a test. This procedure provided even further assurance that the sound source remained constant throughout the test. (Initially, and according to the specifications of the S12.8 test standards (*2*), a reference microphone located above the near barrier was used to monitor the sound source. However, the addition of a far barrier increased the sound level at the reference microphone position, thereby giving a false indication of sound source change. Therefore, the use of the reference microphone was abandoned in favor of using the procedure described to monitor the sound source output). Testing each barrier configuration then involved four microphone positions. The microphone was placed at 25, 50, 75, and 100 ft from the near barrier. These four microphone positions were used for sound source heights of 4, 8, and 12 ft. Therefore, 12 separate combinations were tested for each barrier configuration. Three samples, of 10 sec duration each, were taken at each microphone position. Microphone calibration checks were made at least three times throughout the test. Background levels were monitored before and after each sample and were recorded on the real time analyzer at least four times for each test.

Data Reduction

Raw test data produced by the real time analyzer was in the form of unweighted sound pressure levels for each of the ⅓-octave frequency bands measured. These sound levels were keyed into a spreadsheet for each of the three samples taken, for each of the test runs completed. The sound levels for all three samples at each frequency band were then averaged. The average value was then adjusted for any calibration changes noted during the process of the measurements. Next, the averaged and adjusted values were corrected for any background interference that may have occurred. Finally, the insertion loss was determined by subtracting the adjusted and corrected average sound pressure levels with a barrier from the corresponding sound pressure levels without a barrier.

By preserving the insertion loss as expressed by frequency bands, the usefulness of the data has been maximized. For example, the insertion loss by frequency band, for a particular barrier configuration, can be imposed on any traffic spectrum. By modifying the spectrum according to the insertion loss produced by the barrier at each frequency, a new and different spectrum will result. If desired, the spectrum can then be weighted, added on an energy basis, and expressed as an overall level (i.e., dBA). This procedure has the effect of converting the performance of the barrier for pink noise to that for a selected traffic spectrum (considered as a point source).

Single Barrier Data Analysis

The sound reaching a receiver from a source some distance away is the result of that sound which travels in a straight line between the source and the receiver. By inserting a barrier between the source and the receiver, the sound propagation from that source to the receiver occurs in two major ways. Line-of-sight sound propagation, that is, from the source directly to the receiver, must go through the barrier itself. If the barrier characteristics produce a large transmission loss, then very little sound will actually be transmitted through the barrier itself. The second means of sound propagation from the source to the receiver is along a path in which the sound is diffracted or bent over the top edge of the barrier before it reaches the receiver.

As a general rule, if the transmission loss is at least 10 dB above the attenuation resulting from diffraction over the top of the barrier, the barrier noise reduction will not be significantly affected by transmission through the barrier (less than 0.5 dB) (*3*). The acoustical panels tested on this project produced sufficient transmission loss so that the transmitted sound can be neglected. Therefore, only the diffracted path of the sound needs to be considered when looking at the sound attenuation by the test barriers. To predict the barrier attenuation for a given barrier, the path length difference must first be calculated. The path length difference is the difference in the distance the sound must travel to go from the source to the top of the barrier and back to the receiver compared to the straight line distance between the source and the receiver. The greater the path length difference, the greater the barrier attenuation. The path length can be increased by increasing the height of the barrier or by moving the source or receiver closer to the barrier (*4*).

This simplified theory of barrier attenuation recognizes no difference between a reflective and an absorptive barrier. One of the research objectives was to set up a test in which any differences in performance between absorptive and reflective single barriers could be observed. The reflective barrier test used acoustical test panels in which the solid steel back faced the sound source. For the absorptive barrier test, the perforated front of the acoustical test panels faced the sound source. In this way the same panels were being tested to eliminate any differences from transmission loss, etc.

Before discussing the results of the analysis, it should be noted that the effect of an absorptive single barrier versus a reflective barrier has been addressed by other researchers. According to Simpson (*3*), for diffraction angles greater than 45 degrees, absorptive materials can influence the sound that is diffracted over the top of the barrier. However, in most highway situations it is rare to find a configuration in which the diffraction angle will approach that magnitude. For angles less than 45 degrees, use of absorptive materials is of little advantage in reducing noise levels.

L'Esperance has also studied the effect of single absorptive barriers compared to single reflective barriers. L'Esperance was consulted regarding the test barrier configurations for the possibility of increased attenuation caused by the absorption of the barriers. However, his analysis showed that absorption would have no effect at the small diffraction angles that were used for the test configurations on this project. It would appear then that absorptive barriers could exhibit a benefit over reflective barriers in single applications, when limited to situations where large diffraction angles occur.

The collected data was analyzed, as described earlier, and compared with the barrier attenuation theory described earlier. The graphs shown in Figures 1 and 2 are typical, and indicate the comparison between the calculated barrier attenuation and the experimental results. The solid line represents the calculated insertion loss caused by the barrier, for the particular geometry being considered. As indicated, this insertion loss for the barrier varies by frequency. Because the calculated values represent only the effect of path length difference and not other changes in the sound propagation, a relative comparison is in view here rather than an absolute comparison. Further, it is suspected that some of the scatter in the measured data may be the result of ground effects in

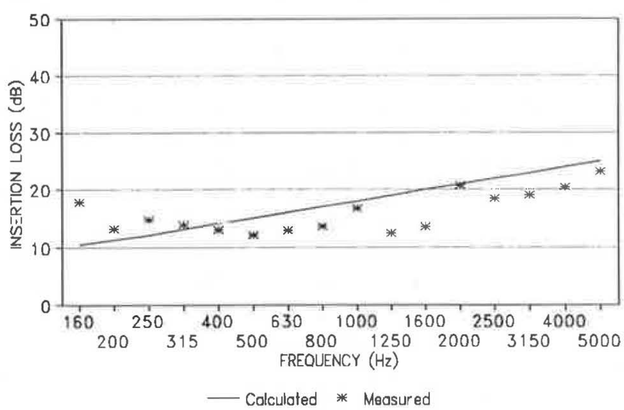


FIGURE 1 An example of single reflective barrier insertion loss test results (source 4 ft high, receiver at 50 ft).

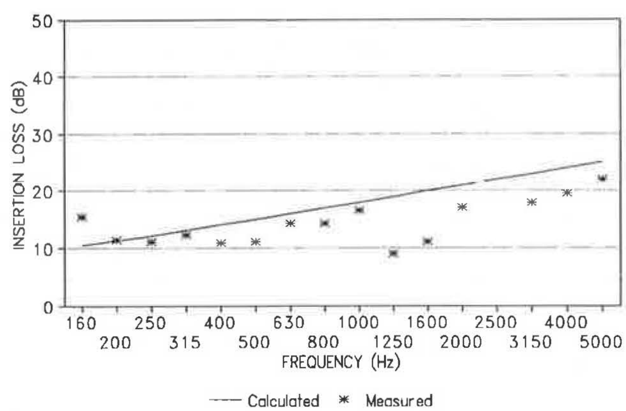


FIGURE 2 An example of single absorptive barrier insertion loss test results (source 4 ft high, receiver at 50 ft).

which both constructive and destructive interference affect the receiver levels for different frequencies.

Several observations can be made from the data in its analyzed form. First, the experimental results generally followed the expectations calculated from theory. Second, there is no significant difference between the absorptive single barriers and the reflective single barriers as shown in the comparison of Figures 1 and 2. Overall, the measured data support rather than contradict barrier attenuation theory. Therefore, on the basis of the analysis of experimental data for this project, there is no basis for assuming that an absorptive single barrier performs better than a reflective single barrier for the range of geometries tested. However, other geometries, particularly those with large diffraction angles may indeed produce greater insertion losses for absorptive barriers. Such geometries may exist in some highway situations. As one example, though not typical, barriers installed close to the near lane may fall into this category. Further, multiple reflections between heavy trucks and such single barriers could be reduced by using absorptive barriers.

Parallel Barrier Data Analysis

As stated earlier, previous research has uncovered a potential problem when a second highway noise barrier is added in parallel to an existing single barrier. It is believed by many researchers that the addition of a second barrier degrades the performance of the first barrier because of an excess noise buildup caused by reflections between the walls of the parallel barriers. Objective 2 involved tests for the parallel barrier degradation phenomena. Parallel barrier tests were conducted for barrier heights of 10, 12, and 14 ft and barrier lengths of 250 ft. The same panels were used in both reflective and absorptive tests. This was done to eliminate any differences in the transmission loss and related areas caused by the use of a different type of panel. For the reflective parallel barrier test, the solid steel side of the panels was exposed to the sound source located in the middle of the roadway. For the absorptive parallel barrier test, the acoustical panels were reversed, allowing the perforated face to be exposed to the sound source located in the middle of the roadway. The performance of the parallel absorptive barriers was then compared with the performance of the corresponding single absorptive barrier, and the performance of the parallel reflective barriers was compared with the performance of the corresponding single reflective barrier.

Parallel barrier analysis involved a determination of performance degradation for the parallel arrangement. The sound pressure level at each frequency band tested for the parallel barriers was compared with the sound pressure level at the corresponding frequency band for a single barrier of the same configuration. Figure 3 shows an example of a graphical plot of the test results for a particular barrier configuration. Each individual data point indicates the actual barrier degradation measured for a particular frequency band. The solid line is a calculated degradation based on a simplified model and provided as a reference. Over 40 of these plots were produced for the degradation tests. Because there were too many to include in this paper, only a sample is shown. This particular plot was chosen as an example, because it was a midrange test, having an 8-ft source height. (Source heights, as mentioned before, were at 4, 8, and 12 ft.)

Figure 3 is typical of the other plots produced and confirms the presence of the barrier degradation effect. For this particular case, with the sound source used, an overall degradation of 1.5 dBA was realized. However, some configurations exhibited less degradation, whereas others produced an overall degradation as high as 4 dBA. A word of caution is given here—the overall levels, which are for the source spectrum used, are merely mentioned as a reference. They are not particularly relevant to the discussion, and provide no direct application to a given highway situation. As stated earlier, the degradation at each frequency band is of interest. It is this degradation that can be used to modify any number of spectra to predict degradation for the same configuration. It is intended that this data be used for further study to support mathematical models that will predict barrier degradation for any configuration.

Not only was the phenomenon of parallel barrier degradation confirmed by the test data, but several trends in the data were also observed. The degradation tended to increase for lower source heights, higher barriers, greater receiver distances, and higher frequencies. However, there were numerous exceptions to these general trends. In addition, as shown in Figure 3, scatter of the data points was observed in each test. One area for further study, as mentioned under single-barrier data analysis, is to consider the possibility of geometric interference caused by ground reflections. Such a phenomena, if it has occurred for the barrier geometry tested, could account for some of these variations, including large negative degradation values.

Figure 4 represents absorptive barriers of the same configuration as the reflective barriers shown in Figure 3. Again, the degradation is the difference between the parallel and single barrier values, and the solid line represents the calculated value, for reference. The absorption coefficients used in the calculations were the actual values measured in the laboratory for each frequency band. For the particular source spectrum and barrier configuration represented in Figure 4, the overall A-weighted degradation was zero.

As stated earlier, this overall degradation is not transferable to other situations, and is only mentioned as a reference. Absorptive barriers were seen to reshape the source spectrum

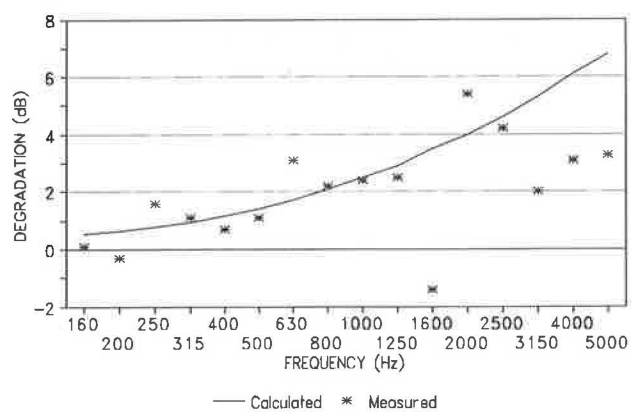


FIGURE 3 An example of parallel reflective barrier degradation test results (source 8 ft high, receiver at 25 ft).

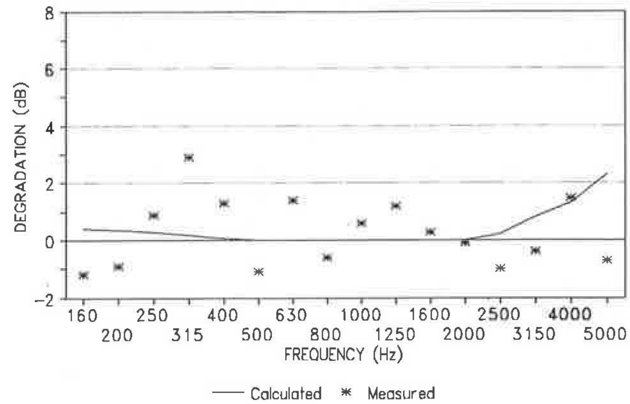


FIGURE 4 An example of parallel absorptive barrier degradation test results (source 8 ft high, receiver at 25 ft).

at the receiver by reducing the sound levels at those frequencies most affected by the absorptive material. By way of application, if a particular absorptive material proved effective at absorbing high-frequency sound but not low-frequency sound, little improvement would be realized by installing absorptive barriers of such material in parallel applications with predominately low-frequency noise. Following the same line of thought, an absorptive material, which provides very high absorption coefficients in the same frequency bands as the noise source, will eliminate the degradation problem. Most applications will fall in the middle of these extremes, with the results depending not only on the source spectrum and the absorptive characteristics of the barrier, but also on the source-barrier-receiver geometries.

In summary, parallel noise barrier testing on this project has demonstrated the performance degradation experienced with a parallel barrier arrangement. In addition, the absorptive barriers tested offer an effective means of reducing this degradation effect. The tests further expose the problem involved in any attempt to predict the degradation at one highway site on the basis of the amount of overall degradation experienced at another highway site. There is no single rule of thumb to be used, or typical degradation experienced. For highway applications, the amount of degradation experienced and the corresponding benefits of absorptive barriers depend on the many traffic, barrier, and geometrical parameters for a particular highway site. In order to predict the degradation for a particular site, a computer model that accounts for all these factors is required. As stated earlier, the test results are intended to serve as a data base to support further computer model development and refinement.

Combination Absorptive and Reflective Parallel Barrier Test

A test was made of a parallel barrier arrangement with one side reflective and one side absorptive. The sound source for this test was 8 ft high. Receiver positions were located on both the east and the west sides of the roadway. Analysis of the data indicated no significant difference in sound levels for the east and west receiver positions. Further, no significant reduction in sound level occurred as a result of one of the barriers being absorptive. Therefore, the case for reducing parallel barrier performance degradation by installing an absorptive barrier on one side of the roadway only, could not be substantiated from the test data. However, the relatively high source height of 8 ft reduced the number of reflections that could have contributed to barrier degradation. Testing with a lower source height may have demonstrated an advantage to having one of the barriers absorptive. A further consideration would be the receiver distances used for this test. Because barrier performance degradation with a parallel arrangement tends to be greater for greater receiver distances, a benefit of having one of the barriers absorptive may be realized but at the greater distances. Because relatively short distances were used in this test, this result could not be substantiated.

RESEARCH CONCLUSIONS

The following conclusions are based on the configurations tested. The results reported for this project are not intended for direct application to highway situations. However, they are intended to supplement existing data bases for future research and prediction model refinement.

1. The addition of a second reflective noise barrier installed parallel to an existing reflective barrier can degrade the performance of the existing barrier caused by multiple reflections between the barriers.
2. The expected performance degradation for various parallel barrier installations cannot be quantified by the same number. It is dependent on the source spectrum and source-barrier-receiver geometries, along with the absorptivity of the barrier itself. The many variables preclude a generalized answer for all cases. Models that account for these variables must be used to predict the results for individual cases.
3. The use of absorptive barriers instead of reflective barriers in parallel situations can significantly reduce and even eliminate the performance degradation effect caused by multiple sound reflections.
4. The use of one reflective barrier and one absorptive barrier in parallel situations does not significantly improve barrier performance compared to parallel reflective barriers. (This conclusion is based on limited testing.)
5. The experimental results for the single barriers tested were in good agreement with the single-barrier theory used, for the range of geometries tested.
6. The use of absorptive barriers instead of reflective barriers in single-barrier situations does not offer a performance advantage for the configurations tested. (This conclusion may not be valid for barriers close to the receiver or the source.)

RECOMMENDATIONS

Barrier Testing

1. A parallel barrier combination, with one barrier reflective and one barrier absorptive, should be tested using lower source heights as found with automobiles and medium trucks. This test may demonstrate a benefit for the combination, particularly at greater receiver distances.
2. Single absorptive and reflective barrier tests should be made for configurations with large diffraction angles to observe the potential performance benefit of absorptive barriers.

Further Study

The effect of ground reflections and their potential for producing both constructive and destructive interference patterns that may modify barrier insertion loss should be studied. This phenomena has the potential of giving explanation for some of the scatter observed in the measured data for the full-scale highway barrier tests.

ACKNOWLEDGMENT

The work performed by North Central Technical College and described in this paper was funded by the State of Ohio's Department of Development (Thomas Edison Program), and Empire-Detroit Steel Division of Cyclops Corporation. The support of both organizations is gratefully acknowledged. Empire-Detroit Steel, in addition, furnished all the acoustical panels for barrier tests along with the barrier structures, and installation.

REFERENCES

1. W. Bowlby, L. F. Cohn, and R. A. Harris. A Review of Studies of Insertion Loss Degradation for Parallel Highway Noise Barriers. *Noise Control Engineering Journal*, Vol. 28, No. 2, March–Apr. 1987, pp. 40–53.
2. *Methods for Determination of Insertion Loss of Outdoor Noise Barriers.* Report ANSI S12.8–1987, American National Standards Institute, New York, 1987.
3. M. A. Simpson. *Noise Barrier Design Handbook.* FHWA-RD-76-58. FHWA, U.S. Department of Transportation, 1976.
4. T. M. Barry and J. A. Reagan. *FHWA Highway Traffic Noise Prediction Model.* FHWA-RD-77-108. FHWA, U.S. Department of Transportation, 1978.

Publication of this paper sponsored by Committee on Transportation-Related Noise and Vibration.

Electrokinetic Soil Processing in Waste Remediation and Treatment: Synthesis of Available Data

YALCIN B. ACAR AND JIHAD HAMED

Electrokinetic soil processing is an innovative in situ technique to remove contaminants from soils and groundwater. The process is an alternative to conventional processes, with significant economic and technical advantages. This paper provides a comprehensive review of the literature on electrokinetic soil processing. The fundamentals of electrokinetic phenomena in soils and their potential use in waste management are presented. A synthesis of available data on the technique is presented. Engineering implications regarding the current-voltage regime, duration energy requirements, and soil contaminant characteristics are provided. This review indicates that the process can be used efficiently to remove ions from saturated soil deposits.

The need to remove contaminants from soil and groundwater, the high cost of current remediation techniques ($50 to $1,500 per cubic yard of soil), and limited resources lead to an everlasting strife to find new, innovative, and cost-effective in situ techniques for removal or separation of contaminants from soils.

The electrokinetic phenomena in soils are envisioned to be useful for removal or separation of organic and inorganic contaminants and radionuclides; for barriers and leak detection systems in clay liners; for diversion schemes for waste plumes; for injection of grouts, microorganisms, and nutrients into subsoil strata; and for in situ generation of hydrogen peroxide for remediation (*1–5*).

In the last five decades since its first application and use (*6*), the mechanics of consolidation by electroosmosis has been extensively investigated by geotechnical engineers. However, studies investigating removal of ions from soils by the electrokinetic phenomena are limited, possibly because of insufficient understanding of the electrochemistry associated with the process. The need to use the process in removal and separation of contaminants necessitates a good understanding of electrochemistry and its relation to mechanical behavior. Recent studies at Louisiana State University provided a better understanding of the electrochemistry and demonstrated that the acid front generated by electrolysis reaction at the anode advances and eventually flushes across the specimen by advection, migration, and diffusion (*4,7*). Hamed et al. (*8*) demonstrated that the movement of this acid front together with migration and advection of the cations and anions under electrical gradients constitute the mechanisms of removing contaminants from soils. The factors influencing the acid-base profile across the porous medium would significantly affect the flow, the flow efficiency, and the extent of ion migration and removal in electrokinetic soil processing.

Department of Civil Engineering, Louisiana State University, Baton Route, La. 70803.

ELECTROKINETIC PHENOMENA IN SOILS

Coupling between electrical, chemical, and hydraulic gradients is responsible for different types of electrokinetic phenomena in soils. These phenomena include electroosmosis, electrophoresis, streaming potential, and sedimentation potential, as shown in Figure 1. Electroosmosis and electrophoresis are the movement of water and particles, respectively, caused by application of a small direct current. Streaming potential and sedimentation potential are the generation of a current by the movement of water under hydraulic potential and movement of particles under gravitational forces, respectively. The effect of coupling becomes more important in fine-grained soils with lower coefficients of permeability.

In electroosmosis, electrodes can be placed in an open- or closed-flow arrangement. Open-flow arrangement constitutes the case when an electrode is sufficiently permeable to admit ingress and egress of water. In the closed-flow arrangement, the electrode is not permeable or porous. Different electrode configurations (open or closed) result in substantial variations in the total matrix potentials across the soil specimen.

Figure 2 shows a schematic diagram of the matrix potentials (pressure head) developed in a soil specimen under constant current (or voltage) conditions. The matrix potential shown in Figure 2 represents the initial conditions and will change with time as the chemistry across the cell changes. In Case I of Figure 2, the water in a reservoir behind the anode would permeate through the soil to displace the pore fluid.

In Case II, where ingress of water is prevented at the anode, negative matrix potentials are generated in the soil, increasing the confining stress proportionally until equilibrium conditions are reached by the flow of water from the anode to the cathode. This arrangement has often been used by geotechnical engineers to consolidate soil deposits. In Case III, positive pore pressures are developed, reducing the confining stress, which may potentially lead to swelling. In Case IV, the matrix potential drops at the anode and increases at the cathode. The water level will rise in the cathode side and drop at the anode side.

The electroosmotic flow rate, q_e, is defined with an empirical relationship,

$$q_e = k_e \, i_e \, A = k_i \, I = \frac{k_e}{\sigma} I \quad (1)$$

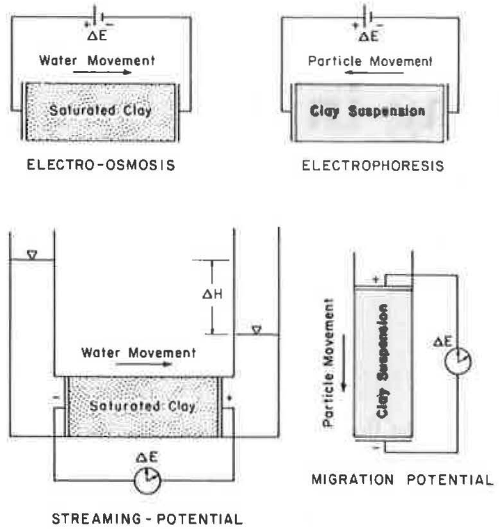


FIGURE 1 Electrokinetic phenomena in soils (*9*).

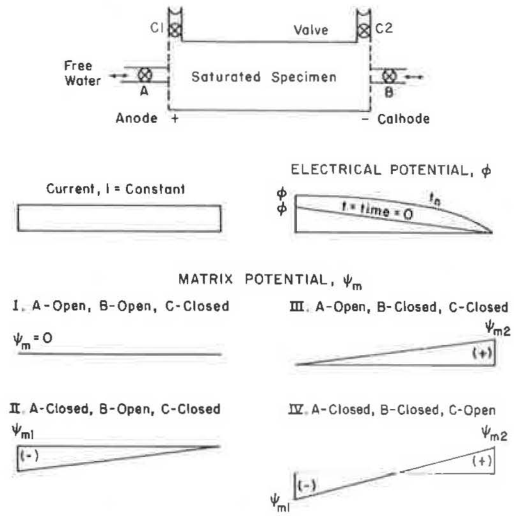


FIGURE 2 Schematic diagram of matrix potential generated with different electrode configurations (*4*).

where

k_e = coefficient of electroosmotic permeability (cm²/sec-V),
k_i = electroosmotic water transport efficiency (cm³/amp-sec),
I = current (amp),
σ = conductivity (siemens/cm),
i_e = electrical potential gradient (V/cm), and
A = cross-sectional area (cm²).

Estimates of electroosmotic flow rates can be made using Equation 1. The values of k_e measured at earlier stages of processing (within hours) vary within one order of magnitude for all soils; 1×10^{-5} to 10×10^{-5} (cm²/V-sec), the higher values being at higher water contents. Further processing results in decreases of k_e values.

Figure 3 shows a schematic diagram of one-dimensional laboratory tests in electrokinetic soil processing. The prevailing electrical gradients and the ion flow are also depicted. A comparison of flow under electrical and hydraulic gradients in clays is also provided. The coefficient of electroosmotic permeability is independent of the size and distribution of pores (fabric) in the soil mass. However, hydraulic conductivity is most affected by the fabric. Therefore, hydraulic conductivity decreases by five to six orders of magnitude (10^{-3} to 10^{-8} cm/sec) from the fine sands to clays. Figure 3 indicates that under equal gradients, electrical potentials in fine-grained soils would result in three orders of magnitude larger flows than hydraulic potentials. Therefore, electroosmosis-induced flow can be considered to be an efficient pumping mechanism in saturated, low-permeability, fine-grained soil.

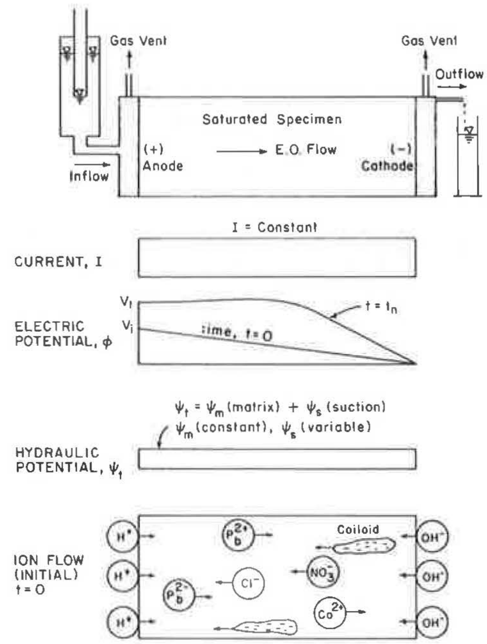


FIGURE 3 Schematic diagram of electrokinetic processing, ion flow, and comparison of flow in clays (*8*).

The efficiency and economics of electroosmotic dewatering is governed by the amount of water transferred per unit charge passed, which is quantified by electroosmotic water transport efficiency, k_i. The parameter k_i may vary from 0 to 1.2 cm³/amp-sec, depending on the electrical conductivity of the porous medium. The conductivity changes with water content, cation exchange capacity, and free electrolyte content in the soil and also because of the prevailing chemistry during electrokinetic processing. An approach to quantifying this amount in soil-water-electrolyte systems is proposed by Gray and Mitchell (*10*). Their study indicated that electroosmotic efficiency decreases with a decrease in water content and an increase in activity of the soil. The electroosmotic dewatering efficiency is independent of variations in electrolyte concentration (sodium ion) for active clays, whereas an increase in electrolytes tends to decrease the efficiency in inactive clays.

Recent studies by Lockhart (*11*) substantiate the conclusions of Gray and Mitchell (*9*). The efficiency parameter k_i increased from 0.32 to 1.20 cm³/amps-sec with the decrease of NaCl and HCl concentrations from 10^{-1} to 10^{-3} M; k_i decreased during an increase in electrical gradients, possibly because of a higher influx of H^+ ions (*8*). Lockhart (*11*) found that a higher electroosmotic efficiency was recorded with H and Cu clays. The efficiency parameter k_i changed in the order of H > Cu > Al > Na > Ca. Higher voltage gradients were required to initiate flow in Al clays.

POTENTIAL USES OF ELECTROKINETICS IN WASTE MANAGEMENT

The four electrode configurations described earlier could potentially be used in the following ways in waste disposal (*4*): (a) dewatering of waste sludge slimes, dredged spoil by first concentrating the solid particles using electrophoresis and subsequent consolidation by electroosmosis (*2*), (b) electroosmotic flow barriers (*12*), (c) leak detection systems for disposal facilities, (d) injection of grouts to form barriers, (e) provision of nutrients for biodegrading microcosm, (f) in situ generation of reactants such as hydrogen peroxide for cleanup or electrolysis of contaminants, and (g) decontamination of soils and groundwater. Figure 4 shows electrokinetic clay barriers, waste plume diversion schemes, and electroosmotic injection.

In situ remediation methods often necessitate the use of hydraulic charge and recharge wells to permeate the decontaminating liquid or stabilization agent through the soil deposit, or to provide nutrients for the biodegrading microcosm. Although such systems may effectively be used in highly permeable soils, they become inefficient and uneconomical in low-permeability silts and clayey deposits. Electrokinetic soil processing with open electrode configuration could also be used to achieve an efficient seepage and decontamination method in such soils.

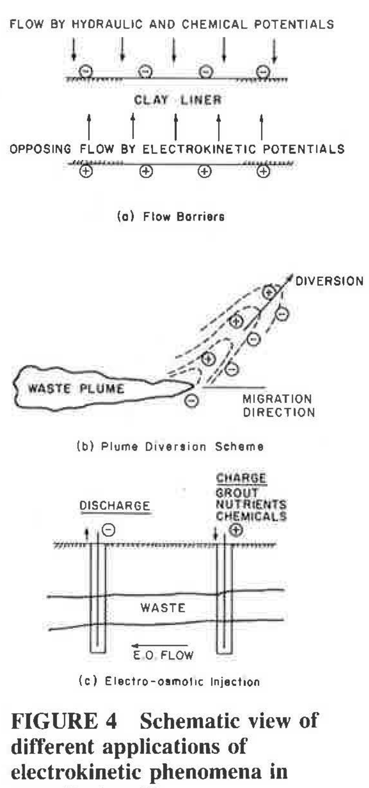


FIGURE 4 Schematic view of different applications of electrokinetic phenomena in remediation (*7*).

Removal of Contaminants by Electrokinetic Soil Processing

Upon application of low-level direct current (in the order of milliamps per cm^2 of electrode area) to the saturated porous medium, the following processes occur:

1. The water in the immediate vicinity of electrodes is electrolyzed. An acid front is generated at the anode while a base front is created at the cathode. Acar et al. (*4*) formalized the development of these fronts. The pH will drop to below 2.0 at the anode and will increase to above 12.0 at the cathode.
2. The acid front will advance across the specimen in time towards the cathode by
   a. Advection of the pore fluid caused by the prevailing electroosmotic flow,
   b. Advection of the pore fluid caused by any hydraulic differences,
   c. Diffusion because of concentration gradients, and
   d. Migration because of the electrical gradients.

Development of these acid and base distributions and movement of ionic species are formalized by Acar et al. (*7*).

3. The migration, diffusion, and advection will also result in movement of cations and anions to respective electrodes in the porous medium (*8,12*).
4. The acid advancing across the specimen exchanges with adsorbed cations in the diffuse-double layer, resulting in their release into the pore fluid and advance towards the cathode by advection and diffusion (*8*).
5. In case the generation of the acid front is not controlled at the anode, the electrolyte concentration inside the porous medium will gradually increase, resulting in increased conductivity in the vicinity of the anode, decrease in electroosmotic flow (*8*), and a corresponding decrease in bulk flow movement by advection.
6. The anion depletion at the cathode region may lead to an increase in voltage and an increase in energy expenditure during the process. The chemistry at the anode and the cathode should ideally be controlled to achieve continued advection while providing sufficient H^+ ions for desorption of contaminants and solubilization of salts. This is possible either by decreasing the current to levels where pH is at a desirable level or by frequent flushing at both ends by a fluid of controlled pH and chemistry.

Synthesis of Available Data

Studies investigating removal of ions from soils by electroosmosis are rare, possibly because of difficulties in understanding the chemistry. Table 1 (*13–20*) provides a synthesis and analysis of laboratory studies that reported some form of data related to ion removal from soils. In an early study by Puri and Anand (*13*), leaching of Na^+ ions was detected in the effluent in electroosmotic consolidation. Puri (*21*) suggested that in electroosmosis monovalent ions will move faster than divalent ions because of the former's higher dissociation from the clay surface. It is also noted that movement of ions was low at low water contents and significantly increases by an increase in water content.

Jacobs and Mortland (*14*) demonstrated that Na^+, K^+, Mg^{++}, and Ca^{++} ions can be leached out of Wyoming ben-

TABLE 1 ANALYSIS OF LABORATORY DATA REPORTED FOR REMOVAL OF CHEMICALS BY ELECTROKINETICS

Soil Type/ Chemical	Concentration ($\mu g/g$) Initial	Concentration ($\mu g/g$) Final	Current Density and/or Voltage (mA/cm^2) or [V/cm]	Duration (hr)	Charge $\frac{amp\text{-}hr}{m^3}$	Energy kWh/m^3	Remarks
1. Puri and Anand (*13*)							A Buchner funnel was used in testing. The diameter was 18 in. Cathode is circular brass plate. Anode consists of five cylindrical bars arranged symmetrically along the circumference. Na normality of percolate increased up to 0.6 N. The effluent was 90% NaOH, 10% Na_2CO_3.
High pH Soil - Na^+	N/A	N/A	9.3-14.9 [20.0]	8 (intermittent)	N/A	N/A	
2. Jacobs and Mortland (*14*)							1-D tests. Cylindrical specimens (D = 0.75 in., L = 1 in.). Circular platinum electrodes. Rate of removal of monovalent ions were directly related to the amount remaining in the specimens. The rate of removal was in the order of $Na^+ > K^+ > Mg^{2+} > Ca^{2+}$. Na^+ was removed more efficiently than all other ions. Tests were discontinued when most Na^+ was removed. Concentrations reported are in symmetry units which is described as normalized concentrations.
5% Bentonite/95% Sand							
Na-Ca	.59/.57	0.0/(N/A)	0.32-0.64 [N/A]	10-140	2-20	N/A	
Na-Ca-Mg	.44/.3/.4	0.0/(N/A)					
Na-K	.65/.41	0.0/(N/A)					
Na	0.78-1.11	0.0					
Ca	0.86-1.0	.26-0.30					
K	0.79	0.0					
3. Krizek, et al. (*15*)							Slurries (w = 100 - 142%) from discharge pipes and contaminated bottom sediments are tested. Cylindrical 1-D consolidation tests (D = 14 cm, L = 25 cm). The concentrations are the initial and final effluent values.
Slurry/Sediment							
No. 1 - Na	320	330	[0.5]	150	N/A	20	
K	7	290					
Ca	65	640					
NH_3-N	138	320					
No. 2 - Na	360	205	[1.0]	150	N/A	65	
K	15	160					
Ca	340	5800					
NH_3-N	172	456					
4. Hamnet (*16*)							A 3-D laboratory model. 35 cm by 8 cm, with 5 cm depth. Carbon rods were used as electrodes. After 12 hours, concentration at anode was a third of the cathode side. The data shows movement of ions to respective electrodes.
Silica Sand - NaCl Heavy Clay	3%	N/A	6.25-16.25 [0.6]	12-20	N/A	N/A	
5. Runnels and Larson (*17*)							1-D tests. Cylindrical specimens (D = 0.75 in., L = 6 in.). Square platinum electrodes (1 in^2). Quartz sand washed with HCl to remove impurities.
Silty Sand - Cu(II)	617	290-543	0.01-0.05 [.165]	24-72	N/A	N/A	
6. Renauld and Probstein (*3*)							1-D cylindrical specmens. 30 cm in length, 8 cm in width, and 5 cm in depth. Tests were conducted to assess electro-osmotic water transport efficiency. Efficiency increased with increasing concentration of acetic acid.
Kaolinite - Acetic Acid	0.5-1.3	N/A	1.47 [1.16]	6.7	328	11.5	
7. Thompson (*18*)							A tubular, 3 section test set-up is used. The middle section contained the chemical in solution, the other two sections contained the soil specimen, and contaminant movement into the cathode section was monitored. Transport was a function of pH of the soil.
Ottawa Sand							
SiO_2 flour $Cu(NO_3)_2$	0.01 N	N/A	N/A [25]	336	N/A		
8. Lageman (*19*)							Details of experiments are not available. From a presentation by Lageman (1990), it is understood that 1-D tests are conducted with rectangular electrodes. The pore fluid at anode/cathode compartment was flushed with a conditioning fluid intermittently in order to control the chemistry at the electrodes and sustain the advection in decontamination.
Peat							
Pb	9000	2400	N/A	N/A	N/A	101	
Cu	600	200	N/A	N/A	N/A	101	
Pottery Clay - Cu	1000	100	N/A	N/A	N/A	25	
Fine Clayey Sand - Cd	275	40	N/A	N/A	N/A	198	

(*continued on next page*)

TABLE 1 (*continued*)

Soil Type/ Chemical	Concentration (μg/g)		Current Density and/or Voltage (mA/cm^2) or [V/cm]	Duration (hr)	Charge amp-hr / m^3	Energy kWh/m^3	Remarks
	Initial	Final					
8. Lageman (*19*) (cont.)							
Clay - As	300	30	N/A	N/A	N/A	207	
Fine Clayey Sand							
Cd	319	< 1	N/A	N/A	N/A	54	
Cr	221	20					
Ni	227	34					
Pb	638	230					
Hg	334	110					
Cu	570	50					
Zn	937	180					
River Sludge							
Cd	10	5	N/A	N/A	N/A	180	
Cu	143	41					
Pb	172	80					
Ni	56	5					
Zn	901	54					
Cr	72	26					
Hg	0.50	0.20					
As	13	4.4					
9. Banerjee, et al. (*20*)							Eight cylindrical 1-D tests were conducted on specimens brought from the field (D = 5.1 cm, L = 2.5 cm to 6.7 cm). Electrodes used were Ni-Cu wire mesh. Hydraulic and electrical potentials were applied simultaneously in order to facilitate removal.
Silty/Silty Clay - Cr	2460	22	N/A [0.1-1.0]	24-168	N/A	N/A	
	2156	12					
	870	50					
	704	19					
	642	22					
	532	37					
	234	3					
	148	10					
10 Hamed, et al. (*8*)							1-D tests. Cylindrical specimens (D = 4 in., L = 4 in. and 8 in.). Circular graphite electrodes. Initial conductivity of specimens 75-86 μs/cm. Rose up to 1000 μs/cm at the anode, dropped to 22 μs/cm at the cathode after the process. Pb(II) movement and electrochemistry across the specimens are reported.
Georgia Kaolinite - Pb(II)	118-145	7-40	0.037 [≤ 2.5]	100-1285	362-2345	29-60	
11 Mitchell and Yeung (*12*)	N/A	N/A	N/A	N/A	N/A	N/A	Investigated the feasibility of using electrokinetics to stop migration of contaminants. Electric field slowed down the migration of cations and increased the movement of anions. k_e did not display a marked change by an increase in backpressure, molding water content and dry density.

tonite by electroosmosis. Figure 5 presents the amount of the ions removed versus the electroosmotic flow in bentonite-sand mixtures. Monovalent ions are removed at a faster rate.

Krizek et al. (*15*) found that the soluble ions content substantially increased in effluent in electroosmotic consolidation of polluted dredgings, while they also noted that heavy metals were not found in the effluent during electroosmosis. Hamnet's tests demonstrated that Na^+ ions move toward the cathode, whereas Cl^- and SO_3^{-2} ions move toward the anode. This study demonstrated that ions could be removed from soils by the process.

Shmakin (*22*) notes that the method has been used in the Soviet Union since the early 1970s as a method for concentrating metals and exploring for minerals in deep soil deposits. He mentioned its use in prospecting for Cu, Ni, Co, and Au. A porous ceramic probe with HNO_3 is placed at the cathode. This acid prevents precipitation of ions at the cathode. Placement of this acid probe is specifically necessary if the base

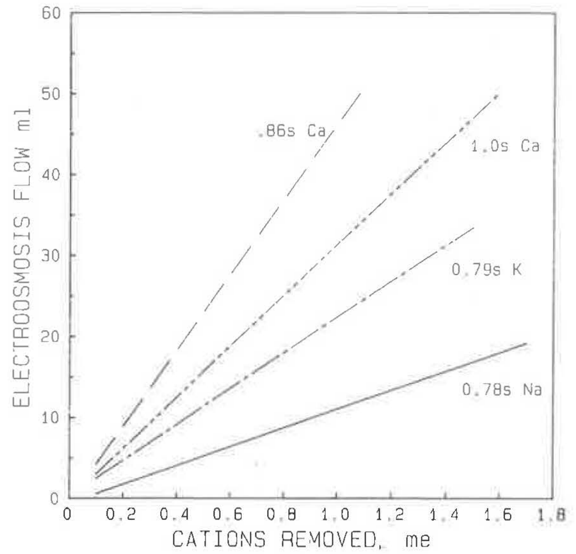


FIGURE 5 Cation removal versus electroosmotic flow (*s* represents symmetry units, an equivalent concentration) (*14*).

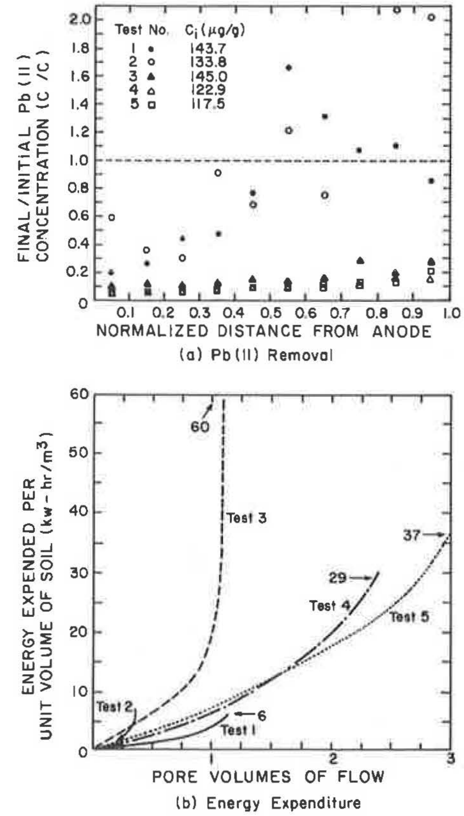


FIGURE 6 Pb(II) removal from kaolinite and corresponding energy expenditure (*8*).

generated at the cathode is not neutralized. The migrating ions are extracted with this probe. The quantity of the extracted metal at the cathode and the rate of accumulation are correlated with the composition of the ore and the distance of the sampling locations to the ores.

The potential of the technique in waste remediation resulted in initiation of several recent studies. Runnels and Larson (*17*) have investigated the potential use of electromigration to remove contaminants from groundwater. The amount of copper removed increased with processing time (total charge passed). However, the current efficiency decreased as the processing time increased, possibly because of the increase in conductivity as noted by Hamed et al. (*8*).

Renauld and Probstein (*3*) investigated the change in the electroosmotic water transport efficiency of kaolinite specimens loaded with acetic acid and sodium chloride. This study indicated that the current efficiency increased with higher concentrations of this weak, organic acid. This implies that when the deposit is contaminated with organic acids, the efficiency of electrokinetic soil processing may increase.

A better understanding of the chemistry in electrokinetic soil processing is achieved by studies at Louisiana State University. Putnam (*23*) investigated the development of acid/base distributions in electroosmosis. Acar et al. (*4,24*) present the theory for acid-base distributions in electroosmosis and compare the predictions of this model with the results of the tests conducted by Putnam (*23*). A good correlation was noted. This theory and the model presented by Acar et al. (*24*) describe the movement of different species in electrokinetics. The study demonstrates the movement of the acid front from the anode to the cathode by advection and diffusion and provides the fundamental basis of the chemistry developed during the process.

A comprehensive subsequent study on removal of Pb(II) from kaolinite is reported by Hamed et al. (*8*). Kaolinite specimens were loaded with Pb(II) at 118 to 145 μg/g of dry kaolinite, below the cation exchange capacity of this mineral. As shown in Figure 6, electroosmosis removed 75 to 95 percent of Pb(II) across the test specimens. The study clearly demonstrated that the removal was caused by migration and advection of the acid front generated at the anode by the primary electrolysis reaction. The energy used in the study to decontaminate the specimens was 29 to 60 kW-hr per cubic meter of soil processed. This study also demonstrates and explains the complicated electrochemistry associated with the process. An interesting finding of this study is electroplating of the removed Pb(II) at the carbon cathode.

Other laboratory studies conducted by Lageman (*19*) and Banerjee et al. (*20*) further substantiate the applicability of the technique to a wide range of contaminants and soils. High removal rates were achieved in soils contaminated with Cr, Cd, Ni, Pb, Hg, Cu, Zn, Hg, and As.

Although these laboratory studies display the feasibility of using electroosmosis to decontaminate soils, limited field studies are available. Table 2 presents a synthesis of field tests investigating or reporting some form of chemical removal from soils. Segall et al. (*25*) present the chemical characteristics of the water accumulated at the electrodes in electroosmotic dewatering of dredged soil. Their study indicated that

1. There was a significant increase in heavy metals and organic materials in electroosmosis effluent above that recorded in the original leachate. The concentrations of zinc, lead, mercury, and arsenic were especially high.
2. Total organic carbon content of the effluent was two orders of magnitude higher than the original leachate. It is postulated that highly alkaline conditions resulted in disso-

TABLE 2 SYNTHESIS OF FIELD DATA REPORTED FOR REMOVAL OF CHEMICALS BY ELECTROKINETICS

Soil Type/ Chemical	Concentration ($\mu g/g$)		Current Density and/or Voltage (mA/cm^2) or [V/cm]	Duration (hr)	Energy kWh/m^3	Remarks
	Initial	Final				
1. Puri and Anand (*13*)						An area of 4.5 m x 4.5 m was first trenched all around, 1.05 m in depth and 0.30 m in width. Anode and cathode were laid horizontally within this area. Anode was a shot of iron, 0.9 m x 1.8 m laid at top. Cathode was a perforated iron tube, 0.10 cm in diameter and 1.8 m long laid at 0.3 m depth. The concentrations reported are for the exchangeable Na in the top 7.5 cm of the soil.
High pH Soil - NaOH	4.8	2.8	1.35 [2.44]	6	N/A	
2. Case and Cutshall (*26*)						An area of 11 m by 5 m was investigated. 1.7 m stainless steel rods were driven in an arc-plus-center point array. The arc consisting of 25 anodes and a central cathode. The concentrations reported are for the effluent in a monitoring well.
Alluvial deposits - ^{90}Sr	8	50	0.3 [0.05]	5,352	N/A	
3. Segall, et al. (*25*)						Concentrations noted are for the effluent in electro-osmotic consolidation of dredged material compared to that of water leached specimens. The distance between electrodes is 3-5 m.
Dredged Material						
Cd, Zn, Pb, As, Fe	<1	0.2-20	[0.01-1.0]	N/A	N/A	
Na	6150	16300				
K	350	510				
OH^-	0	5950				
HCO_3^{-1}	0	979				
Organic Nitrogen	4	15				
Ammonia Nitrogen	67	128				
TOC	3	2000				
4. Lageman (*19*)						
Sandy Clay - Zn	70-5120	30-4470	0.8 [0.4-0.2]	1344	287	An area of 15 m by 6 m studied. Contamination depth: 0.40 m. Temperature rose from 12°C to 40°C. Conductivity increased from 2000 $\mu s/cm$ to 4000 $\mu s/cm$. Voltage gradient decreased. 2 cathodes (vertical) at 0.5 m depth. 33 anodes (vertical), 3 rows at 1.0 m depth. Distances: cathode-anode = 1.5 m; anode-anode = 1.5 m.
Heavy Clay - As	90-385	20-240	0.4 [0.4-0.2]	1200	270	An area of 10 m by 10 m was studied within a depth of 2 m. Cathodes (vertical), 2 rows: 1 row at 0.5 m depth; 1 row at 1.5 m depth. 36 anodes (vertical), 3 rows at 2 m depth; 2 rows of 14; 1 row of 8. Distances: cathode-anode = 3 m; anode-anode = 1.5 m. Temperature rose from 7°C to 50°C.
Dredged Sediment						An area of 70 m by 3 m was studied to a depth of 0.2 m to 0.5 m. Cathode is laid horizontal, anode (vertical). Cathode-anode 3 m; anode-anode 2 m.
Pb	340-500	90-300	N/A	430	N/A	
Cu	35-1150	15-580				
5. Banerjee, et al. (*20*)						Nine field experiments were conducted in an array of electrodes. Combined Hydraulic and electrical potentials were applied. Steel reinforcing bars were used and replaced after each experiment. The results are inconclusive.
Silt/Silty Clay - Cr	N/A	N/A	2-4 [0.2-0.24]	<72	N/A	

lution and release of the organic material. Pesticides came out at the cathode.

Case and Cutshall (*26*) described a field study for control of radionuclide migration in soil by application of DC current. This study demonstrates that it is possible to migrate radionuclides with the technique. Lageman (*19*) reported the results of field studies conducted in the Netherlands to decontaminate soils by electrokinetic soil processing. Figure 7 shows a schematic diagram of the reported field process. An electrode fluid conditioning and purification system is noted. The conditioning (such as buffering of the effluent to avoid precipitation of removed ions) is for the control of the influent-effluent chemistry, whereas purification (such as ion exchange

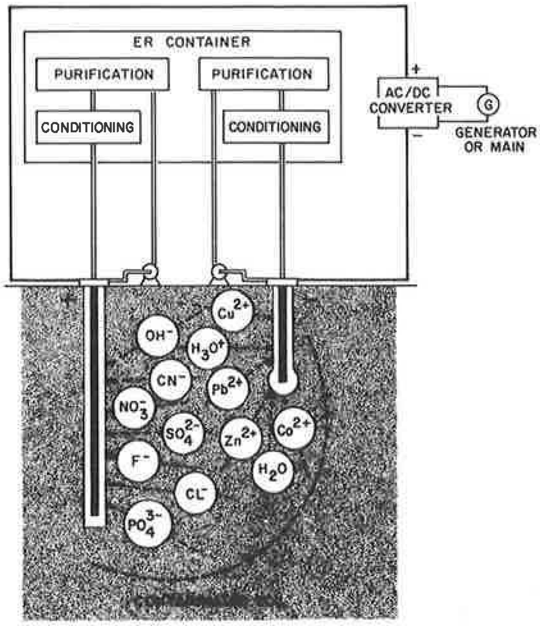


FIGURE 7 Schematic diagram of the field system reported by Lagemann (*19*).

resin columns) is for removing any excess ions in the effluent. A recent study investigating the application of the process to decontaminate a chromium site was reported by Banerjee et al. (*20*). The laboratory studies reported by Hamnet (*16*), Runnels and Larson (*17*), Lageman (*19*), and Hamed et al. (*8*) together with the pilot-scale field studies of Lageman (*19*) display the feasibility of using the process in site remediation. Further pilot-scale studies are necessary to improve the technology and establish the field remediation scheme for different site conditions and chemistry.

Engineering Implications

This review of the present state of knowledge on electrokinetic soil remediation resulted in the following implications:

1. Type of soil. The process results in movement of ions in sandy to clayey soils. It seems that there will not be any major restriction on the type of soil. High-water-content, low-activity soils at low-pore-fluid electrolyte concentrations will result in the most efficient conditions.
2. Type of contaminants. All available data are on ionic forms of inorganic cations, some radionuclides (e.g., ^{90}Sr), and acetic acid. Recent laboratory studies at Louisiana State University demonstrate that it is feasible to remove adsorbed phenol (1,000 ppm) from kaolinite by this technique. Furthermore, release of pesticides and increase in total organic carbon content in the effluent in electroosmotic consolidation indicate that the method has potential in removing organic contaminants. Data regarding acid-base distributions indicate that salts (such as PbO) may also dissolve and migrate because of the advancing acid front. However, there is no factual data to validate this hypothesis. Laboratory data demonstrate that while cations move towards the cathode, anions move towards the anode (*8,12*).
3. Concentration of contaminants. There exists factual data demonstrating removal at levels of up to 10,000 ppm of Cu(II) and 5,000 ppm of Pb(II). As concentrations of contaminants (ionic) increase, removal should be mostly by migration as advection (electroosmotic flow) will substantially decrease. At lower concentrations, both advection and migration will be acting to remove the contaminants.
4. Mixture of contaminants. The data indicate that the process also works on a mixture of contaminants (*19*). The removal amount is related to both the concentration and the mobility of the specific ion under electrical gradients. Monovalent ions may be removed at a higher rate than higher valence ions.
5. Saturation. Mitchell and Yeung (*12*) present data regarding the effect of saturation on k_e, which did not change significantly in specimens compacted at different molding moisture contents. These data suggest that the process may be applicable in partially saturated soils.
6. Depths. The review of literature indicates that there should not be a depth limitation in the process beyond practical problems that may be encountered.
7. Type of electrodes. Inert electrodes such as graphite, carbon, or platinum should be used to avoid introducing secondary corrosion products into the soil mass. Open electrodes allow control of influent and effluent chemistry. It should be recognized that some ions will be electroplated on the cathode. Cost-effective electrodes need to be devised and manufactured.
8. Electrode configuration. The electrodes can be placed horizontally or vertically. The electrical potential gradients generated because of different electrode arrangements affect the flow conditions and hence the removal efficiency. The gradients significantly change by electrode configurations and the depth of individual electrodes relative to the counter electrode. A hexagonal network of electrodes with a central cathode and all electrodes at the same depth is hypothesized to be an efficient configuration. However, it is necessary to design the most efficient electrode configuration by considering the coupling of electrical, hydraulic, and chemical gradients.
9. Electrode spacing. Spacing will depend on the type and level of contamination and the selected current-voltage regime. A substantial decrease in efficiency of the process may result because of increases in temperature when higher voltage gradients are generated. A spacing that generates a potential gradient in the order of 1 V/cm is preferred. The electrode spacing generally is up to 3 m.
10. Current level. The current level reported is in the order of milliamps per square centimeter of electrode area (0.01 to 1.0 mA/cm^2). It can be varied to monitor the influent pH level at the anode and to control the rate of decontamination.
11. Duration. Process should be continued until the desired removal is achieved. The remediation duration is site specific. Acid front generated at the anode will advance to the cathode. In general, in fine-grained soils it is expected for the treatment to continue a number of months. When it is desired to reduce the duration by increasing the electrical potential gradient, the efficiency of the process decreases.
12. Effluent-influent chemistry. It is possible to control the efficiency by controlling the pH and the chemistry of the effluent and the influent. Two alternatives are available: (a) decreasing the current to a level where fewer H^+ ions are

generated, and (b) flushing the anode and cathode by a fluid of known chemistry.

13. Chemistry subsequent to the process. The porous medium becomes acidic on completion of the process. The medium returns to original conditions by diffusion of the acidic pore fluid to the surrounding medium. Cathode effluents may require postchemical treatments (such as ion exchange resin columns for inorganic contaminants) to achieve concentration of contaminants. Cathodes may necessitate treatment with acid to remove the electroplated contaminants, if found necessary.

ACKNOWLEDGMENTS

Studies at Louisiana State University (LSU) investigating electrokinetic soil processing are currently funded by the Board of Regents of the State, the Hazardous Waste Research Center of LSU, and the National Science Foundation. The funds awarded by these agencies are gratefully acknowledged.

REFERENCES

1. Y. B. Acar, R. J. Gale, and H. Olsen. *Decontamination of Soils Using Electro-Osmosis.* Proposal submitted to the Board of Regents of the State of Louisiana, LEQSF Research Development Program Office of Research Coordination, Louisiana State University, 1986.
2. J. K. Mitchell. Potential Uses of Electro-Kinetics for Hazardous Waste Site Remediation. Presented at USEPA-University of Washington Workshop on Electro-Kinetic Treatment and Its Application in Environmental-Geotechnical Engineering for Hazardous Waste Site Remediation, Seattle, Wash., Aug. 1986.
3. P. O. Renauld and R. F. Probstein. Electro-Osmotic Control of Hazardous Waste. *Physicochemical Hydrodynamics*, Vol. 9, No. 1/2, 1987, pp. 345–360.
4. Y. B. Acar, R. J. Gale, G. Putnam, and J. Hamed. Electrochemical Processing of Soils: Its Potential Use in Environmental Geotechnology and Significance of pH Gradients. *Proc., 2nd International Symposium on Environmental Geotechnology*, Shanghai, China, Vol. 1, Envo Publishing, Bethlehem, Pa., 1989, pp. 25–38.
5. J. Steude, S. Viani, and K. Baker. Emerging Technologies for the Remediation of Radioactive Soils. Roy F. Weston, Inc., Walnut Creek, Calif., 1989, 64 pp.
6. L. Casagrande. *The Application of Electro-Osmosis to Practical Problems in Foundations and Earthwork.* Technical Paper 30, Department of Scientific and Industrial Research, Building Research, London, England, 22 pp.
7. Y. B. Acar, R. J. Gale, J. Hamed, and G. Putnam. Electrochemical Processing of Soils: Theory of pH Gradient Development by Diffusion and Linear Convection. *Journal of Environmental Science and Health, Part (a): Environmental Science and Engineering*, Vol. 25, No. 6, 1990, pp. 687–714.
8. J. Hamed, Y. B. Acar, and R. J. Gale. Pb(II) Removal from Kaolinite by Electro-Kinetics. *Journal of Geotechnical Engineering, ASCE*, Vol. 117, No. 2, Feb. 1991.
9. J. K. Mitchell. *Fundamentals of Soil Behavior.* John Wiley, New York, 1976, 422 pp.
10. D. H. Gray and J. K. Mitchell. Fundamental Aspects of Electroosmosis in Soils. *Journal of the Soil Mechanics and Foundation Division, ASCE*, Vol. 93, No. SM6, 1987, pp. 209–236.
11. N. C. Lockhart. Electro-Osmotic Dewatering of Clays, I, II and III. *Colloids and Surfaces*, Vol. 6, 1983, pp. 238–269.
12. J. K. Mitchell and T.-C. Yeung. Electro-Kinetic Flow Barriers in Compacted Clay. In *Transportation Research Record 1288*, TRB, National Research Council, Washington, D.C., 1990.
13. A. N. Puri and B. Anand. *Soils—Their Physics and Chemistry*, Rheinhold Corp., New York, 1936, pp. 134–140.
14. H. S. Jacobs and M. M. Mortland. Ion Movement in Wyoming Bentonite During Electro-Osmosis. *Proc., Soil Science Society*, 1959.
15. R. J. Krizek, F. B. Gularte, and P. B. Hummel. Stabilization of Polluted Dredgings by Electro-Osmosis. Presented at ASCE National Water Resources and Ocean Engineering Convention, San Diego, Calif., April 5–8, 1976.
16. R. Hamnet. *A Study of the Processes Involved in the Electro-Reclamation of Contaminated Soils.* M.S. thesis, University of Manchester, England, 1980, 84 pp.
17. D. D. Runnels and J. L. Larson. A Laboratory Study of Electromigration as a Possible Field Technique for the Removal of Contaminants from Ground Water. *Ground Water Monitoring Review*, pp. 81–91, Summer 1986.
18. R. T. Thompson. *The Effect of Secondary Reactions on the Electrokinetic Treatment of a Silty-Sand Soil.* M.S. thesis, Civil Engineering Department, University of Texas, Austin, 1989, 115 pp.
19. R. Lageman. Theory and Practice of Electro-Reclamation. *NATO/CCMS Pilot Study, Demonstration of Remedial Action Technologies for Contaminated Land and Ground Water*, Copenhagen, Denmark, May 9, 1989, 18 pp.
20. S. Banerjee, J. Horng, J. F. Ferguson, and P. O. Nelson. *Field Scale Feasibility of Electro-Kinetic Remediation.* CR811762-01. Land Pollution Control Division, U.S. Environmental Protection Agency, 122 pp.
21. A. N. Puri, Reclamation of Alkali Soils by Electrodialysis. *Soil Science*, Vol. 42, 1936, pp. 23–27.
22. B. M. Shmakin. The Method of Partial Extraction of Metals in a Constant Current Electrical Field for Geochemical Exploration, *Journal of Geochemical Exploration*, Vol. 23, No. 1, 1985, pp. 35–60.
23. G. Putnam. Development of pH Gradients in Electrochemical Processing of Kaolinite. M.S. thesis, Department of Civil Engineering, Louisiana State University, 1988.
24. Y. B. Acard, J. Hamed, R. J. Gale, and G. Putnam. Acid/Base Distributions in Electro-Osmosis. In *Transportation Research Record 1288*, TRB, National Research Council, Washington, D.C., 1991, pp. 23–34.
25. B. A. Segall, C. E. O'Bannon, and J. A. Matthias. "Electro-Osmosis Chemistry and Water Quality. *Journal of the Geotechnical Engineering Division, ASCE*, Vol. 106, No. GT10, Oct. 1980, pp. 1143–1147.
26. F. N. Case and N. H. Cutshall. Oak Ridge National Laboratory, TN (USA). Conf. 791112-28, Symposium on the Scientific Basis for Nuclear Waste Management, Boston, Mass., Nov. 1979, 5 pp.

Publication of this paper sponsored by Committee on Physicochemical Phenomena in Soils.

Site Remediation by In Situ Vitrification

MICHAEL G. BYERS, VINCENT F. FITZPATRICK, and ROBERT D. HOLTZ

In situ vitrification (ISV) is a treatment technology that uses electricity to heat contaminated soil sufficiently to produce an inert glass product. Organic contaminants are destroyed by pyrolysis or they are stripped out of the soil with the escaping steam and trapped in the off-gas treatment system. Most inorganic contaminants are incorporated into the vitrified zone. The glass product is similar visually and chemically to naturally occurring obsidian, and like obsidian, the ISV product is also durable. The electrical power is delivered to the soil through special graphite-molybdenum electrodes inserted into the contaminated soil region in a square pattern. The ability of ISV to treat hazardous chemical wastes, radioactive wastes, and mixtures of hazardous and radioactive wastes is a significant advantage of this process over many other treatment technologies.

In situ vitrification (ISV) is a thermal treatment technology that uses electricity to heat contaminated soils in situ to produce an inert glass product. ISV has a number of advantages over current technologies, particularly the ability to treat mixed wastes, including soils with buried trash or drums, dewatered sludge, mine tailings, or sediments. Furthermore, after treatment the by-product will remain inert for many years to come.

ISV allows contaminated materials to be treated in their existing location without any excavation or transportation required. This can result in a large cost savings as compared with almost any other treatment alternative that requires excavation or transportation of hazardous materials. Alternatively, if the contamination is present over a large area and in a thin layer, the contaminated materials may be consolidated on site in trenches for treatment. In this manner, ISV can be used as a treatment process at a site or area that would otherwise require extensive excavation.

In the ISV process, the soil is heated by applying an electrical potential across four electrodes inserted in a square pattern in the contaminated soil and connected to an electrical circuit. Because most soils are initially nonconductive, a starter path of flaked graphite and glass frit (pulverized glass) is placed between the electrodes to initiate the melting process. The entire area is covered by a containment hood to collect all off-gas produced during the vitrification process. Figure 1 shows a typical layout of the starter path and electrode placement in relation to the contaminated soil.

Testing has been performed on four different size systems (*1*). The distinguishing characteristics of each system are power level, electrode spacing, and mass of vitrified block produced, as presented in Table 1. Presently, engineering scale testing is being performed at the University of Washington using a variety of different soil types and organic chemicals.

M. G. Byers, Converse Consultants, N.W., Seattle, Wash. 98119. R. D. Holtz, Department of Civil Engineering, University of Washington, Seattle, Wash. 98195. V. F. FitzPatrick, Geosafe Corporation, 303 Parkplace, Suite 126, Kirkland, Wash. 98033.

DESTRUCTION METHODS

The ISV process destroys organic materials by pyrolysis in a strong reducing environment. As contaminants pyrolyze, some components migrate or bubble to the surface of the melt. At the surface, they combust in the presence of oxygen, and remnants of this combustion are drawn into the off-gas treatment system. Most inorganic materials are incorporated into the melt and spread throughout the molten soil by convective currents caused by heating.

In addition to pyrolysis, several mechanisms including molecular diffusion, carrier gas transport, capillary flow, and flow resistance in the soil column contribute to the destruction of high percentages, commonly 99.995 percent, of organic materials contained in the waste (*2*).

PRODUCT DURABILITY

The ISV mechanism of melt production and solidification is similar to the natural reactions that create igneous rocks on earth. A typical soil is primarily composed of silica and aluminum oxides that have melting temperatures between 1,100°C to 1,600°C (2,012°F to 2,912°F) (*3*). When this melt cools, it forms a glass with a random atomic structure and a relatively uniform chemical composition. This uniformity is caused by convection currents caused by the heating of the soil, and the characteristics of the glass are basically the same as obsidian (*4*). Because the ISV-derived glass has a random atomic structure, its fracture mechanism is conchoidal with no preferential planes of weakness. Tests on ISV-derived glass indicate that its durability is similar to that of granite (*3*).

The durability of ISV-derived glass can be estimated by examining the durability of obsidian. The weathering process of obsidian involves the hydration with atmospheric water that is chemically absorbed on the surface. The water then diffuses into the obsidian as a function of time and temperature. In the natural environment, obsidian has a hydration rate constant of about 1 to 20 μm^2 per 1,000 years (Larsen and Langford, 1978). Using the conservative end gives an estimate of less than 1-mm hydrated depth for the ISV-derived glass over a 10,000-year time span (*5*).

ISV-derived glass has been subjected to a variety of leach tests, including the Environmental Protection Agency's (EPA's) extraction procedure toxicity test (EPTOX) and toxic characteristics leach test (TCLP). All of these tests exhibit a uniformly low leach rate for heavy metals of about 5×10^{-4} kg/m²/day or lower (*1*).

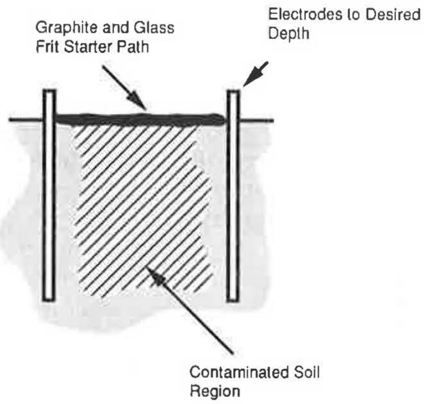


FIGURE 1 Typical ISV layout.

TABLE 1 CHARACTERISTICS OF ISV SYSTEMS

System	Power (kW)	Electrode Spacing (m)	Block Mass (tons)
Bench scale	10	0.11	0.001 to 0.01
Engineering scale	30	0.30	0.05 to 1.0
Pilot scale	500	1.2	10 to 40
Field scale	3,750	5	400 to 800

PROCESS DESCRIPTION

Power System

Three-phase power supplied by the utility grid system or portable generators is delivered to a special adjustable (multitap) transformer that converts it to two single phases and transforms the power supplied to the voltage levels needed throughout the processing. Each electrical phase is connected to a diagonally opposed pair of electrodes in a square pattern; thus a balanced electrical load is created on the secondary side of the transformer. The even distribution of current within the molten soil produces a vitrified product that is almost cubical, a shape that minimizes overlap among adjacent treated volumes. Multiple voltage taps and a balanced load allow for a near-constant power operation; thus run time is shortened and costs minimized.

Adjustable voltage taps are used to maintain a constant input of power to the melt; as more of the soil becomes molten, the conductivity of the soil changes and hence the voltage and current supplied also should be changed to maintain a constant power input. Fourteen adjustment levels permit voltages from a maximum of 4,000 volts to a minimum of 400 volts per phase and currents from a minimum of 400 to a maximum of 4,000 amps per phase. As the melt volume grows downward and outward, power is maintained at a sufficient level to overcome the heat losses from the surface and into the surrounding soil.

In the ISV process, electrical current is delivered to the soil by electrodes. In the large-scale field operation, the electrodes consist of a 50-mm-diameter molybdenum inner electrode core inside a 300-mm-diameter graphite collar (*1*). The graphite collar provides a relatively inexpensive cross section that promotes conduction of current to the melt. It also promotes gas release from the melt by providing a preferential corridor in the vertical direction. The graphite collar is sacrificial in that it is left in the melt when the process is complete. The larger graphite collar also provides a protection layer for the molybdenum inner electrode. This generally means that the more costly inner electrode may be reused, which provides a cost savings.

The electrodes are placed in the soil to the desired treatment depth. The depth of the melt is limited as heat losses from the melt approach the energy level deliverable to the molten soil by the electrodes (*1*). Generally, the melt grows outward to about 50 percent of the spacing of the electrodes. With a typical electrode spacing of 5.5 m, a melt width of 8.5 m would be observed under normal conditions. Full-scale ISV units are designed to operate to depths of 9.1 to 12.2 m. Figure 2 shows how successive melts are accomplished to form a single monolith.

The field-scale ISV system melts soil at a rate of 4 to 6 tons/hr. Accordingly, the rate of melt advance is in the range of 25 to 50 mm/hr. As the thermal gradient advances in the soil, solid and liquid organic materials first vaporize and then pyrolyze (i.e., decompose in the absence of oxygen) into their elemental components. Organic pyrolysis byproducts are typically gaseous; these gases move slowly through the melt toward the upper melt surface. Some of these gases may dissolve

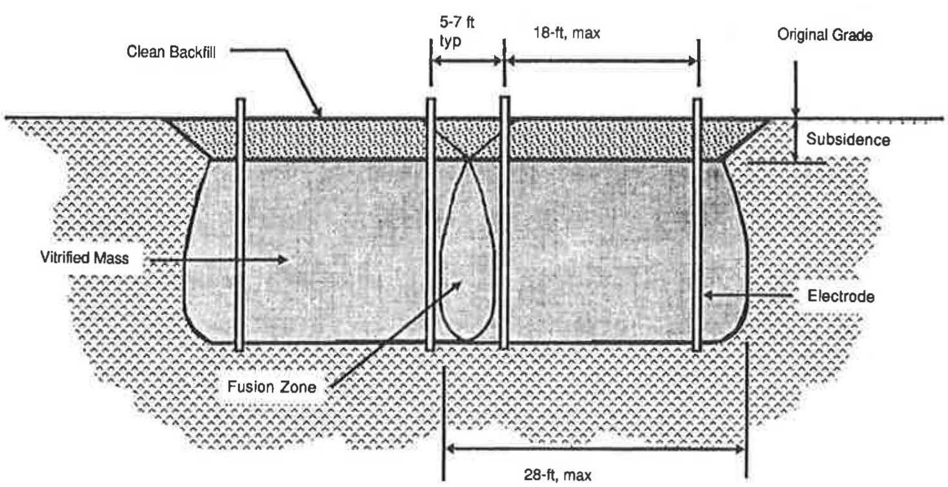


FIGURE 2 Pattern of successive melts.

in the molten mass, but the remainder escape and are collected in the off-gas collection hood for subsequent treatment in the off-gas treatment system.

Individual ISV process settings (with electrode spacings of 5.5 m) may have a mass of 1,000 tons. Adjacent settings fuse together into a single monolith (Figure 2); thus, the residual product presents a minimum of surface area to the surrounding environment. Even when reduced to fine particle sizes, the ISV residual product easily passes EPA's EPTOX and TCLP testing criteria.

Because the soil or other contaminated material is generally nonconductive, a conductive starter path is provided near the surface of the melt zone to provide an initial pathway for the electric current. This starter path consists of a mixture of flaked graphite and glass frit. As the electrical current is applied to the starter path, the glass frit heats up to 2,000°C, which is well above the initial soil mineral melting temperatures of 1,100°C to 1,400°C. Eventually, the starter path is consumed and the energy is transferred into the soil mass. In this way, a molten conductive zone is started in a soil that is initially nonconductive.

The ISV process seeks an equilibrium temperature depending on the fusion temperature of the soil and the particle sizes encountered. In general, the melt can attain temperatures greater than 1,700°C. When a higher fusion temperature layer or a large particle size (greater than 20 mm) layer is encountered, a higher equilibrium temperature is needed to achieve the same downward progression rate for the melt. If necessary, fluxants (e.g., soda ash, lime, and glass frit) injected into the voids of larger grain size materials may be used to enhance the downward melting rate (*6*).

A layer of ceramic insulation tolerant of high temperature is placed over the surface of the treatment area and surrounding soils. This layer minimizes heat loss from the surface of the melt zone and enables a larger and deeper melt to be accomplished with the same available power input.

A percentage of a soil total volume is made up of voids containing either gas or liquid. During the ISV process, the voids present in the treated soil disappear leaving a glass with nearly zero voids; the only exception is some trapped gas bubbles near the surface. As typical porosites in soil vary from 20 to 40 percent, vitrification produces a reduction in volume that is reflected in a surface subsidence. The surface crater is generally backfilled with clean fill to the original grade. This measure also provides some additional protection for the glass monolith as well as a barrier to possible surface disturbance.

Costs

The cost of ISV operations is highly site specific. The major factors in cost include the following:

- The amount of site preparation and staging work required, if any;
- Specific properties of the waste or soil to be processed;
- Volume of material to be processed;
- Depth of processing;
- Water content;
- Unit cost of electricity; and
- Season of the year and weather

Depth of processing is a significant cost variable that relates directly to the ratio of operating time to down time. Because the ISV process is site specific, the equipment must be moved between treatment sites. It is more economical to treat to a deeper depth and thus allow more treating time relative to down or move time. For shallow contamination (e.g., only a few meters), it is usually more cost-effective to remove and stage the material into a deeper configuration for processing.

Water content is one of the most significant variables affecting vitrification costs, which can vary from $50/ton to $75/ton between dry soil and fully saturated soil.

The cost of electricity is also a major variable affecting total ISV costs. Generally, it is more economical to use utility-provided power if it can be obtained for less than $0.07/kW-hr. Above this range, consideration should be given to using a local diesel generator.

The onsite cost of ISV processing usually is in the range of $250 to $350 per ton of material processed (*7*).

Off-Gas System

A large collection hood covers the treatment area to trap any gaseous effluents from the process and to direct them to the portable off-gas treatment system. This hood controls the flow of air, provides a chamber for the combustion of pyrolyzed organics, supports the electrodes, and maintains a negative hood pressure of 125 to 250 kPa. Maintaining a negative hood pressure is accomplished by moving the air through the system using a blower attached to the end of the treatment line and sealing the collection hood as well as possible to the ground allowing only one controlled air inlet.

The effluents exhausted from the hood are cooled and treated in the off-gas treatment system. Much of the heat generated during the process is exhausted through the off-gas system. If there is a high concentration of buried solid or liquid organics present in the contaminated soil, the off-gases can be expected to reach a temperature of 750°C. In general, the off-gas system cools, scrubs, sorbs, and filters the organic chemicals exhausted from the hood in three stages (*1*). First, the gases are scrubbed in two stages, with a quencher and tandem nozzle scrubber. These scrubbers remove particles down through the submicron range. Second, most of the water in the saturated gas stream is removed by a vane separator followed by a condenser and a second vane separator. The gases are then reheated, to ensure that there is an unsaturated gas stream at a temperature well above the dewpoint. In the third treatment phase, the off-gas is filtered with two stages of high-efficiency particulate air filters. Figure 3 is a schematic drawing of the large-scale off-gas treatment system.

FIELD EQUIPMENT REQUIRED

The entire ISV operation, except for the off-gas collection hood and line, is mounted on three standard over-the-road trailers. The off-gas and process control trailer, a support trailer, and an electrical trailer can all be easily hooked together, or unhooked when the system needs to be moved. The off-gas hood and line, which are installed on the site for collection of the gaseous effluents, are dismantled and placed on a flatbed trailer for transport between the sites to be treated. The only other requirement is some method to insert the electrodes into the ground. This can be accomplished by a

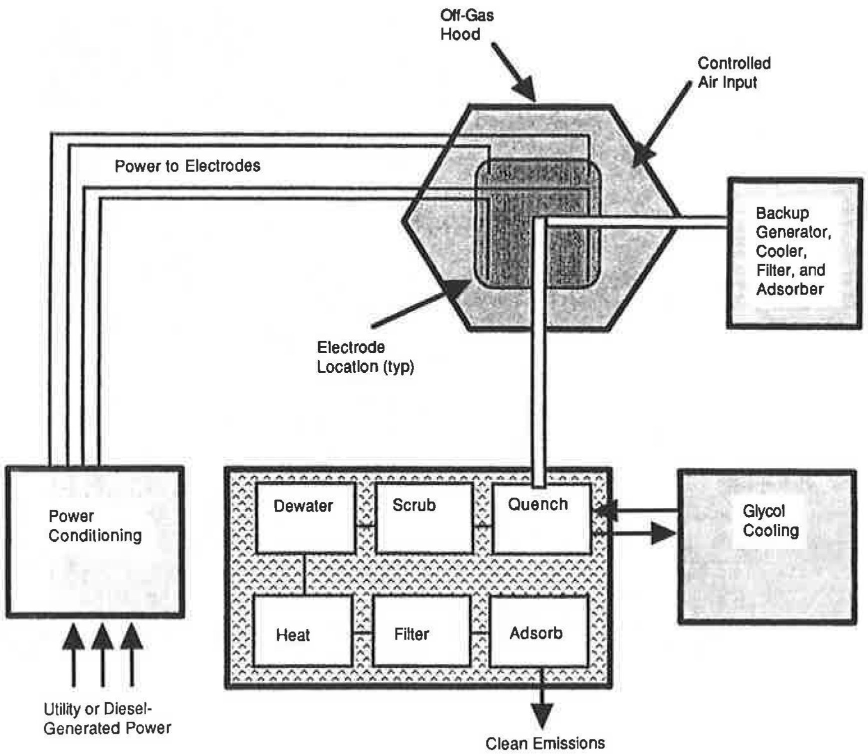


FIGURE 3. Schematic diagram of large-scale off-gas treatment system.

standard drill rig or other similar method. Typical readiness time is about 1 week after arrival (*7*).

The large-scale ISV equipment requires three-phase electrical power usually taken from a utility distribution system at typical transmission voltage of 12,500 or 13,800 volts. The equipment draws about 3,750 kW during operation, which corresponds to a specific power consumption averaging for most soils, of 1 kW-hr/kg of soil treated. If no utility distribution system is available at a particular site, power can be generated on site by diesel generators. Figure 4 shows a field-scale operation.

LIMITATIONS

The ISV process was developed and demonstrated through large-scale projects, and wastes treated have included a variety of hazardous chemical, radioactive, and mixed (hazardous chemical and radioactive) wastes. The ability of ISV to process mixtures of these waste types simultaneously is a significant advantage compared with other remediation technologies. With mixtures of waste types, it is usually necessary to use more than one non-ISV technology, in sequential fashion, to accomplish a total remediation.

Although ISV is broadly applicable, there are limits to its uses. Most significant is the rate of groundwater recharge when the treatment zone is below the water table. As ISV heats the soil, the soil reaches the boiling point of the water and cannot go any higher until all of the water is boiled off. This can cause a significant amount of electricity to be expended boiling off water and not increasing the size of the molten zone. Applications involving a hydraulic conductivity exceeding 10^{-4} cm/sec are considered marginal from an energy consumption standpoint; that is, the use of barrier walls, well points, or some other means to limit recharge would be cost effective compared to removing water by electrical heating (*7*).

Some soils with an organic chemical concentration of more than 5 to 10 percent by weight will cause an overload in the off-gas system. This system is designed to transfer only a certain amount of heat and organics out of the off-gas. Soils with organic concentrations of these levels will overload the system.

Metal concentrations (e.g., drums and scrap metal) in the range of 5 to 16 percent by weight can be processed without special adaptations. Continuous metal contact (e.g., pipeline or rebar) up to 90 percent of the distance between opposed electrodes also is acceptable. With these capabilities, the process may be used to treat contaminated drums and other metallic components.

ISV can also accommodate a significant amount of rubble, debris, or other inclusions within the treatment zone. Figure 5 shows typical inclusions with their maximum acceptable amounts recommended for ISV.

SUMMARY

ISV is a treatment process that uses electricity to treat a variety of contaminated soils resulting in an inert glass product. This glass product is similar to naturally occurring obsidian in composition and therefore should weather similarly to natural obsidian. The glass product has been found to be stable when tested using the EPA leachability test.

ISV has advantages over most other treatment methods available. They include the ability to treat the waste in place with a minimum of excavation and the ability to treat a variety of wastes including mixed wastes, which is often the case

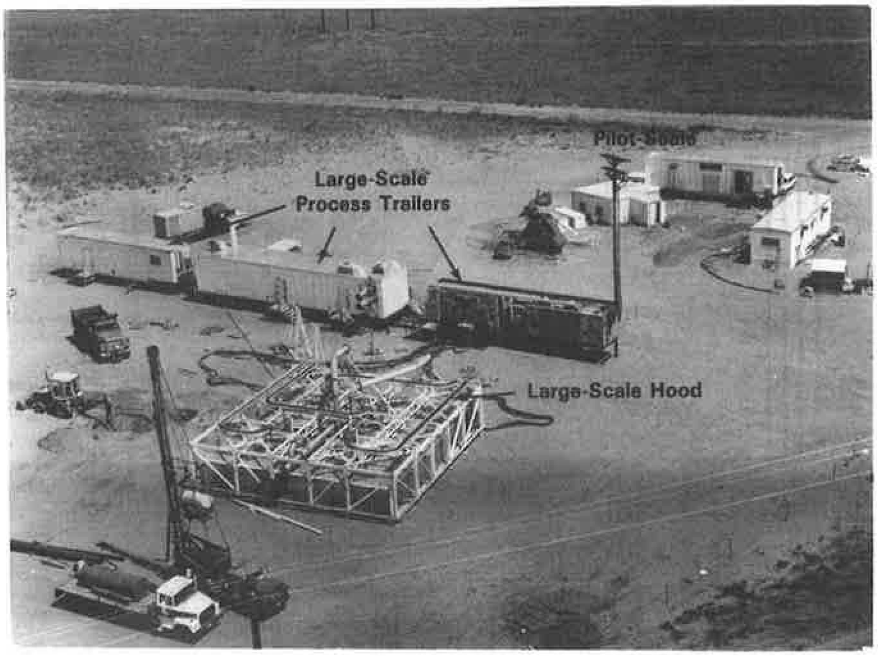


FIGURE 4 Field-scale ISV operations.

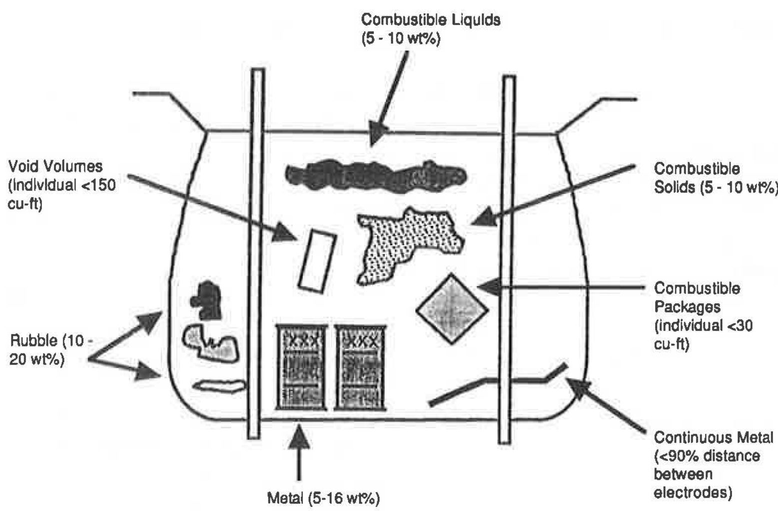


FIGURE 5 General limitations for inclusions within volume to be treated.

found on hazardous waste sites. These advantages make ISV a viable alternative for treatment of contaminated soils.

REFERENCES

1. J. L. Buelt, C. L. Timmerman, K. H. Oma, V. F. FitzPatrick, and J. G. Carter. *In Situ Vitrification of Transuranic Waste: An Updated Systems Evaluation and Applications Assessment*. Report PNL-4800, Suppl. 1. Pacific Northwest Laboratory, Richland, Wash., 1987.
2. D. M. Timmons, V. F. FitzPatrick, and S. C. Liikala. The Behavior of Mercury, Arsenic, and Organo-Chlorine Pesticides when Treated by In Situ Vitrification. *Pollution Engineering*, Vol. 22, No. 6, 1990, pp. 76–81.
3. Electric Power Research Institute. *Remedial Technologies for Leaking Underground Storage Tanks*. Lewis Publishers. Chelsea, Mich., 1988.
4. R. C. Ewing and R. F. Hoaker. *Naturally Occurring Glasses: Analogues for Radioactive Waste Forms*. Report PNL-2776, Pacific Northwest Laboratory, Richland, Wash., 1979.
5. V. F. FitzPatrick, J. L. Buelt, K. H. Oma, and C. L. Timmerman. *In Situ Vitrification: A Potential Remedial Action for Hazardous Waste*. Report PNL-SA-12316. Pacific Northwest Laboratory, Richland, Wash., 1984.
6. J. L. Buelt and W. F. Bonner. In Situ Vitrification: Test Results for a Contaminated Soil Melting Process. Presented at the 1989 Incineration Conference, Knoxville, Tenn., 1989.
7. J. E. Hansen and V. F. Fitzpatrick. In Situ Vitrification: Heat and Immobilization are Cobmined for Soil Remediation. *Hazmat World*, Vol. 2, No. 12, 1989, pp. 30–34.

Publication of this paper sponsored by Committee on Physicochemical Phenomena in Soils.

Field Experience with Electrokinetics at a Superfund Site

SUNIRMAL BANERJEE, J. J. HORNG, AND JOHN F. FERGUSON

The results of a field study concerning the technical feasibility of electrokinetic remediation technique at a Superfund site are discussed. The technique of remediation is based on the migration of ionic contaminants under the action of a dc electric field, but for field application of the remedial action, effective combination of dc field application and the pump-and-treat method is required. A matrix of properly designed extraction wells (from which contaminated groundwater can be withdrawn) and an array of wells surrounding each of these wells in a square or hexagonal pattern are needed. The application of a dc field across the contaminated ground, using electrodes placed in both the central as well as the surrounding wells, can enhance the removal of contaminants from the subsurface system. This technique was explored on the premises of United Chrome Products, Inc., in Corvallis, Oregon, where the shallow subsoils and the unconfined groundwater regime were heavily contaminated with chrome plating wastes. On the basis of the analyses of the acquired data on the concentration of various ionic species, it was concluded that electrokinetic treatment combined with occasional withdrawal of effluents can be more effective than the conventional pump-and-treat method.

The currently available waste management alternatives can be grouped into several broad categories: (a) source control, (b) effluent control, (c) physical removal, (d) encapsulation, (e) stabilization, and (f) in-place decontamination. Needless to say, the first two alternatives, which require prior design of a controlled facility, do not pertain to the superfund sites. At these sites, existing subsurface contamination resulting from uncontrolled point-source discharges and accidental spills, more often than not involves large volumes of ground and groundwater. As such, remediation by physical removal of huge masses of contaminated soils for transport to and disposal at an off-site facility becomes overly expensive, and may sometimes cause greater damage to the environment than if the affected soils were left in place. Encapsulating the contaminated mass with impermeable barriers is useful primarily as a temporary measure for arresting the spread of the contaminants. Stabilization of the ground by application of heat or chemicals, on the other hand, may not be desirable if the land must be restored and rehabilitated.

Regulators have recognized the technical limitations, time requirements, and costs of the above alternatives, and a search is underway (*1*) for in situ techniques that use physical, chemical, thermal, and biological processes to treat the contaminants without removing them from the subsurface system. Currently, a large number of innovative in situ treatment schemes are in the developmental stage. Because there is no unique approach to remediation for the wide variety of contaminants and possible site characteristics, combinations of remedial measures are also being researched and tested.

Department of Civil Engineering, University of Washington, Seattle, Wash. 98195.

One of the most widely used methods for treatment of groundwater contaminated with heavy metals is the removal-recovery or pump-and-treat method. Although pump-and-treat is a relatively low-cost process, the long time frame required to achieve the desired effects may sometimes become unacceptably slow, especially in low-permeability soils and areas where the contaminants tend to sorb. In order to investigate the potential of the electrokinetic technique, a field study (*2*) sponsored by the U.S. Environmental Protection Agency (EPA) was undertaken at a superfund site at Corvallis, Oregon. This technique, based on electrically induced transport of chemicals, was studied both by itself and in combination with the pump-and-treat method. The performance data from bench-scale and field experiments were collected and analyzed.

BACKGROUND

The Superfund site in Corvallis, Oregon, is an abandoned industrial hard-chrome plating facility, previously owned by United Chrome Products, Inc. (UCPI), located next to the Corvallis Airport, south of Airport Road. UCPI operated an industrial hard-chrome plating facility at the site from 1956 to 1985. The facility (Figure 1) consisted of a single building that housed two large plating tanks at the northwest corner, an office in the central northeast end, and a shop that covered the rest of the building. Adjacent to the west side of the building were a few storage tanks, an acid bath, a base bath, and a dry well or gravel pit. As a general practice, the operators discharged the wastes produced at the facility into the dry well constructed specifically for this purpose by digging a hole and backfilling the pit with sand and gravel.

The wastes discharged into the gravel pit generally included floor washing and rinse water collected in a sump within the building, other wastes like spent plating tank solutions, and sludges from the plating tanks. An estimated volume rate of 1,000 gal/year of such waste was discharged into the pit until 1982 when the practice was reportedly stopped.

In 1975, when UCPI submitted an application for an NPDES discharge permit to the Oregon Department of Environmental Quality, the problems at the site came to their attention. Investigations revealed elevated levels of chromium and other heavy metals in the sediment and the water of the nearby drainage ditches. In 1982, the site was determined to be a potential source of contamination of the surface drainage courses as well as the groundwater aquifers unless appropriate steps were taken.

Soon after, investigations of the soil and water conditions at the UCPI site were conducted by Ecology and Environ-

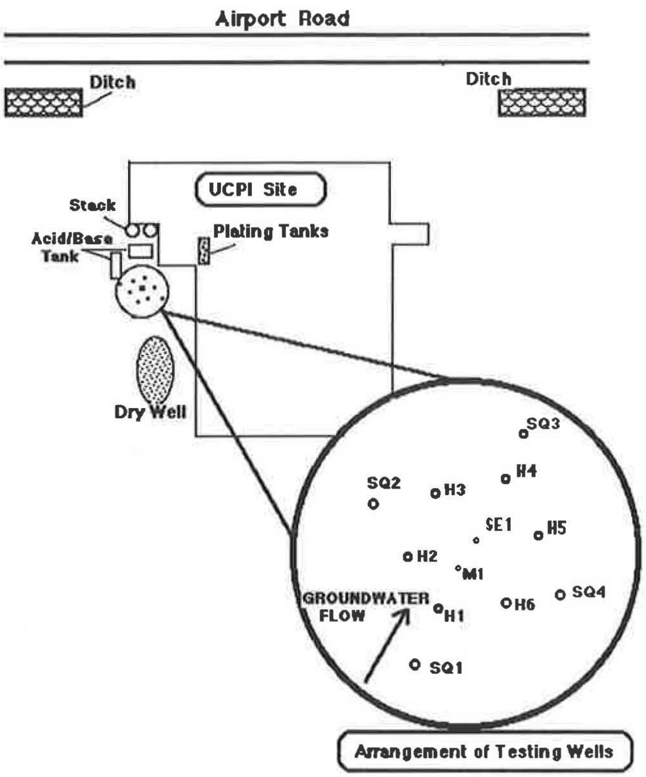


FIGURE 1 Site and well location plan.

ment, Inc. (*3*), on behalf of the Hazardous Site Control Division, EPA, Region X. On the basis of this study and the serious extent of contamination of the subsurface soils and groundwater at the site revealed by these investigations, the UCPI site was classified as a Superfund site.

UCPI terminated operations at the site in May 1985. A CERCLA remedial investigation and feasibility study was also completed by CH2M Hill (*4*).

HYDROLOGIC SETTING

The project site is located on the alluvial plains of the Willamette River Valley and occupies about 1.5 acres of relatively level ground. The surface run-off from the site drains into an open ditch that runs along the northern edge of the site and is carried to Dry Creek located approximately 1½ mi southeast of the site. Dry Creek is a secondary tributary of the Willamette River, so the surface run-off from the site eventually discharges into the river.

The subsurface conditions are generally characterized by unconsolidated deposits of clay, silt, and gravel, which can be broken into three distinct units: an upper aquifer, a lower aquifer, and an aquitard separating the two aquifers. At the drilling locations, the upper soils in the profile consist of about 2 to 3 ft of miscellaneous fill or top soil and 15 to 20 ft of clayey silt to silt, ranging from mottled grayish brown to bright yellow in color and extending to depths of 17 to 21 ft below the ground surface. The less pervious soil layer, which separates the saturated pervious zones, consists of light-blue to dark-gray clay to silty-sandy clay ranging in thickness from 2.5 to 12 ft. The bottom of this layer extends 22 to 29 ft in depth below the ground surface over this site. Finally, the lower aquifer consists of wet, fine-to-coarse sands and gravels and is at least 15 ft thick over the site. A typical subsurface profile is shown in Figure 2. The groundwater table at the site fluctuates seasonally between 0 to 10 ft in depth. During the long, wet winters the site becomes water-logged because of infiltration of rain water. The groundwater table rises to the ground surface and the flow along the local drainage courses comes in contact with the groundwater in the shallow unconfined aquifer. In general, there is little flow of groundwater through the site, and the groundwater table has a slope of about 5 ft/mi (0.0009 ft/ft) in the north-northeast direction.

CHARACTERIZATION OF CONTAMINATION

The activities of UCPI had a serious impact on the local environment. The on-site soils, the groundwater in the upper aquifer, and the surface water are heavily contaminated with inorganic contaminants, viz., arsenic, barium, chromium, copper, iron, and lead. These substances are present in concentrations that exceed the primary drinking water standards. The most important of these contaminations is chromium. Elevated levels of chromium were detected in the vicinity of the gravel pit and the plating baths inside the building. The leaking of the plating tanks and disposal of waste waters in the gravel pit seem to be the primary causes of the soil and groundwater contamination at the site.

Presently, most of the contaminants are confined within the near-surface soils, and the nearly static groundwater regime has not allowed the contaminant plume to extend beyond several hundred feet down gradient from the source point. The boundaries of the plume are not well defined at the present time.

The distribution of contaminants within the affected soils has been found to vary with depth and distance from source points. The soils immediately overlying the confining silty clay layer in the vicinity of the source points show the highest concentrations of chromium (Figure 3). Chromium concentrations as high as 15 000 mg/L have been measured in the groundwater at one of these points. Concentrations about 200 times the primary drinking standard of 0.05 mg/L have been measured in the surface water.

The site poses a significant threat to the environment of the region because in the future, the contaminated groundwater from the site could eventually pollute the existing city wells and the surface drainageways, dangerously elevating the level of pollutants.

METHODOLOGY

Application of a dc electric field across an electrolyte brings about a host of complex interrelated phenomena that are collectively referred to as electrokinetics. In a stationary porous medium, fluid flow through the pores of the medium occurs in conjunction with the excess cations in the diffuse layer, and the process is called electroosmosis. In cases of the dispersion of solid particles in an otherwise stagnant fluid, movement of the solid particles can occur and the phenomena is called electrophoresis. Several other electrokinetic effects are currently used, including streaming current (flow of electric charge under the action of a hydraulic gradient), sedi-

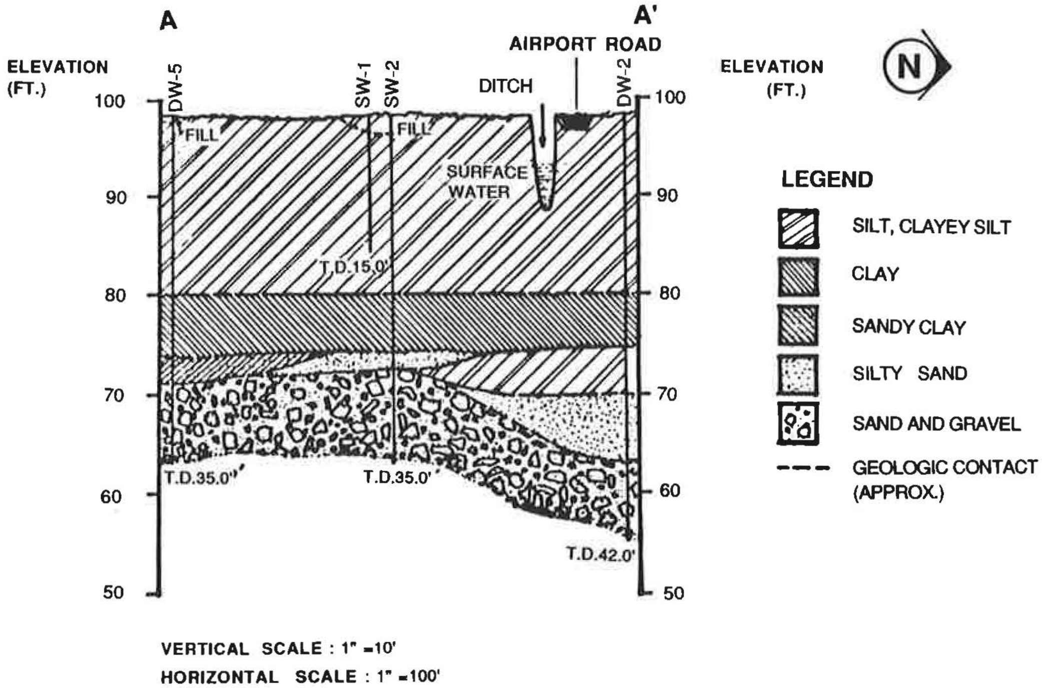


FIGURE 2 Typical subsurface profile at UCPI site (*3*).

mentation potential (electrical potential induced by movement of solid particles in a stagnant fluid) and dielectrophoresis (movement of uncharged matter in a nonuniform electric field). These processes, however, are generally used for purposes other than engineering.

Several basic processes are responsible for the transport of charge and matter caused by the application of a dc field across an electrolyte. These include ionic migration (drift of dissociated ions of the solute through the bulk solution towards the appropriate electrodes), ionic diffusion (movement of ions from regions of higher concentrations to regions of lower concentration), electrodic reactions (chemical reactions required for transfer of charge to and from the electrodes), as well as many other possible electrochemical reactions.

In the past, electrokinetics have been used in electroosmotic dewatering, electrical thickening, electrophoretic separations, and electrophoretic depositions, among other applications. Over the last two decades, extensive research efforts have been focused on seeking more novel applications of electrokinetics. It has been shown that the potential applications could address problems as diverse as injection and control of grouting in soils (*5,6*), electroreclamation of soils (*7–9*), and removal of corrosive chlorides from bridge decks (*10*). These studies suggest that electrokinetic treatment exhibits potential as a new approach to the decontamination of waste sites involving inorganic (heavy-metal) contaminants.

The basic concept involves the migration of ionic contaminants toward collection wells by application of a dc electric field, and the withdrawal of an enriched solution from the well by pumping.

EXPERIMENTAL PROCEDURES

The experimental program was designed to explore the removal of anionic chromium by a treatment scheme that combines both hydraulic advection and electrokinetic migration. Eleven 15-ft-deep wells with 4-in.-diameter slotted PVC casings were installed in an arrangement as shown in Figure 1. The wells labeled from H1 to H6 were placed at the vertices of a regular hexagon at 5-ft spacing. The wells labeled from SQ1 to SQ4 were arranged in a square pattern at 8-ft spacing. The central well SE1 and these other wells were designed to serve as electrode locations. During the experiments, steel reinforcing bars were placed in the appropriate wells to serve as electrodes. Two pairs of monitoring wells, M1 and L1, M6 and L6, were also installed along the lines joining the central well and H1 and H6, respectively. The line connecting SQ1 and SQ3 aligned approximately in the direction of the prevailing groundwater flow.

The dc field was supplied by a dc power supply (Lambda, Model LES-F-03-OV). The potentials and currents were monitored by a multimeter at the output end of the power supply. A microcomputer (AT&T PC 6300) based data acquisition system (Burr-Brown PCI-2000) was also used to keep continuous records of voltage and current. The water level at the wells was measured by a level indicator (Slope Indicator, Model 5145) and recorded manually.

The groundwater was sampled from selected wells by two GEOPUMPs, (peristaltic pumps) at constant depths of 11 and 14 ft below ground surface. The temperature, specific resistance, dissolved oxygen, and pH of the samples were measured immediately after retrieval. Concentrations of total Cr, Cr (VI), and other cations were measured by the Environmental Engineering Laboratory at Oregon State University within 24 hr. Cations and anions were also analyzed later in the acidified samples by atomic absorption (AA) spectrophotometry, ion chromatography (IC), and inductively coupled plasma atomic emission spectrometry (ICP) at the University of Washington in Seattle.

The field experiments involved the determination of initial (background) conditions and the effects of pumping and com-

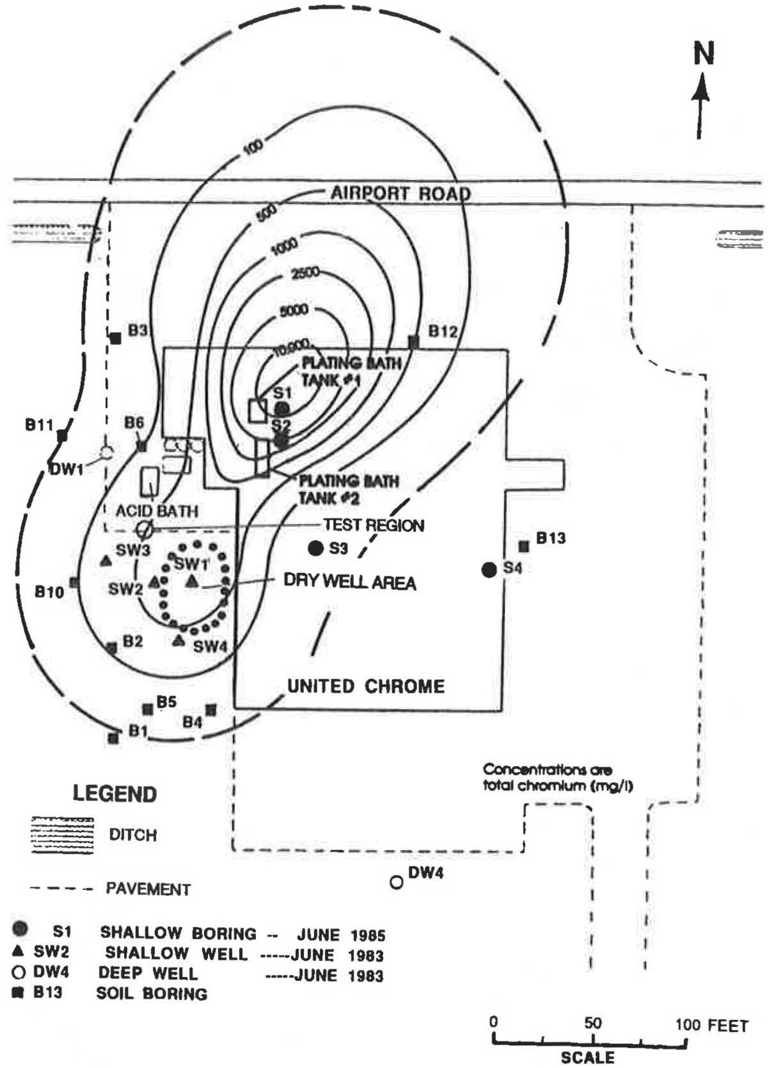


FIGURE 3 Approximate contours of total chromium concentration (*4*).

bined treatment. Two types of combined treatment were tested: (a) electrokinetic treatment with continuous pumping, and (b) electrokinetic treatment with pulsed or periodic pumping. Different pumping rates, electrical currents, and voltage drops were used in these tests. In between successive experiments, all the wells were purged by withdrawing groundwater in the amount of three times the well volume. All but one of the tests used the hexagonally arranged wells (as cathodes) and the central well (as anode) in order to achieve a more radially uniform potential field; the other test was performed with a triangular arrangement of electrodes. All of the wells were used to monitor the treatment effects.

RESULTS AND DISCUSSION

Spatial Distribution of Chromium

Before the beginning of the field tests, two sets of samples from all wells were taken at different depths (11 and 14 ft below the well surface) to ascertain the spatial distribution of chromate. Between the two sets of sampling, one well volume of water was withdrawn from all wells to achieve redistribution of the polluted groundwater. Although total chromium concentration greater than 100 mg/L was found in wells SE1 (390 mg/L), H1 (218 mg/L), H5 (205 mg/L), H6 (500 mg/L), SQ1 (615 mg/L), and SQ4 (245 mg/L), a much lower total chromium concentration (ranging between 1 and 78 mg/L) was found in wells H2 (9 mg/L), H3 (2 mg/L), H4 (18 mg/L), SQ2 (1 mg/L), and SQ3 (78 mg/L). Also, the concentrations in the aquifer were noted to increase with depth, possibly because of higher density of the solution. This nonhomogeneity of the distribution of the contaminant made the interpretation of the test results significantly more difficult.

Aquifer Characteristics

During the pumping tests, the changes of groundwater level in wells indicated nonhomogeneous hydrogeological conditions in the test region. The groundwater level changes were different in the direction of SE-M1-H2-SQ1 than in the direction of SE1-H4-SQ3 (Figure 1). Theoretical drawdown curves were developed from the modified Theis solution using transmissivity (T) of 0.45 and storage coefficient (S) of 10^{-6}. It should be mentioned that these values of T and S are markedly different from T of 0.305 and S of 0.14 (*4*), which were used to make design calculations to establish rate of pumping. In addition, no significant changes in groundwater levels during electrokinetic treatment (by itself) indicated that the influence of electroosmotic flow was small.

Temperature

While dc power was being applied, the water temperature in SE1 was slightly higher (<1°C) than the temperature in the other wells. This increase of temperature was probably caused by the heating effect at the anode. No similar trend was observed at the cathodes. However, the fluctuations of ambient temperature had more significant influence on the temperature of the wellwater than did the heating effect caused by the applied dc power.

Potential Distribution

For cylindrical geometry of the electrode arrangement, the theoretical solution for the potential distribution is available in the literature (*5*). The measured values of the potentials and the theoretical values are compared in the normalized plot (Figure 4). The agreement between the experimental and the theoretical curves is reasonable. Some potential drops (usually called polarization or activation overpotentials) do take place at electrode-solution boundary layers so that effective voltage across the electrolyte is given by the voltage at the power supply minus the sum total of the overpotentials. The potential gradient, which is the driving force for ionic migration, is much higher in the immediate neighborhood of the central anode or cathode than elsewhere. This implies that the ionic migration would be expected to be much more effective in the central region than elsewhere.

Electrode Reactions and pH

The observed changes in pH of the samples exhibited consistent trends in these experiments. A relatively small decrease of pH was noted at the well during pumping tests. But, electrokinetic treatment brought about much larger changes in pH at wells that served as working electrodes. An increase in pH at the cathodes and a decrease at the anodes were invariably associated with the treatment. The pH changes at the electrodes, in fact, occur because of the electrodic reactions. The major reactions occurring at the electrodes were water electrolysis at the anode and the cathode:

$$2H_2O + 2e^- \leftrightarrow 2OH^- + H_2 \qquad \text{(cathode)}$$

$$H_2O \leftrightarrow 2H^+ + \frac{1}{2}O_2 + 2e^- \qquad \text{(anode)}$$

and redox reactions such as the reduction of chromate at the cathode and the dissolution of the anode:

$$HCrO_4^- + 7H^+ + 3e^- \leftrightarrow Cr^{+3} + 4H_2O \qquad \text{(cathode)}$$

$$\left.\begin{array}{l} Fe \leftrightarrow Fe^{+3} + 3e^- \\ Fe^{+3} + 3H_2O \leftrightarrow Fe\,(OH)_{3(s)} + 3H^+ \end{array}\right\}\text{(anode)}$$

These reactions were supported by the pH increase from 5 to above 12 at the cathodes and the rapid increases of iron concentration and iron precipitates at the anode. Anode oxidation had long been noted in many engineering applications of electrokinetics (*11–13*) and electroosmosis and resulted in pH variations at the anode between 2 and 5.5 during the treatment.

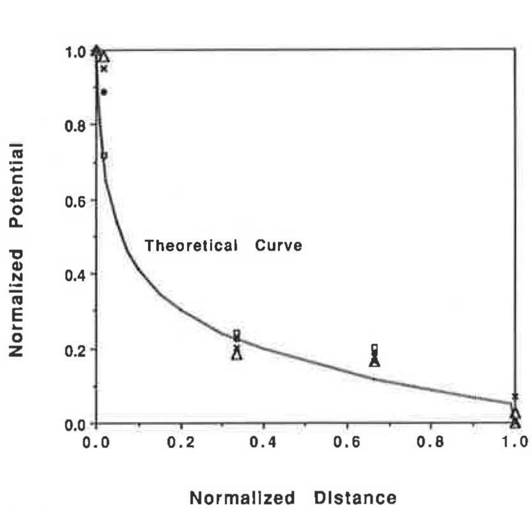


FIGURE 4 Variation of normalized electrical potential with distance from anode.

Chromate Removal

Evaluation of the effectiveness of three different treatment procedures in removing chromate and dichromate from the upper aquifer was of primary concern in these experiments. The treatment procedures explored in these experiments include (a) extraction of contaminants by pumping, (b) combined electrokinetic and pumping treatment, and (c) electrokinetic treatment with occasional withdrawal of effluents. Separate experiments were performed to explore these effects. In the last group of electrokinetic experiments, effluents from the anode were extracted at the nominal rate of 2 gal every 12 hr.

During the continuous pumping tests, it was observed that in comparison to the initial chromium concentration, the chromium concentration of the pumped water was considerably higher (Figure 5). While this higher concentration was maintained as long as the pump was running, at the end of the pumping the concentrations rapidly decreased and gradually increased back to initial concentration. In all of these experiments, continuous pumping caused mixing of the contents of the well and more concentrated solutions from greater depths were extracted during pumping. At the end of pumping, dilution caused by influx of fresh groundwater quickly reduced the concentration. In spite of different rates of pumping (0.17 and 0.26 gal/min), the concentrations in the central well were about the same during both of the pumping tests.

In order to study the combined effects of continuous electrokinetic and pumping treatment, the other test was carried out with a pumping rate of 0.16 gal/min and at a constant dc current of 6 amperes. Again, an immediate increase in total

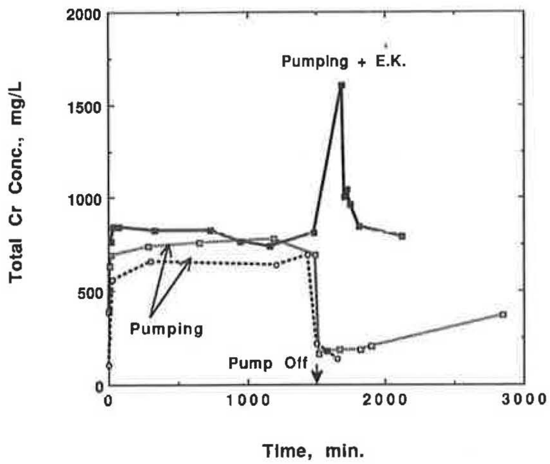


FIGURE 5 Total chromium concentrations during pumping with and without electrokinetic treatment.

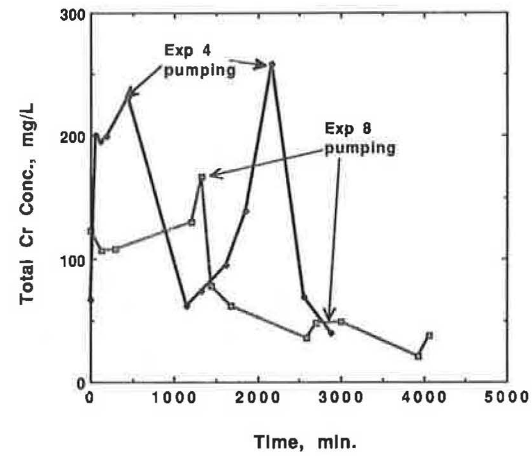


FIGURE 6 Total chromium concentrations during electrokinetic treatment with periodic pumping.

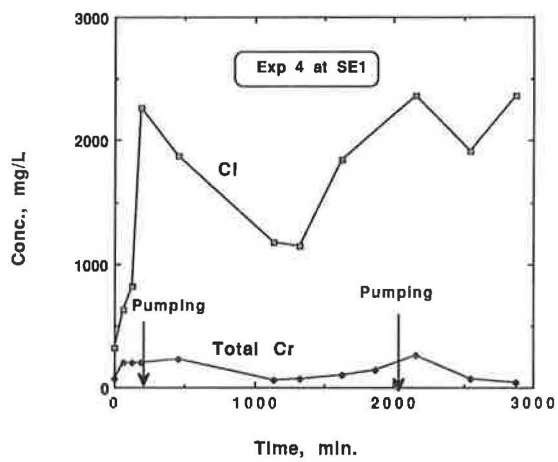


FIGURE 7 Total chromium and chloride concentrations at anode during electrokinetic treatment with periodic pumping.

chromium concentration at the central wall had occurred during this experiment. The concentration sustained during this experiment was slightly higher (800 mg/L) than the corresponding concentration (770 mg/L) noted in the pumping tests. The posttreatment concentrations were different in this test. For several hours after the termination of treatment, a higher level of chromium concentration than in the pumping tests was sustained in this test. This difference seems to suggest that during the treatment period the effects of pumping might have obscured the effects of electrokinetic migration, but with the influx of groundwater towards the central well at the end of treatment, the migrating front of higher concentration advected toward the central well.

The other group of experiments (electrokinetic treatment with occasional extraction of effluents) used the central well as the anode. Figure 6 shows the chromium concentration versus time curves at the anode obtained from these tests. Generally, the trend of concentration changes indicates an increase in concentration during treatment with an abrupt decrease on withdrawal of the effluents. The overall effect, however, was a gradual reduction of concentration. The response in Experiment 4 was particularly pronounced for two reasons: (a) higher initial concentration; and (b) higher applied current. The different rates of response indicate that the electrokinetic migration process is dependent on initial or pretreatment concentration and the total current. The dependence of the treament effect on initial concentration is clearly shown in Figure 7; chloride ions responded more vigorously to the same treatment because of higher initial concentration of chlorides. Meanwhile, the concentration of anionic chromium at the cathodes decreased significantly.

Removal of Chloride and Other Ions

Obviously, advection and migration were the two major transport processes involved in these field experiments, diffusion being of minor importance except in the vicinity of the electrodes. Although ion migration is a discriminative process (in the sense that it drives cations towards the cathode and anions to the anodes), advective transport of both cations and anions occurs in the prevailing direction of overall flow of the electrolyte. On the other hand, in the absence of advection, ionic transport depends strongly on the strength of the applied field (or current density), charge number (or valence), and diffusivities of ions. In other words, ionic species with higher diffusivity and lower charge number (chloride, for instance) can be expected to transport faster than other ions. But, when advective and electrokinetic transport are combined, the effects of migration may be masked by that of advection because the former process is much slower (flow velocity of the order of 10^{-5} cm/sec under a unit field strength). Furthermore, other side effects such as electrode dissolution of the anode and electrolysis of water at the cathode would affect the removal of ions because these effects may introduce additional competing ions.

The evidence of ion migration is conspicuous in the results of the group of experiments in which the two treatment processes were combined. Figure 8 shows the results of pumping and combined tests at the well SE1. The chloride ion migration was greater than that of chromate ions because of their monovalence, higher mobility, and of course, higher concen-

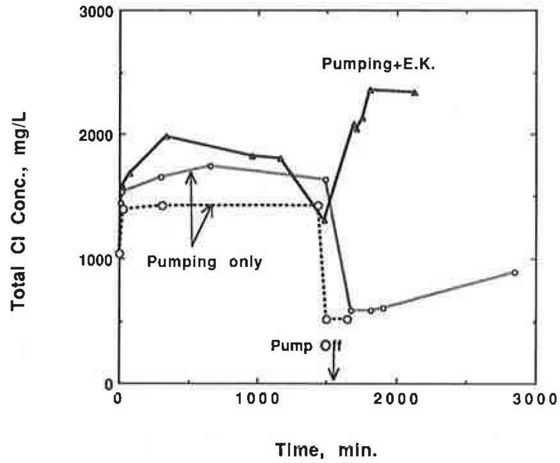


FIGURE 8 Chloride concentrations during pumping with and without electrokinetic treatment.

tration. All of the cationic species except iron decreased in concentration at the anode. The trends of iron and chloride concentration changes were similar, implying that dissolution of the anode was accompanied by migration of anions to the anode.

Figure 9 shows the concentrations of the major ions at the cathode during electrokinetic treatment with occasional withdrawal of effluents. Increase of sodium and calcium ion concentrations and decrease of both chloride and chromate concentrations are also direct effects of ion migration. These effects are relatively smaller than those at the anode because of the lower potential gradient in the vicinity of the cathode as discussed previously.

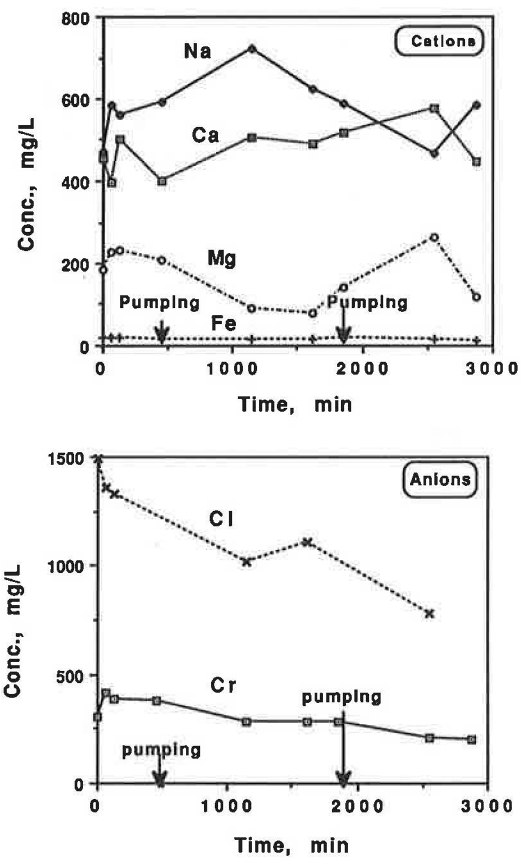


FIGURE 9 Major cation and anion concentrations at cathode during electrokinetic treatment with periodic pumping.

Final Remarks

By comparing the variation of concentration of chloride ions, it can be seen (Figure 10) that the electrokinetic applications of low current (5 amp) yielded less transport of anions than high current (10 amp) did. The use of constant voltage (30 volts) not only yielded higher currents—initial current at 17 amp, 10 amp after 6 hr, and 6 amp after 24 hr—but also induced more migration of chloride ions in 24 hr. If the chromate ions were the major anions, the removal efficiency of chromate by the combined treatment would be much higher.

After several tests, the chromate concentration in the central well SE1 was low (<50 mg/L) even after several pumping and purging cycles. This reduction of chromate near the region of the central well was caused by the combined treatment. In order to assess the effectiveness of the combined treatment schemes, the mass of chromium removed per well volume of effluent withdrawn was computed from the results of the appropriate experiments. The total mass of chromium recovered per well volume of effluents by pumping, combination of electrokinetics with continuous pumping, and combination of electrokinetics with occasional pumping were, approximately, 4.77, 5.15, and 9.23 g, respectively. Finally, passage of electric current may provide the impetus for desorbing offensive ions that may otherwise be firmly attached to the soil solids and would not be removed by pumping.

CONCLUSIONS

The feasibility of electrokinetic treatment to accelerate removal of chromate from soils was evaluated in this study. Although the field heterogeneity had an obscuring effect, proper combination of pumping with the electrokinetic treatment substantially shortened the time frame of treatment. The

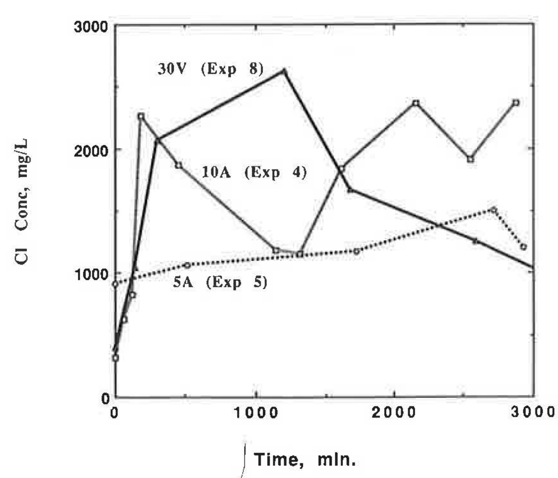


FIGURE 10 Chloride concentrations at anode for constant-current and constant-voltage applications.

major conclusions that can be drawn from this study are the following:

1. In electrokinetic treatment, electrically induced ion migration plays a significant role. However, the process of ion migration is slow and its effect may be enhanced or suppressed by the effects of advective groundwater movement or pumping.

2. The factors that influence the results of electrokinetic applications include (a) geological and hydrogeological setting, (b) nature of pollutants, and (c) operational conditions. The favorable geological and hydrogeological factors include moderate permeability of soils, nearly static groundwater, and moderately large areal extent of the site. The chemistry and the composition of the ion species are decisive factors that control the mobility of the contaminants within the subsurface system. The experimental constraints, e.g., direction of electrical field, the type of application (constant voltage or constant current), the direction of fluid flow, the electrode material, and the electrodic reactions can significantly influence the overall outcome.

3. Two types of pumping operations were used in conjunction with electrokinetic treatment; continuous pumping is suitable for removing pollutants at high concentrations and periodic pumping is more effective for concentrating pollutants at low concentrations.

4. Electrodic reactions may introduce other side effects during treatment, for instance, electrolysis of water may change the chemistry of the ion species, and the pH of water and the dissolution of metal electrodes may introduce other species into the system.

The present knowledge is not sufficiently developed that a mathematical or theoretical model of the treatment effects can be formulated. In addition, the treatability of particulate and colloidal pollutants and the mixtures of inorganic and organic contaminants has not been studied. These research needs must be addressed before the potentials of the technique can be fully realized.

ACKNOWLEDGMENTS

This research was supported wholly or in part by the U.S. Environmental Protection Agency through a cooperative agreement with the University of Washington.

REFERENCES

1. R. D. Hill. Demonstration of New and Innovative Hazardous Waste Treatment Technologies—The SITE Program. *Proc., 61st Annual Conference of the Water Pollution Control Federation*, Dallas, Tex., Oct. 1988.
2. S. Banerjee, J. J. Horng, J. F. Ferguson, and P. O. Nelson. *Field-Scale Feasibility Study of Electrokinetic Remediation.* Risk Reduction Engineering Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio, 1989.
3. Ecology and Environment, Inc. *Field Investigation of Soil and Water Conditions at United Chrome Products, Inc., Corvallis, Oregon.* TDD R10-8308-04. Seattle, Wash., 1983.
4. CH2M Hill. *Remedial Investigation Report—United Chrome Products Site, Corvallis, Oregon.* USEPA 106.OL32.0, U.S. Environmental Protection Agency, Seattle, Wash., 1985.
5. M. I. Esrig. Pore Pressures, Consolidation, and Electrokinetics. *Journal of the Soil Mechanics and Foundations Division, ASCE*, Vol. 94, No. SM4, July 1968, pp. 899–921.
6. T. Yamanouchi and S. Matsuda. Stabilization of Sand Feasible to Liquefaction by Means of a Kind of Grouting Applying Electro-Osmosis. *Proc., Symposium on Recent Developments in the Analysis of Soil Behavior*, University of New South Wales, Australia, 1975, pp. 411–421.
7. H. J. Gibbs. Research on Electro-Reclamation of Saline-Alkali Soils. *Transactions of the American Society of Agricultural Engineers*, Sept. 1966, pp. 164–169.
8. R. Hamnet. *A Study of the Process Involved in the Electro-Reclamation of Contaminated Soils.* M.Sc. thesis, University of Manchester, Manchester, England, 1980.
9. A. N. Puri and B. Anand. Reclamation of Alkali Soils by Electrodialysis. *Soil Science*, Vol. 42, 1936, pp. 2327.
10. J. E. Slater, L. R. David, and P. J. Moreland. Electrochemical Removal of Chlorides From Concrete Bridge Decks. *Material Performance*, Vol. 15, Nov. 1976, pp. 21–26.
11. L. Casagrande. Electro-Osmosis in Soils. *Geotechnique*, Vol. 1, 1949, pp. 1959–1977.
12. B. A. Segall, C. E. O'Bannon, and J. A. Matthias. Electroosmosis Chemistry and Water Quality. *Journal of Geotechnical Division, ASCE*, Vol. 106, No. GT10, 1980, pp. 1148–1152.
13. R. H. Sprute and D. J. Kelsh. *Electrokinetic Consolidation of Slime in an Underground Mine.* Report of Investigation 8190, Spokane Mining Research Center, Spokane, Wash., 1976.

This research does not necessarily reflect the views of the U.S. Environmental Protection Agency and no official endorsement should be inferred.

Publication of this paper sponsored by Committee on Physicochemical Phenomena in Soils.

Altimetric Sensing of Currents: Spatial Averaging and Sampling Impacts on Ocean Routing

Hong Kam Lo and Mark R. McCord

Strategic ship routing impacts of various spatial resolutions associated with satellite-based altimetric sensing of ocean currents are evaluated. Underlying true current patterns are averaged to more aggregate grids, routes are chosen designed to exploit the currents from these aggregate estimates, and the underlying true current patterns are used to simulate relative fuel consumption on these routes. The underlying current patterns are produced from an advanced model of a spatially variable area of the North Atlantic Gulf Stream, and the route origins and destinations are confined to this region. The results indicate that the 5 × 5 degree spatial resolution of available data is too coarse to be of use in strategic routing in this area. However, the improved resolution that would be compatible with satellite altimeter systems would lead to positive fuel savings, and savings that are fairly close to those available if the underlying data were known in their finest detail. Sampling along geographically separated ground tracks can still lead to positive fuel savings, even in this spatially variable region. Impacts for practical applications are discussed.

Weather and waves are the principal environmental inputs to the choice of a commercial ocean vessel's general path (*1,2*). If ocean currents are considered in the attempt to reduce fuel costs or time of passage, it is usually through grossly averaged (in time and space) pilot chart data. Even these gross averages are strong enough and oriented in such a way that they can influence strategic ship routes (*3*). However, currents are dynamic (*4*), and to be of use in strategic routing, estimates would have to be obtained with better frequency than that offered by the pilot charts. But, other than in a few special cases (*5,6*), dynamic estimates of current patterns have not been available.

For these reasons, a collaborative project has been designed to provide timely estimates of current patterns to the shipping industry. The ability to do so is predicated on the technical advances in satellite altimetry and the schedule of proposed altimeter missions (*7*). An altimeter measures sea surface heights, which can then be analyzed to deduce ocean currents. Altimeters offer distinct advantages, compared with infrared sensors, for example. They can obtain measurements in cloudy areas and determine current velocitites, as well as locations. They can also sense current activity below the water surface, which is important for ships of any draft. Simulation studies (*8*) indicate that the 5-cm accuracy of the GEOSAT altimeter and, therefore, the improved 2-cm accuracy of the planned TOPEX/Poseidon altimeter (*6*) are more than sufficient for estimating current velocities, giving variances of less than 0.03 nautical mph. Orbit errors (*6*) would have even less of an effect (*8*), because they occur over much greater arc lengths than those used to determine ocean currents. Moreover, the need to determine only sea surface slopes in the geostrophic analysis (*9*), rather than absolute heights, eliminates much of the spatially correlated error signal (*8*).

However, there remain potential difficulties associated with using altimeters for sensing of ocean currents (*7*). The systems being developed would take advantage of altimeters flown for primarily scientific and military reasons (*5–7*). This strategy would greatly reduce the costs to the shipping community, but the dependence on other owners, operators, and preprocessors increases the institutional risk (*7*) associated with data supply. On the technological and scientific side, errors from estimating the geoid, the equipotential gravitational surface from which sea surface slopes must be referenced for analysis (*6*); missing data, arising primarily from the time it takes the altimeter to lock back into proper measurement angle after flying over land (*10*); the temporal resolution of an orbiting altimeter (*4*); and the spatial coverage offered are presently being investigated. The independent effects of the separate error sources are being investigated to see which ones seem most critical. As an example, measurement accuracy or orbit error are no longer of concern, although this was not the case before conducting of studies (*8*). Moreover, this type of independent error analysis can lead to lower-bound estimates of the combined effects.

Two aspects associated with the spatial resolution of a current estimating system are investigated. The first deals with the inherent spatial variability of current patterns. Any sensing and estimating system will produce current estimates at relatively coarse spatial scales compared to the tens of meters which would affect a ship. Does spatially averaging mask the inherent structure which is important to route selection? The second aspect relates to the sampling schemes of orbiting altimeters. Satellites on which altimeters are flown circle the earth, sampling the ocean along spatially separated ground tracks, For example, the North Atlantic ground tracks of the GEOSAT altimeter (*10*) are shown in Figure 1. The spatial separation of the tracks is 1.53 degrees (approximately 166 km) at the Equator and 1.46 degrees (approximately 123 km) at the midlatitudes, where the variable Gulf Stream becomes important to ship routes. The upcoming TOPEX mission will have even greater separation between ground tracks (*11*). Can such sparse sampling lead to estimates that represent the spatial structure well enough to assist in strategic routing?

In order to answer these questions, the best available data

Department of Civil Engineering, Ohio State University, 470 Hitchcock Hall, 2070 Neil Ave., Columbus, Ohio 43210.

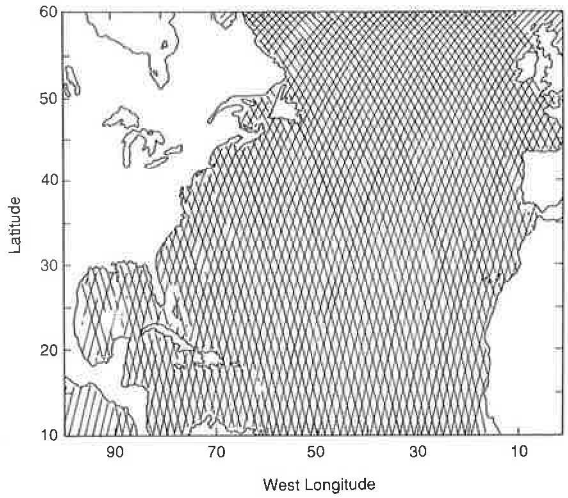


FIGURE 1 GEOSAT ground tracks in North Atlantic.

set on current patterns is used. In response to the first question, these data are aggregated to various levels of spatial resolution, routes are chosen on the basis of the current data, and ship performance is simulated on the basis of the original disaggregate data. To address the second question, sampling is conducted only along satellite ground tracks before aggregating the data.

The methodology is discussed in more detail in the next section. In the third section, the results are presented that show that choosing routes based on data aggregated to levels consistent with planned altimeter missions can still lead to interesting fuel savings. Moreover, the effect of spatially sampling only along ground tracks before aggregating has a slight, but not overwhelming effect. The results also indicate that the spatial resolution of traditional data is too coarse to be of much use for strategic routing, at least in regions as spatially variable as the one used here. One should be careful in interpreting the results, however. Although the results show that the spatial resolution in itself is no problem, the same may not be true when the time necessary to obtain this resolution is considered.

METHODOLOGY AND DATA

The basic design used to investigate the effect of spatially averaging current vectors can be summarized by the following steps: (a) Select a data set of ocean currents in a given region with a given spatial resolution and consider this as the true data; (b) Partition the region into grid cells consistent with a given coarser spatial resolution; (c) For each cell, determine a single representative current vector based on all the vectors in the cell; (d) For each of a set of origin-destination pairs on the boundaries of the region, choose a strategic route through the gridded current data and a route that would be followed if the current data were ignored; (e) Determine a performance measure of the current-based route relative to the route that ignores the current information by simulating the effect of the original current data set on the routes selected; (f) Repeat Steps (b) through (e) with partitions consistent with different spatial resolutions; (g) Repeat Steps (b) through (f) with different origin-destination pairs; (h) Repeat Steps (a) through (g) for different data sets. The design used to investigate the effect of sampling only along ground tracks is the same except that only current vectors in the original data set that intersect the satellite ground tracks (see, e.g., Figure 1) are used when forming the representative currents in Step (c).

The process uses three general inputs—a set of ocean currents, a strategic ship routing model, and a performance measure. Each of these is described in more detail.

Ocean Current Data

Harvard University has been working with the U.S. Navy to develop a forecasting model (*12,13*) for an area of the Gulf Stream (see Figure 2) bordered by (north) latitude and (west) longitude coordinates (39°, 74°), (32°, 72°), (38°, 50°), and (46°, 55°). The model is driven by altimetry, infrared observations, and in situ measurements. Harvard researchers submitted surface current output on a 15- by 15-km grid for two 5-week periods. The fine resolution of this data, coupled with the advanced modeling techniques and the data acquisition systems used, makes this the best and most comprehensive data set compiled over a defined geographical region.

The 15- × 15-km Harvard data were aggregated into grid cells of 0.1° latitude by 0.5° longitude, with a current vector (actually, two orthogonal current velocity components) in each cell. This partition, which is referred to as 0.1 × 0.5 resolution for simplicity, served as the underlying true data. The model Gulf Stream region was also partitioned into cells 0.3° latitude by 1.5° longitude, 0.5° latitude by 2.5° longitude, 0.6° latitude by 3.0° longitude, 1.0° latitude by 5.0° longitude, and 5.0° latitude by 5.0° longitude (see Table 1). The 0.3 × 1.5, 0.5 × 2.5, and 1.0 × 5.0 resolutions correspond to resolutions that would allow only a few, several, and many, observations from the GEOSAT ground tracks, respectively. The 0.6 × 3.0 resolution would allow several observations from TOPEX ground tracks. The 5.0 × 5.0 resolution corresponds to that offered by the traditional pilot charts.

In order to determine the current vectors representing a cell at a given resolution, the arithmetic average of each of the two orthogonal velocity components for each Harvard observation falling into the cell was taken. Figures 3 and 4

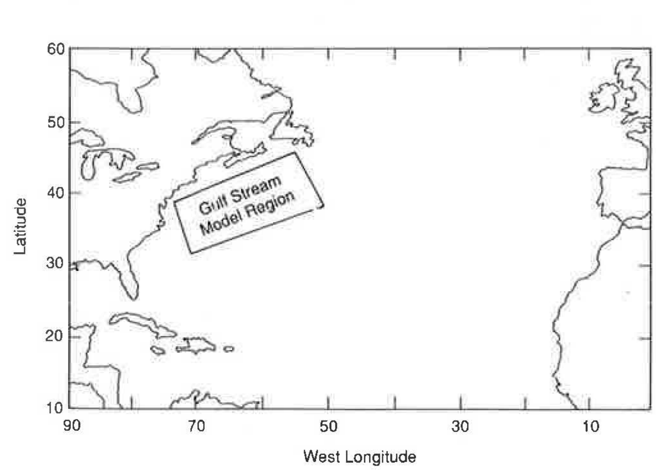


FIGURE 2 Location of Harvard Gulf Stream Model region.

TABLE 1 SPATIAL RESOLUTION OF GRID CELLS USED IN CURRENT AVERAGING

Resolution (degrees latitude × degrees longitude)	Mnemonic
0.1 × 0.5	Underlying "True" Data
0.3 × 1.5	Fine GEOSAT Resolution
0.5 × 2.5	Moderate GEOSAT, Fine
0.6 × 3.0	Moderate Topex Resolution
1.0 × 5.0	Coarse GEOSAT Resolution
5.0 × 5.0	Pilot Chart Resolution

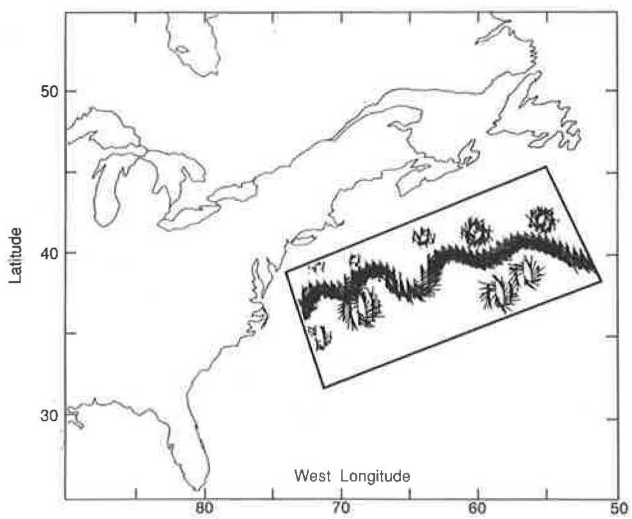


FIGURE 3 Study region current pattern at 0.1 × 0.5 resolution.

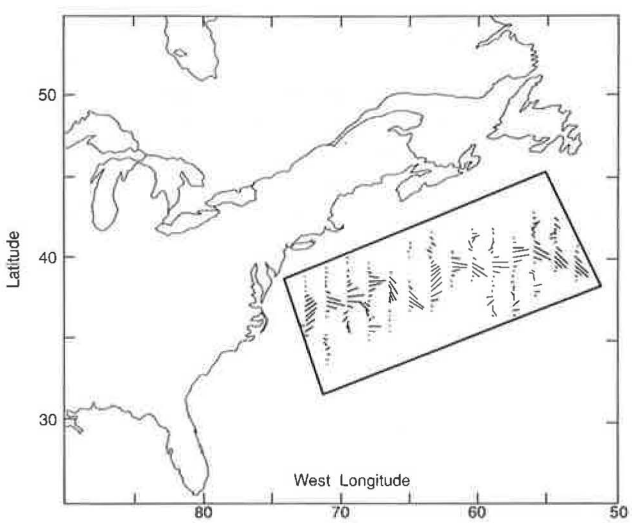


FIGURE 4 Study region current pattern at 0.3 × 1.5 resolution.

show one set of data at 0.1 × 0.5 and 0.3 × 1.5 resolutions. When analyzing the impact of sampling along GEOSAT or TOPEX ground tracks, only cell components of Harvard data that intersected these ground tracks were averaged. In both cases, for any resolution level, there was a single current vector (i.e., two orthogonal current speed components) for each cell at the end of this step.

Routing Model

For each resolution level, two routes were found—one that would be followed to minimize fuel consumption in the absence of current information, and one that would exploit the information in trying to minimize fuel consumption. In both cases, a ship traveling at constant velocity relative to the water and arriving at its destination (the exit point of the region) T hours after departing its origin (the entrance to the region) was considered. The constant relative velocity policy is optimal in the absence of information and is often recommended in practice (*3*). Moreover, Lo et al. (*3*) found relative fuel savings to be insensitive to this policy. Because the study area was in the open ocean (see Figure 2), fixing the time T to be the same along the two routes was equivalent to specifying boundary conditions and putting the two cases on the same footing. It also made determining the performance measure more straightforward and less dependent on arbitrary assumptions between time and fuel consumption tradeoffs.

The route followed in the absence of current information was the shortest-distance route. For a given origin-destination pair, this shortest-distance route would be the same for any level of current resolution and only had to be determined once. For expository purposes, however, it is easier to think of it as a separate route in each case. The time T was found by simulating a passage through the underlying currents over this route at constant velocity Vr relative to the water. A velocity Vr = 16 knots representative of tankers, the class of ships offering the largest potential fuel savings from current routing (*14*), was used. The route selected in the presence of currents was the minimum time route, through the estimated currents, from the origin to the destination. The velocity was that which would ensure arriving at the destination in time T along this route. The marginal impact of currents can, as a first approximation, be added to the effects of weather and waves (*3*), unless there is some reason to believe that these factors are highly correlated with the current activity. Moreover, as found by Lo et al. (*3*), currents shift routes only slightly, finely tuning the routes at much smaller spatial scales than would be consistent with storm-generated wind and wave patterns.

In order to find the minimum time route between the origin and destination, the ocean surface was discretized into a network, where node $N_n = (X_n, Y_n)$ represented a geographic location with longitude X_n and latitude Y_n, and the impedance on each arc was the time to go from its tail node to its head node. Approximating the earth as a sphere, the distance d_{ij} between two nodes $N_i = (X_i, Y_i)$ and $N_j = (X_j, Y_j)$ was calculated by Robinson et al. (*15*) as

$$d_{ij} = \cos^{-1}[(\sin Y_i \sin Y_j) + (\cos Y_i \cos Y_j \cos P)] * R \quad (1)$$

where P was the number of degrees longitude between the two nodes (i.e., $P = |X_j - X_i|$), and R was the radius of the earth (taken as 3,440.4 nautical miles).

An arc could traverse several grid cells, each with different current vectors. The average current velocity, Vc_{ij}, along an arc connecting nodes as N_i and N_j was calculated as the distance-weighted projection of the current vectors in cells through which the arc passed in the direction of the arc. The velocity with which the ship would cover the distance d_{ij}—the over-the-ground velocity Vg_{ij}—was found by vector addition of the current velocity Vc_{ij} to the constant velocity of the ship through the water Vr = 16 knots (i.e., $Vg_{ij} = Vc_{ij} + 16$). Knowing d_{ij} from Equation 1, the time to traverse the link was then

$$t_{ij} = d_{ij}/(Vc_{ij} + 16) \qquad (2)$$

So, given a network configuration and the grid cells of current vectors for a specific resolution, t_{ij} could be uniquely determined for each arc of the network. The minimum time route through the estimated current vectors could then be found using any traditional shortest-path algorithm [see, e.g., Larson and Odoni (*16*)]. To find the shortest distance route, the route that would be chosen in the absence of the current information, simply set all $Vc_{ij} = 0$ and run the same computer programs used to find the minimum time route.

A network was used in which 21 arcs fanned out from each tail node (beginning with the origin node), connecting head notes that were 0.5° longitude toward the destination from the tail node and separated by 0.1° latitude increments from 1.0° north of the tail node to 1.0° south of the tail node. To test the accuracy of this discretized representation of the ocean, shortest-path distances through the network between origins and destinations were determined and compared with great circle distances computed by substituting the coordinates of the origin-destination pairs in Equation 1. The network distances were longer, as would be expected, but within 0.2 percent of the great circle distances. The distances would be longer on both the minimum time and minimum distance paths analyzed, reducing any effect of the discrete approximation.

Performance Measure

The impact of the spatial aggregation was indicated by the relative fuel savings achieved when using current estimates in route selection. Letting FC_w and FC_{wo} represent, respectively, the fuel consumed on a specific passage when using currents in choosing the route and the fuel consumed on the passage when not using currents, the relative fuel savings, RFS, on a specific passage can be written as

$$\text{RFS} = 1 - (FC_w/FC_{wo}) \qquad (3)$$

To determine the ratio between FC_w and FC_{wo} in Equation 3, the relation given by Jansson and Shneerson (*17*) that approximates the rate of fuel consumed as varying with the third power of the velocity Vr(*t*) of the ship relative to the water at that time. The fuel consumed during time *T*, then, would be

$$FC = \int_0^T k[Vr(t)]^3 dt \qquad (4)$$

where *k* is a constant depending on the ship considered.

As mentioned earlier, the time *T* in the study region was assumed to be the same for the routes analyzed when using the currents as inputs to route selection and when not using them as inputs. Because the effect was analyzed on the same ship for two routes, *k* was also constant. Finally, because the ship was assumed to travel at constant velocity through the water, Vr(*t*) = Vr. With these assumptions, substituting Equation 4 in Equation 3 and simplifying,

$$\text{RFS} = 1 - (V_{r,w}/V_{r,wo})^3 \qquad (5)$$

where $V_{r,w}$ and $V_{r,wo}$ are the constant speeds relative to the water on the minimum-time route (that taken when using currents as inputs to route selection) and the minimum-distance route (that taken when not using currents as inputs), respectively. In the study, $V_{r,wo}$ was 16 knots (the same Vr value used to determine the minimum time route), and $V_{r,w}$ was the velocity along the minimum-time route that ensured arriving at time *T*.

The minimum time route was determined with Vr = 16 knots. There is no guarantee that the minimum-time route identified with Vr = $V_{r,w}$ would be the same route. However, $V_{r,w}$ never differed by more than 0.5 knot from the assumed 16-knot velocity, and Lo et al. (*3*) obtained the same minimum-time routes whether 16 knots or $V_{r,w}$ was used. (Also, as proven by Lo et al. (*3*), when the two minimum-time routes are identical, the route identified with this procedure is guaranteed to be the minimum-fuel-consumption route.)

RESULTS

In order to determine origin-destination pairs, the current study area was laid over trans-Atlantic trade routes found by U.S. Department of Transportation (*18*) and four locations on the west wall and four locations on the east wall were selected on the basis of where the routes entered and left the area. The coordinates of these locations are shown in Figure 5. Then, analyzing 16 eastbound routes formed by using each of the four western wall locations as origins and each of the four eastern wall locations as destinations, RFS in Equation

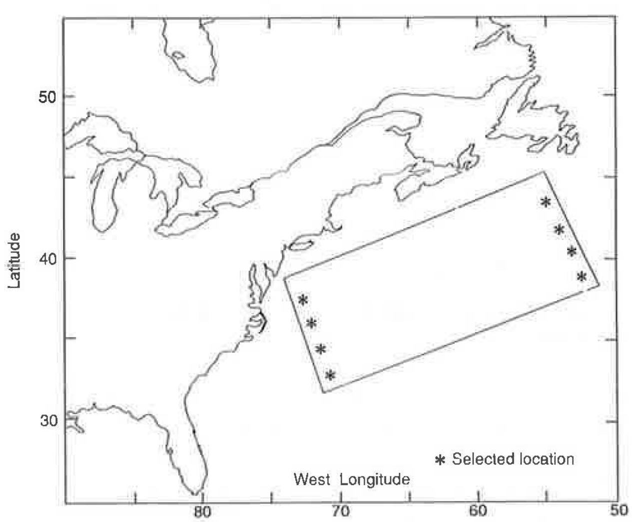


FIGURE 5 Location of points serving as origins and destinations.

5 was determined for each. The 16 westbound routes formed by using each of the four eastern wall locations as origins and each of the four western wall locations as destinations were also analyzed.

As mentioned earlier, two 5-week periods of Harvard model outputs, i.e., 70 daily current patterns, were accessible. From these data, six different dates were chosen for analysis. McCord and Lo (*4*) found that the Gulf Stream reaches about 75 percent of its 6-month variability in current activity after approximately 14 days. Therefore, model outputs separated by 14 days or more could be considered relatively independent. The five dates analyzed were Oct. 28, 1987; Nov. 11, 1987; Nov. 25, 1987; April 20, 1988; May 4, 1988; and May 18, 1988.

For each date and each resolution level, an average fuel savings of the 16 eastbound routes and an average fuel savings of the 16 westbound routes were determined. The impacts of the spatial averaging are presented in Tables 2 and 3. As expected, the performance is best for finest resolution (0.1 × 0.5) and consistently degrades, with minor exceptions, as the resolution becomes coarser. The 5.0 × 5.0 resolution, that consistent with available pilot chart data, would be of little use in strategic routing. The average savings are only 0.1 and 0.5 percent on westbound and eastbound routes, respectively, with several dates offering zero or negative route-averaged savings.

More positively, three of the resolutions consistent with satellite sampling—0.3 × 1.5, 0.5 × 2.5, and 0.6 × 3.0—performed well, always producing positive fuel savings. Moreover, these resolutions, especially the 0.3 × 1.5, produced savings close to the maximum possible savings, those obtained from the true 0.1 × 0.5 resolution. The coarse GEOSAT resolution, 1.0 × 5.0, did not perform as well, but still produced an average 3 percent fuel savings on eastbound routes and 1 percent fuel savings on westbound routes.

The impacts of sampling only along ground tracks before averaging are presented in Tables 4 and 5. GEOSAT gound tracks were averaged to 0.3 × 1.5 and 0.5 × 2.5 resolution and the more spatially separated TOPEX ground tracks to 0.6 × 3.0 resolution. Comparing Tables 4 and 5 with Tables 2 and 3, the savings decreased when sampling along ground tracks. Still, all of these samples produced routes with positive fuel savings compared with the shorter great circle routes. They also produced averages higher than those found in the evaluation studies (*3,4*).

From Tables 2–5, the fuel savings were generally higher when using the April and May data than when using the October and November data. This agrees with Lo et al. (*3*), who found that warmer months performed, on average, better than colder months. This is of practical interest, because route analysts say that business declines as the summer approaches and the weather improves. They hope to use timely current estimates as an additional product to keep more business during this period.

Regardless of the season, though, better fuel savings were obtained by averaging over smaller spatial areas. The type of resolution available with satellite altimeters performed relatively well, however, even when sampling only along ground tracks. The resolution provided by presently available pilot charts did not perform well, at least in the spatially variable Gulf Stream region.

DISCUSSION

These results were pleasantly surprising. The poor performance of the 5.0 × 5.0 resolution was not particularly astonishing. Practitioners do not place much confidence in pilot chart data that are averaged to this resolution. Lo et al. (*3*) mentioned that pilot chart data would probably wash out

TABLE 2 FUEL SAVINGS (%) OF EASTBOUND CURRENT-BASED ROUTES RELATIVE TO NONCURRENT-BASED (GREAT CIRCLE) ROUTES BY CURRENT RESOLUTION AND DATE

Resolution (degrees lat. × degrees long.)	Date						
	10/28/87	11/11/87	11/25/87	04/20/88	05/04/88	05/18/88	Avg.
0.1 × 0.5	9.8	7.2	8.0	9.6	10.1	8.7	8.9
0.3 × 1.5	7.8	5.4	6.2	7.7	8.9	7.7	7.3
0.5 × 2.5	6.3	4.3	4.6	5.9	5.8	5.4	5.3
0.6 × 3.0	6.0	3.0	3.4	6.4	7.1	4.7	5.1
1.0 × 5.0	3.9	−0.2	1.9	3.6	4.6	4.2	3.0
5.0 × 5.0	0.2	0.4	0.0	0.3	1.5	0.6	0.5

TABLE 3 FUEL SAVINGS (%) OF WESTBOUND CURRENT-BASED ROUTES RELATIVE TO NONCURRENT-BASED (GREAT CIRCLE) ROUTES BY CURRENT RESOLUTION AND DATE

Resolution (degrees lat. × degrees long.)	Date						
	10/28/87	11/11/87	11/25/87	04/20/88	05/04/88	05/18/88	Avg.
0.1 × 0.5	5.1	6.0	5.6	8.3	5.9	6.6	6.3
0.3 × 1.5	4.4	5.0	4.7	7.3	4.7	5.9	5.3
0.5 × 2.5	2.9	3.9	3.4	5.4	3.0	4.2	3.8
0.6 × 3.0	2.9	3.1	2.1	4.5	1.9	3.5	3.0
1.0 × 5.0	−0.6	1.8	0.6	2.3	0.7	1.8	1.1
5.0 × 5.0	0.2	0.8	−0.2	0.2	0.0	0.0	0.1

TABLE 4 FUEL SAVINGS (%) OF EASTBOUND CURRENT-BASED ROUTES RELATIVE TO NONCURRENT-BASED (GREAT CIRCLE) ROUTES BY GROUND TRACK SAMPLE AND DATE

Ground Track Sample	Date						
	10/28/87	11/11/87	11/25/87	04/20/88	05/04/88	05/18/88	Avg.
GEOSAT averaged to 0.3 × 1.5	7.8	5.1	5.1	6.6	8.2	7.4	6.7
GEOSAT averaged to 0.5 × 2.5	6.6	4.3	4.3	5.6	6.1	5.7	5.4
TOPEX averaged to 0.6 × 3.0	3.5	1.7	3.4	5.8	7.0	2.5	4.0

TABLE 5 FUEL SAVINGS (%) OF WESTBOUND CURRENT-BASED ROUTES RELATIVE TO NONCURRENT-BASED (GREAT CIRCLE) ROUTES BY GROUND TRACK SAMPLE AND DATE

Ground Track Sample	Date						
	10/28/87	11/11/87	11/25/87	04/20/88	05/04/88	05/18/88	Avg.
GEOSAT averaged to 0.3 × 1.5	4.1	4.8	3.9	6.3	3.8	5.1	4.7
GEOSAT averaged to 0.5 × 2.5	3.2	3.8	2.1	5.1	2.7	3.9	3.5
TOPEX averaged to 0.6 × 3.0	2.5	1.7	1.6	3.5	1.4	2.9	2.3

much of the spatial variability in current patterns that could be advantageous in strategic routing. The results support this claim. What was surprising was the good performance of those resolutions compatible with planned altimeter missions, and especially the fact that sampling only along ground tracks performed so well. The study was limited to an area encompassing only a small percentage of global ocean areas because of a lack of data outside this area. However, results in other areas would be similar or better, because of the recognized spatial variability in this area of the Gulf Stream (*12,13*). Moreover, the Gulf Stream and its eddies are spatially similar to the Kuroshio current, and together these are two of the most studied currents, in part because of their spatial properties, as well as their overall importance (*19,20*).

Although encouraging, the results should be interpreted carefully. They indicate that even in the variable Gulf Stream area, the spatial sampling parameters of satellite coverage can lead to positive fuel savings in a static world. However, currents are dynamic, and the ground tracks shown in Figure 1, and those in the TOPEX mission, can be covered only progressively in time. For example, GEOSAT repeated its entire coverage every 17 days, with only one-seventeenth of the tracks of Figure 1 being covered each day (*21*). McCord and Lo (*4*) found that 17-day coverage does not seem frequent enough to obtain viable estimates, and that an interpolation model or a second altimeter would be needed. TOPEX will have a 10-day repeat cycle, but its temporal resolution is of concern and its impact for strategic routing is being investigated.

Even if the temporal resolution and its combined effect with spatial resolution turns out to be troublesome, the results presented are still encouraging. Options for overcoming limited altimeter resolution include the development of an interpolation model and the supply of more than one altimeter. Interpolation models are being examined. The results presented here indicate that the coverage of the model (both inputs and outputs) could be limited to the ground tracks or to grid size resolutions on the order of those that performed well. This would represent sizable savings in computer storage and computational time. On the supply side, several satellites are expected to carry altimeters in the late 1990s. However, the missions will be under the jurisdiction of different agencies and even different governments. If the commercial routing community could access this data, and more optimistically, if the missions could somehow be coordinated, the temporal resolution would be greatly improved, and the promising results obtained here might be realized in practice.

ACKNOWLEDGMENTS

The authors would like to thank Scott Glenn, Allan Robinson, and the Harvard Physical Oceanography Department for access to their current data; Bruce Douglas and the NOAA Geodetic Laboratory and Lee Fu and the Jet Propulsion Laboratory for the satellite information; and Riyaz Fazal and Jim Mays for insight on routing practices. We are grateful to three anonymous reviewers for their helpful comments. This project was partially funded by NOAA, through the Ohio Sea Grant College Program; NASA and Oceanroutes, Inc., through the Center for the Commercial Development of Space component of the Ohio State Center for Mapping; and an NSF Presidential Young Investigator Grant.

REFERENCES

1. H. Chen. Ship Routing Made Practical. *Sea Technology*, April 1988, pp. 47–57.
2. J. Mays. Shipboard Computers for Navigation and Weather Routing. Presented at Marine Computers '86 Society of Naval Architects and Marine Engineers, New England Section, April 16–17, 1986.

3. H. K. Lo, R. M. McCord, and C. K. Wall. Value of Current Information for Strategic Routing. *European Journal of Operational Research*, Vol. 55, 1991, pp. 1–11.
4. M. R. McCord and H. K. Lo. *Temporal Variability in Ocean Controls*. American Society for Photogrammetry and Remote Sensing (ASPRS) Technical papers. ASPRS–ACSM Fall Convention, Cleveland, Ohio, 1989, pp. 83–90.
5. J. M. Bishop. *Applied Oceanography*. Wiley-Interscience, New York, 1984.
6. I. S. Robinson. *Satellite Oceanography*. Ellis Horwood Ltd., New York, 1985.
7. M. R. McCord and S. Smith. Remote Sensing of Ocean Currents for Improved Routing. *Proc., 14th Ship Technology and Research (STAR) Symposium*, Society of Naval Architects and Marine Engineers, Jersey City, N.J., 1989, pp. S5-4-1–S5-4-8.
8. L. Chase. *Uncertainty of Current Velocity Measurements Due to Altimeter and Orbit Determination Errors*. M.S. thesis, Department of Civil Engineering, Ohio State University, Columbus, 1988.
9. D. Tolmazin. *Elements of Dynamic Oceanography*. Allen and Univin, Inc., Boston, Mass., 1985.
10. *GEOSAT Altimeter Geophysical Data Research User Handbook*. NOAA Technical Memorandum NOS NGS-46. U.S. Department of Commerce, July 1987.
11. M. A. Vincent. *TOPEX/Poseidon Project Orbit Characteristics*. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, June 1990.
12. A. R. Robinson, S. M. Glenn, M. A. Spall, L. J. Walstad, G. M. Gardner, and W. G. Leslie. Forecasting Gulf Stream Meanders and Rings. *EOS*, No. 7, 1989, pp. 1464–1465.
13. A. R. Robinson and L. J. Walstad. The Harvard Open Ocean Model: Calibration and Application to Dynamical Process, Forecasting, and Data Assimilation Studies. *Applied Numerical Mathematics*, Vol. 3, 1987, pp. 89–131.
14. H. K. Lo, M. R. McCord, and C. K. Wall. Estimated Fuel Savings from Current Routing on U.S. Transoceanic Trades. Presented to Transportation Research Forum, Oct. 1990.
15. A. Robinson, R. Sale, J. Morrison, and P. Muehreke. *Elements of Cartography*, 5th ed. John Wiley, New York, 544 pp.
16. R. Larson and A. Odoni. *Urban Operations Research*. Prentice-Hall, Englewood Cliffs, N.J., 1987.
17. J. Jansson and D. Shneerson. *Liners Shipping Economies*. Chapman and Hall, New York, 1987.
18. *United States Oceanborne Foreign Trade Routes*. U.S. Department of Transportation, Sept. 1989, pp. 17–18.
19. H. Stommel and K. Yoshida. *Kuroshio: Physical Aspects of the Japan Current*. University of Washington Press, Seattle, 1972.
20. W. Van Dorn. *Oceanography and Seamanship*. Dodd, Mead & Co., New York, 1974.
21. D. R. McConathy and C. C. Kilgus. *The Navy GEOSAT Mission: An Overview*. Johns Hopkins Applied Physics Laboratory Technical Digest, Vol. 8, No. 2, Baltimore, Md., 1987, pp. 170–175.

The results and opinions presented in this paper are strictly those of the authors.

Publication of this paper sponsored by Committee on Application of Emerging Technology.