TRANSPORTATION RESEARCH

RECORD

No. 1364

Planning and Administration

Transportation Planning, Programming, Land Use, and Applications of Geographic Information Systems

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1992

Transportation Research Record 1364
Price: $35.00

Subscriber Category
IA planning and administration

TRB Publications Staff
Director of Reports and Editorial Services: Nancy A. Ackerman
Senior Editor: Naomi C. Kassabian
Associate Editor: Alison G. Tobias
Assistant Editors: Luanne Crayton, Norman Solomon, Susan E. G. Brown
Graphics Coordinator: Terri Wayne
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.

Transportation planning, programming, land use, and applications of geographic information systems.
p. cm.—(Transportation research record; ISSN 0361-1981; no. 1364)
Reports prepared for the 71st annual meeting of the Transportation Research Board.
ISBN 0-309-05403-6
1. Urban transportation—Planning. 2. Transportation—Management. 3. Geographic information systems. 4. Pavements—Live loads. I. National Research Council (U.S.). Transportation Research Board. II. Series: Transportation research record; 1364.
TE7.H5 no. 1364
[TA1205]
388 s—dc20 92-35094
[388.3′12] CIP

Sponsorship of Transportation Research Record 1364

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Chairman: Sally Hill Cooper, Virginia Department of Transportation

Management and Administration Section
Chairman: Thomas F. Humphrey, MIT Center for Transportation Studies

Committee on Intergovernmental Relations and Policy Processes
Chairman: Sheldon M. Edner, Portland State University
Stephen E. Blake, Joseph F. Coughlin, Joel P. Ettinger, Albert A. Grant, Randall Halvorson, Barna Juhasz, Anthony R. Kane, Hal S. Maggied, Paul F. Maxwell, Bruce D. McDowell, Patricia L. Mokhtarian, Paul L. Posner, Peter L. Shaw, Reginald R. Souleyrette, Richard K. Taube, Mary Lynn Tischer, Edward Weiner, Robert M. Winick

Multimodal Freight Transportation Section
Chairman: Douglas P. Smith, International Strategies CN Rail

Committee on Urban Goods Movement
Chairman: Arun Chatterjee, University of Tennessee
Secretary: Frank Spielberg, SG Associates, Inc.
Jerri Brown, Carol Carter, Dennis L. Christiansen, Siro De Gasperis, S. Earl Dove, Peggy L. Drake, Gordon P. Fisher, Marilyn S. Gay, Lance R. Grenzeback, George C. Harper, Herbert S. Levinson, William Marconi, Kenneth W. Ogden, Marion R. Poole, John P. Reilly, Noreen Roberts, Richard A. Staley, Richard A. Stehr, Carol H. Walters, C. Michael Walton, William M. Wood

Transportation Forecasting, Data, and Economics Section
Chairman: Mary Lynn Tischer, Virginia Department of Transportation

Committee on Transportation Data and Information Systems
Chairman: Arthur B. Sosslau, Comsis Corporation
Secretary: James J. McDonnell, Federal Highway Administration, U.S. Department of Transportation
Suzanne D. Cook, Robert M. Donnelly, Kenneth J. Dueker, Stanely Gee, Brad H. Johnson, Martha M. Johnson, Philip D. Patterson, Alan E. Pisarski, Charles L. Purvis, Phillip A. Salopek, Eddie Shafie, Gordon A. Shunk, Howard J. Simkowitz, Richard P. Steinmann, Ronald W. Tweedie, Marcus Ramsay Wigan, George F. Wiggers

Committee on Transportation Planning Applications
Chairman: Gordon A. Shunk, Texas Transportation Institute
Daniel Brand, Robert Cuellar, Rick Donnelly, Frederick W. Ducca, Jerry Faris, David T. Hartgen, Ron Jensen-Fisher, George T. Lathrop, Eugene J. Lessieu, Richard S. Marshment, George L. Reed, G. Scott Rutherford, Arthur B. Sosslau, Montie G. Wade, W. T. Watterson, Edward Weiner, Debra Ann Whitmore, George V. Wickstrom

Transportation Systems Planning Section
Chairman: Bruce D. McDowell, U.S. Advisory Commission on Intergovernmental Relations

Committee on Statewide Multimodal Transportation Planning
Chairman: Michael D. Meyer, Georgia Institute of Technology
Secretary: Philip I. Hazen, Federal Highway Administration, U.S. Department of Transportation
David Preston Albright, Linda Bohlinger, Kirk P. Brown, James L. Covil, Darryl E. Durgin, John S. Hassell, Jr., David J. Hawes, Max L. Heckman, Carl A. Hennum, Thomas F. Humphrey, Robert P. Jurasin, Lance A. Neumann, Neil J. Pedersen, Catherine L. Ross, Shashi Kumar Sathisan, Roger L. Schrantz, James P. Toohey, Richard A. Torbik

Committee on Transportation and Land Development
Chairman: Elizabeth Deakin, Deakin Harvey Skabardonis Inc.
Raymond J. Burby, G. Bruce Douglas III, Frederick W. Ducca, Robert T. Dunphy, Z. Andrew Farkas, Ralph Gakenheimer, Kathy Gerwig, Jonathan L. Gifford, Gideon Hashimshony, Mary R. Kihl, R. Ian Kingham, George T. Lathrop, Roger Laurance Mackett, Norbert Oppenheim, Jeffrey A. Parker, Poulicos, Prastacos, Richard H. Pratt, Noreen Roberts, Darwin G. Stuart, W. T. Watterson, Debra Ann Whitmore

Committee on Transportation Programming, Planning, and Systems Evaluation
Chairman: Lance A. Neumann, Cambridge Systematics Inc.
Secretary: Dale A. Janik, Illinois Department of Transportation
Abdullah Al-Mogbel, Colin H. Alter, Jon A. Bloom, Priscilla S. Cornelio, James L. Covil, Joel P. Ettinger, William H. Goldstein, Robert A. Gorman, Peter L. Hathaway, Thomas F. Humphrey, Steen Leleur, Cletus R. Mercier, Alex E. Metcalf, Susan Mortel, Norman G. Paulhus, Jr., Kant Rao, Harry Reed, Kumares C. Sinha, Darwin G. Stuart, Antti Talvitie

GROUP 2—DESIGN AND CONSTRUCTION OF TRANSPORTATION FACILITIES
Chairman: Charles T. Edson, New Jersey Department of Transportation

Pavement Management Section
Chairman: Joe P. Mahoney, University of Washington

Committee on Vehicle Counting, Classification, and Weigh-in-Motion Systems
Chairman: Perry M. Kent, Federal Highway Administration, U.S. Department of Transportation
Harold R. Bosch, Bruce A. Broberg, J. Y. Campbell, Wiley D. Cunagin, Peter Davies, Edward K. Green, H. K. Gupta, Barbara Mason Haines, Mark E. Hallenbeck, Lawrence E. Hart, Loyd R. Henion, David L. Huft, Bernard Jacob, Frederick R. Juba, Clyde E. Lee, D. Keith Maki, Bill M. McCall, William A. Mickler, John T. Neukam, Thomas Papagiannakis, T. V. Ramakrishna, Peter Sebaaly, Herbert F. Southgate, Jon P. Underwood, John Van Berkel, Jr., Alex T. Visser

James A. Scott, Elaine King, and Frank R. McCullagh, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1991.

Transportation Research Record 1364

Contents

Foreword

The papers in this Record can be classified in five areas: applications of highway performance monitoring systems (HPMSs); metropolitan transportation planning and systems planning applications; application of geographic information systems (GISs); and weigh-in-motion systems.

Papers on HPMS discuss various aspects of the applications to classifying highway needs, adequacy ratings, and bridge needs and investments. The papers on metropolitan planning focus on the organizational and institutional aspects of regional and metropolitan planning and transportation planning applications.

The papers on transportation system applications discuss transportation network design and levels of service in an era of growth management; factors affecting automobile ownership and use and the emergence of the suburban community pattern; location planning to reduce automobile usage; vehicle trip generation rates for a reconstruction project; and planning under uncertainty.

Four papers discuss applications of GISs for transportation at the county and municipal levels and for superregional transportation modeling and transit service areas. Four papers discuss various aspects of weigh-in-motion and vehicle classification systems for monitoring and calibration; data needed for transportation planning; and automatic vehicle classification systems.

Use of Highway Performance Monitoring System in Reclassifying Rural Highways in Support of National Highway System in Kansas

E. D. LANDMAN

Kansas ranks fourth nationally in total miles of streets and highways. Maintaining such a large system is a burden to all levels of government. There are similarities between Kansas and surrounding states, particularly Iowa and Nebraska. This suggests that the conventional functional classification was done properly. It is evident that national averages do not provide appropriate guidelines for attaining a classification that produces an optimal decision making basis for states such as Kansas. The Highway Performance Monitoring System model proved to be a useful tool for demonstrating the benefits of using different criteria to classify Kansas's highways. An estimate of total needs to attain an assumed level of service was not particularly useful. However, a comparison of the higher user costs associated with the existing functional classification system—which attempts to bring all arterials, no matter how minor, up to arterial standards—to the costs associated with the State Transportation Planning classification system, proved convincing to policy makers.

An explanation is provided of the problem Kansas has had with the existing functional classification and of the measures taken to remedy the problem. Even though the mileage in each class fell within acceptable ranges, the large number of miles of rural local roads inherent to the geography of the state contributed to the large number of miles of arterials. Estimates of necessary costs to achieve arterial standards on many miles of arterials in a relatively low density area were unrealistic. Even though the traffic volumes were typically far below capacity, the minimum arterial standards dictated reconstruction and replacement rather than rehabilitation and resurfacing.

The analysis was conducted using FHWA's Highway Performance Monitoring System (HPMS) to support the reclassification by showing how vehicle operation costs were reduced from the escalated levels in the existing classification. By effectively reducing the mileage classified as arterials, more miles of roadway could be improved and better service on the overall system could be achieved with available funding. The smaller system of principal arterials became the basis for the state's recommendation for the National Highway System (NHS) for Kansas.

Reported here are analyses made during a 6-year period. During that time, there were changes in governors, legislative committees, secretaries of transportation, state transportation engineers, and directors of planning and development. Data used in the analysis also changed. The most up-to-date data available were used at each step as the effort progressed. As an example, approximately 50 mi of the Kansas Turnpike were added to the Interstate system after the 65-mph speed limit was allowed on rural Interstates.

BACKGROUND

The extensive network of transportation facilities in Kansas today is a product of the historical development of the state. Rapid growth in population occurred as a result of the Homestead Act of 1868. Most of the tillable land was homesteaded within the next 20 years with 160-acre (quarter-section) farms. Section-line roads were built throughout the state. During the next decades, as the many farms that resulted from the Homestead Act were incorporated into larger farms, the state's extensive farm population began moving to the cities and even leaving Kansas.

Since all the roads were unsurfaced and the vehicles were horse-drawn, travel beyond the nearest town was very limited. Accessibility to distant markets was provided by rail. Consequently, an extensive rail system developed, along with numerous towns that served both as suppliers for the surrounding community and markets for the crops produced. The rail system served as arterials, and the better roads were, at best, collectors.

Population Trends in Kansas

An analysis of population characteristics and trends is basic to any type of understanding and planning of the state's transportation system. Many counties have lost much of the population since they peaked. Figure 1 shows the losses experienced and the year each county reached its maximum. Correlation between loss of population and the number of farms can be observed by comparing the year of maximum population to the year the maximum number of farms was attained. Although the number of farms has declined, the area farmed has been fairly constant, except for urbanization, resulting in much larger farms and farm equipment.

Division of Planning and Development, Kansas Department of Transportation, Docking State Office Building, 8th Floor, Topeka, Kans. 66612.

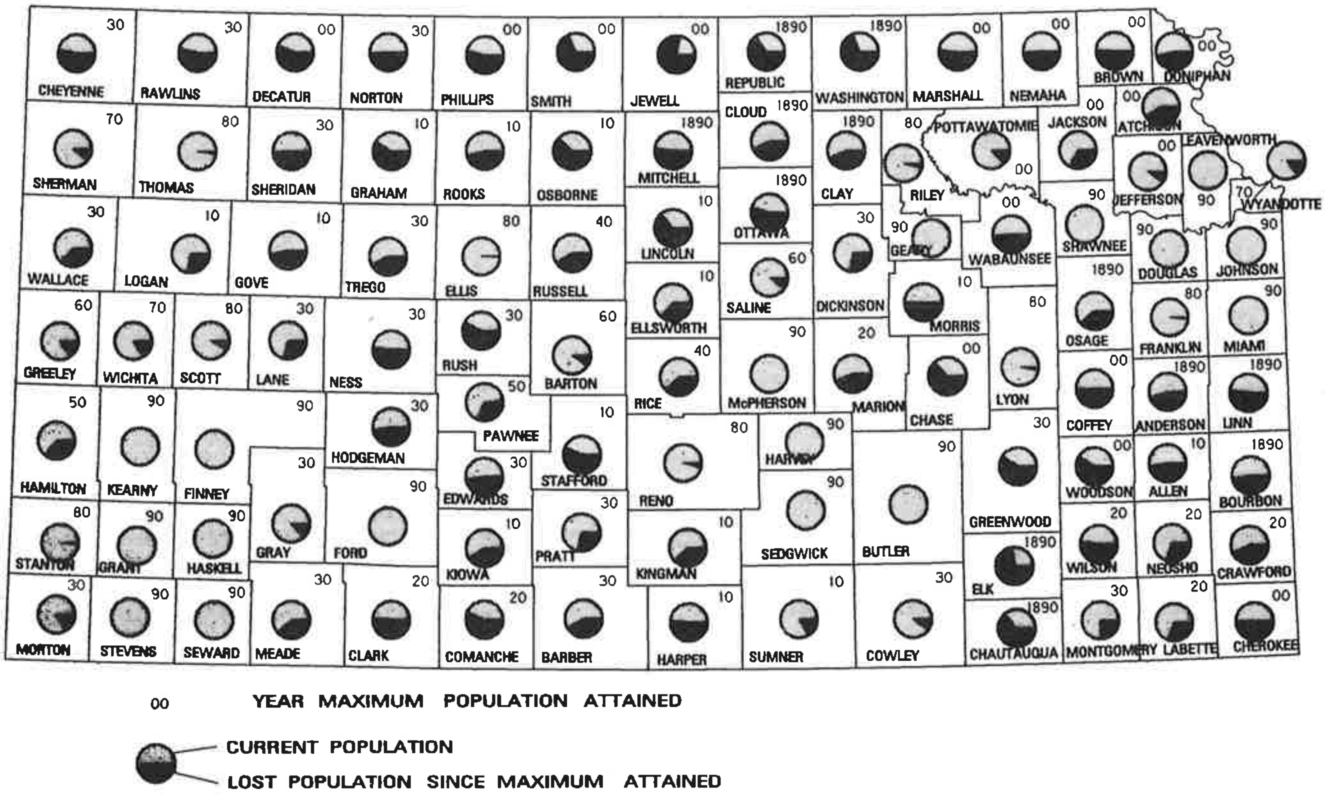


FIGURE 1 Population trends retained and lost.

Terrain

The relatively flat terrain of Kansas contributed to extensive homesteading and to a uniform grid pattern of roads. One road is about as good as any other when so many choices are available on the grid. Another phenomenon of the grid pattern is the large number of bridges. The uniform pattern crosses streams regularly at the point of intersection.

Summary of Existing Classification

Table 1 presents the percentage of miles and vehicle miles by existing federal functional class for all rural roads on the state highway system. The state highway system includes the entire principal and minor arterial system and about 7 percent of the rural major collector system. Only 22 of 9,322 mi (0.2 percent) are minor collectors. No rural local roads are included in the state highway system.

Comparison of Kansas with Surrounding States and Nation

A comparison was made with surrounding states to determine whether the classification was consistent and whether the Great Plains states shared some characteristics. It was recognized at the outset that a state highway system is a political system that varies from state to state.

One of the most important elements facing a transportation agency is that of finances. Funding must be provided to maintain and upgrade the systems for which the agency is responsible. Traditionally, funding has come from fuel tax and vehicle registration. Travel and vehicles owned are both indicators of resources that generate the revenue base that supports the operations of the agency.

The first comparison made was average daily traffic (ADT) on all rural federal-aid highway systems for Kansas and the surrounding states, including the national average. Figure 2 shows that the national average is more than 2.5 times that

TABLE 1 Functional Classification Summary, Rural Mileage and Jurisdiction

CLASS	TOTAL MILES	PER CENT	ACCUM %	ST.SYS MILES	PER CENT	ACCUM %	% OF CLS TOT
I-ST	654	0.5	0.5	654	6.5	6.5	100.0
OPA	3307	2.7	3.2	3307	32.8	39.3	100.0
MA	4505	3.6	6.8	4505	44.7	84.0	100.0
MjC	22628	18.3	25.1	1604	15.9	99.9	7.1
MnC	9329	7.5	32.7	7	0.1	100.0	0.1
Loc	83351	67.3	100.0	0	0.0	100.0	0.0
Total	123774	100.0		10077	100.0		8.1

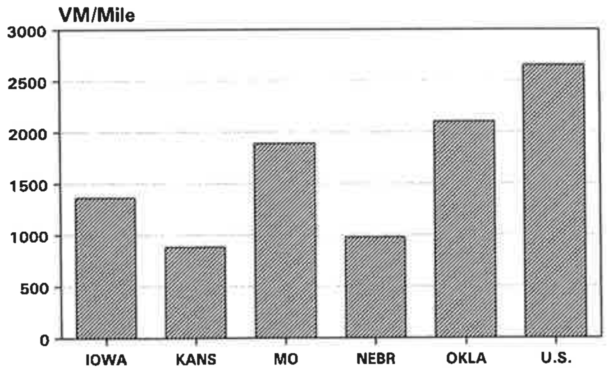


FIGURE 2 Systemwide ADT: rural only, all federal-aid systems.

of either Kansas or Nebraska. This would indicate that it would be more difficult for Kansas, both the state and county governments, to generate enough revenue from highway users to construct federal-aid highways to a national standard.

If only the primary system (not including Interstate) is considered, the conclusion is the same. Figure 3 shows the vehicle miles of travel (VMT) for the rural primary system. On this graph, Iowa falls in the same general range as Kansas and Nebraska; Missouri approaches the national average.

A similar comparison was made in Figure 4 for registered vehicles per mile of road. In Kansas, the vehicle registration fees are dedicated to the state highways only. Again, Iowa and Nebraska were very close to Kansas. As in the two previous graphs showing vehicle miles, both Missouri and Oklahoma had a substantially better revenue base than the three other states, but they were significantly below the national average.

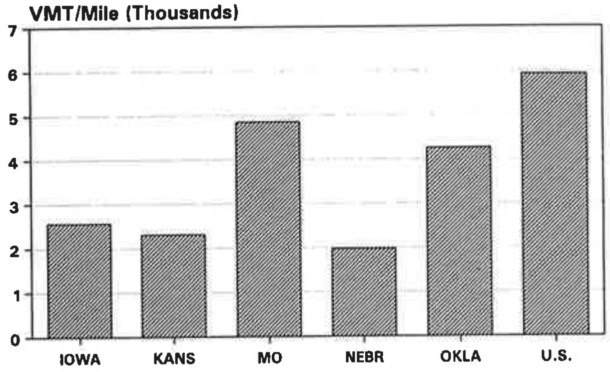


FIGURE 3 Systemwide ADT: primary system, vehicle miles per mile of primary.

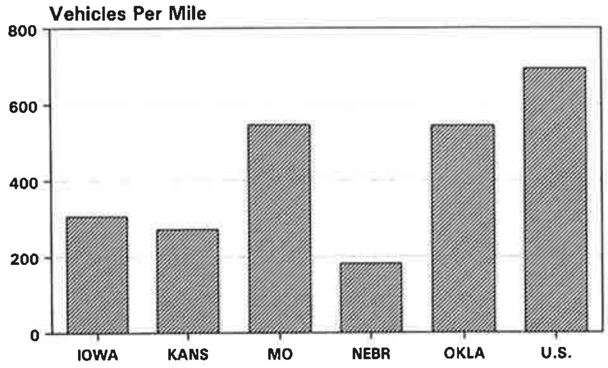


FIGURE 4 State-by-state comparison, registered vehicles per mile of road.

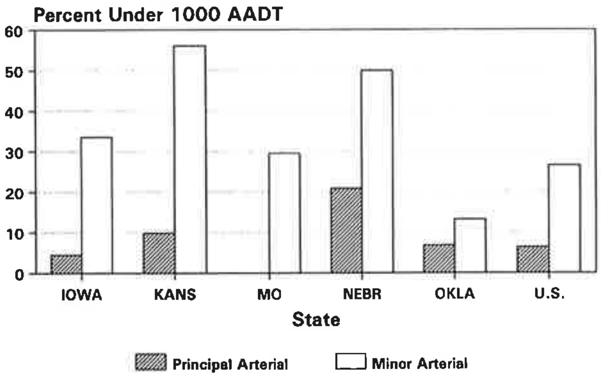


FIGURE 5 State-by-state comparison, AADT under 1,000.

An additional comparison was made to pinpoint the problem more closely. Figure 5 displays the percentage of miles of principal and minor arterials that carry fewer than 1,000 vehicles a day. The value of 1,000 vehicles a day is an arbitrary round number, but it represents the lower traffic volume that will generate enough revenue to support normal maintenance. This chart vividly indicates that Kansas would have to deal with the large number of minor arterials. More than 50 percent of the minor arterials carry fewer than 1,000 vehicles a day; the national average is less than 25 percent.

The charts provided enough evidence to support the use of different criteria for selecting the highways in Kansas that would receive the greatest attention for upgrading and modernization. The adoption of the 1984 AASHTO Green Book lent an even greater urgency to the task, because the Green Book does not distinguish between principal and minor arterials. Therefore, any geometric improvement made to low-volume minor arterials must be made to full arterial standards (this decision was made before Kansas adopted restoration, rehabilitation, and reconstruction standards).

DESCRIPTION OF A–E CLASSIFICATION

The following is a summary of criteria used in the State Transportation Plan (STP) to stratify the routes.

Class A

Class A routes are those routes designated as the Interstate system.

Class B

Class B routes, along with Class A routes, serve the most important corridors of statewide and interstate travel. The trips on the facilities are very long and carry many out-of-state vehicles. Traffic volumes remain relatively constant along major segments of the route. The corridors are spaced widely enough to serve distinctly different trip movements.

Class C

Class C is a part of the statewide arterial system and is integrated with Class A and B routes to provide service to all areas of the state. Major cities not on Class A and B routes are connected by Class C routes. Continuity and long average trip lengths are important, but greater attention is given to coverage of areas not served by other routes.

Class D

Class D contains routes serving a combined role of intercounty movement and access to the arterial routes for county seats and other small urban areas not on a Class A, B, or C route. Through truck traffic should be discouraged, but there may be a large number of trucks serving local industries.

Class E

Class E is made up of stubs and routes whose service is limited almost exclusively to local service. Few nonlocal license plates are observed unless the route serves a well-known park or tourist attraction. The trip lengths are very short: the same people often travel over the same roads several times a day, day after day.

Defining the System

The rural state highway system is made up of facilities displaying a wide variety of characteristics. One of the very important factors to consider in planning a system that will efficiently carry statewide travel is that of route continuity. Corridors should provide a fairly consistent level of service throughout their length so as not to surprise the traveler with unexpected design changes.

As a general rule, the routes carrying the most traffic are those serving the long intercity and interstate trips. However, the classification cannot be based on traffic volume alone. A route with short average trip lengths may also carry high volumes of traffic. Examples of these are commuter routes radiating from the metropolitan areas. Traffic volumes may be some of the highest rural volumes, but the traffic on these routes drops off rapidly beyond the commuting range for the area. Routes serving primarily local trips are not affected by the condition of segments elsewhere along the route. A good test for determining if a corridor serves many very long trips occurs when roads are closed during snowstorms. The function of a route, such as an Interstate system route, can be affected by actions or conditions several hundred miles away in other states.

The reclassification was done from the top down to ensure continuity and integration within each class and with those previously listed. To minimize confusion with FHWA functional classifications and with preconceived notions about what certain words mean, the most generic terms were used. Five classes were defined, with Class A being comparable to the Interstate system and Class E being stubs and other local service routes that should probably not be on the state highway system. County roads and city streets were not included in the reclassification.

A wide variety of tools and resources was used to stratify the rural state system into classes. They include

- Traffic counts,
- Classification counts (breakdown of trucks by type),
- Truck weight surveys,
- Roadside origin-destination surveys,
- Census data,
- Plans of adjacent states,
- Statewide traffic assignment models,
- Highway data base,
- Plans for urban areas,
- Revenue forecasts,
- Uniform design standards, and
- State and federal statutes.

MILEAGE AND TRAVEL SUMMARIES

Table 1 presents summaries of miles, vehicle miles, and percentages for total and heavy-truck traffic for each of the five classes. The Class A routes make up only 7.4 percent of the state's rural mileage but carry a fourth of the total rural travel and a third of all rural heavy-truck travel. Classes A through C encompass half of the mileage and carry 75 percent of total travel and more than 80 percent of heavy-truck travel. This is particularly significant because the remaining half of the system does not generate enough revenue from usage to provide adequate maintenance, let alone to provide for limited modernization.

Figure 6 shows graphically the accumulative values of Table 2. The shape of the curve is typical of those found in functional classification manuals. The lower right-hand curve depicts the percentage of accumulative mileage and total vehicle miles for each class, and the upper left-hand curve depicts the percentage mileage and heavy-truck miles for each class. Although more than half (54.7 percent) of total travel occurs

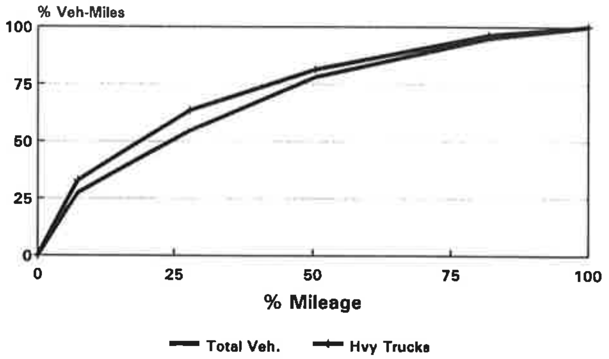


FIGURE 6 STP classification summary of mileage and travel, rural systems only.

TABLE 2 STP A–E Classification Summary, Rural Mileage and Travel

CLASS	MILES	PER CENT	ACCUM %	TOTAL VEH-MI	PER CENT	ACCUM %	TOTAL AVE.*	TRUCK VEH-MI	PER CENT	ACCUM %	TRUCK AVE.#
A	725	7.4	7.4	5480	27.3	27.3	7560	1080	32.8	32.8	1490
B	2021	20.5	27.9	5522	27.4	54.7	2730	1020	30.8	63.6	500
C	2217	22.5	50.4	4699	23.4	78.1	2120	590	17.8	81.4	270
D	3134	31.6	82.0	3428	17.0	95.1	1090	490	15.0	96.4	160
E	1771	18.0	100.0	994	4.9	100.0	560	120	3.6	100.0	70
	9848	100.0		20123	100.0		2040	3300	100.0		330

Vehicle Miles in 1000's per day
1987 Mileage and Travel
* Average Vehicle-Miles per Mile - All vehicles
# Average Vehicle-Miles per Mile - Heavy Trucks

on Class A and B routes, these routes carry 63.6 percent of the heavy-truck travel. The steeper slope of the heavy-truck curve reflects a higher percentage of heavy trucks on a small portion of the system.

A comparison of heavy-truck travel on the Class B and C routes demonstrates one important difference between the two classes: whereas the total traffic is only slightly greater on the Class B routes than on the Class C routes, the heavy-truck traffic on Class B routes is almost twice that on Class C routes.

The STP classification and associated standards were adopted by the agency in 1988. The priority formula used to select candidate projects for the Comprehensive Highway Program was also changed to recognize the A–E standards. A weighting factor incorporated into the formula uses the classification to assign a weight to each section of road.

Obtaining STP Classification Data from Existing HPMS Submittal

The HPMS analytical model was used to compare the existing federal functional classification with the STP classification. Comparisons were made for the agency and for the user. Unmet highway needs were estimated if the available funds were used for projects using the existing federal classification and if the classes in the HPMS model were redefined to the STP standards as

	Class
Interstate (I-St)	A
Other principal arterials (OPA)	B
Minor arterials (MnA)	C
Major collectors (MjC)	D
Minor collectors (MnC)	E

Preparation of Data

Factoring by Classification and Volume Group

The HPMS analytical model uses data compiled from sample sections collected for submittal to FHWA. Because the data are sampled by functional class and factored to the respective classes' total mileage, the sample sections must be refactored to the new STP class totals. The reclassification is complicated slightly because each class is further stratified by volume group. The matrix in Table 3 indicates which functional class and volume groups were involved.

Statistical Tests

Statistical reliability tests were originally used to establish the number of samples that were needed within each class and volume group. The significance of the Interstate and other higher type facilities resulted in a higher confidence level; consequently, a higher sample rate was chosen for these facilities. Fortunately, the reclassification amounted to a downgrading of many of the routes from minor arterials to major collectors. Therefore, it was possible to maintain the necessary statistical reliability without having to collect any additional field data.

Comparison of STP Classification to Existing Classification

As is shown in Figure 7, there has been a shift to the right in the height of the bars from the conventional functional classification to the STP classification. With a few exceptions, all of the minor collectors on the state system were reclassified as Class E routes. As will be shown later, the greatest impact on the system resulted from a large number of minor arterials' becoming Class D routes. Again it must be remembered that this reclassification, to date, has not involved nearly 30,000 mi of nonlocal rural roads that are the responsibility of counties and townships, nor does it include connections of highways through cities and other city streets.

Benefits of STP Classification

The benefit of the reclassification can be seen from the results of runs made using the HPMS analytical model. The model was run using only the rural sections, both using the conventional functional classification and the refactored STP class totals. As can be seen in Figure 8, the 10-year roadway needs have been reduced by more than three times from a previously run in-house needs analysis (HI SCOPE), using the STP classification. The in-house (LO SCOPE) was also nearly twice that of the STP classification. The HI SCOPE and LO SCOPE needs estimates were made using a needs model developed within the agency, and the STP SCOPE was obtained from the HPMS model using the investment-level option. This comparison by itself means little since this is the obvious result of converting a big part of the system to a lower classification with lower standards.

TABLE 3 Expanding HPMS Samples to STP Mileage Totals

		Functional Classification					
		I-St	OPA	Min A	Maj C	Min C	Local
STP Class	Volume Group	Miles	Miles	Miles	Miles	Miles	Miles
A	1	X	X				
	2	X					
	3						
	4						
	5						
B	1		X	X			
	2		X	X			
	3		X				
	4		X				
	5		X				
C	1		X	X			
	2		X	X			
	3		X	X			
	4		X				
	5						
D	1		X	X			
	2		X	X	X		
	3			X			
	4			X			
	5						
E	1		X	X	X	X	
	2			X	X	X	
	3			X	X		
	4						
	5						

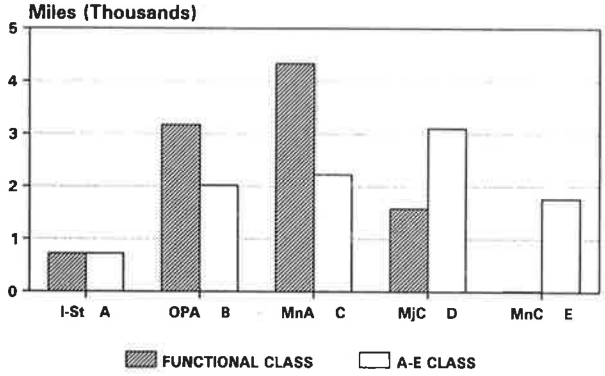


FIGURE 7 Comparison of STP class changes.

A more important comparison is shown in Figure 9. This chart shows the road user cost from the HPMS model at 5-year intervals for both classifications using the fund period analysis option. The funding available in 1987 to finance the 5-year highway construction program was used for each analysis made with the model. The funding was assumed to be level in constant dollars over the remaining three fund periods. The chart shows that user costs more than double over the 20-year period for the conventional functional classification. Part of this increase can be attributed to increased travel over that period. However, during this same period, and with the same funding level, the road user costs increased only slightly when the funds were allocated and used according to the STP classification and standards.

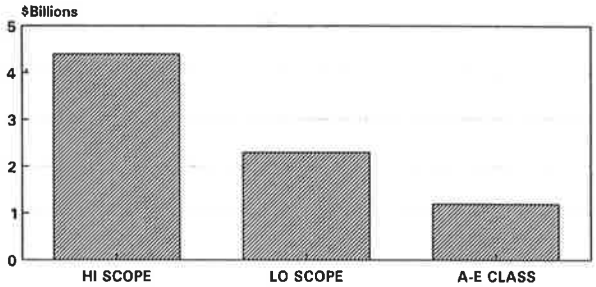


FIGURE 8 Comparison of needs, rural state system (roadway only).

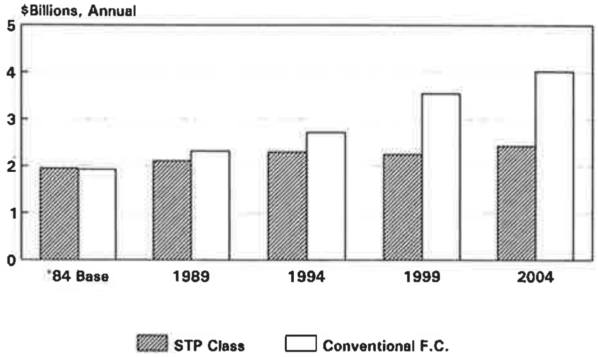


FIGURE 9 Comparison of user costs, rural state system.

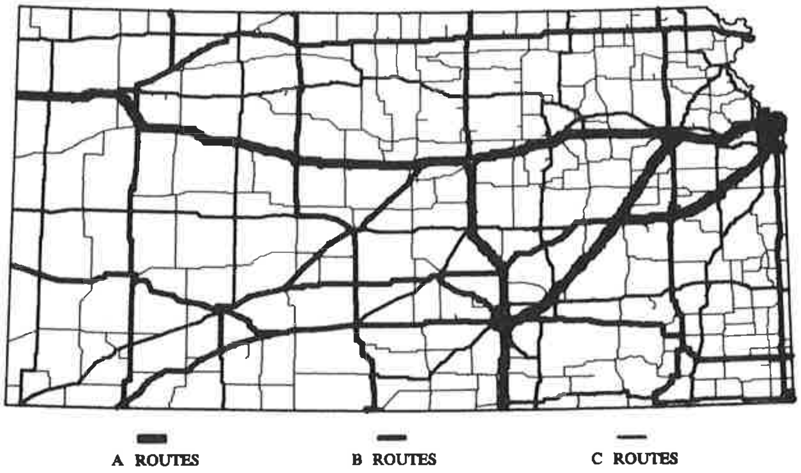


FIGURE 10 Comparison between STP classifications, Class A, B, and C routes.

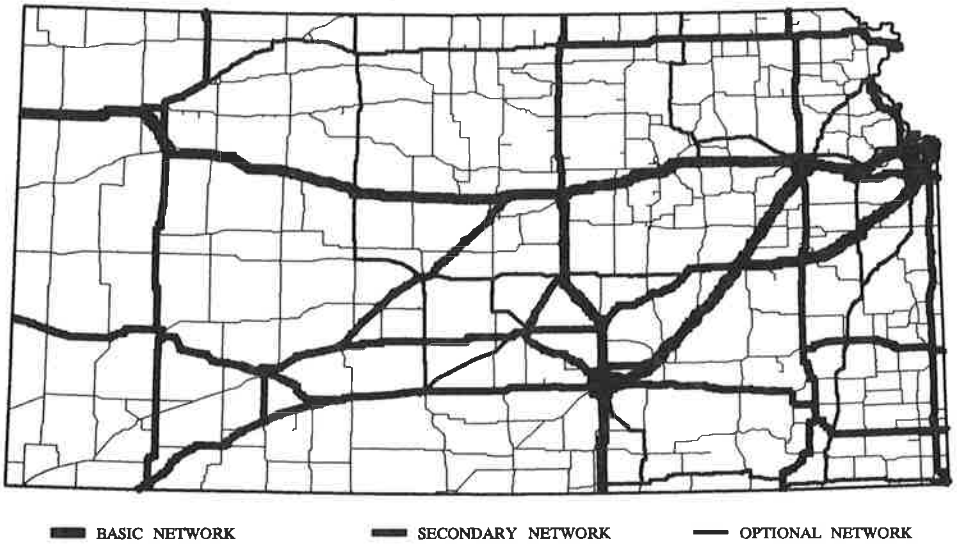


FIGURE 11 Proposed NHS basic and secondary networks.

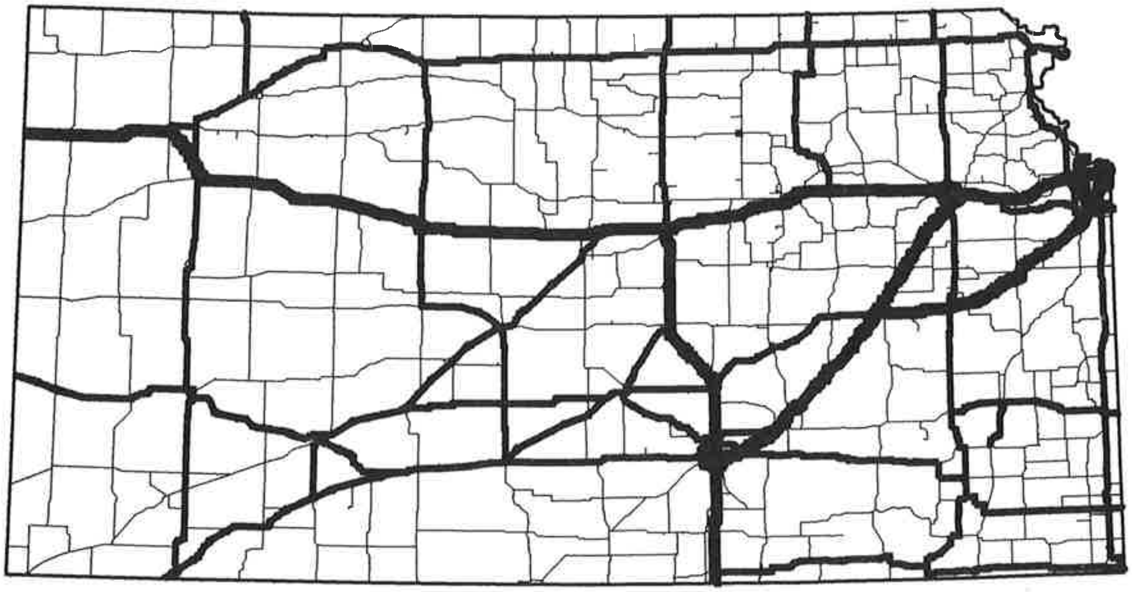


FIGURE 12 Illustrative NHS network prepared by FHWA.

National Highway System Submittal

At the request of FHWA in September 1990, states submitted maps showing highways that should be considered as candidates for the NHS. The NHS was a key element in FHWA's proposed federal reauthorization legislation. Shown in the figures are the comparison between the STP-classified Class A, B, and C routes (Figure 10), the proposed NHS basic and secondary network (Figure 11), and the illustrative NHS network prepared by FHWA (Figure 12).

SUMMARY AND CONCLUSIONS

Kansas ranks fourth nationally in total miles of streets and highways. Maintaining such a large system is a burden to all levels of government. Kansas and surrounding states—particularly Iowa and Nebraska—have some similarities, suggesting that the conventional functional classification was done properly. It is evident that national averages do not provide appropriate guidelines for attaining a classification that produces an optimal decision-making basis for states such as Kansas.

The HPMS model proved to be a useful tool for demonstrating the benefits of using different criteria to classify Kansas highways. An estimate of total needs to attain an assumed level of service was not particularly useful. However, a comparison of the higher user costs associated with the existing functional classification system (which attempts to bring all arterials, no matter how minor, up to arterial standards) to the costs associated with the STP classification system proved convincing to policy makers.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Highway Performance Monitoring System Analytical Process Application to Kentucky's Adequacy Program

A. M. Taqui

The usefulness and applicability of the Highway Performance Monitoring System (HPMS) analytical process to Kentucky's Adequacy Rating program were examined. Noncompatibility of final ratings of highway sections from the Adequacy Rating and the HPMS program was a major concern resulting in the loss of credibility of either program for use in planning and programming. Therefore, the ratings from the HPMS and Adequacy programs were compared using highway sections from the HPMS and Adequacy data files to determine the similarities in methodologies in order to select one common tool for systems- and project-level analysis. Appraisal rates and weights for the different components of rating from the HPMS software were changed in relation to design standards and technical project selection criteria. Four scenarios with different weights and critical accident rate as a separate data element were chosen. Ratings for 21 HPMS samples from each scenario and the Adequacy program were reviewed by an expert panel to select the most appropriate scenario for use by the Kentucky Department of Highways as the standard procedure for the HPMS and Adequacy Rating program. The expert panel recommended using Scenario 1 without the critical accident rate as the standard procedure.

Highway Performance Monitoring System (HPMS) data contain current and accurate information about the geometric and operational characteristics of the highway systems sampled at random to conduct meaningful system analysis with known precision. FHWA uses the national HPMS data for various purposes: for preparing a biennial report to Congress, apportioning I-4R, and determining highway condition, performance, and needs. The use of HPMS data by FHWA to furnish pertinent information to Congress about the nation's highway condition, performance, and needs with different scenarios was very similar to some of the work performed by the Kentucky Department of Highways in providing necessary data to the Kentucky General Assembly. The process of assessing the condition and performance of the system relative to available or expected funding convinced Kentucky Transportation Cabinet management personnel about its potential and usefulness to the cabinet. Therefore, the HPMS program in Kentucky received necessary support for the implementation and expansion to include other systems (state primary and federal aid). The HPMS program was reviewed in detail by the U.S. General Accounting Office (GAO) and found to be a reliable tool for monitoring highway performance using nationally accepted engineering standards.

The HPMS analytical package encompasses most of the geometric and operational data elements substratified and assigned with relative weight, depending on the functional ability of the system. The software computes the overall composite index using the main data elements of condition, safety, and service that have been stratified and assigned numerical weights. In addition, roadway sections are ranked according to existing and future deficiencies with reference to the composite index.

OBJECTIVE

State highway agencies have used adequacy ratings for a long time to help priority rank highway projects. The first ratings were obtained in Kentucky in 1949 and were used for planning a 5-year construction program. Since then updates have been infrequent. A field procedural manual was prepared in 1963, and survey ratings were provided by the highway districts thereafter. Documented revisions to Kentucky's Adequacy Rating procedures were compiled as a research report in 1976. The field inventory form and manual were revised in 1979 for conformance with the HPMS data reporting requirements.

"Adequacy" is defined as being sufficient for a specific requirement or standard. Adequacy ratings are used to provide a systematic procedure for periodically rating highway sections or projects for improvement programming. Using the adequacy rating technique, highway sections or bridges are assigned numerical ratings that indicate their relationship to established design standards. Kentucky's rating procedure includes the elements of condition, safety, service, and operation. Several subelements make up the main elements, and relative weights are represented by maximum points for a total of 100.

The compatibility of Adequacy Ratings with the HPMS ratings became a concern in Kentucky because the system's analysis and needs were determined with HPMS software and because individual project ratings were obtained from the Adequacy Rating program for prioritization. The department's management personnel believed that it would be highly useful and appropriate to use common criteria for both the HPMS and the Adequacy Rating programs. One common tool with uniform criteria further eliminates the possibility of any disagreement between the composite indexes of any roadway segment computed from HPMS and Adequacy Rating. There was also some uncertainty about the appropriateness

---

Division of Planning, Transportation Cabinet, Commonwealth of Kentucky, Frankfort, Ky., 40622.

of the data items and their corresponding weights in the HPMS software for use in Kentucky. Elimination or minimization of personal judgment data items in the Adequacy Rating was considered to be essential in obtaining consistent or uniform ratings (indexes).

The objective of this analysis was to evaluate the HPMS data elements and weighting factors and compare them with the data elements and weighting factors of the Kentucky Adequacy Rating program and to decide whether the department could use HPMS software with appropriate weights for the various data elements as the standard tool for Adequacy Rating in Kentucky. The basic aim was to investigate the feasibility of using one tool for determining performance, analyzing needs, and prioritizing projects.

PROCEDURE

The various attributes used in analyzing a roadway section by the Adequacy Rating and HPMS programs by rural/urban with relative weights are shown in Tables 1 through 4. The Adequacy Rating program contains attributes such as surface condition, base condition, maintenance economy, drainage adequacy, and traffic control devices that are subjectively rated by highway district personnel and are prone to error and inconsistency. The program logic that computes composite index was complex, making desired adjustments difficult. Assignment of uniform weights to the main components of condition, safety, service, and operation, regardless of functional system, was viewed as inappropriate. By contrast,

TABLE 1 Adequacy Rating Elements for Rural Highways

Condition Elements (35 points)	Maximum Points
Maintenance Economy	7
Surface Condition	10
Base Condition	10
Drainage	8
Safety Elements (25 points)	
Accident Experience	5
Vertical Alignment	4
Horizontal Alignment	4
Stop Sight Distance	7
Skid Resistance	5
Service Elements (30 points)	
Ride Quality	5
Pavement Width	10
Shoulder Width	5
Shoulder Type	2
Passing Sight Distance	8
Operational Elements (10 points)	
V/C Ratio	10
Total Maximum Points	100

TABLE 2 Adequacy Rating Elements for Urban Highways

Condition Elements (35 points)	Maximum Points
Maintenance Economy	7
Surface Condition	10
Base Condition	10
Drainage	8
Safety Elements (25 points)	
Accident Experience	10
Traffic Control Devices	
Standardization	5
Effectiveness	5
Maintenance	5
Service Elements (25 points)	
Pavement Width	15
Operating Speed	10
Operational Elements (15 points)	
V/C Ratio	15
Total Maximum Points	100

TABLE 3 HPMS Rural Composite Index Weights Default Values

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Condition		Interstate	Principal Arterials		Minor Arterials		Collectors
		(Total=40)	(Total=40)		(Total=45)		(Total=50)
Pavement Type	1	010	010		015		015
Pavement Condition	2	025	025		025		030
Drainage Adequacy	3	005	005		005		005
Safety		Interstate	Principal Arterials	Principal Arterials	Minor Arterials	Minor Arterials	Collectors
			< 3 Lanes	> 3 Lanes	< 3 Lanes	> 3 Lanes	
		(Total=30)	(Total=30)	(Total=30)	(Total=30)	(Total=30)	(Total=30)
Lane Width	4	010	015	010	015	010	015
Shoulder Width	5	005	005	005	005	005	005
Median Width	6	005	000	005	000	005	000
Alignment Adequacy	7	010	008	010	010	010	010
Service		Interstate	Principal Arterials		Minor Arterials		Collectors
		(Total=30)	(Total=30)		(Total=25)		(Total=20)
Volume/SF Ratio	8	025	025		025		020
Access Control	9	005	005		000		000
Total Points		100	100		100		100

TABLE 4 HPMS Urban Composite Index Weights Default Values

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5
Condition		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
		(Total=40)	(Total=40)	(Total=40)	(Total=45)	(Total=50)
Pavement Type	1	010	010	010	015	015
Pavement Condition	2	025	025	025	025	025
Drainage Adequacy	3	005	005	005	005	010
Safety		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
		(Total=30)	(Total=30)	(Total=30)	(Total=30)	(Total=30)
Lane Width	4	020	020	020	020	030
Shoulder Width	5	005	005	000	000	000
Median Width	6	005	005	010	010	000
Service		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
		(Total=30)	(Total=30)	(Total=30)	(Total=25)	(Total=20)
Volume/SF Ratio	7	025	025	025	025	020
Access Control	8	005	005	005	000	000
Total Points		100	100	100	100	100

Note: Pavement condition data for all HPMS samples are obtained from Mays Ride Meter

the HPMS program has only one subjectively rated data element (drainage adequacy). The pavement condition data for all roadway sections in Kentucky is obtained by the Mays ride meter. Therefore, no subjectivity is involved in its collection. Weighting factors assignment is in accordance with the functional ability of the system. Availability of technical support from the FHWA, easy-to-use new-generation software, periodic enhancements, and flexible weight/appraisal rates are some of the advantages of the HPMS analytical package.

The Transportation Research Center at the University of Kentucky was asked to analyze the two programs' criteria and weighting factors to determine whether the HPMS software could be adopted by the department as the standard tool for analyzing adequacy, determining systems performance, and prioritizing projects. The technique used by the Transportation Research Center in analyzing the methodology and rating procedure of the HPMS and Adequacy Rating program was to determine, summarize, and compare composite indexes for the same highway sections from the HPMS and Adequacy Rating programs' data files.

From a summary of statewide final indexes in which HPMS and Adequacy Rating sections were compared, it was determined that the higher functional systems (Interstate, principal arterials) closely matched in their assigned ratings. The percentage of total matches decreased as the range of final indexes increased. A summary of the statewide final indexes

TABLE 5 Final Index Comparison of HPMS and Adequacy Ratings

	NUMBER OF HIGHWAY SECTIONS BY FINAL INDEX DIFFERENCE											
	<=2		>2 & <=5		>5 & <=10		>10 & <=15		>15 & <=20		>20	
FUNCTIONAL CLASS	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
RURAL												
Interstate	53	30.6%	41	23.7%	48	27.7%	20	11.5%	5	2.9%	6	3.5%
Principal Art.	75	17.9%	119	28.4%	116	27.7%	62	14.8%	21	5.0%	26	6.2%
Minor Arterial	14	6.7%	32	15.2%	37	17.6%	63	30.0%	35	16.7%	29	13.8%
Major Collector	4	2.3%	8	4.6%	21	12.1%	29	16.7%	24	13.8%	88	50.5%
Minor Collector	9	5.3%	14	8.3%	18	10.6%	32	18.9%	31	18.3%	65	38.4%
URBAN												
Interstate	27	27.8%	31	32.0%	21	21.6%	7	7.2%	3	3.1%	8	8.2%
Freeway	7	33.3%	5	23.8%	5	23.4%	3	14.1%	1	4.7%	0	0.0%
Principal Art.	33	14.4%	22	9.6%	53	23.1%	40	17.5%	33	14.4%	48	20.9%
Minor Arterial	25	9.4%	27	10.2%	46	17.3%	34	12.8%	44	16.5%	90	33.8%
Collector	5	7.4%	6	8.8%	10	14.7%	10	14.7%	6	8.8%	31	45.5%
TOTAL	252		305		375		300		203		391	

for ranges of differences for same highway sections from the HPMS and Adequacy data files using the HPMS and Adequacy program software is presented in Table 5. The results show that the ratings from the two programs match closely for the higher functional classes of road. However, the variability increases substantially for most other classes of road. The reasons for this noticeable and wide variation can be attributed to the age and quality of data on the Adequacy Rating file, influence in consistency of the subjectively rated data items, and the weighting process of the data items. The Adequacy Rating data file has not been routinely updated, particularly for lower functional systems, since 1983 because of other priorities and manpower shortages; the HPMS file, which is smaller and more manageable, is more current. Therefore, it is apparent that, although there is noticeable dissimilarity in the data component of the two programs, the final indexes would have shown good correlation because of similar methodology if the Adequacy Rating data file had been more current. Thus, it is conceivable that the HPMS analytical package can be used to determine the adequacy rating and prioritization of all highway sections instead of the Adequacy Rating software.

To analyze the effect of weighting schemes on the indexes and determine the most appropriate weighting factors for departmental use, an expert panel comprising in-house professional staff was formed and chaired by the director of the planning division. The expert panel met several times to review in detail the methodology, various data components, relative weightings, and appraisal rates. Weightings and appraisal rate changes were recommended by the panel in relation to the design standards and technical project selection criteria. Inclusion of accident rate as a separate sub–data element for emphasizing safety was suggested by the panel. Hence, four scenarios (S1, S2, S3, and S4) were developed; they are presented in Tables 6 through 13. The first three scenarios contain accident rates under safety in addition to other data items that are identical to FHWA's analysis process with recommended changes in weighting and appraisal rates by the panel. The appraisal rates are presented in Tables 14 and 15. Scenario S4 is totally different in emphasis, weighting scheme, and data items grouping for arterial and collectors. No changes were suggested for the Interstate system. Twenty-one sample sections comprising all functional systems and rural/urban were selected from the HPMS data file, and the indexes were computed manually for the four scenarios. Those sections with final composite indexes from each scenario and from the Adequacy Rating file are presented in Table 16.

It is apparent from Table 16 that the range of difference in indexes from Scenario S4 is appreciably higher than the other three scenarios because of overemphasis on some data attributes (lane widths, shoulder width, alignment, adequacy, accident) and the unique weighting scheme. Hence, it was unacceptable. The indexes from Scenarios S1, S2, and S3 show good correlation except for an urban collector and a rural collector section that have significant variation because of lane width emphasis. The indexes from the Adequacy Rating source also show good correlation for higher functional systems despite old data. Inclusion of accident rate under the safety item did not cause noticeable change in the final indexes, implying that some sections with very good safety rating have a high accident frequency that is due to access points, driver behavior, and weather conditions. After careful review of the

TABLE 6 HPMS Rural Composite Index Weights, Scenario 1

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Condition		Interstate (Total=40)	Principal Arterials (Total=40)	Principal Arterials (Total=40)	Minor Arterials (Total=45)	Minor Arterials (Total=45)	Collectors (Total=50)
Pavement Type	1	010	010	010	015	015	015
Pavement Condition	2	025	025	025	025	025	030
Drainage Adequacy	3	005	005	005	005	005	005
Safety		Interstate (Total=30)	Principal Arterials < 3 Lanes (Total=30)	Principal Arterials > 3 Lanes (Total=30)	Minor Arterials < 3 Lanes (Total=30)	Minor Arterials > 3 Lanes (Total=30)	Collectors (Total=30)
Lane Width	4	009	013	009	013	009	013
Shoulder Width	5	004	004	004	004	004	004
Median Width	6	004	000	004	000	004	000
Alignment Adequacy	7	008	008	008	008	008	008
Accident		005	005	005	005	005	005
Service		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=25)	Minor Arterials (Total=25)	Collectors (Total=20)
Volume/SF Ratio	8	025	025	025	025	025	020
Access Control	9	005	005	005	000	000	000
Total Points		100	100	100	100	100	100

TABLE 7 HPMS Urban Composite Index Weights, Scenario 1

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5
Condition		Interstate (Total=40)	Freeways/Expressways (Total=40)	Principal Arterials (Total=40)	Minor Arterials (Total=45)	Collectors (Total=50)
Pavement Type	1	010	010	010	015	015
Pavement Condition	2	025	025	025	025	025
Drainage Adequacy	3	005	005	005	005	010
Safety		Interstate (Total=30)	Freeways/Expressways (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=30)	Collectors (Total=30)
Lane Width	4	017	017	017	017	025
Shoulder Width	5	004	004	000	000	000
Median Width	6	004	004	008	008	000
Accident		005	005	005	005	005
Service		Interstate (Total=30)	Freeways/Expressways (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=25)	Collectors (Total=20)
Volume/SF Ratio	7	025	025	025	025	020
Access Control	8	005	005	005	000	000
Total Points		100	100	100	100	100

TABLE 8 HPMS Rural Composite Index Weights, Scenario 2

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Condition		Interstate (Total=40)	Principal Arterials (Total=40)	Principal Arterials (Total=40)	Minor Arterials (Total=30)	Minor Arterials (Total=30)	Collectors (Total=25)
Pavement Type	1	010	010	010	010	010	010
Pavement Condition	2	025	025	025	015	015	010
Drainage Adequacy	3	005	005	005	005	005	005
Safety		Interstate (Total=30)	Principal Arterials < 3 Lanes (Total=30)	Principal Arterials > 3 Lanes (Total=30)	Minor Arterials < 3 Lanes (Total=45)	Minor Arterials > 3 Lanes (Total=45)	Collectors (Total=55)
Lane Width	4	009	013	009	020	015	020
Shoulder Width	5	004	004	004	005	005	005
Median Width	6	004	000	004	000	005	000
Alignment Adequacy	7	008	008	008	015	015	020
Accident		005	005	005	005	005	010
Service		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=25)	Minor Arterials (Total=25)	Collectors (Total=20)
Volume/SF Ratio	8	025	025	025	025	025	020
Access Control	9	005	005	005	000	000	000
Total Points		100	100	100	100	100	100

TABLE 9 HPMS Urban Composite Index Weights, Scenario 2

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5
Condition		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
----------		(Total=40)	(Total=40)	(Total=40)	(Total=30)	(Total=25)
Pavement Type	1	010	010	010	010	010
Pavement Condition	2	025	025	025	015	010
Drainage Adequacy	3	005	005	005	005	005
Safety		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
------		(Total=30)	(Total=30)	(Total=30)	(Total=45)	(Total=55)
Lane Width	4	017	017	017	030	045
Shoulder Width	5	004	004	000	000	000
Median Width	6	004	004	008	010	000
Accident		005	005	005	005	010
Service		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
-------		(Total=30)	(Total=30)	(Total=30)	(Total=25)	(Total=20)
Volume/SF Ratio	7	025	025	025	025	020
Access Control	8	005	005	005	000	000
Total Points		100	100	100	100	100

TABLE 10 HPMS Rural Composite Index Weights, Scenario 3

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Condition		Interstate	Principal Arterials	Principal Arterials	Minor Arterials	Collectors	Collectors
----------		(Total=40)	(Total=40)	(Total=40)	(Total=35)	(Total=35)	(Total=30)
Pavement Type	1	010	010	010	010	010	010
Pavement Condition	2	025	025	025	020	020	015
Drainage Adequacy	3	005	005	005	005	005	005
Safety		Interstate	Principal Arterials	Principal Arterials	Minor Arterials	Collectors	Collectors
------			< 3 Lanes	> 3 Lanes	< 3 Lanes		
		(Total=30)	(Total=30)	(Total=30)	(Total=45)	(Total=45)	(Total=55)
Lane Width	4	009	013	009	020	016	025
Shoulder Width	5	004	004	004	005	005	005
Median Width	6	004	000	004	000	004	000
Alignment Adequacy	7	008	008	008	010	010	010
Accident		005	005	005	010	010	015
Service		Interstate	Principal Arterials	Principal Arterials	Minor Arterials	Collectors	Collectors
-------		(Total=30)	(Total=30)	(Total=30)	(Total=20)	(Total=20)	(Total=15)
Volume/SF Ratio	8	025	025	025	020	020	015
Access Control	9	005	005	005	000	000	000
Total Points		100	100	100	100	100	100

TABLE 11 HPMS Urban Composite Index Weights, Scenario 3

Data Items by Category	Item Number	Group 1	Group 2	Group 3	Group 4	Group 5
Condition		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
----------		(Total=40)	(Total=40)	(Total=40)	(Total=35)	(Total=30)
Pavement Type	1	010	010	010	010	010
Pavement Condition	2	025	025	025	020	015
Drainage Adequacy	3	005	005	005	005	005
Safety		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
------		(Total=30)	(Total=30)	(Total=30)	(Total=45)	(Total=55)
Lane Width	4	017	017	017	025	040
Shoulder Width	5	004	004	000	000	000
Median Width	6	004	004	008	010	000
Accident		005	005	005	010	015
Service		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
-------		(Total=30)	(Total=30)	(Total=30)	(Total=20)	(Total=15)
Volume/SF Ratio	7	025	025	025	020	015
Access Control	8	005	005	005	000	000
Total Points		100	100	100	100	100

TABLE 12 HPMS Rural Composite Index Weights, Scenario 4

Data Items by Category	Item Number	Group 1	Group 2 < 3 Lanes	Group 3 > 3 Lanes	Group 4 < 3 Lanes	Group 5 > 3 Lanes	Group 6
Condition		Interstate (Total=40)	Principal Arterials (Total=40)	Principal Arterials (Total=40)	Minor Arterials (Total=35)	Minor Arterials (Total=35)	Collectors (Total=30)
Pavement Type	1	010	010	010	010	010	010
Pavement Condition	2	025	025	025	020	020	015
Drainage Adequacy	3	005	005	005	005	005	005
Safety		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=30)	Minor Arterials (Total=30)	Collectors (Total=30)
Lane Width	4	009	010	007	000	000	000
Shoulder Width	5	004	003	003	000	000	000
Median Width	6	004	000	003	000	008	000
Alignment Adequacy	7	008	007	007	015	007	000
Accident		005	010	010	015	015	030
Service		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=35)	Minor Arterials (Total=35)	Collectors (Total=40)
Volume/SF Ratio	8	025	010	014	007	005	000
Access Control	9	005	000	000	000	000	000
Lane Width	4	000	008	010	010	010	015
Shoulder Width	5	000	007	003	010	010	015
Alignment Adequacy	7	000	005	003	008	010	010
Total Points		100	100	100	100	100	100

TABLE 13 HPMS Urban Composite Index Weights, Scenario 4

Data Items by Category	Item Number	Group 1	Group 2 < 3 Lanes	Group 3 > 3 Lanes	Group 4 < 3 Lanes	Group 5 > 3 Lanes	Group 6
Condition		Interstate (Total=40)	Principal Arterials (Total=40)	Principal Arterials (Total=40)	Minor Arterials (Total=35)	Minor Arterials (Total=35)	Collectors (Total=30)
Pavement Type	1	010	010	010	005	005	005
Pavement Condition	2	025	025	025	020	020	015
Drainage Adequacy	3	005	005	005	010	010	010
Safety		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=30)	Minor Arterials (Total=30)	Collectors (Total=30)
Lane Width	4	017	010	010	005	000	005
Shoulder Width	5	004	005	005	000	000	000
Median Width	6	004	000	005	000	008	000
Accident		005	015	010	025	015	025
Service		Interstate (Total=30)	Principal Arterials (Total=30)	Principal Arterials (Total=30)	Minor Arterials (Total=35)	Minor Arterials (Total=35)	Collectors (Total=40)
Volume/SF Ratio	8	025	020	020	020	020	020
Access Control	9	005	000	000	000	000	000
Lane Width	4	000	005	005	015	015	020
Shoulder Width	5	000	005	005	000	000	000
Total Points		100	100	100	100	100	100

TABLE 14 HPMS Rural Appraisal Rates, Scenarios 1–4

Functional Class			Interstate	Principal Arterials		Minor Arterials		Collectors		
AADT			ALL	> 6000	<= 6000	> 2000	<= 2000	> 1000	400-1000	< 400
Data Items by Category	Item Number	Row Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Condition										
Pavement Type	1									
High-Flexible		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
High-Rigid		02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Intermediate		03	0.40	0.40	0.40	0.40	0.70	0.70	1.00	1.00
Low		04	0.20	0.20	0.20	0.20	0.50	0.50	0.70	1.00
Gravel		05	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.20
Graded & Drained		06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pavement Condition	2									
4.6 - 5.0		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4.1 - 4.5		02	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
3.6 - 4.0		03	0.85	0.85	0.85	0.85	0.90	0.90	0.90	0.90
3.1 - 3.5		04	0.75	0.75	0.75	0.80	0.85	0.85	0.85	0.85
2.6 - 3.0		05	0.70	0.70	0.70	0.75	0.80	0.80	0.80	0.80
2.3 - 2.5		06	0.50	0.50	0.50	0.70	0.70	0.75	0.75	0.75
2.1 - 2.2		07	0.40	0.40	0.40	0.50	0.60	0.70	0.70	0.75
1.9 - 2.0		08	0.20	0.20	0.20	0.30	0.50	0.50	0.50	0.70
1.6 - 1.8		09	0.10	0.10	0.10	0.10	0.30	0.30	0.30	0.40
1.1 - 1.5		10	0.00	0.00	0.00	0.00	0.10	0.15	0.15	0.20
<= 1.0		11	0.00	0.00	0.00	0.00	0.00	0.05	0.05	0.10
Drainage Adequacy	3									
Good		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Fair		02	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
Poor		03	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Safety										
Lane Width	4									
>= 12		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
11		02	0.60	0.85	0.90	0.95	1.00	1.00	1.00	1.00
10		03	0.30	0.70	0.80	0.85	0.80	0.80	0.90	1.00
9		04	0.00	0.50	0.50	0.50	0.70	0.70	0.80	0.90
<= 8		05	0.00	0.00	0.00	0.00	0.30	0.00	0.30	0.50
Shoulder Width	5									
>= 12		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10-11		02	0.90	0.95	1.00	1.00	1.00	1.00	1.00	1.00
8- 9		03	0.70	0.80	0.90	1.00	1.00	1.00	1.00	1.00
6- 7		04	0.50	0.70	0.75	0.90	0.95	0.95	1.00	1.00
4- 5		05	0.30	0.50	0.60	0.70	0.80	0.80	1.00	1.00
2 -3		06	0.10	0.30	0.30	0.40	0.50	0.60	0.80	1.00
1		07	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.50
0		08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median Width	6									
> 60		01	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
40-60		02	0.85	1.00	1.00	1.00	1.00	0.00	0.00	0.00
30-39		03	0.70	0.70	0.70	0.85	0.80	0.00	0.00	0.00
20-29		04	0.60	0.60	0.60	0.70	0.70	0.00	0.00	0.00
10-19		05	0.40	0.40	0.40	0.50	0.50	0.00	0.00	0.00
5- 9		06	0.10	0.10	0.10	0.30	0.20	0.00	0.00	0.00
1- 4		07	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00
0		08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

(continued on next page)

TABLE 14 ***(continued)***

Functional Class			Interstate	Principal Arterials	Minor Arterials	Collectors
Data Items by Category	Item Number	Row Number	Group 1	Group 2	Group 3	Group 4
Alignment Adequacy	7					
Horizontal						
1		01	1.00	1.00	1.00	1.00
2		02	0.70	0.70	0.70	0.90
3		03	0.40	0.40	0.40	0.70
4		04	0.00	0.00	0.00	0.20
Vertical						
1		01	1.00	1.00	1.00	1.00
2		02	0.70	0.70	0.70	0.90
3		03	0.40	0.40	0.40	0.70
4		04	0.00	0.00	0.00	0.20

Functional Class			Interstate	Principal Arterials		Minor Arterials		Collectors		
AADT			ALL	> 6000	<= 6000	> 2000	<= 2000	> 1000	400-1000	< 400
Data Items by Category	Item Number	Row Number	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Service										
Volume/SF Ratio	8									
< .20		01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
.20-.39		02	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
.40-.49		03	0.90	0.90	0.90	0.90	0.95	0.95	0.95	0.95
.50-.59		04	0.85	0.85	0.85	0.85	0.90	0.90	0.90	0.90
.60-.64		05	0.80	0.80	0.80	0.80	0.85	0.85	0.85	0.85
.65-.69		06	0.75	0.75	0.75	0.75	0.80	0.80	0.80	0.80
.70-.74		07	0.70	0.70	0.70	0.70	0.75	0.75	0.75	0.75
.75-.79		08	0.60	0.60	0.60	0.60	0.70	0.70	0.70	0.70
.80-.84		09	0.50	0.50	0.50	0.50	0.60	0.60	0.60	0.60
.85-.89		10	0.40	0.40	0.40	0.40	0.50	0.50	0.50	0.50
.90-.94		11	0.20	0.20	0.20	0.20	0.30	0.30	0.30	0.30
>=.95		12	0.10	0.10	0.10	0.10	0.15	0.15	0.15	0.15
Access Control	9									
Full		01	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
Partial		02	0.00	0.60	0.80	0.00	0.00	0.00	0.00	0.00
None		03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 15 HPMS Urban Appraisal Rates, Scenarios 1–4

Functional Class			Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
Data Items by Category	Item Number	Row Number	Group 1	Group 2	Group 3	Group 4	Group 5
Condition							
Pavement Type	1						
High-Flexible		01	1.00	1.00	1.00	1.00	1.00
High-Rigid		02	1.00	1.00	1.00	1.00	1.00
Intermediate		03	0.40	0.40	0.40	0.70	0.85
Low		04	0.00	0.00	0.00	0.40	0.70
Gravel		05	0.00	0.00	0.00	0.00	0.30
Graded & Drained		06	0.00	0.00	0.00	0.00	0.00
Pavement Condition	2						
4.6 - 5.0		01	1.00	1.00	1.00	1.00	1.00
4.1 - 4.5		02	0.90	0.95	0.95	0.95	0.95
3.6 - 4.0		03	0.80	0.85	0.85	0.90	0.90
3.1 - 3.5		04	0.70	0.75	0.75	0.85	0.85
2.6 - 3.0		05	0.60	0.70	0.70	0.80	0.80
2.3 - 2.5		06	0.50	0.50	0.60	0.70	0.75
2.1 - 2.2		07	0.40	0.40	0.40	0.60	0.70
1.9 - 2.0		08	0.20	0.20	0.20	0.40	0.50
1.6 - 1.8		09	0.10	0.10	0.10	0.25	0.30
1.1 - 1.5		10	0.00	0.00	0.00	0.15	0.20
<= 1.0		11	0.00	0.00	0.00	0.05	0.10
Drainage Adequacy	3						
Good		01	1.00	1.00	1.00	1.00	1.00
Fair		02	0.70	0.70	0.70	0.70	0.70
Poor		03	0.10	0.10	0.10	0.10	0.10
Safety							
Lane Width	4		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
>= 12		01	1.00	1.00	1.00	1.00	1.00
11		02	0.70	0.70	0.90	0.90	0.90
10		03	0.50	0.50	0.70	0.80	0.80
9		04	0.25	0.25	0.60	0.70	0.70
<= 8		05	0.00	0.00	0.00	0.00	0.00
Shoulder Width	5		Interstate	Freeways/Expressways			
>= 12		01	1.00	1.00			
10-11		02	0.90	0.90			
8- 9		03	0.70	0.70			
6- 7		04	0.50	0.50			
4- 5		05	0.30	0.30			
2- 3		06	0.10	0.10			
1		07	0.00	0.00			
0		08	0.00	0.00			
Median Width **	6		Prn. Arts. Barrier	Principal Arterials No Barriers	Minor Arterials Built-Up	Minor Arterials Outlying	Collectors
> 60		01	1.00	1.00	1.00	1.00	0.00
40-60		02	1.00	0.85	0.85	0.90	0.00
30-39		03	0.70	0.70	0.70	0.80	0.00
20-29		04	0.70	0.50	0.70	0.75	0.00
10-19		05	0.70	0.30	0.70	0.70	0.00
5- 9		06	0.70	0.20	0.70	0.50	0.00
1- 4		07	0.70	0.10	0.70	0.30	0.00
0		08	0.70	0.00	0.70	0.00	0.00
Service							
Volume/SF Ratio	7		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
<= .20		01	1.00	1.00	1.00	1.00	1.00
.21-.40		02	0.95	0.95	0.95	1.00	1.00
.41-.60		03	0.90	0.90	0.90	0.95	0.95
.61-.70		04	0.85	0.85	0.85	0.90	0.90
.71-.75		05	0.80	0.80	0.80	0.85	0.85
.76-.80		06	0.75	0.75	0.75	0.80	0.80
.81-.85		07	0.70	0.70	0.70	0.75	0.75
.86-.90		08	0.40	0.40	0.40	0.70	0.70
.91-.95		09	0.10	0.10	0.10	0.30	0.30
> .96		10	0.05	0.05	0.05	0.10	0.10

(continued on next page)

TABLE 15 *(continued)*

Data Items by Category	Item Number	Row Number	Group 1	Group 2	Group 3	Group 4	Group 5
Access Control	8		Interstate	Freeways/Expressways	Principal Arterials	Minor Arterials	Collectors
Full		01	1.00	1.00	1.00	0.00	0.00
Partial		02	0.00	0.30	0.80	0.00	0.00
None		03	0.00	0.00	0.40	0.00	0.00

** - One-way streets always get a rating of 1.00

** - The code for median type is not "2"

TABLE 16 Summary of Composite Indexes for Scenarios 1–4 and Adequacy Rating

ROUTE	FUNCTIONAL SYSTEM	Adq.	Sc. 1	Sc. 2	Sc. 3	Sc. 4
I-64	Urban Interstate	86.00	88.35	88.35	88.35	88.35
I-75	Urban Interstate	89.40	81.30	81.30	81.30	81.30
I-64	Rural Interstate	88.00	88.35	88.30	88.35	88.35
I-75	Rural Interstate	83.00	73.35	73.35	73.35	73.35
I-75	Rural Interstate	83.00	70.85	70.85	70.85	70.85
KY 4	Urban Freeway	71.00	85.90	85.90	85.90	91.50
US 25	Urban Princ. Art.	56.00	82.30	82.30	82.30	76.15
US 60	Urban Princ. Art.	78.00	82.50	82.50	82.50	88.50
US 60	Rural Princ. Art.	71.00	81.60	81.60	81.60	61.30
US 68	Rural Princ. Art.	59.00	75.60	75.60	75.60	79.25
KY 353	Urban Minor Art.	49.00	78.80	75.20	75.90	84.50
KY 1681	Urban Minor Art.	39.00	82.55	81.25	81.00	90.00
US 60	Urban Minor Art.	76.00	90.55	91.95	90.40	91.50
US 460	Rural Minor Art.	72.00	89.00	86.25	88.50	77.05
US 460	Rural Minor Art.	54.00	81.70	78.15	77.80	59.05
KY 1267	Urban Collector	58.00	85.75	83.50	84.25	87.25
KY 1968	Urban Collector	49.00	65.75	51.00	54.75	68.40
KY 922	Rural Major Coll.	49.00	82.00	76.60	72.15	51.75
US 60	Rural Major Coll.	84.00	95.50	98.50	97.75	97.75
KY 1966	Rural Minor Coll.	55.00	77.00	72.00	66.25	66.25
KY 1268	Rural Minor Coll.	58.00	82.30	77.50	78.00	77.00

weighting factors, appraisal rates, and data attributes of the three scenarios, the panel decided to adopt S1 without accident attribute as the standard common criteria for the Adequacy Rating and the HPMS programs. The direction from the expert panel eliminated the inconsistency and variation in the final composite indexes from the two programs that were of major concern among management personnel. The first biennial Adequacy Rating report by functional system using the aforementioned criteria from Scenario S1 was to be published in December 1991.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Planning for Texas's Needs Using Highway Performance Monitoring System Analytical Process

ARTIE V. ELLIOTT, JR.

The Highway Performance Monitoring System (HPMS) analytical process was developed by FHWA to assess highway conditions and estimate national highway investment needs. It is structured to provide information about the effects of alternative standards, policy strategies, and program allocations on highway needs and performance. It is capable of revealing trends in highway capacity, condition, and service as well as trends in driver safety and environmental effects of airborne pollutants. The Texas Department of Transportation used HPMS to develop its 1989 Strategic Mobility Plan. This 20-year needs assessment for 1990 through 2009 pivoted around HPMS's ability to estimate backlog needs, project rehabilitation and reconstruction needs, and augment other information about new construction and right-of-way needs. To apply HPMS in Texas, changes were made to tailor the results to reflect Texas's standards and practices. The sample size was increased to facilitate use of the model at the district level. Highway improvement costs were factored to reflect differences between national average costs and those in Texas. Rural and urban traffic growth rates were adjusted downward to be more consistent with recent growth trends and default values for minimum tolerable conditions and design standards were modified. HPMS has minimized the employee hours required to identify funding needs, provided statewide consistency in assessing needs, prevented wish-list tendencies, and allowed for the "global" assessment of funding requirements. It is a tool that allows examination of diverse scenarios and alternative uses of funds.

The Texas Department of Transportation's (TxDOT's) use of the FHWA Highway Performance Monitoring System (HPMS) is perhaps better understood by looking first at the size and diversity of the state and at the department's historical approach to estimating needs. This overview highlights the complexities of and supports the department's shift to statistical sampling and computer modeling for needs estimation.

SIZE AND DIVERSITY

In area, Texas is the second-largest state in the union; it ranks third in population. Texas has 262,145 mi^2 of land and 16.8 million people. TxDOT's 12 tourist bureaus welcome more than 3 million visitors each year. Geographical and climatic variety encompass mountains and deserts in the west, snow and freeze-thaw cycles in the north, piney woods with abundant rain in the east, and Gulf Coast ocean and sand to the south. Eleven urban areas have populations over 200,000, and 19 smaller urban areas have populations over 50,000. More than 80 percent of the people live in urban areas. The state has 13.9 million registered vehicles and 11.4 million active licensed drivers, which log more than 107 billion vehicle-mi of travel a year. All of this occurs on an excess of 305,000 centerline-mi of roads, 46,000 bridges and culverts, 14,000 railroad crossings, 10 ferries, and 1 tunnel. With more than 30,000 bridges, Texas leads the nation in the number of bridges on a state highway system. The Interstate system has almost 3,200 centerline-mi; another 31,500 mi are U.S. or state highways; 41,400 mi make up the farm-to-market system; and about 1,000 mi are metropolitan highways.

TxDOT must address Texas's great socioeconomic, demographic, and geographical diversity in its responsibility for designing, constructing, and maintaining the roughly 77,000 centerline-mi of highways that have an estimated replacement value of $100 billion. To accomplish this, the department has an annual construction budget of about $1.5 billion and approximately 15,000 employees in 18 divisions and 18 field districts. There are 288 maintenance sections and 133 resident engineer offices around the state.

HISTORICAL APPROACH TO NEEDS DETERMINATION

Before the adoption of HPMS, TxDOT used committees to determine personnel, equipment, materials, and highway needs for its 20-year Strategic Mobility Plan (SMP). The information was obtained from the districts and divisions and involved historical data, a department computer model (RENU) that projected pavement maintenance and rehabilitation requirements, and the Project Development Plan (PDP). The PDP provides the basis for programming and scheduling individual construction projects. This process had significant shortcomings in that it was labor-intensive, it reflected the inconsistencies and biases of the many participants, and it was viewed by certain political figures as a wish-list. To improve the process, a decision was made to use the computer modeling and statistical projection capabilities of the HPMS analytical process to identify highway needs for TxDOT's 1989 SMP. HPMS-computed roadway-related needs amounted to approximately $33 billion, which is 40 percent of the department's total identified needs. It was used to forecast rehabilitation, reconstruction, additional capacity, new construction, and right-of-way (ROW) requirements. Needs for human resources, equip-

---

Texas Department of Transportation, 125 East 11th Street, Austin, Tex. 78701.

ment, highway maintenance, bridges, loops and bypasses around cities, and new construction projects identified in the PDP were estimated outside HPMS to produce total fiscal requirements for TxDOT.

HPMS PARAMETER ADJUSTMENTS

To estimate more accurately highway infrastructure needs, HPMS was tailored for use at the state level (*1*). In 1987 the HPMS sample size was increased to enhance statistical validity at the district level. Highway improvements costs were factored to reflect differences between national average costs and those in Texas. Resurfacing costs for rural roads were also adjusted. Rural and urban traffic growth rates were adjusted downward to be more consistent with recent growth trends. Default values for minimum tolerable conditions (MTCs), present serviceability ratings (PSRs), and design standards were modified to be more consistent with standards and practices in Texas. The maximum number of lanes and ROW requirements were constrained for forecasting needs. Details of the tailoring follow.

Roadway Samples

From the inventory of 22,559 universe sections, the number of sample sections was increased from 2,363 to 5,702. Universe miles went from 72,749 to 77,244, and the corresponding sample miles were expanded from 8,810 to 17,771. These adjustments stemmed from a 3-year research project sponsored by FHWA and TxDOT and was conducted by the Texas Transportation Institute (TTI) of Texas A&M University (*2*). Because HPMS data were being collected and reported to FHWA on a yearly basis, it was believed that the number of samples could be expanded for effective use within the state. To further increase the validity of the sample, TxDOT conducts surveys of the HPMS sample sites. A TxDOT employee and a representative from FHWA visit one district each month to assess individual district input as well as uniformity of statewide input. This frequency results in an 18-month validation cycle and is believed to be adequate for gauging the quality of field input.

Costs

Adjustments were made to highway improvement costs used in the analytical package. National ROW and construction costs from FHWA were factored in the aggregate to reflect costs in Texas. Since the costs in the data base are in 1986 dollars, they were adjusted to 1988 constant dollars to coincide with the beginning year of the 20-year plan. A highway construction cost index (HCI), calculated annually by TxDOT, was used in making these adjustments. Table 1 presents the adjusted costs that were calculated using the dollar price index (DPI) program parameter (output control card). In addition, for rural roads receiving seal-coat treatment in lieu of rehabilitation, resurfacing costs were adjusted with a program parameter card linked to a program (input) control card to properly reflect this anomaly (*1*) (Table 2). Those costs not related to construction were adjusted from 1986 dollars to 1988 dollars by use of the state and local government index.

Traffic Growth Rates

During development of the SMP in 1988, the unadjusted average annual daily traffic (AADT) values in HPMS produced a statewide, 20-year compounded growth rate of 2.94 percent a year. At the same time, the low and high population growth rates projected by the Texas Water Development Board for these 20 years were 1.38 and 2.26 percent, respectively. In view of the uncertainty regarding the traffic growth for the 20 years and considering the state's declining growth rate of recent years, TxDOT adopted conservative, low-end growth projections for use in the SMP. To accomplish this, the AADT growth was adjusted by a program parameter card coupled with program control cards. The HPMS analytical model facilitated detailed projections of growth in Texas by allowing

TABLE 1 DPI Adjustments

Year	1986	1987	1988
TxDot HCI	1.233	1.071	1.079

DPI = (1.079/1.233) x .845 = .74

Note: Texas's 1988 costs = 0.845 of the 1988 national average costs.

TABLE 2 Rural Road Seal-Coat Costs Adjustments

	INT STATE			OTH P ART			MI ART			MAJ COL			MI COL		
Resurfing	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M
Fed Deflt	72	74	98	51	53	92	40	42	73	15	16	27	11	18	24
TX Costs	58	60	80	41	43	75	33	34	59	8	9	15	6	6	6

Legend: (Roadway Functional Classes: Int State - Interstate; Oth P Art - Other Principal Arterials; Mi Art -Minor Arterials; Maj Col - Major Collectors; Mi Col - Minor Collectors); (Resurfng - Resurfacing; F, R, M - Flat, Rolling, and Mountainous Terrains); (Fed Deflt - FHWA National Default Costs); (Tx Costs - Texas' Costs); NOTE: All costs are in thousands of dollars.

TABLE 3 Texas HPMS Defaults, Rural/Urban AADT Adjustment Factors

RURAL FACTORS:		INT	PRIN ART	MI ART	MAJ COL	MI COL
		0.75	0.75	0.75	0.75	0.75
URBANIZED FACTORS:		INT	OFE	OPA	MI ART	COL
Area Code	015	0.41	0.41	0.41	0.41	0.41
Area Code	028	0.48	0.48	0.48	0.48	0.48
Area Code	066	0.56	0.56	0.56	0.56	0.56
Area Code	090	0.61	0.61	0.61	0.61	0.61
Area Code	096	0.28	0.28	0.28	0.28	0.28
Area Code	120	0.24	0.24	0.24	0.24	0.24
Area Code	122	0.34	0.34	0.34	0.34	0.34
Area Code	135	0.01	0.01	0.01	0.01	0.01
Area Code	137	0.39	0.39	0.39	0.39	0.39
Area Code	139	0.01	0.01	0.01	0.01	0.01
Area Code	140	0.06	0.06	0.06	0.06	0.06
Area Code	151	0.50	0.50	0.50	0.50	0.50
Area Code	166	0.33	0.33	0.33	0.33	0.33
Area Code	174	0.21	0.21	0.21	0.21	0.21
Area Code	197	0.19	0.19	0.19	0.19	0.19
Area Code	201	0.64	0.64	0.64	0.64	0.64
Area Code	205	0.79	0.79	0.79	0.79	0.79
Area Code	208	0.14	0.14	0.14	0.14	0.14
Area Code	211	0.25	0.25	0.25	0.25	0.25
Area Code	213	0.57	0.57	0.57	0.57	0.57
Area Code	230	0.83	0.83	0.83	0.83	0.83
Area Code	232	0.20	0.20	0.20	0.20	0.20
Area Code	248	0.64	0.64	0.64	0.64	0.64
Area Code	249	0.52	0.52	0.52	0.52	0.52
Area Code	250	0.39	0.39	0.39	0.39	0.39
Area Code	277	0.69	0.69	0.69	0.69	0.69
Area Code	282	0.31	0.31	0.31	0.31	0.31
Area Code	361	0.22	0.22	0.22	0.22	0.22
Area Code	362	0.69	0.69	0.69	0.69	0.69
Area Code	363	0.33	0.33	0.33	0.33	0.33

Note: National Default AADT Adjustment Factor value is 1.00 for each area code. Texas values differ from National Defaults and are highlighted with shading.

Legend: INT - Interstate; PRIN ART - Principal Arterial; MI ART - Minor Arterial; MAJ COL - Major Collector; MI COL -Minor Collector; OFE - Other Freeway/Expressway; OPA - Other Principal Arterial; COL - Collector.

growth rates by functional class of road for both rural and 30 urban areas (*1*). The actual AADT adjustment values used in the analytical model are presented in Table 3.

Minimum Tolerable Conditions and Related Changes

Individual MTCs were changed in the model with a program parameter card tied to program control cards so that the decision threshold between rehabilitation and reconstruction would represent standard practice in Texas (*1*). Some rural and urban shoulder widths were changed, as were a few pavement condition ratings and surface type definitions. Also, some volume-capacity ratios were changed for the urban areas to represent the use of freeway control systems that optimize traffic flow. Additionally, certain rural lane widths and shoulder type definitions were changed. Finally, lower functional class rural and urban roadway PSRs were adjusted with program parameter cards to reflect reconstruction practices in Texas (*1*). Adjustments used are presented in Tables 4 through 11.

Design Standards

Design standards were changed by the use of a program parameter card connected to program control cards (*1*). This facilitated the change of some rural lane widths along with rural and urban surface type definitions. Changes were also made to a small number of urban shoulder widths and average highway speeds. These modifications are presented in Tables 12 through 15.

Number of Lanes

There is an upward limit to the number of highway lanes that are acceptable in existing locations. Therefore, a lane limit roughly 50 percent greater than the default values of 8/10 lanes in the HPMS analytical model was used (*3*). This 12-lane-limit philosophy embraced vertical and horizontal capacity expansion beyond the FHWA program default lanes. In Texas, capacity demand beyond 12 lanes must be met by parallel corridors or other transportation solutions. In view of environmental concerns and the high cost of urban ROW, it was determined that long-range solutions for the latent capacity demand will require the use of increased mass transit, travel demand management, and traffic systems management.

APPLICATION ASSESSMENT

In assessing the tailoring of HPMS for Texas, as well as reviewing activities related to data gathering and computer mod-

TABLE 4 National HPMS Defaults, Rural/Urban Options in Effect

ANALYSIS MODE: FUNDING PERIOD
BASE YEAR: 1989
LENGTH CYCLE: 1 YEAR(S)
DOLLAR PRICE INDEX: 1.00
ADJUST FUTURE ADT: NO
DESIGN ADT FOR RURAL FULL ACCESS CONTROL: 10000

TARGET YEAR: 2009
PAVEMENT CYCLE AHEAD: 5 YEARS
FUTURE ADT YEAR: 2009

	INT	PRIN ART	MI ART	MAJ COL	MI COL
Rural maximum number lanes:	10	10	10	8	8
Rural PSR to determine reconst:	2.0	2.0	1.5	1.1	0.8

		INT	OFE	OPA	MI ART	COL
Urban maximum number lanes:						
	Built-up	8	8	8	8	8
	Outlying	10	10	10	10	8
Urban PSR to determine reconst:		2.2	2.0	1.8	1.1	1.0

SURFACE TYPE:	H FLEX	H RGD	INTR	LOW
Rural Improved PSR for Reconstruction:	4.6	4.6	4.4	4.2
Rural factor added to existing PSR for resurf:	1.8	1.8	1.8	1.8
Rural maximum improved PSR for resurfacing	4.3	4.3	4.2	4.0
Urban improved PSR for reconstruction:	4.6	4.6	4.4	4.2
Urban factor added to existing PSR for resurf:	1.8	1.8	1.8	1.8
Urban maximum improved PSR for resurfacing:	4.3	4.3	4.2	4.0

Note: National Defaults are provided for reference.

Legend: INT - Interstate; PRIN ART - Principal Arterial; MI ART - Minor Arterial; MAJ COL - Major Collector; MI COL -Minor Collector; OFE - Other Freeway/Expressway; OPA - Other Principal Arterial; COL - Collector; H Flex - High Flexible; H RGD - High Rigid; INTR - Intermediate.

TABLE 5 Texas HPMS Defaults, Rural/Urban Options in Effect

ANALYSIS MODE: FUNDING PERIOD
BASE YEAR: 1989
LENGTH CYCLE: 1 YEAR(S)
DOLLAR PRICE INDEX: 0.74
ADJUST FUTURE ADT: YES
DESIGN ADT FOR RURAL FULL ACCESS CONTROL: 10000

TARGET YEAR: 2009
PAVEMENT CYCLE AHEAD: 5 YEARS
FUTURE ADT YEAR: 2009

	INT	PRIN ART	MI ART	MAJ COL	MI COL
Rural maximum number lanes:	12	12	12	12	12
Rural PSR to determine reconst:	2.0	2.0	1.5	1.5	1.5

		INT	OFE	OPA	MI ART	COL
Urban maximum number lanes:						
	Built-up	12	12	12	12	12
	Outlying	12	12	12	12	12
Urban PSR to determine reconst:		2.2	2.0	1.8	1.5	1.5

SURFACE TYPE:	H FLEX	H RGD	INTR	LOW
Rural Improved PSR for Reconstruction:	4.6	4.6	4.6	4.5
Rural factor added to existing PSR for resurf:	1.8	1.8	1.8	1.8
Rural maximum improved PSR for resurfacing	4.3	4.3	4.2	4.0
Urban improved PSR for reconstruction:	4.6	4.6	4.6	4.5
Urban factor added to existing PSR for resurf:	1.8	1.8	1.8	1.8
Urban maximum improved PSR for resurfacing:	4.3	4.3	4.2	4.0

Note: Texas values that differ from National Defaults are highlighted with shading.

Legend: INT - Interstate; PRIN ART - Principal Arterial; MI ART - Minor Arterial; MAJ COL - Major Collector; MI COL -Minor Collector; OFE - Other Freeway/Expressway; OPA - Other Principal Arterial; COL - Collector; H Flex - High Flexible; H RGD - High Rigid; INTR - Intermediate.

TABLE 6 National HPMS Defaults, Urban MTC

	INTERSTATE	OTH F/EXP	OTH P ART	MINOR ART	COLLECTOR
Volume-To-Capacity Ratio	0.95	0.95	0.95	0.95	0.95
Lane Width	12	11	10	8	8
Surface Type	2	2	2	3	4
Pavement Condition	3.2	3.0	2.8	2.4	2.0
Shoulder Type	1	1	2	3	3
Right Shoulder Width	8	8	6	6	6

Note: *Widths are in feet.*
National Defaults are provided for reference.

Legend: OTH F/EXP - Other Freeway/Expressway; OTH P ART - Other Principal Arterial; MINOR ART - Minor Arterial.

Shoulder Type Codes:	*Surface Type Codes:*
1 - Surfaced	*1 - High Flexible*
2 - Stabilized	*2 - High Rigid*
3 - Earth	*3 - Intermediate*
4 - Curbed	*4 - Low*
	5 - Gravel

TABLE 7 Texas HPMS Defaults, Urban MTC

	INTERSTATE	OTH F/EXP	OTH P ART	MINOR ART	COLLECTOR
Volume-To-Capacity Ratio	1.05	1.05	0.95	0.95	0.95
Lane Width	12	11	10	10	10
Surface Type	1	1	3	3	4
Pavement Condition	2.5	2.3	2.0	2.0	2.0
Shoulder Type	1	1	2	3	3
Right Shoulder Width	8	8	4	2	0

Note: Widths are in feet.
Texas values that differ from National Defaults are highlighted with shading.

Legend: OTH F/EXP - Other Freeway/Expressway; OTH P ART - Other Principal Arterial; MINOR ART - Minor Arterial.

Shoulder Type Codes:	*Surface Type Codes:*
1 - Surfaced	*1 - High Flexible*
2 - Stabilized	*2 - High Rigid*
3 - Earth	*3 - Intermediate*
4 - Curbed	*4 - Low*
	5 - Gravel

TABLE 8 National HPMS Defaults, Rural MTC (Interstates, Arterials)

	INTERSTATE			OTHER PRINCIPAL ARTERIALS						MINOR ARTERIALS					
ADT	ALL ADT			> 6000			< OR = 6000			> 2000			< OR = 2000		
Terrain	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M
Lane Width	12	12	12	11	11	11	11	11	11	10	10	10	10	10	10
Rshld Width	8	8	6	8	8	6	8	8	6	6	6	4	6	6	4
Shld Type	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3
Pave Cond	3.0	3.0	3.0	3.0	3.0	3.0	2.8	2.8	2.8	2.4	2.4	2.4	2.4	2.4	2.4
V/C Ratio	0.75	0.90	0.95	0.75	0.85	0.95	0.75	0.85	0.95	0.75	0.85	0.95	0.75	0.85	0.95
Surf Type	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
Horiz Align	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Vert Align	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Note: *Widths are in feet.*
National Defaults are provided for reference.

Legend: Terrain types F, R, M are Flat, Rolling, Mountainous.

Shoulder Type Codes:	*Surface Type Codes:*	*Horizontal/Vertical Alignment Codes:*
1 - Surfaced	*1 - High Flexible*	*1 - All curves/grades meet design standards*
2 - Stabilized	*2 - High Rigid*	*2 - Some curves/grades below design standards*
3 - Earth	*3 - Intermediate*	*3 - Curves/grades with reduced speed*
4 - Curbed	*4 - Low*	*4 - Several curves unsafe/significant reduction of speed on grade*
	5 - Gravel	

TABLE 9 National HPMS Defaults, Rural MTC (Major and Minor Collectors)

	MAJOR AND MINOR COLLECTORS								
ADT	> 1000			400 - 1000			< 400		
Terrain	F	R	M	F	R	M	F	R	M
Lane Width	10	10	10	8	8	8	16	16	16
Rshld Width	4	4	4	2	2	2	0	0	0
Shld Type	3	3	3	3	3	3	3	3	3
Pave Cond	2.0	2.0	2.0	2.0	2.0	2.0	1.8	1.8	1.8
V/C Ratio	0.75	0.85	0.95	1.00	1.00	1.00	1.00	1.00	1.00
Surf Type	3	3	3	4	4	4	5	5	5
Horiz Align	2	2	2	3	3	3	3	3	3
Vert Align	2	2	2	3	3	3	3	3	3

Note: Widths are in feet.
National Defaults are provided for reference.
MTC shown for lane width on collectors group 3 are for surface width.

Legend: Terrain types F, R, M are Flat, Rolling, Mountainous.

Shoulder Type Codes:
1 - Surfaced
2 - Stabilized
3 - Earth
4 - Curbed

Surface Type Codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

Horizontal/Vertical Alignment Codes:
1 - All curves/grades meet design standards
2 - Some curves/grades below design standards
3 - Curves/grades with reduced speed
4 - Several curves unsafe/significant reduction of speed on grade

TABLE 10 Texas HPMS Defaults, Rural MTC (Interstates, Arterials)

	INTERSTATE			OTHER PRINCIPAL ARTERIALS						MINOR ARTERIALS					
ADT	ALL ADT			> 6000			< OR = 6000			> 2000			< OR = 2000		
Terrain	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M
Lane Width	12	12	12	11	11	11	11	11	11	11	11	11	10	10	10
Rshld Width	8	8	6	8	8	6	4	4	4	4	4	4	4	4	4
Shld Type	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
Pave Cond	2.5	2.5	2.5	2.4	2.4	2.4	2.3	2.3	2.3	2.2	2.2	2.2	2.0	2.0	2.0
V/C Ratio	0.75	0.85	0.90	0.75	0.85	0.90	0.80	0.85	0.90	0.80	0.85	0.90	0.85	0.90	0.95
Surf Type	1	1	1	3	3	3	3	3	3	3	3	3	3	3	3
Horiz Align	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3
Vert Align	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3

Note: Widths are in feet.
Texas values that differ from National Defaults are highlighted with shading.

Legend: Terrain types F, R, M are Flat, Rolling, and Mountainous.

Shoulder Type Codes:
1 - Surfaced
2 - Stabilized
3 - Earth
4 - Curbed

Surface Type Codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

Horizontal/Vertical Alignment Codes:
1 - All curves/grades meet design standards
2 - Some curves/grades below design standards
3 - Curves/grades with reduced speed
4 - Several curves unsafe/significant reduction of speed on grade

eling of needs, Texas's use of HPMS has been found to be quite satisfactory. A private consulting firm, university research studies, a governmental agency study, and in-house studies have found our efforts to be on target in a number of areas, to require enhancements in a few areas, and to have the potential for use in other areas.

On Target

- Statistical analysis indicates that the error due to HPMS's being run on a sample of highway sections rather than on all highway sections is small when HPMS is used at the state level. Specifically, a 1991 analysis indicates that, at the state

TABLE 11 Texas HPMS Defaults, Rural MTC (Major and Minor Collectors)

	MAJOR AND MINOR COLLECTORS								
ADT	> 1000			400 - 1000			< 400		
Terrain	F	R	M	F	R	M	F	R	M
Lane Width	10	10	10	9	9	9	18	18	18
Rshld Width	4	4	4	2	2	2	0	0	0
Shld Type	2	2	2	3	3	3	3	3	3
Pave Cond	2.0	2.0	2.0	2.0	2.0	2.0	1.8	1.8	1.8
V/C Ratio	0.85	0.90	0.95	1.00	1.00	1.00	1.00	1.00	1.00
Surf Type	3	3	3	4	4	4	4	4	4
Horiz Align	3	3	3	3	3	3	3	3	3
Vert Align	3	3	3	3	3	3	3	3	3

Note: *Widths are in feet.*
MTC shown for lane width on collectors group 3 are for surface width.
Texas values that differ from National Defaults are highlighted with shading.

Legend: Terrain types F, R, M are Flat, Rolling, and Mountainous.

Shoulder Type Codes:	*Surface Type Codes:*
1 - Surfaced	*1 - High Flexible*
2 - Stabilized	*2 - High Rigid*
3 - Earth	*3 - Intermediate*
4 - Curbed	*4 - Low*
	5 - Gravel

Horizontal/Vertical Alignment Codes:
1 - All curves/grades meet design standards
2 - Some curves/grades below design standards
3 - Curves/grades with reduced speed
4 - Several curves unsafe/significant reduction of speed on grade

TABLE 12 National HPMS Defaults, Urban Design Standards

	FREEWAY/EXPRESSWAY		OTHER DIVIDED	
	BUILT-UP	OUTLYING	BUILT-UP	OUTLYING
Average Highway Speed	55	65	-	-
Median Width	16	24	-	-
Lane Width	12	12	12	12
Right Shoulder Width*	10	10	10	10
Left Shoulder Width*	6	6	6	6
Surface Type	2	2	2	2

	UNDIVIDED ARTERIALS		UNDIVIDED COLLECTORS	
	BUILT-UP	OUTLYING	BUILT-UP	OUTLYING
Average Highway Speed	-	-	-	-
Median Width	-	-	-	-
Lane Width	12	12	12	12
Right Shoulder Width*	8	10	6	10
Left Shoulder Width*	-	-	-	-
Surface Type	2	2	3	3

Note: *Widths are in feet.*
National Defaults are provided for reference.
Dash in table (-) indicates data not applicable.
Average Highway Speed is defined as the Weighted Average Design Speed in Miles per Hour.

*Legend: * For Facility Which Is Not Curbed*
Surface Type Codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

TABLE 13 Texas HPMS Defaults, Urban Design Standards

	FREEWAY/EXPRESSWAY		OTHER DIVIDED	
	BUILT-UP	OUTLYING	BUILT-UP	OUTLYING
Average Highway Speed	60	70	-	-
Median Width	16	24	-	-
Lane Width	12	12	12	12
Right Shoulder Width*	10	10	10	10
Left Shoulder Width*	4	4	4	4
Surface Type	1	1	1	1
	UNDIVIDED ARTERIALS		UNDIVIDED COLLECTORS	
	BUILT-UP	OUTLYING	BUILT-UP	OUTLYING
Average Highway Speed	-	-	-	-
Median Width	-	-	-	-
Lane Width	12	12	12	12
Right Shoulder Width*	8	10	4	8
Left Shoulder Width*	-	-	-	-
Surface Type	3	3	3	3

Note: *Widths are in feet.*
Dash in table (-) indicates data not applicable.
Texas values that differ from National Defaults are highlighted with shading.
Average Highway Speed is defined as the Weighted Average Design Speed in Miles per Hour.

Legend: ** For Facility Which Is Not Curbed*
Surface Type Codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

TABLE 14 National HPMS Defaults, Rural Design Standards

	INTERSTATE			OTHER PRINCIPAL ARTERIALS						MINOR ARTERIALS					
Design ADT	ALL ADT			> 6000			< OR = 6000			> 2000			< OR = 2000		
Terrain	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M
Shld Width	12	10	8	10	10	8	10	10	8	8	8	8	8	8	6
Surf Type	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3
Median Wdh	64	64	16	40	40	16	40	40	16	40	40	16	0	0	0
Lane Width	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Av Hwy Spd	70	70	55	70	65	55	70	65	55	70	60	50	65	55	45

	MAJOR AND MINOR COLLECTORS								
Design ADT	> 1000			400 - 1000			< 400		
Terrain	F	R	M	F	R	M	F	R	M
Shld Width	8	8	6	4	4	4	2	2	2
Surf Type	3	3	3	4	4	4	4	4	4
Median Wdh	0	0	0	0	0	0	0	0	0
Lane Width	12	12	11	11	11	11	10	10	10
Av Hwy Spd	65	55	45	60	50	40	50	40	30

Note: *Widths are in feet.*
National Defaults are provided for reference.
Av Hwy Spd: Average Highway Speed is defined as the Weighted Average Design Speed in Miles per Hour.

Legend: *Terrain types F, R, M are Flat, Rolling, and Mountainous.*
Surface type codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

TABLE 15 Texas HPMS Defaults, Rural Design Standards

	INTERSTATE			OTHER PRINCIPAL ARTERIALS						MINOR ARTERIALS					
Design ADT	ALL ADT			> 6000			< OR = 6000			> 2000			< OR = 2000		
Terrain	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M
Shld Width	12	10	8	10	10	8	10	10	8	8	8	8	8	8	6
Surf Type	1	1	1	1	1	1	1	1	1	1	1	1	3	3	3
Median Wdh	64	64	16	40	40	16	40	40	16	40	40	16	0	0	0
Lane Width	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Av Hwy Spd	70	70	55	70	65	55	70	65	55	70	60	50	65	55	45

	MAJOR AND MINOR COLLECTORS								
Design ADT	> 1000			400 - 1000			< 400		
Terrain	F	R	M	F	R	M	F	R	M
Shld Width	8	8	6	4	4	4	2	2	2
Surf Type	3	3	3	4	4	4	4	4	4
Median Wdh	0	0	0	0	0	0	0	0	0
Lane Width	12	12	11	12	12	11	10	10	10
Av Hwy Spd	65	55	45	60	50	40	50	40	30

Note: *Widths are in feet.*
Av Hyw Spd: Average Highway speed is defined as the Weighted Average Design Speed in Miles per Hour

Legend: *Terrain types F, R, M are Flat, Rolling, and Mountainous.*
Surface type codes:
1 - High Flexible
2 - High Rigid
3 - Intermediate
4 - Low
5 - Gravel

level, highway needs estimated through the use of the HPMS sample will be within 10 percent of highway needs estimated for all sections 99 percent of the time (*4*).

- A 1991 review of Texas's controls over HPMS data collection found them to be reasonable, especially since FHWA's monitoring of state data gathering provides a further check on data quality (*4*). These findings were consistent with the findings of the U.S. General Accounting Office (GAO) review of HPMS use in six other states. In 1987 GAO affirmed HPMS to be statistically sound in representing the various types of highways at the national aggregate level (*5*).
- GAO found the analytical model to be a reasonable tool for presenting useful information on the condition and capital investment needs of the nation's highways (*5*). At the state level, the Texas State Auditor's office completed a study in 1991 and endorsed the use of HPMS as a useful planning tool for TxDOT.
- For those MTCs codified in the *Texas Highway Design Manual*, a 1991 study found that there is substantial agreement between HPMS inputs and the manual (*4*). However, because of restrictions on the availability of funds, the state has been unable to undertake many projects that meet its criteria for funding. As a result, the standards and practices represented in the input to HPMS are somewhat stricter than those currently used in the field.
- The procedure used by HPMS to identify highway improvements was found in a 1991 examination to produce generally reliable results with some minor exceptions noted (*4*).

Future Enhancements

One private consultant firm's review suggested the following improvements for the HPMS process (*4*):

- Vigilance in recording HPMS projections for new construction needs in the SMP so that there is no overlap with funding requirements identified in the PDP.
- Careful instruction of district staff in the coding of pavement items to prevent understatement of pavement-related needs that HPMS identifies for the state as a whole.
- Prudent coding of HPMS resurfacing costs and threshold values to prevent overstatement of some farm-to-market road needs.

Future Utilization of HPMS

- HPMS does not explicitly develop estimates of maintenance needs or take into account the effects of maintenance expenditure levels and policies. It may be possible to reflect the effects of increased (or decreased) maintenance expenditures on resurfacing and reconstruction needs by adjusting factors within the HPMS model that determine pavement deterioration rates (*4*).
- HPMS does not specifically consider the effects of changes in public transportation service levels and policies. If changes in transit levels or policies are contemplated, they can be

represented in HPMS through ad hoc adjustments to traffic growth rates for specific urbanized areas or by temporary programming adjustments to sample data of specific locations (*4*).

- HPMS default AADT growth rates were adjusted from 3 to 1.4 percent a year for the 20-year planning period of the SMP. This rate was adopted intentionally to provide conservative, low-end estimates. Recent review of the default and adjusted growth rate values, along with the sensitivity of HPMS output to traffic growth rates, indicates that a number of growth rates should be used in future SMP evaluations. Use of several growth rates for future projections will be used to compensate for the high degree of uncertainty of traffic growth rates in the coming years.
- HPMS will be used for tracking the status of the highway system.
- The environmental effects of vehicle emissions, particularly in highly congested urban areas, will be evaluated.
- HPMS will be used to establish an approximation of needs for non-Interstate highways of national significance. This will be accomplished by using a TxDOT developed interface program that targets selected routes and corridors. This is accomplished by stripping all sample data not related to the selected route from the data base and by redefining appropriate expansion factors for the new universe.
- HPMS and the federally mandated Pavement Management System (PMS) currently under development address some overlapping areas: HPMS identifies rehabilitation, reconstruction, and some new construction needs, and PMS targets maintenance and reconstruction. Isolation of these overlapping areas as well as needs identified by the two different data bases will be devised so that each analytical package can enhance the other.
- Good correlation between the HPMS PSR scores and the Pavement Evaluation Scores used by PMS has been found through TxDOT and TTI research (T. Scullion and R. Smith, unpublished data). The department is currently moving toward one data collection effort that will satisfy HPMS and PMS requirements.
- As a result of deferred improvements because of lack of funds, HPMS will be used to evaluate increases in highway improvement costs.
- HPMS will be used to assess changes in pavement-related costs due to changes in truck volumes and truck weights.
- HPMS will be used to determine funds required to maintain the infrastructure. This will be accomplished by setting urban volume-capacity ratio thresholds so that the costs of construction and reconstruction for major widening and adding lanes are not simulated by the program.
- The 72nd Texas Legislature directed all state agencies to develop a 6-year strategic plan. HPMS will contribute significantly toward the development of this plan.

SUMMARY

The HPMS analytical process minimizes the employee hours required to estimate funding needs, provides statewide consistency in identifying needs, prevents wish-list tendencies, and minimizes political influence by globally assessing funding requirements in a non-project-specific manner. It can assess the highway infrastructure by estimating current and future deficiencies and improvement costs. HPMS is a policy planning tool that allows examination of a diversity of scenarios to best use available funds. It helps to position TxDOT to provide safe, dependable, and efficient movement of people and goods through effective and efficient use of funds.

REFERENCES

1. *Highway Performance Monitoring System Analytical Process*, Vol. III, Version 2.1. FHWA, U.S. Department of Transportation, Dec. 1987, pp. II-8–II-29.
2. J. L. Memmott. *Sample Size and Accuracy of Highway Performance Monitoring System*. Research Report 480-1. Texas Transportation Institute, Texas A&M University, College Station, May 1986.
3. *Strategic Mobility Plan*. Texas Department of Transportation, Austin, Feb. 1989, pp. 4–10.
4. *Review of the Highway Performance Monitoring System*. Cambridge Systematics, Inc., Jack Faucett Associates, Cambridge, Mass., March 1991, pp. 2-1–5-3.
5. *An Evaluation of DOT's Process for Assessing the Nation's Highway Needs*. GAO/RCED–87–136. U. S. General Accounting Office, Aug. 1987.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Use of Highway Performance Monitoring System and Bridge Needs and Investment Process for Reporting Conditions, Needs, and Performance Trends

LARRY W. MCPHERSON, MARION R. POOLE, AND MARVIN P. STRONG

The estimation of current highway and bridge needs along with the future performance and conditional effects resulting from alternative policies and funding scenarios is an important engineering function. This function was accomplished in North Carolina by using the statewide data base and analytical procedures in the Highway Performance Monitoring System (HPMS). Various study applications of HPMS have been tested to verify and identify performance characteristics, operational statistics, trend lines, and improvement needs over time. HPMS studies were also used to establish investment-performance relationships for each highway functional classification. Collectively, HPMS study results provide a general assessment of the condition, safety, and service components for the North Carolina system of roadways and bridges for any given period of analysis. HPMS study findings also provide an important informational source for highway administrators and decision makers responsible for policy evaluation and improvement program development. For that reason, an HPMS-based administrative status report of performance characteristics is scheduled to be provided annually to North Carolina highway officials.

The use of the Highway Performance Monitoring System (HPMS) to study and analyze North Carolina highway and bridge systems relative to their physical condition, operational characteristics, and needs requirements is described. Study result summaries can give officials an important factual basis for assisting the development of a cost-effective highway improvement program. HPMS roadway analyses have been performed annually since 1980 and have focused on the existing arterial and collector systems. Local roadways were not analyzed because the HPMS data base did not contain samples of local roads and streets.

For more than a decade, HPMS could estimate both current and future roadway needs on existing collector roads and arterial highways as well as the operational effects resulting from different funding levels and scenarios. It could not be used for analyzing and estimating bridge needs, conditions, and funding level effects. Recently, however, FHWA developed the Bridge Needs and Investment Process (BNIP), which can be used to estimate the physical condition, safety, and operation of bridges.

North Carolina was one of seven states selected by FHWA to test the usefulness of BNIP for conducting a statewide bridge system analysis. Provided elsewhere (*1*) are the details for that testing and a description of how BNIP analytical procedures were subsequently used to perform needs and investment analysis of North Carolina bridge data as contained in the National Bridge Inventory (NBI) file. The BNIP study projected bridge conditions and needs from 1988 through 2000. Deficiencies were identified, improvements selected, and improvement costs estimated to maintain the 1988 level of service over the analysis period. Study summaries were constructed for the federal-aid and functional classification systems.

DATA SOURCES

Data files that serve as input to HPMS and BNIP analytical processes are available from FHWA. State highway agencies and FHWA developed these files jointly and are responsible for their integrity, currency, support, and maintenance (*2,3*).

HPMS

The HPMS inventory files contains two types of data. They are usually referred to as "universe data" and "sample section data."

Universe data define the extent of roadway mileage by functional system and jurisdiction. Over and above the universe data are the sample section data, which are routinely collected on randomly selected sections from the complete arterial and collector highway systems. The sampled highway sections are spatially fixed over time and are homogeneous relative to highway geometric characteristics.

Individual sample sections are selected in accordance with the *Highway Performance Monitoring System Field Manual* and collectively provide the physical and operational data file from which the performance of the highway system can be evaluated (*2*). It is the only data file required as input to the HPMS models. State highway agencies are responsible for the annual update of the HPMS sample section data file (*2*).

BNIP

The NBI file is the primary source of data required by the BNIP. In addition to the NBI data file, the BNIP requires

---

Department of Highways, North Carolina Department of Transportation, P.O. Box 25201, Raleigh, N.C. 27611.

traffic growth rates and *K*-factors from the HPMS in order to predict bridge conditions and deficiencies.

The NBI file contains records for each bridge structure in the United States; it is kept current with data reported biannually by the states. NBI contains all state highway bridges and culverts but no tunnels. More than 16,000 highway structural records are contained in the file for North Carolina, and each record contains data that can be analyzed for structural, functional, and conditional deficiencies. State highway agencies are responsible for the biennial update of the NBI (*3*).

METHODOLOGY

HPMS

Several HPMS models may be used to perform analyses required for a particular highway study. These models have been designed to analyze the sample section data file and establish relationships between various levels of capital investment and the resultant performance of the highway system.

Figure 1 shows a hierarchical listing of HPMS models that have been used in North Carolina for base year and investment-performance analyses. Study references provide for a complete documentation and description of these models, ranging from an overview of their use in yielding performance information to a detailed discussion of their analytical potential as policy planning tools (*4*).

BNIP

The BNIP contains several model types that may be used to perform analyses required for special bridge studies. The BNIP consists of a main computer program (BRIDANAL) and five subprograms (BRIDEDEV, BRIDNEED, BRIDIVST, BRIDDRCS, and BRIDCTAB) that are dynamically called by user-supplied options. Such user options determine the number and sequence in which the subprograms are called.

Base Year Analysis

- ■ Complete Index Analysis
- ■ Multiple Deficiency Analysis
- ■ Impact Analysis

Investment/Performance Analyses

- ■ Investment Level Analysis
    - ▲ Needs Analysis
    - ▲ Investment Analysis
- ■ Funding Period Analysis
    - ▲ Needs Analysis
    - ▲ Investment Analysis
- ■ Summary Analyses
    - ▲ Composite Index Analysis
    - ▲ Deferred Cost Analysis
    - ▲ Multiple Deficiency Analysis
    - ▲ Impact Analysis

FIGURE 1 HPMS analytical process hierarchy (*4*).

The subprogram BRIDEDEV must be called initially to accept input records from the NBI file and create the master bridge file called BRIDMAST. Other BNIP subprograms can analyze BRIDMAST for a particular time period and output a systemwide estimate of highway bridge needs along with the investment level required to address those needs.

Specific information relative to the function and philosophy of the BNIP is explained elsewhere (*5*), with an executive summary for that process (*6*).

DEFINITIONS

The performance of a highway system can be defined in terms of the safety, economy, and efficiency of the flow patterns resulting from the movement of people and goods on the system. A measure of highway performance is defined to be an indicator of highway service derived from the condition, usage, operation, and physical characteristics at a particular point in time (i.e., past, present, or future).

Important examples of highway performance measures are speed, volume-to-capacity ratios, pavement condition, roadway cross sections and alignments, system mileage and travel, accidents, and user costs.

EXTENT AND USE OF NORTH CAROLINA HIGHWAY SYSTEM

As stated previously, the basic function of any highway system is to provide for the safe and effective movement of people and goods. It follows, then, that a highway system must be planned and designed with that function in mind. Several measures of effectiveness (MOE) including safety, convenience, efficiency, economy, mobility, and accessibility should be monitored closely to determine if the system is optimally serving its purpose. Such indicators are now being monitored conveniently on the North Carolina highways by the HPMS.

Mileage and Travel

HPMS estimates of mileage and travel on North Carolina's 1989 highway system are presented in Table 1. Data in this table are stratified by the federal functional classification system (*7*); they show that there were 149,266,000 daily vehicle miles traveled (DVMT) distributed over 94,619 highway-mi. It should be noted that the distribution of travel is not directly proportional to mileage. For example, Table 1 shows that rural highway mileage is nearly 80 percent of the statewide total but only 60.3 percent of the total travel.

The percentage of rural and urban travel is expected to change in the next few years. It is anticipated that urban travel will increase and rural travel will decrease. These changes will be due primarily to the redefinition of rural to urban areas in the 1990 census. The HPMS is useful in tracking such jurisdictional changes in mileage and travel over time.

Performance Relative to Condition, Safety, and Service

Performance of the North Carolina highway system is related to many physical and operational characteristics. Some of

TABLE 1 North Carolina Highway System Mileage and Travel in 1989

Functional Classification	Miles	% of Total Miles	Daily Vehicle Miles Traveled (DVMT)	% of Total DVMT
Rural				
INTERSTATE	621	0.7	16,297,000	10.9
OPA	2,061	2.2	15,881,000	10.6
MA	2,047	2.2	8,985,000	6.0
MAJ COLL	10,473	11.1	29,107,000	19.5
MIN COLL	9,058	9.5	10,191,000	6.8
LOCAL	51,140	54.0	9,606,000	6.5
Subtotal	75,400	79.7	90,067,000	60.3
Urban				
INTERSTATE	226	0.2	9,655,000	6.5
OF&E	213	0.2	6,198,000	4.2
OPA	1,732	1.8	24,338,000	16.3
MA	2,193	2.3	14,443,000	9.7
MC	1,370	1.5	3,224,000	2.1
LOCAL	13,485	14.3	1,341,000	0.9
Subtotal	19,219	20.3	59,199,000	39.7
Total	**94,619**	**100.0**	**149,266,000**	**100.0**

these characteristics are relatively fixed (e.g., lane width and alignment), and some can change rapidly (e.g., pavement condition). HPMS data that can be used to define performance are organized into the categories of condition, safety, and service. Condition data contain information on pavement type, pavement condition, and drainage adequacy. Safety data contain information on the adequacy of roadway cross section (i.e., lane, shoulder, and median widths) and alignment. Service data contain information on volume-capacity ratios and access control.

The HPMS analytical models provide information on highway system performance by measuring and reflecting changes in the system condition, safety, and service performance indicators that over time provide the factual basis for trend-line analysis.

Needs Estimate

The three interrelated variables of present conditions, future travel, and investment levels will determine future needs and conditions on the existing arterial highways and collector roads. These variables can be used in HPMS analyses as a basis for establishing investment-performance relationships. The HPMS user can tailor analyses for the evaluation of specific policies and situations by selecting appropriate minimum tolerable conditions (MTCs), construction improvement types, design standards, travel projections, and funding strategies (*5*).

Several types of analyses and studies have been performed on the North Carolina highway system; they include the estimation of needs, development of investment-performance relationships, and determination of future cost of travel as a function of investment.

Needs Determination

HPMS defines needs on the existing arterial highways and collector roads in terms of the funding level required to maintain a highway system equal to or above MTCs. Dropping below the chosen MTC values implies a state of deficiency. The level of funding necessary to correct all deficiencies as they occur is called full needs funding. The HPMS needs model determines full needs funding by first identifying deficiencies and then simulating the type and cost of capital improvements required to correct those deficiencies. Such funding level estimates are objectively based on a cost to maintain the highway system level of performance defined by MTC values.

TABLE 2 Other Principal Arterial Needs (1989–1998)

	Backlog (1989)		Accruing (1990-98)	
	Miles	Cost	Miles	Cost
Reconstruct to Freeway	107	281,904	300	856,357
Reconstruct w/more Lanes	0	0	0	0
Reconstruct w/wider Lanes	52	53,266	13	16,613
Pavement Reconstruction	1	417	0	0
Pavement Reconstruct w/Align	1	270	0	0
Major Widening(add lanes)	24	48,908	27	48,217
Minor Widening	12	6,225	0	0
Resurfacing w/Shoulder Improv.	47	12236	323	75593
Resurfacing	41	8879	327	73422
Resurf w/ Align & Shoulder Improv.	3	2027	76	62194
Resurfacing w/Align Improv.	2	1344	22	25424
Total	**289**	**415,476**	**1,087**	**1,157,820**

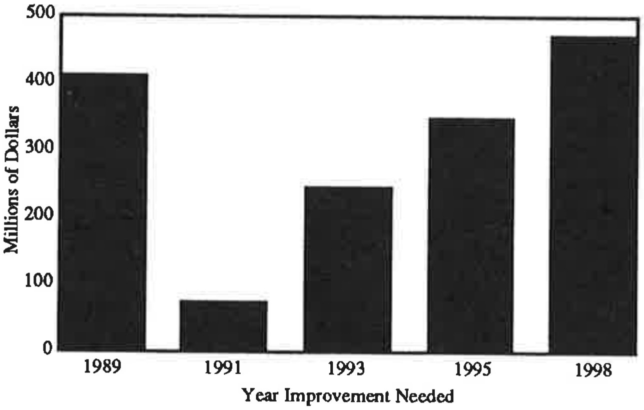


FIGURE 2 Rural OPA needs (1989–1998).

The needs model requires three major types of look-up tables. These tables contain MTC values, design standards, and costs for right-of-way and construction. System default values and standards are national averages. System default values are used for the North Carolina study.

An example of the output of an HPMS roadway needs study is presented in Table 2 and graphically shown in Figure 2. The costs shown are in 1989 dollars. It should be noted that HPMS needs assessment is accomplished without regard to revenue availability, user cost distribution, jurisdictional responsibility, or other subjective factors that actually determine highway program direction and investment levels. It should also be noted that the assessment of needs is the first step to be accomplished in investment-performance analyses.

Development of Investment-Performance Relationships

Highway investment-performance relationships output by the HPMS analytical procedure are based on theoretical and empirical research conducted by federal and state governments, AASHTO, and several leading universities (*4*). These relationships permit modeling of future highway system performance when given investment patterns and levels, future travel estimates, and applicable design standards. Such investment-performance models can be developed for each functional class system from the HPMS investment-level analyses.

The two types of investment-performance analyses that HPMS can accomplish are known as investment level and funding period. Both types were used in the North Carolina study. Investment-level analysis was used to determine total highway needs on the existing arterial and collector road systems and to estimate base year conditions and vehicle performance impacts. Funding-period analysis was used to forecast target year conditions and vehicle performance impacts.

It should be noted that either base or target year conditions can be analyzed by the impact model but that target year conditions and impacts can be analyzed only during a funding-period analysis (*4*).

The HPMS investment-level model simulates seven funding levels ranging from full needs investment to no investment at all. The full needs investment (or 100 percent funding) level simulates highway system effects for all improvements selected by the needs model. In terms of full needs investment, the next six funding levels simulated have the respective percentages of 80, 70, 60, 40, 10, and 0. The lower funding levels simulate only a portion of total required improvements, depending on the relative amounts of funds available. The zero funding level simulates the effect of making no capital improvements during the analysis period, which can be a maximum of 20 years.

After completion of the investment-performance analysis, the HPMS models output seven coordinate points for each of the safety, service, condition, and composite curves. The shapes of these curves are discretely and uniquely determined by those point sets.

Investment-performance graphs are developed by plotting the composite index values versus the dollars allocated for each of the funding levels (i.e., 100, 80, 70, 60, 40, 10, and 0 percent). Such graphs are being used by the North Carolina Department of Transportation (NCDOT) to answer many highway programming and budgeting questions. For example, management may desire an estimate of how anticipated funding changes could affect the safety, service, condition, and composite performance indicators over time. The investment-performance models can be used to answer that question.

Investment-performance models have been prepared for each functional classification in the North Carolina highway system and will serve as the basis for evaluating the future impact of different funding strategies and improvement programs.

Highway User Costs as Function of Highway Performance

Changes in pavement conditions, traffic congestion levels, vehicle operating characteristics, and roadway geometry affect the costs of using the North Carolina highway system. Costs for using that system can be estimated from data output by the HPMS simulation procedures for vehicular operation (*8*). The procedures simulate the way that highway conditions affect vehicle performance.

HPMS simulation results are expressed as vehicle performance indicators, which include speed, fuel consumption, operating costs, emissions, and accident rates. With the exception of accident rates, all of the performance measures are summarized by vehicle type.

Vehicle performance indicators provide a flexible means by which cost of travel estimates can be compared for different highway program, policy, or investment strategy scenarios. For example, vehicle performance indicators can be converted to user costs units and (a) compared between base (or existing) and target (or forecast) years to obtain the effects of a single program over time or (b) compared at the target year with alternative improvement programs to obtain relative cost of travel differences.

For comparison purposes, the HPMS analytical procedures produce highway travel cost components for both base and target years. Costs components reported by these procedures are average travel speed, accidents, and operating costs. The respective measurement units for these components are miles per hour, accidents per 100 million vehicle-mi traveled (VMT), and dollars per 1,000 VMT. These units can be equated to the single unit of money and subsequently combined to yield an economic basis for comparing alternative highway improvement programs.

The task of assigning money values to the variables of travel time or accidental death and injury is usually very difficult and subjective in nature. This task must be accomplished, however, before conducting any analysis designed to yield total user cost differences among highway improvement program alternatives.

Future Costs of Travel as Function of Investment

From a cost/benefit perspective, the most interesting and useful study using the HPMS had the objective of estimating the future cost of travel on the North Carolina highway system as a function of funding level. The study objective was accomplished in two steps.

First, the vehicle performance measures were determined by simulation for the target year for each of three funding levels. The second step involved calculating the total cost of travel based on the vehicle performance measures as determined in the first step. The methodology used to calculate those costs is outlined in the appendix cited elsewhere (*9*); the study results are also reported elsewhere (*8*).

NORTH CAROLINA HIGHWAY BRIDGE CONDITIONS, NEEDS, AND INVESTMENT REQUIREMENTS

The BNIP was conceptually designed as an analytical tool that could assess the operation, safety, and physical condition of highway bridges at the state or federal level. North Carolina was one of several states chosen to field-test and evaluate the analytical potential for that system.

Computer runs thoroughly tested each type of analysis including base year, full needs, and investment level. Each testing scenario was developed by selecting specific user options relative to analysis type, investment levels, design standards, and MTCs. User control statements may be used to specify a particular type of analysis and optionally override system default values (*5*).

BNIP testing results were compared with output data from the North Carolina Bridge Management System (*10*). The bridge needs and condition trends output by the two systems were very similar.

Evaluation of the testing showed that the BNIP provides a powerful analytical method for determining existing or future highway structure conditions, estimating backlog needs, and predicting investment requirements on a statewide basis. The analytical results can be output for either the federal-aid or the federal functional systems. Therefore, for a given period of analysis, bridge conditions for the base year can be compared to those in the target year to provide an estimate of conditional effects for different funding levels. In addition, BNIP summary tables can be conveniently used for trend-line analysis.

The versatility and usefulness of BNIP for determining bridge needs and condition trends on the highway system was clearly demonstrated by the study. Subsequently, it was decided to use the system to perform a statewide structure investment analysis and study to determine the annual funding required to maintain the 1988 base structure conditions through the year 2000. The findings, conclusions, and recommendations from that study are given in another paper (*1*).

Condition of North Carolina Highway Bridges in 1988

A BNIP analysis was used to determine highway bridge conditions that existed in the 1988 inventory year. The subprogram BRIDNEED first determined the number of structural, functional, and conditional deficiencies from the input NBI file. BRIDNEED summarized the frequency of condition values for the deck, superstructure, substructure, culverts, retaining walls, and safe load capacity. Summaries of the number of structurally deficient bridges that are open to all traffic, load- or speed-posted, or closed to traffic are also provided by BRIDNEED. The subprogram BRIDCTAB tabulated the structure condition summary tables for either the federal-aid or the functional classification systems. Results of the 1988 BNIP base year analysis of conditions and deficiencies for North Carolina bridges are presented in Table 3.

North Carolina Highway Bridge Needs from 1988 to 2000

BNIP will determine future bridge needs up to 20 years. After identifying the structural, functional, or conditional deficiencies by use of the MTC method, other subprogram models will select an improvement type that is dependent on the deficiency category. For example, if a bridge is structurally deficient, then the improvement type selected to correct that deficiency is "replacement." If a bridge is functionally deficient, the improvement type selected may be "widening" or "replacement." The bridge improvement type for a conditional deficiency is "rehabilitation."

TABLE 3 1988 Bridge Conditions on North Carolina Highway System

Federal-Aid System

	Number of Bridges				Condition
Classification	STR DEF	FUN DEF	FUN ADE	Total	Index
Interstate	37	199	560	796	7.3
Primary	113	195	1,355	1663	7.0
Secondary	507	386	1,314	2,207	6.5
Fed-Aid Urban	133	132	604	869	6.7
Non Fed-Aid	3,075	4,080	3,274	10,429	6.1
Grand Total	**3,865**	**4,992**	**7,107**	**15,964**	**6.3**

Federal Functional Classification System

	Number of Bridges				Condition
Rural	STR DEF	FUN DEF	FUN ADE	Total	Index
Interstate	11	109	328	448	7.2
Prin.Arterial	25	61	656	742	7.2
Maj. Arterial	49	66	309	424	6.7
Maj. Collector	523	397	1,338	2,258	6.5
Min. Collector	870	327	667	1,864	6.2
Local	1,994	3,461	1,983	7,438	6.0
Total	3,472	4,421	5,281	13,174	6.2

	Number of Bridges				Condition
Urban	STR DEF	FUN DEF	FUN ADE	Total	Index
Interstate	26	90	230	346	7.3
Free/Expressway	27	63	213	303	7.1
Prin. Arterial	58	76	538	672	6.9
Min. Arterial	101	105	320	526	6.6
Collector	41	30	97	168	6.6
Local	140	207	428	775	6.3
Total	393	571	1,826	2,790	6.7
Grand Total	**3,865**	**4,992**	**7,107**	**15,964**	**N/A**

After selecting the improvement type, BRIDNEED estimates improvement cost, assuming unlimited funding. Finally, the model determines future physical conditions and operational characteristics as if the simulated structural improvement were to be made. The results from the analysis of bridge needs for the North Carolina highway system from 1988 to 2000 are presented in Table 4.

North Carolina Highway Bridge Investment Analysis from 1988 to 2000

A BNIP investment analysis can be used to determine the funding level required to simulate improvements (i.e., as identified from the needs analysis) necessary to accomplish a future level of service for the highway bridge system.

The subprogram BRIDIVST will evaluate the future physical condition and operational effects that different funding levels would have on highway bridge structures. BRIDIVST accomplishes such an evaluation by building on output from the BRIDNEED model and priority ranking the improvement as a function of constrained investment levels specified by the engineer/user. Based on user funding specifications and time horizon, the subprogram BRIDIVST will produce future bridge condition values summarized by either the federal-aid or the functional classification system.

Investment analyses have been performed to determine an annual funding level that would maintain the 1988 bridge conditions until 2000. Four analyses were performed over that period using the respective yearly investment levels of $35 million, $55 million, $85 million, and $110 million. Condition values resulting from each of those investment levels are graphically illustrated in Figure 3.

TABLE 4 Number of Deficient Bridges and Costs by Improvement Type and Year of Deficiency (In Millions of Dollars)

	Backlog		1989-1993		1994-2000		Total	
	No. of Bridges	Cost	No. of Bridges	Cost	No. of Bridges	Cost	No. of Bridges	Cost
RURAL								
Rehabilitation	3	0	31	4	278	70	312	75
Widening	398	30	14	2	17	2	429	34
Replacement	7,331	1,112	20	7	149	129	7,500	1,248
URBAN								
Rehabilitation	8	1	23	8	326	97	357	104
Widening	45	5	1	1	2	0	48	6
Replacement	749	280	13	8	101	59	863	347
Total	**8,534**	**1,428**	**102**	**29**	**873**	**357**	**9,509**	**1,814**

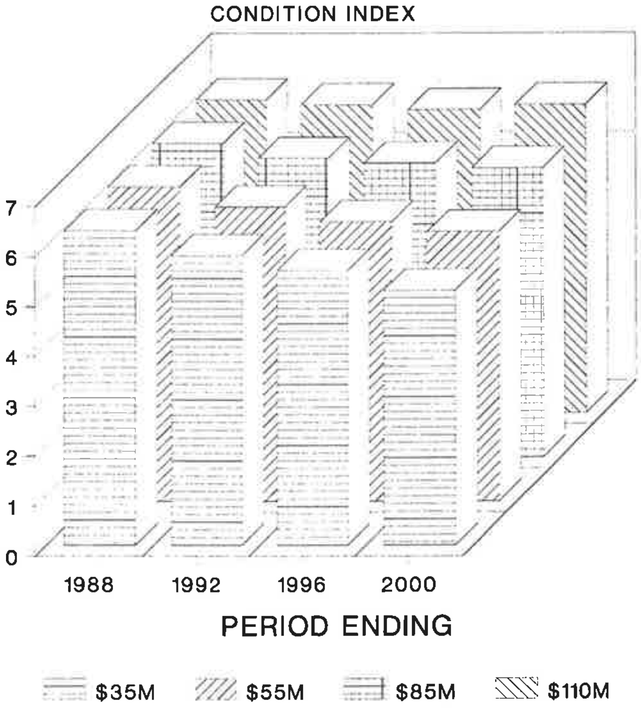


FIGURE 3 Bridge condition index as function of funding level (1988–2000).

ANTICIPATED USE OF HPMS DATA AND ANALYSIS PROGRAMS

The NCDOT business plan completed in 1990 recommended that the department produce a periodically updated report on statewide transportation needs. It also recommended that a committee be formed of representatives at the working level, drawn from participating groups, to make decisions on report format, report content, schedule, and participating group responsibilities. The committee would make decisions on modal and service coverage, standardization of forms and time periods covered, technical problems with information and data resources, and coordination between responsible groups.

In March 1991 NCDOT staff met to discuss the production of the annual report. Major conclusions from this meeting were as follows:

1. The annual report should be initiated as a report on the status of the highway system.
2. The HPMS analytical package had been run on the HPMS data set for each year beginning with 1982. This would provide a means for selecting key data to be portrayed in the report. The first report will illustrate what has been happening with the highway system during the 1982–1991 period.
3. Data from the bridge analytical package would also be incorporated into the report.
4. Data to be presented should be easily understood by lay persons.
5. The report should be short.
6. The Statewide Planning Group, Planning and Environmental Branch will have responsibility for producing the report. The research unit of the Planning and Environmental Branch will assist in developing the first report.
7. The first report should be completed by December 1991.

SUMMARY

There is a national recognition of the necessity to assess highway and bridge systems periodically relative to their extent, physical condition, efficiency, economy, and safety. In addition to such a general assessment, it is sometimes desirable to evaluate the effects that various highway programs and policies may have on those systems.

The HPMS analytical process has been used to assess needs on the existing arterial highways and collector road system and evaluate the effects that different levels of investment would have on the highway system for particular time horizons. However, bridges could not be analyzed in an analogous manner and provided some of the motivation for developing the BNIP.

The NCDOT was among the first to test the analytical capability of BNIP and evaluate its potential for performing bridge needs assessment and for predicting bridge conditional and operational characteristics as a function of constrained funding. Testing and evaluation showed that BNIP could provide a powerful yet convenient method for analyzing bridge conditions, operational characteristics, and bridge needs on a statewide basis.

The proper management of highway systems requires that administrators have current estimates of system conditions, performance, and needs. The decision maker should also be given estimates of the impact of different investment levels and strategies on the performance of roadways and bridges over time. It has been concluded that these two systems can provide an informational support basis to help management evaluate policy, analyze need, develop improvement programs, and allocate money to maintain optimally the North Carolina highway and bridge systems.

REFERENCES

1. *Determination of Bridge Conditions, Needs, and Investment Requirements on the North Carolina Highway System by Using the Bridge Analytical Process.* Planning and Environmental Branch, North Carolina Department of Transportation, Raleigh, Nov. 1988.
2. *Highway Performance Monitoring System Field Manual for the Continuing Analytical and Statistical Data Base.* FHWA, U.S. Department of Transportation, Jan. 1984.
3. *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges.* FHWA, U.S. Department of Transportation.
4. *Highway Performance Monitoring System Analytical Process*, Vols. I, II, III; Version 2.1. FHWA, U.S. Department of Transportation, Dec. 1987.
5. *Bridge Needs and Investment Process Technical Documentation and User's Guide*, Version 1.2. FHWA, U.S. Department of Transportation, Aug. 1989.
6. *Bridge Needs and Investment Process—Executive Summary.* FHWA, U.S. Department of Transportation, Aug. 1989.
7. *Highway Functional Classification: A Management Tool.* FHWA, U.S. Department of Transportation, Nov. 1982.
8. L. W. McPherson and M. R. Poole. Use of the Highway Performance Monitoring System To Determine Needs and Travel Cost on North Carolina Highways. In *Transportation Research Record 1156*, TRB, National Research Council, Washington, D.C., 1988.
9. *The Status of the Nation's Highways: Conditions and Performance.* Report to Congress. Office of the Secretary, U.S. Department of Transportation, June 1985.
10. *Bridge Management Systems.* Draft Report FHWA-DP-71-02. FHWA, U.S. Department of Transportation, 1988.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Transportation Network Design Using a Cumulative Genetic Algorithm and Neural Network

YIHUA XIONG AND JERRY B. SCHNEIDER

Currently available algorithms for finding optimal solutions to the discrete transportation network design problem are deficient in two ways. First, their computing time requirements are very large, which makes them infeasible for processing large networks. Second, they cannot process multiple criteria simultaneously—that is, only one objective value can be optimized in one run and, therefore, only one final solution can be obtained. A neural network in the optimal solution search process to replace the trip assignment algorithm for the computation of total travel time is employed. Before a neural network is used, it must be trained and tested with solutions obtained from a user-equilibrium trip assignment model. Experiments show that the trained neural network can predict total travel times quickly and accurately. Next, this neural network is used in combination with a genetic algorithm to search for optimal network designs. The original genetic algorithm did not work well for the problem. However, an analysis of its results suggested improvements that led to the creation of a very powerful search algorithm: the cumulative genetic algorithm. Experiments show that the cumulative genetic algorithm can seek and find system optimal designs extremely fast, using two criteria simultaneously. A full set of optimal solutions can be obtained to construct a trade-off curve for the two criteria. This trade-off curve, composed of optimal solutions, is the boundary of one side of the entire solution space.

The discrete transportation network design problem (DTNDP) involves the selection of new facilities (links) to add to a transportation network or to determine a set of capacity enhancements for some existing links so that the system performance and capital investment costs are optimal. Unfortunately, DTNDP is very difficult to solve because it is NP-hard (*1*). Currently available methods for finding optimal solutions are deficient in two ways. First, their computing time requirements are very large, which seriously limits the size of the network that can be processed (*2*). Some methods, such as branch-and-bound, are effective but can handle only very small networks. Heuristic algorithms can handle larger networks and have been used in many applications (*3*), but their computing requirements are still high. The final solutions are often locally but not globally optimal. Second, none of the currently available methods can handle multiple objectives. They try to minimize either the areawide total travel time on the network under a nonadjustable budget limit or the construction cost in relation to a specific total travel time goal. A third approach is to define first a trade-off relation between the total travel time and cost and then minimize the objective function defined as a weighted sum of the two (*4,5*). Using these methods, the overall relationships among the various performance values associated with the optimal solutions cannot be determined, because only one optimal or near-optimal solution can be generated. Many optimal solutions should be seen to obtain a general overview of the problem's solution space and to assist the search for a preferred network development plan.

These difficulties can be greatly reduced by using some new analytical techniques that have become available only recently. The first problem can be addressed using a neural network. An improved genetic algorithm that uses the neural network is used to address the second problem.

Previous studies have shown that the user-equilibrium trip assignment technique can produce a network flow pattern that matches actual trip flows quite well (*6*). Most network design studies have used this technique to compute the total travel times. However, because the total travel must be calculated for every possible solution encountered in the search process and because the solution space is extremely large, the assignment algorithm must be invoked again and again as alternative network designs are generated and compared with previous solutions. This multiple trip assignment computing requirement is a major bottleneck in the search for optimal designs for networks of reasonable size.

To address this problem, a short list of some possible solutions is generated, and a user-equilibrium trip assignment model [the Frank-Wolfe (F-W) algorithm] is used to calculate the total travel time for each of them. Then, this list is used as a training set to train a neural network. Because a neural network can respond much faster than the F-W algorithm, it replaces the algorithm in the network design optimization process. The tests of the trained neural network have shown that its total travel time predictions become very accurate after a reasonably long training period.

Next, using the neural network that had been trained and tested, an improved genetic algorithm is employed to conduct the network design process; it has been found capable of processing multiple optimization criteria and reaching a set of optimal solutions very quickly. By examining these solutions, the overall relationship between the total travel time and the project costs as well as the distribution of the optimal solutions in the solution space can be understood clearly. Thus, a better understanding of the basic properties of DTNDP and the identification of a preferred network development plan become possible.

Department of Civil Engineering, University of Washington, Seattle, Wash. 98195.

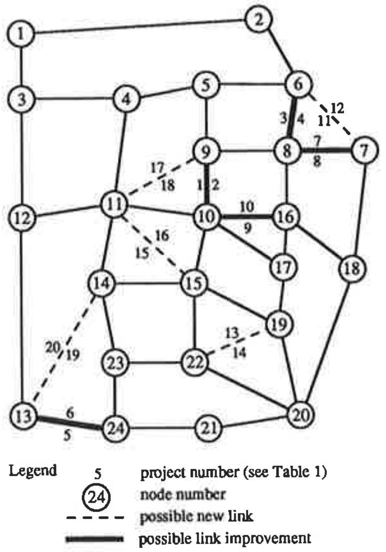


FIGURE 1 Test network.

DATA SET

Throughout this study, a synthetic network, with a fixed travel demand matrix, and 20 proposed improvement projects, each with a specific cost, have been used (Figure 1). This network has been used as the test network in several previous papers (*7,8*). The travel time along each link is computed using FHWA's travel time delay function. That is, if the traffic volume on the link l is v_l, its travel time from the starting node to the end node of the link l will be

$$t_l(v_l) = t_{0,l} \cdot \left[1 + 0.15 \cdot \left(\frac{v_l}{c_l}\right)^4\right] = a_l + b_l \cdot v_l^4 \qquad (1)$$

where $t_{0,l}$ is link l's travel time under free-flow condition.

TABLE 1 Link Parameters and Costs for Proposed Projects

Proj no.	Link	Parameters a(×10^{-2})	Parameters b(×10^{-5})	Cost (×\$1000)	New?
1	9-10	1.60	.0037	625	No
2	10-9	1.60	.0037	625	No
3	6-8	1.30	.1562	650	No
4	8-6	1.30	.1562	650	No
5	13-24	2.20	.2678	850	No
6	24-13	2.20	.2678	850	No
7	7-8	1.50	.0355	1000	No
8	8-7	1.50	.0355	1000	No
9	10-16	2.70	.3240	1200	No
10	16-10	2.70	.3240	1200	No
11	6-7	3.00	.0321	1500	Yes
12	7-6	3.00	.0321	1500	Yes
13	19-22	1.00	.0042	1650	Yes
14	22-19	1.00	.0042	1650	Yes
15	11-15	1.50	.0411	1800	Yes
16	15-11	1.50	.0411	1800	Yes
17	9-11	2.14	.0028	1950	Yes
18	11-9	2.14	.0028	1950	Yes
19	13-14	1.00	.0160	2100	Yes
20	14-13	1.00	.0160	2100	Yes

LeBlanc gives the network link parameters (a's and b's) and the travel demand between each possible node pair (*7*).

There are 20 proposed projects for this network: 10 are new link constructions and 10 are existing-link improvements. Table 1 presents their assumed (new) parameters and expected construction costs.

DEVELOPMENT OF NEURAL NETWORK TO REPLACE USER-EQUILIBRIUM TRIP ASSIGNMENT ALGORITHM

The neural network is a recently developed analytical technique that mathematically simulates the connections of the biological neural system in the human brain with respect to how it reacts to changes in the outside environment. A neural network can be trained by giving it some examples (inputs) and the corresponding responses (outputs) required. During the training process, the neural network's interior connections are adjusted so that the network can gradually predict the correct responses. The fundamentals of neural networks are described in several books (*9*). To obtain the total travel time for any network solution quickly, a neural network is used in the search process as a replacement of the trip assignment model. The neural network can respond much faster than the F-W algorithm. Experimental results show that the trained neural network performs very well in this role.

Neural Network Structure

To ensure a high level of accuracy, three layers in the neural network have been used. There are 20 proposed projects, so the neural network has 20 input variables, each of which can be either 0 (not constructed or improved) or 1 (constructed or improved). The first layer is the input layer: it distributes each of the 20 input variables to each neuron in the second layer. The second layer affects the neural network's capacity the most. Therefore, this layer includes 61 neurons, which may be more than necessary. And, since there is only one output (total travel time), the third layer (output layer) contains only one neuron and its output is just the neural network's output. Each variable in the first layer is an input to every neuron (except the last one) in the second layer, and each neuron's output in the second layer is an input to the neuron in the third layer. The last neuron in the second layer has no input; it is used to correct the bias on the output (*10*). Altogether, there are 62 neurons and 1,261 weights in this neural network (Figure 2).

In the neural network, if a neuron's inputs are x_i ($i = 1, \ldots, n$) and their corresponding weights are w_i, its output will be

$$z = \frac{1}{1 + \exp(-y)} = z(y) \qquad (2)$$

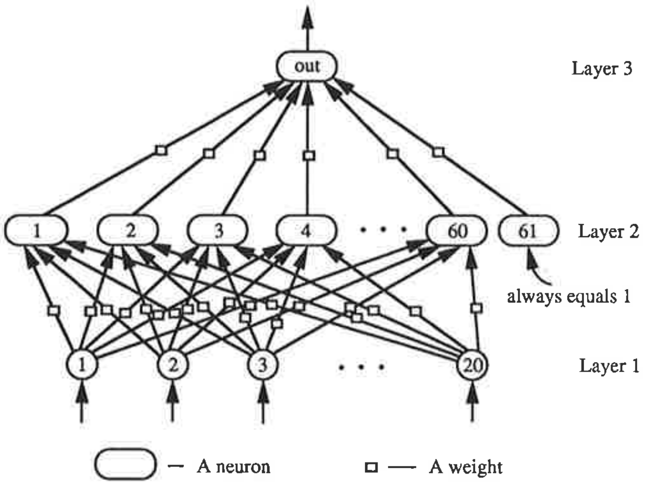


FIGURE 2 Neural network structure used to make total travel time predictions.

where

$$y = \sum_{i=1}^{n} x_i w_i \tag{3}$$

The output value of the neuron in the third layer is restricted to interval [0.25, 0.75], the linear part in the middle of the curve $z = z(y)$. The transformation relation between the output of this neuron and the total travel time it predicted is set as a linear mapping:

$$[0.25, 0.75] \leftrightarrow [t_{\min}, t_{\max}] \tag{4}$$

Here, $t_{\min}$ and $t_{\max}$ are predefined minimum and maximum possible values of the total travel time on this network. Although the interval [0.25, 0.75] and linear mapping are used here, it is not certain if they are the best selections.

Neural Network Training and Testing

There are 20 proposed projects, so there are $2^{20} = 1{,}048{,}576$ combinations. Each combination of projects is a solution (or a design) in the solution space. To train the neural network, 2,500 solutions were selected at random and the F-W algorithm was used to calculate their total travel times. The first 1,000 solutions were used for training, and the other 1,500 solutions were used for testing. Although 1,000 training solutions were used here, a 100-solution training set can produce a quite accurate neural network.

Note that there are two different errors involved here. First, the F-W algorithm is an iterating procedure. The stable value t for a solution can only be obtained after a large number of iterations. But if we use the algorithm directly in the network design optimization process, we can calculate only a few iterations and will obtain an approximate value, t_n. Therefore, the first error is computed as $e = (t_n - t)/t$.

A second error describes the performance of the trained neural network. Once a neural network has been trained, it must be tested to evaluate its predictive accuracy. If the value it predicts is t_{nn}, this error will be $e = (t_{nn} - t)/t$. Obviously, the neural network's prediction error depends on how accurate the training set is. Therefore, to decrease the neural network's prediction error, a highly accurate training set should be used. In this study, the error for each solution in the training and testing set has been reduced to almost zero (smaller than 0.001 percent) by using a large number of iterations (more than 25). These values, therefore, can be interpreted as being true values (t).

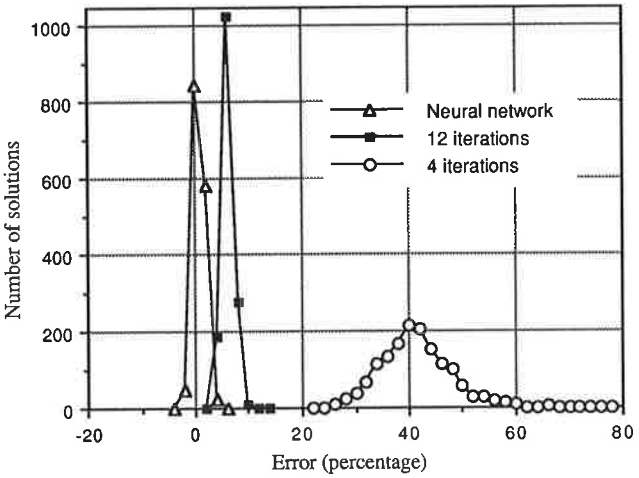


FIGURE 3 Neural network's prediction error versus F-W trip assignment algorithm's error.

After the training set was ready, the neural network was trained using the back-propagation algorithm, a supervised training program widely used in many applications (*9*). The training algorithm applies each solution in the training set to the neural network sequentially and adjusts its weights until its prediction error is acceptable. Next, this neural network was tested using the testing set. The results from the test are shown in Figure 3, along with two other plots that represent the errors obtained from the same 1,500 solutions but by using the F-W algorithm directly with 4 and 12 iterations. The left plot shows that the errors from the neural network are very small.

Training Time

Although the neural network's predictive errors are very small, it would not be worthwhile to devise it if the training time was longer than the time required to use the F-W algorithm in the search process directly. A neural network's training time depends on the neural network's complexity defined by the numbers of layers and neurons and their connectivity, the size of the training set, and the precision requirement to decide when to stop the training process. On an Apollo 4500 workstation (whose computing speed is much slower than a mainframe computer), we used 7,200 sec for training. If added with the training set generating time (3.1 · 1,000 sec) and the predicting time (1,200 sec, used in the genetic algorithm as described later), the total would be 7,200 + 3,100 + 1,200 = 11,500 sec.

Figure 3 shows that the neural network's prediction precision (standard deviation = 1.00 percent, mean = 0.75 percent) is equivalent to 12 iterations (standard deviation = 1.07 percent, mean = 6.10 percent) of the F-W algorithm. Experiments show that each 12-iteration computation needs about 2.03 sec on an Apollo 4500. If the trip assignment algorithm had been used in the genetic algorithm directly, the total computing time would have been 2.03 · 80,000 = 162,400 sec, which is more than 15 times longer than using the neural network. Here, 80,000 is the number of times the neural network was called by the genetic algorithm. Since the F-W algorithm computes very slowly, many previous studies have employed only four iterations (*11*), and the error (standard deviation = 6.69 percent, mean = 41.26 percent) for such results appears to be unacceptable.

The size of the neural network and its training time are related to the number of proposed projects, not to the transportation network size. This is a fundamental difference between a neural network and the conventional trip assignment algorithm. Therefore, neural networks may be especially useful in dealing with very large transportation networks that have only a few potential projects that need to be examined. This is often the type of problem encountered in practice.

DEVELOPMENT OF CUMULATIVE GENETIC ALGORITHM FOR FINDING OPTIMAL SOLUTIONS

A genetic algorithm is a general-purpose stochastic search technique applicable to a broad range of optimization problems. Its development has been derived from the study on natural evolution. During the search process, a genetic algorithm works from solution set to solution set simultaneously, not just sequentially—from one solution to another—as in conventional algorithms. Each solution set is called a generation, and population size is the number of solutions in the set. This generation-to-generation method enables the search process to escape a local optimum and eventually reach global optimum. There are three basic operations within a genetic algorithm: reproduction, crossover, and mutation. Once every operation has been applied sequentially to the current generation, a new generation consisting of offspring will be obtained. This new generation will replace the old entirely and become the current generation; the three operations will apply to it, and we will have another generation. In this repeating procedure, all generations are numbered sequentially and will be called the generation number: for example, the first generation, the second generation, and so forth. The interested reader is referred to Goldberg's book for a more detailed description (*12*).

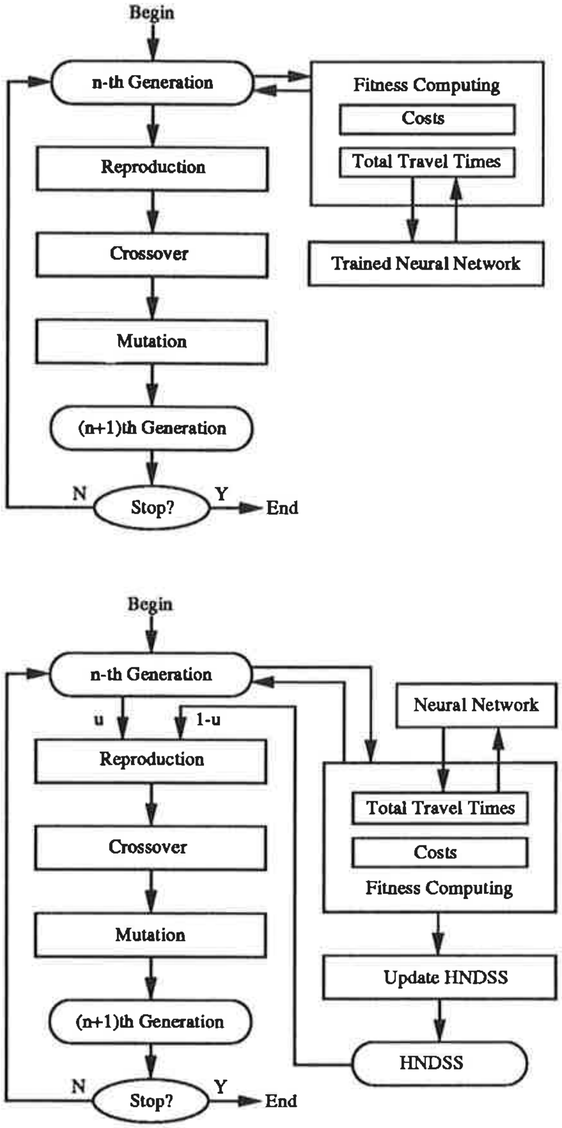


FIGURE 4 Genetic algorithms: *top*, original; *bottom*, cumulative.

Original Genetic Algorithm

A genetic algorithm that uses the trained neural network described previously was implemented on an Apollo 4500 [Figure 4 (*top*)]. Here, each solution is represented as a binary string with a length of 20, which is the same as the input variable format of the neural network. Each solution has a construction cost and total travel time, but in the genetic algorithm each solution can have only one objective (fitness) to represent its performance. To use these two values to define a solution's fitness, the domination comparison method was used. A solution v dominates solution w if v's performance values are not worse than w's for both criteria. If a solution is not dominated by any other solution, it is nondominated. For each solution in a generation, a count is made of the number of times that it is dominated by the other solutions in the same generation. The more times it is dominated, the lower its fitness value will be. Therefore, each solution's fitness value is calculated as

$$\text{fitness} = C - \text{number of times dominated} \tag{5}$$

where C is the largest "number of times dominated" among all the solutions in that generation. Obviously, the nondominated solutions will have a fitness C. It is hoped that the nondominated solution subsets within each generation improve as the evolution proceeds and that the subset of non-

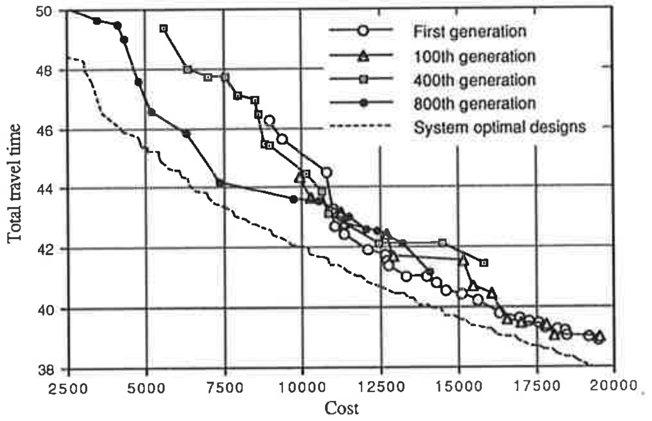


FIGURE 5 Solutions generated by original and cumulative genetic algorithms.

dominated solutions from the last generation will be the final optimal solutions.

However, this algorithm's performance was unsatisfactory. Figure 5 shows all the nondominated solutions in the 1st, 100th, 400th, and 800th generations, using a population size of 100. We can see that many "best" solutions in the 100th, 400th, or even 800th generation are even worse than those in the 1st generation. Even the 800th generation is not a clear improvement on the 1st generation. This tells us that the algorithm did not make progress during its search for optimal solutions. Given such results, it is unknown whether the computing process can find the optimal solutions or when the search process can be stopped.

It was determined that this poor performance was due to the fact that each generation passed some good chromosomes to its offspring, but not all of them. And, after a number of generations, some good chromosomes were found to have totally disappeared. To solve this problem, the original genetic algorithm was improved and a cumulative genetic algorithm (CGA) was devised. The solutions obtained by using CGA are clearly superior to the solutions from the original genetic algorithm (Figure 5).

Cumulative Genetic Algorithm

Transferring good chromosomes is just like transferring knowledge to children. The knowledge transfer cannot be expected to be based on oral communication only, because even if everything new is remembered, other things are often forgotten after a while. So, we often record our knowledge in book form. In this way, children will be able to learn and later review it if they forget. They can also modify the books by adding new knowledge or updating the old. Now, the children can learn either by reading the book or by asking questions of members of their parent's generation. These two options make the process of accumulating knowledge more efficient and faster; the book acts as a knowledge-accumulation device. In this algorithm, the book concept was implemented as a solution set called the historical nondominated solution set (HNDSS). At any specific moment during the search process, HNDSS includes all of the solutions that have never been dominated (as compared with all of the solutions generated by the algorithm up to that moment). This improved genetic algorithm was named the CGA; its operation procedure is shown in Figure 4 (*bottom*).

In CGA, the reproduction operation was modified so that it picks up solutions randomly not only from the previous generation (parents), but also from HNDSS. The probabilities of being selected for the solutions in the previous generation are proportional to their fitnesses, whereas the probability of being selected for every solution among HNDSS is equal. Then, after the reproduction, crossover, and mutation operations, if some better solutions are found in the new generation, the algorithm will add them to HNDSS and delete all old solutions that are dominated by the newcomers. When CGA stops, the solutions in HNDSS, instead of those in the last generation, will be the final results. Here another problem is encountered: how much time should be spent on books (HNDSS) and how much on communication with parents (previous generation)? Nine possible proportion values (u) have been tested using a population size of 100. Since each solution in HNDSS was contributed by a specific generation in the search history, we call its generation number the solution's contributing generation number. The term "average contributing generation number" is defined as the averaged value of the contributing generation numbers of all solutions in the final HNDSS. When $u = 1.0$; the algorithm becomes the original genetic algorithm, and, as shown in Figure 6, the average contributing generation numbers are closely related to the total number of generations carried out. This reveals that the solutions in HNDSS were almost evenly contributed by all generations and thus the process cannot be stopped after 800 generations. On the other hand, the best value is clearly located around 0.5. Its average contributing generation number is the lowest, which represents the shortest computing time. Also, the average contributing generation numbers are the same for both 800 and 1,600 generations. This means that the second 800 generations have contributed no solutions to HNDSS; all the optimal solutions have been generated by the first 800 generations. Therefore, $u = 0.5$ (half from the previous generation and half from HNDSS) has produced the best results for this problem.

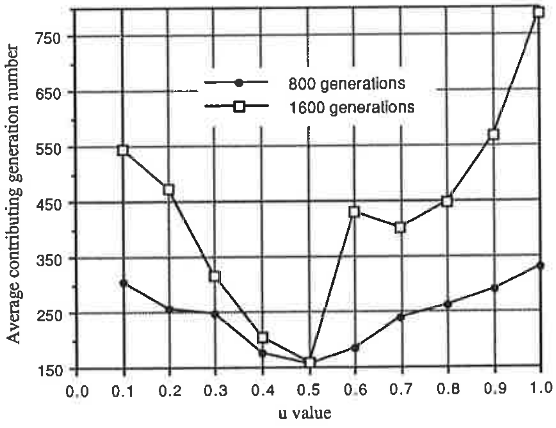


FIGURE 6 Contributing generation number for different *u*-values.

Population Size Versus Number of Generations Needed

To implement CGA for a DTNDP, there are still two parameters yet to be determined: population size (p) and the number of generations (g) needed. To find the most suitable parameters for our network, population sizes of 10, 20, 40, 100, 200, 300, 500, and 2,000 have been tested sequentially; their associated performance values are presented in Table 2. We can see that the p has almost no impact on the algorithm's performance and that g is inversely proportional to the population size (p). This indicates that for a DTNDP, the computing time ($g \cdot p$) is constant.

As shown in Figure 7, the distribution of the number of final solutions contributed by every generation matches a Poisson distribution well, with their means and medians inversely proportional to population size. This shows again that after a certain number of generations, most of the optimal solutions have been generated and further searching becomes unnecessary. For this problem, most of the final optimal solutions are contributed by the first half of the entire search process; that is, if the population size is p, 40,000/p generations is enough to obtain 95 percent of the optimal solutions.

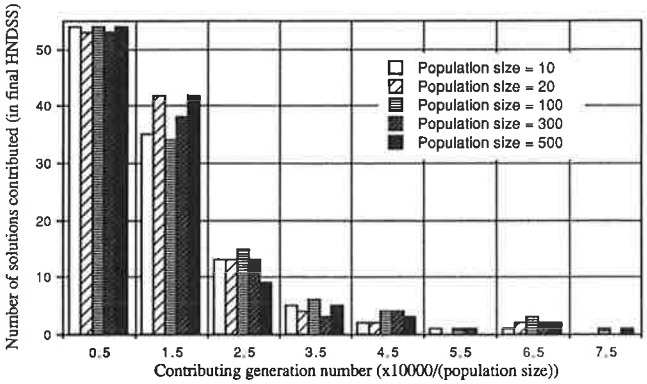


FIGURE 7 Final optimal solutions in HNDSS: number from every generation.

TABLE 2 Various Population Sizes and Their Contributing Generation Numbers Using CGA

Pop-size	Contributing generation numbers		
	Mean	Median	Std.dev
10	1335	1006	1108
20	685.4	542.5	568
40	373.2	308	313.6
100	156.9	116	144.3
200	64.1	41	60.28
300	46.5	34	41.65
500	27.5	21	25.61
2000	7.9	7	5.33

Final Optimal Solutions for Testing Transportation Network Using Two Criteria

Using CGA, 117 optimal solutions were found in HNDSS after 800 generations with a population size of 100 (Figure 8). Furthermore, additional runs were made using different population size values and the final optimal solution sets were found to be identical. This reveals that the final solutions obtained are real system optimums, not just local optimums like those generated by heuristic algorithms.

DISCUSSION OF RESULTS

If the original genetic algorithm had worked, a population size larger than the number of total optimal solutions would have been used, because the population size is fixed in the process and the last generation must be capable of containing all final optimal solutions. But, how can the number of op-

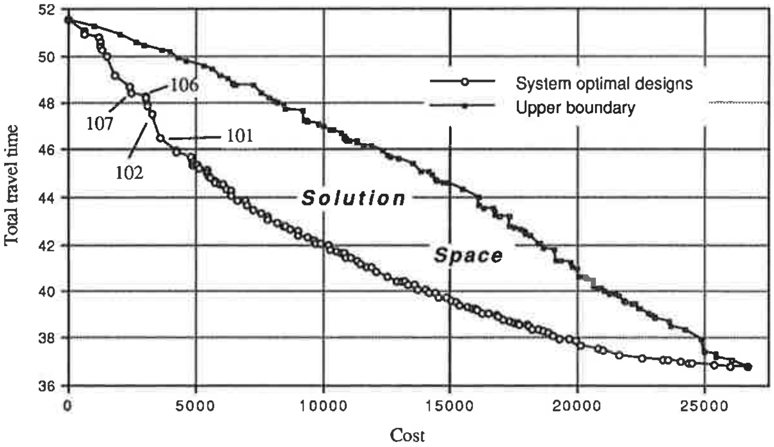


FIGURE 8 Final optimal solutions in HNDSS: location in solution space as its lower boundary.

timal solutions be known, even an approximate one, before the algorithm is run? CGA removes this question because the size of HNDSS is not fixed during the search process.

If those solutions that have the highest costs or longest total travel times can be found, a set of worst solutions using the same CGA with an inverted domination comparison can be obtained. These solutions form another curve, which shares two end points with the optimal solution curve selecting-no-project and selecting-all-projects. Together, these two curves form the entire boundary of the solution space (Figure 8).

The selection of a preferred transportation network design alternative is a complex process and will not normally be limited to only two criteria (total travel time and cost), although they are often the most important. If CGA is to be capable of processing more performance criteria, the domination comparison method can be modified and more performance comparisons included. If, for example, three criteria are used, the solution space would be a solid in three-dimensional space and the final optimal solution set will be a surface of this solid.

For the network design problem, an examination of the optimal solutions shows that the addition of some projects can reduce the total travel time more than others while costing less. These projects are the healthy chromosomes and are contained in most optimal solutions. Among the 117 final designs, Projects 15 and 16 both appear more than 100 times, and Projects 7 and 8 appear only 6 and 4 times, respectively (Figure 9). The addition of the healthy chromosomes will add vertical segments to the solution curve, and the addition of the bad chromosome will add horizontal segments. Therefore, the left part of the curve is more vertical because there are many "healthy" chromosomes left to select, and the right part is more horizontal because all the healthy chromosomes have already been included and only the bad chromosomes are left to select. Also, the curve is smooth in most but not in all locations. In Figure 8, Solution 101 is obtained by using Project 15 to replace Project 11 in Solution 102. Because Project 15 is the healthiest and Project 11 is relatively unhealthy, this replacement costs only $300,000 but saves 1.014 hr of travel time. But, from Solution 107 to 106, the healthy Projects 3 and 16 are replaced by two relatively unhealthy projects, 11 and 12. This replacement costs $550,000 but saves only 0.171 hr of time. Thus, the former replacement is represented by a relatively vertical line segment, and the latter replacement is very horizontal.

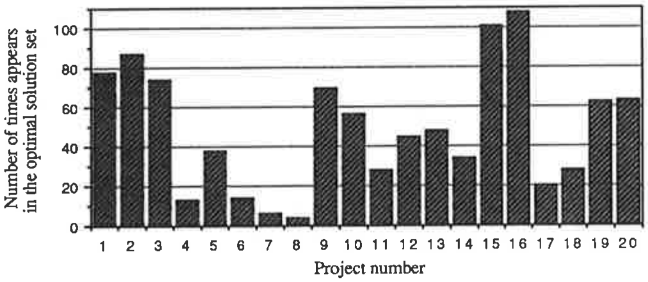


FIGURE 9 Number of times each project appears in set of final solutions.

CONCLUSIONS AND RECOMMENDATIONS

The neural network is shown to be a technique that can speed up optimization computing dramatically, especially when the accuracy requirement (compared with the user-equilibrium trip assignment algorithm) is not very high (e.g., an error of less than 2 percent). This method also establishes a way of using other trip assignment models in the network evaluation and design process. Since a neural network can be trained with results from the F-W trip assignment model, it is certainly possible to train it using results from other trip assignment models. However, from our experience, if the requirement for the predictive accuracy is very high (e.g., an error of less than 0.2 percent), the neural network's training time would become much longer.

Although the performance of the original genetic algorithm was not satisfactory, CGA worked very well for a DTNDP, even using two criteria. This is a step forward, as multiple criteria have not been used by previous algorithms, and a set of optimal solutions, not just one, has never been generated before. Using the results from CGA, it is clearly seen how the solutions are distributed on the cost-travel time plane for any transportation network. Further, the solutions generated are system-optimal. The trade-off curve defined by these optimal solutions clearly illustrates how one performance value changes when the other changes. This trade-off information for the entire range of potential solutions can greatly assist a decision-making process designed to identify a preferred alternative.

However, this study is just a beginning. It is recommended that further work be undertaken as follows:

1. Larger transportation networks of various shapes should be used to test the performance of the method, including their neural network training times and error characteristics, CGA optimization speed, and the shape of their trade-off curves.
2. Methods for coping with various types of constraints may be applied to the solution space, because constraints are very common in practice. (For example, it might be stated that Project X is not feasible unless Projects Y and Z are both selected, or Project A and B cannot be both selected).
3. Methods should be devised for including other criteria in the genetic algorithm (e.g., the link volume-capacity ratios) and incorporating them in the analyses of the results.
4. Further efforts should be made to define the basic properties of DTNDP in relation to the specific characteristics of a range of transportation network sizes and shapes.

REFERENCES

1. M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman and Co., San Francisco, Calif., 1979.
2. T. L. Magnanti and R. T. Wong. Network Design and Transportation Planning: Models and Algorithms. *Transportation Science*, Vol. 18, No. 1, 1984, pp. 1–55.
3. B. N. Janson, L. S. Buckels, and B. E. Peterson. Network Design Programming of U.S. Highway Improvements. *Journal of Transportation Engineering*, Vol. 117, No. 4, 1991, pp. 457–478.
4. B. N. Janson and A. Husaini. Heuristic Ranking and Selection Procedures for Network Design Problems. *Journal of Advanced Transportation*, Vol. 21, No. 1, 1987, pp. 17–46.

5. B. H. Immer and P. H. Mijjer. Optimization of Transport Networks. Presented at 20th International Scientific Conference on Transport Planning and Traffic Engineering, Budapest, Hungary, April 1989.
6. D. M. Chang and G. B. Dresser. *A Comparison of Traffic Assignment Techniques.* Research Report 1153-3. Texas Transportation Institute, Texas A&M University, College Station, 1990.
7. L. J. LeBlanc, E. K. Morlok, and W. P. Pierskalla. An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem. *Transportation Research*, Vol. 9, No. 5, 1975, pp. 309–318.
8. H. Poorzahedy and M. A. Turnquist. Approximate Algorithms for the Discrete Network Design Problem. *Transportation Research*, Vol. 16B, No. 1, 1982, pp. 45–55.
9. P. D. Wasserman. *Neural Computing: Theory and Practice.* Van Nostrand Reinhold, New York, N.Y., 1989.
10. P. K. Simpson. *Artificial Neural Systems: Foundations, Paradigms, Applications and Implementations.* Pergamon Press, New York, N.Y., 1990.
11. S.-I. R. Tung and J. B. Schneider. Designing Optimal Transportation Networks: An Expert Systems Approach. In *Transportation Research Record 1145*, TRB, National Research Council, Washington, D.C., 1987, pp. 20–27.
12. D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Reading, Mass., 1989.

Publication of this paper sponsored by Committee on Transportation Programming, Planning, and Systems Evaluation.

Diffusion of Transportation Planning Applications in Metropolitan Planning Organizations: Results of National Survey

J. Scott Lane and David T. Hartgen

The University of North Carolina at Charlotte conducted a survey of metropolitan planning organizations (MPOs) throughout the United States to determine the extent and causes of diffusion of transportation applications in present practice. The survey quantified the current use and plans for the expanded use of these applications, documents the diffusion process, and shows how innovation is related to funding, creativity of managers and employees, agency independence, and other factors. The analysis also discusses the implications of these trends for shifts in the power structure of transportation planning. The results show that MPOs are very computer-literate, having purchased an average of eight access points for $44,000 over the past 5 years and planning even more purchases in the near future. Common hardware is IBM and IBM-compatible; commonly used software includes spreadsheets, word processors, and data base management. Specialized packages for transportation modeling are also commonly used. MPOs cite a lack of funding and computer knowledge as the key obstacles and improved agency efficiency and user demands as the key motivators to computer innovation. MPOs use primary contacts with peers and staff as the key data-gathering mode for system information. Large agencies have adopted systems 2 years earlier—on average—than other agencies, but this appears to be related to funding and knowledge constraints. No evidence was found that the characteristics of agency managers influenced the adoption of computer systems. It is concluded that rapid diffusion of computer technology through MPOs will fundamentally change the balance of power in and between planning agencies, opening up the process and encouraging cooperative technical analysis at many levels.

There are approximately 321 metropolitan planning organizations (MPOs) in the United States. Each MPO serves as the primary forum for transportation planning efforts in the area that it serves. While performing these functions, the MPO must provide leadership to, and cooperate with, city and county planning groups in its service area. The existence of these agencies is not only crucial to the structured development of an area, but also mandated by law. When an area becomes urbanized (urbanized being defined by a variety of characteristics, but all having the trait of having at least 50,000 inhabitants residing in the main population concentration), federal regulations stipulate that an MPO be formed to help manage the growth of the area.

Like many other government institutions, these MPOs in large part operate—with the exception of an occasional professional conference or workshop—in a near vacuum with regard to their peers. Little has been done to consider these organizations as a national group. Recently, the University of North Carolina at Charlotte (UNC Charlotte) surveyed the transportation activities of MPOs across the country, focusing on planning applications and software.

The university was interested primarily in the transportation planning sector of the MPOs. Often, transportation in a growing area is the first part of the infrastructure to feel pressure from the local citizens, media, and political bodies. Planning for and responding to these changes has become a data- and model-intensive prospect for any group that wishes to examine transportation patterns and growth in an urbanized area. The techniques that transportation planners use have not changed much, even though the demand for analysis has grown in response to the growth of population and commerce. Fortunately, computer technology has developed more rapidly and has kept up with the demands of the MPOs. The major problem with the new technology, or so it was theorized, was that the tight budgets of many MPOs were limiting their access to the new technology. This appears to be particularly true for the smaller organizations, which might also be hampered by another resource constraint, that of a lack of knowledgeable staff to operate the increasingly powerful models. The ultimate goal of this research was to learn more about the nature of computer technology, its use, and its growth inside MPOs across the United States.

REVIEW OF PLANNING APPLICATIONS TECHNOLOGY

The past several years have seen several studies and reports reviewing recent computer applications technology in transportation planning. TRB's Task Force on Transportation Planning Applications, initially formed to encourage dissemination of techniques, has sponsored two major conferences on this topic and plans a third conference in Dallas. The first, held in Orlando, Florida, in 1987 (*1*), focused primarily on early versions of urban transportation planning systems (UTPS) software and their applications. The second, also held in Orlando (*2*), displayed a healthy familiarity with emerging geographic information system (GIS) procedures, an expanded market penetration of UTPS-type models, and a near-universal awareness of spreadsheet-based systems. Plans for the third conference anticipate significant advances in these and other areas.

J. S. Lane, South Central Regional Council of Governments, North Haven, Conn. 06473. D. T. Hartgen, Department of Geography and Earth Sciences, University of North Carolina, Highway 49, Charlotte, N.C. 28223.

General descriptions of software have become quite commonplace. McTrans has continued this tradition, originated by the U.S. Department of Transportation (DOT), of periodically publishing a software and systems guide; the latest, in looseleaf form, covers more than 200 commercially available and public-sector applications (*3*). McTrans's newsletter fills in during interim periods (*4*). This newsletter has recently taken over a transit-oriented service operated by Vanderbilt University. From the University of Kansas, PC Trans publishes a periodic newsletter focusing on traffic applications (*5*).

Despite the wide availability of software, surprisingly few comparative assessments of software have been published or conducted—at least for private consumption. In an early study, the New York State DOT reviewed the characteristics of about a dozen UTPS-type systems (*6*). A later assessment by Michigan described the characteristics of systems but "hid" the brand names (*7*). ITE has recently undertaken to develop a criterion checklist of software selection features, also brandless (*8*). Essentially, agencies interested in learning about software characteristics in an objective atmosphere have no readily available source. Instead, they must rely on vendor information or word of mouth.

DESCRIPTION OF SURVEY INSTRUMENT

A survey about transportation planning and computer technology was sent to all 321 MPOs in the United States. The mail-out survey was three pages long (the first page was an introduction). The following pages questioned the respondent on such topics as

- Characteristics of the MPO (size, service area),
- Functions of the MPO,
- Characteristics of computer systems and their users,
- Plans for future systems, and
- Characteristics of the manager of the MPO (years with the agency, computer expertise).

Response to the survey was considered to be quite good: nearly half (154) of the agencies responded. Ninety percent of the questionnaires were returned within 40 days after mailing. The surveys were answered with varying degrees of completeness, the vast majority being completed entirely.

The following documents the results of an analysis of data gleaned from the survey. No statistical methods more complex than standard deviations are used, but the results do show some interesting contrasts among MPOs, their staffs, and their current (and future) state of computer involvement.

FINDINGS

Size

In any discussion of the capabilities of a transportation department or MPO, the organization's size will almost invariably be brought into play: the larger the organization, generally, the more functions it serves. MPOs are thought to vary in size in a direct relationship with the size of the population that they serve. The study performed by UNC Charlotte used several indicators of MPO size: (a) population of the planning area, (b) number of employees at the MPO, and (c) the amount of funding that they received in FY 1990. Table 1 illustrates the size characteristics of the sample MPOs.

Two of the three measures presented in Table 1 vary quite widely across the sample: the size of the service population and the amount of funding. The average number of employees did not vary nearly as much. This can be seen by the standard deviation of the three variables: whereas population and funding had standard deviations that were more than two and three times greater than their means, the number of MPO employees had a standard deviation that was only half again as great as the mean. This suggests that although the assumption stated earlier, that MPO size (at least, as measured by the number of employees) follows the size of the service population, is correct, the relationship follows a fairly short logistical curve. The size of the MPO in terms of employment does not increase proportionately as the population of the service area rises.

Functions

The number of functions that the transportation department of an MPO can perform in-house—versus those that are often farmed out to consultants or other public agencies—is a measure for determining that MPOs range of capabilities. As is shown in Figure 1, not all transportation departments are equally equipped to perform some of the more infrequent or computer-intensive functions that are normally associated with such an agency. This does not mean that an MPO that must farm out much of its work ends up with worse results than a product done in-house; indeed, the opposite can be true. Figure 1 does show that transit system design and environ-

TABLE 1 Size of MPOs

RESPONDENTS	1980 POPULATION*	NO. EMPLOYEES	FUNDING, 1990
SUM	67,879,278	1,148.0	$59,171,705
AVERAGE	484,852	7.5	$402,529
STANDARD DEVIATION	1,606,562	11.1	$1,020,866

**Source:* Transportation Planning Data for Urbanized Areas

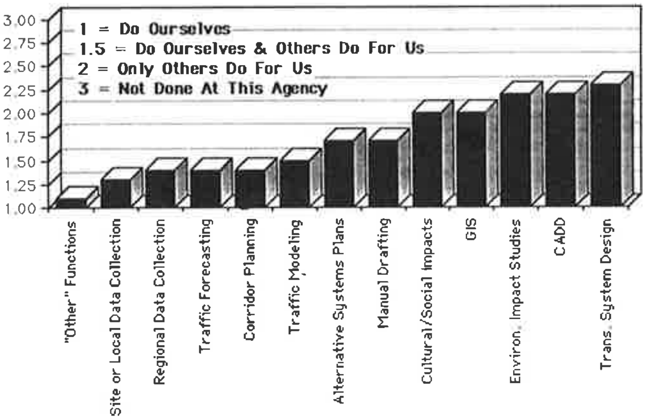


FIGURE 1 Degree of autonomy.

mental impact studies (federal EIS and EA) are performed less frequently at the MPO than data collection, corridor planning, and traffic forecasting. MPOs view the former as requiring skills that demand a higher degree of specialized knowledge or more legwork than would be practical or economical to have on hand constantly.

The various roles filled by the MPO and the emphasis each MPO places on them will necessarily affect some of these functions. All MPOs do not fulfill the same roles in every urbanized area; some are called on to do more as their skills and reputation for a certain type of project improve. All the MPOs that we surveyed said that they performed as area MPOs (Figure 2), which is to be expected. As MPOs, it is also expected that they would engage in site planning and be involved in other transportation planning efforts; about 50 percent of the respondents included these in their repertoire of functions. Traffic operations and construction of transportation facilities were activities of very few MPOs, only 17 and 2 percent, respectively. Computer technology has done more than probably any other single factor to allow smaller MPOs to perform more functions and play a larger, more diversified role in their areas.

Computers perform a variety of roles in the MPO, from the mundane, like generating a report, to the complex and power-hungry transportation modeling effort that are required of today's modern transport system analysis methods. Making all of these devices work, and work well together, produces a whole new set of problems and frustrations for the staff of an MPO. These agencies, however, would be severely crippled if not inoperable without a variety of computer systems.

One of the items that concerned the researchers was the inroads that various computer systems have made into the MPOs in recent years. Conversely, it was believed that some types of computer technology were being abandoned, such as mainframe computers. The graph in Figure 3 shows the types of systems that are being used in MPOs today. IBM and IBM-compatible machines are by far the most common type of system used, with 840 terminals being used in the 154 sample MPOs (5.5 computers per agency). However, other brand names are making a dent in the IBM dominance in microcomputing. There are 171 micros being used in the sample that are not IBM or IBM-compatible. The gap between IBM and its competitors is still quite large, but if the survey had been taken only 10 years ago, it might have been much greater. Local-area networks (LANs) have also increased. IBM mainframes and minicomputers are not nearly so numerous as their desktop counterparts. Only 70 workstations were present in the sample, and just 41 terminals for mainframes and minicomputers from companies other than IBM were recorded. The power and advantages of flexibility of microcomputers versus macrocomputers are shown clearly in Figure 3. It is very probable that within a short time, the desktop computer will compare favorably with the mainframes in power and number-crunching capability. When this happens, there will be another plunge in the use of mainframe computers that will leave only the largest and most complex of urban areas still needing mainframe capabilities.

Of course, computer technology comes at a price, and only those agencies that can allocate sufficient funds toward the acquisition of computer systems will have access to new technology. Table 2 presents the amount of spending that MPOs have concentrated on computer systems in the past 5 years, as well as the amount they expect to spend in the next 5 years. More than $6 million was spent on computer hardware and software from 1985 through 1990 in the sample of MPOs that were surveyed. The average computer expenditure is expected to increase almost two-thirds in the next 5 years ($29,474

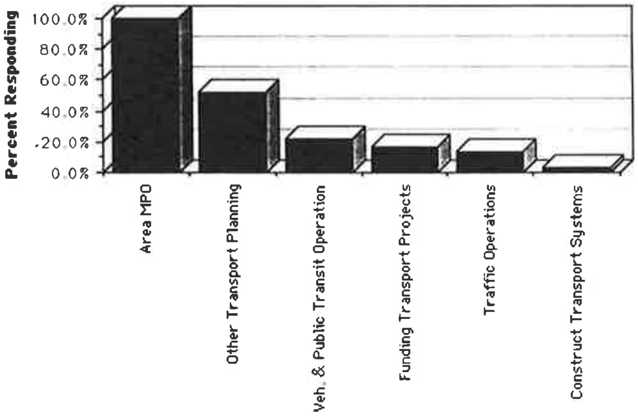


FIGURE 2 Agencies' roles in transportation planning.

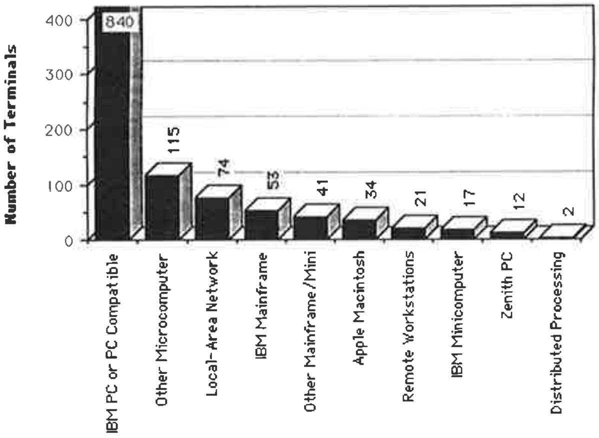


FIGURE 3 Types of systems in MPOs.

TABLE 2 Present and Future Computer Expenditures

RESPONDENTS	COMPUTER SYSTEM ALLOCATIONS	NUMBER OF ACCESS POINTS
	LAST FIVE YEARS	NOW
SUM	$6,177,900	1,136
AVERAGE	$44,128	8
	NEXT FIVE YEARS	IN FIVE YEARS
SUM	$3,713,700	1,383
AVERAGE	$29,474	10

added to $44,128). But the number of access points, a surrogate for all computer hardware (i.e., terminals, workstations, and desktop computers), will increase only by about 20 percent in the next 5 years. This indicates a relatively advanced state of commitment to computer technology in the MPOs. The majority of the initial funding was focused toward buying basic hardware, such as workstations and printers. The dramatic increase in spending will occur because future acquisitions will be geared toward acquiring more software and more efficient hardware. Both software and hardware will be bought as the needs of an expanding area population and development demand them.

What were the initial factors that contributed to some of the $6 million of MPO computer spending? What were those factors that might have hindered such spending in the past? Figures 4 and 5 address these questions. These two figures show both primary and secondary considerations when computer purchases are made.

Improving the efficiency and productivity of the MPO and its staff was the number one answer given for the primary contributing factor in computer purchases. Demands made by the staff of the MPO followed closely. Strangely, demands created by software use were the foremost secondary reason given for computer purchases. The highly touted claim of lower prices in the computer realm was not rated as high as one might have supposed.

Although increased staff computer expertise and available funding seemed to be of only moderate importance in contributing to computer purchases, these two issues were by far the greatest obstacles for most of the MPOs that were surveyed. Most of the MPO managers (93, or 60 percent) claimed that a lack of funding was the primary reason that kept their organizations from purchasing new computer systems. The second most important factor was the absence of trained computer operators. The next three most important hindering factors were obstruction by governing bodies, limited amounts of time and money available to train personnel in computer operation, and the unwillingness of top management personnel to invest in computer technology. Several of the MPOs claimed that a *loss* of productivity would result if new computer systems were installed. Compatibility with existing computer systems was seldom a hindering factor, which is some-

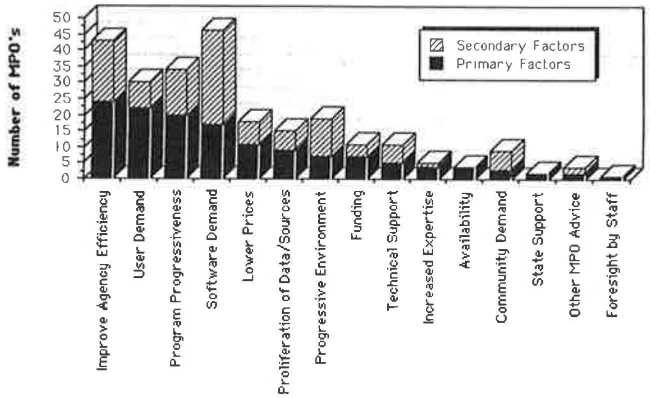


FIGURE 4 Factors contributing to computer purchasing decisions.

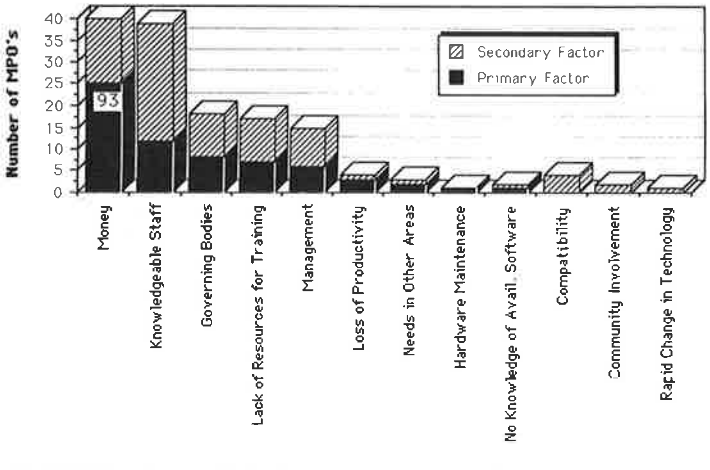


FIGURE 5 Factors hindering computer purchasing decisions.

what surprising, but it should be encouraging to those agencies considering new computer system purchases.

Figures 6, 7, and 8 show the most popular brands of the most frequent computer applications in the transportation sectors of MPOs. Word processing is the most-often-used computer application in almost any organization; competition between different manufacturers of word processors is intense. The top two word processors used by MPOs are WordPerfect and Wordstar; the former is a relatively new entry into the market, and the latter is one of the very earliest word processors to achieve a wide popularity. WordPerfect is by far the more commonly used of the two, with about half of the respondents citing WordPerfect as their word processing software. Lotus 1-2-3 enjoys a similarly wide lead over its nearest competitor in the computerized spreadsheet comparison shown in Figure 7. Ninety-six of the respondents stated that they used the Lotus 1-2-3 spreadsheet. The second choice was Quattro Pro, a recent option in the field of computerized spreadsheets. Again, of these top two contenders, one is a well-established product, and the other is a new one. Unlike the situation with word processors, the older program is the more popular.

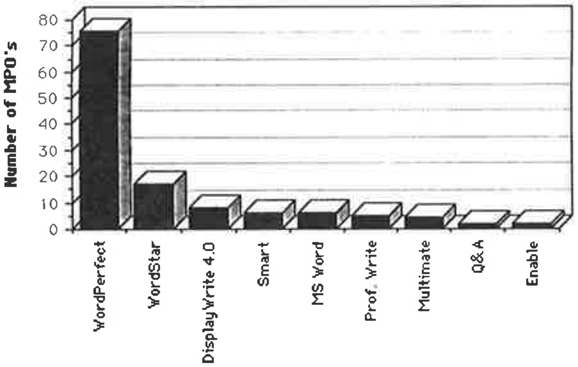


FIGURE 6 MPOs' commonly used word processors.

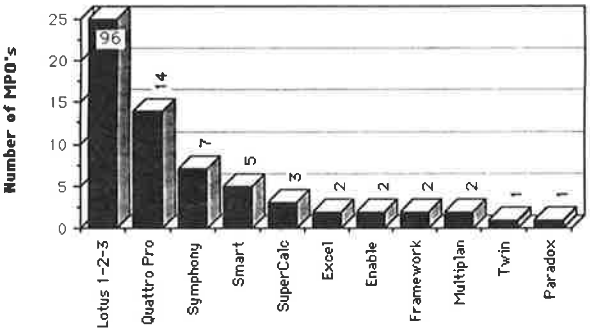


FIGURE 7 MPOs' commonly used computerized spreadsheets.

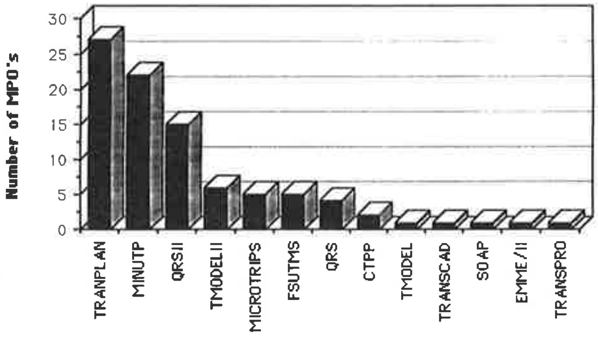


FIGURE 8 MPOs' commonly used microcomputer transport models.

Three microcomputer-based transportation simulation models dominate the MPO markets: TRANPLAN, MINUTP, and QRS II own a commanding lead in this area over their nearest competitors, TMODEL II and MICROTRIPS (Figure 8). Earlier versions of two of these programs, QRS and TMODEL, were also given as answers. Table 3 presents the types of software found most frequently in the MPOs and the average year that they were acquired. Each software type is ranked by the number of employees that work with that software. As mentioned earlier, word processors and computerized spreadsheet programs are very popular and there are quite a few knowledgeable users in the MPO. There are an average of 7.5 employees per organization in the sample, and about 7 of these know how to use word processors and spreadsheets. Data base management, statistics, and project management programs are also used by a large percentage of the people employed at MPOs. Conversely, site-planning software and traffic-flow models are familiar to fewer employees. Interestingly, all but one of the current software programs were acquired, on average, between 1985 and 1988. UTPSs were acquired, on average, in 1977—a full 8 years before the next earliest pieces of software (either financial tracking or statistics programs, acquired in 1985). This venerable modeling tool is being replaced rapidly by the microcomputer-based systems mentioned earlier.

Many issues should be resolved both before and after the purchase of new computer hardware or software systems. Often, the effective resolution of these issues means more to their MPOs than the specific type of software or hardware or the dollar amounts spent on them. The researchers asked MPO managers how well they believed their organization was performing on specific issues relating to their computer systems. Two of these issues—operating and training personnel to operate new systems and financing needed improvements or updates in computer applications—were repeatedly ranked quite low by the surveyed MPOs (Figure 9). Again, the issue of acquiring funding is seen to be a source of frustration for those who know that computer technology is important to their organizations. Further, it suggests that more attention should be given to training personnel and to finding innovative ways of financing new technology. Identifying and priority ranking computer needs was not seen to be a problem for many of the MPOs, nor were they dissatisfied with their overall performance in developing computer applications. Some dissatisfaction was apparent with the way in which older, outdated systems are replaced or updated.

Another item that influences the decision-making process in public (and private) agencies is how information about various computer systems is gathered. Other agencies, liter-

TABLE 3 Description of Software Used by Transportation Division in MPOs

SOFTWARE SYSTEMS	AVG. YEAR OF ACQUISITION	AVG. NUMBER OF USERS	% MENTIONING (PENETRATION)
SPREADSHEET	1986	7.0	89
WORD PROCESSOR	1987	6.8	83
DATABASE MANAGEMENT	1986	4.3	76
MICRO-BASED U.T.P.S.	1988	2.0	48
HIGHWAY CAPACITY MODELS	1987	2.1	45
FINANCIAL RECORDS/ACCOUNTING	1985	2.8	38
G.I.S.	1988	2.4	34
C.A.D.D.	1988	2.8	29
STATISTICS PACKAGES	1985	4.1	26
SITE PLANNING SOFTWARE	1988	1.8	23
TRAFFIC FLOW MODELS	1988	1.7	21
MAINFRAME UTPS	1977	3.3	12
PROJECT MANAGEMENT	1988	3.7	10
TRANSIT OPERATIONS	1987	2.0	9

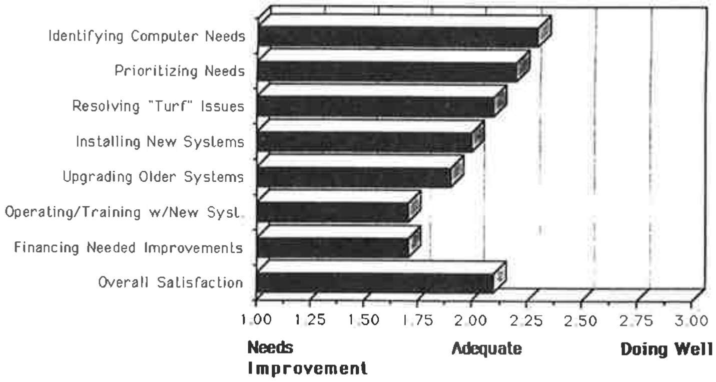


FIGURE 9 MPO performance in developing computer system applications.

ature reviews, and word-of-mouth are some traditional methods of gaining insight to computer markets. Figure 10 shows how the surveyed MPOs ranked various contacts that provide information about computer technology. Personal contacts and correspondence were ranked the most useful method of getting information about hardware and software. This was closely followed by word-of-mouth with planners and other peers, experts and agency "champions" inside their own MPO, and contacts at professional conferences. A very strong element of personal, informal contacts is a common factor with each of these methods. Four supply-side sources of information were ranked as less important to the MPOs: user groups, trade shows, consultants and vendors, and trade magazines. Since the cost of these systems is considerable, it is interesting to note that the directors of agencies put the most faith on the experiences of peers when they are getting information about new computer applications. As noted earlier, there are few comparative, up-to-date sources of such information.

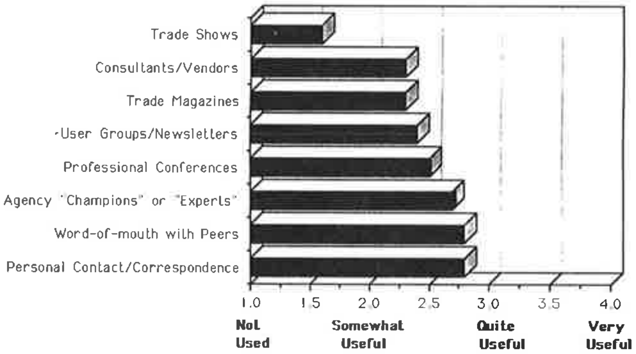


FIGURE 10 Usefulness of information sources for computer technology.

Regardless of what method the MPO used to gain knowledge about new applications, when the time came to actually purchase the new systems, the agencies responded that a formal review process was the most instrumental factor in the decision. Compatibility with existing equipment was ranked low on the list of factors hindering computer development, but it was the second most highly regarded method used in computer procurement. Similarly, a review of pertinent literature on the subject of computer hardware and software, rated low on a list of ways of getting information about new systems, ranked high in the procurement process. Only two agencies responded that no primary method was used to decide on computer purchases.

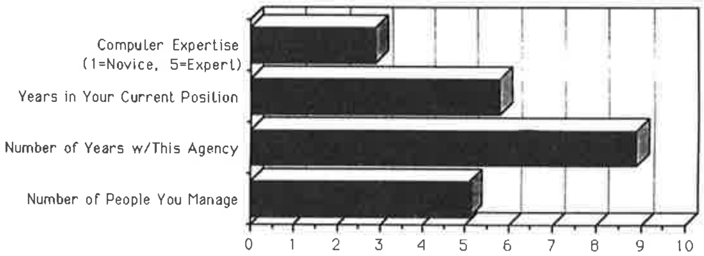


FIGURE 11 Average MPO manager characteristics.

Manager Characteristics

The manager of MPO activities will usually but not always play the most important part in decisions about computer technology. Figure 11 examines the characteristics of MPO managers. Most managers thought that they were slightly better than average in regard to computer expertise. They have spent an average of nearly 9 years with the same MPO, and more than 5.5 years in the same position in their agency. Because the average number of employees in an MPO is about 7.5, it is not surprising to see that managers supervise an average of 5 employees.

The researchers hypothesized that managers' length of employment at the same agency might negatively affect the willingness of the managers to adopt new technology quickly and the computer expertise of the managers themselves. When the managers' length of employment was cross-tabulated with these two variables, these hypotheses were found to have no support in our study sample. Less than a half-year of agency employment separated the computer novices from the computer experts. Similarly, the date of adoption of new computer systems was found to have little relevance to a manager's stay of employment at the same agency. This means either that a long-time manager varies very little in computer attitudes from a new person at the same post or that MPO managers have little to do with the actual procurement of computer systems in their agencies.

POLICY IMPLICATIONS

This study suggests a number of important policy implications for transportation planning:

1. These findings depict, in total, a remarkable diffusion of computing technology through a population of agencies in just 7 years. That diffusion is still proceeding, but the penetration rates for both hardware and software in conjunction with substantial past and planned investments show that the MPOs as a group have seized microcomputer technology with a firm grip.
2. Microcomputer software for transportation planning has now become indispensable to this trade. Its widespread use has created a de facto standard: no credible analysis can be complete without the use of such tools.
3. Rapid changes in software, particularly specialized software, can be expected to evolve further the complexity of planning. However, good transparent techniques are likely to emerge victorious.
4. Large agencies that previously held near monopolies on planning analysis can expect to see their status and power eroded. Smaller MPOs will continue to follow the lead of their larger partners, breaking dependence on mainframes and their operators. Towns and smaller counties, in turn, can be expected to develop their own analysis capabilities in parallel with their parent MPOs. With that capability will come disagreement over assessments and forecasts (stemming from different assumptions) unless agencies and analysis groups develop cooperative ways to share information and findings.
5. With planning the shift of importance from the federal government to state and local bodies in areas of hands-on activities with communities will occur. The federal government is effectively out of the modeling business, having been supplanted by the private sector. Its roles now—trainer, fund support source, and regulator—will intensify, leaving the analysis field to local governments.
6. There may be a need for promulgation of "good practice" manuals on such topics as modeling. In earlier times, such techniques were taught routinely for mainframes; now they are needed even more as the proliferation of procedures makes such practice more difficult to define.
7. Our findings roughly parallel those of an earlier study of the state highway departments (*10*) that found that agency funding was the primary deterrent to computer literacy and that larger agencies were able to adopt the technology more rapidly than smaller ones. Managers in smaller MPOs need to recognize the relationships between computer literacy and agency performance and act to ensure that their staff has the requisite equipment and skills to take advantage of these trends.
8. In the future, intensive computer use can be expected for LANs and distributed processors. But MPOs are not likely to adopt such systems in great numbers, because the nature of computing in those agencies does not require joint use of large systems planning tools. Minicomputers and workstations are not likely to replace the micro, even though the micro will become more powerful.
9. Perhaps surprisingly, a remarkable base of objective information or consumer reports was found regarding the experiences of agencies with most of these systems. There has, apparently, been no assessment of the characteristics or features of these systems. The apparent reliance of managers on word-of-mouth and agency champions in system selection is disturbing: let the buyer beware.
10. The trends in hardware appear to be toward IBM and IBM-compatible equipment, toward LANs, and away from other machines. Software trends appear to be consolidating

toward a handful of UTPS models, supported by spreadsheet and word processing systems. Students would be well advised to be familiar with these procedures. Given the investments in software training as well as data files, these locked-in behaviors will be hard to shift. Newcomers will have to offer transparent adaptability and a very significant cost or functional advantage in order to compete. GIS-based software may provide one avenue for change.

SUMMARY

The computer has changed its role in the MPO—from one of a gate that controlled doable functions to one of a conduit that expands capability. It continues to be one of the cutting edges of change in urban transportation planning, pushing out toward a more open, distributed, decentralized decision-making process. Let us hope that our institutions are up to the test. Let's make sure our software is neither Caesar's Rubicon nor Hannibal's impedimenta.

REFERENCES

1. U.S. Department of Transportation. *Proc., 1st Conference on Transportation Planning Applications*, Orlando, Fla., April 1987.
2. University of North Carolina at Charlotte. *Proc., 2nd Conference on Transportation Planning Applications*, Orlando, Fla., April 1989.
3. *Software Sourcesbook*. McTrans, University of Florida, Gainesville, 1990.
4. *PC Newsletter*. McTrans, University of Florida, Gainesville, 1990.
5. *PC Trans-Mission*. Transportation Center, University of Kansas, Lawrence, 1990.
6. *Transportation Software*. Phase I report. New York State Department of Transportation, Albany, 1988.
7. *Microcomputer Traffic Forecasting Model Evaluation*. Michigan Department of Transportation, East Lansing, 1989.
8. *ITE Committee Proceedings*. ITE, 1990.
9. *Traffic Forecasting Guide*. New York State Department of Transportation, Albany, 1987.
10. J. S. Lane and D. T. Hartgen. Factors Affecting Adoption of Information Systems in State Departments of Transportation. In *Transportation Research Record 1271*, TRB, National Research Council, Washington, D.C., 1990.

Publication of this paper sponsored by Committee on Transportation Programming, Planning, and Systems Evaluation.

Starting a Regional Transportation Planning Organization

NEAL D. FLOYD, TRACY NEWSOME, DAVID DRYMON, DAVID T. HARTGEN, AND H. E. TIMMERMAN

As urban areas grow, they eventually fill up the central core cities and often subsume neighboring towns and cities that were previously separate. Cross-border travel patterns thus defy service at the municipal scale, requiring instead that large regions containing many jurisdictions work together. The process is described by which the jurisdictions of the 13-county, 100-mi region surrounding Charlotte, North Carolina, formed a new superregional transportation planning organization, larger than metropolitan planning organizations and counties, across two states. The emerging need for such agencies is reviewed, and the Charlotte organization—the Carolinas Transportation Compact (CTC)—is described in detail. A strong project-oriented work program, a neutral host, advocacy instead of operating roles, local funding, and galvanizing issues are necessary to success. The future of CTC is bright because it has developed solid complementary working relationships with state and local governments and provides a forum for the pursuit of cooperative regional solutions to regional problems.

As many urban areas grow and expand, they fill up the area included in the traditional urban transportation planning process and often subsume neighboring areas that were previously separate towns and cities. The new region, bigger and more complex, requires more complex spatial planning at the regional scale and introduces greater complexity into the political context. It also provides an opportunity for the area to address transportation problems in a truly comprehensive way, because it brings the planning process into balance with the multicounty environment used by citizens and businesses in their daily travel patterns. In a region in which one metropolitan area dominates the landscape, the political and economic structure may vary significantly from one part to another. The economy and dynamics of such a region will generate transportation problems, as surrounding towns and even the central city become interdependent parts of the larger metropolitan region. Citizens commute across political boundaries without second thought as to what effect this has on the economy and on the transportation system. But the political structure of the region may not be as developed as the economic structure. This problem can be resolved if there is a sound regional planning organization that can act according to the pressures and problems caused by a growing multigovernmental region. The intent of this paper is to describe the processes that led to a large superregional transportation planning organization in the Charlotte, North Carolina–South Carolina, area. Discussion will center on the case study of Charlotte and the efforts made and methods used to coordinate a regional transportation planning organization.

EMERGING GROWTH PATTERNS

The provision of transportation facilities and services on a regional scale is important for several reasons. Economic growth and the corresponding travel generated from this growth do not generally respect geographical or political boundaries. Across the United States, most recent growth has occurred in areas outside the traditional control of the central city, resulting in a diverse regional pattern of development. These regional development patterns are both a by-product of and a cause for our current transportation systems. The widespread use of the automobile in the post–World War II era as well as an advanced highway system has made it possible for people to live many miles from their jobs. Therefore, transportation systems have increased the American society's mobility and have promoted increased, dispersed development patterns throughout the region.

Dispersed growth, as much as we appear to like it, can create special problems. Land use patterns are less dense and make most types of public transit service impractical. Suburban land use patterns have also made increased automobile use necessary, which has added to the pollution, traffic congestion, and energy consumption of the nation. Much of the nation's regional growth has occurred in the neighboring counties of central cities. This decentralization of the urban population has become a dominating force shaping the urban region in the post–World War II era. Growth in the outlying rings has accelerated, but growth in many central cities has stabilized or declined. Between 1940 and 1970 the growth rates of central cities and suburban fringes have differed by as much as 20 percent. Between 1960 and 1980 the differences in this growth assumed a regional pattern. In Northeast and North-Central regions, central city losses have increased and suburban growth itself has begun to slow. In the West, growth in both areas has increased but is now beginning to slow. It is in the South where this decentralization in the metropolitan area has continued at a strong pace. The differences in the outlying and central city growth in the South are a function not only of increases in the number of metropolitan counties but also of net migration balances between central cities, ring or edge communities, and nonmetropolitan counties. In the South, the flows in and out of the central city and urban areas were more heavily biased to suburbs than were other regions (Table 1) (*1*). This spillover has brought pressure to create

---

Department of Geography and Earth Sciences and the Urban Institute, University of North Carolina, Highway 49, Charlotte, N.C. 28223.

TABLE 1 Percentage Change in Populations: Center City and Suburban Ring by Region, 1960–1980

	Percent Change in Population			
	1960-1970		1970-1980	
Region	Center City	Suburban ring	Center City	Suburban ring
Northeast	-1.9	19.2	-7.3	13.5
North Central	-0.3	26.8	-4.4	23.3
South	11.2	35.6	18.2	69.0
West	18.0	37.1	19.3	34.5

Source: US Bureau of Census

public planning entities that are larger than those of the traditional county.

Transportation planners have, of course, been aware of the regional transportation phenomenon for almost as long as the automobile has been in existence. In 1916 Congress passed the Federal Road Act, which mandated that each state establish a highway department to rank and choose transportation projects to receive federal funding. During the 1930s and 1940s Congress often debated the need for intercity connectivity. The 1956 Federal Aid Highway Act established the Interstate Highway System. The roots of transportation planning go back to the early 1900s, but regional multimodal transportation planning, in the urban-suburban-rural sense, was first mandated by the Federal Highway Act of 1962. This act required that metropolitan areas of 50,000 or greater establish by July 1, 1965, a "coordinated, comprehensive, and continuing" transportation planning process for the urbanized area and the surrounding land area likely to become urbanized. This act mandated local and state governments to begin cooperative transportation planning functions for an entire urban area, which usually meant planning for more than one jurisdiction. During the early 1960s, the forerunner of the metropolitan transportation planning organizations were established, and planning activities were undertaken for almost all of the more than 250 urbanized areas. For those areas in which the metropolitan area encompassed more than one state, cooperative ventures were formed. The 1973 Federal Highway Act was the start of federally mandated metropolitan planning organizations (MPOs). MPOs were to be designated by the governor of each state for metropolitan areas to "perform metropolitan planning" (*2*). The MPO was envisioned to cover a growing urban area and to deal with the problem of traffic congestion on an areawide rather than geopolitical basis.

This organizational structure served quite well through the 1970s and into the 1980s. But in the late 1970s regional growth patterns began to change in many ways. A rapid development of rural areas and cities outside the urban core was propelled by a healthy economic environment, lower land values at the fringe, and increasing interstate transportation access. Much of this growth occurred in the Southeast and Southwest. The space inside large metropolitan regions began to fill as regional development spread across the area. Metropolitan superstrips, covering several hundred miles in some cases, evolved on a relatively large scale. Even "rec-opolises" began to emerge, containing high levels of activity for certain periods of the year, underpinned by a narrowly based economy dependent on recreational development. Each of these changes accelerated demands on the transportation system to handle the travel it generated and put pressures on the organizations responsible for managing and planning for change. Congress realized that economic growth and transportation congestion do not respect political boundaries. A 1980 publication from the U.S. Department of Transportation states: "Ideally, the MPO is in a position to coordinate the various elements of the transportation system to shape orderly development of the metropolitan area" (*2*). This report went on to say that the MPO has failed to reach its ideal because it has no statutory operating or construction authority.

The 1990s are likely to see continued evolution of the regional nature of transportation planning. Pressure to expand current urban boundaries (based on the 1990 census) include more participants and undertake more cooperating planning balanced by declining real funds for planning, a possible refocus of the federal rôle away from urban areas, and increasing competition between the suburbs and the city. In many areas, disagreements between factions have fragmented the cooperative base of planning; in the New York City metropolitan area, the region has dissolved its regional planning structure in favor of several subregional structures focusing on state and county portions of the area. These tensions imply that the evolution of a regional system of transportation planning is by no means ensured in many areas, perhaps not even likely in some.

REGIONAL PLANNING ORGANIZATIONS

Various kinds of regional planning agency exist today. At one end of the spectrum are the large superagencies performing region-building and operating functions for multicounty areas or corridors. Examples are the Portland (Oreg.) and Washington (D.C.) metropolitan transportation authorities. The Washington Metropolitan Area Transit Authority (WMATA) was formed in 1967 to plan, finance, and construct a rapid transit system for Washington, D.C. (*3*). Its makeup included elected officials from Washington and the neighboring states of Maryland and Virginia. Later, in 1973, WMATA gained operating authority for the transit system as well as power to acquire and operate bus services, so that now it operates the whole region's system. WMATA has no dedicated fund source from the states; its revenues come from fares and concessions and from local government contributors. In Portland the regional operator, TRIMET, is responsible for service provision in a three-county area, and funds come from a payroll tax. TRIMET also evolved from a transit bus operator to build a transit line. The regional planning agency expanded its functions to become METRO, a regional authority responsible for land use planning and solid waste disposal, separate from TRIMET. Another one of the first regional agencies was the Metropolitan Transportation Commission, formed in 1970 by the California Legislature for the San Francisco Bay Area. It is responsible for overall transportation planning for the area and is currently empowered to review and approve the allocations of transportation funds for projects.

One of the more common forms of an areawide planning organization is the MPO. These groups are responsible for general-purpose transportation planning for the urbanized area surrounding the central city. Most MPOs cover just this area, but a few—primarily in New York City, Arizona, Massachusetts, and Connecticut—cover larger regions of the state (K. E. Heanue, unpublished data, 1989). Because federal aid and requirements had declined from 1978 to 1983, MPOs were becoming increasingly isolated from decision making, more special-purpose funded, less comprehensive, shorter range in focus, and under strong pressure to subregionalize rather than to broaden areas of coverage (*4*). MPOs may try to be regional planning organizations, but in most cases they do not have the power or authority to function as such.

CASE STUDY: CHARLOTTE, NORTH CAROLINA–SOUTH CAROLINA, AND THE CAROLINAS TRANSPORTATION COMPACT

Charlotte and Region

Charlotte, North Carolina–South Carolina (NC-SC), provides a good example of a region in which growth (economic, population, infrastructure, etc.) is not limited to the boundaries of one city or even one county. Charlotte and Mecklenburg County are the center of the Charlotte–Gastonia–Rock Hill metropolitan statistical area. The Charlotte metro region, however, consists of 13 counties that include both North and South Carolina (Figure 1). The core city, Charlotte, is the largest city in the two Carolinas and serves as the major business hub for the two states (*5*). Charlotte has a 1990 population of more than 400,000, ranked 37th in the nation. Mecklenburg County has a population of 511,433 (*6*). First Union Corporation and North Carolina National Bank (NCNB) are headquartered in Charlotte, making it the sixth-largest financial center in the nation. NCNB's merger with C&S Sovran has created NationsBank, the nation's fourth-largest bank; its headquarters is in Charlotte. In 1987 Charlotte was the sixth-largest wholesale trade center (in terms of sales) in the country. The region has a population of approximately 1.6 million, which ranks 28th out of all other metropolitan areas (*5*). Population in the region grew by 15.9 percent between 1980 and 1990. Table 2 illustrates the growth of the Charlotte, NC-SC, region by county. As shown in Figure 1, the central Mecklenburg County core is ringed by cities that have their own MPO planning process (Rock Hill; Gastonia; Concord/Kannapolis; Hickory) and other developing cities such as Monroe, Albemarle, Salisbury, Statesville, and Shelby.

The substantial economic vitality that this region has experienced has come tied not just to the city but to the region, which stands poised to grow (or fail to grow) as a whole. Many large activity sites and projects depend on promoting the region, not just Charlotte. Many facilities and organizations serve the region, including Charlotte/Douglas International Airport, the National Basketball Association's Charlotte Hornets, the Charlotte Motor Speedway, the Charlotte Coliseum, a possible National Football League stadium and franchise, museums, orchestras, and lakes and recreational facilities. In addition many issues are receiving attention at a regional level: economic development, water and lake systems, and transportation.

One of the prime reasons for the growth of the Charlotte region comes from its location at the crossroads of two major Interstates (I-85 and I-77) and from the proximity of I-40. Much of the growth enjoyed by the region occurred during the automobile era, with all the attendant characteristics of such: urban development spreading into surrounding suburbs, growth of ring cities, and development of previously rural areas. These land use characteristics encouraged continued reliance on the private auto for travel. In-commuting to Charlotte is substantial because of the size and centrality of the city to the region. There are 90,000 to 100,000 net in-commuters who come into Mecklenburg County each day. Interstate travel volumes approaching Charlotte range from 50,000 to 70,000 vehicles daily, and traffic is increasing rapidly. Traffic problems in and around the Charlotte area have become a major issue facing not only Mecklenburg County but all of the region: a *U.S. News and World Report* article, quoting an FHWA study, stated that Charlotte would be the most congested urban area in the United States by the year 2005 (*7*). Even though the statement was later retracted, Charlotte's reputation suffered; traffic problems helped to defeat the incumbent mayor's reelection in 1987. Vehicle miles traveled are expected to double between 1985 and 2000. Officials say it will be physically impossible to build all the necessary roads to handle the traffic growth, even if money were available. Even with the apparent traffic problems in Charlotte, it is recognized that the problem has regional proportions. There is congestion caused by commuters along main arterials among the surrounding towns, cities and counties, and the city itself (Figure 2).

The region has a good radial highway system in most directions but suffers from a lack of circumferential service. Traffic congestion exits throughout the entire region but is extremely heavy in the south and southeast areas. Improvements under way and planned on key Interstates, coupled with other planned highway improvements, should ensure that the highway system will perform well into the next century. The region's role in international and intercity travel is also substantial. Charlotte/Douglas International Airport is one of the busiest in the nation. The Charlotte region has an excellent rail system that provides overnight freight and passenger service to much of the eastern seaboard.

Transportation access is a critical factor in determining economic growth. A recent assessment of growth rates in North Carolina shows that those counties with the highest accessibility are those that are growing the most rapidly. Thus, a major issue in the Charlotte region is how to balance economic expansion, growth, quality of life, and lifestyle of the living environment.

Formation of Carolinas Transportation Compact

Local elected city and county officials began to recognize the need for a regional multicounty, bistate transportation plan-

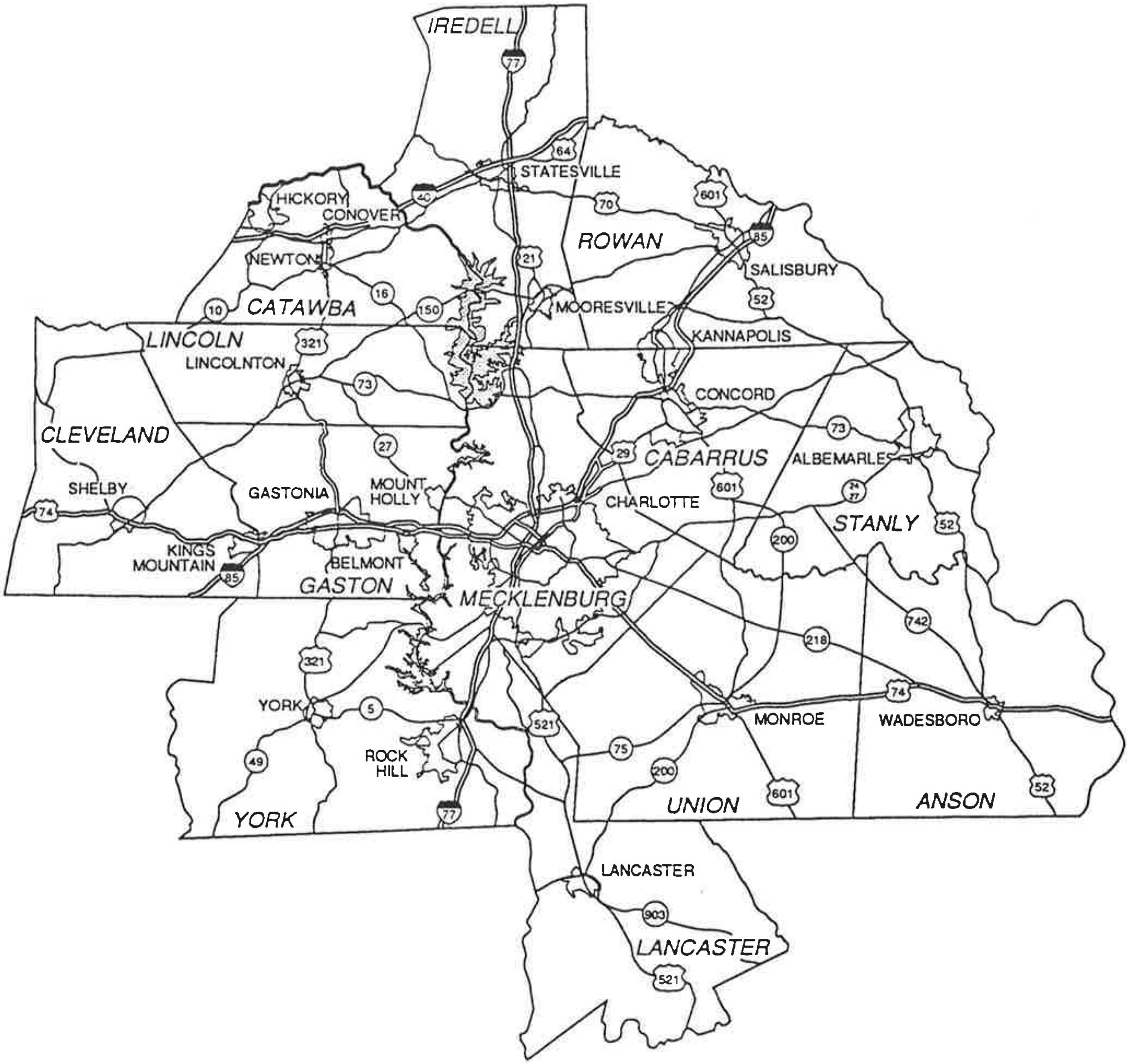


FIGURE 1 Charlotte metropolitan region: Hub of the Carolinas.

ning agency for Charlotte and the region. This perceived need, after building for some time, began to take shape with the establishment of the Carolinas Counties Coalition in 1985. This group, which consisted of elected officials, was formed to pursue regional agendas for the Charlotte, NC-SC, metropolitan area and the neighboring counties. The group identified transportation access as one of its critical issues for the region. There was considerable interest in a second-tier ring road around Charlotte and Mecklenburg County, which would later become known as the Carolinas Parkway. This road was envisioned to circumscribe Charlotte at 30 to 40 mi limits, connecting the counties in the outer ring (Figure 3). This loop would provide access to the region's Interstate systems and connect north-south and east-west Interstates. Within the coalition, a transportation task force was formed to undertake all transportation-related issues. The coalition recognized the linkages between continued transportation improvements and continued regional growth and quality of life. This task force was formed to determine precisely what transportation issues should be addressed by the group and to recommend ways to further the agenda.

In the early days, the task force was represented by each of the eight counties in the region immediately surrounding Charlotte, including South Carolina. Each county contributed three members to the task force. These members were also elected or appointed officials from county governments of major cities. The task force established a three-person executive committee and elected a chairperson. The task force was charged with developing an administrative and organi-

TABLE 2 Carolinas Transportation Compact, Population and Growth from 1980 to 1990

COUNTY	1980 POP.	1990 POP.	% CHANGE	1990 POP. % OF TOTAL
COUNTIES IN N.C.				
ANSON	25,649	23,474	-8.48%	1.48%
CABARRUS	85,895	98,935	15.18%	6.23%
CATAWBA	105,208	118,412	12.55%	7.46%
CLEVELAND	83,435	84,714	1.53%	5.33%
GASTON	162,568	175,093	7.70%	11.03%
IREDELL	82,538	92,931	12.59%	5.85%
LINCOLN	42,372	50,319	18.76%	3.17%
MECKLENBURG	404,270	511,433	26.51%	32.21%
ROWAN	99,186	110,605	11.51%	6.97%
STANLY	48,517	51,765	6.69%	3.26%
UNION	70,435	84,211	19.56%	5.30%
Sub Total NC Counties	*1,210,073*	*1,401,892*	*15.85%*	*88.29%*
COUNTIES IN S.C.				
LANCASTER	53,361	54,516	2.16%	3.43%
YORK	106,720	131,497	23.22%	8.28%
Sub Total SC Counties	*160,081*	*186,013*	*16.20%*	*11.71%*
	==========	==========	==========	==========
TOTAL ALL COUNTIES	1,370,154	1,587,905	15.89%	100.00%

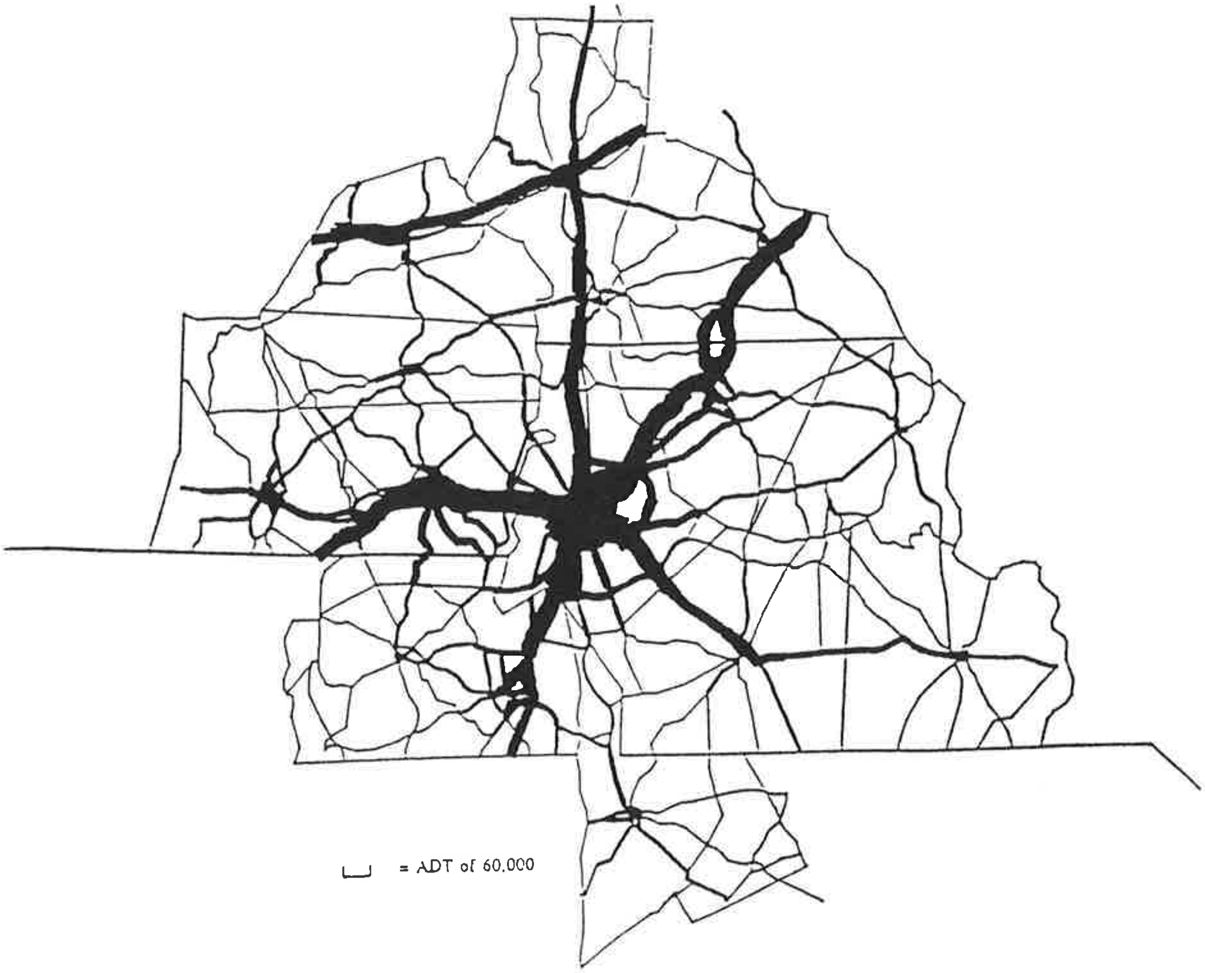


FIGURE 2 Traffic volumes within the Charlotte metropolitan region.

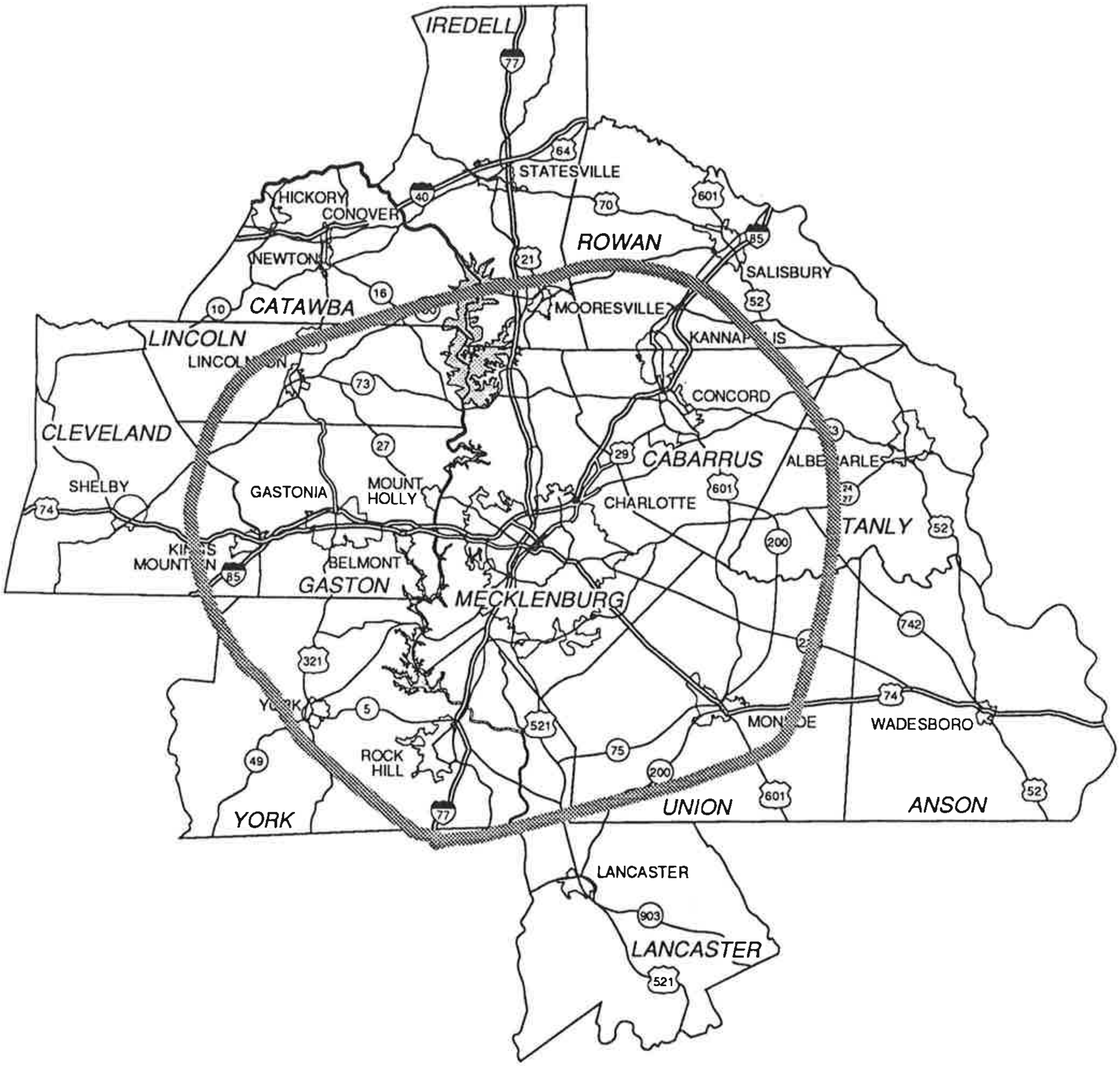


FIGURE 3 Proposed Carolinas Parkway, Charlotte metropolitan region.

zational plan for establishing a transportation planning and funding organization for the region.

At this time, the University of North Carolina at Charlotte (UNC Charlotte) became involved with the Carolinas Counties Coalition. The university offered assistance by organizing and hosting a series of meetings, over a 6-month period, for task force members. Faculty associates of the Urban Institute, a university-affiliated think tank, were able to assist the task force by structuring its issues, helping it review materials, maintaining national and state perspectives, and otherwise providing technical expertise. The university also has provided a convenient central meeting place, available space and facilities, and a readily available administrative pool for preparing reports and materials. The task force began its work in January 1989; the first four major items to be undertaken were mission statement, goals and objectives, organizational structure, and funding options. The group concluded its mission by recommending that a separate regional transportation planning organization be established to plan and promote all modes of transportation in the 13-county region.

The Carolinas Transportation Compact (CTC) is the regional transportation planning organization that grew directly out of the task force recommendations. The CTC was formed to advocate regional transportation needs in the Charlotte, NC-SC, region. The CTC consists of elected representatives from 13 counties and a representative from each of the two state transportation agencies. The counties involved are Anson, Cabarrus, Catawba, Cleveland, Gaston, Iredell, Lincoln, Mecklenburg, Rowan, Stanly, and Union from North Carolina and Lancaster and York from South Carolina (Figure 1). The mission statement adopted for the CTC is "to serve the individual counties by establishing coordinated, continuing, comprehensive, and proactive efforts to acquire federal, state and local resources for planning, constructing, operating and maintaining adequate regional transportation facilities that enhance the quality of life in each county and the economic opportunity of the region." The development of goals and objectives were undertaken along with the development of the mission statement. Recognizing the responsibilities of other agencies in the region to plan, build, and operate transportation systems, CTC agreed upon a set of goals that includes proactive planning but leaves the construction of transportation systems in the hands of the North Carolina and South Carolina departments of transportation (DOTs). The goals encourage existing organizations to plan at the superregional scale in cooperation with the CTC. Within counties and cities, the CTC leaves local governments the role of preparing internal transportation systems plans but recognizes that such activities should be coordinated with the activities of other contiguous and noncontiguous areas. The goals are to

- Work with local, state, and national elected and appointed officials to promote consideration and funding of regional transportation facilities;
- Conduct a feasibility study and corridor plan for the Carolinas Parkway;
- Develop a regional high investment plan;
- Develop a regional highway plan with DOTs, councils of government, and MPOs;
- Develop regional plans for airport facilities, public transportation service, rail service, and carpool-vanpool systems with cities and DOTs;
- Encourage and assist counties and cities in preparing and implementing coordinated thoroughfare plans;
- Identify and encourage the preservation of right-of-way for regional transportation facilities; and
- Monitor, report, and predict regional growth trends.

The source of the funds, organizational structure, and the specific activities of the compact were discussed at considerable length. The final structure selected for the CTC consists of two elected officials from each county, with votes and local funding proportional to population. Each state transportation agency has one representative. Initial proposals prepared by the university staff suggested funding be split between a staff and technical studies. The funding level was targeted at $225,000 a year. Members thought that the organization should start small and build capabilities as needed, beginning with just an executive director. An executive director was selected and began work in April 1991. In July 1991, the CTC "put some meat on its bones" (*8*). The members of the CTC approved an interlocal agreement between the counties to formalize the structure of the CTC, bylaws, and a work program for the coming year.

CURRENT CTC ACTIVITIES

The establishment of the CTC encouraged comprehensive regional planning and implementation within the area of the 13-member counties. For success, the CTC must cement itself as a viable regional transportation planning agency, both through its dealings with other agencies and through its own activities. To this end, the CTC is engaged in three major ongoing projects: an "outer-outer" belt feasibility study, a regional transportation authority feasibility and organizational study, and a railroad right-of-way preservation study. Other projects are in development.

Carolinas Parkway

The first of CTC's major projects deals with the feasibility and impacts of an outer-outer belt road through the outlying CTC counties. This belt is described as a second-tier ring road located 30 to 40 mi outside Charlotte (Figure 3). It is a second-tier road because Charlotte is currently constructing an outer belt within Mecklenburg County and is wrestling with proper alignment of that route (*9*). As an expression of the truly regional character of the outer-outer belt, it was named the Carolinas Parkway during an August 1990 meeting of the CTC.

From the beginning, the Carolinas Parkway project has served as the main galvanizing force of the CTC counties. Since the parkway will run through several outlying CTC counties, it is thought that benefits from the project will be spread around rather than serve only the interests of larger, more urban counties. In addition, planning for the parkway is truly regional as the local governments share in the future growth and development of the entire region. Finally, the circumferential design is more acceptable to the outlying counties as a group than would be a radial design with suburban nodes.

Progress has been steady on the Carolinas Parkway study. Up to this point, work has been conducted by research associates under the auspices of the Urban Institute at UNC Charlotte. Although the overall study deals with the feasibility and the effects of the parkway, work so far has concentrated on providing necessary input to those components. Specifically, researchers have conducted extensive data gathering, input, and analysis, which have culminated in traffic forecast models for the region (*10*). These will be used as inputs to the next stages of the project, the feasibility and impact studies.

The two state highway departments have begun a high-level feasibility study of the costs and benefits of the parkway. Calls for proposals for the feasibility study itself went out in March 1991, and a consultant was selected in July 1991. Estimates place the cost of the study at $500,000. This shows the willingness of the various participants to move forward with area megaprojects. It also shows that the state transportation agencies and the counties are comfortable with the role of the CTC, as they are now coming forward as partners in a major CTC project that should encourage continued progress.

Essentially, progress in the Carolinas Parkway project has been fast-paced. Beginning in August 1990, CTC staff developed the traffic models necessary to examine present and future regional traffic flows (*10*). Once proposals for the next stage are examined and a consulting team is chosen, CTC staff will continue to assist in location assessment and impact analyses of the project.

Regional Transportation Authority

The CTC is also studying the feasibility of establishing a two-state regional transportation authority (RTA). Both state transportation agencies have funded a feasibility study that began in the fall of 1991. This would allow local governments in the region the opportunity to decide the organizational and funding structure, as well as the type of service provided (*11*). As envisioned, the RTA would be similar to that recently legislated for the Research Triangle region of North Carolina and would plan for, fund, and perhaps operate future public transportation (transit) in the region.

At least five cities in the region provide intracity transit service, but there is little coordination of transit, even though there is significant evidence of intercity commuting, particularly from surrounding towns into Charlotte. What services do exist are mostly in the form of organized carpools and vanpools, although some shuttle buses have been attempted in the past. Now, someone commuting to work in Charlotte has almost no choice but to travel by car. For these reasons, regional service could become a reality soon in the next century (*12*). The RTA under consideration for the CTC region would examine the need for and plan for many types of services, including bus and eventually rail services.

To this point, most of the CTC's work on the RTA project has dealt with organizational issues related to such an authority. The CTC developed an "issues" brochure spelling out some of the major questions that must be addressed before creation of the RTA. These issues were discussed during a special forum held on the UNC Charlotte campus in October 1990. The issues discussed included the need for an RTA; governance and organization; geographic scope, powers, and services; special problems of a two-state RTA; and funding options. The last issue is of particular importance because it has already presented some problems for the Research Triangle Transportation Authority. The North Carolina Legislature did not provide operating funds for the Triangle authority until July 1991, even though legislature was passed to create the authority in 1989 (*13*).

The CTC has also conducted some preliminary needs analysis for an RTA. A rough cost estimate of service, based on general estimates of demand between cities, has allowed different service levels to be evaluated in terms of costs and revenues. This will become more important in the following months as the CTC goes through the necessary steps to have appropriate legislation passed in both the North Carolina and South Carolina legislatures and in Congress. With the Research Triangle RTA setting a precedent in North Carolina and a history of such agencies in South Carolina, passing in appropriate legislation should be direct. Clearly, continuation of the project will require more in-depth needs analysis, particularly as the RTA begins to plan services. The tasks completed on the RTA project have shown that more discussion and evaluation of the major issues are needed. The October 1990 forum did not create a consensus that an RTA was needed for the CTC region. Questions were raised about the need to create another level of government (*14*). This seemingly negative response can be explained in part by the fact that participants in the forum got caught up in discussion of funding and taxation issues. Clearly, the CTC must address these issues, but this should be considered separately from the need for an RTA. As one CTC executive committee member put it, "this region will mire up in its own mud if we don't work together and create other modes of transportation" (*15*). Charlotte and Mecklenburg County, along with Rock Hill, Gastonia, and York counties, have recently passed resolutions endorsing the CTC's efforts to conduct a feasibility study for an RTA.

Railroad Right-of-Way Preservation

The CTC's third major study deals with preserving existing railroad rights-of-way, an issue of increasing importance for many regions. Existing rights-of-way for corridors no longer used by the railroads are lost for future transportation use if they are converted back to other land uses by private interests. These rights-of-way could be very important to the public for use in any future rail transit or other transportation in the region.

The Charlotte-Mecklenburg Planning Commission recommended in its *Generalized Land Use Plan 2005* that countywide transportation services be provided and in addition, mandated a study of light rail transit (LRT) (*16*). That study, conducted in 1989, concluded that transportation services should focus first on expanding bus service and then plan for LRT as a future possibility (*17*). These recommendations were based on current figures. However, the county could eventually require LRT as part of a viable regional transit system. Therefore, preserving existing rights-of-way is vital for future transit needs.

In keeping with its role as the center of the transportation region, Charlotte is progressing with preparations for LRT, partly through negotiations with NCDOT to share costs of acquiring rights-of-way (*18*). The state, with more explicit statutes for right-of-way acquisition, can be a more forceful negotiator; combining resources can allow more acquisition of rights-of-way.

Even with increased resources from the state, Charlotte can only negotiate for rights-of-way within the city limits and lobby for preservation of those just outside the city limits. The CTC provides an obvious forum for identifying and preserving those corridors for which there is not a county or MPO to assume a leadership role, or those that the state does not deem as strategic for preservation (L. Purnell, unpublished data, 1991). In addition, the CTC has become a leader in helping South Carolina establish a coordinated preservation effort. The executive director has been working with a South Carolina House subcommittee in drafting legislation for rail preservation.

In an attempt to achieve some of the tasks just outlined, the CTC is currently collecting data on existing railroad rights-of-way. It is hoped that these data can identify those corridors that are most strategic and, therefore, most important to preserve. The CTC is surveying existing rail lines in the region to include data on current mileage and abandoned mileage. Response rates to the survey were not good; a follow-up may be necessary.

Other Projects

While the three main projects of the CTC just discussed are already well under way, the CTC is planning additional projects. The CTC is developing a "vision document" that would spell out and demonstrate the region's future transportation characteristics. This document could be used as a comprehensive planning tool for the region's future transportation needs. A regional transportation plan is also scheduled to follow the vision document. In addition, the CTC is also interested in investigating the possibilities of "smart car" technologies in the CTC region. This represents the cutting edge of transportation technology and could be useful on a regional scale. A fourth possibility is the analysis of a cargo airport for the region. These are but four possible directions in which future CTC research may proceed, in addition to continuation of existing CTC activities and service to member governments.

Role of University

UNC Charlotte continues to play a support role for the CTC. UNC Charlotte, as a regional entity, provides a base for CTC activities. The CTC and its executive director are housed on campus, and the technical staff is provided through the university. In addition, the university provides facilities for larger CTC activities, such as meetings and the October RTA forum. The CTC has hired a permanent executive director to oversee CTC activities. The executive director provides the CTC with a leader to keep the regional transportation agenda on track and to act as a liaison between the various agencies involved in CTC-related interests.

EVALUATING CTC'S SUCCESS

Earlier papers outlined several necessary ingredients of a successful superregional transportation agency (*5*). These were put forth as informal guidelines for similar agencies:

- The need for pragmatism and compromise,
- The need for "champions,"
- The importance of galvanizing issues,
- The need for neutrality in the host agency,
- Support for other agencies,
- Issues of exclusivity and inclusivity,
- The need to select "real players."

It is possible, through examining these goals, to evaluate the CTC's success. Most of these goals have been met, at least to some extent, in the early part of the agenda.

The CTC continues to act as a forum for compromise for the various participants in the region. But, ". . . a rapid pace (as desired by the participants) simply could not be maintained, and a scale as broad as that initially perceived could not impose on the existing organizational structure" (*5*). The CTC is a tempered approach to regional transportation issues, but it still recognizes the needs of the various cities, counties, and other agencies involved. An example of this is the forum conducted on RTA issues. This type of discussion demonstrates the need to let all members be involved in the process yet clearly shows that not everyone will always agree.

The CTC clearly has champions. The activities of the CTC have gained fairly broad political support, partly through the active involvement of a recently elected North Carolina state senator, who stated at the RTA forum: "We cannot fail in the endeavor—we won't be able to compete in this country, much less globally, unless we create a working regional transportation authority" (*15*). At the same forum, the deputy secretary of NCDOT stated that the department "applauds the local efforts already begun and that the NCDOT is ready and willing to help" (*19*). In addition, funding support for CTC activities have come from the counties and both states, and both North Carolina and South Carolina are investing in the Carolinas Parkway feasibility study. The larger cities and counties have passed resolutions requesting that the CTC undertake the RTA feasibility study.

As another stated element of success, the CTC has become a very important galvanizing forum. The Carolinas Parkway continues to serve as a rallying point for CTC member counties and the state agencies. At the August 1990 meeting of the CTC, one observer stated that he had never seen such a large region (so many counties) work together on one issue so well. Again, from the beginning, the parkway has rendered service to member governments and has been supported strongly by all the CTC members.

In terms of "finding a supportive niche," the CTC continues to focus on supporting the roles of other agencies. The CTC's role is to advocate and plan regional transportation issues, not to take over the roles of other agencies already involved in transportation. For example, NCDOT clearly maintains the responsibility for planning and constructing roads in the state. The fact that the deputy secretary of NCDOT was so supportive of CTC efforts for an RTA indicates a general level of goodwill that would not exist had the CTC encroached upon the state's roles. Another aspect of the CTC's success hinges on its ability to select "real players." Hartgen and McCoy pointed out some initial problems with meeting attendance and dissemination of information (*4*).

This discussion has indicated the continued attempts by the CTC to maintain and improve the necessary ingredients for success. Even with the success mentioned, some problems still exist. These range from the different views on the need and funding for an RTA. Even with the problems mentioned, the CTC shows signs of being a successful approach to regional transportation planning. Most of the successful components are in place and functioning. The CTC has garnered broad political support and increasing public interest in the issues it presents, particularly the issues of regional transit and the Carolinas Parkway. Public support should increase as traffic and suburban ring growth continue to attract public attention. Partly because of this, the CTC and its activities have received excellent media coverage to promote the regional agenda.

In addition to champions in the legislature's influence, other groups are bringing the whole issue of regional topics to the forefront. For instance, the Urban Land Institute (*20*) pointed out that "the region should adopt a broad-based approach to transportation. Implementation of a regional transportation element is critical so that a regional distribution of employment and housing opportunities can occur." Finally, the work already completed on the major CTC projects and the hiring of an executive director, in conjunction with efforts to begin new research activities, suggest that the CTC is congealing into a viable and vital organization concerned with meeting regional transportation needs.

The future of the CTC looks bright. The CTC has now established a track record for advocating regional transportation through its past activities. The CTC plans to continue work on the major projects already discussed. In addition, plans for future projects are already under way, and the CTC will continue to try to meet the goals put forth in "Uncharted Waters" (*4*), particularly in regard to working with other agencies to create a truly regional approach to planning and implementing appropriate regional transportation.

REFERENCES

1. F. L. Knox et al. *The United States: A Contemporary Human Geography*. John Wiley and Sons, Inc., New York, N.Y., 1988.
2. *Transportation Decisions: The Role of the Mayor: State/Local/Regional Transportation Cooperation from the Mayor's Perspective*. Office of the Secretary, U.S. Department of Transportation, 1980.
3. R. F. Kirby. Case Study of WMATA. In *Special Report 217: New Organizational Responses to the Changing Transit Environment*. TRB, National Research Council, Washington, D.C., 1980.
4. D. T. Hartgen and W. J. McCoy. Uncharted Waters: The Super-Regional Transportation Agency. *Transportation Quarterly*, July 1990.
5. J. W. Clay and A. Stuart. *Metrolina Atlas*. University of North Carolina Press, Chapel Hill, 1990.
6. *The Charlotte Observer*. Charlotte, N.C., Jan. 26, 1991.
7. C. P. Work et al. Jam Sessions. *U.S. News and World Report*. Sept. 7, 1987.
8. *The Mecklenburg Gazette*. Charlotte, N.C., July 10, 1991.
9. *The Charlotte Observer*. Charlotte, N.C., Nov. 2, 1990.
10. P. Gallimore, D. Hartgen, and Y. Li. Applications of GIS-Transportation Packages to Super-Regional Modeling. Presented at the 3rd National Conference on Transportation Planning Applications, Dallas, Tex., April 1991.
11. *The Charlotte Observer*. Charlotte, N.C., July 18, 1991.
12. *The Charlotte Observer*. Charlotte, N.C., Oct. 15, 1990.
13. *The Charlotte Observer*. Charlotte, N.C., July 1989.
14. *The Charlotte Observer*. Charlotte, N.C., Oct. 16, 1990.
15. J. Blackmon. Speech. Fall 1990 Carolinas Transportation Compact meeting, Charlotte, N.C., Oct. 15, 1990.
16. *Generalized Land Use Plan 2005*. Charlotte-Mecklenburg Planning Commission, Charlotte, N.C., Nov. 1985.
17. *Charlotte Transit Corridor Study*. Barton-Aschman Associates, Inc., Washington, D.C., June 1989.
18. *The Charlotte Observer*. Charlotte, N.C., Jan. 4, 1989.
19. T. S. Rhodes. Speech. Fall 1990 Carolinas Transportation Compact meeting, Charlotte, N.C., Oct. 15, 1990.
20. *An Evaluation of Development and Marketing Strategies for the Carolinas Partnership*. Urban Land Institute, Charlotte, North Carolina Region, Washington, D.C., July 1990.

Publication of this paper sponsored by Committee on Transportation Programming, Planning, and Systems Evaluation.

Roadway Levels of Service in an Era of Growth Management

REID EWING

The tendency in growth management is to focus on roadway level-of-service standards. However, the methods used to determine roadway levels of service may affect conclusions about road adequacy as much as do the standards to which they are compared. The specific method used to analyze roadway levels-of-service can make at least a two-letter grade difference in the outcome; so can the choice of analysis period or peak hour. Although harder to quantify, the effect of averaging levels-of-service across facilities could be of comparable magnitude. In determining roadway levels of service, most Florida jurisdictions go by the book. They analyze the 30 highest hourly volumes roadway by roadway, using methodology from the 1985 *Highway Capacity Manual.* A few jurisdictions have opted for innovative but unconventional approaches. Although it is tempting to reject these approaches as "not professionally accepted," the "book" was written for applications other than areawide growth management. This relatively new area of application requires fresh thinking. In the context of growth management, the use of the following is recommended: (a) simple regression models to estimate average travel speeds and, from them, arterial levels of service; (b) average levels of service to determine adequacy of facilities within travel corridors; and (c) the 100th rather than the 30th highest hourly traffic volumes as the basis for roadway level-of-service determinations.

Roadway levels of service play a central role in Florida's efforts to manage growth. The state's landmark growth management law embraced a "pay as you grow" philosophy, commonly known as concurrency. Adequate infrastructure must be available concurrent with the effects of development. Adequacy is defined by level-of-service standards, which are adopted by local governments as part of their comprehensive plans. No development order or permit may be issued if levels of service will be degraded below the adopted standards.

As one observer noted, "five of the infrastructure elements have posed few problems for local governments. But the sixth category, roads, is proving to be a nightmare" (*1*, p. 6). Roads are the infrastructure element most likely to trigger public dissatisfaction, growth moratoria, and legal challenges under the growth management law. Thus, it is crucial that roadway level-of-service determinations be accurate and results be interpreted meaningfully.

SCOPE OF INQUIRY

In determining roadway levels of service, most Florida jurisdictions go by the book. They analyze 30th highest hourly volumes roadway by roadway using methodology from the 1985 *Highway Capacity Manual* (*2*).

A few jurisdictions have opted for innovative but unconventional approaches. Although it is tempting to reject these approaches as being not professionally accepted, it must be remembered the "book" was written for applications other than areawide growth management. This relatively new area of application requires fresh thinking.

Accordingly, three old methodological issues are addressed anew in this paper:

1. What methods should be used to analyze levels of service?
2. When should levels of service be averaged or otherwise aggregated?
3. What peak period or peak hour should be analyzed?

METHODS OF ANALYSIS

Methods of analyzing roadway levels of service may be arrayed according to data and analytical requirements and corresponding precision of estimates. It is usually assumed that the simplest methods are the least precise, the most complex methods the most precise. To the author's knowledge, this assumption has never been field-tested.

Assessment of Standard Methods

To test standard methods of analysis, traffic and travel time data for three arterials were acquired from consulting firms. Two of the arterials, Kirkman Road and Turkey Lake Road, are in Orlando. The former has high traffic volumes and low signal density, and the latter has relatively low traffic volumes and higher signal density.

For each arterial, traffic counts and travel time runs were increased during the same peak hours on the same weekdays. Thus, by design, actual travel speeds (derived from travel time runs) and estimated travel speeds (dependent on traffic counts) relate to the same periods.

Data in hand, intersections were first analyzed with HCS (Highway Capacity Software). Liberal assumptions were made about

- Saturation flow rates at intersections on these arterials (1,850 vehicles per hour after adjustments),
- Amount of green time devoted to arterial through movements (the maximum possible, given the timing plans of these semiactuated traffic signals),

---

Center for Urban Transportation Research, College of Engineering, University of South Florida, Tampa, Fla. 33620.

- Arrival types of vehicle platoons (the best possible progression, given signal spacing and signal timing offsets), and
- The peak-hour factor (a value of 1.0 was assumed, as if flow rates were constant during the peak hour).

A peak-hour factor of 1.0 was assumed to achieve a measure of consistency between HCS results and travel time runs. If actual peak-hour factors had been used instead, HCS would have analyzed the peak 15-min period of the peak hour, whereas travel time runs were averaged over the entire peak hour.

After intersection delays were computed, they were fed into the HCS arterial analysis program. The program adds intersection delays to running times between intersections to arrive at estimates of overall travel speed.

Assumptions from HCS analyses were later carried over to ART-PLAN; ART-PLAN is a simplified version of HCS distributed by the Florida Department of Transportation (FDOT). This meant that outputs of both programs could be compared with travel time runs with some assurance that all were measuring the same conditions.

Estimated and actual average travel speeds are compared in Figures 1 and 2. Given two arterials, two peak periods, and two directions, eight comparisons can be made. It appears that actual travel speeds are significantly higher than estimated speeds in nearly all cases. They are 5 to 10 mph higher in most cases than HCS estimates and a letter grade higher in level of service.

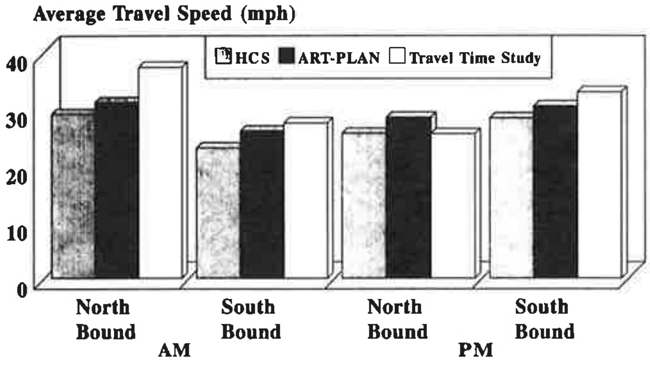


FIGURE 1 Average travel speeds, Kirkman Road.

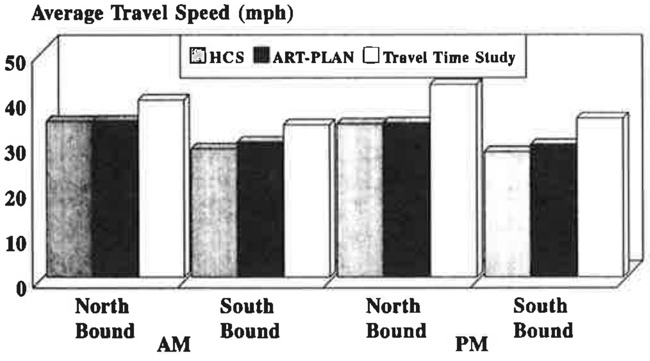


FIGURE 2 Average travel speeds, Turkey Lake Road.

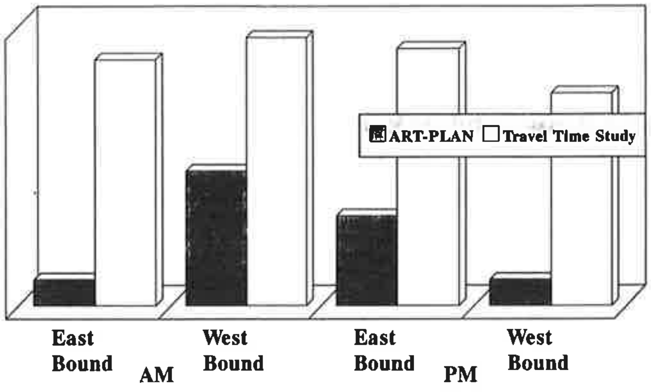


FIGURE 3 Average travel speeds, Broward Boulevard.

The other arterial analyzed was Broward Boulevard in Fort Lauderdale–Plantation. Broward Boulevard has higher traffic volumes on side streets than Kirkman Road and thus can claim a smaller portion of the signal cycle to accommodate its heavy traffic volumes.

Average travel speeds and peak-hour volumes for Broward Boulevard were gathered on comparable weekdays of the same month. Portions of the cycle devoted to through movements were observed at the same time that traffic counts were taken. Thus, although green ratios for Broward Boulevard appear very low, there is no reason to doubt the validity of the values supplied by the consulting firm.

Estimated travel speeds on Broward Boulevard are a fraction of actual speeds (see Figure 3). If results for Kirkman and Turkey Lake roads suggest that standard methods of analysis underestimate travel speeds, results for Broward Boulevard indicate that standard methods break down entirely when demands are too heavy relative to intersection capacity. Methods of analysis seem to break down before the road system does in practice.

Why Estimates Differ from Actual Travel Speeds

To help explain why actual travel speeds are higher than estimated speed, results for Kirkman and Turkey Lake roads were analyzed by roadway segment and by component of total travel time, the components being delay at intersections and running time between intersections. Typical results (for northbound a.m. movements) are presented in Tables 1 and 2. Actual stopped delays are mostly shorter than those estimated with HCS and ART-PLAN. Differences are greater for intersections with long delays. Actual running speeds are significantly higher than those estimated with either program and, as such, account for most of the difference between actual and estimated average travel speeds. Following convention, the posted speed limit was taken as the free-flow speed in HCS and ART-PLAN analyses. However, roads are often designed for higher speeds than are posted, and, as casual observation suggests, drivers tend to drive at design speeds on long, uninterrupted segments with moderate traffic volumes.

There is another reason that actual running speeds are higher than estimated speeds. HCS and ART-PLAN compute some

TABLE 1 Average Stopped Delay in Seconds: Northbound a.m. Movements

	HCS	ART-PLAN	Travel Time Study
Kirkman Road			
Carrier to International	10.0	11.4	12.15
International to Major	7.3	7.3	6.0
Major to Vineland	9.8	11.1	11.2
Vineland to Conroy	29.3	28.8	6.4
Average	14.1	14.7	9.0
Turkey Lake Road			
Sand Lake to Wallace	3.4	3.4	12.3
Wallace to Panther	2.6	2.6	2.6
Panther to Paw	1.0	1.0	0
Paw to Hollywood	1.9	1.8	0
Hollywood to Production	1.2	1.2	5.7
Production to Vineland	0.9	0.8	0
Average	1.8	1.8	3.5

TABLE 2 Average Running Speed in Miles per Hour: Northbound a.m. Movements

	HCS	ART-PLAN	Travel Time Study
Kirkman Road			
Carrier to International	32.1	31.8	29.7
International to Major	36.8	46.7	51.5
Major to Vineland	33.2	33.1	39.1
Vineland to Conroy	35.2	35.3	42.8
Average	35.2	39.6	44.6
Turkey Lake Road			
Sand Lake to Wallace	38.9	38.9	49.6
Wallace to Panther	38.2	38.2	46.2
Panther to Paw	32.4	32.4	48.4
Paw to Hollywood	29.8	29.9	30.8
Hollywood to Production	30.5	30.6	29.0
Production to Vineland	36.1	36.1	38.9
Average	36.6	36.7	44.7

stopped delay at all intersections. Hence, they assume acceleration and deceleration at all intersections. Although delay associated with deceleration is set proportional to stopped delay, and thus is small when stopped delay is small, delay associated with acceleration is a function of segment length (*2*, Table 11-4). The latter can significantly depress estimated running speeds on short segments.

In reality, no stops or delays are experienced during most travel time runs at intersections with high green ratios and good progression. Thus, contrary to HCS and ART-PLAN, vehicles travel at free-flow speeds over entire roadway segments or even sections.

Application of Travel Time Studies

Few jurisdictions conduct travel time studies as a method of determining arterial levels of service. One reason is the high cost of such studies; this factor may be rendered moot eventually by advances in automatic vehicle location technology. The more important reason is the perception that travel time study results apply only to the specific period when travel time runs are done—that they cannot be used to predict levels of service under standardized conditions, as required in growth management. This perception is incorrect.

Travel time study results can be used to calibrate HCS, ART-PLAN, and other programs. Default values assumed by these programs may not be applicable to a particular locale. Programs can be run with progressively higher saturation flow rates or free-flow speeds until estimated intersection delays, running speeds, and overall travel speeds better approximate travel time runs. The better-calibrated programs can then be used to predict levels of service under standardized conditions.

Alternatively, travel time study results can be correlated directly with traffic volumes and other variables in statistically derived models. To illustrate this approach, average peak-hour travel speeds were acquired for 17 two-lane roadways in Seminole County, Florida. With a.m. and p.m. peak hours, and northbound and southbound directions, speed data were available for a total of 68 movements.

Average peak-hour travel speeds were regressed on peak-hour traffic volumes, numbers of signalized intersections per mile, and posted speed limits. Both linear and nonlinear forms of the regression equation were tested. The best fit to the data was obtained with a linear equation in two independent variables: peak-hour traffic volume and number of signalized intersections per mile (see Figure 4). The speed-volume relationship is known to become nonlinear as road capacity is approached. Apparently, Seminole County roads operate in a flow range that is adequately represented by a linear equation.

The explanatory power of the model estimated for Seminole County is probably inadequate for use in predicting average travel speeds and levels of service. The standard error of the estimate, 5.3 mph, could result in a one- or even two-letter grade difference between estimated and actual levels of service. Nonetheless, with 55 percent of the variation in average travel speeds explained by only two independent variables, it appears likely that a good predictive model could be developed with a richer data base (including such independent variables as the green ratio, arrival type, and percentage of turns from exclusive lanes).

The regression model's simplicity should not be viewed as a shortcoming. The complicated models and multitude of parameters used in the 1985 *Highway Capacity Manual* only give the appearance of precision (*2*). In light of results for Kirkman

$$\begin{array}{l} \text{Average} \\ \text{Travel} \\ \text{Speed} \end{array} = 44.7 - \underset{(3.12)}{0.0087} \;\text{X}\; \begin{array}{l} \text{Peak} \\ \text{Hour} \\ \text{Traffic} \\ \text{Volume} \end{array} - \underset{(6.65)}{7.74} \;\text{X}\; \begin{array}{l} \text{Signals} \\ \text{Per Mile} \end{array}$$

R-squared = 0.55

Standard Error = 5.3

Number of Observations = 68

Degrees of Freedom = 65

t - statistics shown in parentheses

FIGURE 4 Regression equation for average travel speed: Seminole County, two-lane roads.

Road, Turkey Lake Road, and Broward Boulevard, added complexity need not translate into added precision.

AREAWIDE LEVELS OF SERVICE

The *Florida Engineering Society Journal* (*3,4*) featured a debate over the merits of averaging roadway levels of service within a corridor, district, or entire urban area. One author contended that averaging could result in a "glossing over of transportation problems" (*3*, p. 20). Another countered that requiring each roadway link to operate at a minimum acceptable level of service causes "short-term incremental improvements rather than long-term comprehensive improvements" (*4*, p. 24).

Both authors are right. The challenge is to devise level-of-service measures and standards that encourage a long-term comprehensive approach to transportation improvement programming while still addressing localized traffic problems.

Current Practice

It is routine in traffic impact studies to estimate levels of service for

- A lane group at an intersection,
- An entire intersection,
- A roadway segment from intersection to intersection, and
- A section of roadway with multiple intersections along its length.

However, we enter uncharted waters when combining levels of service of various facilities into one overall level of service. There is no standard, professionally accepted level-of-service measure for a travel corridor, a traffic district, or an entire road network.

Concepts Underlying Areawide Levels of Service

Two distinct concepts justify and guide the development of areawide level-of-service measures. The first is the concept of typical trips. Over the course of a day, or even a single trip, a person may travel on scores of roadway links and dozens of different roads. Presumably, a traveler's perception of roadway conditions is based on an entire trip or even an entire day's worth of travel, not the delay at one intersection or congestion on one roadway segment. Therefore, roadway levels of service might reasonably be combined to reflect common travel patterns and trip lengths.

Areawide levels of service may also be justified by the concept of alternative routes. Where a well-developed road network exists, an individual may have many routes available for a given trip. If any route provides an acceptable level of service, government may have met its responsibility to the individual trip maker. Hence, roadway levels of service might reasonably be combined for parallel routes within a travel corridor.

Alternative Areawide Approaches

There are at least three ways to combine the levels of service of various facilities into one overall level-of-service measure. All have precedents in Florida's local comprehensive plans.

The first approach is to sum traffic volumes and capacities for roads in a given area (where capacities are equal to maximum volumes at adopted levels of service). If the sum of traffic volumes is less than the sum of capacities, the area might be deemed to meet level-of-service standards. Lee County, Florida, sums traffic volumes and roadway capacities within traffic districts and uses any net capacity to justify degradation of already "backlogged" roads (see Figure 5).

A second approach is to average levels of service across facilities of a given type in an area. Although averaging in this context is novel, averaging travel speeds on arterials has been an accepted practice since the 1985 *Highway Capacity Manual* was released. It is not difficult conceptually or methodologically to go from averaging speeds on arterials to averaging speeds across arterials. The Brevard County Comprehensive Plan allows levels of service to be averaged in one specific travel corridor.

A third approach is to adopt a performance summary for roads in an area that specifies the percentage of roads at or above given levels of service. A standard is applied to centerline miles, lane miles, or vehicle miles collectively rather than to each road individually. An example of this approach is found in the Orlando Comprehensive Plan (see Table 3).

All three approaches—summing volumes and capacities, averaging levels of service, and adopting performance summaries—allow local governments to finance the most cost-effective system improvements rather than isolated roadway improvements dictated by minimum operating standards. How does one choose among them?

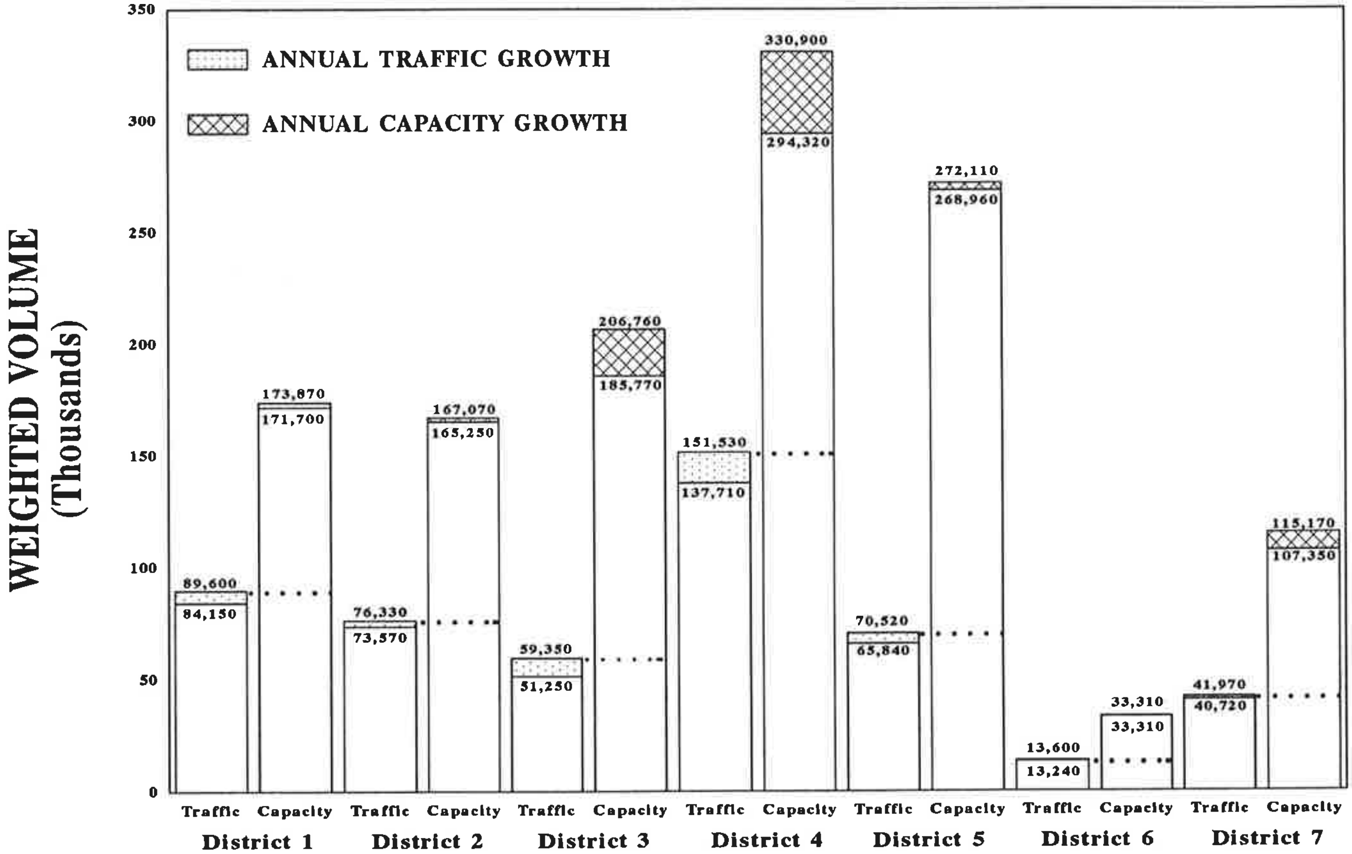
The adoption of performance summaries conforms to standard engineering practice, whereas the other two approaches extrapolate from such practice. By continuing to analyze roads individually, performance summaries avoid methodological leaps of faith.

Even so, the averaging method may be preferred for growth management purposes. Travel speeds fall precipitously as traffic volumes approach capacities. With areawide averaging, local governments, concerned about maintaining average travel speed, will have a considerable incentive to fix traffic hot spots. Less incentive is provided by the other approaches.

The fact that areawide averaging is not standard engineering practice practice represents an opportunity instead of a constraint. Standard practice could change with an update of the *Highway Capacity Manual*, as level-of-service standards are increasingly applied to growth management. Even if standard practice remains tied to individual facilities, areawide averaging will gain all the legitimacy required for growth management if it is sanctioned by regulatory agencies.

Weighting Factors

Whichever method is chosen, roadways must be assigned weights that reflect their contributions to overall levels of service. Lee County weights traffic volumes and capacities of roadway segments by their respective lengths (i.e., by cen-



Source: 1990 Amendment to the Lee County Comprehensive Plan, Volume 1 of 3, September 1990, Traffic Circulation Issues, Exhibit VI-6, page VI-10.

FIGURE 5 Traffic volume versus capacity by district, Lee County.

TABLE 3 District Performance Criteria, City of Orlando

Traffic Performance District	Lane Miles Operating at or Above Level of Service Standard	
	1995	2010
1	73%	75%
2	69%	73%
3	33%	52%
4	32%	34%
5	79%	88%
6	81%	84%
7	88%	88%
8	80%	95%
9	60%	66%
10	55%	50%
11	59%	62%
12	89%	97%
13	100%	100%
14	85%	91%
15	51%	58%

terline miles). Brevard County also uses segment lengths as a weighting factor. Pasco County weights its performance summary by the number of vehicle miles traveled on various roads; Orlando uses the number of lane miles; and Tampa uses centerline miles in one performance summary and vehicle miles in another.

Use of vehicle miles accounts for the volume of traffic exposed to various traffic conditions. Since it is the "average" experience of travelers we wish to capture in an overall level-of-service measure, not the average condition of roadways, vehicle miles appears to be the preferred weighting factor. Use of other weighting factors could encourage improvements to low-volume roads simply to meet regulatory requirements, whereas higher-volume roads go unattended.

Delineation of Travel Corridors

Localities will require some guidance as they delineate corridors or districts within which levels of service are averaged or otherwise combined. This discussion will refer to such areas generically as transportation concurrency management areas (TCMAs), a name coined by Florida's state planning agency.

If TCMAs are too large, traffic problems will be glossed over and development decisions will be subject to challenge. Property owners near the edges of large TCMAs might be expected to challenge project disapprovals prompted by traffic congestion at central locations or opposite edges.

If TCMAs are too small, flexibility to respond to system-wide needs will be sacrificed. In the extreme, TCMAs will cease to reflect motorists' experiences on typical trips or their choices among alternative routes and simply become surrogates for individual facilities.

For guidance in delineating TCMAs, the concepts of typical trips and alternative routes may be combined in the following general guideline: TCMAs should be drawn so as to encompass alternative routes available for common peak-hour trips. How this guideline is put into operation is best left to local planners; let it suffice to say that the guideline could be put into operation. For example, regional travel models could be used to generate tables of trip interchanges between traffic zones, and from them, common origin-destination pairs could be identified. Because level-of-service standards apply to peak hours, primary consideration might be given to work trip interchanges. Boundaries could be drawn so that traffic zones between which a majority of trip interchanges occur are part of the same TCMAs.

CHOICE OF PEAK HOUR

Florida's administrative rules require that levels of service be analyzed for peak-hour conditions. Use of peak-hour volumes is consistent with standard engineering practice in facility design, traffic operations, and traffic control.

However, as McShane and Roess note, "If peak-hour volume is to be used as a common focus of design, operations, and control analyses, it is critical to understand *which* peak hour is being used" (*5*,p.63). Among the multitude of choices are the single highest hour of the year, the 30th highest hour of the year, the average peak hour of the peak season, and the annual average peak hour.

Interplay of Peak Hour and Level-of-Service Standards

The choice of peak hour cannot be divorced from the setting of level-of-service, or LOS, standards. The effect will be the same if a lower standard is applied to a higher-volume hour, or if a higher standard is applied to a lower-volume hour. In its comprehensive plan, Lee County adopted two standards: LOS D for the annual average peak hour and LOS E for the average peak hour of the peak season. The lower standard (LOS E) applied to the peak season may be more restrictive than the higher standard (LOS D) applied to the entire year.

Does this mean that there is no preferred peak hour for growth management purposes? Hardly; it does mean that the choice of peak hour must be made on some basis other than the desire to foster or restrict growth (which can be accomplished with any peak hour by simply lowering or raising the level-of-service standard).

30th Highest Hour

Use of the 30th highest hourly volume is near universal in roadway design. The practice dates back to the 1950 *Highway Capacity Manual* (*6*). Traffic studies of that era observed extreme variations in traffic flow on facilities from hour to hour, day to day, and season to season. When hourly traffic volumes during a 1-year period were plotted in order of descending magnitude, the resulting curves often dropped sharply at first and leveled off quickly. The "knee" of the curve, where the slope changed markedly, often corresponded to the 30th highest hourly volume.

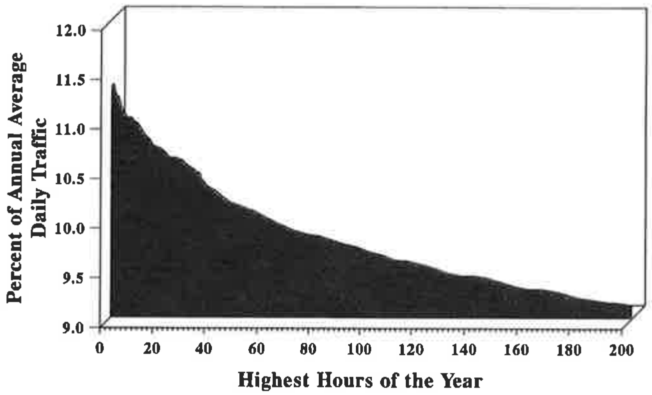


FIGURE 6 Small city route: Marianna US-90, Station 117.

Based on this early work, it has become conventional wisdom that

- The 30th highest hourly volume is the point of diminishing returns in roadway design.
- It is uneconomical to design for volumes to the left of the 30th highest hour, because a great deal of capacity is required to meet demands that occur only a few times a year.
- It is shortsighted to design for volumes to the right of the 30th highest hour, because little additional capacity is required to meet demands that occur frequently.

This need not be the case. Hourly traffic volumes in many localities do not follow the indicated pattern. Hourly volume curves tend to flatten out rather than remain static as areas become more developed; the knee of the curve becomes a moving target or disappears entirely. Even if hourly volume curves have predictable turning points, these points have no economic significance; the optimum design of a facility can be determined only by comparing the costs of alternative designs with the benefits to motorists (*7*).

Plots of hourly traffic volumes for 20 representative FDOT permanent count stations illustrate the arbitrariness of the 30th highest hour (see, for example, Figure 6). Several of the curves never level off and some have no point at which the slope changes markedly. Even where there is a discernible "knee," it seldom corresponds to the 30th highest hour.

The choice of design hour is ultimately a political rather than a technical matter. It involves balancing the public's desire to hold down road user taxes (which means more traffic congestion) against their desire to avoid traffic congestion (which means higher user taxes).

100th Highest Hour?

If statewide level-of-service standards are to have meaning, they must apply to the same peak hour everywhere. FDOT has proposed a shift from the 30th to the 100th highest hour as the basis for level-of-service determinations. Is this shift warranted and in the right direction?

Use of the 30th highest hour ties level-of-service standards to the exceptional travel experience. It could be argued that

TABLE 4 Traffic Counts at Various Peak Hours

Station	30th Highest Hour	100th Highest Hour	Average Peak Hour/Weekdays/ Peak Season
117	1890	1736	1658
87	2217	2069	2157
166	1444	1345	1145
13	2105	2049	2024
161	4905	4815	4756
96	2022	1958	2005
145	594	548	554
149	252	229	219
38	2162	2048	2130
118	1938	1748	1604
47	879	811	644
66	1913	1772	1594
94	3494	3403	3419
105	1424	1309	1326
113	4571	4425	4481
151	3494	3359	3421
159	2222	2137	2231
160	1076	989	948
164	2167	2039	1697
165	3350	3123	3320

standards should instead reflect a more typical travel experience—"typical" at least of peak periods.

Table 4 presents traffic counts at 20 permanent FDOT count stations. For some roads, the 100th highest hourly volume is lower than the average peak-hour volume on weekdays during the peak season. For others, it is somewhat higher. However, in general, the 100th highest hour is roughly equivalent to the average weekday peak hour during the peak season, except on recreational routes. (On recreational routes, weekend peaking causes the 100th highest hourly volumes to far exceed weekday peak-hour volumes. This is the case at FDOT Count Stations 66, 164, and 166.)

The 100th highest hour volume would be easy to estimate, assuming this rough equivalence is borne out. It would be necessary to take only one 24-hr count on a typical weekday during the peak season. The highest hourly count for that 24-hr period could be taken as an estimate of the 100th highest hourly volume. This would improve on the practice in many localities of applying a generalized *K*-factor to a single, seasonally adjusted 24-hr traffic count.

Additionally, the 100th highest hour volume would be relatively easy to project. Standard regional travel models forecast traffic volumes for the average weekday during the peak season. To obtain estimates of the 100th highest hourly volumes, it would be necessary only to apply a peak-to-daily ratio to model outputs. At present, modelers must first convert model outputs to annual average daily traffic volumes and then apply a generalized *K*-factor to the result.

Peak Period Instead of Peak Hour?

Daily peaks tend to spread out as urban areas grow and traffic congestion causes motorists to adjust their travel hours. Indeed, the largest cities do not have a peak hour per se but a 2- to 3-hr period in the morning and afternoon when commuting is heaviest. Roads become capacity-constrained, and *K*-factors come to be determined by supply rather than demand. We can expect even more spreading of the peak as traffic congestion worsens and communities seek to better manage travel demand.

With the state's approval, Dade County and the city of Miami based levels of service in their comprehensive plans on average hourly traffic volumes for the two highest consecutive hours of the average weekday. Although analysis of a 2-hr peak period flies in the face of time-honored design convention (which uses a single design hour), the convention may prove too limiting.

It is not the spreading of the peak that, in time, will justify a shift from peak-hour to peak-period analysis. This spreading is already reflected in hourly traffic counts and should be reflected in the *K*-factors used in traffic projections. Instead, it will be the adoption of policies that encourage commuters to adjust their times of travel.

Let us say a locality adopts a trip reduction ordinance requiring employers to institute flextime. Employees would then have the option of commuting at less congested hours. Such a policy could justify the averaging of traffic volumes over flexible starting and ending hours. As when alternative routes are made available, a case could be made that with flextime in place, government no longer has responsibility for accommodating every trip maker's choice of travel hour.

CONCLUSION

The tendency in growth management is to focus on roadway level-of-service standards. However, the methods used to determine roadway levels of service may affect conclusions about road adequacy as much as do the standards to which they are compared.

The specific method used to analyze roadway levels of service can make at least a two-letter grade difference in the outcome. So can the choice of peak hour. Although it is harder to quantify, the effect of averaging levels of service across facilities could be of comparable magnitude.

Thus, even adopting the same level-of-service standards, level-of-service determinations for, say, the city of Miami and Jefferson County, Florida, have entirely different implications for motorists. Jefferson County goes by the book, comparing the 30th highest hourly traffic volumes on individual roads to the maximum volumes at different levels of service based on *Highway Capacity Manual* methodology. In contrast, Miami has adopted an innovative but unconventional approach, comparing person-trip volumes for the two highest hours on the average weekday to the practical capacities of multimodal transportation corridors.

In the context of growth management, three innovations seem particularly promising: (a) using simple regression models to estimate average travel speeds and, from them, arterial levels of service; (b) using average levels of service to determine adequacy of facilities within travel corridors; and (c) using 100th rather than 30th highest hourly traffic volumes as the basis for roadway level-of-service determinations.

ACKNOWLEDGMENTS

The traffic and travel time data used in the arterial analyses were supplied by Glatting Lopez Kercher Anglin, Inc., Orlando; Barton-Aschman Associates, Inc., Fort Lauderdale; and the Seminole County Public Works Department.

REFERENCES

1. J. Koenig. Down to the Wire in Florida. *Planning*, Vol. 56, No. 10, Oct. 1990.
2. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 1985.
3. R. Hall. Concurrency Management for Transportation. *Florida Engineering Society Journal*, Vol. 44, No. 1, Sept. 1990.
4. T. Jackson. Transportation Concurrency: How Can It Be Achieved? *Florida Engineering Society Journal*, Vol. 44, No. 1, Sept. 1990.
5. W. McShane and R. Roess. *Traffic Engineering*. Prentice-Hall, Englewood Cliffs, N.J., 1990.
6. *Highway Capacity Manual*. U.S. Department of Commerce, 1950, pp. 130–132.
7. M. Wohl and B. Martin. *Traffic System Analysis for Engineers and Planners*. McGraw-Hill Book Company, New York, N.Y., 1967, pp. 168–175.

Publication of this paper sponsored by Committee on Transportation Programming, Planning, and Systems Evaluation.

Decentralization of Jobs and Emerging Suburban Commute

JONATHAN C. LEVINE

Large-scale suburbanization of employment has dramatically changed transportation and land use planning. Intersuburban commuting now dominates regional highway networks, and the automobile has replaced mass transit for many commutes. A study was undertaken to examine one aspect of the debate on the effects of employment decentralization on regional mobility: the impact of growing suburban employment on the commutes of people from various income groups. The study suggests that suburban employment centers with high levels of multifamily housing will exhibit commute patterns in which household income and commute distance are largely independent. In contrast, in suburban areas in which the development of dense housing has not kept pace with employment growth, it is hypothesized that new commute patterns are emerging wherein lower-income households commute greater distances than their upper-income counterparts. This pattern would reverse the prediction of monocentric urban models for central city employment. These hypotheses are tested for San Francisco Bay Area communities using data from 1981 and 1989. Bivariate analyses generally supported the predicted effects of community employment base and housing stock on commute patterns by income. Nested logit models of the household residential location decision were estimated for workers in San Ramon and in northern Santa Clara County on the basis of 1989 data. The models appeared to demonstrate a positive effect of the availability of multifamily housing on the residential location decisions of low- to moderate-income households. Forecasts of commute patterns using the estimated models indicated a potential for reducing long-distance commutes by low- to moderate-income households through a policy encouraging multifamily housing construction in the vicinity of suburban employment centers.

Rapid employment growth in many suburban areas during the 1980s markedly affected the range of available transportation planning strategies. Suburb-to-suburb commutes currently outstrip both intraurban and suburb–to–central city journeys to work; nationwide, a plurality of metropolitan commutes now begin and end outside central cities (*1*). Transit's mode share declines sharply with employment suburbanization because suburban employment locales are virtually impossible to serve by conventional transit because of scattered trip ends (*1–3*). Even ride sharing may become more difficult when increasing numbers of people work in sites with fewer nearby workers than in more traditional downtown employment settings.

Trends toward intersuburban commuting and greater reliance on the private automobile associated with employment suburbanization are not in great dispute. Much more controversial are the implications of these trends for transportation and land use planning as well as for the prospects for metropolitan areas in general. One point of view is that large-scale employment suburbanization harms the long-term viability of metropolitan areas by reinforcing automobile dependency and promoting environmental destruction through excess land consumption and air pollution (*4*). An alternative viewpoint is that decentralization is the force that renders accessible large metropolitan areas. By eliminating the need to commute from the metropolitan periphery to the central business district (CBD), employment suburbanization has kept commute distances in larger urban areas from growing to unmanageable proportions (*5*).

A third view accepts the inevitability of large-scale employment suburbanization but points to a systematic separation between suburban workplaces and suburban residences as a continuing impediment to regional mobility (*6–8*). Despite the traditional conception of the "suburb," some suburban communities have a large employment base relative to a limited housing stock, whereas others contain the reverse. Intersuburban commuting, now the dominant form of metropolitan journeys to work, is the result. The "jobs-housing balance" approach seeks to identify those economic and political forces that lead to deficits of housing near suburban employment centers, and to develop structures for planning and development that would generate a better geographic match between employment and housing. The geographic matching would presumably obviate the need for much of the intersuburban commuting that has been observed recently.

Separation of jobs and housing is a product of affordability and not merely space; imbalances between jobs and housing are most important for those households unable to afford scarce housing near suburban employment centers. The effects of employment suburbanization are thus best analyzed with specific reference to the various income groups affected. This study examines one aspect of the debate on the effects of employment suburbanization on regional mobility: the impact of growing suburban employment on the commutes of people from various income groups. The study poses three major questions: (a) Does suburban employment tend to favor the commutes of one income group or another? (b) If income-related commute patterns are evident among commuters to suburban employment centers, is there a relationship between observed commute patterns and characteristics of the particular suburban employment center being analyzed? and (c) Can policies allowing or encouraging development of higher-density housing near suburban employment centers reduce long-distance commutes by low- to moderate-income commuters?

These questions are analyzed with reference to the San Francisco Bay Area. Regional commuting characteristics are

---

College of Architecture and Urban Planning, 2000 Bonisteel Boulevard, University of Michigan, Ann Arbor, Mich. 48109.

first explored through a descriptive analysis of the Bay Area Travel Survey (BATS), a 1981 home interview survey of some 7,200 households (*9*). The effects of local conditions on commute patterns are then analyzed through the estimation of a discrete choice model of residential location based on a 1989 workplace survey of workers at selected large employers in northern Santa Clara County, the region's high-technology manufacturing center, and in San Ramon, a suburb on the Bay Area's eastern edge that expanded rapidly in office employment over the 1980s.

The 1989 survey was administered by mail and distributed to samples of employees of selected large firms in San Ramon, Mountain View, Sunnyvale, Cupertino, and San Jose. The overall response rate was 56.2 percent. Survey respondents were significantly wealthier than the population at large; the mean household income of survey respondents ranged from 90 to 200 percent of the population mean household income for their communities of residence (median = 130 percent). Because of income biases in the sample, caution must be used in interpreting results.

MODELS OF URBAN DECENTRALIZATION

The analysis of income-related patterns in metropolitan commutes was a central topic in much location theory of the 1960s and beyond (*10–12*), a literature that has been ably reviewed elsewhere (*13–17*). Under the standard assumptions of employment locating at the center of the metropolitan area, and an income elasticity of demand for residential space exceeding the income elasticity of commute costs, the prediction of this "monocentric" theory was a pattern of concentric rings of residential zones of increasing income radiating from the metropolitan center. Higher income was thus associated with longer journeys to work.

Later models analyzed decentralized or polycentric regions but diverged in their style of analysis. To a great extent this lack of a unified approach is attributable to the intractability of polycentricity (*17*). Some authors continued to derive monocentric models (*18–20*), whereas others restricted their prototypical urban area to one dimension (*17,21*) or allowed for only two centers (*22*). Researchers in the spatial interaction tradition focused on interzonal travel flows or systemwide transportation and location optima (*23,24*). With the exception of White (*25*), most nonmonocentric modelers did not deal explicitly with the question of spatial distribution of residences by income, a feature that was so prevalent in their monocentric precursors.

Another tradition of locational analysis emphasizes neighborhood characteristics and service differentials (over the commute distance–land price tradeoff) as determinants of residential location (*26–28*). This approach derives support from surveys in which households consistently rank factors such as quality of schools and safety and general appearance of neighborhoods as more important than workplace accessibility in determining their choice of residential location (*29,30*). This style of analysis is well suited to the description of dispersed urban areas because it refrains from making assumptions on the commute in the first place. But for most workers the commute remains a determinant of residential location, at least on a macro scale, and is logically a component of residential location analysis.

It has been suggested that urban decentralization models need to fuse the tradition of Alonso (*10*), which emphasizes elasticities of commute cost and space, with that of Tiebout (*26*), which emphasizes local service differentials (*13,31*). This is a realm in which the discrete choice family of models can excel and perhaps provide an empirical bridge between these two theoretical approaches. This is a result of the ability of discrete choice location models to analyze jointly two sets of characteristics: (a) attributes of potential residential locations such as public service levels; and (b) attributes arising from the interaction of individual households and communities, such as housing affordability or commute distance. This capacity represents a crucial difference between the discrete choice approach and the family of regression-based tools that tend to focus on the household or the community but rarely on both simultaneously.

For problems with more than one possible outcome (e.g., the selection of a single community from many possible communities), the most commonly applied discrete choice technique is the multinomial logit model. Following the notation of Ben-Akiva and Lerman (*32*), the multinomial logit model assumes that the probability that the utility of community i exceeds the utilities of all other communities j for household n (i.e., the probability that i is the chosen alternative) equals

$$P_n(i) = \frac{e^{Vin}}{\sum_{j \varepsilon Cn}} e^{Vjn}$$

where V is the deterministic component of a choice's utility and is a function of the attributes of the communities (e.g., school quality, tax rate) and attributes arising from the interaction of households and communities (e.g., income-housing price ratio, commute distance), and C_n is the set of communities available to the household or the "choice set."

The first major application of the discrete choice approach to locational modeling was Lerman's model of household locational choice between 145 census tracts in Washington, D.C. (*33*). In assuming a joint selection among so many alternatives, the pioneering work demonstrated the feasibility of a multinomial logit model of residential location but also showed the need for alternative representations of choice to comply with the "independence from irrelevant alternatives" (IIA) limitation of the multinomial logit model. To overcome this limitation, later models employed a hierarchical modeling structure (the nested logit model) in which secondary choices are modeled as conditional on primary-level choices. These hierarchical models have taken various forms, such as a locational decision conditional on a decision to move or to stay in place (*34*), or mode to work conditional on vehicle ownership, which in turn was conditional on residential location (*35,36*). Quigley modeled the choice of housing unit conditional on neighborhood selection, which was modeled conditionally on choice of town (*37*). In each of these cases, the utility of a location to an individual was modeled as a function both of attributes of the location and an interaction of locational and household characteristics.

Anas developed a model of the Chicago area rental market that differed from those referred to above in two important ways (*14*). First, the model used U.S. census data aggregated into 0.25-mi^2 zones over the Chicago metropolitan area rather than the disaggregate household level data used in other studies. The second difference of Anas's work is its analysis of both the demand and supply sides of the rental housing market. Using a utility-maximizing model for households and a profit-maximizing model for landlords, Anas derived a partial equilibrium model of the housing market, an accomplishment that previously had been the domain of primarily bid rent analyses.

COMMUTE PATTERNS IN BAY AREA SUBURBS

The view associating large-scale employment suburbanization with enhanced overall regional mobility is largely based on two premises. First, it is assumes that commutes to suburban locations are shorter in distance than those ending at the metropolitan center. The second premise is that employment suburbanization benefits different income classes equally, or that the ability of a household to reside close to its suburban workplace is largely unaffected by its income status.

Shortened Commutes in Suburbs?

The first premise, that of shortened commutes in the suburbs, is generally supported by the 1981 BATS data. Figure 1 presents median commute distances by superdistrict (34 aggregations of travel analysis zones encompassing the entire Bay Area). As expected, the longest commutes end in downtown San Francisco, with a median distance of 19.3 mi. The San Francisco CBD's position at the tip of a peninsula lengthens commutes significantly; the commutes of the Oakland workers represent more typical center city commutes, with a median distance of 12.2 mi.

Workers in suburban areas tend to enjoy shorter commutes; median trip distances are less than 8 mi in most areas. The commute benefits of suburbanization are not universal, however. Commutes of near–center city length are found among workers in the industrial suburbs of northern San Mateo County (12.1 mi) and Richmond (9.5 mi). Both areas contain concentrations of heavy industry and populations of low or lower-than-average incomes and may be viewed as being similar to center cities in their commute patterns. In contrast, the longer commutes among workers employed in northern Santa Clara County, the heart of the Silicon Valley high-technology manufacturing region (medium = 8 mi), occur against a backdrop of relatively cleaner electronics manufacturing and a higher-income population. It may be that the long median commute in this area is affected by the acute surplus of jobs over housing in the area and the necessity for much in-commuting (7).

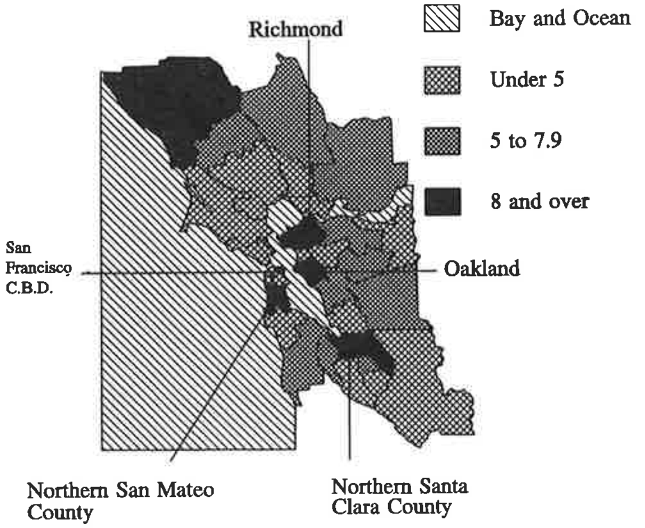


FIGURE 1 Median commute distance (mi) by superdistrict of employment, 1981.

Suburban Employment and Commutes by Income

The prediction of the monocentric model of increasing incomes of central city commuters with increasing commute distance is supported by the 1981 BATS data. Figure 2 maps the Pearson correlations (*r*) between household income and commute distances for primary workers employed in each of the Bay Area's 34 superdistricts. Although the magnitude of the correlations indicates that income has little explanatory power in predicting commute distances, the correlation between commute and income is positive as expected for the downtown areas of San Francisco (*r* = .28) and San Jose (*r* = .21) and for the city of Oakland (*r* = .08). The low correlations indicate that there remains a great deal of unexplained variation of income over the entire commuting range, not a surprising result given the great variety of settlement patterns in the cities as well as in the suburbs.

Although the three central cities all exhibit the expected positive correlation between commute distance and income, the pattern of relationships between income and commute varies more widely in suburban and exurban areas. Positive correlations appear in the fringes of the metropolitan area, such as eastern Alameda County (*r* = .35), eastern Solano County (*r* = .22), and the Santa Rosa area (*r* = .20), as well as in some of the industrial suburbs: Richmond (*r* = .27), northern San Mateo County (*r* = .25), and northern Santa Clara County (*r* = .09).

Within the Bay Area's inner ring, most suburban superdistricts exhibited independence between household incomes and commute distances in 1981. Exceptions to this were found in the same areas that exhibited longer-than-typical suburban

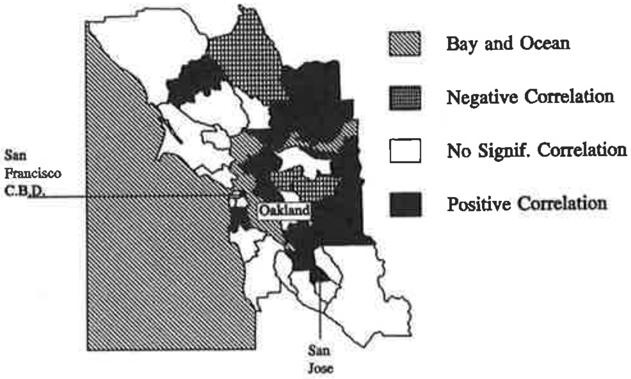


FIGURE 2 Correlations between household income and primary worker's commute distance, by superdistrict of employment, 1981.

commutes and thus may be somewhat "urban" in their commute characteristics: Richmond area (r = .27), northern San Mateo County (r = .25), and northern Santa Clara County (r = .09). Data from 1989 revealed no significant relationship between incomes and commutes for northern Santa Clara County.

In contrast, in two of the Bay Area's suburban superdistricts, higher-income workers lived closer to their workplaces than did those from lower-income households. These areas included the Interstate 680 corridor communities of Walnut Creek and Lafayette (r = −0.18) and Danville and San Ramon (r = −0.29).

Despite incompatibilities of data sources, 1989 and 1981 survey data revealed similar patterns; the correlation between income and commute distance remained negative and significant for the San Ramon workers, although once again without much explanatory power (r = −0.11) because of a high variance of commute distance at all income levels. Despite the low explanatory power of income on commute distance, a commute distribution histogram (Figure 3) reveals a clear pattern of commute distances by income; among the highest-earning households, 26.9 percent lived within 4 mi of their San Ramon workplace, whereas only 16.5 percent of those households earning up to $50,000 lived within so close a commuting range. The opposite pattern emerges when one considers the longest commute of 40 mi or more.

Thus the hypothesis of an income-neutral effect of employment suburbanization appears to hold in some areas and not in others. Given intersuburban differences, blanket statements about the effect of employment suburbanization on metropolitan commutes may be misleading. Instead, a finer-grained approach relating local conditions to the commute patterns to which they give rise may be more instructive from both a theoretical and policy standpoint. In particular, it may be that the effect of suburban employment on the commutes of people from different income groups is largely a product of local housing stock conditions. Initial analysis of housing stock conditions in the San Ramon area and in northern Santa Clara County are revealing; Silicon Valley communities have a significantly higher proportion of their housing stock in multifamily units (between 41 and 72 percent, compared with 27 percent for San Ramon). It is hypothesized that differences in the income-commute relationship between the two areas (negative for San Ramon and insignificant or positive for northern Santa Clara County) are in part a product of these differences in housing stock conditions.

MODELING APPROACH

Commute patterns are primarily the result of locational decisions by households and employers. This study models household locational decisions given employment at a fixed workplace or two fixed workplaces in the case of the dual-worker household.

Choice Set Development

The capacity of the multinomial logit model to analyze both characteristics of the individual and the community requires an explicit delineation of the set of communities from which the individual chooses. The feasible set of communities for all households in the sample—the choice set—was assumed to be those communities within a 60-min driving radius of the workplace, generating choice sets of 69 communities for the San Ramon employees and 52 communities for the northern Santa Clara County employees. The massive data sets generated by the large numbers of alternatives in the choice sets were reduced through a random sampling procedure proposed by McFadden (*38*).

Modeling a household's locational choice between such a large number of communities would strain the behavioral interpretation of the model; selection of one of 69 commu-

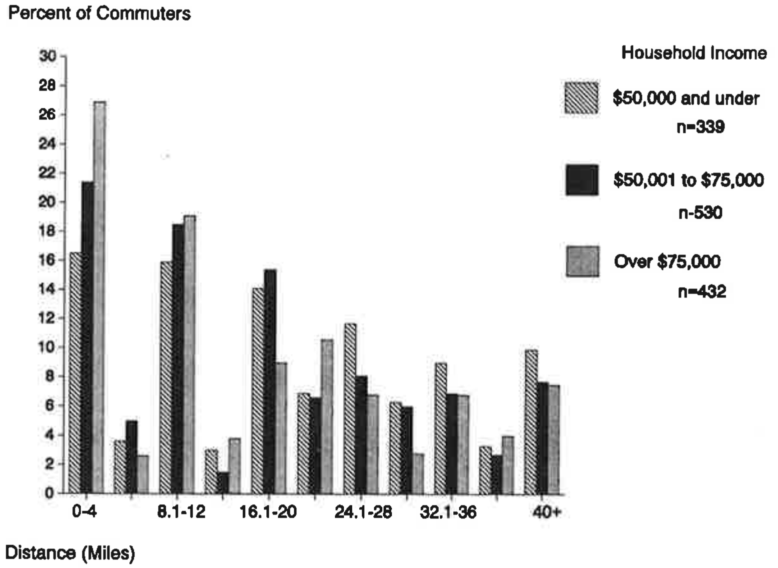


FIGURE 3 Commute distance distribution by household income group, San Ramon workers, 1989.

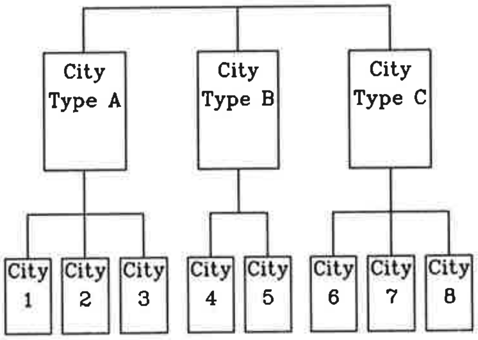


FIGURE 4 Diagram of nested structure of communities within community clusters.

TABLE 1 Grouping of Communities for Nested Analysis, Choice Set for San Ramon Workers

Median Single Family Home Price	Percent of Housing Stock in Multifamily	
	33 percent and under	Over 33 percent
$227000 and under	Antioch	Concord
	Benecia	Fairfield
	Brentwood	Hayward
	El Sobrante	Pleasant Hill
	Livermore	Richmond
	Manteca	San Leandro
	Martinez	San Pablo
	Newark	SJ: Alum Rock
	Oakley	SJ: Downtown
	Petaluma	Union City
	Pinole	
	Pittsburg	
	San Lorenzo	
	Suisun	
	Tracy	
	Vacaville	
	Vallejo	
$227,001-$307,000	Castro Valley	Alameda
	Dublin	Albany
	El Cerrito	Berkeley
	Fremont	Campbell
	Half Moon Bay	Colma
	Kensington	Daly City
	Milpitas	Hercules
	Pleasanton	Oakland
	San Ramon	Santa Clara
	SJ: Berryessa	SJ: Evergreen
	SJ: Cambrian/Blsm Hill	S San Francisco
	SJ: Zipcode 95125	Walnut Creek
Over $307,000	Alamo	Belmont
	Danville	Cupertino
	Lafayette	Foster City
	Los Altos	Los Gatos
	Orinda	Moraga
	Piedmont	Mountain View
	San Carlos	San Francisco
	Saratoga	San Mateo
		SJ: Westgate
		Sunnyvale

(SJ: City of San Jose)

nities is beyond the cognitive skills of most households. Moreover, modeling the locational decision in such a matter would undoubtedly violate the IIA limitation of multinomial logit. For these reasons, selection of community was modeled as conditioned on a prior selection of a generalized community type, or cluster (Figure 4). Clusters of communities were developed on the basis of housing price (i.e., three levels of median single-family housing prices) and density (i.e., two levels of multifamily housing stock) as described in Table 1. These dimensions were used on the assumption that price and density form the primary components of housing affordability; communities within a cluster should thus have similar affordability characteristics. As discussed later, results of the analysis confirm the appropriateness of this as a valid nested framework for modeling residential location decisions.

Modeling Results

Variables used in the analysis are described in Table 2; the choice modeled was residence in one of 69 communities. The models in Table 3 are designed to assess the impact of multifamily housing on the utility of a particular community to low- and moderate-income households. Independent variables include those pertaining to access, affordability, and community attributes. Some variables vary by community (SCHOOL, CRIME, CENTERDUMMY), others vary by household and by community (the commute time variables HTIME and LTIME, the affordability variables $SQFT/INC, %MULT:LO, %MULT:MED, and %MULT:HI, as well as TENURE/$ and MFCHILD), and others vary by household and by community cluster (MED$/INC, LOGSUM).

On basis of these categories of variables, the total utility to a household of a particular community and community cluster together is equal to

$$U_{imn} = \bar{V}_i + \bar{V}_{in} + \bar{V}_{mn} + \bar{\varepsilon}_i + \bar{\varepsilon}_{in} + \bar{\varepsilon}_{mn}$$

where

U_{imn} = total utility of community i and community cluster m to household n;
$\bar{V}_i$ = systematic component of utility common to all elements of C_n, using community i (the "‾" refers to a subset of the entire choice set C_n);
$\bar{V}_{in}$ = remaining systematic component of utility specific to combination (i,n);
$\bar{V}_{mn}$ = remaining systematic component of utility specific to combination (m,n);
$\bar{\varepsilon}_i$ = unobserved components of total utility attributable to community choice;
$\bar{\varepsilon}_{in}$ = unobserved components of total utility attributable to interaction of households and community choice; and
$\bar{\varepsilon}_{mn}$ = unobserved components of total utility attributable to interaction of households and community cluster.

Three models were estimated for San Ramon employees and three for workers in northern Santa Clara County. Nested logit models were estimated using the LIMDEP statistical package compiled on the Berkeley Cray X-MP/14 under un-

TABLE 2 Variables Used in Multinomial Logit Models

Access Variables	
HTIME:	Peak hour automobile travel time from accepted or rejected place of residence to place of work of the highest wage earner in the household.
LTIME:	Peak hour automobile travel time from accepted or rejected place of residence to place of work of the second wage earner in the household.
Affordability Variables	
$SQFT/INC:	Median 1989 price per square foot of single family homes in a community divided by total annual household salary (in thousands).
$MED/INC:	Median home price for all communities (within a cluster of communities) divided by household income.
%MULT:	Proportion of housing stock in a community in non-single family homes. This includes duplexes, apartment, condominiums and mobile homes. This variable was partitioned into three according to household income level: %MULT:LO equals %MULT for households with incomes of up to $50,000 and 0 otherwise; %MULT:MED was constructed similarly for households of incomes between $50,000 and $74,999; and %MULT:HI for households of incomes of $75,000 and above. Each of these variables was designed to measure the utility of multifamily housing in a community to households of a different income stratum.
TENURE/$:	For homeowners, equal to the number of years of residence at their current address divided by community median home price. The variable is equal to 0 for renters. This variable is designed to capture the effect of an inflating housing market; long term owners may tend to live in higher priced communities than newcomers.
Community Service and Amenity Variables	
SCHOOL:	Aggregated test results from California Assessment Program standardized testing for 1989. SCHOOL was the median of six scores: statewide percentile rankings for third, eighth and twelfth grades in reading and mathematics.
CRIME:	Residential and commercial burglaries per capita, 1989.
MFCHILD:	Equal to %MULT for households with children present, 0 for other households. This variable is designed to measure any disutility of multifamily housing in a community to household with children.
Centerdummy:	A center-city dummy variable, equalling 1 for residence in San Francisco and Oakland and 0 for all other cases.
Variable Used in the Estimation of Nested Logit	
LOGSUM:	A variable used in the nested logit model (also referred to as the inclusive value, or I), equivalent to the expected utility of the community level nest:

$$I_{mn} = \quad \ln \sum_{i \epsilon D_m} {}_{e}(V_i + V_{in})$$

where D_m is cluster of communities *m* available to household *n* (i.e., a subset of the full choice set C_n). The coefficient of this variable, μ, becomes an indicator of the appropriateness of the hierarchical structure assumed, with values of μ greater than 0 and less than 1 validating the postulated relationship.

The probability of a household *n* choosing community cluster *m* and community *i* becomes:

$$P_n\,(i,m) = \quad \frac{(e^{(V_{mn} + \mu I_{mn})})}{\sum_{m \epsilon D_m} {}_{e}(Vmn + \mu Imn)} \quad \frac{(e^{(V_{in} + V_i)})}{\sum_{i \epsilon Cn} {}_{e}(Vin + Vi))}$$

constrained maximum likelihood sequential estimation. The starting point for the modeling is the place of work; thus the "San Ramon model" refers not to a model of San Ramon residents but to San Ramon workers who may live virtually anywhere in the Bay Area and beyond. In each case, Model 1 represents a model using the full set of variables; Model 2 drops insignificant variables, as well as LTIME, the commute time of the secondary earner in the household. This accounts for the possibility that residential location decisions may be made with reference to a primary worker's place of employment, with a secondary worker seeking employment close to home (i.e., opposite direction of causation). The preferred models restore the LTIME variable, as well as significant variables from Models 1 and 2.

Model Evaluation and Interpretation

The initial hypothesis of a nested structure with community clusters forming the higher-level nest and individual communities forming the lower level was validated by modeling results. The coefficients of the LOGSUM variable (μ) were statistically discernable from unity in both cases. Thus joint

TABLE 3 Alternative Nested Logit Model Specifications

	San Ramon			Santa Clara		
Variable	Model 1	Model 2	Preferred Model	Model 1	Model 2	Preferred Model
Lower level nest (community choice) variables (t-statistics)						
HTIME	-0.0672	-0.0754	-.0677	-0.0687	-0.0725	-0.0680
	(-13.8)	(-17.0)	(-14.3)	(-12.5)	(-14.3)	(-12.71)
LTIME	-0.06310		-.0618	-0.0497		-0.0494
	(-13.3)		(-13.4)	(-7.95)		(-8.07)
$SQFT/INC	-0.6847	-0.524	-.6513	-0.5073	-.5605	-.5399
	(-5.68)	(-5.1)	(-5.86)	(-2.88)	(-3.59)	(-3.34)
%MULT:LO	3.38540	2.9235	3.4334	4.2524		4.4481
	(3.230)	(2.93)	(3.37)	(2.089)		(2.234)
%MULT:MED	not estimable			-0.6612		
				(-0.46)		
%MULT:HI	2.0393			-0.5383		
	(1.80)			(-0.39)		
Centerdummy	2.0813	2.1344	2.0287	3.7664	4.1276	3.6321
	(7.31)	(8.55)	(7.46)	(5.623)	(8.07)	(6.52)
SCHOOL	0.0102	0.0100	0.0122	0.02547	0.0266	0.0256
	(2.34)	(2.71)	(3.05)	(4.255)	(4.69)	(4.312)
MFCHILD	0.5729			-4.5821	-4.941	-4.981
	(0.44)			(-2.70)	(-3.7)	(-3.41)
TENURE/$	-0.253			2.5153		
	(-0.3)			(0.528)		
CRIME	-13.73			-1.069		
	(-1.2)			(-0.05)		
Model statistics: Lower level nest						
L*(0):	-1494.1	-1494.1	-1494.1	-570.6	-570.6	-570.6
L*(ß'):	-898.26	-1024.5	-901.96	-364.5	-406.6	-364.8
rho^2:	0.3988	0.3143	0.3963	0.3612	0.2874	0.3606
rho(bar)^2:	0.3914	0.3110	0.3923	0.3520	0.2787	0.3487
Upper level nest (choice of community clusters) variables						
$MED/INC	-0.100		-.1390	0.0840		
	(-3.9)		(5.39)	(3.13)		
LOGSUM	0.5988	0.8417	0.5979	0.3258	0.3398	0.3089
	(17.2)	(16.9)	11.65	(7.38)	(7.16)	(16.38)
(The t-statistics of the LOGSUM test H_o: ß=1 rather than H_o: ß=0.)						
Model Statistics: Upper level nest						
L*(0):	-1167.0	-1167.0	-1167.0	-536.0	-536.0	-536.0
L*(ß'):	-939.81	-965.98	-934.18	-495.3	-505.7	-504.4
rho^2:	0.1947	0.1723	0.1995	0.0759	0.0565	0.0589
rho(bar)^2:	0.1927	0.1714	0.1977	0.0722	0.0547	0.0570
Summary statistics for both levels						
L*(0):	-2661.1	-2661.1	-2661.1	-1106.6	-1106.6	-1106.6
L*(ß'):	-1838.1	-1990.5	-1836.1	-859.79	-912.3	-869.23
rho^2:	0.3093	0.2520	0.3100	0.2230	0.1756	0.2145
rho(bar)^2:	0.3044	0.2497	0.3071	0.2113	0.1702	0.2064
No. of Observ.	1,475	1,475	1,475	480	480	480

(i.e., non-nested) models of locational choice would have violated the crucial IIA assumption of the logit model and would thus have biased parameter estimates.

The coefficients of HTIME and LTIME are both negative and significant, indicating a disutility of commute time in community selection. It appears that although locational factors such as neighborhood amenity and school quality may rank uppermost in the locational considerations of households, the search for adequate environments is conditioned on the acceptability of the commute. The other half of the accessibility-price tradeoff is captured in $SQFT/INC, the median single-family price per square foot divided by income in thousands of dollars. The coefficient of this variable is negative and significant; controlling for commute distance and locational characteristics, households prefer housing that requires a lower proportional expenditure of income.

The statistical significance of LTIME, along with the fact that it perturbs neither the direction of HTIME nor its statistical significance, appears to indicate the importance of secondary workers' job locations as independent factors in households' residential decision making. Undoubtedly many decision-making patterns exist in households, including those that limit the location of the secondary worker's job according to a previously determined household location. But for a large number of dual-worker households, residential location seems to be determined with reference to both workplaces.

The positive coefficient of the CENTERDUMMY variable can be interpreted in two ways. First, it appears that center city living (i.e., in Oakland and San Francisco) on balance constitutes a draw given the variables measured here. A major part of the apparent attraction may be that the low standardized school test scores (SCHOOL) for Oakland and San Francisco are irrelevant to many of the well-paid workers in this survey because upper-middle-class members of these communities commonly send their children to private elementary and secondary schools. Another factor influencing CENTERDUMMY's significant positive coefficient is that the two San Ramon firms surveyed relocated to San Ramon from central locations in 1984. A number of people who still live close to their former workplaces in the central Bay Area may be unable or unwilling to move. Nonetheless, the effect of central city location remained positive in the Santa Clara County model, which did not include recently relocating firms.

The most important policy variable in the model is $MULT:LO, equal to percentage of a community's housing stock in multifamily housing for households with up to $50,000 income (and 0 otherwise). The significant and positive sign of $MULT:LO is interpreted as indicating the importance of multifamily housing in a suburban community to low- to moderate-income households. Although these results should be viewed with caution because of sample biases, the results appear to imply that when controlling for other factors, such as housing price and school quality, an increase in multifamily housing levels in a suburban community increases the likelihood of selection of that community by low- to moderate-income households. If the community is a job center, changes in the housing stock may also reduce commutes by these households. These effects are modeled in the next section.

Despite the apparent commute-reducing potential for such residential development, there are many sides to the jobs-housing balance complex. As evidenced by the negative utility of MFCHILD in the Santa Clara County model, increased residential densities can tend to repel households with children. Thus, a policy of increasing housing density may eventually suffer from decreasing or even negative marginal returns in its potential to reduce commuting. Replicating urban levels of multifamily housing in suburban employment centers may eventually incur the costs of central city–style development in congestion and in-commuting by larger households without the crucial transportation advantages of a central location. Results of this study are speculative in this regard because of income biases in the sample; further research on the potential deterrent effect of suburban density is needed. Yet the potential of denser development to repel some people from seeking lower-density environments may be mitigated by adequately planning and zoning multifamily housing. Such planning will ensure open space for residents, as well as privacy sufficient to afford them some of the amenities of the more remote single-family house they may now be foregoing.

CHANGES IN COMMUTES IN RESPONSE TO CHANGES IN HOUSING STOCK

One of the most important uses of the multinomial logit model in land use and transportation modeling is its potential as a forecasting tool. Using already-calibrated coefficients, attributes of the choice sets (or the households themselves) may be manipulated to predict roughly the range of potential land use and transportation system responses to policy stimuli. For the purposes of this analysis, travel times are assumed to remain unchanged even with changes in the housing stock; this factor is clearly a simplification because a densifying housing stock near suburban employment centers would surely create some localized congestion even if it were to reduce long-distance commuting. Single-family housing prices are similarly assumed to remain unchanged by policy changes.

Two such housing policy options are tested for San Ramon. The first raises the levels of multifamily housing in San Ramon and neighboring Dublin by 10 percentage points, to the point that multifamily housing represents 36.7 and 38.6 percent of the housing stock, respectively (no addition to the single-family housing stock is assumed). This is equivalent to adding 1,815 multifamily units in San Ramon and 1,095 units in Dublin, on the basis of California Department of Finance figures (*39*). The model is then used to explore the question, Under these revised conditions, how many more households may be expected to locate in San Ramon or Dublin than would locate there under current conditions? What other communities would be expected to house fewer San Ramon workers if the San Ramon housing area stock were changed? The second policy experiment entails boosting the multifamily proportion of these communities' housing stocks to 50 percent, a figure typical of the communities studied in Santa Clara County.

Results of the simulation are presented in Table 4, which summarizes changes for all those communities forecast to gain or to lose San Ramon workers under the alternative housing stock scenarios described earlier. The first column represents the forecast when San Ramon's multifamily housing stock equals 36.7 percent and Dublin's equals 38.6 percent of the total. The values in Column 1 represent that percent of the sample forecast to live in that community that did not reside there before; thus 1.7 percent of the sample (over and above those currently living there) would be forecast to opt for living in San Ramon under the new housing stock conditions. Similarly, half of 1 percent of the sample of San Ramon workers is forecast to opt against living in Benecia in response to the housing stock change.

TABLE 4 Cities Forecast To Gain or Lose San Ramon Workers Under Increased Multifamily Housing Scenarios

City	Percent Change, add 10% Multifamily	Percent Change, 50% Multifamily	Auto Travel Time, Minutes	Travel Distance, Miles
San Ramon	1.70	3.08	4	2
Dublin	1.03	1.95	10	6
Hayward	-0.06	-0.06	30	12
Antioch	-0.07	-0.20	39	29
Pleasanton	-0.09	-0.08	14	10
Pittsburg	-0.13	-0.28	35	26
San Leandro	-0.13	-0.13	25	18
Martinez	-0.14	-0.28	33	21
Livermore	-0.14	-0.15	24	17
Concord	-0.20	-0.19	29	17
Pleasant Hill	-0.39	-0.86	20	13
Benecia	-0.50	-0.80	31	23
Walnut Creek	-0.90	-1.73	15	9

Column 2 represents the forecast locational response when both San Ramon and Dublin include 50 percent multifamily housing in their housing stock. Columns 3 and 4 represent approximate automobile travel time and travel distance from each community to San Ramon.

The expected result of this simulation was a diversion of commuters from affordable but remote communities. With the exception of Walnut Creek, all communities forecast to lose San Ramon workers are at least 10 mi removed from San Ramon, indicating a considerable potential for the reduction of commutes. Even the exception to this rule—Walnut Creek—presents an interesting case. Walnut Creek is 9 mi from San Ramon; it offers expensive housing (median single-family home price = $300,000) but a high proportion of multifamily housing as well; 59 percent of Walnut Creek's housing stock is in multifamily housing. Thus, despite its high cost of housing, it represents a relatively affordable community from which moderate-income households would be drawn if more affordable housing were to be constructed in San Ramon and Dublin. The result tends to confirm the jobs-housing balance approach to transportation planning that presumes that people would select housing that reduces their commutes if it were available at affordable prices. The findings presented above may underestimate for three reasons the commute-reducing potential of multifamily construction in these communities. First, as described, communities falling beyond a 60-min drive from San Ramon were excluded from the analysis so as not to perturb the results of the model with potential commutes that represented relevant options for only a tiny minority of San Ramon commuters. Commuters from these communities were excluded from the forecasting as well so as not to impute characteristics on untested data. An explicit forecast as to the behavior of these groups when faced with increased supplies of multifamily housing would require a housing type model to capture explicitly the tradeoff between affordable but dense housing nearby and larger, remote single-family homes. But to the extent that they come from moderate-income households, these long-distance commuters may be amenable to living in potential multifamily housing near their workplaces.

Second, the results presented below are for the household's primary worker only. Those multiworker households that would be most amenable to nearby higher-density living would be the ones for whom the move to Dublin or San Ramon reduces both commutes, not just one.

Finally, the models understate the commute-reducing potential of multifamily development because they include only those primary wage earners actually working in San Ramon. A new condominium unit in San Ramon occupied by, for example, a Dublin worker in all likelihood represents a shortened commute compared with commutes from alternative residential locations. The model simulates the behavior of San Ramon workers only and thus fails to capture these potential effects.

POLICY IMPLICATIONS

Cervero asserts that "[t]he principal reason for jobs-housing mismatches is that ad hoc market forces have generally shaped suburban growth in most U.S. metropolitan areas" (7). He hypothesizes five forces leading to the imbalance, two of which are demographic trends (two-wage-earner households and job turnover) and three of which are the product of planning and public decision making rather than the market: fiscal and exclusionary zoning, growth moratoria and worker earnings/housing cost mismatches generated by fiscal zoning and growth ceilings. The problem thus does not appear to be not enough planning; instead, it appears to be a planning style that seeks a localized kind of environmental quality (defined as large-lot, single-family development) without full regard for more regional concerns. The problem may in fact be too little rein given to the market rather than too much. General plans and zoning ordinances typically do not define minimum residential densities but rather maxima. The policy expressed in the San Ramon Housing Element of restricting high-density housing to just one of the city's eight planning subareas may be seen as one example of this phenomenon (*40*). It may be that allowing developers to build more densely near suburban job centers is all the incentive needed to produce significant residential densification near many suburban employment centers.

Results of this study are in accord with the jobs-housing balance approach to metropolitan transportation planning; when concentrations of suburban employment are matched with sufficient affordable housing, households seek to reduce commutes. Importantly, this approach is strictly a voluntary, incentive-based system; it is based on harnessing individuals' own desire to reduce commutes instead of imposing travel or mode restrictions that would be politically unpopular and intrusive on individuals' lives.

The commute-reducing potential of affordable housing depends on its occupancy by employees of nearby job sites. Results discussed above imply that San Ramon workers may actually occupy less than a quarter of new housing in the San Ramon vicinity. Policies to generate acceptance of nearby housing on the part of local workers can hold significant benefits in commute-reducing potential. When developers build housing in a community, they may be indifferent to its occupancy by local workers or by commuters. In contrast, to the extent that commute reduction is a public goal, it is in the community's interest that housing be occupied by local workers rather than out-commuters. It is not difficult to envisage community-developer agreements that would stipulate the nature and extent of locally targeted marketing for newly constructed housing to attempt to boost the proportion of housing occupied by local workers.

CONCLUSION

The conditions modeled in this study represent a single point in time, a snapshot in a continuum of development patterns that are constantly evolving. Many of those communities that appear to be lacking in alternatives to the single-family home today already have added considerable amounts of multifamily housing over the past decade and may in fact be on a path toward development patterns that will afford households the kind of choices referred to here.

The future course of these communities is still open. It is becoming increasingly clear that decision makers in suburban communities must recognize trade-offs between high employment levels, low-density residential environments, and uncongested highways; these three traditional goals may not be achievable simultaneously. When communities offer a range of dwelling and commuting choices, individuals and households will respond in ways that can meaningfully improve the quality of living in metropolitan areas.

ACKNOWLEDGMENTS

The author wishes to acknowledge the support of the University of California Transportation Center for this research. Many thanks also to Elizabeth Deakin, Robert Cervero, William Garrison, and John Landis for comments on drafts of the doctoral dissertation on which this paper is based (*41*).

REFERENCES

1. A. E. Pisarski. *Commuting in America: A National Report on Commuting Patterns and Trends.* Eno Foundation for Transportation, Inc., Westport, Conn., 1987.
2. P. W. Daniels. Transport Changes Generated by Decentralized Offices. *Regional Studies*, Vol. 6, 1972, pp. 273–289.
3. P. W. Daniels. *Office Location and the Journey to Work: A Comparative Study of Five Urban Areas.* Gower, Kettering, England, 1971.
4. *Reviving the Sustainable Metropolis: Guiding Bay Area Conservation and Development into the 21st Century.* Greenbelt Alliance. San Francisco, Calif., 1989.
5. P. Gordon, A. Kumar, and H. W. Richardson. Congestion, Changing Metropolitan Structure and City Size in the United States. *International Regional Science Review*, Vol. 12, No. 1, 1989, pp. 45–56.
6. R. Cervero. *Suburban Gridlock.* Center for Urban Policy Research, Rutgers University, New Brunswick, N.J., 1986.
7. R. Cervero. Jobs-Housing Balancing and Regional Mobility. *Journal of the American Planning Association*, Vol. 55, No. 2, Spring 1989, pp. 136–150.
8. R. Cervero. *America's Suburban Centers: The Land Use Transportation Link.* Unwin Hyman, Boston, Mass., 1989.
9. *1981 Bay Area Travel Survey.* Crain and Associates, Menlo Park, Calif., 1981.
10. W. Alonso. *Location and Land Use.* Harvard University Press, Cambridge, Mass., 1964.
11. J. Kain. The Journey to Work as a Determinant of Residential Location. *Papers and Proceedings of the Regional Science Association*, Vol. 9, 1962.
12. R. F. Muth. *Cities and Housing.* University of Chicago Press, Chicago, Ill., 1969.
13. W. Wheaton. Monocentric Models of Urban Land Use: Contributions and Criticisms. In *Current Issues in Urban Economics* (P. Mieszkowski and M. Straszheim, eds.), Johns Hopkins University Press, Baltimore, Md., 1979.
14. A. Anas. *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Policy Analysis with Discrete Choice Models.* Academic Press, New York, N.Y., 1982.
15. M. Fujita. Urban Land Use Theory. In *Location Theory* (J. J. Gabszewicz et al., eds.), Harwood Academic Publishers, Chur, Switzerland, 1986.
16. T. De La Barra. *Integrated Land Use and Transportation Modeling: Decision Chains and Hierarchies.* Cambridge University Press, Cambridge, England, 1989.
17. H. Ogawa and M. Fujita. Equilibrium Land Use Patterns in a Nonmonocentric City. *Journal of Regional Science*, Vol 20, No. 4, 1980, pp. 455–475.
18. R. F. Muth. Models of Land Use, Housing and Rent: An Evaluation. *Journal of Regional Science*, Vol. 25, No. 4, 1985, pp. 593–606.
19. B. Brown. Modal Choice, Location Demand and Income. *Journal of Urban Economics*, Vol. 20, No. 5, 1986, pp. 128–139.
20. H. Cremer. Residential Choice and the Supply of Local Public Goods. *Journal of Urban Economics*, Vol. 27, 1990, pp. 168–187.
21. M. J. Beckmann. Spatial Equilibrium in the Dispersed City. In *Mathematical Land Use Theory* (G. Papageorgiou, ed.), D.C. Heath and Company, Lexington, Mass., 1976.
22. K. F. Wieand. An Extension of the Monocentric Urban Spatial Equilibrium Model to a Multicenter Setting: The Case of the Two-Center City. *Journal of Urban Economics*, Vol. 21, 1987, pp. 259–271.
23. S. H. Putman. *Integrated Urban Models: Policy Analysis of Transportation and Land Use.* Pion, London, England, 1983.
24. S. H. Putman. Mathematical Programming Formulations of Transportation and Land Use Models: Practical Implications of Recent Research. In *Transportation Research Record 1125*, TRB, National Research Council, Washington, D.C., 1987, pp. 39–47.
25. M. J. White. Location Choice and Commuting Behavior in Cities with Decentralized Employment. *Journal of Urban Economics*, Vol. 24, 1988, pp. 129–152.
26. C. Tiebout. A Pure Theory of Local Expenditures. *Journal of Political Economy*, Vol. 54, 1986, pp. 416–424.
27. P. Rossi. *Why Families Move.* The Press, Glenco, Ill. 1955.
28. J. Persky. Suburban Income Inequality: Three Theories and a Few Facts. *Regional Science and Urban Economics*, Vol. 20, No. 1, 1990, pp. 125–137.
29. M. A. Stegman. Accessibility Models and Residential Location. *Journal of the American Institute of Planners*, Vol. 35, 1969, pp. 22–29.
30. D. P. Varady. Influences on the City-Suburban Choice. *Journal of the American Planning Association*, Vol. 56, No. 1, 1990, pp. 22–40.
31. G. Palumbo, S. Sacks, and M. Wasylenko. Population Decentralization within Metropolitan Areas: 1970–1980. *Journal of Urban Economics*, Vol. 27, 1990, pp. 151–167.
32. M. Ben-Akiva and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press, Cambridge, Mass., 1985.
33. S. R. Lerman. *A Disaggregate Behavioral Model of Urban Mobility Decisions*, Ph.D. dissertation. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1975.
34. G. E. Weisbrod, S. R. Lerman, and M. Ben-Akiva. Tradeoffs in Residential Location Decisions: Transportation Versus Other Factors. *Transport Policy and Decision Making: An International Journal*, Vol. 1, No. 1, 1980, pp. 13–26.
35. J. C. Shin. *Residential Location Choice: A Nested Logit Model with Transportation and Housing Characteristics.* Ph.D. dissertation. Department of Civil Engineering, Stanford University, Stanford, Calif., 1985.
36. G. Harvey. *Residential Location and the Journey to Work in Suburban Households.* Department of Civil Engineering, Stanford University, Stanford, Calif., 1980.
37. J. M. Quigley. Consumer Choice of Dwelling, Neighborhood and Public Services. *Regional Science and Urban Economics*, Vol. 15, 1985, pp. 41–63.
38. D. McFadden. Modelling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models* (A. Karlqvist et al., eds.), North Holland, Amsterdam, the Netherlands, 1978.
39. *Municipal Housing Data Base.* California Department of Finance, Sacramento, 1989.
40. *1990 Housing Element Update.* Community Development Department, Planning Services Division, San Ramon, Calif., 1990.
41. J. C. Levine. *Employment Suburbanization and the Journey to Work.* Ph.D. dissertation. Department of City and Regional Planning, University of California, Berkeley, 1990.

Publication of this paper sponsored by Committee on Transportation and Land Development.

Location Planning for Companies and Public Facilities: A Promising Policy To Reduce Car Use

Erik J. Verroen and Gijsbertus R. M. Jansen

One promising instrument for reducing car travel is the coordination of land use and infrastructure planning. Traditionally, this coordination has been tried by encouraging high-employment densities near public transportation stations. A more sophisticated strategy is based on the observation that companies generate a mobility of persons and goods that varies with the type of company and, naturally, its location. Companies are classified according to their potential use of public transportation. The land use strategy presented essentially consists of locating companies with a high potential use of public transportation near public transportation facilities, and locating those with a low potential use near highway exits. The results of an empirical investigation into the relationship between the mobility generated by companies (the mobility profiles), the type of company, and the accessibility characteristics of the locations (the accessibility profiles) are presented. These profiles have been elaborated for practical use in regional planning by the Netherlands Organization for Applied Scientific Research. An overview of the main results of these studies is given: first, a tentative classification of firms is introduced; then, the typology of locations is defined and operationalized. Evaluation and demonstration results of the developed profiles are presented. It is concluded that, with the use of the profiles, more integrated and comprehensive policies for land use and transportation planning can be developed.

One of the main goals of current Dutch transport policy is to reduce the growth in car traffic. A promising way to achieve this is to encourage use of public transportation through better coordination between the planning of transportation facilities and land use, particularly for jobs. Industrial plants, public facilities, and offices for business or government all generate the mobility of persons and goods. The amount of mobility generated and the use of various transport modes depend heavily on the characteristics of these companies and their locations. It is well known that locating employment near railway stations and other public transit facilities enhances the use of public transportation. Many examples can be found that demonstrate the influence of the location of a company on the mode choice of commuters. The recent location of the University of Utrecht in the Netherlands, for instance, which moved from the city center to a peripheral suburban location (the "Uithof"), has led to an increase in the car share from 25 to more than 60 percent.

Clearly, suburbanization of employment may lead to a dramatic increase in car use. Locating companies near public transit facilities can reduce this undesirable effect. However, stimulating the development of locations well served by public transportation requires a balanced policy: a promising and innovative land use strategy exploits the differences between companies as to the mobility they generate. Attention should therefore be paid to the large variation among companies with respect to their potential use of public transportation and the role of the car in business travel and freight transport. Because space near public transit stations is limited, and because some companies depend heavily on road facilities, locations with excellent public transit facilities should be reserved mainly for companies with a high potential for using public transportation. Companies with low potential that are heavily dependent on road transportation and business travel by car can be better located near highway exits.

To implement this location policy in urban regions, a key instrument for regional planning has been developed in the Netherlands [see a general overview by van Huut (*1*)]. The planning instrument is based on two classifications: one of locations with respect to their multimodal accessibility characteristics (the accessibility profile) and another one of companies according to their mobility characteristics (the mobility profile).

To establish optimal locations for each type of company, several types of locations are distinguished. In the first concept of the planning instrument, the classification distinguished three basic location types:

- *A-locations*: locations that are highly accessible by public transportation. Examples of A-locations are major public transportation nodes such as central stations in the larger urban areas.
- *B-locations*: locations that are reasonably accessible by both public transportation and car.
- *C-locations*: locations that are defined as typical car-oriented locations. Examples can be found near motorway exits in fringe areas having poor public transportation access.

Figure 1 illustrates the difference between these A-, B-, and C-locations.

This paper presents the key results of various empirical studies carried out by the Institute of Spatial Organization, Netherlands Organization for Applied Scientific Research (INRO-TNO) that constitute the basis of this so-called ABC location planning instrument. Most of these studies have been commissioned by the Dutch Ministry of Transport and the Ministry of Housing and Planning. In the paper, we will pre-

---

Traffic and Transportation Unit, Institute of Spatial Organization, Netherlands Organization for Applied Scientific Research, P.O. Box 6041, 2600 JA Delft, the Netherlands.

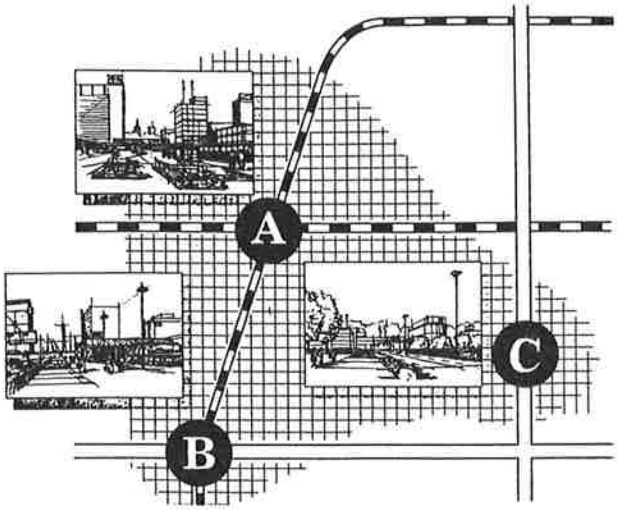


FIGURE 1 Concept of A-, B-, and C-locations.

sent the developed typology of companies and locations. The effectiveness of the ABC location planning instrument will be demonstrated on the basis of a simulation study in The Hague. The value of this planning instrument is evaluated on the basis of the experiences so far. Suggestions for further refinements are described.

MOBILITY PROFILES OF COMPANIES

Definition

A mobility profile describes the mobility generated by a company. The characteristics of commuting travel, visitor's travel, and freight transport are taken into account. Important indicators constituting the composite measure of a mobility profile are

- Employment density,
- Potential modal split shifts from car to public transportation and bicycles,
- Expected average level of car use in commuting travel,
- Car dependency of workers for business trips,
- Number of visitors, and
- Importance of the truck for freight transportation.

Research in the past has shown that these indicators depend heavily on the characteristics of the companies and their activities. Key factors are

1. The business activities at the particular establishment of the company, such as goods handling, type and amount of visitors, and business travel by staff, and
2. The socioeconomic characteristics of the workers (age, sex, income, level of education, working hours), which are also related to the type of company.

In principle, the mobility profile of a company does not depend on its location. The mobility profile refers to the average values. The influence of locational aspects is expressed in the indicators describing the modal split margins between A- and C-locations. The mobility profile is also independent of the size of a company. The aspect of space consumption is elaborated only in relative terms with the indicator for employment density (in square meters per employee).

Typology of Companies

Mobility profiles have been determined for different homogeneous classes of companies in the Netherlands (*2*) with comparable mobility characteristics. The starting point was a standard classification of companies and public organizations developed by the Dutch and Central Bureau for Statistics. Because mobility characteristics can vary substantially within one class of companies—for example, the combination of production and office activities within an industrial company—the standard classification has been modified somewhat for this study. This resulted in a classification with 62 types of companies, public organizations, and public facilities. For each of these company types, the values of the different indicators of the mobility profiles are estimated. As a first step, these values were based on general statistical data and relevant literature. In the second step, extensive empirical research was carried out among various companies and their personnel (e.g., interviews and surveys). The results were used to adjust the first concept resulting from Step 1. Finally, the developed mobility profiles were evaluated on their reliability and applicability.

To make the typology easier to handle, the 62 company classes were further clustered into a typology of 11 main company types, using the method of single pairwise clustering. Table 1 gives an overview of these company types and the estimated values of the various indicators of their mobility profile. The final estimations are based on a combination of empirical data, such as the Dutch National Travel Survey and additional surveys among many companies. The table also illustrates the indicators we were able to elaborate in the profiles with the available data. Because most of the data were related to commuting travel, most indicators describe the travel according to numbers of personnel. Travel by visitors and freight transportation are not worked out in detail. Clearly, further research on travel components in these categories would be welcome to improve the mobility profiles on this point.

The values in Table 1 are meant to roughly differentiate between the main company types. The indicators for the modal split margins between A- and C-locations have broad confidence intervals, and the modal split figures can vary with the service level of the public transportation system in various urban regions. In Table 1, the values are presented for the larger urban areas in the western part of the Netherlands (the Randstad).

ACCESSIBILITY PROFILES OF LOCATIONS

Definition

The accessibility profile describes the accessibility of the location for personnel, visitors, and goods with various travel modes. In view of the policy goals of the ABC location planning instrument, our main concern was to describe the ac-

TABLE 1 Characteristics in Mobility Profiles of 11 Major Company Types (*2*)

	Main company type:										
Indicator:	1	2	3	4	5	6	7	8	9	10	11
Employment Density (m2 per worker)	200	500	30	200	30	30	30	30	30	60	60
Public Transport Share (perc)	15	15	15	15	17	19	19	17	19	19	21
Public Transport Margins (perc)	15	15	15	15	25	30	30	25	20	25	15
Car Share (perc)	65	60	65	60	65	55	60	60	50	60	55
Car Margins (perc)	15	20	15	30	30	10	25	15	15	15	15
Car Dependency Workers (perc)	15	15	15	50	50	15	15	5	5	5	15
Share Slow Modes (perc)	24	24	24	32	24	28	24	24	36	28	32
Average Commuting Distance (km)	19	17	19	19	19	17	19	19	13	15	15
Visitors intensity (m2 per visitor)	450	900	150	450	150	450	50	50	150	15	50
Car Share Visitors (perc)	90	90	90	10	90	90	90	70	90	50	70
Importance Road Freight Transport	Great	Great	Mod.	Great	Mod.	Great	Small	Small	Mod.	Mod.	Small

Note:
1. Industrial plants with low density
2. Agricultural firms
3. Trade companies
4. Transport companies
5. Business offices with high car dependency
6. Industrial plants with high density
7. Business offices with low car dependency
8. Governmental offices
9. Social services
10. Public facilities
11. Medical facilities

Some definitions:
- The share of a mode is defined as the percentage of commuting trips made with this mode.
- The margins are defined as the differences in the share of a certain mode in commuting travel between A-type and C-type locations. For instance, if an average the PT share of a company type at A-type locations is 40 per cent, and at C-type locations 12 per cent, then the PT margins for this company type are 28 per cent.
- The car dependency of workers is defined as the percentage of the workers of a company which need the car during their working time for business trips.

cessibility by public transportation and by car. Slow modes were not taken into account explicitly in this stage.

We found that the distinction in A-, B-, and C-locations was too limited to give a meaningful and exhaustive categorization of all employment locations. Therefore, two additional location types were added to the typology: Al (A-local) locations, which are defined as locations reasonably accessible by public transportation and poorly accessible by car, and R-locations, which are considered to be poorly accessible by both public transportation and car. The resulting typology of locations by accessibility profile is summarized in Table 2.

Accessibility Indicators

During the past decades, an impressive number of methods has been developed to measure the accessibility of locations (*3,4*). To keep the accessibility indicators easy to apply in daily planning practice, we have chosen criteria based on the "egress" aspects of locations. The position of locations with respect to highway exits and public transit nodes are the most important indicators of the accessibility profiles. The criteria chosen can be briefly described as follows:

1. Easily accessible by public transportation:
   –Near a railway station with high service levels for both local and interlocal public transit.

TABLE 2 Profile of Typology of Locations by Accessibility (*2*)

	Accessibility by Public Transport:		
Accessibility by car:	Well	Reasonable	Poor
Well	A-type	B-type	C-type
Poor	A-type	Al-type	R-type

–No more than 800 m or 15 min away from a station entrance by urban public transport.

2. Reasonably accessible by public transportation:
   –Near a public transit node with several local public transit (rail) lines with high frequencies.
   –No more than 800 m away from a station entrance.
3. Easily accessible by car:
   –No more than 500 m from an urban arterial road.
   –No more than 2000 m from a motorway exit.
   –Limited parking restrictions on B-locations (at least one parking space for every four workers) and no parking restrictions on C-locations.

These criteria vary among various urban areas in the Netherlands, taking into account the general level of service of the public transportation system. In this paper, the criteria for the larger urban areas in the Randstad are presented.

A great advantage of the chosen method is that it takes only a look on a road map, a time schedule for public transit, and an inventory of the parking restrictions to determine the accessibility profile of a location. Clearly, this gives only a rough description of the actual accessibility, because total travel times, congestion, and travel costs for potential travelers to the location are not considered explicitly. Later we will evaluate the validity of the approach in greater detail.

RIGHT COMPANY AT RIGHT LOCATION

Given the mobility profiles of companies and the accessibility profiles of locations, we now face the question of what type of company ideally should be located at what type of location, given the policy goals we want to achieve. Which strategy will yield a maximum reduction of "avoidable" car travel and will guarantee the accessibility by car for companies that depend heavily on business travel by car or road freight transportation, or both?

This is the key question in the ABC location planning instrument. The answer is complicated. First, both governmental policy targets and company objectives must be considered in combination. These two can lead to conflicts of interest. Second, the locational behavior of companies is a complex process. Many factors are taken into account when companies choose their locations, of which accessibility is only one. Locational choice of companies can be influenced only partially by public policy.

To keep the method transparent, judgment of the suitability of certain combinations of companies and locations is based on the expected travel demand effects only. Regional economic and financial aspects are not considered in this stage. Given the different goals that must be achieved, a multicriteria approach is chosen. The suitability of a company type for a certain location type is determined by designing an "ideal" mobility profile for each location type. The more the mobility profile of a certain company matches this ideal profile, the more suitable this company is for the specific location.

Ideal profiles are theoretical mobility profiles that are considered to fit optimally with the accessibility profile of a certain location type. For instance, the ideal mobility profile of an A-location is a profile with high scores on indicators such as employment density, public transit shares and margins, car use margins, visitors per worker, and travel distances, and low scores on indicators such as car share, car dependency, car use by visitors, and the importance of road freight transportation.

TABLE 3 Preferable Locations for 11 Major Company Types (*2*)

	Preferable location type:	
Company type:	First priority	Second priority
1. Ind. plants, low dens.	C	-
2. Agricultural firms	C	R
3. Trade companies	B	C
4. Transport companies	C	-
5. Bus. offices, high car dep.	B	-
6. Ind. plants, high dens.	B	Al
7. Bus. offices, low car dep.	A	Al/B
8. Governmental offices	A	Al
9. Social services	B	Al
10. Public facilities	A	Al
11. Medical facilities	B	Al

Many versions of this multicriteria approach have been tested using various ideal profiles and weight factors for the 11 selected indicators in the mobility profiles. The results indicated that the chosen method is not very sensitive to these assumptions. For further details see Verroen et al. (*2*). Table 3 presents the results of the version that was judged best and that has been used in Dutch planning practice since the conclusion of our investigations. The table shows that the preferred location type for the 11 main company types varies substantially. This confirms the usefulness of the more sophisticated approach of the land use and transportation planning coordination as realized by the ABC planning instrument.

DEMONSTRATION OF EFFECTIVENESS IN REAL-WORLD CASES

The ABC location planning instrument has been developed to improve the coordination between land use and infrastructure planning for employment. Given this objective it is interesting to determine the distribution of companies in the Netherlands over the various location types in the current situation: How many companies are located on less suitable locations from the perspective of the transportation policy goals? Which company and location types are causing the greatest discrepancies? What are the alternative planning strategies to improve the adjustment of mobility and accessibility profiles? Which strategies are preferable?

To answer these questions, we carried out several simulation studies in different urban regions. As an example, the results of one case study for the region of The Hague (*5*) are summarized in Figure 2. The region of The Hague is in the highly urbanized western part of the Netherlands (the Randstad) and accommodates about 1 million inhabitants. Figure 2 illustrates the size of the problems in the Dutch urban areas today. In The Hague, it appears that more than half of the employment (57 percent) is located on unsuitable locations.

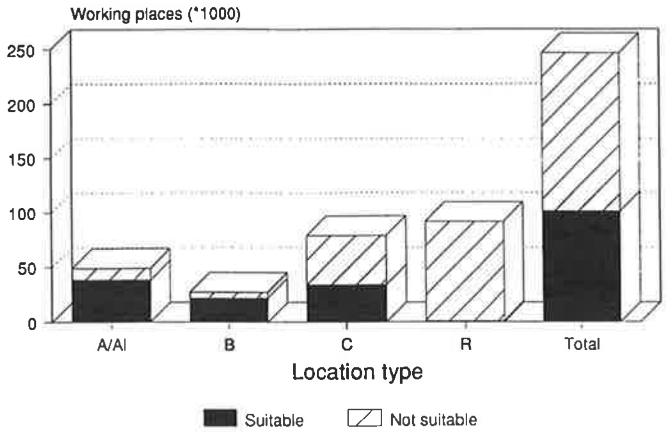


FIGURE 2 Current distribution of employment in The Hague over different location types and their suitability for these locations.

A minority of the jobs can be found at locations that are easily accessible by public transportation (A-, Al-, and B-locations). The majority of the employment is on C- and R-locations. Office buildings for business companies and governmental organizations (Types 7 and 8) and facilities for higher education (Type 10) are causing the adjustment problems between companies and locations. In terms of the ABC planning concept they should be located at least near public transit facilities. Case studies in other urban areas have shown similar results.

Given the problems in the current situation in The Hague, it is interesting to analyze ways to reduce the share of employment located on unsuitable locations. In principal, there are two policy options:

1. Infrastructure planning—improving the accessibility of companies at their current locations.
2. Land use planning—regulation of locational choice for new or relocating companies.

Figure 3 shows the possible effects of these two options for the region of The Hague in the next 15 years. It becomes clear that especially the assumed development of new A-, Al-, and B-type locations by improving the public transportation network with the investments planned for the region can have a substantial effect. The share of employment at unsuitable locations will decrease by about 20 percent, from 57 to 37. If this policy is combined with a strict land use planning strategy, allowing companies to relocate only in suitable areas according to Table 3, an extra reduction in poorly located employment of 10 percent can be achieved.

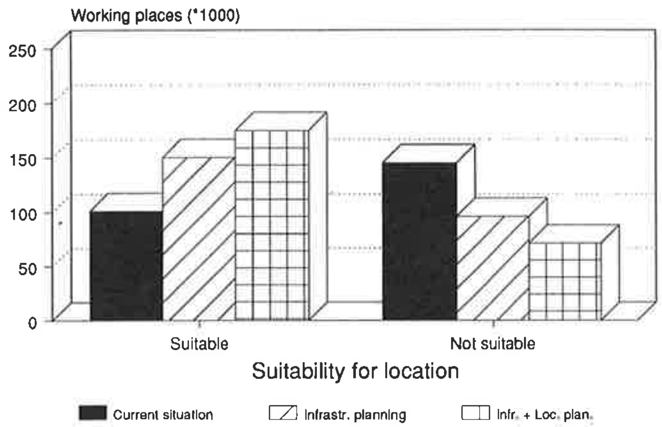


FIGURE 3 Estimated effects of two policy options on share of employment on unsuitable locations.

At this stage, it is not easy to accurately estimate the overall effects of the policy options on the share of the car in commuter travel in The Hague. Rough calculations indicated that the reduction could be about 8 percent (from 50 to 42 percent).

EVALUATION OF ABC LOCATION PLANNING INSTRUMENT

Modal Split Effects

The main objective of the ABC location planning instrument is to stimulate a shift in the modal split from car to public transportation. The effectiveness of the instrument depends heavily on the margins in the use of these modes (differences in average share) between company and location types. Table 4 gives an impression of the margins we found in our surveys in The Hague. In the table, the companies we analyzed are categorized into three groups: companies with mobility profiles most suitable for A-locations (Types 7, 8, and 10), companies with mobility profiles suitable for B-locations (Types 3, 5, 6, 9, and 11), and those suitable for C-locations (Types 1, 2, and 4). AL- and R-locations were not taken into account.

The table shows substantial differences in both companies and locations in the use of the car and public transportation. The average share of the car varies by about 20 percent between both company types and location types, proving that it is relevant to incorporate both company and location characteristics in location planning strategies. The table also shows that the differences in modal split between A- and B-locations are much larger than the differences between B- and C-locations. Apparently, car use is influenced more by accessi-

TABLE 4 Comparison of Commutes by Car and Public Transport for Various Company and Location Types in The Hague (*2*)

	Accessibility profile:			
	A	B	C	Average
Car				
Mobility profile most suitable for location type (see table 3):				
A	32	58	55	48
B	44	72	54	57
C	59	71	76	69
Average	45	67	62	-
Public Transport				
Mobility profile most suitable for location type (see table 3):				
A	41	15	12	23
B	37	17	14	23
C	13	11	6	10
Average	30	14	11	-

bility to cars than by accessibility to public transit. This conclusion is in accordance with other research findings in the Netherlands.

Intraclass Variations

Mobility Profiles

Our analyses show that in practice there is a considerable variation in the characteristics of individual companies and locations. The typology developed only explains a part of this variation. The results of multivariate analyses indicate that the classification of companies explains about 50 to 70 percent of the total variation in the different indicators that are part of the mobility profile (Figure 4). The typology offers only a rough classification. We conclude that the mobility characteristics of individual companies can differ substantially from their class average. Restrictive planning regulations should therefore be based ideally on the individual scores of companies with respect to indicators and not only on the average values of their group.

Accessibility Profiles

The same conclusion about variation is valid for the accessibility profiles developed. We conclude that the homogeneity of the classification can be further refined. Surveys among the personnel of companies at various locations showed that the ratio between car and public transit travel times for commuters may vary strongly within one location type (see, for example, Figure 5). Similar variations were found with respect to the parking facilities for personnel (Figure 6). The indicator for parking restriction used in Figure 6 is the sum (divided by 3) of the percentages of the personnel who have no parking spaces on private grounds, who must pay for their parking spaces, and who must walk more than 250 m to their workplaces. Although the average parking restrictions tend to be as expected, the restrictions vary substantially between companies at the same location type.

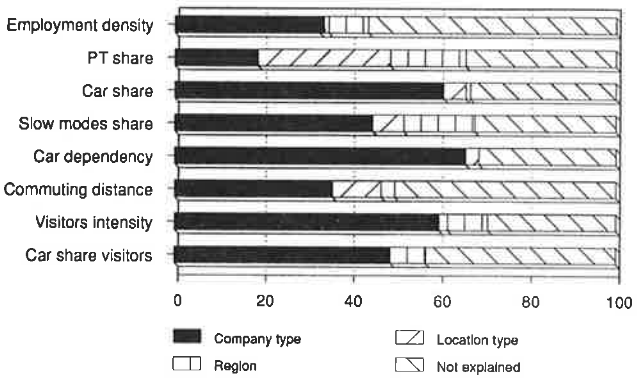


FIGURE 4 Variation in some indicators of mobility profiles with classification of companies.

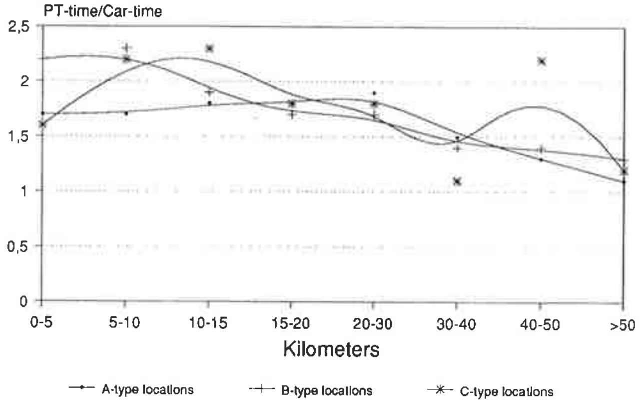


FIGURE 5 Travel time ratio between car and public transportation in The Hague for various location types and distance classes.

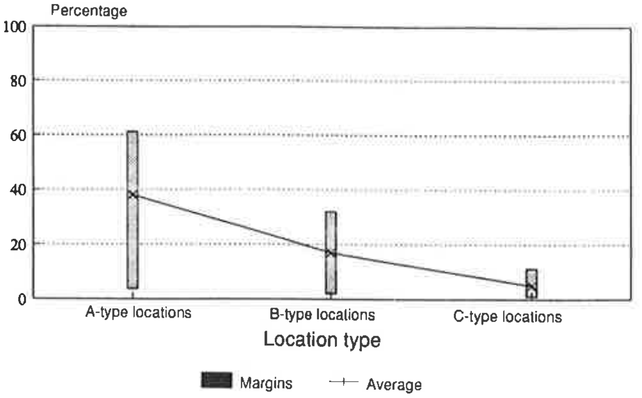


FIGURE 6 Parking restrictions on various location types in The Hague.

Suggestions for Further Research on Accessibility Profiles

There are several ways to further refine the measurement of accessibility. We already indicated the need for differentiation by type of urban area. There is a large variation in the level of service in public transportation between urban areas. In regions with poor public transportation facilities, public transportation tends to be no alternative for the car. In these regions, the stimulation of slow modes (e.g., cycling and walking) and carpooling appear to be promising additional strategies to reduce car traffic. It is therefore interesting to take these modes into account more explicitly in the accessibility profiles.

Another refinement can be achieved by incorporation of travel distances in the accessibility profiles. This would make it possible to deal with locational strategies that are based not only on modal split shifts but also on the reduction of car kilometers by reducing the home-to-work distances, for instance by improving the proximity of employment locations to the major population centers.

Given these possibilities for further improvement, more sophisticated methods can be elaborated to measure the accessibility of locations. At the moment of writing, INRO-TNO is working out more accurate criteria for accessibility profiles. As far as commuting is concerned, these methods are based on centrality indicators for accessibility (in travel times) and proximity (in travel distances) of locations. The following indicators are used:

$$B_{jvm} = \frac{\sum_i I_i * p_m(T_{ij}) * T_{ijv}}{\sum_i I_i * p_m(T_{ij})}$$

$$D_{jm} = \frac{\sum_i I_i * p_m(T_{ij}) * d_{ij}}{\sum_i I_i * p_m(T_{ij})}$$

where

B_{jvm} = average accessibility of zone i with mode v for purpose m,
D_{jm} = average travel distance to zone i for purpose m,
I_i = number of inhabitants, zone i,
$p_m(T_{ij})$ = distance disutility for average travel generalized time between zone i and zone j for purpose m,
T_{ijv} = travel time between zone i and zone j with mode v, and
d_{ij} = distance between zone i and zone j.

With these indicators, the competitiveness of alternative modes to the car and expected travel distances can be determined for use in classification criteria of locations, as well as their suitability for certain company types. For business travel of personnel and visitors, the measurement of accessibility will be based on graph theory (*6*). The network position of locations in relationship with other economic centers will be described with the Shimbel index.

APPLICATION OF PLANNING INSTRUMENT

Appropriate Policy Measures

The example of The Hague proves that both the improvement of public transportation supply and land use control can be effective. The ABC location planning instrument can be used as a planning method for improving the public transportation system and selecting infrastructure improvements. The instrument can be implemented in several ways in land use planning. First, Dutch urban and regional planning laws offer several possibilities for land use regulation. The activities permitted on certain locations can be restricted, although it is juridically not easy to distinguish among various types of employment. It is also possible to regulate land use by building regulations, such as space-floor indexes, building heights, parking facilities, and so on. Still another option is a pricing policy for selling and hiring land, to influence market price developments.

So far, it has not become clear which measures are most appropriate for implementing the ABC location planning instrument in daily planning practice. Local and regional governments are just starting to work with the instrument. At the moment, their main concern is to identify the various location types within their region. At this stage, local authorities still have a lot of freedom in how they should interpret and implement the planning instrument.

Political Acceptance

The introduction of the ABC location planning instrument triggered some interesting political discussion in the Netherlands. Companies and local governments were especially critical, each arguing from the perspectives of their own interests. Companies and chambers of commerce stated their concerns over limitations to their freedom to select the location they regard most appropriate for their activities. An inadequate location is seen as a threat to the profitability of the companies. Local governments are especially afraid of conflicts over regional economic development plans because of the important role these plans have in the acquisition of companies in competition with other regions. Land use restrictions based on transportation policy goals can (in the short term) damage the competitiveness of certain regions and cities.

Discussions are still ongoing about the ways in which the ABC location planning instrument can be further improved. In particular, the definition and role of the B-locations cause some concern. Because of their accessibility by car, B-locations are, on the one hand, more popular with companies (and therefore local governments) than are A-locations. On the other hand, stimulating B-locations is less effective because of the small differences in modal split compared with C-locations. A-locations should do the job but are not attractive enough. An oversupply and liberal use of B-locations may therefore decrease the effect of the planning instrument.

Given these political and methodological backgrounds, it might be interesting to refine the instrument with regulation methods based on mobility impact analysis. This means the development and setting of standards for the generation of car kilometers by various company types in various regions. Whether and by what means they fulfill the standards—by locational choice or other measures—is the responsibility of the companies. Mobility impact fees can stimulate the companies to meet their standards. This approach has much in common with policies for transportation demand management in certain parts of the United States [for an extensive overview, see Ferguson (*7*)]. It is interesting to learn from experience when further developing the ABC location planning instrument in the Netherlands.

CONCLUSIONS

The ABC locational planning instrument for companies and public organizations appears to be a promising and challenging concept. Proper application in infrastructure and land use planning may result in a substantial reduction in car use in commuting travel. Compared with traditional land use control, the use of mobility and accessibility profiles offers the possibility of a more balanced and sophisticated location planning strategy that takes into account the large variation between companies and locations. The concept has been put into operation and tested empirically for the Dutch situation. We are convinced, however, that the general approach also can be of great use to other countries that are confronted with environmental problems and congestion in urban areas.

Practical experience indicates that the ABC location planning instrument has several possibilities for further refinements. The accessibility profiles especially deserve further elaboration. Nevertheless, there will always be political opposition against the planning instrument because it tries to limit the freedom of choice of companies and local governments. This will also be the case when the planning instrument is aimed more explicitly at the main policy goal (reduction of car kilometers)—for example, use of mobility impact fees and mobility impact standards.

ACKNOWLEDGMENTS

The research on which this paper is based was performed under contract to the Projectbureau for Integrated Transport Studies, Ministry of Transport and Public Works and the National Planning Agency, Ministry of Housing and Planning. We would like to thank our colleague Hans Hilbers for his assistance in the preparation of the paper.

REFERENCES

1. H. van Huut. The Right Business in the Right Place. Presented at the PTRC 19th Summer Annual Meeting, University of Sussex, Brighton, England, 1991.
2. E. J. Verroen, M. A. de Jong, W. Korver, and G. R. M. Jansen. *Mobility Profiles of Companies and Public Facilities*. Research Report 1990–03. INRO-TNO, Delft, the Netherlands, July 1990.
3. J. M. Morris, P. L. Dumble, and M. R. Wigan. *Accessibility Indicators for Transport Planning*. Australian Road Research Board, Victoria, 1978.
4. J. G. Koenig. Indicators of Urban Accessibility: Theory and Application. *Transportation*, Vol. 9, 1980, pp. 145–172.
5. E. J. Verroen and H. D. Hilbers. *Accessibility Profiles in the Metropolitan Areas of The Hague and Eindhoven: A Survey on the Employment at A-, B-, and C-Type Locations*. Research Report 1990–09. INRO-TNO, Delft, The Netherlands, August 1990.
6. P. Hagget. *Geography, A Modern Synthesis*. New York, N.Y., 1979.
7. E. Ferguson. Transportation Demand Management: Planning, Development, and Implementation. *APA Journal*, Aug. 1990, pp. 442–456.

The opinions and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the Dutch authorities.

Publication of this paper sponsored by Committee on Transportation and Land Development.

Areawide Estimation of Vehicle Trip Generation Rates for Route 9A Reconstruction on Manhattan's West Side

JOHN C. FALCOCCHIO AND ROBERT M. MICHEL

The Route 9A reconstruction project on the west side of Manhattan, New York City, required the projection of trip generation by land uses that would be developed by the project build year and beyond. Because new development would vary greatly in its accessibility to public transportation, socioeconomic character, size, and other attributes, it was decided that available, standard trip generation rates were inadequate to capture trip generation variations within a major land use category. In Manhattan, residential and office land uses account for 75 percent of existing developed square footage and 90 percent of the new square footage to be developed by the build year. By isolating zones of pure residential or office land use (95 percent of developed square footage in the respective land use), it was possible to analyze existing trip generation rates developed from origin-destination surveys and to infer trip generation rates. These rates were correlated with various zone characteristics such as subway access to develop formulas and trip generation rates that more accurately capture the trip-generating characteristics of developments in various locations. The trip generation rates developed are currently in use on the Route 9A project.

Trip generation factors are one of the most important elements of the traffic forecasting process. Sophisticated models for assigning vehicular traffic (to the highway network) and assessing roadway operating conditions are ultimately only as accurate as the input fed into them.

The most common usage of trip projection estimates is for facilities planning or environmental assessment. For any type of land development project, features such as parking facility size, access points, driveway size, and even project location may depend on traffic generation projections. For transportation projects, trip projections serve to size facilities and determine geometry and the potential number and characteristics of access points. Trip generation estimates in the context of growth management and environmental assessments determine where a significant impact may occur. If estimates are too low, unforeseen effects may occur that would affect a community's health, development, and general well being. When estimates are too high, projected effects may be mitigated by measures that are inappropriate, excessive, or unnecessarily expensive.

For individual land development projects, wide variations are possible in the calculation of vehicle trip generation rates. The total number of trips generated in the peak hour is dependent on the interaction of several factors: the total daily trip generation, the temporal and directional distributions, the modal split, and the vehicle occupancy rate. For each of these factors there is a range of reasonable values. When used in combination to derive a trip generation rate, the range of rates resulting from the factor's use is also broad. Typically an analyst will use factors and rates from surveys conducted by others or, when a small number of projects are being studied, the analyst may choose to do spot surveys of trip generation at existing similar developments.

Large-area or corridor studies with diverse land use, neighborhood, and transportation conditions, however, demand a different approach. Typically a single trip generation rate from a general source such as the *ITE Trip Generation Handbook* might be used for a land use under a variety of conditions. The Route 9A replacement study in the west side of Manhattan, New York City (NYC), provided a test case for deriving more appropriate trip generation rates based on the results of an origin-destination survey and on other generally available data. The Route 9A study required that trip generation rates be developed for land uses in a wide variety of neighborhoods and under a variety of conditions. Modeling to predict the potential impact of the construction of Route 9A would require determining what factors were important determinants of trip generation rates and then measuring how these factors would affect the trip generation of a specific type of land use. This paper presents the process used to derive those estimates and the results of those efforts.

BACKGROUND

Manhattan's original West Side Highway was an elevated structure running from the Battery at the southern tip of the island north to West 72nd Street. Deterioration in the highway became significant over the years and ultimately led to the collapse of a portion of the structure. In 1974, the West Side Highway was closed from the Battery to 59th Street. Subsequent efforts to identify alternatives for a replacement roadway resulted in a range of options developed by the New York State Department of Transportation (NYSDOT) in concert with the NYC departments of Transportation and City Planning. A task force appointed to look at the options also developed an alternative of its own.

Beginning in 1988, a team of consultants began preliminary design and environmental studies. The purpose of these studies was to produce the data that, among other purposes, would explore preliminary design issues, flesh out the potential al-

---

Urbitran Associates, Inc., 71 West 23rd Street, New York, N.Y. 10010.

ternatives, indicate the need for modifications or additions to the alternatives proposed, and assess the potential effects the proposed alternatives had on the environment. Central to each of these issues was the question of how much vehicular traffic would have to be accommodated.

The study of traffic for the Route 9A environmental impact statement required the collection of massive amounts of data. The data collection process was as follows:

1. The entire Manhattan central business district (CBD) study area was divided into just over 500 zones, generally containing two to four blocks and having approximately 1.9 million ft^2 of development on average. Origin and destinations outside the study area were also aggregated into zones.
2. Existing automobile trips generated by each zone were determined. Surveys were conducted of drivers entering the study area at its boundaries. Overall, 14.8 percent of the 484,996 drivers crossing the cordon line during the study period responded to the survey. The data were tabulated so that zone destination could be identified. The sample data were then expanded based on the sampling proportion of total trips. Internal-internal trips were obtained separately by coordinating surveys within the study area of sampled parking facilities.
3. Taxi trips were surveyed from trip sheets for medallion cab companies, reflecting 15 percent of the total daily trips. The origin and destination zone of each trip, with or without fare, was tabulated. Trips to or from each zone were expanded to reflect the sampling proportion.
4. The automobile and taxi survey data were supplemented by a survey of commercial vehicles. The survey logs had space to record a complete day's trips. The logs were distributed to and completed and returned by drivers as they crossed the cordon line into CBD locations.
5. NYC Department of City Planning records were used to compile a zone-by-zone listing of land use square footage in each of 20 general categories.
6. NYC Department of Transportation records were used to compile a listing of the parking facilities and number of spaces available in each zone.

Although census data and maps showing transit access points were also used, the data formed the core of information for determining vehicular trip generation rates.

THEORETICAL FRAMEWORK AND METHODOLOGY

Within the Manhattan CBD study area approximately 75 percent of the existing developed floor area falls into either the residential or the office category. The remaining floor area is divided between 18 other categories. Projected new development, for which the trip generation rates will be used, shows an even greater concentration (90 percent) in residential and office uses. Those two land use categories account for so much of the developed area that it is possible to identify zones in which there is virtually only one land use. For the purposes of the study, zones in which approximately 95 percent of the built floor area was office or residential were defined as "pure use zones."

Taking the trip tables developed from the survey data and the land use square footages for the pure use zones, trip generation rates were calculated for each zone. Specific zone characteristics that would explain the variation in zone trip generation rates were then sought.

CALCULATION OF PURE RESIDENTIAL ZONE TRIP RATES

Automobile Trips

Method

The Route 9A study area contains 49 pure residential zones. The relationship between the number of automobile trips generated in each zone and the square footage of development is shown in Figure 1*a*. It is evident from the figure that a given amount of development can produce a wide range of automobile trips. Among the pure use zones, for example, the p.m. peak-hour trip rate could fall in the range of from 0.024 to 0.227 trips per 1,000 ft^2. We sought to reduce the amount of unexplained variance by testing the effects of subway access, parking availability, and neighborhood type (as a proxy for income-socioeconomic factors). Various variables were tested, both singly and in combination. Where appropriate, regressions were used to seek relationships. Our testing program included the following:

Parking Availability/Density We hypothesized that parking availability should be correlated with increased trip generation rates and tested different parking conditions:

- Parking versus no parking in zone;
- High versus low parking density in zone;
- Number of spaces in the zone;
- (Number of spaces in the zone)/(developed square feet in the zone);
- (Number of spaces in the zone and adjacent zones)/(developed square feet in the zone and adjacent zones); and
- (Number of spaces in the zone and adjacent two zones)/(developed square feet in the zone and adjacent two zones).

No relationship could be established at either the zone or neighborhood level of detail between automobile trips and parking. To investigate why there might be none, we looked at the balance between parking space demand and supply. The demand-supply comparison is presented later.

Subway Accessibility We correlated subway accessibility of zone with trip generation rates under various measures of subway accessibility on the assumption that with convenient subways, automobile use would drop. The subway accessibility measures examined included

- Average number of subway stations within ¼ mi of each zone in a neighborhood. Each separate subway station is counted, although some may serve more than one line, whereas others may serve lines that can be accessed at multiple locations within the ¼-mi radius.

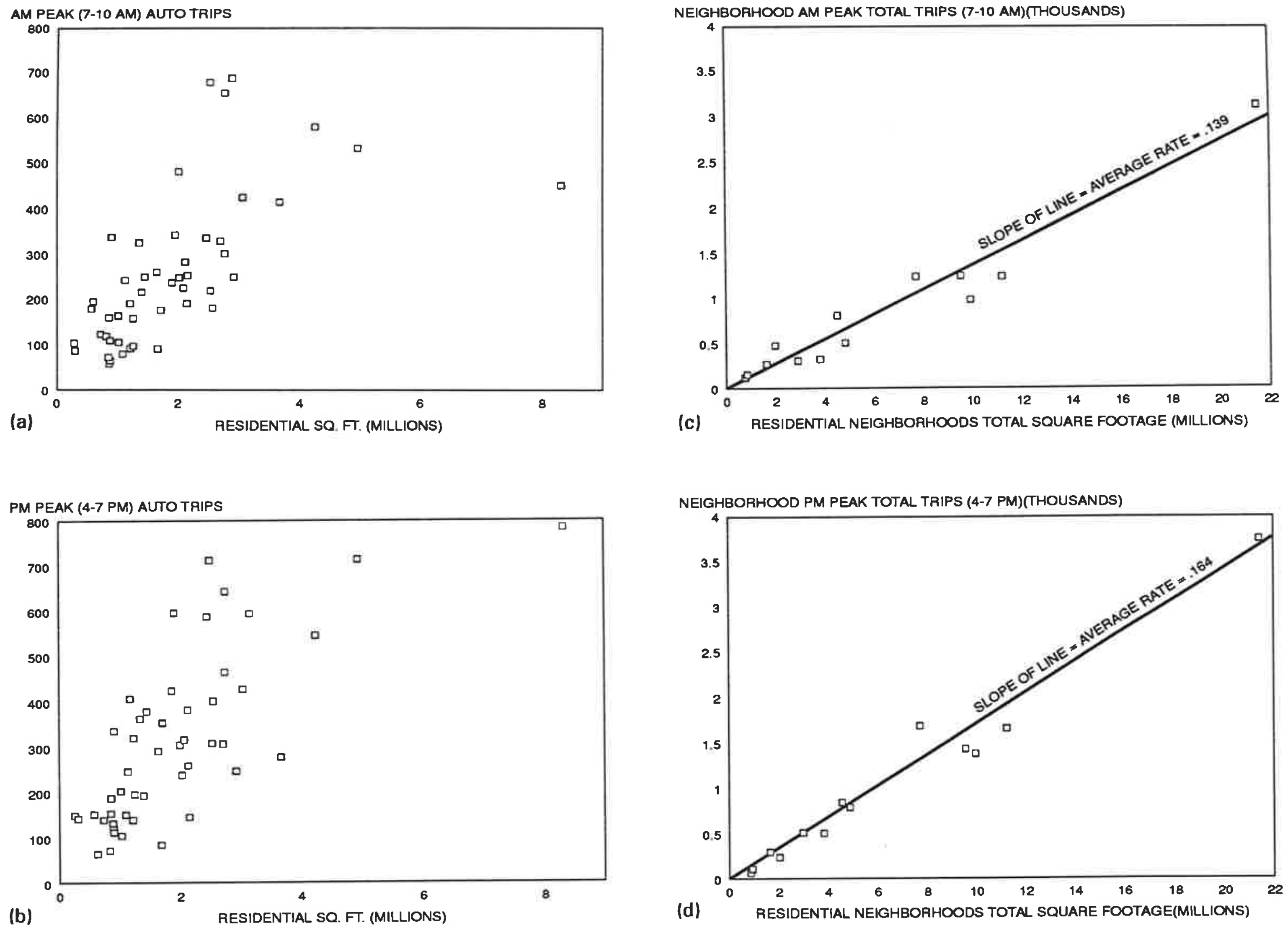


FIGURE 1 Relationship between number of automobile trips generated in each zone and square footage of development; (*a*) pure residential zones, a.m. peak automobile trips; (*b*) pure residential zones, p.m. peak automobile trips; (*c*) residential neighborhoods, a.m. peak trips; (*d*) residential neighborhoods, p.m. peak trips.

- Average number of subway lines accessible within ¼ mi of each zone in a neighborhood.
- Subway station access versus no subway station access within ¼ mi of zone.

It was found that those zones without subway access had a markedly higher average trip generation rate than those with subway access.

Neighborhood Groupings Neighborhood groupings were established as proxy for income and socioeconomic factors. The tests described earlier were performed for various groupings:

- 18 neighborhood areas centered around clusters of pure use office or residential zones.
- Uptown (north of 14th Street) versus downtown (south of 14th Street).

The neighborhood groupings of zones resulted in less variability in trip rates and more easily discernable patterns (compare Figures 1*a* and 1*b* to 1*c* and 1*d*).

Results

A.M. Peak Hour For zones with subway access, the most accurate estimates of existing trip ends were derived by using an average trip generation rate of 0.070 trips/1,000 ft² for zones without subway and a rate of 0.053 trips/1,000 ft² for zones with subway (see Figures 2*a* and 2*b*).

P.M. Peak Period Because the p.m. regression against subway station density yielded weak results, we used average trip generation rates for the 1-hr peak period shown in Figures 2*c* and 2*d*. Scatter patterns fit the average with moderate success when neighborhoods are stratified by subway versus no subway. For zones with subway, we used the average rate of 0.060 from Figure 2*c*. For zones without subway we used 0.069 (also from Figure 2*d*) as the average rate.

Taxi Trips

Method

Taxi trip rates were thought to be related to three factors that we had the ability to measure. These factors are subway accessibility, location of zone on a major taxi cruising corridor (e.g., a midtown avenue), and zone income. As with automobile trips, we tried to reduce the amount of unexplained variation by addressing the following factors:

Subway Accessibility The effect of subway accessibility was tested for various measures of accessibility in combination with various neighborhood groupings. The distribution of trip rates for uptown zones with and without subway access is shown in Figure 3. When trips are aggregated according to uptown and downtown, there appears to be a relationship between taxi trip rate and subway accessibility. Uptown zones with subway access have lower average taxi trip generation rates than uptown zones without subway access. Downtown zones, however, display the opposite pattern: zones with subway access have higher taxi trip generation rates than zones without. Although superficially counterintuitive, the pattern is most likely reflective of taxi cruising habits. South of 14th Street, the residential zones without subway access are either in lower-income areas or away from avenues with heavy taxi cruising. Hence, in these zones, lower demand for and lower availability of taxis combine to produce lower taxi use.

Neighborhood Groupings Variances in zonal trip generation rate were reduced by grouping zones into neighborhood aggregations. These aggregations, in conjunction with subway access, were a significant explanatory factor. The uptown-downtown dichotomy in taxi use is most likely attributable to income differences.

Results

A.M. Peak Hour (7:00–8:00 a.m.) Based on our findings, we used average rates for various conditions to project taxi trip generation. Four separate rates were used, as appropriate:

Route	*Trip/1,000 ft²*
Uptown with subway	0.039
Uptown without subway	0.063
Downtown with subway	0.034
Downtown without subway	0.017

P.M. Peak Hour (5:00–6:00 p.m.) Four separate rates were used, as appropriate:

Route	*Trip/1,000 ft²*
Uptown with subway	0.043
Uptown without subway	0.047
Downtown with subway	0.027
Downtown without subway	0.015

Truck Trips

Focus

When the O-D surveys were conducted, commercial vans were included as trucks and are therefore included in our analysis as well. The traffic survey counts indicate that the ratio of vans to trucks is approximately 0.8 to 1.0.

We did not anticipate a relationship between truck trips and parking or subway availability because trucks typically do not park in commercial lots or garages. Also, if costs of truck transportation rise, a business cannot shift to moving bulk goods by subway. Parking costs are also not relevant to trip generation because truck trips generally will be for the purpose of loading or unloading goods and will not require parking for extended periods in commercial lots. Our efforts were therefore focused on reducing the amount of variation

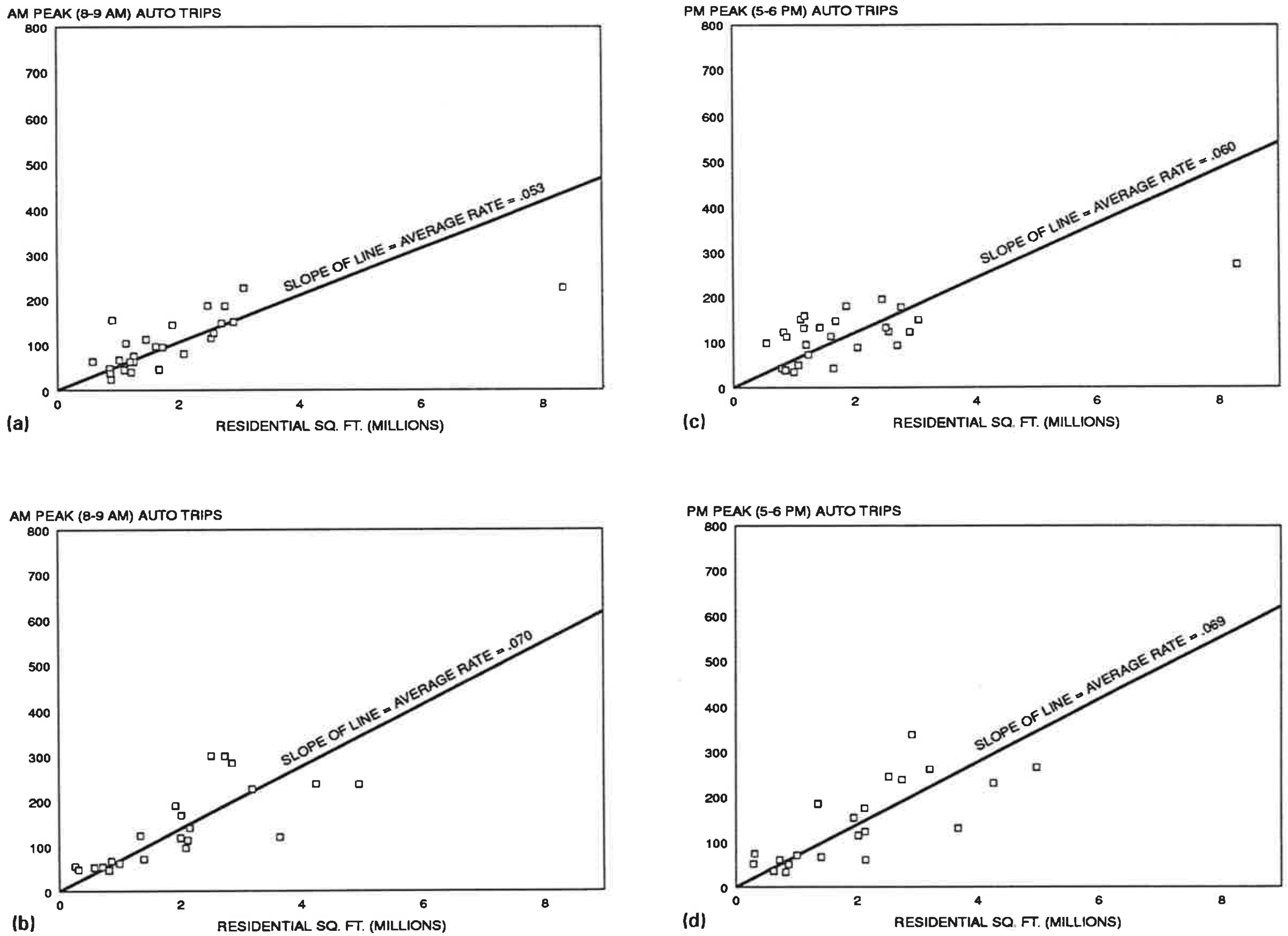


FIGURE 2 Comparison of residential zones with and without subway accessibility: (*a*) with subway, a.m. automobile trips; (*b*) without subway, a.m. automobile trips; (*c*) with subway, p.m. automobile trips; (*d*) without subway, p.m. automobile trips.

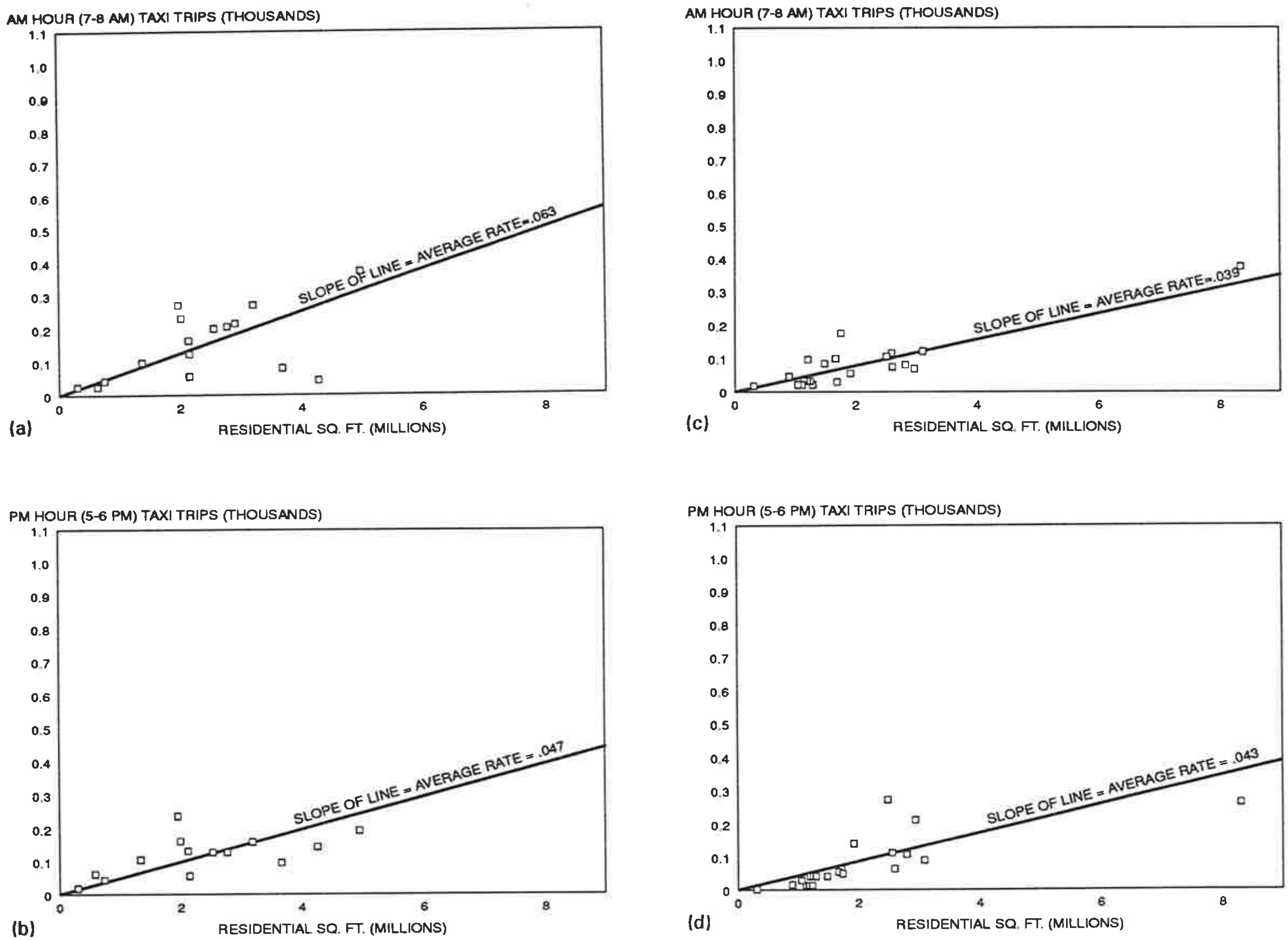


FIGURE 3 Distribution of trip rates for uptown residential zones with and without subway accessibility: (*a*) with subway, a.m. taxi; (*b*) without subway, a.m. taxi; (*c*) with subway, p.m. taxi; (*d*) without subway, p.m. taxi.

in trip generation rates by grouping into neighborhoods and by looking at 3-hr peak periods, as well as 1-hr peak periods.

Our findings indicate that when trips are aggregated at the zone level of detail, there is a large variability in trip rates between zones that may be reduced by aggregating into neighborhoods. Trip rates are also less variable when studied at the 3-hr instead of the 1-hr interval.

Results

Considering our findings, we used simple averages for trip generation (see Figure 4).

A.M. Peak Period We used a 3-hr trip rate of 0.046 trips per 1,000 ft^2 of development and converted it to a 1-hr rate

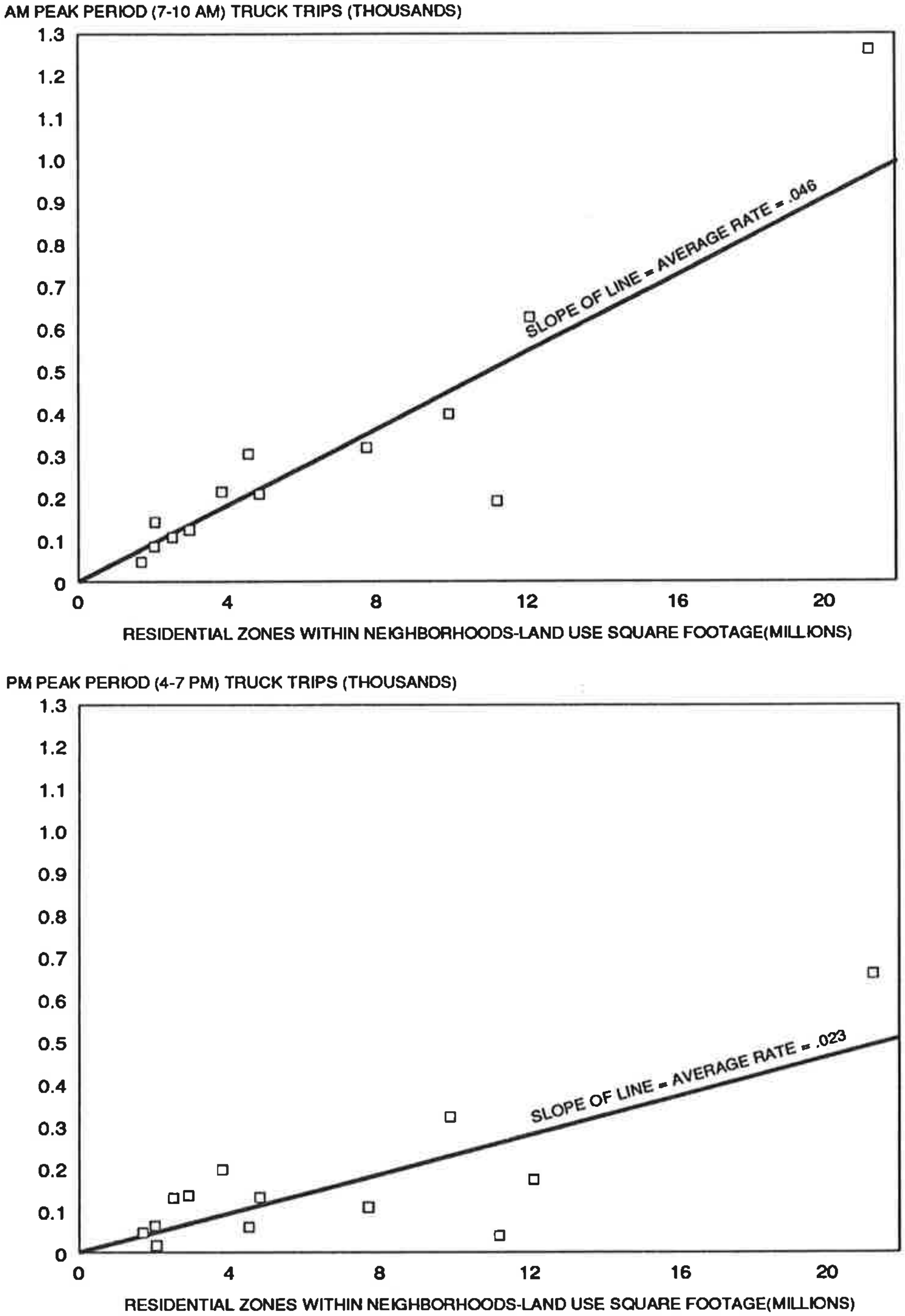


FIGURE 4 Comparison of a.m. (*top*) and p.m. (*bottom*) truck trips in residential neighborhoods.

by multiplying it by the factor 0.35 (total truck trips in peak hour per total 3-hr peak period trips).

P.M. Peak Period We used a 3-hr trip rate of 0.023 trips per 1,000 ft^2 of development and converted it to a 1-hr rate by multiplying it by the factor 0.31.

CALCULATION OF PURE OFFICE ZONE TRIP RATES

Automobile Trips

Focus

The relationship between automobile trips generated in each of the 35 zones defined as pure office and the square footage of development is shown in Figure 5. Zones containing City Hall and the World Trade Center were omitted because their trip generation patterns were different from most zones. Evening peak-hour automobile trip generation rates for pure office zones ranged from 0.013 to 0.104, whereas a.m. peak-hour rates ranged from 0.017 to 0.091 trips per 1,000 ft^2. Our program to reduce the amount of unexplained variation was based on the same premises as those for the residential zone analyses. The findings for pure office zones follow.

Parking Availability/Density When zones are aggregated at the neighborhood level, we find that most neighborhoods have the same parking density. As with residential zones, no relationship could be established at the zone level between automobile trips and parking. Parking availability was examined when defined both as the number of spaces per square foot in an area consisting of the individual zone and a single zone-width band around it and as an area consisting of an individual zone and a double zone-width band around it. Neither measure provided good results. A more detailed discussion is provided later.

Subway Accessibility Virtually all office zones are served by subway. When trips are aggregated at the neighborhood level, the relationship between subway accessibility and trip rates can be established. The relationship is stronger when subway accessibility is measured as the average number of stations within ¼ mi of a zone than when it is measured as the average number of different lines within the ¼-mi radius.

Neighborhood Groupings In general, variability of trip rates is reduced when trips are aggregated to the neighborhood level of detail.

Results

A.M. Peak Period For zones with subway access, the most accurate estimates were derived from the regression for the 3-hr peak period displayed in Figures 5*c* and 5*d*. The formula for this regression is

$$\text{trip generation rate} = 0.249 - 0.028X$$

where X is the number of subway stations within ¼ mi. For zones without subway access we used 0.24 as a "cap" on the trip rate. For zones with subway access, we used 0.083 (the right-hand intercept) as a minimum rate. The factor for deriving 1-hr trip rates was calculated as the ratio of automobiles in the peak hour (6,198) to total automobiles in the peak 3-hr period (19,460). The resulting conversion factor would be 0.32.

P.M. Peak Period For zones with subway access, we used the regression for the 3-hr peak period displayed in Figures 5*c* and 5*d*. The formula for this regression is

$$\text{trip generation rate} = 0.208 - 0.024X$$

For zones without subway access, at the left-hand side of the graph, we used 0.21 as a cap on the trip rate. This cap is an extrapolation from the regression line back to the vertical axis. For zones with subway access, we used 0.070 (the right-hand intercept) as a minimum rate. The p.m. residential automobile conversion factor (from the 3-hr to the 1-hr rates) would be 0.45.

Taxi Trips

Focus

Our program of actions to reduce observed variations in trip generation rates included grouping zonal data into neighborhoods to service as proxies for socioeconomic and locational factors as well as subway accessibility. We found that there appears to be a difference in trip generation rates for uptown and downtown zones. However, no distinction can be made between those areas with and without subway access, since virtually all pure office zones are in areas close to subways. As in the other analyses, grouping zones into neighborhoods reduced the amount of variance between the trip generation rates.

Results

As noted previously, offices are seldom located in zones not served by subways. Therefore, no with/without subway distinction is possible. Figure 6 shows the average rates for uptown and downtown taxi trips.

A.M. Peak Hour We used two separate rates as appropriate:

Zone	*Trip/1,000 ft^2*
Uptown	0.042
Downtown	0.019

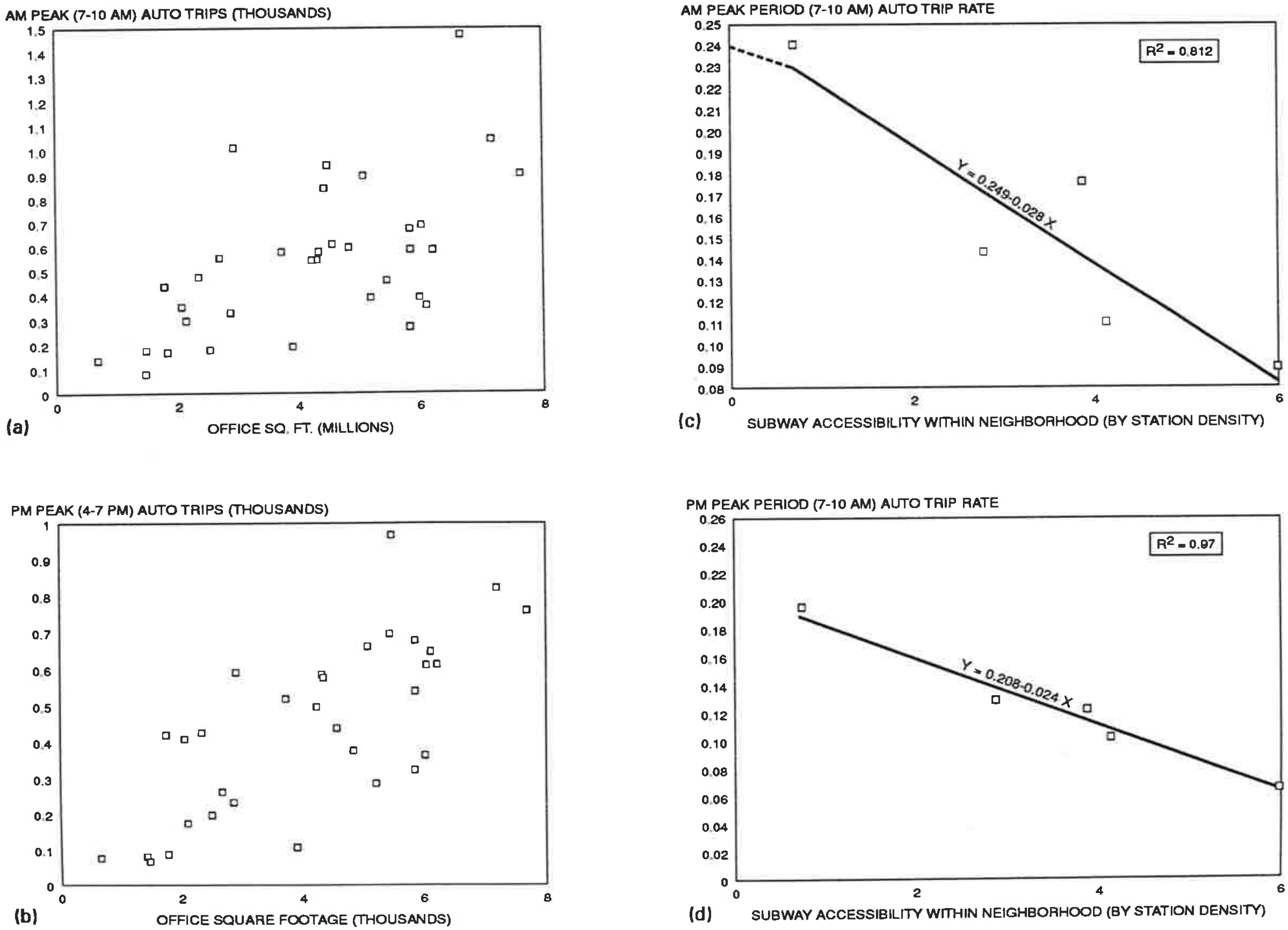


FIGURE 5 Relationship between automobile trips generated in each of 35 zones defined as pure office and square footage of development: (*a*) all office zones, a.m. automobile trips; (*b*) all office zones, p.m. automobile trips; (*c*) office neighborhood, a.m. automobile trips (r^2 = .812); (*d*) office neighborhood, p.m. automobile trips (r^2 = .97).

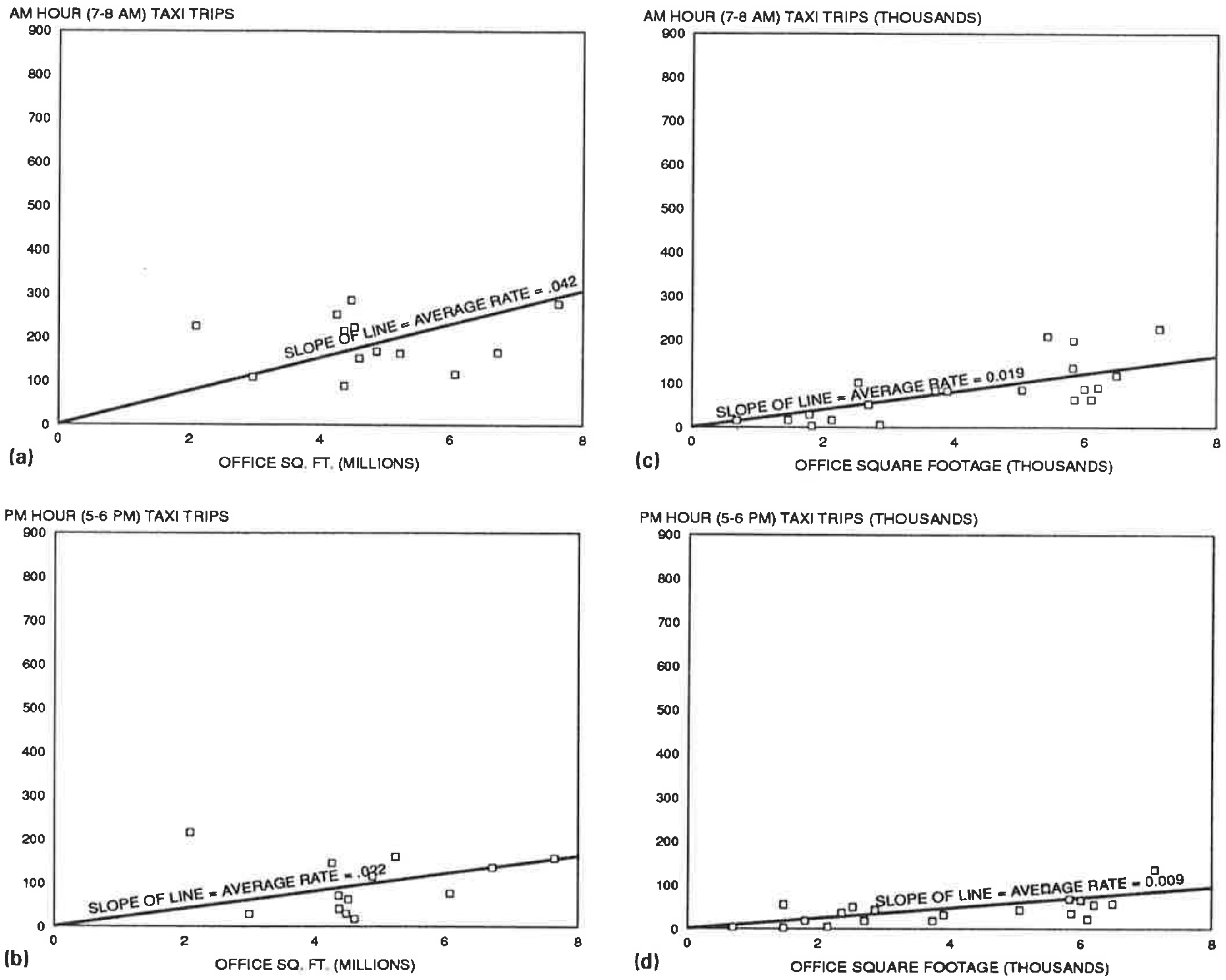


FIGURE 6 Average rates for taxi trips: (*a*) uptown, a.m.; (*b*) uptown, p.m.; (*c*) downtown, a.m.; (*d*) downtown, p.m.

P.M. Peak Hour We used two separate rates as appropriate:

Zone	*Trip/1,000 ft²*
Uptown	0.022
Downtown	0.009

Truck Trips

Discussion

Truck trip generation for office uses followed the same pattern as for residential land use. We did not, therefore, anticipate a relationship with parking or subway availability. Our analysis findings were also similar, to the extent that when trips are aggregated at the zone level of detail, a large variability in trip rates between zones is reduced by aggregation into neighborhoods and the 3-hr trip rates show less variability than the 1-hr rates.

Results

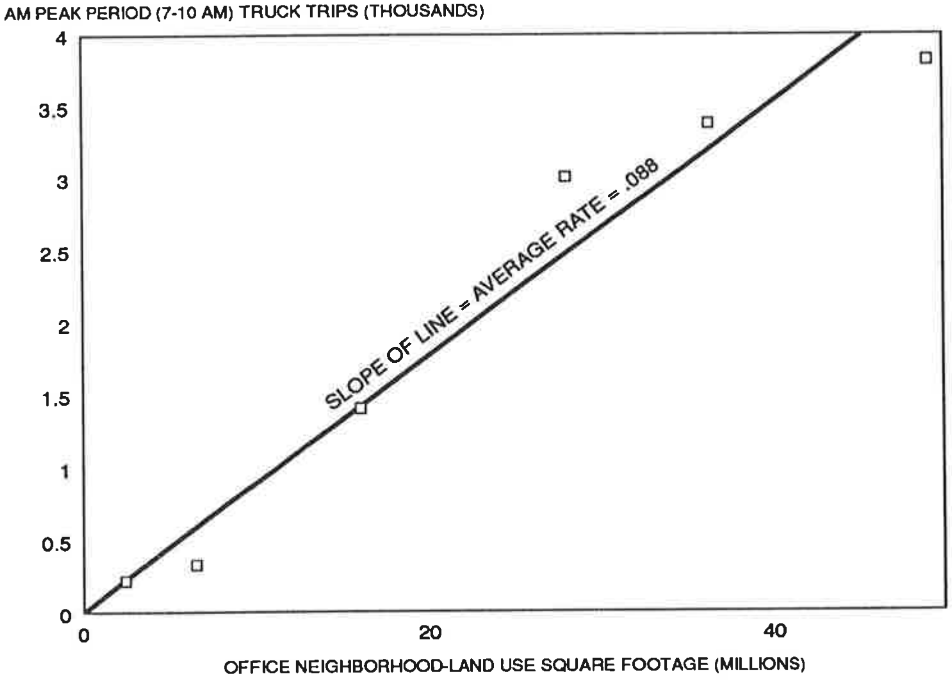
We used simple averages for trip generation (see Figure 7).

A.M. Peak Period We used a 3-hr trip rate of 0.088 trips per 1,000 ft² of development and converted it to a 1-hr rate by multiplying it times the factor 0.37 (total peak hour trips/total 3-hr peak period trips).

P.M. Peak Period We used a 3-hr trip rate of 0.034 trips per 1,000 ft² of development and converted it to a 1-hr rate by multiplying it times the factor 0.34.

PARKING DEMAND/SUPPLY CALCULATION

The research into automobile trip generation for residential and office uses yielded results that, in some cases, had run



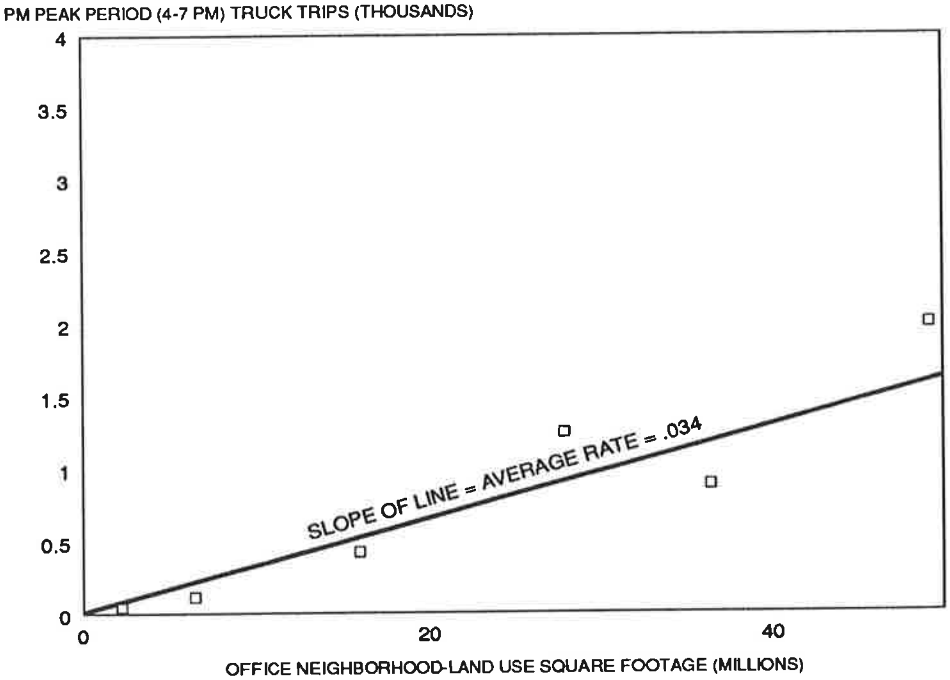


FIGURE 7 Comparison of a.m. (*top*) and p.m. (*bottom*) truck trips in office neighborhoods.

counter to our expectations. New York City has a policy of discouraging the construction of new public parking facilities in the Manhattan CBD. The intent is to limit the number of parking spaces in an area in which there is already perceived to be a shortage and thereby further discourage people from driving into the city. It was our expectation, and that of policy makers, that there would be a correlation between the supply of parking and an area's peak-hour vehicular trip generation rate. The analysis of pure residential- and office-use zones did not find a clear correlation between the density of parking spaces and trip generation rates.

We set out to investigate a possible explanation for the lack of correlation. We compared parking demand and supply to determine if there was an excess supply of spaces in most of the study area zones. If this were the case, then parking availability would not be a strong factor in a commuter's decision to drive or not.

We measured parking demand as the net of the trips into a zone minus the trips out during the morning peak period. Parking supply was simply the number of parking spaces in a study zone. We acknowledge that this gives only an imperfect view of the supply/demand balance because it does not

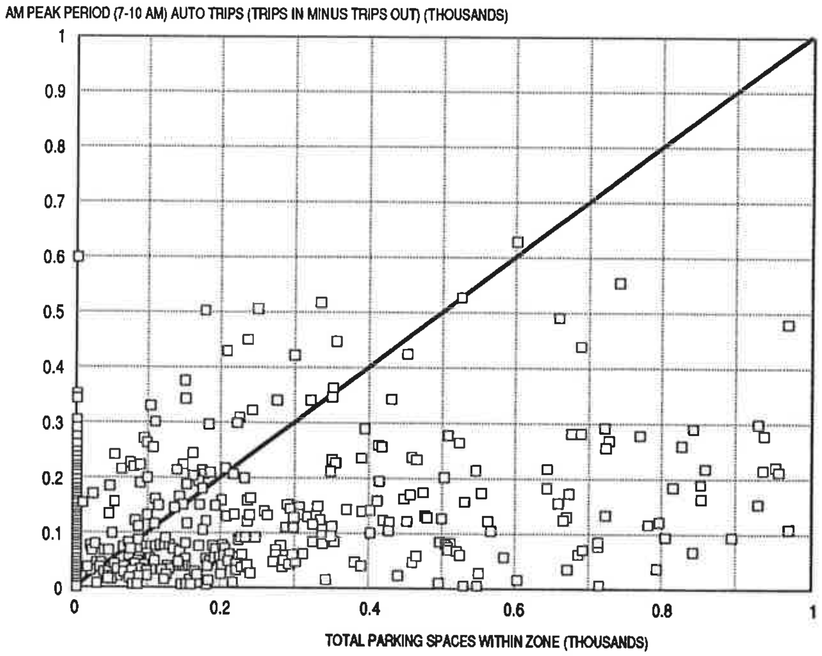


FIGURE 8 Automobile trips for all zones, a.m. peak.

TABLE 1 Final Trip Generation Rates for All Land Uses with Areawide Adjustments (trip ends per 1,000 ft² unless otherwise noted)

Land Use	Auto Trip Ends AM Peak Hour	Generation Rates PM Peak Hour	Taxi Trip Ends AM Peak Hour	Generation Rates PM Peak Hour	Truck Trip Ends AM Peak Hour	Generation Rates PM Peak Hour
Retail:						
Specialty	.0	.200	.0	.122	.03	.0
Dept. Store	.0	.078	.0	.051	.09	.0
Convenience	.081	.233	.163	.500	.02	.0
Hotel	.029	.022	.054	.051	.013/room	.006/room
Hospital	.135	.177	.056	.086	.004	.003
Museum	.0	1.532	.0	1.449	.0	.0
Health Club	.0	.133	.054	.122	.016	.009
Cinema	.0	.017/seat	.0	.026/seat	.0	.0
Theater	.0	.0	.0	.004	.0	.0
Schools	.046	.004	.043	.004	.007	.0
Colleges	.148	.249	.034	.063	.007	.0
Manufacturing/						
Automotive	.059	.021	.055	.019	.05	.05
Warehouse	.066	.173	.012	.065	.0	.0
Churches	.013	.071	.012	.065	.0	.0
Library	.113	.546	.121	.502	.007	.0
Community Ctr	.261	.317	.243	.292	0.015	0.007
Residential						
with Subway	.055	.058	0.038[b] 0.034[c]	0.039[b] 0.024[c]		
without Subway	.072	.067	0.062[b] 0.017[c]	0.042[b] 0.013[c]		
Office						
with Subway	0.081-0.009X[a]	0.096-0.011X[a]	0.045[b] 0.020[c]	0.018[b] 0.007[c]	0.029	0.010
without Subway	0.081	0.096	----	----	----	----

[a] Independent variable X = Number of subway stations within ¼ mile of a zone. Each separate subway is counted, although some stations may serve more than one line, while others may serve lines which can be accessed at multiple locations within the ¼ mile radius.

[b] North of 14th Street

[c] South of 14th Street

account for such factors as on-street spaces or long-term parking in garages. Figure 8 does, however, give a strong indication that in most zones the demand for parking is less than the supply of spaces (those zones to the right of the diagonal line). Hence most commuters, if they are willing to pay the cost of parking, would not be influenced in their decision to drive by concern over availability of spaces. Since the analysis was limited to a 3-hr peak period in the morning, it does not have implications for noncommutation trips. Available parking supply continues to diminish as the midday approaches, and drivers arriving later may have much more trouble finding parking spaces.

TRIP GENERATION RATES FOR MISCELLANEOUS LAND USES

Miscellaneous land uses that are present in the study area and that are called out in the land use data base are divided into 16 categories, including, for example, department stores, hotels, schools, churches, hospitals, theaters, and factories.

We obtained peak-hour vehicular trip generation rates for these land uses by two methods. First, where possible, we calculated trip generation rates that would typically be used in NYC environmental impact assessment work. This was accomplished by assembling common vehicle occupancy rates, modal splits, temporal distributions, and person trip generation rates. Second, for more unusual land uses, we resorted to trip generation rates assembled by ITE, with modifications we thought were appropriate for New York City. The ITE trip generation rates are derived from studies of communities that are smaller and generally less congested than New York. All miscellaneous land use trip generation rates and their sources are summarized in Table 1.

TRIP GENERATION RATES FOR SPECIAL LAND USE ZONES

Special use zones are those zones that are so unique that we thought they should be handled individually. Either the use in the zone did not fall into one of the broad categories for which trip generation rates were available (such as courthouses), or the use was large enough or unique in some other way that would preclude the use of standard rates. Trips resulting from growth in the unique uses of these zones would be estimated by using the trip rate of that particular zone.

ADJUSTMENTS TO PURE USE ZONE TRIP GENERATION RATES

It was necessary to adjust the trip generation rates calculated for the pure use zones downward to account for the 5 percent, on average, of land use square footage that was categorized as miscellaneous. Using the trip generation rates compiled for miscellaneous land uses and the derived residential and office trip rates, a new trip total was calculated for all land uses. This new trip total was divided into the trip total for the office and land uses alone to obtain an adjustment factor. The office/residential trip generation rates were then multiplied by the factor. Table 2 shows the results of the adjustment.

TABLE 2 Comparison of Original and Adjusted Trip Generation Rates

	Original Trip Generation Rate		Adjusted Trip Generation Rate	
	AM	PM	AM	PM
Residential Auto				
W/Subway	.053	.060	.047	.052
W/out Subway	.070	.069	.062	.060
Residential Taxi				
Uptown W/Subway	.039	.043	.035	.038
Dwntwn W/Subway	.063	.047	.057	.042
Uptown W/out	.034	.027	.031	.024
Downtown W/out	.017	.015	.015	.013
Residential Truck	.016	.007	.015	.007
Office Auto				
W/Subway	.080-.009X	.094-.011X	.070-.008X	.086-.010X
W/out Subway	.080	.095	.070	.086
Office Taxi				
Uptown	.042	.022	.042	.018
Downtown	.019	.009	.019	.007
Office Truck	.033	.012	.029	.009

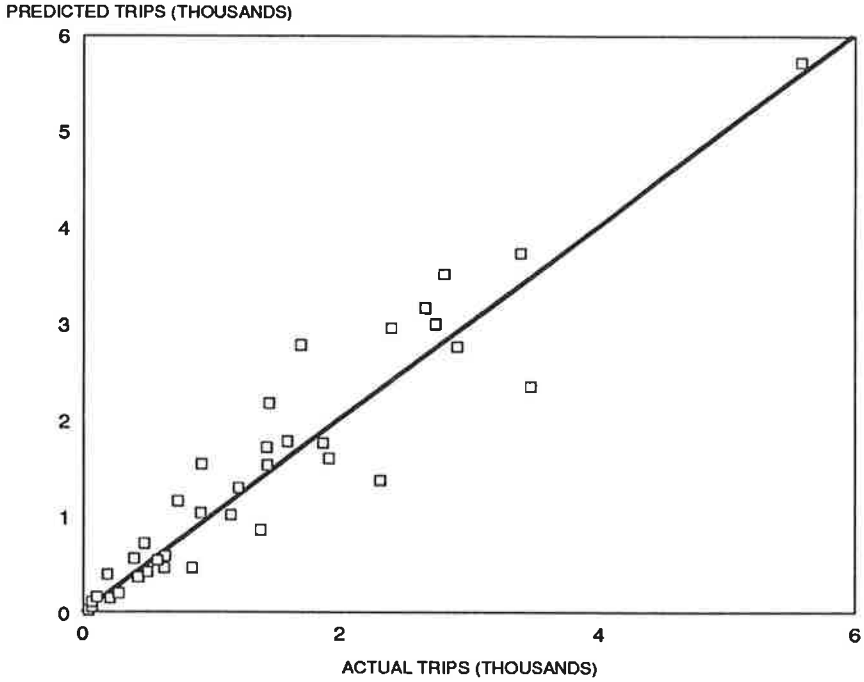
VALIDATION/CALIBRATION RESULTS

Using the process described in the methodology section, trip generation was projected for each zone in the study area. The ratios of actual trips to predicted trips when totaled across all zones for automobiles, taxis, and trucks were then calculated separately. The ratios determined were as follows:

	Ratio of actual to predicted trips	
	a.m.	*p.m.*
Automobiles	1.16	1.11
Taxis	1.09	1.02
Trucks	.92	1.08

These ratios were in turn used as a final adjustment factor for trip generation rates of all land uses. The final trip generation rates are given in Table 1.

Actual trips, by zone, were compared against predicted trips (using the rates in Table 1) at the zone level and at the neighborhood level. Figure 9 shows predicted versus actual trips for each zone. The diagonal line indicates where predicted and actual trips are equal. The figure demonstrates that the predicted trip generation rates are relatively good approximations of actual trip generation at the aggregate level.



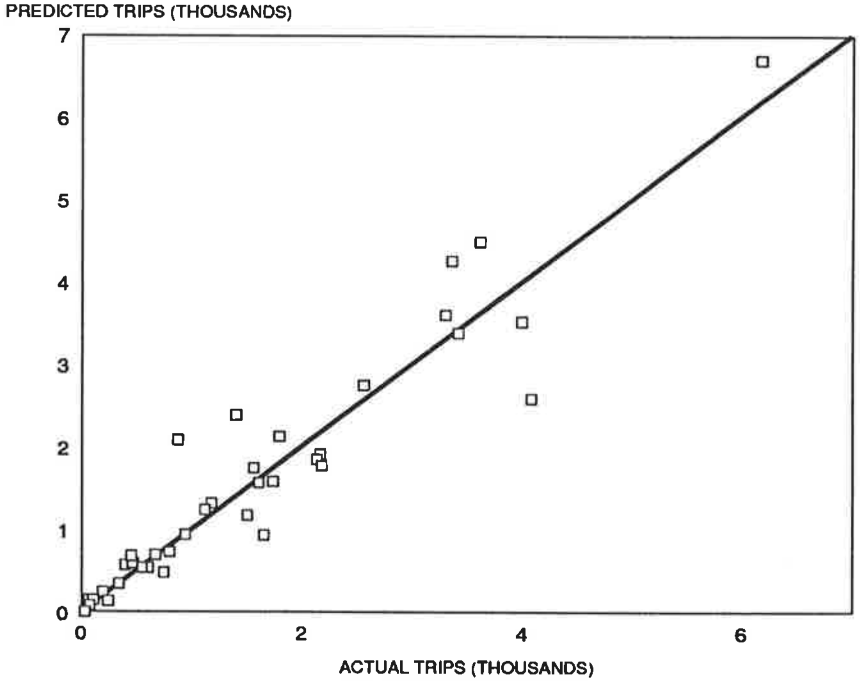


FIGURE 9 Comparison of actual and predicted trips for each zone: a.m. (*top*); p.m. (*bottom*).

CONCLUSIONS

The Route 9A project for the replacement of Manhattan's West Side Highway necessitated development of an alternative methodology for the calculation of trip generation rates. For the 9A project, it was not as important for trips from each new building or land use to be projected accurately as it was for the overall estimate of future trips to be as true as possible. The approach taken, however, provides greater accuracy than total reliance on secondary sources. The *ITE Trip Generation Handbook* provides little guidance on which rates to use when transit is a viable alternative to most vehicular trips. The handbook, for example, states that for the a.m. peak hour, office trip generation rates range from 0.03 to 0.56 trips per 1,000 ft^2 (depending on the modal split assumptions) and provides no rule for which rate to apply and where. For the Manhattan CBD, the expected automobile trip generation rates range from 0.027 to 0.081 trips per 1,000 ft^2 and are explicitly related to transit services. For residential apartments, the ITE handbook provides a.m. peak hour rates that typically range from 0.06 to 1.64 trips per 1,000 ft^2 (again depending on the modal split). In contrast, the Manhattan

CBD rates range from 0.055 to 0.076 trips per 1,000 ft^2 and are explicitly related to transit service.

The approach is one that (a) focuses on the dominant land uses, thereby avoiding wasted effort to estimate more accurate trip generation rates for land uses that will only slightly affect the margin of error; (b) allows the use of available data without extension field work beyond the establishment of the trip table; and (c) through the use of small zones, allows for subarea summaries that serve as proxies for income, transportation access, and other characteristics of zones.

ACKNOWLEDGMENTS

This study was one in a series sponsored by NYSDOT as part of the Route 9A replacement study. The authors wish to thank the NYC departments of Transportation and City Planning and NYSDOT for their help on this project and, in particular, the members of the technical review committee: Gerard J. Cioffi and Lee Maynus of NYSDOT, Michael Graham of NYC Department of Transportation, Dick Chudd of the NYC Department of City Planning, and Daniel Greenbaum and Jerry Nielsten of Vollmer Associates. Their input and commentary were invaluable in the analysis of results and in development of the trip generation methodology.

Publication of this paper sponsored by Committee on Transportation and Land Development.

Geographic Information System–Based Transportation Program Management System for County Transportation Agency

MICHAEL D. MEYER AND WAYNE SARASUA

A prototype geographic information system for transportation (GIS-T) that was developed for a suburban county transportation agency in metropolitan Atlanta is described. The main objective in developing the prototype was to design a system to help the county agency to better manage its transportation program. An organizational assessment was conducted to determine the different types of GIS-T applications that would most benefit the agency. These applications were ranked on the basis of greatest need, and five application modules were chosen for inclusion in the prototype to address these needs. The modules were: (a) an integrated accident record system, (b) traffic engineering, (c) pavement management, (d) transportation planning and land use, and (e) transit. The key finding of this project is that GIS-T provides a strong management decision support capability for officials in this agency. In particular, the prototype provides more than just a data management capability. It also provides a means by which various types of analysis can be performed more efficiently than they were in the past.

Transportation agencies are facing increasing pressures to manage transportation programs with fewer resources. For some years, many agency managers have examined the potential role that information systems technology could play in providing more productive program management. Such management information systems have become an important element of today's management process, although the level of importance varies from one agency to the next, depending on the type and magnitude of implementation. Geographic information systems (GIS) provide an important enhancement of these information systems, especially for those organizations, such as transportation agencies, that deal extensively with spatially defined services and facilities.

This paper describes a prototype GIS for transportation (GIS-T) that was developed for a suburban county transportation agency in metropolitan Atlanta, Georgia. In particular, this prototype system was designed to help the agency better manage its transportation program; a concerted effort was made to better understand the types of data currently collected in the agency, the manner of collection, and possible improvements in the use and handling of these data.

This project began with a literature search that examined the application of information system technologies to public works and transportation agencies similar in structure and function to the county transportation agency. Individual meetings held with agency personnel helped to establish the types of data and uses of information that existed within the agency. The responsible official for each unit within the agency was interviewed to determine current uses of data, desired types of data that should be collected, and possible software applications that would make the agency more efficient and effective. Where possible, standardized forms that are currently used to collect or input data, or both, were obtained.

ORGANIZATIONAL ASSESSMENT

The county transportation agency consists of six divisions. The Traffic Engineering Division is responsible for safe and efficient traffic operations on county streets and highways through the design, installation, and maintenance of traffic control devices. The division is divided into transportation planning, field operations, and traffic engineering. One of the major tasks of this division is to analyze development proposals for their potential effect on the traffic operations of county roads. The Engineering Division is responsible for engineering design, design contract management, acquisition of right-of-way, construction management, and inspection. This division is organized into engineering services, right-of-way, and construction management. The Roads Division is responsible for maintaining 1,800 mi of county roads and 193 bridges under county jurisdiction. This division is organized into administration, construction, and maintenance sections. The division is expected to maintain accurate data on the costs of maintaining and repairing the county's roads and bridges. The Administration and Control Division is divided into general administration, project coordination, and archaeology. This division is responsible for the overall financial administration of the agency and for keeping track of payment requests; these are activities that lend themselves quite well to some form of automated information systems. The Transit Division is responsible for the short- and long-range planning of a suburban transit service as well as rideshare and vanpool operations. The Aviation Division is responsible for the operation of a small county airport.

The organizational assessment examined the current data management activities that were considered most important to the operation of the agency. The most relevant type of data were to be included as part of the prototype program management GIS. The following sections describe how data were collected and managed within the various divisions. A discussion is included on how data for the prototype GIS were obtained.

School of Civil Engineering, Georgia Institute of Technology, Atlanta, Ga. 30332.

Traffic Engineering Division

The Traffic Engineering Division had two major types of information management activities that were important for this project: accident records and traffic signal management. For its accident records system, the Traffic Engineering Division uses a computer program for the input, management, and analysis of accident information. The data base was constantly updated using information provided from police accident reports. The system could perform statistical analysis and could produce a variety of hard copy reports. The system could also produce simple collision diagrams. Although the system adequately met the needs of the division from an efficiency standpoint, the system had a number of limitations. First, the display capabilities were limited. Diagrams produced by the system were limited to individual intersections. There was no capability to produce large-scale plots of collision diagrams at consecutive intersections on arterials. Second, only rudimentary analysis capabilities were possible with the system. Although it was possible to use the accident records system to list intersections of high accident incidence, there was no facility for producing a map that displayed various categories of accident incidence. Furthermore, there were no graphic capabilities that would make it possible, for example, to automatically create a pie chart for a particular intersection that displayed the percentage of accidents under various weather conditions.

The Traffic Engineering Division is also responsible for maintaining a number of traffic signals throughout the county. For each of the signalized intersections, the division keeps records for a variety of data, including hardware information, signal timing plans, intersection geometrics, and count data. However, not all of the signalized intersections in the county were under the jurisdiction of the county transportation agency. Some of the intersections fell under the jurisdiction of various cities.

The first type of data collected by the county to be included in the prototype GIS was the geometrics of the various intersections. For county intersections, these data were obtained in part through examination of bluelines, plans, and freehand sketches. Many of the signal plans were not available for some intersections. The geometrics for these intersections were developed from field visits.

The second type of data collected was the timing plan for each of the signalized intersections in the target area. Most of the intersection timings were available in digital form because these intersections were part of one of the 18 closed-loop systems connected to a central computer located in the signal shop. Timings for these intersections were quickly downloaded to diskette or to hard copy output. For those intersections not on the central system, timing plans were usually available on handwritten forms.

The third type of intersection data collected related to the various hardware at each of the signalized intersections. This included such items as controller type, inductive loops, and sensors. Many of the signal plans were either incomplete or out of date. Maintenance information was obtained for almost all of the intersections. The format of this information varied greatly between the different jurisdictions involved. One city had extensive maintenance data available on computer. Maintenance data provided by the county were in hard copy form and were limited. Additionally, traffic characteristics such as traffic count data and intersection levels of service were obtained from recent studies performed by a private consultant.

Engineering Division

In the Engineering Division, all data summaries and form handling were done manually. The most important data included the status and content of engineering contracts, right-of-way appraisals, information on right-of-way negotiations, and status of construction contracts. These documents were organized by project number and by funding category. If someone wanted to obtain information about a particular parcel of land, the division staff had to go to a file and manually retrieve the information. The specific type of information in these project files is extremely important for division managers and for the agency director to have retrieved in a timely manner. Right-of-way information was divided in parcel level information and included property costs, associated legal fees, appraisal fees and negotiation fees. Construction contract information included information on specific items, such as aggregate cost, and included a payment schedule. Engineering design contract information included similar types of data.

Roads Division

The only pavement distress information stored on-line was an aggregate score of pavement condition. All of the components that defined the aggregate score (i.e., alligator cracking, transverse cracking, oxidation, etc.) were kept in hard copy form only. Additionally, the existing computer data base had a lot of missing or incomplete information for many of the streets in the study area. Many streets in some parts of the study area were not in the data base and thus did not even have an aggregate distress rating. Another drawback of the data base, and the system in general, was a lack of segmentation for a number of streets. Streets 1 to 2 mi long or greater were listed in the data base as one continuous segment. Field investigations of several roads determined that the pavement surface varied substantially over the entire length of the segment.

Administration and Control Division

The Administration and Control Division included two data areas of interest: archaeology and project control. The county has an extensive inventory of historical artifacts. This inventory is augmented with site investigation and project management information to allow response to requests for archaeological and historical evaluation of projects. Different types of data exist in various formats, including hard copy and digital data bases. Some data, such as the location and characteristics of sensitive burial grounds, are classified. The archaeology section also reviews planning and zoning requests and all road projects. These reviews were done using a word processor. The projects control section mainly processes invoices for projects. This section recorded project financial information on forms that were filled out manually.

There appeared to be a great deal of similarity in the types of project information needed by this section and the Engineering Division. A project management system that monitored not only construction progress but also financial status appeared likely to be an important capability for both divisions.

Transit Division

The operating data for the transit system are collected by a private transit management firm. The data collection effort was very weak in maintenance costs and in socioeconomic and land use density data that would be necessary for the planning of any service expansions. Financial information came from the county finance department, but there was some concern that the information was not provided on a timely basis or that it was not categorized in a way that would be most useful for day-to-day management.

In summary, data acquisition techniques appeared for the most part to be manual and labor-intensive in nature. Standardized paper forms were completed by various staff members and circulated within the agency. The data collection appeared to be oriented toward business or accounting and did not include engineering data. Items such as maintenance records for existing facilities (roads, bridges, and traffic signals) were not computerized. Even within the financial section, various kinds of data or aggregation of existing data seemed absent. It was this status of data collection that led to the development of a prototype GIS that would help agency officials to better manage the transportation program.

PROTOTYPE GIS-T

A prototype GIS-T was developed for the county transportation agency using the TransCAD software. TransCAD was chosen for a number of reasons. First, TransCAD is designed for use on personnel computers and can be installed and tested on existing computers already available throughout the transportation department. Second, TransCAD has a number of specialized tools already in the software that are designed to address transportation-related problems. These tools include shortest-path algorithms, travel forecasting capabilities, and several tools to solve operations research problems related to transportation. Functional applications can be developed in a relatively short time. Furthermore, TransCAD is relatively easy to learn.

The roadway basemap for the prototype was developed by converting U.S. Census Bureau TIGER files into a TransCAD line data base for the entire county. A subset of this data base was used for the prototype. Because this prototype was intended to be used only to illustrate GIS capabilities, the line data base that represented the street system was not checked for accuracy. If the system were to be implemented for the entire county, one of the first major tasks would be to correct errors in the base map.

Using this line data base, it was possible to create additional data bases for signalized and unsignalized intersections, transit routes, bus stops, and accident locations. The only locational data that had to be digitized or entered manually were existing traffic analysis zone (TAZ) boundaries and future development sites.

The prototype GIS-T is divided into five application modules: (a) an integrated accident record system; (b) traffic engineering; (c) pavement management; (d) transportation planning and land use; and (e) transit.

Integrated Accident Record System

The prototype GIS-T has the capability to store, retrieve, and analyze accident data. The design of the digital data base is patterned after the existing agency accident record system. The various attributes for a particular accident record in the prototype GIS-T are listed in Table 1. The GIS-T data base was built by translating the existing accident system dBase files for 1990 into Lotus 1-2-3. TransCAD can import data directly from Lotus 1-2-3. All accident record data stored in the prototype are associated with an intersection. Accidents are described as being at an intersection or some distance away from the intersection along a particular approach.

The analysis and display capabilities of the accident module are extensive. A number of descriptive maps identifying, for example, those intersections with unusually high incidences of left-turn–related accidents can be generated using the prototype GIS. Another possibility is using the charting capability of the TransCAD to produce a pie chart that shows, say, the percentage of accidents that occurred at each intersection in the study area during 1990. Figure 1 shows a bar chart, also produced by TransCAD, indicating the various types of accidents that occurred in the study area in 1990.

Spatial queries and spatial analysis are capabilities unique to GIS technology. Examples of spatial queries that were tested in the prototype GIS-T included listing all accidents that occurred within a given distance from an intersection and producing a map showing the incidence of midblock left-turn accidents on streets with a two-way left-turn lane under the county agency's jurisdiction. Using spatial analysis techniques, it may be possible to identify trends in accidents. By cross-referencing accident incidence with other types of information, such as roadway inventory, a correlation may be found between high rates of accident occurrence with, for example, poorly maintained roadway pavement markings or missing signs.

Figure 2 illustrates a proposed integrated accident record system for the county. It proposes that the system be made up of three components, the first of which is the GIS-T. The accident module of the prototype GIS-T was designed to complement the county's current accident record system (called SCARS) rather than possibly replace it. The SCARS system represents the second component of the integrated accident record system. The purpose of SCARS would be to perform various statistical analyses on the accident data. Eventually, the analysis capabilities of SCARS could be programmed into the GIS-T, eliminating the need for the SCARS component.

The final component of the system is AutoCAD. Its purpose would be to produce detailed collision diagrams. Data from the SCARS or the GIS component would pass directly to AutoCAD, which would then generate the collision diagram.

TABLE 1 Accident GIS-T Attribute Fields

TRANSCAD FIELD	DATA TYPE	COLUMN WIDTH
Intersection ID	Number	12
Accident Date	Character	10
Time	Number	5
Vehicle,Injured, Killed	Number	3
Accident Number	Number	12
Route	Number	5
Direction-Distance	Character	5
County Node	Number	6
Site	Coded	3
Vehicle 1 Type	Coded	25
Vehicle 2 Type	Coded	25
Direction 1	Character	2
Direction 2	Character	2
Damage 1	Number	5
Damage 2	Number	5
Damage 3	Number	5
Vehicle-Pedestrian 1	Character	1
Vehicle-Pedestrian 2	Character	1
Description	Number	2
Movement 1	Number	2
Movement 2	Number	2
Location 1	Number	2
Location 2	Number	2
Event	Coded	25
Control	Coded	25
Lighting Condition	Coded	25
Surface	Coded	25
Cause 1	Character	3
Cause 2	Character	3
Road ID	Coded	25
Violation 1	Number	4
Violation 2	Number	4
North-South Street	Character	20
East-West Street	Character	20

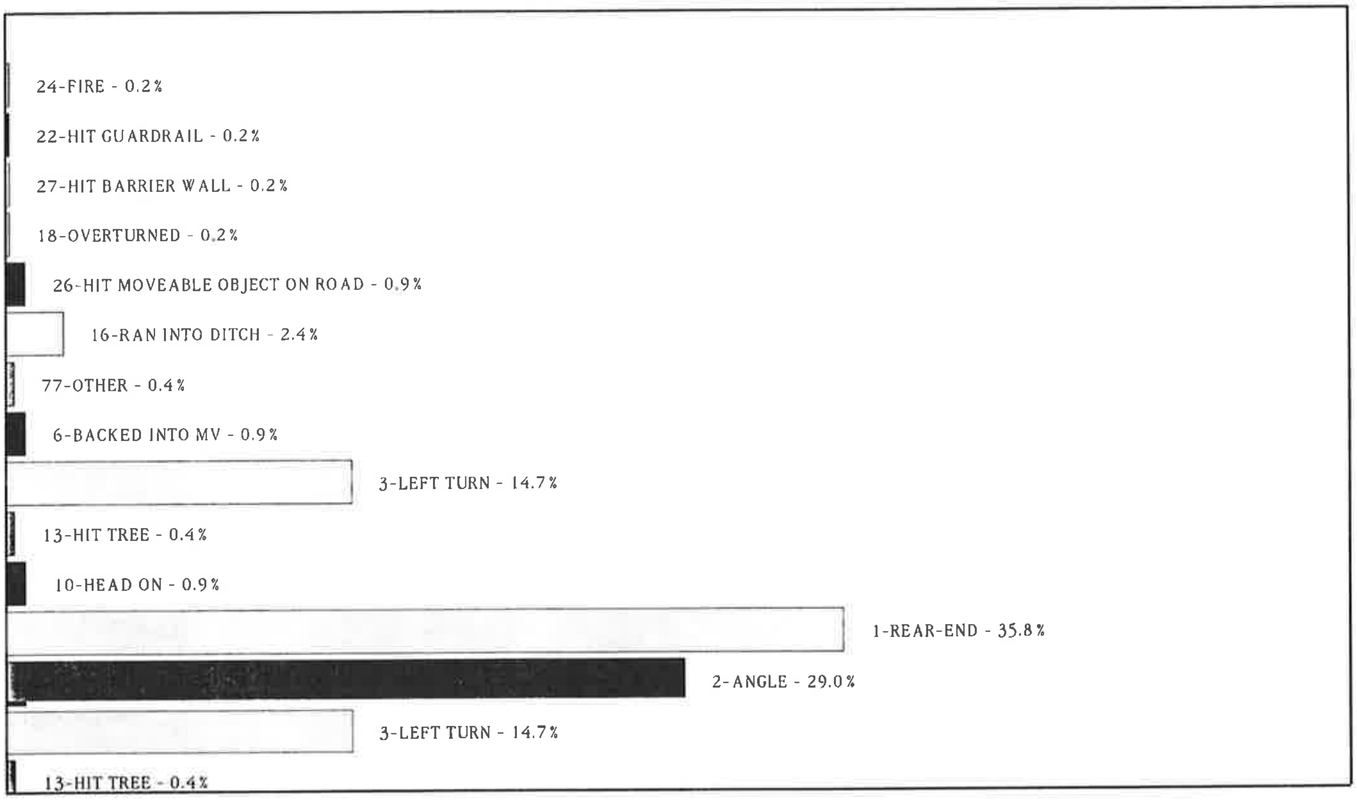


FIGURE 1 Accident records bar chart from GIS-T, all data (455 observations).

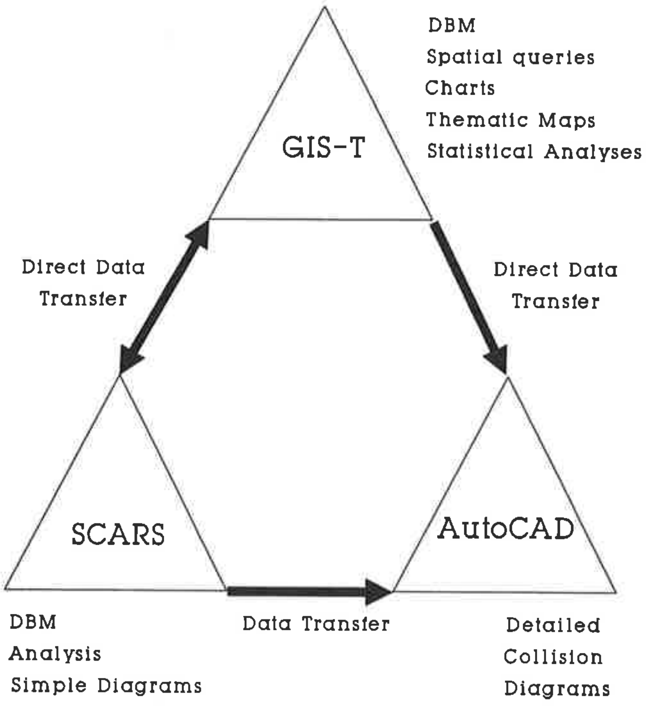


FIGURE 2 **Integrated accident record system.**

Traffic Engineering

Traffic Signal System Management

Specific information related to signalized intersections was incorporated into the prototype GIS-T. The intent of the signal module was to help automate the process of managing and maintaining the traffic signal system. The system provides a number of capabilities:

1. Provides quick and easy access to signalized intersection information.
2. Allows for swift identification of "troubled" intersections.
3. Keeps maintenance activities up to date to help spot problems before they occur.

The signal module of the prototype GIS-T is able to provide quick and easy access to signalized intersection information such as controller type, signal heads, detectors, sampling loops, intersection geometrics, and travel characteristics. Queries can be made simply by selecting an intersection from the map display. The associated text information that results from the query can be reviewed, printed, edited, or any combination thereof.

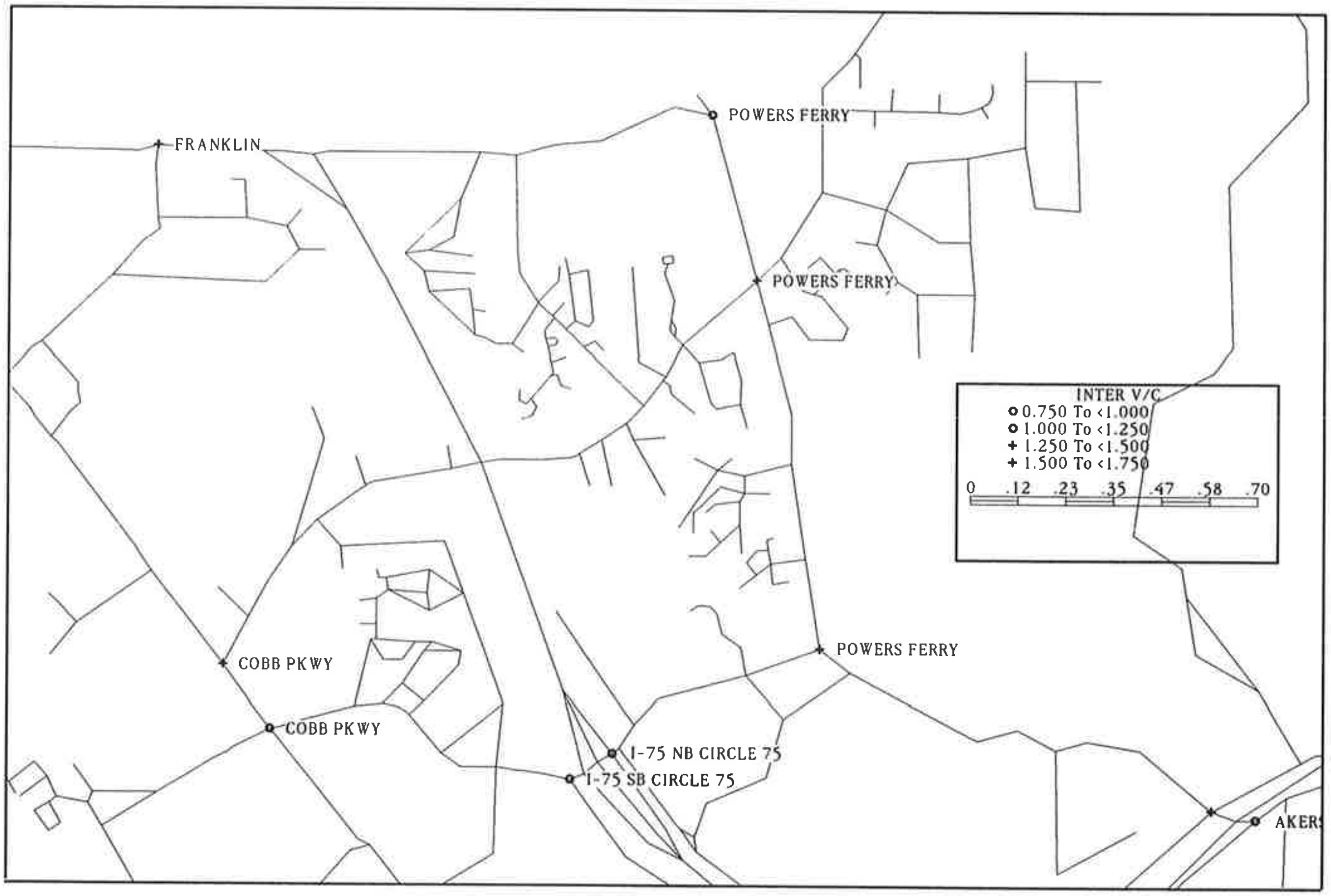


FIGURE 3 **Thematic map showing *v/c* ratios.**

The prototype GIS-T allows for swift identification of troubled intersections. These troubled intersections may be defined as those with large volume-to-capacity (*v/c*) ratios, enormous delays, unacceptable levels of service, intersections with a high number of maintenance activities in a relatively short period of time, or any combination of these variables. These troubled intersections can be displayed on high-quality thematic maps. Thematic maps are an intricate part of any GIS. A theme is a classification of entities in a layer based on the value of any one attribute. When users create a theme they also choose any or all of the following: colors, styles, or icons that are used to distinguish the classification on the map display or plotted output. An example of a thematic map produced by the prototype GIS-T that displays *v/c* ratios is shown in Figure 3.

If the user prefers maps with labels (which is normally the case when colored output is not possible), this is easily accomplished. Figure 4 shows a map of levels of service for the signalized intersections in the study area.

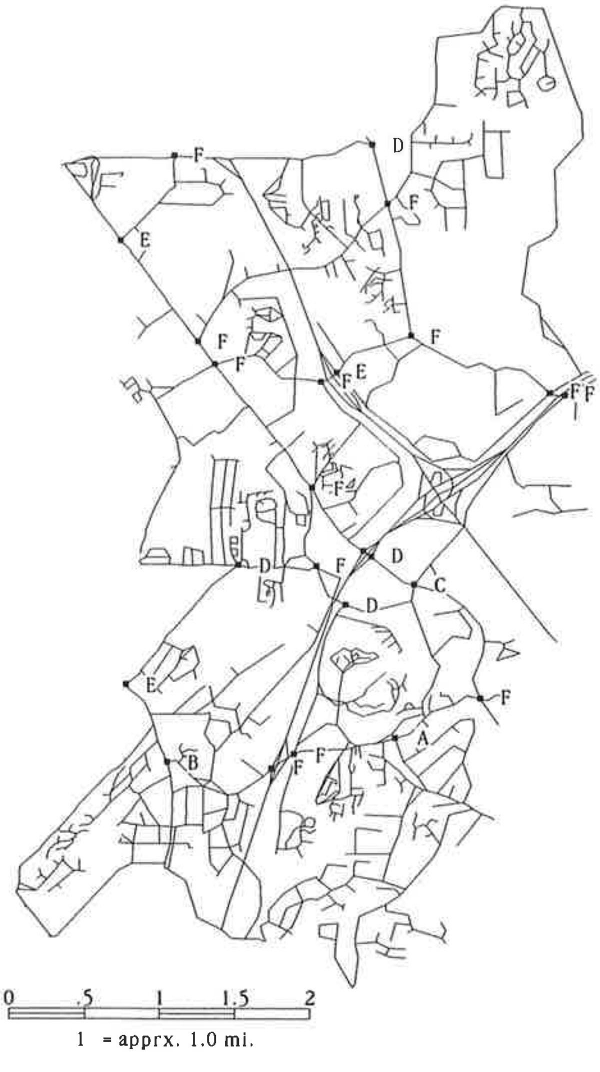


FIGURE 4 Thematic map showing levels of service.

Maintenance Management

The GIS-T prototype also keeps maintenance activities up to date, which can help spot problems before they occur. Included in the prototype is the type of last maintenance activity, date of last maintenance activity, and the number of maintenance activities in the last year. Using this information, it is possible to produce, for example, a map that identifies all intersections with more than four maintenance activities a year. A partial listing of information included in the signal prototype is as follows:

- *I.D. Number*—A three digit identification number that comes from TIGER files.
- *N-S Street, E-W Street*—name of the cross streets at the intersection.
- *Controller Type*—A description of the controller type at the intersection.
- *Plans Approved, Plans Revised*—Date (MMDDYY) that the signal plans were approved, revised, or both.
- *Signal Heads*—Character description of all types of signal heads from all directions at intersection.
- *Loops*—Physical description of loops actually located in the field.
- *Detectors*—Type or mode of loops at that intersection. Types are pulse, presence, pedestrian, and system.
- *Sampling Loops*—Number and location of loops used to sample traffic for occupancy and volume.
- *(NB-SB-EB-WB) Lanes*—Description of the approach lanes in each direction. Possible types are L-T-R-LT-TR-LTR-LR.
- *Study Date*—Date that the "field study" was conducted to gather filed data.
- *(NB-SB-EB-WB) (L-T-R) v/c*—Volume-capacity ratio for each approach lane in each direction. For shared lanes (i.e., LT or TR), the through (T) lane was used.
- *Intersection v/c*—Volume-capacity ratio for entire intersection.
- *(NB-SB-EB-WB) Delay*—Delay, measured in seconds.
- *Intersection Delay*—Delay, measured in seconds.
- *Intersection Lost Time*
- *(NB-SB-EB-WB) LOS*—Level of service for each directional approach based on directional approach delay. Levels of service are defined as follows:

LOS	*Delay (sec)*
A	0–5
B	5–15
C	15–25
D	25–40
E	40–60
F	>60

- *Intersection LOS*—Level-of-service for each intersection based on levels of service for all approaches.
- *(NB-SB-EB-WB) ADT*—Represents the 24-hr count of traffic for each directional approach on the specified study date.
- *Last Maintenance Activity*—A record of the last maintenance problem that occurred at the individual intersection. The possibilities are as follows:

Code	*Choice*
1	Bulb out
2	On flash
3	Defective loop
4	Replace controller
5	Signal head/pole hit
6	Rearrange/install signal heads
7	Preventive maintenance
0	Other

- *Date of Activity*—MMDDYY of the last maintenance activity performed at the individual intersection.
- *Number of Activities in Year*—Number of maintenance activities in the last year.

A potentially powerful tool that was developed as part of the prototype GIS-T is automating the process of signal coordination. Much of the information used by traffic signal coordination packages (e.g., TRANSYT-7F, PASSER) can be stored in a GIS. Software was developed that automatically accesses the GIS data base and generates input files for use in TRANSYT-7F (*1*,*2*). Implementation of this capability into an expanded countywide system will help make the process of coordinating signals more economical.

A desirable capability that is not possible with the current version of TransCAD is to be able to visually manage information at the intersection level. Although it is possible to access all of the attribute information for an intersection by scrolling across fields, this can be cumbersome. A better way would be to use a GIS intersection manager. This system would allow the user to visually see a wide variety of attribute information for an intersection on a single screen. For example, geometrics, turning movement volumes, signal head locations, and loop detector information could be included on a single graphical display.

Pavement Management

One transportation application of GIS that is getting a great deal of attention in public agencies is pavement management (*3*,*4*). Unfortunately for the demonstration project, pavement data were limited for streets located in the study area. Therefore, to better illustrate the capabilities of a pavement management GIS, imaginary data were added for those roads that lacked pavement data. Additionally, a number of fields were included in the data base even though data for these fields were unavailable. A list of the fields included in the pavement module are shown in Table 2.

One of the initial steps in the pavement management process is to collect and record the condition of roadway segments. The prototype GIS-T, with its visual display capabilities, can show the segments color-coded (e.g., a thematic map) by various attributes that would greatly facilitate the process of data entry and editing. Omissions in the data input would be immediately apparent from segments in the roadway showing no data. Furthermore, errors in measurement or coding would also be readily apparent. One such example would be where a long stretch of roadway that had recently been resurfaced included a segment that was coded as having severe transverse cracking. The prototype GIS-T can produce a series of color-coded map displays that would identify differing segments.

By using the conditional querying capabilities of TransCAD, it is possible to produce thematic maps that highlight distressed roadways. These maps can be used in system assessment. All of the graphical products produced as a part of the assessment can be easily understood by management, political figures, and citizen groups, helping to clarify issues and obtain needed support. Figure 5 shows a simple pavement rating map produced using the prototype GIS-T.

TABLE 2 Pavement GIS-T Attribute Fields

TRANSCAD FIELD	DATA TYPE	COLUMN WIDTH
TIGER File ID	Number	10
Length	Number	4
Direction	Coded	2
Street Prefix	Character	1
(i.e. N,S,E,W)Road Name	Character	25
Type (i.e. St.)	Character	6
Suffix (i.e. SE)	Character	2
Code	Character	4
Maintaining Jurisdiction	Character	10
Total Score	Number	3
Alligator Cracking	Number	3
Pothole/Patching	Number	3
Transverse Cracking	Number	3
Longitudinal Cracking	Number	3
Rutting	Number	3
Raveling	Number	3
Oxidation	Number	3
Bleeding	Number	3
Pavement Width	Number	3
Shoulder Width	Number	3
Functional Classification (i.e. Arterial,Collector)	Coded	2
Estimated ADT	Number	7
Pavement Evaluator	Character	20
Date of Evaluation	Date	8
Posted Speed	Number	3
Travel Time	Number	5

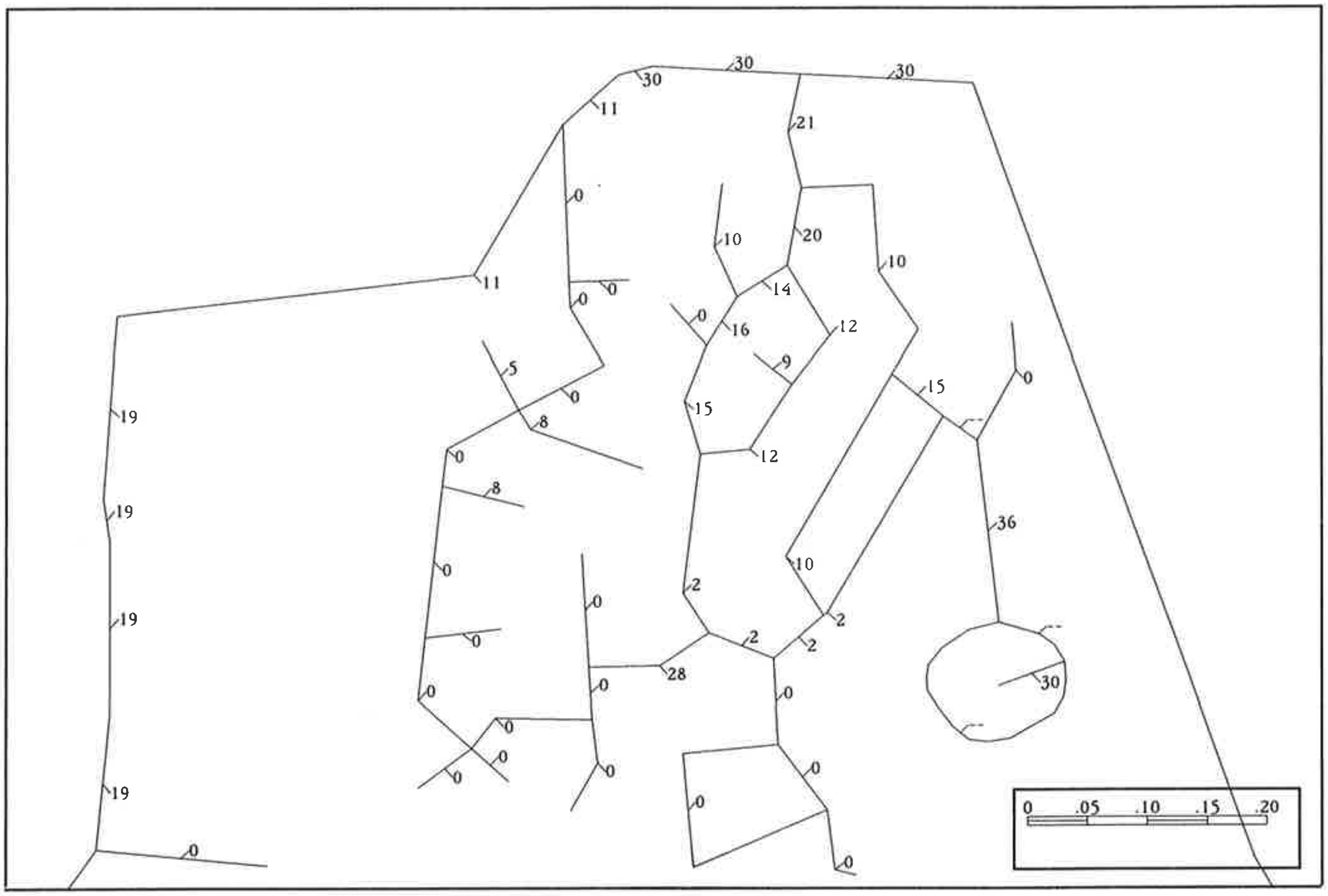


FIGURE 5 **Pavement rating map.**

One area of extension of the prototype GIS-T's capabilities is the addition of expert system technology to the pavement GIS. By adding "intelligence" to the GIS, maintenance strategies could be automatically identified on the basis of the types of distress of a particular roadway.

Transportation Planning and Land Use

Microcomputer-based travel forecasting usually involves the creation of lengthy ASCII files for input. The TRANPLAN transportation demand modeling package used by the county agency is one example of a travel forecasting software that requires ASCII text input files. Conventional data base packages are one means of simplifying the process of creating the ASCII input files. A characteristic of conventional data base packages that make them poorly suited to manage TRANPLAN input files is the cryptic nature of having to rely heavily on identification numbers (nodes, links, TAZs, etc.). Input file management is the perfect domain for a GIS because attribute information stored in a GIS is linked to spatial objects. Thus, it is possible to edit information for a TAZ by selecting the graphical representation of the TAZ from the computer screen. Likewise, link and node data can be edited by selecting the appropriate intersection or roadway from the screen. For the prototype GIS-T, a TAZ data base was developed to manage land use data.

As mentioned earlier, an ideal application of a GIS is to manage TRANPLAN input files. By developing certain procedures, it is possible to translate land use information that is included in the TAZ data base into productions and attractions that can be used by TRANPLAN. A translator can also be developed to extract link and node data included in the GIS and produce a network file compatible with TRANPLAN.

Transit

The transit module of the prototype GIS-T currently is very limited. Nevertheless, it can be used to illustrate a number of capabilities. The potential for additional capabilities is vast if more data fields were to be included in the data base.

Two layers of information are included in the transit module. The first layer contains bus stop information. The bus stops in the study area are shown in Figure 6. This figure was used for illustrative purposes only and does not necessarily represent actual bus stop locations. Ridership information associated with bus routes is included in the data base. Using this information, it is possible to produce a variety of descriptive thematic maps that can show, for example, those bus stops that have the greatest number of boardings and alightings during any period of the day.

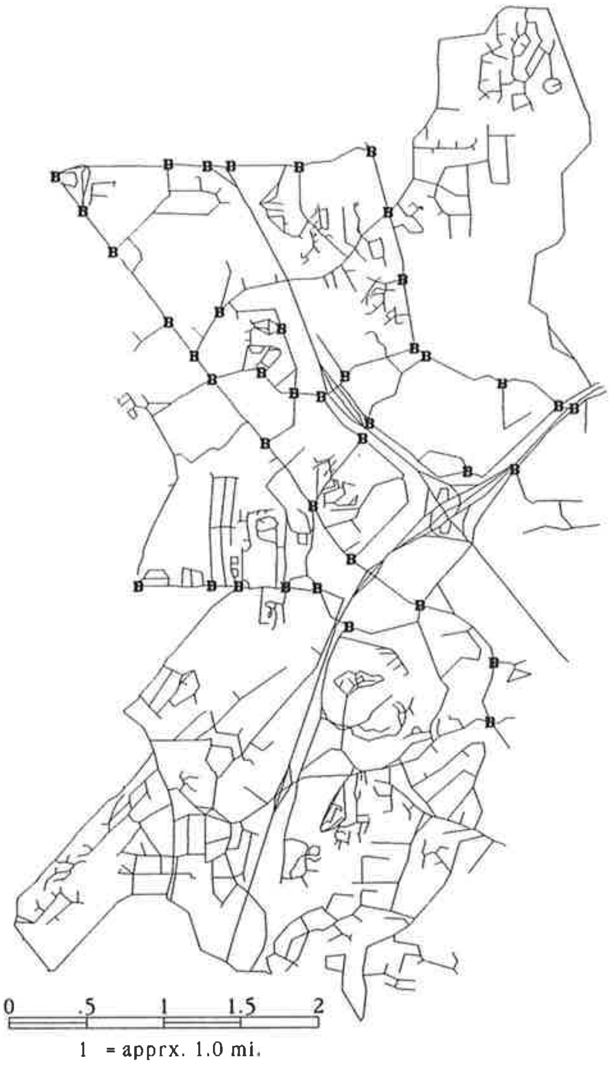


FIGURE 6 County bus stops.

FUTURE EXPANSION OF PROTOTYPE GIS-T

After an extensive evaluation of the prototype GIS-T, county officials have decided to implement an expanded GIS-T for the entire county. This GIS-T will be developed as part of an integrated multipurpose countywide GIS. The GIS platform (e.g., TransCAD/GisPlus, ARC/INFO, etc.) for this multipurpose GIS has yet to be determined.

In developing the prototype GIS-T, care was taken in the design of specialized functions (e.g., the TRANSYT-7F interface) to be as portable as possible to other GIS platforms. Thus, the programs were developed in the C programming language and designed to operate external to the GIS platform. The only functions that would be unique to a particular GIS platform would be how information is passed to and retrieved from the specialized functions. Information would probably be passed using customization tools built into the platform.

CONCLUSIONS AND OBSERVATIONS

This paper has described a prototype GIS-T that has been developed for a county transportation agency. The GIS-T provides a strong management decision support capability for the officials in this agency. In particular, it provides more than just a data management capability. As shown in the integrated accident records system, the development of TAZ input data into TRANPLAN and the automated development of input files for TRANSYT-7F, the prototype GIS-T is taking an important step in using GIS technology to provide more efficient analytical procedures. Clearly, given the current manual operation of the data collection and management procedures in the county agency, the GIS-T data management capabilities will also provide an important benefit to the agency.

The implementation strategy adopted for this project is also worthy of note. The project team consciously started with a small geographic area of the county to show the constraints that might be found in a countywide application, while at the same time illustrating in concrete terms what benefits could be obtained. This small-area prototype strategy has been well received by agency management and has been used to show midlevel managers the capabilities that could exist in their agencies.

REFERENCES

1. W. A. Sarasua and C. C. Hatton. *GIS-T7F: A Geographic Information System–Based Data Input Module for TRANSYT-7F.* Georgia Institute of Technology, Atlanta, 1991.
2. C. C. Hatton. *Development of a GIS-Based System for Developing Input Files into TRANSYT-7F.* M.S. thesis. Georgia Institute of Technology, Atlanta, 1991.
3. H. J. Simkowitz. Using Geographic Information System Technology To Enhance the Pavement Management Process. In *Transportation Research Record 1261,* TRB, National Research Council, Washington, D.C., 1990.
4. D. R. Fletcher et al. *Pavement Management Decision Support Using a Geographic Information System.* Report FHWA-DP-90-085-006. FHWA, U.S. Department of Transportation, 1990.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Creating a Municipal Geographic Information System for Transportation: Case Study of Newton, Massachusetts

MARC BAILEY AND SIMON LEWIS

The potential of geographic information systems (GISs) as depositories of urban data has been made clear by a number of authors. As a next step, further consideration should be given to developing the best methods for accessing and manipulating such data for transportation applications. Building simple transportation tools directly within a general GIS package may be an appropriate avenue for smaller agencies without sufficient staff and resources to support multipackage solutions. A case study in the city of Newton, Massachusetts, demonstrated the modification of a GIS to provide an additional range of transportation functionality, including traffic assignment, vehicle routing, reapportionment of traffic zones, and location of centers (such as fire stations) on a network. An important aspect of this case study is that only readily available sources of data were used. Such an approach may be particularly appropriate for simplified "sketch planning" purposes at the local level.

A broad range of transportation modeling software is available to transportation planners at the local level. Relative to the graphics display standards established by new geographic information system (GIS) software, most of the specialized transportation packages appear inherently limited in their general graphical editing and data query capabilities.

Conversely, most GIS packages historically have offered only minimal utilities for transportation analysis. These have included tools for geocoding and rudimentary shortest-path and network allocation capabilities. However, much of the data frequently integrated into a GIS are of considerable use for transportation planning, modeling, and highway inventory. For example, zonal data stored in a GIS may be used for trip generation and distribution; street data may be used to create link files for a transportation model; and highway inventory records may be used to assist street maintenance and emergency response evaluations. Unfortunately, although full-featured GIS software is becoming increasingly accessible, even to municipal-level transportation professionals with limited budgets, most of the available GIS software does not offer specialized transportation tools.

To date, software development in GIS for transportation (GIS-T) has focused heavily on linking existing transportation packages to GIS. Another approach has been to build GIS tools with a strong transportation focus. TransCAD, produced by the Caliper Corporation, is an example of a package developed using such an approach (*1*).

Existing transportation models may access GIS data bases through either "cold links" (data transfer through ASCII files) or "hot links" (interactive use of a common data base). Also, a number of intermediate alternatives are possible—that is, the data links between packages may be more or less automated. Although the potential benefits of such tools are often referred to, a number of problems with linking alternatives have not yet been thoroughly addressed: for example, one street network link in the GIS data base may not translate into one link in the transportation data base.

At the municipal level, a key problem with this linking approach is that the average city traffic engineer or planner must learn a major GIS system as well as a major transportation package and the linking software. The user must learn not only two command sets, each of which may consist of hundreds of directives, but also different command sequences, interface standards, and overall design nuances. This situation may be feasible in a large transportation agency with dedicated staff and resources. However, in the context of technical staff with time and resource constraints, it is important that software tools be as integrated as possible in form and style, even at the possible cost of some marginal functionality.

There are significant advantages to building transportation analysis tools within a GIS. For instance, GIS packages have sophisticated built-in features for display and mapping to facilitate the visualization of model inputs—a loaded transportation network, for example—and the direct comparison of analysis results. Integration of analysis tools within a GIS permits a more direct conversion of geographic data to the specific input data required by analysis models. Such integration is particularly advantageous if the GIS used is what is commonly referred to as a "full-featured" GIS, offering a complete range of facilities for cartographic input, query, and analysis. In addition, in various regions and states, certain GIS products have been adopted as standards for use; if these products can be used, associated benefits are greater still.

If transportation analysis tools are provided within a GIS environment, they may be used in a similar style and form to the more standard GIS tools. The syntax of the commands also may be similar, so that the user may perform all required tasks within a single, familiar environment. Incorporation of transportation functions within GIS may be critical if more sophisticated transportation applications are to be learned and used effectively, particularly at the local level.

The institutional advantages of building transportation applications tools within a comprehensive municipal GIS framework are also clear. Sharing of coverages (digital maps with attribute data) among and across agencies within a common

GIS/Trans, Ltd., 675 Massachusetts Avenue, Cambridge, Mass. 02139.

software environment potentially removes redundancies in data collection and maintenance efforts, while strengthening data quality and analytical consistency. Transportation planning may be significantly improved by direct access to land use and population information from other departments, such as assessing. Other departments, such as planning, might benefit greatly from direct access to transportation department information, such as network modeling results.

The work summarized in this paper concerned the development of a GIS-T framework for transportation analysis. Project technical papers provide more detailed review of this effort, as well as references to the many works in the field on network data and analysis applications. It is not the intent of this paper to describe the particular algorithms and models used in detail or the intricacies of the particular GIS—in principle, it would have been possible to complete the work in several alternative fully functional GIS packages.

BACKGROUND

In Massachusetts, the state agencies under the Executive Office of Environmental Affairs provide GIS maps, digital data, and analyses to all levels of government and the private sector. This program, called MassGIS, is extremely successful and recently won several prestigious national awards.

Within this context of intensive state GIS activity, GIS/Trans, Ltd., is working with the city of Newton, Massachusetts, to develop a case study for the integration of transportation applications within a municipal GIS. Some of the lessons learned in Newton may be largely transferable to other cities and towns interested in investigating their capacity for transportation GIS.

Newton is a suburb west of Boston, straddling the metropolitan area's main east-west road and rail routes. The city's 84,000 inhabitants are concentrated around a number of village centers, where congestion and other pressures on traffic management are increasing. At the same time, resources to address these problems are severely restricted.

The city has attempted to be forward thinking in evaluating how to better manage and analyze its planning and engineering data and is currently developing a citywide GIS. This initiative began in part through the running of a semester-long workshop in conjunction with the computer resource laboratory of the Massachusetts Institute of Technology (MIT). In this workshop, MIT students surveyed the different city departments and completed a preliminary information systems analysis. From this work, a view of the potential for GIS was identified and an implementation plan was drafted. GIS hardware and software were acquired, and a parcel-level data base is currently being digitized.

The work discussed here originated in a number of research initiatives at MIT. In further pursuing these directions, GIS/Trans, Ltd., has assisted the city traffic engineer over the past 3 years in preparing to establish various transportation GIS applications for the city. The goal is to improve traffic data management and to allow the execution of transportation analyses that previously were very difficult, if not impossible. A fundamental benefit of the work is the provision of a data set for broad consideration and GIS analysis.

The work has also been completed as a case study representing the range of applications for which GIS-T might be used in a municipal setting. It is used in a 3-day National Highway Institute GIS-T course that will be presented at each state transportation agency over the next 3 years (*2*). The work to date includes activities in three main areas.

Matching TIGER files with Roadway Inventory Files

Some innovative programming was required to match U.S. Census Bureau TIGER records with Massachusetts Department of Public Works (MDPW) road inventory data. The resulting data set provides both a geographic network and street attributes for Newton.

Implementing Link to Transportation Planning Package

The data from the file matching were first entered into a transportation planning package, QRS II. This package was chosen because of its low cost and ease of use. Links to other packages such as TRANPLAN also have been developed and demonstrated. An origin-destination (O-D) matrix was generated by "cordon compression" from the Massachusetts Central Transportation Planning Staff (CTPS) statewide data base, using tools within the Urban Transportation Planning System (UTPS) suite of programs.

Development of Prototype Applications for Routing and Traffic Assignment

Newton was used to test the validity of incorporating transportation tools within a GIS environment (as opposed to merely linking GIS to transportation packages such as QRS II or TRANPLAN). The example applications built within a GIS environment include traffic assignment, routing applications (such as school bus route generation), network location problems (such as fire station location), and traffic zone reapportionment.

This approach differs from the cold and hot links described earlier, as the analytical tools reside as modules within the GIS. Essentially, the analytical tools are composed of the core algorithms and make direct use of the various facilities and programming services provided within the GIS. Thus, although the algorithms are functionally complete if provided with input in appropriate formats, they depend on the GIS's lower-level routines for file input and output, screen display, and other operating system tasks. By improving the various file input and output hooks developed below, other users would be encouraged to develop and incorporate algorithms to handle various analytical tasks of their own.

OPPORTUNITIES FOR LINKS BETWEEN GIS AND APPLICATION TOOLS

There are a number of alternatives for linking GIS and transportation analysis tools. These tools could range from the

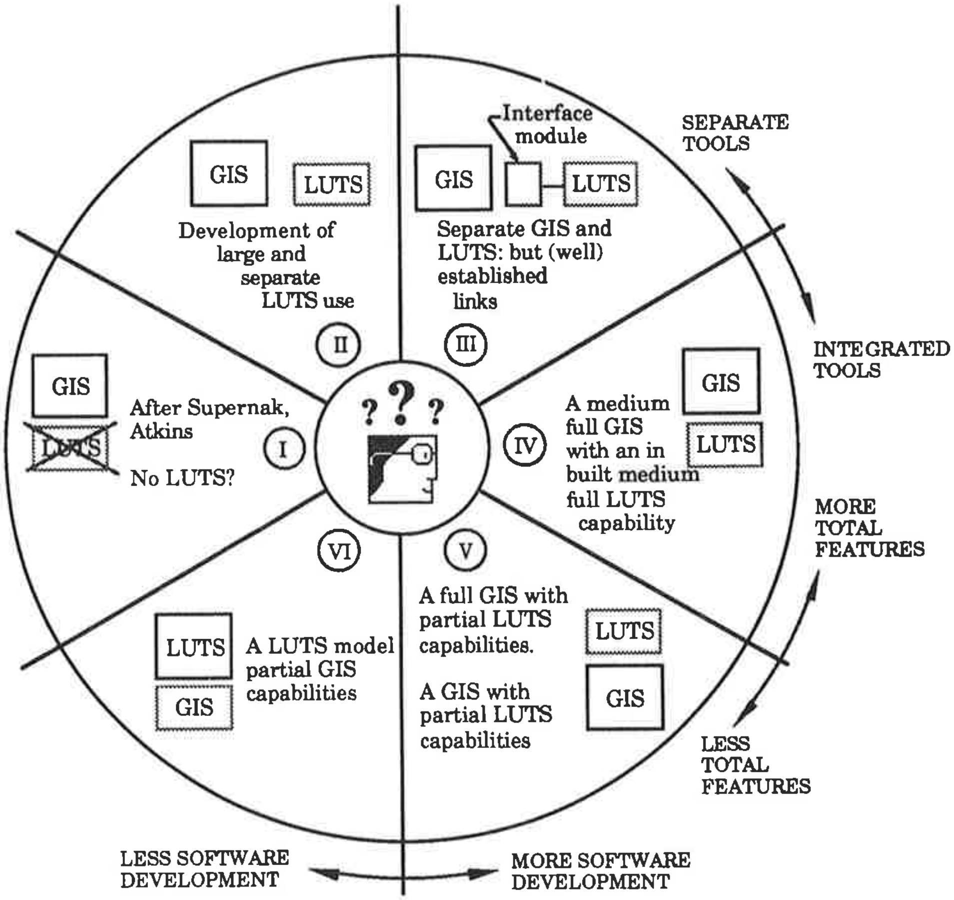


FIGURE 1 Links between GIS and transportation application tools.

traditional transportation models through to routing packages for school buses or snow removal. These alternatives have been described in detail elsewhere (*2–4*), so a full description is not given here.

In brief, Figure 1 summarizes some of these opportunities. The alternatives indicate various trade-offs system designers have made between the integration and separation of various tools. They also show whether the delivered systems have more or less utility and whether there is a greater or lesser degree of software development.

Models I through III assume that GIS toolboxes and transportation planning toolboxes are essentially separated. Model I asserts that there is a decline in use of the classic UTPS family of tools, given the criticisms directed at these models in various "before and after" studies, whereas the use of GIS continues to expand. Model II assumes the separate growth of GIS and transportation modeling tools. This assumption was the case until fairly recently; the GIS and transportation communities maintained their distance with separate professional organizations, conferences, and other activities. However, this situation has changed over the last 2 years. Model III allows for the provision of links between GIS and transportation modeling tools by the use of interface modules. These interface modules may use either hot or cold links for the exchange of data between packages.

Models IV through VI assume a closer linking of GIS and transportation planning code. Model IV is best described as the addition of graphical capabilities to many transportation packages. These additional graphical capabilities have become increasingly more sophisticated, allowing, for example, the maintenance of topological connectivity. Model V is the state of many existing fully functional implementations of GIS. Finally, Model VI represents the "transportation analysis toolbox" approach, in which many tools are disaggregated into component parts within a common GIS platform. The effort summarized in this paper was essentially concerned with an initial review of Model III, which is followed by a more extensive review of Model VI.

BUILDING THE DATA BASE

The first and most essential concern in developing a GIS-T is to generate or secure adequate data. To accommodate the Newton GIS-T model, two distinct and previously independent data bases were joined. The first of these is the TIGER/Line, developed and maintained by the Bureau of the Census as a cartographic information base to assist with the execution of the 1990 national census. TIGER is crucial to the Newton GIS-T because it provides spatial geometry for the various intended analytical tasks and a structure to which data base attributes can be linked. Additionally, TIGER provides certain roadway characteristic data, such as street addresses and intersection longitudes and latitudes. (A number of vendors provide "cleaned" versions of TIGER data, with missing street names and addresses inserted, positional accuracy checked, and so forth.)

The other key data file for the Newton GIS-T is the MDPW Roadway Inventory, a computerized record of geometric and structural data for approximately 33,800 mi of roadway in Massachusetts. Most of the data describe the dimensions of each street, including lane, shoulder, and sidewalk widths, with additional information on surface types and conditions, access control, federal aid, and categorical estimates of traffic flows. The inventory was developed to facilitate effective transportation planning and bookkeeping and to bring into accord the needs and resources of local, state, and federal agencies.

The greatest challenge was the actual matching of TIGER and MDPW records. Having evaluated several approaches to coupling these files, it was ultimately decided to create a single file containing, for each roadway segment defined in the original TIGER file, all the related information from both the original TIGER and MDPW files. These efforts generated a GIS-T data base structure flexible enough to be used in many applications.

In the integration of TIGER and MDPW files, one-to-one mappings based on street name are rare. TIGER street records are segmented according to intersections, whereas MDPW data for each street are linearly referenced and independent of other features. Furthermore, the TIGER and MDPW files have many missing or conflicting address ranges, rendering these virtually useless for matching.

For each street segment, the MDPW inventory provides an odometer distance measured to the hundredth of a mile. Although TIGER does not include a comparable data field, makeshift lengths were generated by using basic geometry from TIGER street segment endpoint longitudes and latitudes. Programs were developed to tie individual MDPW records to the corresponding groups of TIGER records on the basis of street names and these newly calculated link lengths.

The matching program indexed each record in TIGER by adding a field to the data base containing the number of the corresponding record in MDPW. In a first pass, about 86 percent of the records in TIGER were identified for updating with Roadway Inventory data. There are several explanations for the missing links. First, the MDPW inventory does not include entries for private roads, whereas TIGER does. Next, both TIGER and MDPW contain many "unnamed road" entries. Some subset of these records probably describes the same roads, but they cannot be matched using merely segment length (independent of street name). Finally, there are spelling discrepancies between the two files (e.g., Lindberg Ave. versus Lindbergh Ave.). A Soundex function, as well as some manual processing, was used to deal with many of these discrepancies. After the problems that could be identified were addressed, the indexing program was reexecuted and produced a 94 percent match. A separate paper will address the wider issues of coupling networks from different sources and the approach taken in detail in Newton.

The successful creation of an integrated data base for Newton provides some indication of the potential value of such efforts to the modeling field. Most important, it demonstrates that a process for merging such disparate sources as TIGER and state road inventory data may be developed inexpensively. By using the same steps as those taken for Newton, one could establish a similar data base for any city or town in Massachusetts. Furthermore, roadway data foundations analogous to the MDPW inventory might be used to produce similar GIS-T base data for municipalities outside Massachusetts.

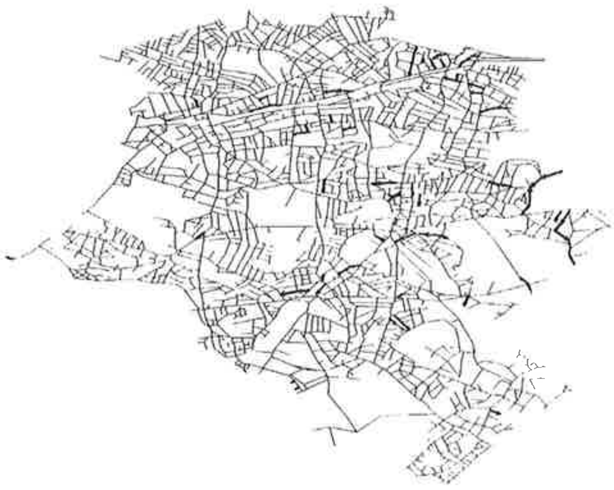


FIGURE 2 One-way streets in Newton.

With access to road attribute data in a GIS, a city traffic engineer can respond quickly and efficiently to queries about current road and traffic conditions of the city. For example, Figure 2 shows a map of the city of Newton with one-way links distinguished graphically. Many other graphical displays of road information can be generated, including dimensions, structural condition, and state aid status.

Considerable work is required to maintain any road data base. Geometric and attribute information change constantly. A GIS is currently most useful for graphically tracking the status of attribute data and verifying new attribute data when it is merged into the data base; GISs are less useful for merging geometric data.

APPLICATIONS

Traffic Assignment

A valuable tool in municipal-level transportation planning and management is traffic assignment. A traffic assignment may be used to predict the effect on traffic flows of changes in demand for travel or in the structure of the transportation network. The goal of the tools provided in this study was to meet the traffic engineer's needs for quick first-order assessments, or "sketch planning." For example, what would be the first-order effect on traffic flows across the Newton network of a bridge collapse that made a certain link unpassable? The tool is especially valuable for such "quick-response" situations when there is insufficient time to use a more complex analytical package.

Another potential use for the tool is concurrency management. For example, what additional traffic strain would a new shopping center in Analysis Zone x place on the network?

The basic types of input information required for traffic assignment are

- An analysis network, representing major highways and streets, which generally must be abstracted or reduced from a GIS coverage including all streets in the study area. Vehicle flow capacities and traffic volume-delay relationships for each road represented in the analysis network must also be specified.
- A set of internal and external traffic analysis zones (TAZs), defined for the city and its immediate surroundings. The volume of traffic generated is typically determined using the socioeconomic characteristics of these zones.
- An O-D trip table, which describes the demand for travel between all pairs of TAZs in the study area. Such a table is usually developed from a combination of past travel survey data and traffic flow observations. (Tools were also developed to directly generate these from traffic count data in the GIS, although these are not discussed in this paper.)

A city-level transportation analysis network and an O-D trip table were extracted from a UTPS-based regional transportation model for metropolitan Boston operated and maintained by the Central Transportation Planning Staff, an adjunct agency of Boston's MPO. The cordon compression technique was used to build the trip table, summarizing trips entering and leaving the city and allocating them to the surrounding ring of external zones. This technique—when combined with the use of GIS technology—can make transportation modeling cost effective and useful for many local agencies that previously did not have the resources to engage in a longer-term modeling effort.

With the data available in appropriate format, four basic steps are executed using the traffic assignment module developed within the GIS:

1. Read in the link data representing the analysis network, converting each two-way street into a pair of one-way analysis links. The travel impedance and capacity for each link are defined based on the coverage link attribute table item values.
2. Read in the O-D trip table data.
3. Perform the traffic assignment. Separate options are provided for "all-or-nothing," user equilibrium, and incremental assignment.
4. Write out the assignment results—final traffic flows and costs in each direction—as new items on the link attribute table.

Programs were developed to produce a graphical display of results for a single assignment (see Figure 3) or a comparison of traffic volumes for two scenarios (see Figure 4). These procedures define a set of line symbols of varying widths and a look-up table to associate each level of traffic flow with a correspondingly wide symbol. When two scenarios are com-

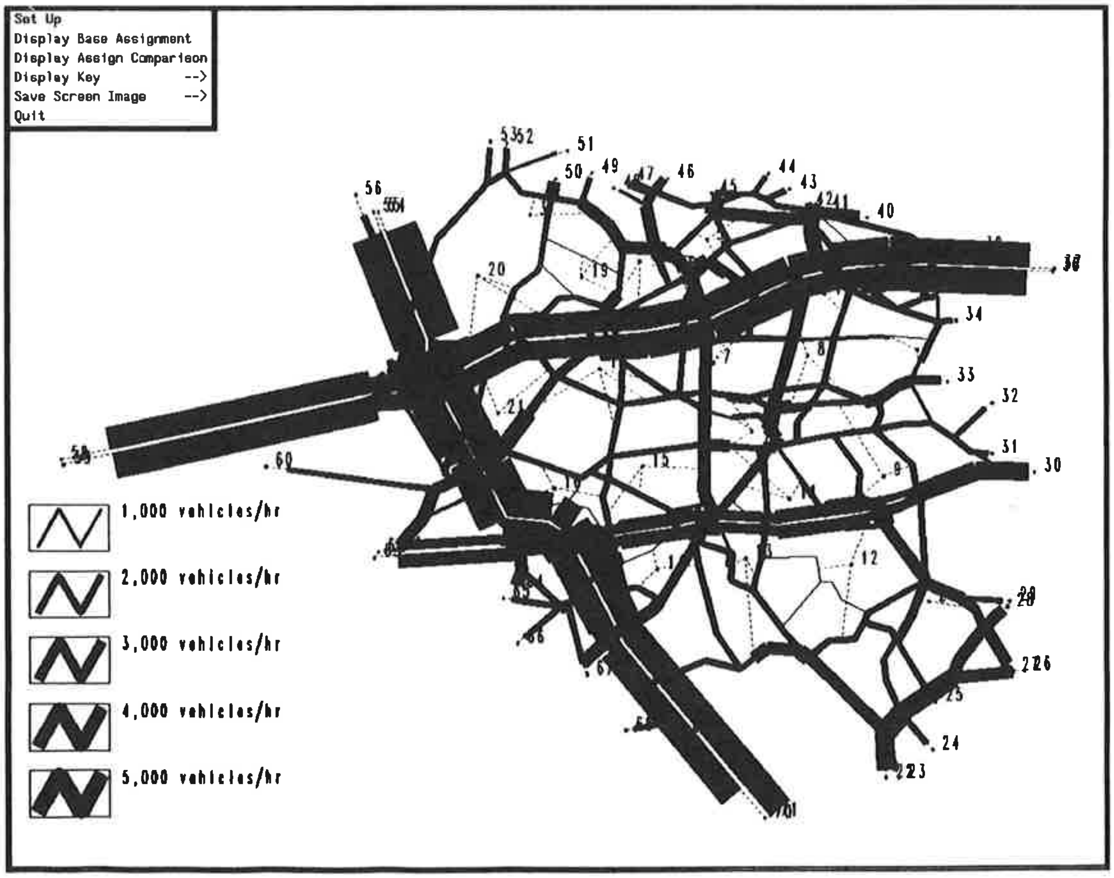


FIGURE 3 Traffic assignment: results for base assignment.

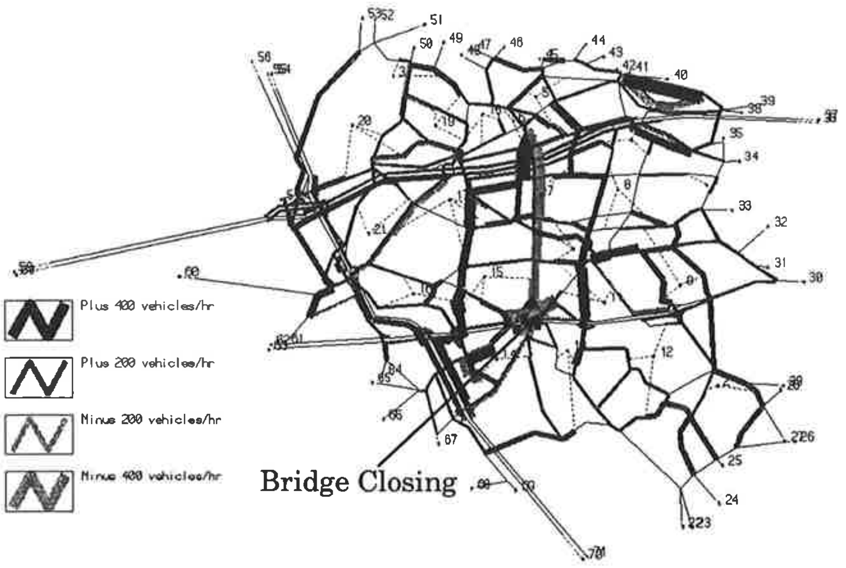


FIGURE 4 Comparison of two assignments—bridge closing.

pared, the display employs red and green line symbols to denote net increases and decreases in volume, respectively.

The ability to modify the transportation network (or trip table) and present modeling results in a GIS environment gives transportation planners the power to generate and thoroughly evaluate many "what if" scenarios within a limited time and resource budget.

Routing Applications

Many city services involve employees or work crews visiting customers, fixed facilities, or other landmarks at specific points. Some component of the total service cost relates to the time spent traveling between successive customers or work sites. Thus, by sending servers along the shortest route to all work sites, the city may realize a reduction in service costs. This is the basis of the well-known traveling salesman problem (TSP). Relatively simple examples might include a single person or crew performing mail collection or servicing street lights.

A related problem considers servicing sets of links rather than points. Typical applications include street sweeping or garbage collection. A current practical application in Newton is to calculate an optimal order in which to visit all streets in the city, so street lights may be inventoried and incorporated into the GIS.

More complicated vehicle routing problems—such as school bus transportation or dial-a-ride service—involve multiple service vehicles, time constraints on pickups and drop-offs, real-time response, or all of the above. The key to solving most, if not all, of these routing problems is to determine the solution to one or more TSPs.

Even for simple versions of the TSP, the difficulty in finding an optimal ordering of the stops is magnified enormously as the number of stops increases. Thus, for problems as small as a few hundred stops, a heuristic procedure is usually applied to find a good solution within a reasonable period of time rather than exhaustively determining the optimal shortest tour.

The GIS module to determine good solutions for TSPs requires the following input information:

- A link coverage defining the transportation network. The link attribute table must include attributes (e.g., distance, travel time) that will be used to specify travel cost.
- A set of points within the network to be visited.

To solve a TSP using this module, the following basic steps are performed:

1. Read in the link coverage, converting each street segment to a pair of one-way links.
2. Read in the list of points to be visited.
3. Build a table of travel costs between each possible pair of visit points. This is done with a shortest-path algorithm using any specified link attribute as the measure of link impedance.
4. Execute an algorithm to find a good ordering of the points in the tour.
5. Write the final order of stops to a data file.

The output results of the TSP module can then be read with a macro to produce a graphical display of the solution route. Figure 5 displays a solution tour to 18 school locations in Newton.

School Bus Applications

Routing school buses in the city of Newton offers an interesting development of base tools to demonstrate some GIS functionality in solving a real transportation problem. A macro-based menu system was developed for student clustering and bus routing. The menu is invoked from within the module used for displaying final bus routes.

Input information was provided by the Newton Schools Department, which maintains a computerized student data

FIGURE 5 TSP: optimized visit of Newton elementary schools.

base for school assignment and bus transportation planning.

The steps toward solving a school bus problem are to

1. Read the street network.
2. Read the school location and bus stop locations (stored as separate point coverages) into the module.
3. Cluster the student bus stops by invoking a specially written C-language program from the menu. The user is prompted to enter the maximum number of students that may be picked up per bus trip.
4. For each cluster of stops, determine the best route by calling the TSP module.
5. Display each route using a separate line color.

Figure 6 shows the final results of a clustering and routing analysis applied to students of the F.A. Day Junior High School. Each bus route visits up to 20 bus stops before returning to the school.

Traffic Zone Reapportionment Application

In the transportation community, reapportionment is often used almost interchangeably with redistricting. Reapportionment refers to the process of defining and redefining TAZs for transportation analysis. TAZs are areas from which traffic is generated. Various information is used to build a regional O-D matrix, which in turn is used as input data to traffic assignment or routing algorithms.

There are two main criteria for the creation of new TAZs. First, equality of population—the population of each district should be distributed as uniformly as possible. Major deviation from the target population can be justified only by extraordinary local circumstances, such as data problems, racial distribution, county lines, and shapes. The second criterion is compactness, or contiguity of districts.

In the past, redistricting has simply been performed manually. More recently, computerized automation of the redistricting process has been explored. User-specified criteria and conditions are input that allow the program to make judgments and decisions based on the statistics it computes at every stage of the process.

The reapportioning process used in this study has the following characteristics:

1. It is a heuristic approach: it presents one of the many possible solutions on a redistricting problem. Other solutions may be equally practicable, but this system produces a result that is as close to user specifications as possible.
2. It combines the power of GIS and the capabilities of the C programming language in the UNIX environment. The GIS provides access to raw data from census files, a user interface, and graphic display of the area to be redistricted, whereas the C-module performs the actual redistricting.
3. Currently it uses only the two major criteria of redistricting mentioned previously.

The GIS generates three tables: a population table, a correspondence table, and a small adjacency table. The population table contains information on the population at the block level; the correspondence table lists the blocks in each district; and the small adjacency table describes the adjacency relations among the blocks. The module is able to deduce the adjacency relations among the large districts from the correspondence table and the small adjacency table.

The main output file generated by the module is an updated correspondence table, with the new correspondence relations between the districts and the blocks. The summary file calculates the mean population difference of the districts from the target population and provides statistics on the extent of improvement in terms of population distribution after the redistricting process.

Perhaps the most important advantage of such a redistricting system is its automated nature. As a result, there is less human work involved, a great reduction in tedious and mechanical work, and fewer errors. Further description of this application is given in Lin (*5*).

Network Allocation

Optimal siting of emergency response facilities, such as fire stations, is a traditional and typical network decision problem. Partly because of the lack of available network data, many existing approaches have used areal allocation measures without reference to the underlying detailed network.

In a base form, the stated problem may be "to determine the number and location of a given number of centers given a maximum response time." Variants of the problem allow for certain centers to be fixed. Typically, this type of location decision problem is a "brute force problem." A number of simplifications may be made so that a computer-tractable algorithm can be constructed, such as reducing the problem so that only network nodes are considered or assuming that all centers have the same capacity.

The GIS was used to complete the spatial ordering of the data so that a specially prepared network allocation algorithm

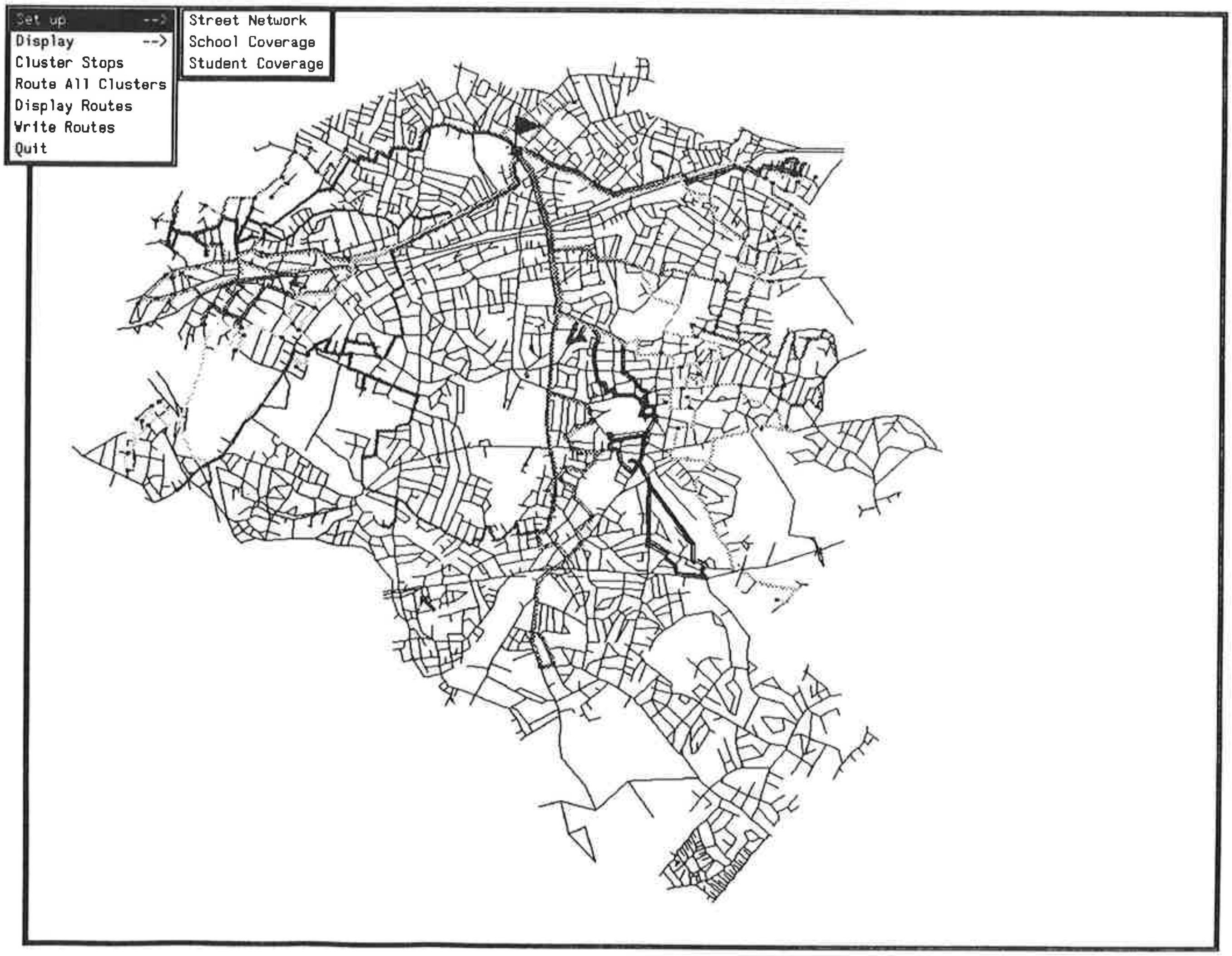


FIGURE 6 Results for school bus routing.

FIGURE 7 Network allocation.

could be run within the GIS. Once the algorithm executed successfully, the GIS was used to organize and display the network data. See Figure 7 for results of network allocation. Further description of this activity is given in Yuan (*6*).

CONCLUSIONS

The assemblage of a GIS-T toolbox approach to planning and analysis has been reviewed; the work described practically outlines the incorporation of one set of transportation applications within GIS. Some focus was made on integrating network data from multiple sources because, although they are technically less challenging to some reviewers, such issues have been major stumbling blocks to GIS-T implementation, both technically and practically. Second, the generation of a useful set of core transportation tools within a GIS is described. The purpose of this paper was to propose and demonstrate the framework, not to describe its instantiation in detail. This is a topic of continuing detailed research.

The simple transportation applications described for the city of Newton are a modest beginning. Greater sophistication will soon be included in the clustering and routing programs in the TSP applications, the interface will be further simplified, and many other applications will be added. The purpose of the initiative thus far has been to explore the general feasibility and value of developing transportation tools (and, in particular, simple tools at the local level) within a GIS package, rather than merely to link them to a GIS. Naturally, there will still be many applications for which more sophisticated and dedicated tools will be required, as was demonstrated in the first part of this work.

The research done here establishes a foundation for future GIS-T work, tailored to the local transportation engineer in particular and to others who may require simple, cost-effective tools for sketch-planning or quick-response applications.

REFERENCES

1. H. J. Simkowitz. Geographic Information Systems: An Important Technology for Transportation Planning and Operations. In *Transportation Research Record 1286*, TRB, National Research Council, Washington, D.C., 1991.
2. GIS/Trans, Ltd. *Introduction to Geographic Information Systems for Transportation. Participant Workbook*, Version 1.05. FHWA, U.S. Department of Transportation, March 1992.
3. S. Lewis. Use of Geographical Information Systems in Transportation Modeling. *ITE Journal*, March 1990.
4. S. Lewis, P. Cook, and M. Minc. Comprehensive Transportation Models: Past, Present, and Future. *Transportation Quarterly*, Vol. 44, No. 2, April 1990; pp. 249–265.
5. C. Lin. Redistricting with GIS: An Automated Approach. Presented at the New England URISA Conference, White River Junction, Vt., March, 1992.
6. W. Yuan. Fire Station Allocation Within a GIS Framework. Presented at the AASHTO GIS-T Symposium, Portland, Oreg., March 1992.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Applications of Geographic Information System–Transportation Analysis Packages in Superregional Transportation Modeling

W. PAUL GALLIMORE, DAVID T. HARTGEN, AND YUANJUN LI

"Superregions" are large areas 50 to 100 mi across containing several large cities that function as an integrated group. Transportation planning for such regions has typically been minimal, limited to the planning efforts of the separate urban area metropolitan planning organizations. As the area grows, however, interarea commuting patterns and integrated economies increase the need to analyze the entire region as a unit. The use of new Geographic Information System (GIS)-transportation packages to conduct such an analysis for the Charlotte, North Carolina, area, an emerging superregion of 1.6 million people, is discussed. The impetus for the study is a proposal for a 150-mi (or more) road around the region, called the Carolinas Parkway. Using the GIS package TransCAD, a sketch network for the region is developed by merging data from a variety of sources. Traffic is simulated over the network using a doubly constrained gravity model technique, calibrating simulated traffic to existing traffic counts. Preliminary forecasts of the parkway traffic are then made. An additional procedure, LANDSAT imagery, is being used to identify and categorize land uses in several alternative corridors for the parkway. The problems and opportunities presented by superregional modeling are discussed, and ways by which transportation planning will be changed by both the need for such models and the availability of the software to build them are suggested.

Historically, regional planning traces its roots back to the 1930s. The Federal Highway Act of 1962 mandated that "regional" (i.e., urban area) multimodal transportation planning be carried out in urban areas; this was the forerunner of local metropolitan planning organizations (MPOs). In most communities, these MPOs consisted of representatives of the major city, the major county, other major towns in this central county, and the city's extensions into rapidly growing rural areas and suburban towns in nearby counties. Cooperative state ventures were formed in cases in which the metropolitan areas encompassed more than one state.

Through the 1970s and the early 1980s, such organizational structures served well. However, in the late 1970s, changes in regional growth patterns began. Propelled by a generally healthy economy, lower land values on the fringe, and increasing interstate transportation system access, rapid development on urban fringes began to occur and space inside the large metropolitan areas began to fill. These changes accelerated demands on the ability of the metropolitan area to handle commuter traffic on its transportation network. Thus, additional pressure was placed on the organizations responsible for the management and planning of the affected transportation system. From this experience, the older 1960s style of urban planning process was noted to be ill equipped to plan effectively for the transformed metropolitan area or new "superregion." Increasingly, metropolitan areas recognized the existence of a broader scale of influence extending far from the city proper.

The superregional area can be roughly defined as an area encompassing several counties 50 to 100 mi across, often spreading over several states. It may be thought of as the "influence" area of the major city, larger than the MPO boundary, more like the "television market" or "maximum commuting market." Charlotte, North Carolina, an emerging superregion, is the center of a 13-county transportation planning coalition known as the Carolinas Transportation Compact (CTC). A study initiated by this group, the Carolinas Parkway Study, is of the superregional scale and was studied and analyzed using TransCAD, a newly developed geographic information system (GIS)-based transportation network analysis package, and ERDAS, a satellite imaging system for land use planning.

With the advent of such packages, different levels of analysis are now possible when conducting studies. Projects that were limited to analysis on mainframe systems can now be solved in the microcomputer or minicomputer environment at a fraction of the cost and time required by the larger systems.

With the need for an increase in superregional planning, the demand for the tools to handle superregional analysis increases concurrently. As this need increases, computer programs designed to analyze these issues continue to improve. Costs (measured in time required for analysis, person hours needed for such projects, software expense, and processing requirements) have consistently been decreasing over the last 20 to 30 years. Additionally, the ability to work with more complex analyses on smaller machines has been growing. Examples of new transportation analysis programs to work in the microcomputer environment included TRANPLAN, TMODEL, Microtrips, and TransCAD. TransCAD is a transportation analysis program that is GIS-based; that is, data in the form of point, line, and polygon layers can be tied to sets of coordinates that are tied to the transportation network being analyzed. TransCAD also has the ability to read TIGER files that are generated by the U.S. census, convert these files into usable formats, and perform various transportation network analysis procedures on the basis of such data. Through the use of this program, a high-level superregional analysis of the Charlotte, North Carolina, area is being conducted.

Department of Geography and Earth Sciences, University of North Carolina, Highway 49, Charlotte, N.C. 28223.

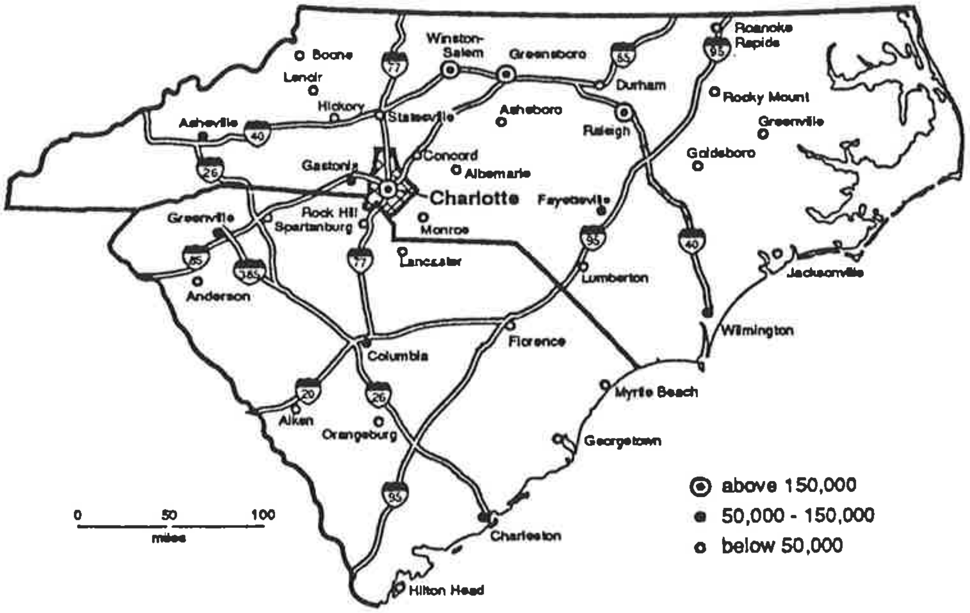


FIGURE 1 Charlotte-Mecklenburg and Carolinas region.

FORMULATION OF CAROLINAS PARKWAY STUDY

The greater Charlotte metropolitan area encompasses two states and covers a broad multiple-county area of the Piedmont region in south-central North Carolina and north-central South Carolina (Figure 1).

Growth was rapid in this region during the 1980s: the region's population grew by 15.89 percent. Such growth was not confined to the central cities or their suburbs but occurred throughout the 100-mi region. As expanding cities and their associated urban areas began to touch, boundaries began to blur, resulting in the formation of integrated economies. This new form of metropolitan area, known as a superregion, congealed, with growth occurring not only on the fringes but also within the region (*1*). As this growth occurred, a radial character for this region evolved along with an increasing imbalance in the jobs-to-housing distribution. Commuting patterns, which once were almost entirely within the individual counties, shifted to an intercounty pattern, and larger cities served as employment magnets and smaller communities acted as bedroom communities. This commuter-oriented radial traffic began to engulf previously adequate radial Interstates and increasingly came into conflict with cross-town and cross-region circumferential traffic using inadequate circumferential road systems. Thus, pressures from this growth encouraged a regionally thinking "mind set" for transportation planning. Existing urban area MPOs are responsible only for their respective areas, not the entire region, so some new form of organization was needed. Such thinking manifested itself in the December 1989 formation of the CTC and the planning areas for the individual MPOs located throughout the region (see Figure 2).

High on the agenda of the CTC is improvement in the circumferential access of the region. The placement of a regional ring road that encircles the region would improve this type of access. Envisioned as an interstate, limited-access expressway, this road would circumscribe Charlotte about 30 to 50 mi out, connecting the major towns and cities of the region (Figure 3). With an estimated length of about 140 to 180 mi, construction costs of the road have been estimated to range from $2 to $4 billion.

The idea for a Carolinas Parkway has encouraged the cooperative assessment of its impact. Beginning in fall 1991, the North Carolina Department of Transportation and the South Carolina Department of Highways and Public Transportation with the help of a feasibility study are expected to begin studying the feasibility of a ring road that encircles the region. This study will include corridor analysis, alternatives, corridor scanning, effect of facilities, and estimated costs of construction. The planning process for modeling the parkway contained eight elements: (a) very long modeling time frame (20–40 years); (b) great uncertainty in population and employment forecasts; (c) undefined alternatives; (d) separate multiple urban areas; (e) no common base-year data; (f) limited regional travel information; (g) large and distant major nodes (e.g., other cities) at the edge of the region; and (h) unknown feedback between transportation and economic growth. These features introduce great uncertainties in the analysis but also permit more rapid sketch planning.

BUILDING THE MODEL'S GIS DATA BASE

Data necessary for model development consisted of several types: network description and related data, population and employment data, and trip data. Traffic forecasting was accomplished by (a) coding in a base network for the region, (b) developing trip "load nodes" (generated from population and employment data) on the network, (c) generating flows between these origins and destinations through the use of a gravity model, (d) assigning the traffic to the network, (e) calibrating the base model by using observed traffic counts, and (f) assigning projected traffic on the future networks.

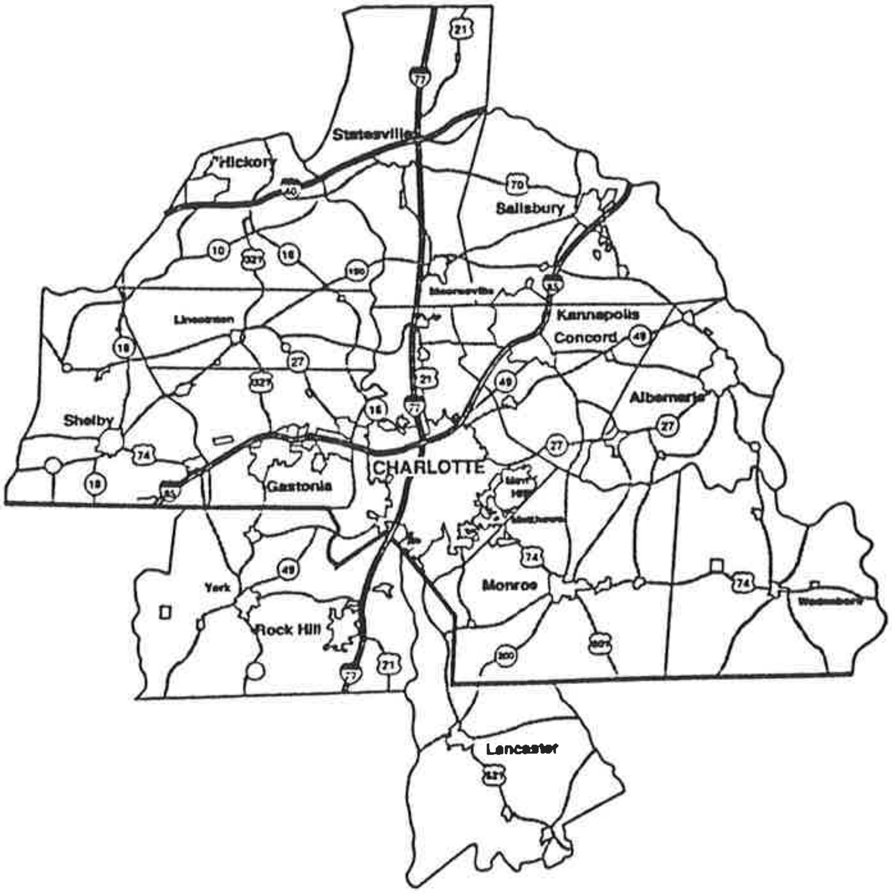


FIGURE 2 Carolinas Transportation Compact.

FIGURE 3 Possible location for Carolinas Parkway.

Network Data

The strategy for modeling was to create several networks reflecting the construction of several transportation projects. These networks can be summarized into the following categories: (a) the 1988 base network, (b) the formalized transportation improvement plan (TIP) network additions, (c) additional possible transportation improvements under discussion but not yet on the TIP, and (d) the Carolinas Parkway corridors and their alternatives. Data describing the base network and its additions (e.g., location, lanes, and traffic counts) were supplied by local agencies, state highway departments, and the TransCAD vendor.

The initial starting point in building the GIS data base consisted of developing a base transportation network. Two options were available to accomplish this task. One option was to begin with digitized TIGER files describing the region's complete road network and strip out the unnecessary links. The other option was to begin with a simplified regional network by digitizing additional links. This initial base network for both states consisted of parts of the U.S. Interstate and highway systems in addition to parts of the state highway systems and originated from Oak Ridge National Labs. The second option, network tailoring by densification, was chosen because the analysts wanted specific control over which link sections to incorporate. Basic data for each link included speeds and lengths, number of lanes, capacity, and traffic.

One useful feature of TransCAD is its ability to store many (literally hundreds of) characteristics describing links included within a network. One link characteristic variable used in the model was link travel costs or link impedance factors, intended to account for node delays as highways pass through towns. During the coding of the regional transportation network, each link was assigned to a link-type category. These categories consisted of U.S. highway rural links, U.S. highway urban links, Interstate highway rural links, Interstate highway urban links, state highway and local rural links, and state highway and local urban links. For each link type, a corresponding link-type travel time penalty (LTTTP) was estimated and added to the data base. The travel time was then calculated as

$$\text{TRAVEL TIME}_l = [\text{LENGTH}_l/(\text{SPEED}) - 5)]$$

$$+ (\text{LENGTH}_l * \text{LTTTP})$$

where

TRAVEL TIME_l = travel time for a given link l (hr),
LENGTH_l = length of link l (mi), and
SPEED_l = posted speed of link l (mph).

LTTTPs were estimated by assigning a time delay factor associated for each link type. These factors were initially estimated and subsequently adjusted slightly during the calibration phase of modeling. The final values used in the model are listed in Table 1.

TABLE 1 Link-Type Delay Factors

Link Type	Delay Factor in hours/mile (sec/mile)	
US Rural	.00416	15
US Urban	.01000	36
Interstate Rural	.00100	3.6
Interstate Urban	.00800	30
State/Other Rural	.00416	15
State/Other Urban	.01666	60

As an example, consider a 1-mi section of urban U.S. highway with a posted speed of 45 mph. The travel time on this link would be (1.0/(45 − 5) + (1.0 × 0.0100) = 0.035, implying an effective average speed (with delays) of 28.6 mph. Thus, the travel time penalty effectively accounts for delays that otherwise would not be observable in a high-level sparse network.

Figure 4 shows the final sketch network. Note that it is very sparse, showing only major routes. Also note that some towns are shown as single nodes. This is consistent with the high-level and long-range nature of the forecast.

Population and Employment

In addition to the transportation network, a method is needed to include population and employment data. Such data are used to generate trips for assignment to the network. Traditionally, this is accomplished by delineating traffic analysis zones (TAZs), and both population and employment data are then aggregated to the TAZ. The centroid "loading node" for each zone is then defined and the data are attached to this point, which is in turn tied to the network using loading links. In the development of data for this regional model, population and employment data were directly tied to loading nodes found on the network. This technique eliminated the need for zones and zonal loading links and helped to eliminate the calibration problems associated with using these pseudolinks.

The 1988 population estimates from the Census Bureau were used as the control data source for base population futures. This data source provided population data by county

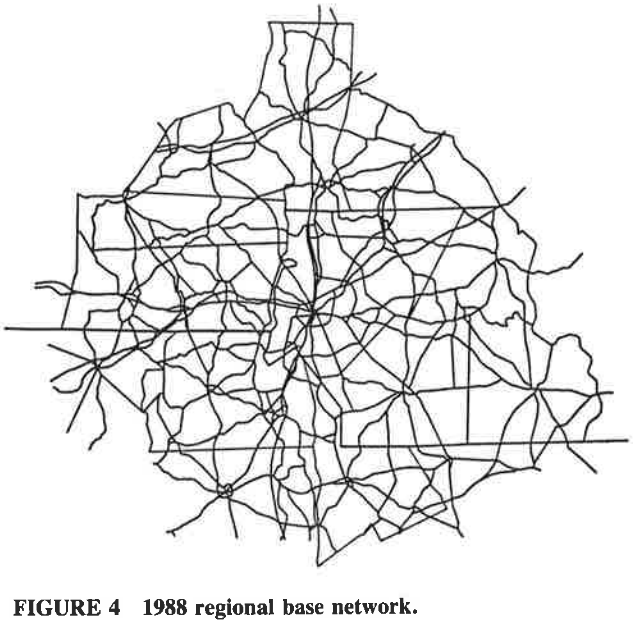


FIGURE 4 1988 regional base network.

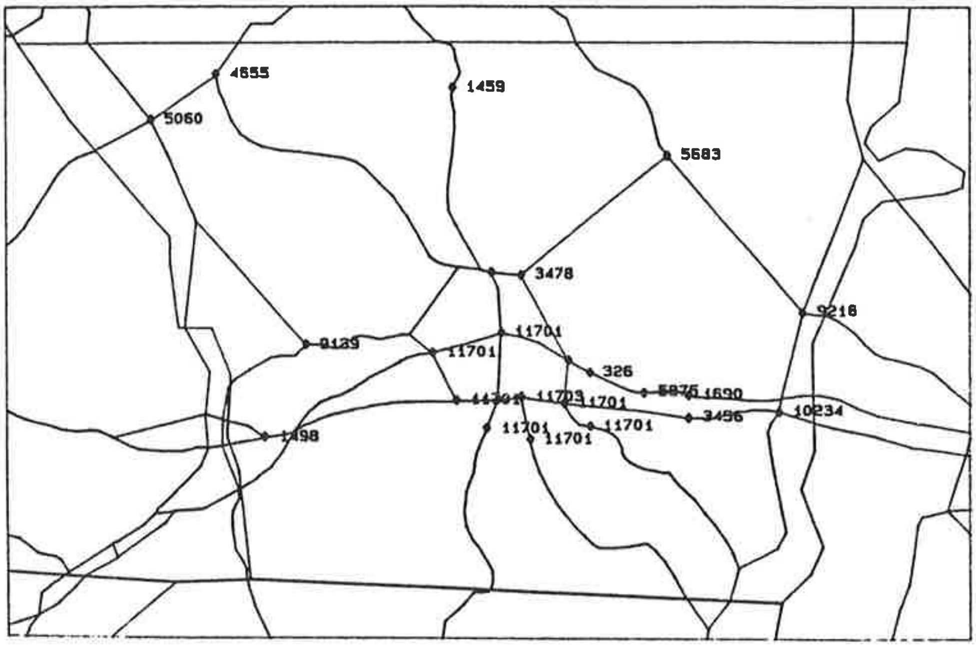


FIGURE 5 Population loading nodes for Gaston County.

and incorporated place for 1980 and 1988. Later checks with 1990 data show good agreement. From these data, each place's share of a county's total population was determined and used to distribute the remaining county population found outside the incorporated places within the county. The rationale behind this procedure was that in the area surrounding these incorporated places is a band of population living just outside the incorporated limits. Population density maps were used to help in "attaching" a town's population to the network notes that make up the town. Figure 5 shows the population loading nodes for Gaston County, North Carolina.

Future baseline trend population distributions were projected using a shift-share analysis: that is, historical changes in the share of population in each incorporated place within a county were assumed to continue in the future. North and South Carolina county population projections were used as controls for distributing total population growth or reductions. Sources for these county control projections were *North Carolina Population Projections: 1988–2010* and *Population Projections for the Census County Divisions in South Carolina: Number of Inhabitants, 1985, 1990, 1995, 2000* (*3*).

Unlike population, employment data by place of work are not publicly available between census years. The source for these data was the North Carolina Economic Security Commission (*4*). These data were distributed on a subcounty level using ZIP code distributions of retail and nonretail employment provided by a private data vendor—Equifax, Inc. (*5*). By overlaying a ZIP code map with a map of the network, the employment by ZIP code could be assigned to loading nodes. Employment data were broken down into retail and nonretail categories, using the classifications in an NCHRP report (*6*). Retail employment was defined as any industry that attracts commercial or trade traffic off the streets (e.g., retail establishments, professional service offices, government service offices), and nonretail was defined as everything else. Figure 6 shows employment assigned to loading nodes in Gaston County.

Future baseline employment projections were developed using current county population-to-employment ratios and applying such a ratio to projected future county populations. These projections were then distributed to subcounty nodes proportionally according to each node's proportion of the current employment.

Trip Generation

In development of the regional model, trip production and attraction estimates were used for each loading node on the transportation network. Person productions and attractions for these nodes were estimated using the areawide procedures from an NCHRP report 187 (*6*), based on dwelling units (derived from population counts), retail employment, and nonretail employment. These estimates of productions and attractions were developed for home-based work trips, home-based nonwork (HBNW) trips, and non-home-based (NHB) work trips. Vehicle occupancy rates by trip purpose (e.g., HBW, 1.13; HBNW, 1.55; and NHB, 1.44) were used to convert person trips to vehicle trips. Total vehicle productions and attractions for each trip type were added and converted to vehicle trips per node for use in the sketch network. A sketch factor was required to "scale down" vehicle trips to fit the sketch network on which they were assigned (i.e., there was not enough network to show the full vehicle miles traveled (VMT) unless productions and attractions were also adjusted). This conversion was made by using FHWA's Highway Statistics, 1989 (*7*) to develop a ratio of federal system VMT to total state VMT. Federal system VMT was used because the sketch network consists largely of the federally supported streets. A sketch factor of 0.721 was calculated for North

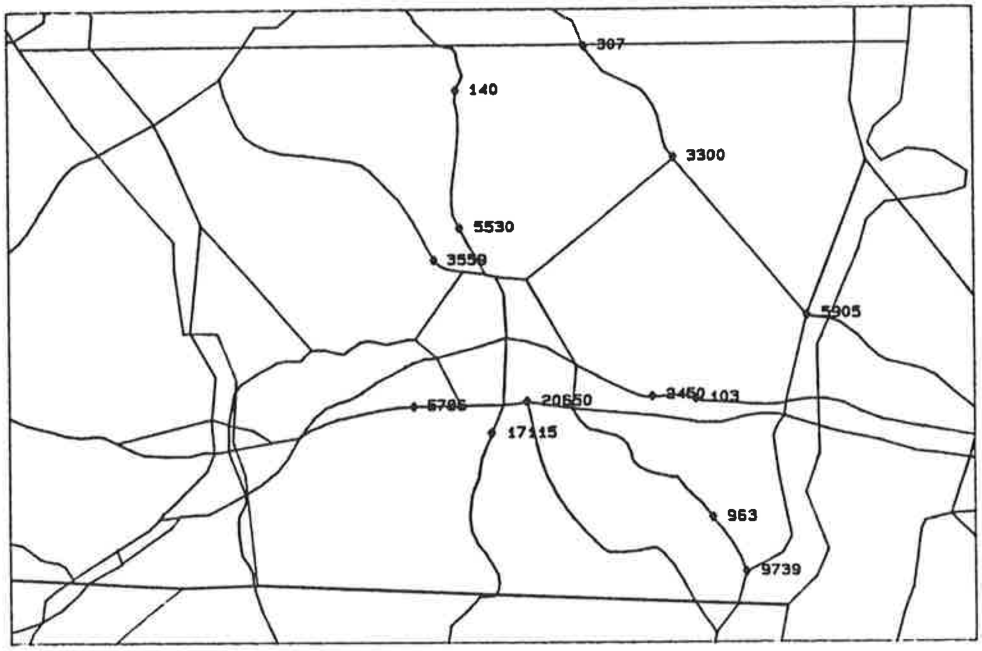


FIGURE 6 **Assignment of employment to Gaston County loading nodes.**

Carolina and then applied to the vehicle trips. For all trip calculations (person, vehicle, and sketch), total productions and attractions equaled each other for the region. By balancing these productions and attractions, a doubly constrained gravity model was used in developing an origin-destination (O-D) flow table. Table 2 shows the regional summary of productions and attractions.

An additional issue involved trips with origins or destinations outside the region. External-cordon O-D traffic surveys are suggested in the literature in syntheses of travel movements (*8*). However, this region does not have an external traffic survey. With prohibitive cost and time constraints of conducting one, and the reduced need to use such a survey because of the region's large size (about 5 percent of trip ends are external), an alternative method of synthesizing external trips was used. External trips were estimated by treating the nodes on the edge of the region because loading nodes have both productions and attractions. These productions and attractions were set such that the annual average daily traffic (AADT) on these links, which connected them to the rest of the network, was accurately estimated.

A possible result of not having an external O-D matrix involves likely shorter trip lengths and the possible underestimation of traffic on major roads and Interstates throughout the region. As external-external traffic enters from the edge, it would find destinations within the region rather than travel directly across the network as an external-external trip. In an analogy illustrating this effect a superregion is compared with a lake with many islands scattered throughout. If a rock is dropped into the lake at the edge and waves are generated, they travel across the lake's surface until they are absorbed or deflected by the islands. The net result is that very little, if any, of the wave is able to reach the far side of the lake or region. The problem would be most severe for interstate or other long-distance travel routes—possibly such as an outer belt.

TABLE 2 **Summary of Superregional Sketch Travel Data**

County Name	Vehicle Productions 1988	Vehicle Attractions 1988
Anson	39450	34380
Cabarrus	124074	99226
Catawba	193980	229613
Cherokee (partial) *	77047	55028
Chester *	57257	46498
Chesterfield (partial) *	33800	25406
Cleveland	118987	93105
Davie (partial) *	14550	14550
Gaston	225385	190402
Iredell	132317	115476
Lancaster	81247	55278
Lincoln	62749	38075
Mecklenburg	631412	912627
Rowan	159638	124971
Stanly	71732	56377
Union	108680	83797
York	164315	121811
Total	**2296620**	**2296620**

* Not a Member of the Carolinas Transportation Compact

CALIBRATION OF REGIONAL MODEL

Calibration of the regional model consisted of balancing the simulated traffic generated by the model on the base network with actual measured AADT found on the network. Traffic counts rather than a trip length distribution were used, because of their available, low-cost nature. This is different from many simulations that calibrate to an O-D matrix or to a trip length distribution: in superregional modeling, such a matrix generally will not be available because the areas cover more than just the urban region. Our procedure of calibrating to the AADT approximates the procedure suggested by Willumsen but does not produce the optimum O-D matrix or optimum calibration (*9,10*). Instead, we develop a reasonable (not

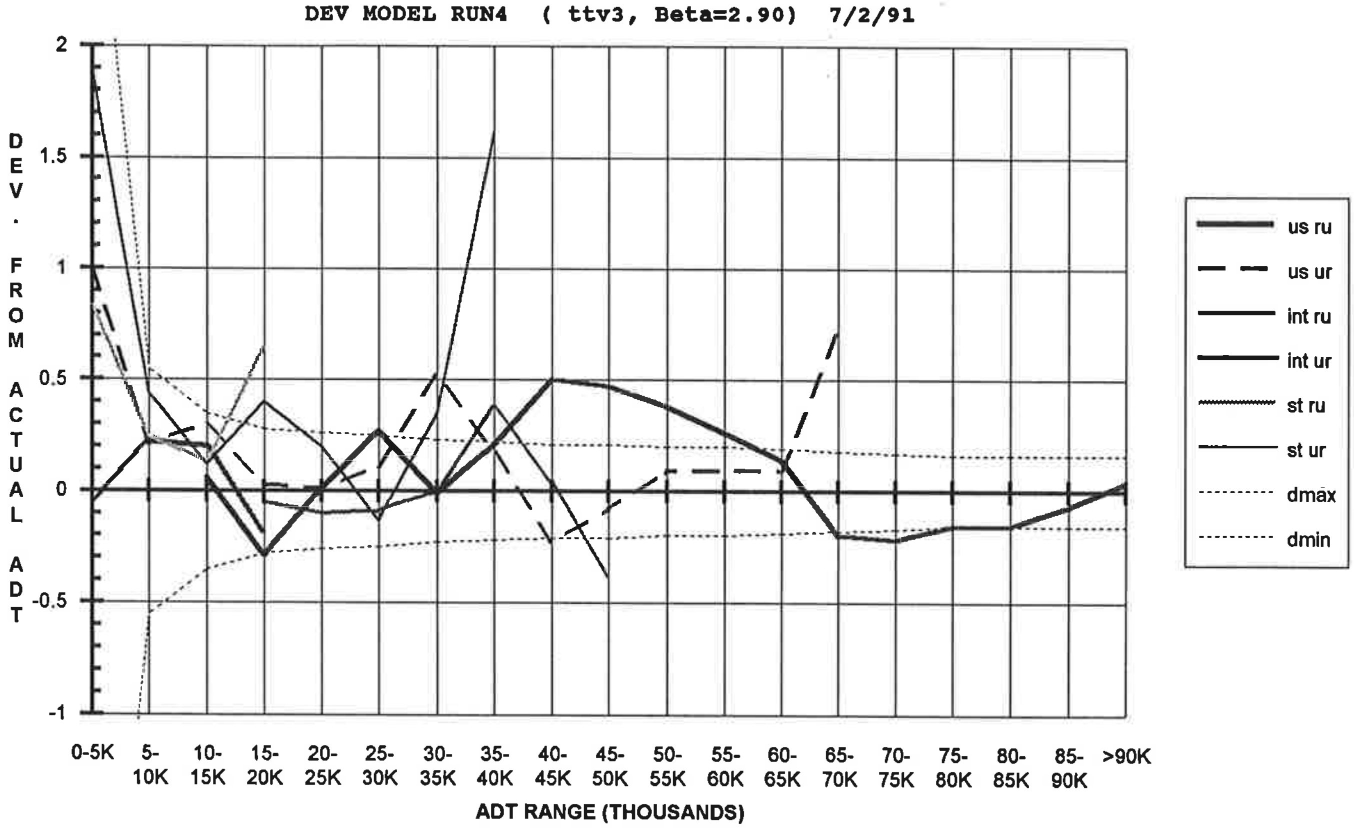


FIGURE 7 Link-type estimated ADT deviation by volume ranges for model 4.

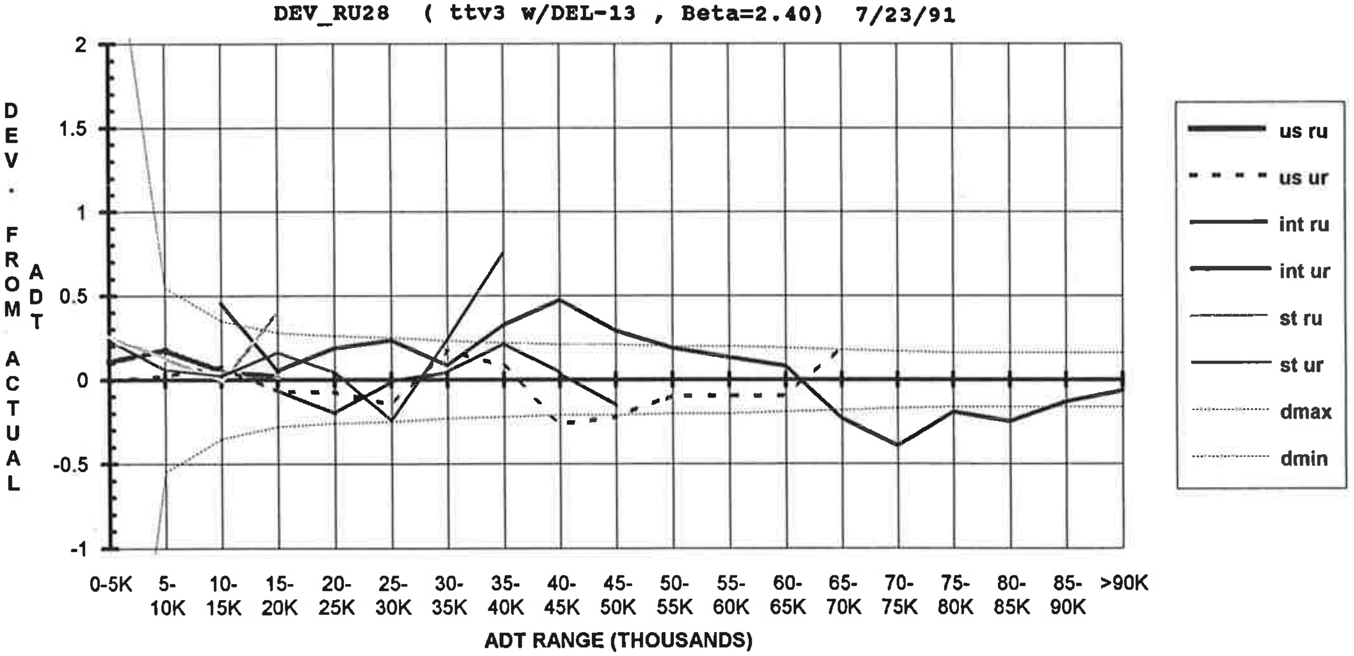


FIGURE 8 Link-type ADT deviation by volume ranges for Model 28.

optimum) simulation through iteration using different beta (decay) coefficients inserted into the gravity model. To compare actual and estimated traffic, an NCHRP report was used to develop acceptable deviation ranges for links with different volume ranges (*11*). Such tables were produced for each link type. By overlaying each of these link-type deviation graphs, a tool was developed to help calibrate the model. Figure 7 displays the deviation graph of an early run of the model. Figure 8 shows a later run of the model with adjustments made to both the beta value and the LTTTP.

By minimizing these deviations between estimated and actual values, additional corrections to the model were made. These corrections consisted of adjusting several parameters of the model, which included changing the beta value of the gravity model, and changing link-type travel time impedance values for individual links. Individual link deviation review also led to finding errors in network coding and more even distribution of the population and employment data. Over a number of trials, a decay coefficient of 2.4 was ultimately selected. Summary tables of VMT and vehicle hours traveled (VHT) deviation by link type, county, screen line, and region were used to assist in the calibration effort. The resulting final individual link deviations (ratios of estimated to actual VMT) were then used as pivot points to adjust model output for future traffic forecasts.

NETWORK ALTERNATIVES

For the analysis for proposed highway systems, including preliminary alternative corridors of the Carolinas Parkway, an additional set of links and nodes was added to the data base. The preliminary locations for the parkway were determined by the two state highway departments. Included in these additions were TIP improvements (i.e., new facilities and capacity upgrades), post-TIP proposals currently under discussion, and alternative alignments of the proposed Carolinas Parkway. For each of these three categories, unique codes were developed within the data base allowing various combinations of network improvements to be selected for analysis. Figure 9 shows the base network and all the proposed network additions.

PRELIMINARY SIMULATION RESULTS

A simple uniform balloon expansion of population and employment was used to show a preliminary estimate of traffic likely to use the Carolinas Parkway in 2010 (Figure 10). The parkway appears to be serving circumferential traffic, but the analysis is too preliminary for firm conclusions.

EVALUATION OF EFFECTS

Procedures used for evaluation of user effects will be based on traffic forecasts, using criteria familiar to analysts. User effects are broken down into travel time costs, operating costs, and accident rates. These effects are developed by converting VMT and VHT data into estimates of operating costs, fuel use, and user savings in a spreadsheet environment.

Indirect effects include changes in land use, business activity, and environmental effects. Induced effects are developed by converting VMT and VHT data into estimates of operating costs, fuel use, and user savings in a spreadsheet environment.

Indirect effects include changes in land use, business activity, and environmental effects. Induced effects consist of population and economic development within the region, the development of second-order traffic, environmental changes within the region, and change in land use patterns.

Both indirect and induced effects will be evaluated using economic impact models in conjunction with satellite imagery data. SPOT and LANDSAT satellite imagery will be used to classify land use in each likely corridor according to land cover. High-resolution satellite imagery has been widely used

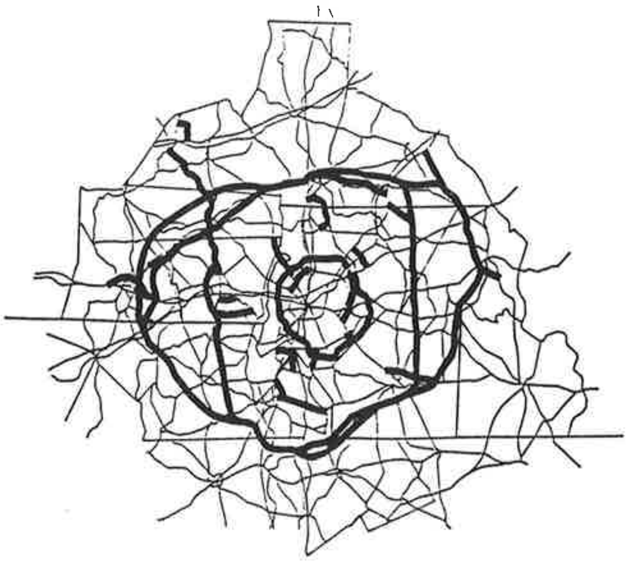


FIGURE 9 Regional base network with proposed network alternatives.

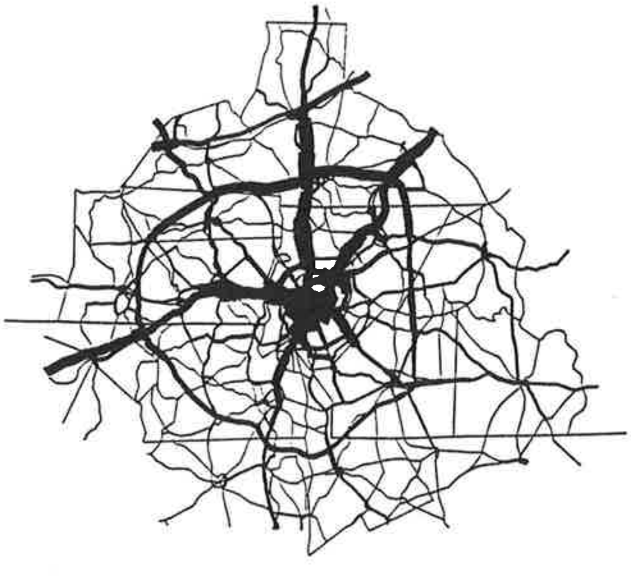


FIGURE 10 Preliminary estimated 1,010 AADT projected for Carolinas Parkway.

for land use and land cover analysis. It provides both timely and accurate information.

The objectives of this part of the study are to develop and apply a method of classifying land cover and analyze land use in the proposed corridors of the parkway by using satellite images. The procedure is divided into two phases: (a) develop a method of analysis in a test area, and (b) apply the method in the whole corridor area. Initially, the data require preprocessing, which involves image registration. Registration is done by rectifying the digital data to ground control points (GCPs) common to both U.S. Geological Survey 1:24,000 scale topographic maps and the digital image itself. From this registered set of data, a subset image can be drawn for further analysis and serve as the test area. For this study, the initial area chosen for analysis is located in both Rowan County and Cabarrus County, northeast of Charlotte, and includes the towns of China Grove, Landis, Kannapolis, and Concord.

ERDAS image processing software, hosted on a Sun workstation, is the platform used in the project. Both supervised and unsupervised classification algorithms were used to classify the land cover. The land cover is being classified into seven types: water, softwood forests, hardwood forests, croplands and grasslands, wetlands, open areas, and urban or built-up areas. To obtain more detailed land use classes, both visual interpretation and field verification will be needed since the resolution of the multispectral image is only 20 m. Existing aerial photographs are a good reference for assisting in determining conventional land use categories (e.g., residential, commercial, agricultural, transportational, etc.).

IMPLICATIONS AND CONCLUSIONS

The use of GIS-based packages in transportation planning and analysis has altered the scale, scope, methodology of analysis, and, with that, the relative power of planning organizations. Smaller cities or towns that had been analyzed individually can now be integrated into larger areas for analysis called superregions. Included in these aggregated regions are areas that could be missed during previous studies or that would be analyzed with differing underlying data and assumptions.

With the aggregation, information can be shared among different organizations and towns. An example would be step-down transportation planning for counties along a rim or crescent of the region, or more detailed modeling for urban areas. The use of GIS transportation packages allows and promotes the integration of data and analysis together. The data base for a given project is now an integral part of the GIS programs that works with those data. Additionally, the GIS package itself is now part of the evaluation and analysis process used in transportation research and problem solving. The use of study boundaries and research methodology can be thought of as the horizontal integration of study areas and the vertical integration of analytical functions.

GIS-based superregional modeling also has spatial, technical, and procedural implications. Since regions must work together to conduct the analysis, compromises of technique, data, base year, future year, alternatives designation, and a host of other issues are required. The price of regional modeling is therefore loss of independence and control over the separate forecasts once done in isolated fashion. The gain, however, is a much better understanding of how the region functions and a corresponding increase in the probability of building regional concerns on transportation futures. In this sense, superregional modeling can be viewed as an opportunity to enhance the region's competitiveness and improve long-term performance. In a nutshell, the goal of the Carolinas Parkway Study is to evaluate broad future visions for the region as a whole, setting aside the historic view of an isolated set of superregional areas.

ACKNOWLEDGMENTS

This research was performed by the Urban Institute, University of North Carolina at Charlotte, and has been supported by the Carolinas Transportation Compact. Additional thanks go to Al Stuart for assistance in developing population and employment projections.

REFERENCES

1. D. T. Hartgen and W. J. McCoy. Uncharted Waters: The Super-Regional Transportation Agency. *Transportation Quarterly*, Vol. 44, No. 3, July 1990, pp. 363–387.
2. *North Carolina Population Projections: 1988–2010*. Management and Information Services, North Carolina Office of State Budget and Management, Raleigh, 1988.
3. E. L. McLean. *Population Projections for the Census County Divisions in South Carolina: Number of Inhabitants, 1985, 1990, 1995, 2000*. Working Paper Series 031885. The Strom Thurmond Institute, Clemson University, Clemson, S.C., 1985.
4. *North Carolina Employment by Four-Digit SIC Codes by County, 1988*. North Carolina Employment Security Commission, Raleigh, 1988.
5. *County Business Counts and County Employment Counts by Zip Codes 1990*. Equifax Marketing Decision Systems, Inc., 1990.
6. A. B. Sosslau, A. B. Hassam, M. M. Carter, and G. V. Wickstrom. *NCHRP Report 187: Quick-Response Urban Travel Techniques and Transferable Parameter Users Guide*. TRB, National Research Council, Washington, D.C., 1978.
7. T. D. Larson. *Highway Statistics, 1989*. FHWA, U.S. Department of Transportation, 1989.
8. M. R. Poole and J. T. Newman, Jr. North Carolina Procedure for Synthesizing Travel Movements. In *Transportation Research Record 1139*, TRB, National Research Council, Washington, D.C., 1987, pp. 28–38.
9. L. G. Willumsen. Simplified Transport Models Based on Traffic Counts. *Transportation*, Vol. 10, Sept. 1981, pp. 257–278.
10. O. Z. Tamin and L. G. Willumsen. Transport Demand Model Estimation from Traffic Counts. *Transportation*, Vol. 16, 1989, pp. 3–26.
11. N. J. Pendersen and D. R. Samdahl. *NCHRP Report 255: Highway Traffic Data for Urbanized Area Project Planning and Design*. TRB, National Research Council, Washington, D.C., Dec. 1982.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Analysis of Transit Service Areas Using Geographic Information Systems

WENDE A. O'NEILL, R. DOUGLAS RAMSEY, AND JACHING CHOU

The transit route location and analysis problem requires the estimation of a population within the service area of a route. A route's service area is defined using walking distance or travel time. A procedure for performing service area analysis on transit routes using a geographic information system (GIS) is compared with the more common technique of buffering, and implementation strategies for three GIS packages—namely, ARC/INFO, TransCAD, and SPANS—are discussed. A case study is performed for Logan, Utah, where a new fixed-route transit system currently is being planned. The case study compares the technique proposed here with that from two other approaches. Effects of the socioeconomic data source on the accuracy of these approaches are discussed.

User-oriented transit service is designed to meet particular needs of a selected group of travelers. Transit routes are located to provide convenient linkages between a user's origin and destination in such a way that out-of-vehicle time, such as access and transfer time, is minimized. The planning of transit routes requires understanding demographics, land use, and travel patterns of an area. The dynamic nature of these systems necessitates regular review and analysis to ensure that the transit system continues to meet the needs of the population it serves.

Geographic information systems (GISs) provide a flexible framework for planning and analyzing transit routes and stops. Socioeconomic, demographic, housing, land use, and traffic data may be modeled in a GIS to identify efficient and effective corridors in which to locate routes. However, very little research has been done on transit applications using GIS, except for studies currently under way at the Massachusetts Institute of Technology (*1*).

Part of the route location and analysis problem requires estimating the total population in the service area of a route. A route's service area is defined using walking distance or travel time and indicates the route's accessibility to the public. Here we consider only the problem of estimating persons within walking distance of a route. The problem of identifying service areas for park-and-ride or automobile or bus users is assumed analogous to walk or bus trips. Once the service area for a transit route is identified, population information for this area is used as input to travel demand models for estimating ridership.

This paper describes a procedure for performing service area analysis on transit routes using GIS. Aerial photographs are used to verify the suitability of this method. This procedure is compared with the more common technique of buffering, and implementation strategies for three GIS packages, namely ARC/INFO (Workstation Version 5.0), TransCAD (Version 2.0), and SPANS (Version 4.0), are discussed. A case study is performed for Logan, Utah, where a new fixed-route transit system is being planned. The service area estimation technique proposed here is demonstrated using census data. Other data sources, such as postal delivery route and aerial photographs combined with census data, are considered with regard to the potential for reducing error in the estimation process. Results of this technique are compared with results generated using the standard buffer methodology.

PROBLEM DESCRIPTION

If pedestrian travel is assumed to take place on an isotropic plane, lines of equal walking time, or isochrons, may be constructed around a transit route to define its service area. The population within these lines represents potential transit users.

The scale of analysis for service areas ranges from entire transit systems to individual routes to bus stops. For instance, a transit system may be analyzed by constructing isochrons around each route to evaluate system coverage and duplication of service. In a GIS framework, the entity of interest in this analysis is a series of connected line segments that form the transit network. Microanalysis may be performed for each individual stop along a route. An investigation of this nature is made to ensure that bus stops are sufficiently placed along a route to provide adequate accessibility to the public. However, stops should be spaced far enough apart to minimize redundancy in coverage as well as prevent increased travel time, which results from stops too close together. The entity of interest in this analysis is a point that represents the bus stop location.

In general, the problem to be solved is to count occurrences of a data attribute within some proximity of an entity. Manual techniques using transit routes overlaid on aerial photographs are standard practice (*2*). In defining transit service areas, an attribute may be population or households and the entity may be a transit route or bus stop. Solutions to this general problem typically are achieved using GIS buffering operations. Buffers represent lines of equal distance around either point or line entities. Areas or polygons defined by the buffer line may be overlaid onto other polygon layers in a data base, such as census blocks, and the value of the attribute is determined on the basis of the amount of intersection between the areas of the polygons.

W. A. O'Neill, J.-C. Chou, Department of Civil and Environmental Engineering, Utah State University, Logan, Utah 84322. R. D. Ramsey, Department of Geography and Earth Resources, Utah State University, Logan, Utah 84322.

Service area delineation problems are more complicated for transit applications. Assuming that average walking speed is reasonable for determining access time to a transit route, service areas may be defined using distance. Walking, or pedestrian activity, typically occurs along a street network. Consequently, euclidean, or straight-line, measurement of distance in a buffering operation is not appropriate in this analysis. The common practice in transit planning is to use 0.25 mi as the upper limit on walking distance to a transit route (*2,3*). However, non-euclidean distance metrics may be more appropriate for locating buffer lines. For instance, consider an urban area with a predominant grid layout of streets. The Manhattan metric,

$$d = |x_1 - x_2| + |y_1 - y_2| \tag{1}$$

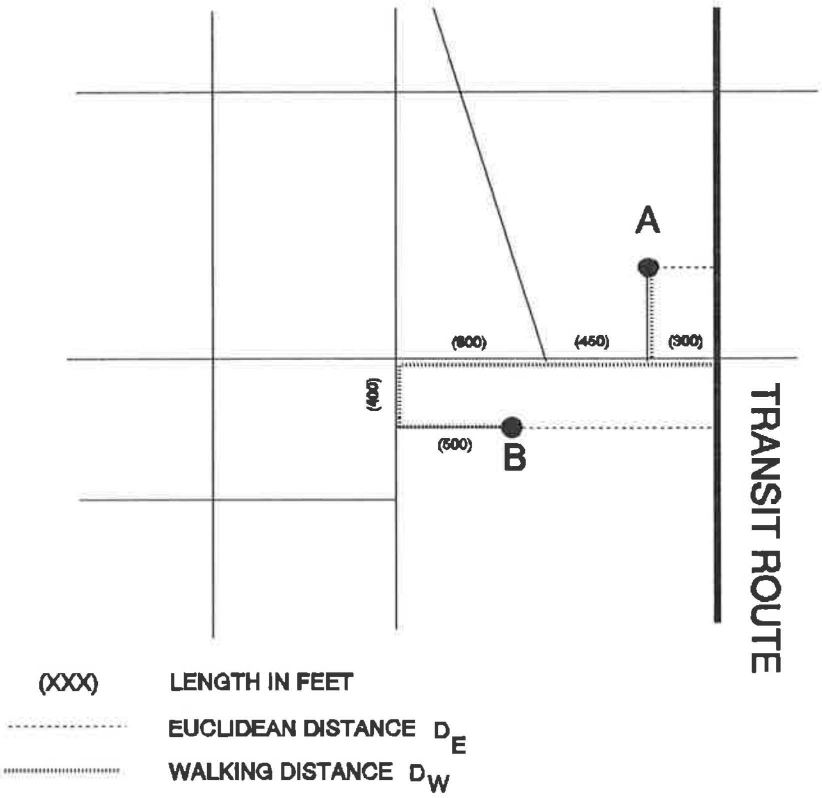
(where d is distance, and (x_i,y_i) is the cartesian location of ith point) more accurately describes walking distance than the measure of euclidean distance:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{2}$$

However, the accurate analysis of transit service areas requires models that can determine distance along paths in the road network, as opposed to buffering algorithms that use calculation of euclidean or other metric equations for distance. The procedure described here for estimating population in a service area uses specialized network analysis procedures in conjunction with standard buffering of entities.

Figure 1 graphically depicts the difference in results between euclidean and network distance measurements. Points A and B are quite close to the transit route when measured in straight-line euclidean distance. However, persons walking to the route must cover considerably more ground. Persons living at Point B may not consider themselves within the route's service area because of the long walk. In fact, as shown in Figure 2, the entire street block on which Point B is located is outside the 5-min walk (assuming a 3.5–mph pace) area for this route.

Figure 2 also demonstrates the difficulty of developing a buffer area or polygon enclosing the population within a certain distance. One procedure may be to identify each street segment within a specified distance and buffer it. Next, all



	A	B
D_E	300	850
D_W	750	2250

FIGURE 1 Network distance versus straight-line distance.

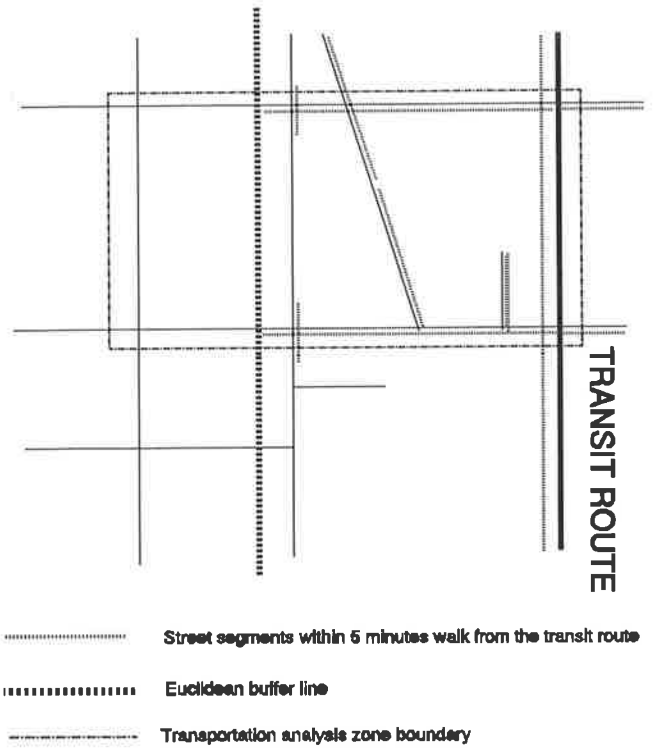


FIGURE 2 Streets in service area of transit route.

buffer areas may be joined to form one service area polygon, similar to the euclidian buffer polygon. These polygons may be overlaid onto the socioeconomic variable polygons to estimate population. But this is a tedious procedure, and there are no appropriate rules for defining the width of the buffer area for any particular segment. If population estimates are based on intersecting areas, the buffer width controls the estimated values.

To summarize, although the problem of delineating transit service areas is similar to the problem of creating buffers around entities in a GIS, an analysis based on euclidean buffer areas is inappropriate since the number of streets and people may be overestimated. Further, methods that attempt to buffer street segments, determined to be within a certain walk distance from a route, introduce error in population estimates.

REVISED APPROACH TO SERVICE AREA ANALYSIS

Problems with previously described methods may be overcome by eliminating the buffer process for individual segments and redefining the procedure for estimating population within service areas. The common approach for estimating populations within a service area uses euclidean buffering and polygon overlay techniques. Most often, polygon coverages of an area are formed by transportation analysis zone (TAZ) or census block, block groups, or tract boundaries. Attributes for these polygons include population, housing, and possibly, employment data. Polygons associated with socioeconomic data are referred to as analysis zones. A single value for each of the attributes is associated with each analysis zone. Consequently, a uniform distribution of the attribute within the area is assumed. In many instances, a polygon generated from the buffer of a route or stop intersects several polygons, either fully or partially. The following equation, called the area ratio method, is typically used to estimate population in the buffer area:

$$p_{w_i} = \frac{a_{b_i}}{a_{z_i}} * p_i \tag{3}$$

where

p_{w_i} = population in service polygon i within walking distance of the transit route,
a_{b_i} = area of service polygon formed from intersection of buffer polygon with analysis zone polygon i,
a_{z_i} = area of analysis zone polygon i, and
p_i = population of analysis zone polygon i.

This equation is appropriate only when the underlying street network is an evenly spaced fine mesh grid. A more appropriate measure—namely, the network ratio method—for allocating population uses network distances, as follows:

$$p_{w_i} = \frac{m_{w_i}}{m_{z_i}} \times p_i \tag{4}$$

where m_{w_i} is the total street miles within walking distance from the transit route of streets in analysis zone polygon i and m_{z_i} is the total street miles of streets in analysis zone polygon i.

An assumption of this model is that the number of houses on a street is proportional to the street length. Furthermore, this model assumes that houses are uniformly distributed on streets in a zone. In residential areas this assumption is relatively good. In mixed residential zones or zones with retail, industrial, and recreational activities, this assumption is weak but no weaker than those of the previous model.

These two models were applied on three neighborhoods in Logan. Six aerial photographs, two for each neighborhood, at a scale of 1 in. = 100 ft were used. These photographs represent the most accurate data base available on which to test the models. The small-scale photos enabled us to count the number of houses without field verification. Commercial buildings are not counted in these tests, and it is not possible to distinguish between single- and multiple-family dwellings.

The first neighborhood to which these models were applied is in downtown Logan. The street network is predominantly an evenly spaced grid in this older part of the city. The land use is primarily commercial along Main Street, where the bus route is located, and residential on adjacent and parallel streets. The second neighborhood chosen is primarily residential housing circa 1950; it is in the northeast section of the city next to the university. The street network is predominantly grid; however, blocks are more rectangular than square and an occasional cul-de-sac or dead-end street exists. The third neighborhood selected is the newest and has the highest housing costs of the three, containing residential development from the 1980s and 1990s. It is located directly south of the second neighborhood, and its street network has no apparent pattern.

For each photograph, the street with the transit route is identified, and an arbitrary rectangular analysis zone approximately 2,100 ft × 2,000 ft is drawn such that one edge lies along the transit route and the opposite edge is parallel to the route. The euclidean buffer line, parallel to the transit route, is drawn at approximately 1,600 ft. This distance is slightly longer than the 1,540 ft used to determine streets within walking distance along the network to include those houses directly across the street from houses within a 1,540-ft euclidean buffer. The number of houses in the analysis zone and buffer zone was determined from the photos. The area of the analysis zone and buffer zone was calculated. Total street miles in the analysis zone and within the 1,540-ft network distance were calculated. The area ratio model and network ratio model were used to estimate the number of houses within the service area defined by both the euclidean buffer and the network walking distance method.

Results are presented in Table 1. The network ratio method is shown to be more accurate for determining the number of households in primarily residential, modified grid street network areas. It is unclear whether the larger error in the downtown, regular grid neighborhood can be attributed to the network structure or the nonuniform spacing of houses along streets caused by the mixed residential-commercial land use.

To perform network ratio service area analysis, the following tools are needed in a GIS:

1. A procedure to total street miles within each analysis zone polygon,
2. A procedure to identify all street segments, or parts of streets, within a specified distance from a point or line entity,
3. A procedure to total street miles for those streets identified in Task 2 above for each analysis zone polygon, and
4. The capability of creating new data attributes from the mathematical manipulation of others.

TABLE 1 Results from Aerial Photographs

NEIGHBORHOOD	1A	1B	2A	2B	3A	3B
TOTAL HOUSES IN ANALYSIS ZONE	253	248	184	126	91	153
AREA IN ANALYSIS ZONE	42 x10^5	42 x10^5	42 x10^5	42 x10^5	359 x10^4	389 x10^4
AREA IN BUFFER ZONE	308 x10^4	308 x10^4	308 x10^4	308 x10^4	274 x10^4	296 x10^4
HOUSES IN SERVICE AREA (AREA RATIO METHOD)	186	182	135	92	69	117
HOUSES IN SERVICE AREA (AERIAL PHOTOGRAPHS)	216	212	135	122	85	150
ERROR RATE (%)	13.8	14.2	0	24.6	18.8	22.0
TOTAL STREET MILES IN ANALYSIS ZONE	2.42	2.57	2.25	2.08	1.28	2.26
TOTAL STREET LENGTH IN WALKING DISTANCE FROM TRANSIT ROUTE	1.45	1.54	2.25	1.78	1.21	1.62
HOUSES IN SERVICE AREA (NETWORK RATIO METHOD)	152	149	184	108	86	110
HOUSES IN SERVICE AREA (AERIAL PHOTOGRAPHS)	195	195	184	117	86	124
ERROR RATE (%)	22.1	23.6	0	14.5	0	11.3

EXISTING GIS SOFTWARE CAPABILITIES

Most GIS software packages available on the market satisfy Tools 1 and 4. Those that do not can usually interface with user-designed software and off-the-shelf data base management packages.

ARC/INFO's network software, combined with its data base management tools, is fully capable of performing service area analysis. A detailed description of the steps required to accomplish data import and data base construction follows.

TIGER data can be imported into ARC/INFO using procedures outlined in the user manuals. Separating the various themes (i.e., roads, census tracts, block groups, etc.) into individual coverages involves relating arcs within the data base to the TIGER tabular data set. Roads are identified and separated using codes in TIGER files that indicate the facility type of an arc. These arcs are extracted from the data base and placed into a separate road layer. Census tracts and block groups are extracted in a similar manner and placed into corresponding layers.

Road networks are created by splitting road line segments into 5-m increments. This provides a 5-m resolution to use in identifying service areas. The ALLOCATE utility in the network portion of ARC/INFO is used to identify possible transit stops. For example, let each stop service a 1,540-ft distance (impedance) (490.44 m) in all directions. The GROW subroutine identifies road segments no farther than the specified distance and, consequently, the area serviced by each stop.

To identify service areas based on a simple euclidean buffer, the BUFFER routine is applied to each stop for a maximum distance of 490.44 m. The buffer polygon is used to CLIP any road segments located within the boundary. Each euclidean buffer totally encloses all roads identified by the ALLOCATE utility.

Census block groups are clipped using the same buffer area. Area estimates for each block group are calculated within the buffer zones and compared to the total area of each block group. Only those block groups that fall within the buffer area are examined.

Road networks and block groups within the buffer area are combined using the IDENTITY subroutine in ARC/INFO to identify roads associated with each block group. This provides an estimate of total road length within each block group falling within the buffer area.

Necessary steps for determining transit service area populations using ARC/INFO may be summarized for each of the procedures previously outlined.

1. Procedure to total the street miles within each analysis zone polygon.
    a. Clip roads with euclidian buffer zone. Clip census block groups with same buffer zone.
    b. Use *Identity* to join roads to census block groups using the block groups as the identity coverage and the line option to preserve arc topology within each road segment.

   c. Within *Info*, reselect for each census tract, block group combination located within the buffer area. Use *Report* with the *Total* option to tally all road segment lengths associated with that block group.

2. Procedure to identify all street segments, or parts of streets, within a specified distance.
   a. Use the *Densifyarc* subroutine with the *ARC* option to split each road segment at a specified interval (we used 5 m).
   b. Enter the *ALLOCATE* network utility and position *CENTERS* at selected points along the network. Set the *IMPEDANCE* for each center based on a predetermined walking distance (we used 1,540 ft).
   c. Use the *GROW* routine to identify those road segments located within the specified impedance (distance).
   d. *WRITE* arc data to the network coverage arc attribute table (AAT). This will identify which road segment belongs to a particular center.
3. Procedure to total the street miles for those streets identified in Step 2 for each analysis zone polygon.
   a. In *INFO* select road coverage AAT. *Reselect* all roads that have been assigned to a center.
   b. For each segment assigned to a center, *Reselect* for an individual census tract block.
   c. Run *Report* with the *Total* option for the length of each road to sum all road segments associated with a center and one census tract block.
   d. Repeat Steps *a* through *c* until all census tract blocks have been summed.
4. Creating new data attributes from the mathematical manipulation of others.
   a. Manipulate numeric fields using the *Calc* command in *Info*. Preparation for the mathematical manipulation includes the addition of pertinent data (i.e., total number of households per block group, and total number of individuals per block group) and the addition of an *Item* to hold the output.
   b. Use the *Additem* command in *Arc* to add the items needed to perform the calculation.
   c. Once in *Info*, use the *Reselect* command to select only those records that fit a particular description (i.e., arcs belonging to a particular block group or center).
   d. Use the *Calc* command, which uses standard algebraic notation, to calculate the output value for each selected record.

Analysis of service areas is somewhat complicated using ARC/INFO, but the sophistication of the software enables users to perform a wide variety of applications.

SPANS is capable of euclidean buffering only. Network procedures in SPANS are not well developed. Users may create specialized procedures to integrate with SPANS. However, because most transportation planners are familiar with vector models, a conversion to a vector structure from quadtrees is helpful first.

MODEL IMPLEMENTATION

The rest of this paper focuses on implementing transit service area analysis using TransCAD. TransCAD has a euclidean buffering procedure (*Query buffer zone* command) that buffers a single line entity. Several problems exist with this procedure for use in analyzing transit networks. First, the end of a line segment is not buffered, and the buffer line is drawn through the endpoint (node) instead of creating a semicircle with a set radius from the endpoint [see Figure 3 (*top*)]. Second, the algorithm handles only nearly straight line entities. For example, if the transit route runs south, then turns west for a few blocks, then turns north, the buffer is drawn as shown in Figure 3 (*bottom*). Third, the tallied results displayed in a pop-up window, as well as the buffer polygon boundaries, cannot be saved to the data base. To use the results for each transit route, a user has to print the contents of the window and enter them in the data base later. Finally, each transit route has to be represented as a single entity in the data base to buffer the whole route. Otherwise, each street segment in the route must be buffered separately and the results totaled manually. If the route is not straight, some areas are counted many times and population is overestimated.

TransCAD contains tools to accomplish previously mentioned Steps 1, 3, and 4. Further, this system facilitates developing user-programmed procedures using any programming language desired. Consequently, a simple tree-search algorithm was created to identify all links within a specified walking distance of each node along a transit route. The output of the procedure is a value between zero and one for each link in the network. This value indicates the proportion of a link within walking distance from the route (e.g., 0 indicates that none of the street segment has access to the route; 1 indicates that the whole street segment is within walking distance, and 0.5 indicates that half of the street segment has access to the transit route).

This new procedure for selecting links within a specified walking distance of a point enables TransCAD to be used

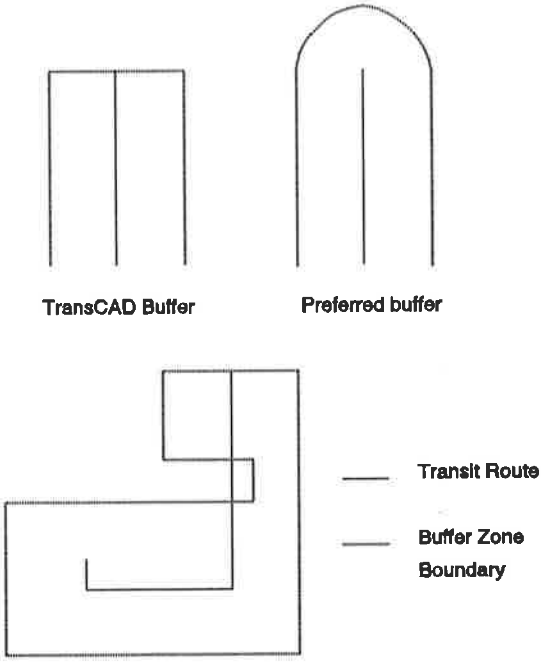


FIGURE 3 Example of buffering problems in TransCAD.

effectively in analyzing transit service areas. The necessary steps in performing this analysis are listed as follows in the order of the procedures defined above. Words and phrases in italic are from the TransCAD software.

1. Procedure to total the street miles within each analysis zone polygon:
    - *a.* In the *Map Display: current layer* = links (street segments). *Select* (by *pointing*) links on a bus route.
    - *b. Current layer* = analysis zones (i.e., block groups, TAZs). *Select (intersecting) by vicinity* (0.25-mi tolerance) of the *selected* links.
    - *c.* In the *Data Editor: display selected only*. In a blank data column (real values) use *edit column aggregate* to *sum* the length of links in each analysis zone.
2. Procedure to identify all street segments, or parts of streets, within a specified distance:
    - *a.* In the *Map Display: current layer* = links (street segments). *Select (intersecting) by vicinity* (0-mi tolerance) of the *selected* analysis zones.
    - *b.* Choose the *procedure* to *make* a *network* from the selected set of links.
    - *c.* Invoke the new *procedure* (transit service area analysis) to identify links from the network that are within walking distance from the transit route. Note, users must provide the walking distance in feet or walking time in minutes and average pace. Also, users must point to each intersection along the transit route for which the set of links in the service area is desired. A blank data column (real values) must be specified for output.
3. Procedure to total the street miles for those streets identified in Step 2 for each analysis zone polygon:
    - *a. Current layer* = analysis zones (i.e., block groups, TAZs).
    - *b.* In the *Data Editor: display selected only*. In a blank data column (real values) use *edit column aggregate* to compute (*weighted sum*) the length of links in each analysis zone weighted by the value returned in Step 2 indicating the proportion of each link within walking distance.
4. Creating new data attributes from the mathematical manipulation of others: In the *Data Editor: display selected only*. In a blank data column (integer values) use *edit column formula* to enter Equation 4. Note p_i may represent any variable attribute in a data base. For instance, the number of elderly people in the service area, the number of renters, or the number of children may be determined if these variables are part of the data base.

CASE STUDY

The process described here for delineating transit service areas and estimating the population within these areas is demonstrated for Logan, Utah. This city (population, 33,000) is in the process of developing a fixed-route transit system. Reasons cited for developing a transit system include

1. Increasing mobility for university students and households with one car,
2. Providing cost-effective transportation for the increasing elderly population,
3. Reducing traffic volumes, particularly on Main Street, and
4. Increasing accessibility to downtown businesses that suffer from limited parking availability.

The purpose of this case study is solely to demonstrate the potential of GIS for transit analysis. Results presented here are not derived from a comprehensive planning study and are not intended to recommend any particular route. In essence, enough data for demonstration purposes were available, so we used them to show how our model may be applied.

Data Base

The TIGER files for Cache County serve as the spatial data component of the GIS data base. Particularly, road segments, identified by codes starting with an A, and census block group boundaries constituted line and area (polygon) layers in TransCAD.

Attribute data were collected at two resolutions. For the lower resolution, population and housing information reported at the block group level was found in the Census Bureau's Summary Tape File 1A (STF1A). Note, data reported at the block level are not yet available. A subset of these variables was read from the tape into a TransCAD data base using a menu-driven basic program. Selection of variables, which are listed in Table 2, for the data base was influenced by a previous discussion of important transit planning factors (*4*) as well as the city's goals for the transit system.

Maps of census block boundaries and data base variables were generated during the study to assist in locating the stu-

TABLE 2 Census Block Group Area, Road Miles, and Population Data

TRACT NO.	BLOCK GROUP	AREA (SQ. MILES)	ROADS (MILES)	POPULATION
4	2	6.11	9.22	201
5	3	.19	3.04	1415
7	1	.44	5.77	2081
7	2	.46	9.03	2494
7	3	1.50	11.06	2040
7	4	.19	4.14	619
8	1	.13	2.34	1196
8	4	.12	2.06	526
8	5	.14	2.74	1158
9	1	.27	4.37	1620
9	3	.21	3.70	896
10	1	.15	2.50	784
10	5	.22	2.68	1156
11	1	1.83	9.79	1430
11	2	.53	5.43	332

dent, lower-income, and elderly populations to be served by the transit system. During route planning, maps of major traffic flows also are prepared. Other maps generated using GIS show the locations of major traffic generators in the city.

Ideally, a disaggregated data base enumerating individual households provides the greatest accuracy in estimating service area populations. Collection of detailed information is costly, both in gathering the data and building the data base. Even if housing structures are digitized from aerial photographs, field work is required to determine whether the structure is a single- or multiple-family house or is being used for commercial purposes. However, once the number of houses within walking distance is determined, from counting digitized house centroids along an arc, this value can be multiplied by the average household size for the block group to generate a more accurate estimate of persons in the service area. Increased accuracy results from relaxing the assumption that houses are uniformly distributed on a street and that the number of houses is proportional to the street length.

Another source of readily available data is information on the number of residences and businesses on each carrier route kept by post offices. Use of these data entails assigning a postal carrier route number to each street in a data base. Groups of street segments may be formed on the basis of postal route number. For each group, residences are distributed to streets on the basis of street length. An underlying assumption of this process is that the number of residences on a street is proportional to street length.

The impact of data sources on service area analysis is under investigation. It is hoped that the trade-off between the level of accuracy obtained from a particular data source and the cost required to build the data base will be discovered.

Results

For test purposes, service areas for six potential transit stops are shown in Figure 4. These stops form three service areas: one in downtown Logan on Main Street, one in a residential area covering the Island area development and Cliffside, and one near the university. An interesting feature of the second service area is that the buffer lines include streets that are inaccessible to the transit stop because of a steep barrier cliff. Barriers must be taken into account in service area delineation. Also, much of the land incorporated into the larger census tracts on the east side of town, particularly in Service Area 3, is undeveloped or agriculture or mountainous land. Residences are clustered on the western edges of these tracts.

Tables 2 and 3 contain information about service area population estimates for the block groups adjacent to the transit stops shown in Figure 4. Results are presented for three scenarios, namely,

1. Using euclidean buffers and the area ratio method (Equation 3),
2. Using the new procedure for identifying links within walking distance, based on network distance, census block group data, and Equation 4, and
3. Using a modified procedure to estimate total street miles within the buffer area in each block group, census data, and Equation 4 modified for total buffer street miles instead of street miles within walking distance.

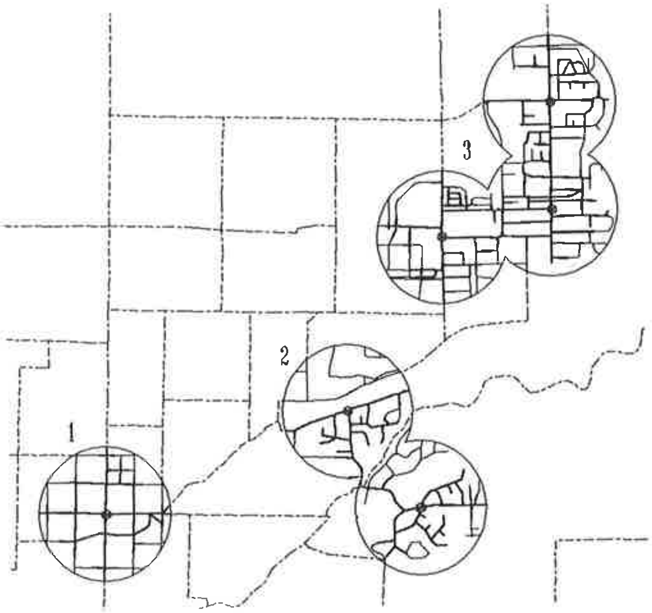


FIGURE 4 Map of example service areas (circles are euclidean buffers, dashed lines are block group boundaries, thick lines are streets within walking distance from the stop).

Table 2 contains information on the total area, total road miles, and total population for each group affected by the analysis. Table 3 shows the numeric data used to determine population in each service area for each of the methods described above. The final population estimates for each service area as summarized from Table 3 are as follows:

Service Area	*Total Population in Service Area (by Method)*		
	1	*2*	*3*
1	1,411	1,036	1,556
2	1,171	605	1,452
3	2,786	3,408	4,826

As shown, results vary based on the methodology and data used. Field verification is required to determine which approach yields the most accurate results. Notice, however, that the proposed service area analysis technique, that is, using walking distances, provides the most conservative estimates of population except for Tract 7, Block Group 3 of Service Area 3. Even in this instance, the result is believed to be more accurate than the buffer estimate. The area in the buffer polygon represents only a small portion of the total area of the block group, hence the low estimate. However, as stated previously, because of the mountains and undeveloped land, the population in this tract is not spread throughout but clustered in the area near the buffer. Because roads are not built in the undeveloped area, the ratio of roads within walking distance to total road mileage is more representative of the spatial organization of the population within the block group. Another noteworthy result occurs in Service Area 2. Streets, and consequently persons, are eliminated from consideration in Service Area 2 when walking distance is used in the analysis.

TABLE 3 Data Tabulated for Each Block Group in Each Service Area for All Methods

SERVICE AREA	TRACT NO.	BLOCK GROUP	BLOCK GROUP AREA WITHIN BUFFER	PERSONS IN SERVICE AREA	ROAD LENGTH IN ALLOCATED AREA	PERSONS IN SERVICE AREA	ROAD LENGTH IN BUFFER AREA	PERSONS IN SERVICE AREA
1	5	3	.0039	28	.04	10	.04	11
1	8	4	.0789	359	.85	205	1.30	316
1	8	5	.0032	27	.04	15	.06	26
1	9	1	.0009	5	.00	0	.00	0
1	9	3	.0619	266	.77	331	1.20	516
1	10	1	.1028	521	1.28	402	1.88	591
1	10	5	.0387	204	.51	74	.66	96
2	7	3	.1238	168	1.22	338	1.61	444
2	7	4	.0763	243	.00	0	1.15	589
2	8	1	.0135	126	.00	0	.24	36
2	9	1	.0002	1	.00	0	.00	0
2	9	1	.0660	404	.89	215	1.19	289
2	11	1	.2914	228	2.41	53	4.22	92
2	11	2	.0003	0	.00	0	.02	1
3	4	2	.0711	2	.55	254	.82	381
3	7	1	.1467	691	1.72	317	3.04	560
3	7	2	.3114	1691	5.15	1421	6.90	1906
3	7	3	.2965	402	3.93	1415	5.49	1979

People in Block Groups 7-4 and 8-1 cannot access these stops without descending a steep brush- and rock-covered cliff. No roads connect these areas, so walking paths cannot be built, and these people should not be considered in the analysis. As expected, the least variation in results occurs for Service Area 1, which most resembles a regularly spaced grid network.

CONCLUSIONS

A model that uses the tools in GIS is developed for delineating transit service areas and estimating population within these service areas. Network distance, as opposed to euclidean distance, is preferred for identifying streets with access to a transit system. The model was tested in Logan, Utah, using two different data sources. Although population estimates found using the procedure developed here seem reasonable, field verification is needed to justify the use of this approach over others. The use of network paths in service area definition has advantages over the common techniques because travel barriers are recognized and unevenly distributed populations are considered. Future areas for research include investigating the accuracy of estimates using a variety of data sources at different spatial resolutions.

REFERENCES

1. K. T. Azar and J. Ferreira. 1991, GIS for Transit Passenger Information Services. *Geographic Information Systems for Transportation Symposium*, Orlando Fla., AASHTO, 1991.
2. *Analyzing Transit Options for Small Urban Communities—Analysis Methods.* UMTA, U.S. Department of Transportation, Jan. 1979.
3. *Transit Corridor Analysis—A Manual Sketch Planning Technique.* UMTA, U.S. Department of Transportation, April 1979.
4. *Mass Transit Management: A Handbook for Small Cities*, 2nd ed., revised, Part III, UMTA, U.S. Department of Transportation, Sept. 1980.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.

Matching Planning Styles and Roles in Transportation to Conditions of Uncertainty and State of Technological Know-How

C. Jotin Khisty

In the process of transportation planning, the planning need be conducted not just within the limits of the rational planning model, but in the context of planning styles that match the type and characteristics of the system, as well as the degree of uncertainty involved in the process. There are three interdependent key variables: (a) the means (technologies); (b) the ends (goals and objectives); and (c) the degree of uncertainty attached to the means and ends. Three levels of the state of the technological know-how—that is, known, developing, and unknown—are considered and connected with three degrees of uncertainty attached to the goals and objectives of a planning agency. This connection results in a matrix of nine cells, each representing a particular type of planning problem, matching a unique planning style or a combination of planning styles. An agenda for action is proposed to aid the planner in selecting an appropriate planning style.

Most countries in the developed and developing world are currently identifying future directions for conducting transportation planning. In the United States, for example, the 2020 Transportation Consensus Program, initiated by AASHTO in collaboration with other organizations, is attempting to determine the nation's surface transportation requirements over the next 30 years (*1*). Recognition of the importance of innovation and new technology in such diverse fields as vehicular navigation, control and location, robotic systems and automation, real-time control, artificial vision, and logistics and communication science results in a great deal of enthusiasm in adopting these technological innovations. However, comparatively little has been discussed as far as planning, evaluating, and implementing these advanced technologies within the framework of current planning organizations. One of the key characteristics of the future is that planners will have to grapple with uncertainty in dealing with not only new, unproven, and uncertain technology, but also with highly complex organizations having fuzzy goals and objectives.

This paper deals with three interdependent key variables—means, ends, and degree of uncertainty. If we consider three levels of the state of technological know-how—that is, known, developing, and unknown—and connect these three levels with three degrees of uncertainty attached to the goals and objectives of a planning agency, we end up with a matrix connecting the means or the state of technological know-how with ends or goals adopted by an agency, resulting in a typology of nine cells, each representing a particular type of planning problem and needing a unique planning style or a combination of planning styles.

To grapple with these uncertainties, urban transportation planning in the future will tend to be conducted in the context beyond the limits of the rational planning model (RPM), a style widely used by transportation engineers and planners over the last 30 years (see Figure 1). This means that to effect plans planners will adopt the following planning styles and roles: incremental, transactive, advocacy, radical, and reflective, among others. Such styles and roles will probably be defined by planners themselves or by society at large. In any case, as the planning process grows more complex, active citizen participation will be essential, if only to ensure that bureaucracies are responsive to the public they serve. This paper describes and discusses the possibility of adapting appropriate planning styles and roles in transportation planning.

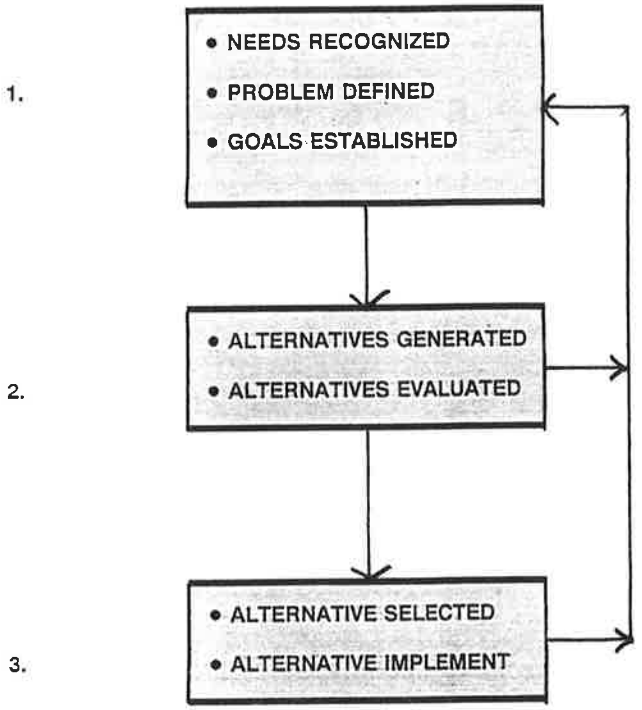


FIGURE 1 RPM.

Department of Civil and Environmental Engineering, Washington State University, Pullman, Wash. 99164.

NATURE OF TRANSPORTATION TECHNOLOGY: PLANNING PROCESS

Profound advances in transportation engineering technology have occurred in the last decade, and these advances continue unabated. Adopting, adapting, extending, and refining these new technologies in current planning and engineering practice is no easy task. Since a great deal of transportation planning is concerned with the medium- and long-range future, it is characterized by a high level of uncertainty. Almost all transportation facilities are capital-intensive and, once built, are not easily amenable to major changes and rebuilding without enormous expenditure and disruption. What do these basic facts mean in terms of planning for the future?

One of the most fundamental tasks is to discover, assess, and address this uncertainty. However, the issues and problems faced by transportation planning agencies appear to defy resolution with the help of the decisions, tools, and traditional planning techniques that have been developed over the past 40 years. These techniques were designed primarily to solve problems in which both ends and means were pretty well known. This inherited professional legacy that attempts to address almost all planning problems by means of the RPM has led to serious dilemmas in recent years (*2–4*). Given the great complexities that face planners and decision makers, particularly now effective planning and evaluation can best be undertaken if planners and decision makers are aware of a variety of planning theories and corresponding planning styles to address a range of uncertainties occurring in the adoption of new technologies.

Transportation systems continually present critical problems long before these systems actually break down or lose their capacity to serve the goals and objectives they were meant to satisfy. Naturally, the conceptualization, research, development, planning, and selection of alternatives required to meet the needs of today and the future must be projected on the basis of several factors, one of the most important of which is time. Although many hundreds of amazing innovations will become available in the future, the questions that continuously arise are, Which ones are proven and well established? Which ones are currently considered to be operationally practical for adoption? Which ones are technologically feasible? and, Which ones are glamorous and exotic yet conceptual and therefore unproven? (*5*). As stated earlier, the high investment in transportation facilities tends to make adoption of technology a slow and incremental process.

In summary, transportation planning problems, as they appear today, are highly complex, conflictual, and interconnected, involving multiple perspectives and assumptions. They can never be solved, as some planners believe, but possibly may be resolved following the "law of the indestructibility of wicked problems."

APPROPRIATE STYLES OF PLANNING: TECHNOLOGICAL KNOW-HOW

Although the RPM is the most favored and common process adopted by transportation planners, this style of planning has led to ineffective results, particularly when means and ends are poorly defined. Many other planning theories and corresponding planning styles can deal with various degrees of uncertainty (*4*). Unfortunately, transportation planners have discussed little about alternative planning styles to deal with uncertainty, and transportation agencies have not tried to adopt planning styles other than the RPM.

	OBJECTIVES		
STATE OF KNOWLEDGE	SINGLE AGREED	MULTIPLE AGREED	NOT AGREED
KNOWN	A	B	C
DEVELOPING	D	E	F
UNKNOWN	G	H	I

FIGURE 2 Matrix connecting state of technological knowledge and goals adopted by agency.

There has always been uncertainty and ambiguity over means and ends in real life. According to Christensen, the key variables that prompt the planner to adopt a particular style of planning are means, ends, and degree of uncertainty (*6*). A matrix connecting the state of technological knowledge (means) and the goals adopted by an agency (ends) was developed on the basis of this idea (Figure 2). Each dimension is further subdivided according to the degree of certainty or uncertainty. Obviously, no real-life problem can be exactly placed in a particular cell created by this matrix for the simple reason that technology itself is value laden and goals can be biased by the technologies adopted. Thus, the matrix consists of nine cells, and there are several grey areas at the boundaries of these cells. A description of each cell follows.

Cell A—Known Technology, Agreed Goals or Single Goal

When a planning agency agrees on what it wants (through public representation), and the technology to achieve this goal exists, then it is possible that a fair amount of certainty prevails and the RPM can be applied through standard, routine procedures. The U.S. military procurement program is an appropriate example (*4*). It uses the latest analytical techniques, resulting in a high degree of formal rationality. Another major application of the RPM is found in municipal planning and engineering work, such as replacing a stop-sign–controlled intersection with a signal. Despite its elegance and simplicity, the RPM has been criticized by various quarters and for myriad reasons. Limits to rationality, methodology, and professional expertise are some of the important ones (*7*).

On the plus side, the RPM, which requires a known technology (means) and an agreed goal (end), is predictable, accountable, efficient, and effective, although it may often be bureaucratic.

Cell B—Known Technology, Multiple Agreed Goals

Cell B represents a pretty common situation faced by planners in which the technology is known and the public has accepted a battery of goals that are not necessarily compatible with one another. The conditions in this case are not as clear-cut as those in Cell A but certainly not as ambiguous as when the goals are not agreed on (Cell C). What one observes therefore is a case of "bounded" rationality, as suggested by Simon (*8*). He emphasized the fact that decision makers can never be completely rational in the sense of having total knowledge of a situation and the alternatives available to them, nor do they have the time, resources, and intelligence to sort out the consequences of the multiple objectives adopted by the public. Under the circumstances, the best one could do was not to seek optimization, but simply to manage to "satisfy" major organizational values; to satisfy was all that one could reasonably expect within the limitations prescribed. In other words the course of action would be one that appeared to be good enough; the first and perhaps not the least important action would be to apply the test of common sense. Simon was convinced that the choice among alternatives is simplified if we replace the goal of maximizing (or optimizing) with the goal of satisfying (*9*).

Cell C—Known Technology, No Agreed Goals

The Cell C situation, frequently encountered, is one in which proven methods are available but the public has not agreed on the goals. For example, a city may have just the right population density for introducing a light-rail system with the technology for adopting light rail, but the city may not have goals, thus precluding the adoption of this proposed light-rail system over an extended portion of the city.

One way out of such a situation is through a bargaining process. Consensus building can be established through communication, particularly for those communities that seem to hold out against continuous rail lines through their turf. Bargaining in any case is difficult because it goes against the very principles of bureaucracy. Consensus building can be crucially important, partly guided from above by controlling societal groups and partly voluntaristic (*10*). This process could also be considered as "interwoven planning" embracing the twin processes of consensus formation and societal guidance. Here the accommodation of multiple preferences through bargaining seems to be the key issue (*4*).

Cells D and E—Developing Technology, Simple or Multiple Goals

Cells D and E represent the case in which the technology is developing and the public has agreed on single or multiple goals. The general approach to these two boxes would call for the planner to work through trial and error, trying something, receiving feedback, and making modifications and adjustments until the situation became acceptable.

Incremental planning as suggested by Lindblom (*11,12*) attempts to achieve this kind of adjustment, adopting decision-making strategies to the limited cognitive capacities of decision makers. Some of the major characteristics of incrementalism are that only a few alternatives are considered and that only a few consequences are considered and evaluated. Also changes are made in increments, considering that we are trying to shoot a moving target. Lindblom's "muddling through" tactic is most apt.

It would appear at first sight that incrementalism and satisfying are almost the same. However, there is a subtle difference. Although both operate within bounded rationality, incrementalism need not stop short at finding a satisfying solution—it may attempt at optimization within its bounded context. This difference is all the more possible when goals are well defined or when only a single goal is being considered (*9*). The concept of "mixed scanning" as proposed by Etzioni is also appropriate because it synthesizes large (synoptic) and small (incremental) decisions into a single framework. Here, an in-depth investigation of selected problems is done that merits special attention (*10*).

Cell F—Developing Technology, No Agreed Goals

Cell F, although similar in some respects to Cell C, emphasizes uncertainty to a much larger extent because the technology is in the developmental stage. Hence, planners take their clients' points of view and set forth proposals in their clients' interests. Promoting more active and diverse citizen participation could also be considered. In short, this is a case in which advocacy planning could be considered. One consequence of advocacy planning is to shift social policy formulation into the open. It also calls for developing plural plans, involving a host of interest groups. The ideal role of the advocate is to help the client organization clarify its ideas and goals and, thus, plan correctly (*13*).

Cell G—Unknown Technology, Agreed Goals

When there is a commitment on the part of the government to attack a pressing problem but no proven technology is available, experts may try to propose plausible solutions. The immediate problem is to search for the missing knowledge and come up with a technology. A good example is one concerned with mitigating traffic congestion through high technology. Obviously, innovation is needed and hence anything routine or bureaucratic has no place in such a situation. In this case the focus is on methods to generate the necessary knowledge to understand what is truly needed. Experimentation is therefore the name of the game. Such experimentation may lead to innovation. One should be cautious in applying this experimentation or innovation by applying the fuzzy technology in a pilot project, which in itself is not extensive but large enough for figuring out the sensitivity of the various parameters entering the system. Intelligent vehicle navigation systems and artificial vision and logistics are examples of new transportation technology that is available for experimentation and limited use.

Cell H—Unknown Technology, Multiple Goals

The Cell H box resembles that of Cell G, except that its conditions of uncertainty are much more confused and unstable—although certainly not as bad as those in Cell I. To some, this box may represent a case of near chaos, requiring radical means of solving the problem. Indeed, radical planners believe that education should come from the everyday life of local communities—a form of social learning or reflection in action (RIA) popularized by Schön and his colleagues. RIA is an improvisational problem-solving, interactive experimentation undertaken on the spot, using local knowledge. Schön's emphasis is to teach local participants the esoteric skills of "learning to learn." The learning-to-learn paradigm may even result in discovery and innovation and has been put into practice in recent years with much success (*14–16*).

Cell I—Unknown Technology, No Agreed Goals

The Cell I box represents situations that are not uncommon. Goals can be nonexistent or at best nebulous, combined with "solutions" that need a technology that is still unrefined. This situation can be all the more aggravating when an organization, faced with such a dilemma, does not have a leader. In such cases, conditions of uncertainty over both means and ends result invariably in chaos. Rittel and Weber have addressed this and other situations that convincingly explain some of the characteristics of "wicked" problems (*2*).

It is not clear how one could extricate an organization from this chaotic situation. In some cases it may be possible to sift through a bundle of vague goals, a process suggested by Friedman under the heading of "social learning" (*4*). Beginning and ending with action, social learning is a complex, time-dependent process that involves social practice.

DISCUSSION OF RESULTS

The nine cells were adopted for sheer convenience and should not be construed as watertight compartments. Combining planning styles is the answer. Uncertainty appears to be the dominant characteristic facing most planning organizations now and will prevail perhaps more intensively in the future. Depending on the chemistry of the situation, a range of planning-solving approaches and styles needs to be adopted by planners who recognize the difficulty of practicing the rational planning methodology so common over the past decades. The various approaches suggested in the last section envision the planner as neither the pure technician nor the value-free implementor of others' decisions. In fact, the complexity of emerging problems puts transportation planners at the crossroads of engineering, planning, and sociopolitics, and much will therefore depend on their ability to integrate the principles of planning theory in day-to-day practice (*16*).

Radford has suggested four broad specifications for dealing with complex decision making (*17*):

1. The procedure should include the most appropriate characteristics of existing approaches developed in the analytical, behavioral, and political sciences.
2. The process should be readily comprehended by the public.
3. The process should be sufficiently broad based and flexible for application in a wide range of problem situations.
4. The process should be one that can be introduced unobtrusively into an agency, with a minimum amount of disruption.

In recent years planners have made deliberate effort toward development of a "contingency" role in practicing their profession. They have changed their strategies to suit the situation and have realized in no uncertain terms that political concerns generally take precedence over technical concerns for all but the most simple situations (*18*).

Schön's RIA concepts have also gained recognition, particularly in those fuzzy areas of planning in which the situation appears chaotic. Ideas for moving from the hard, high ground of the theorist down to the dark, boggy swamp of the practitioner are gaining momentum (*19*).

In recent years, some radical changes have been initiated by notable planners across the world who are trying to manage problems similar to those that are occurring in transportation planning. Of particular interest is the work of Checkland, whose published literature focuses on the approach to plural rationality through "soft" systems methodology. As opposed to "hard" systems methodology (similar in many respects to RPM), which is used for tackling real-world problems in which an objective or end to be achieved can be taken as given, soft systems methodology, based on a phenomenological stance, tackles real-world problems in which ends that are known to be desirable cannot be taken as given. Checkland's methods have been applied with much success in scores of planning and management situations around the world (*20,21*).

The involvement of citizens in governing society is the subject of history itself. During the past 40 years, the level and effectiveness of citizen participation in the planning process have most often been stimulated and enhanced when existing social problems are complicated and the level of uncertainty is high. Such enhancement is also observed when citizens are skeptical of official solutions. It is therefore anticipated that as the planning process grows more complex, active citizen participation and control will become commonplace, thus ensuring that bureaucracies are responsive to the public they serve. Ultimately, all plans are really political statements; indeed, all attempts to implement them are political acts (*22*).

In closing, a word should be said about evaluating methods. To achieve a high degree of acceptance by the many participants in the planning process, it is suggested that this evaluation process be divided into two basic stages:

1. Modeling of scenarios under various situations, and
2. Assessment of effects and consequences, both tangible and intangible, for a variety of factors using cost-benefit, cost-effectiveness, or utility analysis.

This evaluation is the most important part of the decision process, but it is a topic that should be addressed in a separate paper. Suffice it to say that the multidimensional implications embedded in the decision-making process will be reflected in the selection of the objective functions (*23*).

AGENDA FOR AGENCY BUILDING

If there is one nagging theme that haunts us through our discussion on planning, it is that of rationality. The bottom line is that in attempting to be rational, do we make things better or worse? The complexities of planning forces planners to view problem solving as a process of social interaction, trial and error, successive approximation, and social learning. Such an approach induces planning institutions and agencies to move away from the Weberian model as described by Friedman (*4*).

Under the circumstances it is very likely that transportation planning agencies will be obligated to reconstitute and restructure their organizations to adopt some or all of the following characteristics:

1. *Technical capability:* the ability of agencies to deliver technical services and sift through technological know-how for guiding society regarding technological innovations and possible adoption.
2. *Normative commitment:* the ability of agencies to internalize innovative ideas and practices for the betterment of society.
3. *Environmental image:* the ability of agencies to attain favorable recognition from society, on the condition that they respect environmental concerns when adopting innovative ideas.
4. *Equity concerns:* the ability of agencies to effectively address questions of equity at the micro and macro levels. Distributive justice is as important as the adoption of new technology.
5. *Citizen participation:* the ability to engage the participation of system members in contributing to the collective knowledge of the system. The more complex the problem, the greater is the need for localized solutions and value innovations—both of which call for broadly based citizen involvement in the decision process.
6. *Accountability:* the ability of agencies to recognize that, under conditions of uncertainty, errors and mistakes are not only likely—they are to be expected. The concept of policy making as social experimentation requires that a project be planned and implemented in such a way that errors and mistakes can be uncovered as the project proceeds. It can then be redesigned and revised incrementally. This point is highly significant because planning agencies are notorious for suppressing mistakes and errors, and they have been known to punish managers—sometimes wrongly. Fear of making mistakes discourages correction, redesign, and redirection and inhibits creativity, innovation, flexibility, and experimentation—the very core of successful planning and implementation.

CONCLUSIONS

Traditionally, transportation planning has worked under the assumption that both the technology to be adopted and the goals and objectives set forth by an agency are well known, in which case the RPM is well suited for application. Although this assumption is theoretically true, the real world does not operate so tidily. With the tremendous strides made in almost every area of transportation technology, such as electronic guidance systems, automatic vehicle control, and communication science, transportation planners and decision makers must deal with technology that is constantly in transition. Coupled with this problem is the one connected with goal formation and adoption: uncertainty in both dimensions of the matrix—means and ends—is difficult to comprehend. However, this is a fact that will become more and more prevalent in more and more transportation planning agencies. Planners and decision makers must face this uncertainty by tailoring the planning process according to the degree of uncertainty embedded in the technological knowledge base and the goals adopted.

REFERENCES

1. *A Look Ahead: Year 2020.* TRB, National Research Council, Washington, D.C., 1986.
2. H. Rittel and M. Webber. Dilemmas in a General Theory of Planning. *Policy Sciences,* No. 4, 1973, 155–169.
3. A. A. Dzurik and R. L. Feldhaus. Evolution of Planning Theory and Practice: Engineering Implications. *Journal of Urban Planning, ASCE,* No. 12, 1986, pp. 37–45.
4. J. Friedman. *Planning in the Public Domain.* Princeton University Press, Princeton, N.J., 1987.
5. C. J. Khisty. *Transportation Engineering: An Introduction.* Prentice-Hall, Englewood Cliffs, N.J., 1990.
6. K. S. Christensen. Coping with Uncertainty in Planning. *APA Journal,* No. 1, 1985, pp. 63–73.
7. S. L. Hart. Steering the Path Between Ambiguity and Overload: Planning as Strategic Social Process. In *Interdisciplinary Planning Center for Urban Policy Research,* (M. J. Dluhy and K. Chen, eds.), Center for Urban Policy Research, Rutgers, N.J., 1986.
8. H. A. Simon. *Administrative Behavior*, 3rd ed. Free Press, New York, N.Y., 1976.
9. G. Chadwick. *A Systems View of Planning.* Pergamon Press, Oxford, England, 1971.
10. A. Etzioni. *The Active Society.* Free Press, New York, N.Y., 1968.
11. C. E. Lindblom. The Science of Muddling Through. *Public Administration Review*, No. 19, 1959, pp. 79–88.
12. C. E. Lindblom. Still Muddling, Not Yet Through. *Public Administration Review*, No. 39, 1979, pp. 517–526.
13. P. Davidoff. Advocacy and Pluralism in Planning. *AIP Journal*, No. 2, 1965, pp. 331–338.
14. D. Schön. *Beyond the Stable State.* W. W. Norton, New York, N.Y., 1971.
15. D. Schön. Some of What a Planner Knows. *APA Journal*, No. 2, 1982, pp. 351–364.
16. D. Schön. *The Reflective Practitioner.* Basic Books, New York, N.Y., 1983.
17. K. Radford. *Complex Decision Problems: An Integrated Strategy for Their Resolution.* Basic Books, New York, N.Y., 1977.
18. J. Bryson and A. Delbecq. A Contingency Approach to Strategy and Tactics in Project Planning. *AIP Journal*, No. 2, 1973, pp. 179–187.
19. C. J. Khisty and L. L. Khisty. Reflection in Mathematical Problem-Solving and Design. Presented at 15th Conference on Psychology of Mathematics Education, Assissi, Italy, 1991.
20. P. Checkland. *Systems Thinking.* John Wiley, Chichester, England, 1981.
21. P. Checkland. *The Approach to Plural Rationality Through Soft Systems Methodology.* Lecture Notes in Economics and Mathematical Systems 248. Springer-Verlag, Berlin, Germany, 1985.
22. D. A. Rondinelli. *Development Projects as Policy Experiments.* Methuen, London, England, 1983.
23. H. Keller. Comprehensive Evaluation of Advanced Technologies in Transportation Engineering. *Proc., 1st International Conference on Applications of Advanced Technologies in Transportation Engineering* (C. Hendrickson and K. Sinha, eds.), San Diego, Calif., 1989.

Publication of this paper sponsored by Committee on Statewide Multimodal Transportation Planning.

Phoenix Commercial Vehicle Survey and Travel Models

EARL R. RUITER

The primary objectives of the Phoenix urban truck travel model project, which was conducted for the Arizona Transportation Research Center, Arizona Department of Transportation, were to conduct a travel survey of commercial vehicles operating within the Phoenix metropolitan area and to use the data collected in this survey to develop commercial vehicle trip generation, distribution, and traffic assignment models. The models are designed to be incorporated into the Urban Transportation Planning System-based travel model system maintained by the Maricopa Association of Governments Transportation and Planning Office, which predicts highway and transit system usage throughout the Phoenix metropolitan area. The urban truck travel model project, including the methods used to collect commercial vehicle travel data, the types of information provided by the survey, and model development using the survey data, is summarized, and the issue of the transferability of the results of this project to other urban areas is discussed. Thus, the commercial vehicle travel patterns identified in Phoenix, and the travel forecasting models based on these patterns, may also be useful in other urban areas that are similar to Phoenix with respect to their mix of commercial and industrial activities and their history of growth and development into major metropolitan regions.

This paper presents the results of the Phoenix urban truck travel model project, which was conducted for the Arizona Transportation Research Center, Arizona Department of Transportation. The primary objectives of the project were to conduct a travel survey of commercial vehicles operating within the Phoenix metropolitan area and to use the data collected in the survey to develop commercial vehicle trip generation, distribution, and traffic assignment models. A full discussion of the project details and results is given elsewhere (*1*).

SURVEY METHODS

The Phoenix commercial vehicle survey provides detailed information on 3,402 trips made by 606 commercial vehicles registered in Maricopa County or used by the U.S. Postal Service (USPS) in the county. Each surveyed trip has its origin and its destination within the Maricopa Association of Governments (MAG) transportation study area. The survey does not include any commercial vehicles registered outside Maricopa County. In the Phoenix travel forecasting system, most of the trips made by these vehicles are included in external commercial vehicle trip tables. The purpose of this survey was to develop new models for internal commercial vehicle trips only.

Two sources of data were used to determine the total population of commercial vehicles to be sampled in the survey. The first was a computerized file of approximately 157,000 commercial vehicles registered in Maricopa County in 1989. This file, obtained from the Department of Motor Vehicles (DMV), contains truck type identifiers and owners' names and addresses. The second was a listing, by garaging location in Maricopa County, of the 2,300 vehicles owned by the USPS but not registered in Arizona.

The data collection procedure used for vehicles selected from the DMV file was a combined telephone and mail method. This approach was adopted after low response rates were obtained in an initial pretest that relied entirely on a mailout-mailback method. The following procedure was used:

- *Telephone contact:* Vehicle owners for which telephone numbers could be obtained were called, initial screening questions were asked, and cooperation was requested in the mail portion of the survey.
- *Mail contact:* A mail-back questionnaire including a 1-day trip diary was mailed to those who agreed to participate in the survey and to selected owners who could not be contacted by telephone.

For USPS vehicles, with the assistance of the manager of fleet operations for the Phoenix postal district, vehicles were sampled by weight class and garaging location. Then, for the sampled vehicles, USPS forms detailing daily itineraries were obtained and translated into the format of the trip diary used for vehicles obtained from the DMV files.

Data Collection Forms

During the telephone portion of the survey, a script was used to introduce vehicle owners to the survey, elicit their cooperation, and obtain the following information on their registered vehicle, which had been chosen to be included in the survey:

- For vehicles leased by another firm or individual, name and address of the lessor;
- For vehicles not used on a specified survey day, the reason for no usage (no work, vehicle not operational, or other) and the registration number for a replacement vehicle, if any; and
- Person to whom the mailout questionnaire should be sent.

The mailout questionnaire used for the truck survey was patterned after that used previously by the Chicago Area

---

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

Transportation Study for a major commercial vehicle survey in the Chicago area (*2*). It was designed to obtain the following data for each surveyed commercial vehicle:

- Starting and ending addresses on survey day;
- Vehicle type, based on number of axles and body style;
- Estimated gross weight;
- Vehicle usage for transportation between home and work and for work-related purposes; and
- Total number of one-way trips on the survey day.

The DMV file also provides data items that were used along with the survey data. These items include the ZIP code of the owner and the registered vehicle weight.

In addition, the travel diary requests the following information on the first 10 one-way trips made by each vehicle on the selected survey day:

- Start and stop times;
- Stop odometer readings;
- Name and address of stop;
- Driver and vehicle activity at stop;
- Land use at stop; and
- Vehicle type and total axles during trip (to determine trailer pick-up and drop-off locations.)

Sample Design

DMV-Registered Vehicles

Stratified samples were selected from the DMV registration file and from the list of USPS vehicles. In both cases, the stratification was on the basis of vehicle weight. The initial sampling rate for light vehicles (under 8,000 lb) was 1 in 79. This rate increased to one in four for the heaviest vehicle class (over 64,000 lb). The total sample was designed to provide 4,000 vehicles distributed by vehicle weight to include light vehicles as 40 percent and each of the remaining three weight categories as 20 percent of the total sample.

By sorting the entire DMV file by ZIP code before sample selection, subsamples were obtained in which all geographic areas are represented in proportion to their vehicle weight category–specific distribution in the total population.

USPS Vehicles

USPS vehicles to be included in the survey were selected by the postal service's manager of fleet operations. All USPS vehicles in Phoenix fall into the two lightest-weight categories used in this project (2,180 in the under-8,000-lb category and 101 in the 8,000- to 28,000-lb category). The selection process provided 1 in 40 postal vehicles in the light category and 1 in 10 in the next heavier category.

Data Collection

DMV-Registered Vehicle Survey

After pretest and pilot surveys, which were essential in refining the data collection strategy, combined telephone and mail procedures were used to conduct the main survey. Tables 1 and 2 provide information on the levels of survey responses obtained in the main portion of the truck survey. After telephone contact, 37 percent of the total subsample agreed to participate in the survey. Most of those who did not agree to participate either could not be contacted or owned vehicles that did not qualify for the survey.

TABLE 1 Survey Results, Telephone Portion

Category	Number		Percentage	
Total subsample		3855		100
No telephone contact possible		1393		36
Vehicles not qualified		538		14
No work/no alternative vehicle	214		6	
No information available	103		3	
Non-commercial vehicle	169		4	
Lessee name not available	35		1	
Out of state owner	17		0	
No agreement to participate		498		13
Agreement to participate – surveys mailed		1426		37

TABLE 2 Survey Results, Mail Portion

Category	Number		Percentage	
Total surveys mailed to those agreeing to participate		1426		100
Not returned or not completed		1006		71
Responses		420		29
Trips made	358		25	
No trips made	62		4	
Total surveys mailed to other vehicle owners		300		100
Not returned or not completed		195		65
Responses		105		35
Trips made	60		20	
No trips made	45		15	
Total surveys mailed (1.1 + 2.1)		1726		100
Not returned or not completed (1.2 +2.2)		1201		70
Responses (1.3 + 2.3)		525		30
Trips made	418		24	
No trips made	107		6	

The results for the mail portion of the main survey (Table 2) show that the response rate was 29 percent for those who agreed to participate. The overall response rate to the mailed questionnaires was 30 percent. Before geocoding, the total number of survey responses was 720: 195 from the pilot survey and 525 from the main survey. Of these 720 responses, 527 (73 percent) represent vehicles that made commercial trips on the survey day.

USPS Vehicle Survey

Data collection for USPS vehicles was much simpler than for the vehicles in the DMV file because the cooperation of the manager of fleet operations for the Phoenix postal district was obtained before subsample selection. Travel diary data for the 62 selected vehicles were obtained from existing USPS forms and transferred directly to the vehicle and trip data sets used in the project.

Data Coding and Factoring

Vehicle Factors

The 1989 DMV file used to obtain a survey sample contained 156,645 commercial vehicle registration records, and the Phoenix postal district reported a total of 2,281 vehicles. The breakdown by vehicle weight class is shown in Table 3.

The telephone portion of the survey revealed that only 75.7 percent of the vehicle owners contacted via the DMV data set reported that their vehicles were available for use for commercial purposes within the Phoenix metropolitan region on a typical travel weekday. This fraction was used to obtain an initial estimate of the total population of qualified vehicles, subject to adjustment in later stages of the project (see the section on calibration). Thus, the survey data were expanded to represent 118,645 DMV vehicles in operation and 2,281 USPS vehicles.

Expansion factors for the DMV vehicles were developed separately by vehicle weight class and by owner's ZIP code as contained in the DMV file. USPS vehicles were weighted using a similar strategy, expanding to match the USPS totals by weight/postal garaging location category and to match totals by weight class. Overall, the survey represents a 0.5 percent sample of all commercial vehicles based in Maricopa County; the average vehicle expansion factor is 203.7.

Trip Factors

Because the commercial vehicle drivers responding to the survey were asked to report individual information for a maximum of 10 trips on their survey day, additional truck-specific expansion factors were required to account for each truck's unreported trips. These factors account for the reporting of trips to and from the overnight garaging location, which generally was available for each truck, and for the partial re-

TABLE 3 Summaries of Survey Responses

	Vehicle Weight (lbs)				
Vehicle Type	0-8,000	8-28,000	28-64,000	64,000+	Total
Total Population					
DMV Vehicles	127,427	19,440	4,830	4,948	156,645
Postal Service Vehicles	2,180	101	0	0	2,281
All Commercial Vehicles	**129,607**	**19,541**	**4,830**	**4,948**	**158,926**
Percentage of Total	81.6%	12.3%	3.0%	3.1%	100.0%
Daily Vehicle Trips					
All Commercial Vehicle Trips	**726,889**	**188,545**	**32,659**	**19,742**	**967,835**
Percentage of Total	75.1%	19.5%	3.4%	2.0%	100.0%
Average Trips per Vehicle	5.6	9.6	6.8	4.0	6.1
Daily Vehicle Mileage					
Average VMT per Vehicle	79.0	56.2	74.0	156.8	78.5
Average Miles per Trip	11.0	4.7	9.2	33.4	10.2

porting of all other trips by trucks making more than 10 trips per day. The factors also correct for unusable trip records.

The average total trip factor is 284.5, implying an overall trip sampling rate of 0.35 percent. When the trip factors are applied to the 3,402 usable reported commercial vehicle trips, an estimate of 967,835 total daily trips is obtained. The numbers of trips by weight class are provided in Table 3. Both the largest and the smallest weight classes have proportionately fewer trips than they have commercial vehicles. This is true because the trucks in the middle weight categories (8,000 to 64,000 lb) reported making more trips per day than those in the smallest and largest weight categories.

Survey Results

The weighted survey records were summarized statistically to provide information on a number of commercial vehicle and truck travel characteristics in the Phoenix region. All data summaries were obtained by vehicle weight category. Information on the following vehicle characteristics was obtained:

- Average vehicle weights per surveyed vehicle;
- Vehicle types;
- Vehicle usage for home-to-work travel;
- Vehicles not used on a typical weekday;
- Time of first daily trip;
- Vehicle trips per day; and
- Vehicle mileage per day.

Statistical summaries of this information are provided in the project's final report (*1*). For the two final data categories (vehicle trips and vehicle mileage per day), the totals reported by vehicle weight category are included in Table 3. The vehicle mileage results, obtained from measurements provided by odometer readings at the start and end of the day, are inversely related to the number of trips made, also summarized in Table 3. This apparently anomalous result is explained by the differences in average miles per trip by vehicle category. Vehicles in the 8- to 28,000-lb category make many short trips, typically for such activities as refuse pickup and package delivery. Vehicles in the heaviest category make a few long trips and in so doing generate many more vehicle miles per day than the lighter vehicles generated. As in the case of trips per day, the remaining vehicle classes exhibit average vehicle mileage and trip lengths similar to the overall averages. These averages are 78.5 mi/day and 10.2 mi/trip.

Information was also obtained on the following characteristic of commercial vehicle trips:

- Average vehicle weights per commercial vehicle trip;
- Trips by vehicle type;
- Time-of-day distributions;
- Activities at trip ends;
- Land uses at trip ends;
- Activity and land use linkages at trip ends;
- Stop locations (on- or off-street);
- Trips by time duration; and
- Trips by travel distance.

As in the case of vehicle characteristics, these results are reported in detail in the final report (*1*).

URBAN COMMERCIAL VEHICLE TRAVEL MODELS

Trip Generation

Because the commercial vehicle survey includes information on land uses at trip ends and the MAG zonal data include the number of residents and employment by land use category, it was possible to determine trip rates by land use category. As defined for use in this project, these rates have the following form:

truck trip rate for land use category *i*

$$= \frac{\text{total study area trips to land use category } i}{\text{total study area employment at land use category } i}$$

The five categories of land use available in the MAG zonal data, and the corresponding categories used in the truck survey, are retail, industrial, public, office, and other.

An additional land use category—residential land—was included in the survey. For trips to and from this category, the trip rate was defined as:

$$\frac{\text{total study area trips to residential land}}{\text{total study area households}}$$

Each of the trip rates defined includes a minor specification error, because the reported trips include those made to construction sites but the land use data do not identify these sites explicitly. However, since construction activity is also not predicted explicitly for future years and all present and future commercial vehicle trips must be accounted for, this misspecification was required as part of a long-range forecasting model.

The trip generation modeling results obtained for the two heaviest vehicle weight classes combined indicated that, from the standpoint of predicting trip generation, this combination of two weight categories was preferable to keeping these weight categories separate. A final decision on this issue was not made, however, until the average travel times, based on MAG's network data, were determined for each category and preliminary distribution model results were obtained. Thus, models based on both classification strategies were initially developed and used at the beginning of the trip distribution modeling task. Because the combination strategy proved to be preferable for trip generation and trip distribution modeling, only the combined model is shown in Table 4.

The final estimated trip generation models are presented in Table 4, which contains the coefficients associated with each independent variable for each model. The following equation, for commercial vehicles less than 8,000 lb, illustrates how the rates shown in Table 4 are used in the trip generation models:

$$\begin{aligned} \text{TRIPS}_i = {} & 0.15433 * \text{TOTHH}_i + 0.59091 * \text{RETEMP}_i \\ & + 0.64087 * \text{INDEMP}_i + 0.29491 * \text{PUBEMP}_i \\ & + 0.30925 * \text{OFFEMP}_i + 0.76348 * \text{OTHEMP}_i \\ & + 0.04004 * \text{RESHH}_i \end{aligned}$$

TABLE 4 Trip Generation Models

Independent Variable	Vehicle Weight (lbs) 0-8,000	8-28,000	28,000+
Total households	0.15433[a]	0.06859	0.01260
Retail employment	0.59091	0.13253	0.03685
Industrial employment	0.64087	0.09972	0.04991
Public employment	0.29491	0.00596	0.02398
Office employment	0.30925	0.02119	0.00320
Other employment	0.76348	0.10567	0.07527
Resident households	0.04004	-	0.00288
Group quarter households	-	7.52348	-
Total area (acres * 100)	-	-	0.00365
Vehicles	-	-	0.00062

[a] Commercial vehicle one-way trips per one unit of the independent variable.

Note: The coefficients shown here do not reflect the results of the calibration/assignment phase of the project. See the Calibration/Assignment section for a discussion of the final regional factor used to estimate total commercial vehicle trip generation.

where

TRIPS_i = total average weekday commercial vehicle trips for vehicles less than 8,000 lb originating in (and the total destined for) zone or district i,
TOTHH_i = total households in zone or district i,
RETEMP_i = total retail employees in zone or district i,
INDEMP_i = total industrial employees in zone or district i,
PUBEMP_i = total public employees in zone or district i,
OFFEMP_i = total office employees in zone or district i,
OTHEMP_i = total other employees in zone or district i, and
RESHH_i = total resident (nongroup quarters, nontemporary, and nonseasonal) households in zone or district i.

Trip Distribution

Network-Based Average Trip Times

Zonal level trip tables for each of the three final vehicle weight categories were developed using the weighted truck travel survey data. A table of zone-to-zone off-peak highway skimmed travel times for Phoenix's existing highway system was obtained from MAG. This table was combined with the three truck trip tables to obtain travel time distributions and averages by vehicle class. The average times by vehicle weight category are much less than those obtained from vehicles' reported stopping times per trip, reflecting the elimination of stopped time from the averaging process and the differences between times based on minimum paths in a highway network and times reported by vehicle drivers.

Model Structure

For consistency with the MAG Transportation and Planning Office's (MAGTPO's) person trip models and with the state of the modeling practice in many U.S. metropolitan areas, the standard gravity-type model structure was selected for commercial vehicle trip distribution modeling in the Phoenix metropolitan area. This structure is one in which trips for a particular category (in this case, commercial vehicle trips by weight category) between a production zone i and an attraction zone j are directly proportional to the total number of trip productions in zone i, attractions in zone j, an attractiveness factor based on the impedance (in this case, off-peak highway travel time) from i to j (although this factor decreases for larger values of impedance as an attractiveness measure should, it is termed a friction factor in the transportation literature; to avoid confusion, the standard terminology is used here) and, optionally, an adjustment factor (K-factor) that varies by origin and destination superdistrict (no K-factors are included in these commercial vehicle models). Because a share formulation is used, the number of trips between zones i and j is inversely proportional to the numbers of attractions in all other zones, to the friction factors from i to each of these zones, and, optionally, to *K*-factors from zone *i* to each of these zones. The friction factors are normally estimated iteratively for each trip category using a gravity model calibration program that attempts to match the observed impedance distributions.

Gravity Model Calibration

Comparisons of the predicted and observed trip time distributions from initial calibration runs for all three vehicle weight categories revealed significant variations, even when average trip times were very nearly matched. Furthermore, increases in the number of calibration iterations did not improve these initial results. A careful review of the calibration algorithm used in the TRANPLAN package revealed that, as in other applications involving trip length distributions with higher-than-typical numbers of very short and very long trips, its friction factor smoothing process was apparently responsible for these results; by fitting a smooth log-linear function to the adjusted friction factors, the required adjustments were being canceled out on each iteration.

This problem with the available gravity model calibration program was overcome by switching to an iterative application of the TRANPLAN gravity model calculation program, supplemented by a spreadsheet to help make manual friction factor adjustments. As in the TRANPLAN calibration process before smoothing, the manual adjustments involved reestimating each friction using a correction term equal to the desired fraction of trips in a travel time range divided by the previously estimated fraction in this range. Rather than using constant travel time ranges of 1 min, the travel time ranges were selected to ensure that the resulting friction factors would always decrease as travel times increase. This procedure converged after just three to five iterations (beginning with the results of a five-iteration run of the calibration program) to models with acceptable travel time averages and distributions. Table 5 provides comparisons of the observed and predicted averages for these final models.

TABLE 5 Observed and Predicted Average Trip Times for Final Distribution Models

Vehicle Weight (lbs)	Average Trip Times (minutes)		Percentage Error
	Observed	Predicted	
0-8,000	16.4	16.1	-2.0
8-28,000	11.9	12.2	+2.6
28,000+	18.8	18.8	+0.2

Calibration and Traffic Assignments

Prior Internal Commercial Vehicle Travel Forecasting

Before the development of the commercial vehicle models, MAG's travel modeling process, as updated in 1988, included a trip generation model for a single category of internal truck trips representing all weight classes and one gravity model. The trip generation model was borrowed from the forecasting system developed for the Detroit metropolitan area by the Southeast Michigan Council of Governments transportation staff. The gravity model was developed using Phoenix data collected more than 15 years ago. The internal commercial vehicle trips estimated by this gravity model were added to all other vehicle trips, including external truck trips that are estimated on the basis of a recent external vehicle trip survey and assigned to the Phoenix highway network using a network equilibrium procedure.

During the 1988 model updating process, these internal truck generation and distribution models were considered as temporary "place holders," to be replaced by the models developed in this project. However, they were also used to provide the final adjustments required to calibrate the complete vehicle trip modeling system to match current vehicle miles of travel (VMT) data for the entire Phoenix metropolitan region. Thus, a regionwide factor of 1.38 was applied to the results of the old trip generation and distribution models as these trips were added to all other vehicle trips before the traffic assignment step. The overall adjustment of 38 percent provided by this factor represents the total effect of each of the following components of changes in internal truck travel:

- The expansion of truck vehicle trips to the equivalent number of two-axle counts, as measured by the automatic traffic recorders used to estimate the total VMT in the Phoenix area.
- The adjustment of internal truck travel estimated with the current models used in Phoenix to represent the actual internal truck travel in the Phoenix area.
- The expansion of Phoenix internal truck travel to compensate for any underreporting in the latest Phoenix travel survey or underestimation in the updated nontruck Phoenix models.

Only the first component, accounting for internal truck axles rather than trucks, can be determined accurately using the available data. The other two factors cannot be isolated to determine the relative importance of the adjustments caused by model transfer and those caused by underreporting of nontruck travel.

Adjustments of New Truck Models

It was necessary to incorporate the first and third adjustment components in the calibration process for the new models. In addition, although no adjustments are required because of model transfer, an adjustment was required to account for the fact that only the trips made by commercial vehicles registered in Maricopa County (the Phoenix metropolitan study area) are included in the models developed in this project. Because these models are integrated into the MAG forecasting system, they must be adjusted to represent all internal commercial vehicle trips, including those made in the study area by vehicles registered outside Maricopa County. As in the case of the current models, the net effect of the factors related to vehicle registration location and underreporting can be determined, but the separate factors making up this total adjustment cannot be isolated. Thus, the calibration process for the new models consists of two steps:

- Expanding the commercial vehicle trips by weight class to account for the average number of axles per vehicle in each class.
- Expanding total commercial vehicle trips so that total estimated and observed VMT in the Phoenix region are equal. This expansion factor represents the net effect of internal trips by all commercial vehicles versus those by vehicles registered in Maricopa County, and of any underreporting or underestimation in any of the Phoenix models that affect the number of truck and nontruck vehicle trips.

When the average travel time statistics by weight class and the overall average speed for the entire expanded Phoenix travel survey are applied to the survey's total commercial vehicle equivalent two-axle trips, an estimate of 7.182 million vehicle mi is obtained. This value compares with 11.659 million VMT, the difference between the total observed two-axle VMT in the Phoenix area and the total estimated by all current Phoenix models, except the temporary internal truck trip model. Thus, the combined registration-underreporting factor, the ratio of the latter number to the former, is 1.623.

Model Implementation

The adjustment factors described were combined with the trip generation models listed in Table 4, the trip distribution models described in the previous section, and MAG's current vehicle trip assignment procedure to fully implement the new commercial vehicle models as an integral part of the total Phoenix travel forecasting system.

Model Transferability

Because travel patterns are often found to vary largely from one urban area to another, the safest means of using the results of this project in another city would be to repeat the travel survey and model development tasks using the procedures found to be most effective in this project. The information requirements of this strategy would be within the usual capabilities of local and regional agencies responsible for transportation planning. These requirements include

- A file from the state vehicle registration agency of all commercial vehicles registered to owners in the planning agency's study area;
- The ability to geocode street addresses to traffic analysis zones;
- Current zonal data on households and employment by type on vehicles and land area;
- A matrix of zone-to-zone off-peak highway travel times in the year of the commercial vehicle travel survey;
- An existing model system to which truck travel models can be added, or in which existing truck travel procedures can be replaced; and
- Estimates of regional VMT by commercial vehicle type and by private automobiles.

Although the information requirements of this strategy for transferring the procedures used in this project to other areas are reasonable, the costs of doing so will be significant. Thus, it is important to explore less expensive means of transferring the modeling strategies developed in this project to other urban areas. Recognizing the inherent trade-offs between the reductions in costs and possible reductions in precision and accuracy involved in alternative approaches, several possible approaches are described briefly. They are ordered from the least costly to the most costly in terms of resource requirements and development time.

Complete Transfer of Phoenix Models

If planners in another urban area have no current tools to predict commercial vehicle travel, they would be able to use the models developed in this project, including its modeling strategies, model parameters, and UTPS travel forecasting procedures and setups. In this way they could implement a complete new set of commercial vehicle models. To the extent that commercial vehicle travel patterns in Phoenix are representative of local conditions, this approach would provide a useful tool for local planning at a relatively small cost.

This approach would be reasonable for a number of large and growing cities in the South and West whose current or expected future levels of commercial vehicle travel are similar to those in Phoenix.

Adjusting Phoenix Models To Match Local Data

The previous strategy could be improved at low cost by adjusting the Phoenix models to match local information on commercial vehicle registrations or VMT for the entire study area, or both, following the strategies used for Phoenix. In this project, information on total vehicles registered by weight class was used to provide preliminary expansion factors for both vehicles and trips. Changes in registrations per employee could thus be used to adjust the Phoenix trip generation models for application in other cities. Similarly, changes in the resulting models could be adjusted to match total commercial VMT. As in the case of Phoenix, VMT data can be obtained from local vehicle classification counts. Thus, after the Phoenix trip generation models are revised, a set of commercial vehicle models calibrated to local regionwide data can be obtained.

The Phoenix trip distribution models can also be adjusted if local data on average commercial vehicle trip lengths are available. However, this information is not likely to be available unless a recent commercial vehicle travel survey has been conducted.

Model parameters that are revised to reflect local conditions in other urban areas require changes of various types in the programs that implement the models. These changes include

- Revisions to the trip generation models implemented in Fortran to reflect coefficient changes required to match local measures of vehicle registrations or VMT, or both; and
- Revisions of the friction factors input to the trip distribution models implemented in AGM to reflect changes required to match local data on average trip lengths.

Development of National Model

Perhaps the ultimate extension of the models developed in this project to other urban areas would involve generalizing them to create a national model, taking the quick-response system (*3*) as a pattern. This would involve combining the existing models with information in the FTA reports *Characteristics of Urban Travel Demand* and *Characteristics of Urban Travel Supply* to provide tables of each of its parameters as these are likely to vary by urban area type and size. Although this would involve a great deal of effort, it would provide all urban areas with versions of the models developed in this project that, in the absence of local data and model estimation, could be used to estimate commercial vehicle travel with acceptable levels of accuracy for sketch planning purposes, such as performing initial feasibility assessments of new highway facilities with or without features designed for exclusive use by either automobiles or commercial vehicles.

ACKNOWLEDGMENTS

The Phoenix urban truck travel project was conducted for the Arizona Transportation Research Center, Arizona Department of Transportation. MAGTPO, the metropolitan planning organization for the Phoenix area, provided technical monitoring of the project. The project was carried out by Cambridge Systematics, Inc., Cambridge, Massachusetts, with subcontractor support for data collection from O'Neil Associates, Inc., Phoenix.

REFERENCES

1. Cambridge Systematics, Inc. *Final Report: Phoenix Urban Truck Travel Model Project.* Arizona Department of Transportation, Phoenix, 1991.
2. F. G. Rawling and E. Ryan. CATS Commercial Vehicle Survey of 1986: A Discussion of Project Management Issues. *CATS Research News,* Vol. 26, No. 1, Feb. 1987.
3. A. B. Sosslau et al. *NCHRP Report 187: Quick-Response Urban Travel Estimation Techniques and Transferable Parameters: User's Guide.* TRB, National Research Council, Washington, D.C., 1978.

Publication of this paper sponsored by Committee on Urban Goods Movement.

Factors Affecting Automobile Ownership and Use

PANOS D. PREVEDOUROS AND JOSEPH L. SCHOFER

Factors affecting automobile ownership, effects of the availability of company cars on automobile ownership and travel behavior, and effects of short-term household evolution on automobile ownership were explored using data from two low-density, outer-ring growing suburbs and two high-density, inner-ring stable suburbs. Major findings are summarized. The average number of automobiles per household in the sample was 2.02. The average share of trips by automobile (versus transit) was 84.6 percent for work trips and practically 100 percent for nonwork trips. Life-cycle stage is a major determinant of automobile ownership; its effect is highly nonlinear. Changes in life-cycle stage alone may cause automobile ownership to increase (or decrease) substantially. Also, aging of children tends to increase automobile ownership. The number of persons eligible to drive and the number of workers affect automobile ownership positively, whereas the use of transit is associated with reductions in automobile ownership. Location of residence or workplace, or both, and residence relocation into outer-ring, low-density suburbs affect automobile ownership positively, whereas workplace locations in the central city affect automobile ownership negatively because of the availability of high-quality commuter rail service. The availability of company cars increases automobile ownership by 9 percent. The travel characteristics of those with company cars are affected significantly by the availability of an essentially free vehicle. There is some evidence that the increasing availability of company cars may increase daily vehicle miles of travel and, consequently, worsen traffic congestion.

The analyses in this paper are based on a sample of 1,420 respondents to a 1989 survey of Chicago suburbs. The following analyses are presented: (a) a static analysis to identify factors that currently affect or explain household automobile ownership; (b) an exploration of the effects of company car availability on automobile ownership and travel; and (c) a dynamic analysis of the effects of short-term household evolution on household automobile ownership.

The data include extensive demographic and socioeconomic information about households as well as detailed information on the automobile fleet available to each household. The data also include limited retrospective information on these variables for 2 years before the time of the survey (i.e., 1987); these were gathered to address issues of short-term dynamics of household automobile ownership. In addition, a weekday travel diary was completed by almost each respondent (typically an adult household member).

Four suburbs were included in the survey: two outer-ring, low-density growing suburbs (Naperville and Schaumburg) and two inner-ring, high-density (i.e., 9,700 residents per mi^2) stable suburbs (Park Ridge and Wilmette). Low density is defined as 4,600 residents per square mile, and high density as 9,700 residents per square mile. This classification of suburbs and the survey selection were taken from earlier findings (*1–4*).

OVERVIEW OF SAMPLE

Compared with nationwide averages, the sample (Table 1) contains marginally older people [i.e., United States average was 31.4 in 1984 (*5*) and is expected to be 34.1 in 1989 (*6*)] and larger households [i.e., the 1985 nationwide average was 2.75 and dropping (*5*)]. The households in the sample earn a much higher income than the nationwide average, which was about $30,000 in 1989. (All dollar figures in this paper reflect 1989 conditions.)

Younger, larger, and less-affluent households reside in outer-ring, low-density growing suburbs. More outer-ring suburban households have children living at home. Outer-ring suburban households have more workers, and their automobile ownership is higher in terms of number and worth of automobiles owned than inner-ring suburban households. The differences in all the characteristics between outer- and inner-ring suburbs are significant at the 95 percent level (Table 1).

Automobile ownership is high and steadily increasing for all four suburbs surveyed; the percentage of households owning two or more automobiles was 62.2 percent in 1980 (census), 72.2 percent in 1987 (retrospective part of survey), and 75.1 percent in 1989.

Most household vehicles reported are passenger cars (83.3 percent), whereas 11.0 percent are station wagons or passenger vans, and 5.7 percent are cargo vans or pickup trucks. The U.S. manufacturers have the lion's share (71.5 percent); Japanese brand-name automobiles come next (18.7 percent) and European automobiles follow (9.2 percent). The average age of vehicles owned is low (4.4 years) compared with the nationwide average [6.9 years in 1983 and climbing (*4,5*)].

Automobile ownership increased roughly by 1.5 percent a year from 1987 to 1989—a high rate given that automobile ownership in the sampled suburbs is already high. Part of the reason is that the period between 1987 and 1989 was marked by extraordinary financial incentives from domestic and foreign manufacturers. Since then, new car sales have decreased substantially. Another part of the reason is that about half of the sample contains people who either reside in or relocated from inner-ring suburbs, the central city, or other areas outside the Chicago metropolitan area to outer-ring, lower-density suburbs; such locations have a strong positive asso-

---

P. D. Prevedouros, Department of Civil Engineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822. J. L. Schofer, Department of Civil Engineering, Northwestern University, Evanston, Ill. 60208.

TABLE 1 Characteristics of Suburbs

Characteristics	OUTER RING	INNER RING	ALL
Population Age	31.7	37.8	34.5
Household (HH) Size	2.9	2.8	2.85
Number of Workers	1.7	1.5	1.6
% of HHs with Children	54.3	49.2	52.0
Average HH Income	59,500	67,200	63,000
Automobiles per HH	2.07	1.96	2.02
Median Worth of HH Autos	13,000	12,000	12,700

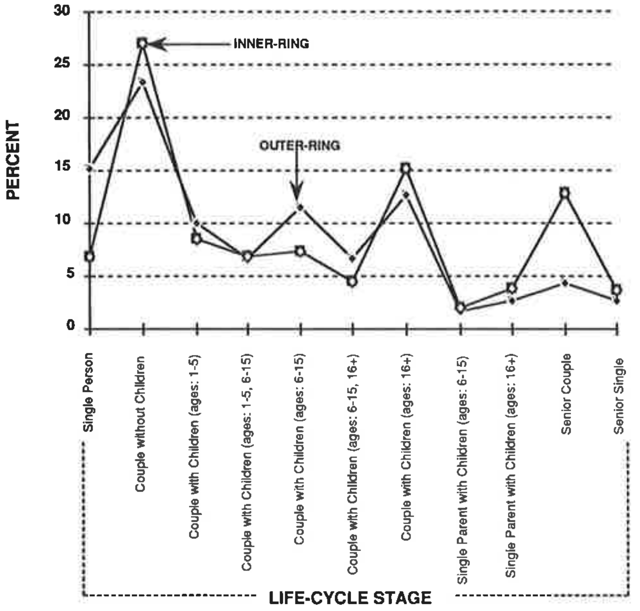


FIGURE 1 Life-cycle stage distributions in outer- and inner-ring suburbs.

ciation with increases in automobile ownership, as discussed later in detail.

The life-cycle stage of a household describes its position in the cycle of life. The life-cycle stage distributions are substantially different between outer- and inner-ring suburbs, as Figure 1 indicates. In Figure 1, several children-age categories for the Single Parent with Children life-cycle stage are not present because of lack of data.

STATIC ANALYSES OF AUTOMOBILE OWNERSHIP

In the static analyses section, descriptive and variance and regression analyses are used to identify associations between potential causal factors and automobile ownership. The last part of this section presents a detailed analysis on the associations and effects of company cars on automobile ownership.

Factors Affecting Automobile Ownership

The number of automobiles and its variants (i.e., automobiles per driver and total worth of household automobiles) are analyzed extensively. These three variables include owned and leased cars and company cars available to the household.

The independent variables used are defined as follows:

- LC STAGE 11 is the original list of life-cycle stages (Figure 1);
- LC STAGE 4 is an aggregate list of life-cycle stages compiled from the original list of 11 stages as follows:
  - –Early (Stages 1 and 2),
  - –Child-development (Stages 3, 4, 5, and 8),
  - –Child-independent (Stages 6, 7, and 9), and
  - –Senior (Stages 10 and 11).
- WORK FULL is the number of full-time workers in the household;
- HHSIZE is the household size (range, 1 to 6);
- DRIVERS is the number of household members eligible to drive (i.e., at the age of 16 or older) irrespective of the extent that the privilege is exercised;
- INCOME CAT represents total household income categories specified as follows:
  - –Less than $32,500,
  - –At least $32,500 but less than $57,500,
  - –At least $57,500 but less than $100,000, and
  - –$100,000 or more.
- TRANSIT represents the use of public transportation (1 if at least one trip in the 1-weekday travel diary was made by mass transit, and 0 otherwise);
- DENSITY represents the type of community in which the household resides (1 = low-density, outer-ring, and 0 = high-density, inner-ring);
- SUBURBS is the number of household workers who work in a suburb other than the one in which they reside (range = 0 to 4);
- SAME is the number of household workers who work in the suburb in which they reside (range = 0 to 2);
- CITY is the number of household workers who work in the non–central business district (CBD) central city of Chicago (range = 0 to 3);
- CBD is the number of household workers who work in Chicago's CBD (range = 0 to 2).

Analysis of variance indicated that the number of household members eligible to drive had by far the largest effect on automobile ownership; it overshadowed the contribution of most other variables. Excluding this variable from specifications revealed that (a) life-cycle stage has the highest contribution in the amount of variance explained (12 percent), (b)

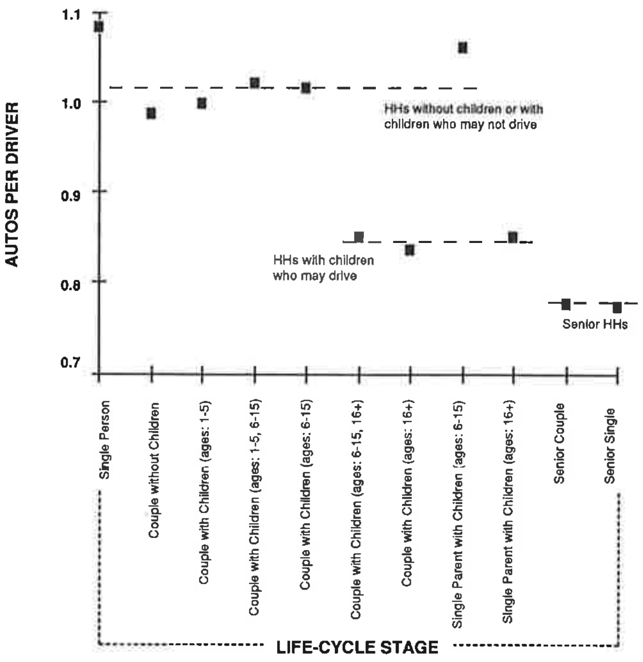


FIGURE 2 Distribution of ratio of automobiles per driver across life-cycle stages.

aggregate life-cycle stage (4 stages) does not work as well as the original 11-stage classification (5 percent), (c) residence location and location of work are important and significant contributors (DENSITY, 6 percent; all workplace variables combined, 11 percent), (d) public transportation use has a large and significant effect (9 percent), and (e) size (4 percent), number of full-time employed workers (7 percent), and household income (3 percent) all have a significant effect on household automobile ownership (*4*).

Of all workplace variables, SUBURBS has by far the largest contribution in the portion of variance explained, which signifies the need for more automobiles if the workplace is in a suburb, since currently public transportation service is limited among Chicago suburbs.

The most powerful explanatory factors—life-cycle stage, drivers, and income—are discussed further. Figure 2 shows that automobile availability is low (i.e., approximately 0.85 automobile per driver) at life-cycle stages describing families with children eligible to drive (i.e., child-independent stages 6, 7, and 9); it takes a relatively stable value around 1.0 automobile per driver over stages describing either families with children who cannot drive (i.e., child-dependent stages 4, 5, and 8) or early-stage households (HHs) (i.e., Stages 1, 2, and 3); and it reaches a minimum around 0.77 automobile per driver at very advanced life-cycle stages (i.e., senior couples or singles).

The trend of automobile availability across life-cycle stages suggests a maturing over time with respect to automobile acquisition, that is, households in advanced stages (i.e., 6 to 11) seem to be better at rationing and using their automobile fleet by assigning roughly 0.90 automobile per driver, whereas in early stages there is a tendency to own more automobiles than appear to be necessary (i.e., in several stages more than one automobile is available per driver). At advanced stages, with more drivers in the household, families may come up against income constraints that prevent them from maintaining the automobile-per-driver levels achieved in earlier stages.

The number of drivers and income are important variables that partially explain household decisions concerning automobile ownership. Table 2 presents the level of automobile ownership by household income categories and number of drivers. Income per person is used instead of household income not only because household size is implicitly accounted for, but also because this is a more objective representation of a household's financial ability given that total income must cover the needs of a variable number of persons (e.g., a two-person household earning $50,000/year can afford higher expenditures than a four-person household earning the same income). Among other things, the former household is financially better equipped to buy an automobile and to participate in activities, which in turn increases mobility needs.

The results in Table 2 indicate that the number of drivers (i.e., people who can fulfill their transportation needs by driving) is a much stronger force driving automobile ownership than income, which has a positive but marginal effect on automobile ownership. Economics of scale are clearly identifiable: households with one or two drivers own roughly one or two automobiles, respectively, and households with three or more drivers own roughly 20 percent fewer automobiles than the respective number of drivers.

The reason that income has only a marginal effect on household automobile ownership may be that for this relatively high income sample, the ability to buy automobiles is less important than the need to use them. Therefore, if there is little need for additional automobiles, it is not likely that an additional automobile will be bought, regardless of the fact that there may be no income constraint. This is supported by the presence in the data set of affluent single-person households

TABLE 2 Automobiles by Household Income per Person and Drivers

	INCOME PER HOUSEHOLD PERSON ($ thousands)			
	INC<10	10≤INC<20	20≤INC<30	30≤INC
DRIVERS				
1	0.89 (28)	0.93 (29)	@	1.09 (144)
2	1.83 (83)	1.92 (300)	1.93 (303)	2.01 (168)
3	@	2.63 (97)	2.61 (33)	2.63 (57)
4+	@	3.18 (66)	3.58 (47)	@

__NOTES:__ (#) = number of households;
@ = insufficient number of observations (0-9).
Figures are from 1989 calculations.

TABLE 3 Worth of Automobiles by Household Income per Person and Drivers (in Thousands of Dollars)

	INCOME PER HOUSEHOLD PERSON ($ thousands)			
DRIVERS	INC<10	10≤INC<20	20≤INC<30	30≤INC
1	4.4 (28)	4.9 (29)	@	8.1 (144)
2	10.9 (83)	12.6 (300)	14.8 (303)	17.6 (168)
3	@	16.6 (97)	23.2 (33)	21.7 (57)
4+	@	18.6 (66)	23.1 (47)	@

NOTES: *(#) = number of households;*
@ = insufficient number of observations (0-9).
Figures are from 1989 calculations.

and double-income–no-kids households with upper-bracket incomes and relatively low automobile ownership.

Another perspective on the role of income is revealed in Table 3, in which the worth of household automobiles is tabulated with respect to household income per person and number of drivers. It is obvious that larger-income households spend a substantially larger amount of money on automobiles than do lower-income households.

An interesting example of resource allocation is furnished by the second column of Tables 2 and 3 (income per person between $10,000 and $20,000). One-driver households (mostly households having a single person, a senior single person, or a single parent with young children) have no one to share expenses (i.e., rent or mortgage, home equipment). As a result, less capital is left for automobile acquisition: the average worth of each automobile owned is $5,300. Larger households with two or more drivers are able to realize economies of scale. As a result, more capital becomes available for automobiles: the average worth of each automobile owned is $6,500.

The aforementioned interpretations deal with behavior that accounts for the transportation utility of automobile ownership; the value of automobiles owned as well as the potentially confounding role of prestige involved in automobile ownership remain unaccounted for. These issues have been explored previously (*7*).

Automobile Ownership Models

The next step of the analysis was modeling of household automobile ownership. The specifications include all factors that could help in understanding more about the process affecting household automobile ownership decisions.

The model displayed in Table 4 offers a good fit with the data ($R^2 = .54$) that improves substantially ($R^2 = .67$) after segmentation across the categories involving the age of the head of the household.

The inclusion of the life-cycle stage variable in the specification is bound to fail because the numerical values assigned to various life-cycle stages do not correspond in any way to the natural meaning of the life-cycle stage concept and its effect on household automobile ownership. The use of dummy variables to represent life-cycle stages resulted in an inferior model fit with a large number of insignificant parameter estimates (see also *Dynamic Analysis of Automobile Ownership*).

The best results for nonsegmented models were obtained by including CHILD DEP, a dummy variable that represents households in child-dependent stages (i.e., stages in which children exist and are all younger than 16 years of age); and SENIOR, a dummy variable that represents senior households in which at least one member is 65 years or older. Another variable used in the model is HHINC$, which is the total household income in thousand dollars.

TABLE 4 Explanatory Household Automobile Ownership Model: Best-Fit Estimation and Best-Fit Segmentation

Model Segment	A L L	Age of Head of Household Categories 35<	35-49	50-64	65+
Dependent	NUMBER OF HOUSEHOLD AUTOMOBILES				
R-squared	0.54	0.67			
Cases	1372	207	520	92	153
Constant	0.23	0.07*	0.68	0.94	0.49
DRIVERS	0.58	0.89	0.51	0.45	0.53
DENSITY	0.14	0.24	0.13	-	-
TRANSIT	-0.28	-0.20	-0.24	-	-
WORK FULL	0.11	-	0.15	-	-
SUBURBS	0.11	-	-	0.24	0.26
CHILD DEP	0.11	-	-	-	n/a
SENIOR	-0.13	n/a	n/a	n/a	n/a
HHINC$	0.004	-	-	-	-

Note: ***all perameters significant at 95%, except (*)=not sign.***
(all estimations with 1989 data)

All factors displayed in the model have a significant impact on household automobile ownership. The number of drivers has the largest positive effect, and use of public transportation has the largest negative effect on automobile ownership. Residence locations in a low-density, outer-ring suburb as well as the number of workers who are employed in the suburbs (with little public transportation for suburb-to-suburb commutes) increase household automobile ownership. All variables except TRANSIT and SENIORS have a positive effect on automobile ownership.

Being at a child-dependent stage increases automobile ownership, whereas being at a senior stage decreases automobile ownership. There are many more households in child-dependent stages in outer- (36.4 percent) than in inner-ring suburbs (29.0 percent), and many more senior households in inner- (16.6 percent) than in outer-ring suburbs (6.9 percent). The significant difference in automobile ownership between outer- (i.e., 2.07 automobiles per household) and inner-ring suburbs (1.96 automobiles per household) can be partly explained by the various underlying life-cycle stage distributions in each type of community.

The model specifications in Table 4 show, however, that despite the inclusion of major socioeconomic, locational, and transportation supply variables, the DENSITY variable remains significant. The variable may account for other factors in outer- and inner-ring suburbs, such as intra-household age structure differences, underlying cultural differences, or differences in spatial arrangements of land uses and the transportation systems connecting them.

Several estimations across various population segments were explored (e.g., age of household head, life-cycle stage, household income categories and residence location). Segmentation by the age of the head resulted in a most substantial improvement in model fit (see Table 4). Some variables from the full specification were omitted from this final estimation, either because they are not applicable (n/a) or because their parameter estimate in earlier specifications was not significant.

Segmentation by age of the head of household reveals that the effect of the number of licensed drivers tends to decrease as the household ages. The effects of residence location and use of public transportation on automobile ownership are significant only for younger households. Also, the number of workers employed in the suburbs has a significant effect on automobile ownership only in older households, in which multiple workers are more common.

Company Cars

This section presents a detailed investigation of the effects of company car availability on individual and household transport behavior characteristics.

In a study conducted in the United Kingdom, Bates et al. point out that employer financing or subsidy of automobile costs has important effects on the household's decisions about automobile ownership and use (*8*). The availability of company cars increases the actual household automobile ownership through two causal chains: (a) the household disposable income is effectively increased, and (b) restrictions in the use of company-financed vehicles force households to maintain additional vehicle(s). Bates et al. also noted that company car users tend to live farther from work and that automobiles are provided predominantly according to status instead of to job performance requirements. These results may carry cultural or historical biases (1969–1975 data from the United Kingdom).

The analysis strategy matched groups of households with and without company cars in terms of the number of drivers, number of workers, number of automobiles available, and income bracket. The only difference is that the automobile ownership of the one group is equal to *n* owned or leased automobiles, whereas the automobile ownership of the other group is equal to *n* − 1 owned or leased automobiles plus one company car; so both groups of households have an identical number of automobiles available.

The results of this investigation are summarized in Table 5. The sample includes 67 households that had one company car available and 410 similar households that had no company

TABLE 5 Effects of Company Cars on Transportation Behavior of Individuals Who Use Them

#CARS	DRV	WRK	HH INCOME	Cases	% DRIVE ALONE TO WORK	% WORK in SUBURBS	DISTANCE by AUTO (miles)
1+C	2	1	$57,500-	12	*100.0	81.8	*48.1
2	2	1	$99,999	86	66.7	52.9	28.3
1+C	2	2	$57,500-	21	*90.5	86.4	*66.1
2	2	2	$99,999	147	82.7	81.9	37.9
1+C	2	1	$100,000	7	*100.0	71.4	N42.6
2	2	1	or more	56	69.8	44.2	37.8
1+C	2	2	$100,000	20	*100.0	70.0	@49.2
2	2	2	or more	107	75.8	58.1	35.0
2+C	3	3	$100,000	7	@100.0	71.4	@58.4
3	3	3	or more	14	69.2	64.3	36.1

***NOTES**: DRV=number of household drivers; WRK=number of household workers; (*)=comparison significant at 95%; (@)=comparison signif. at 85%; (N)=comparison not signif.*

car available. The respondent is the primary driver of the company car in 65 of the 67 selected households that have a company car available. The actual number of households with company cars in the sample is 164 (11.5 percent of the sample), but the socioeconomic profiles are so widespread that only 67 belonged to groups of sufficiently high membership to support statistical analysis.

The results in Table 5 suggest that the availability of a company car causes several substantial and, in many cases, significant differences in transport behavior. To begin with, the commute to work by those with company cars is being made exclusively by automobile; in most cases it is drive alone (i.e., 100 versus 67 percent for the first group in Table 5, or 97 versus 73 percent overall—weighted average of all groups in Table 5). Also, the overwhelming majority of company car users work somewhere in the suburbs (including the suburb in which they reside). This proportion is much lower for respondents from similar households without a company car (i.e., 82 versus 53 percent for the first group in Table 5 or 78 versus 64 percent overall—average of all groups in Table 5).

The total distance traveled by automobile on the 1 weekday for which the respondent reported is in most cases significantly higher for owners of company cars than for those without company cars. This finding is largely because company car owners need to travel more for work-related purposes: company car owners travel 29.4 mi for such purposes, whereas respondents from similar households without a company car travel 11.0 mi (averages from all groups in Table 5).

The latter may be partly attributed to the substantial difference in occupations between the two groups, as the following distributions of occupations indicate:

Occupation	*Respondent with Company Car*	*Respondent Without Company Car*
Managerial/business owner	48.6	29.2
Sales	31.4	10.9
Professional/technical	12.9	45.2
Other	7.1	14.7
Total	100	100

The proportion of salespersons in the group with the availability of a company car is three times higher, and the proportion of managers or business owners is nearly two times higher than the group without the availability of a company car. Thus, company cars are seemingly more available to travel-intensive occupations. (This does not preclude use of the company car for personal trips. It is possible that a number of household trips are diverted to the company car from the owned automobiles.) In contrast to the study by Bates et al. (*8*), which noted that company cars are given mostly to high-status employees, in the United States (the Chicago suburbs to be exact) company cars seem to be available mostly for employees who need to travel extensively. Of course business owners may have the power to purchase company cars independent of their business travel needs.

The degree of substitution between company and privately owned automobiles is revealed in the following automobile ownership model (in which all parameters are significant at 95 percent):

$$\begin{aligned}\text{HH AUTOS} = {} & 0.39 + 0.27 * \text{COMPANY} \\ & + 0.10 * \text{DENSITY} + 0.13 * \text{CHILD DEP} \\ & + 0.14 * \text{WORK FULL} \\ & - 0.19 * \text{TRANSIT} + 0.10 * \text{SUBURBS} \\ & + 0.59 * \text{DRIVERS} \\ & - 0.14 * \text{SENIOR} \qquad R^2 = .53\end{aligned}$$

This model indicates that the availability of a company car increases the automobile ownership of households by 0.27 automobile, which corresponds to an average increase of about 9 percent. In other words, only 9 percent of the households maintain the regular household fleet plus the company car; the rest of them (91 percent) substitute a household-owned automobile with the company car—that is, if a company car is available, they do not buy an additional car. This means that the cost of owning and operating a second or third household automobile (the company car) is partly or entirely absorbed by a company; therefore, the annual household income effectively increases by the amount of the subsidy, thereby increasing the overall ability of the household to participate in activities and to travel.

The Internal Revenue Service requires that company cars be reported for taxation as fringe benefits. One method for estimating the contribution of a company car to a household's income is the *annual lease value method*, according to which the fair market value of the company car is assessed when given to the employee. Then an annual lease value for the first 4 years is computed; for example, a car worth $11,000 in 1990 has an annual lease value of $3,350 for the years 1990 through 1993. This value is to be added to the household's income. For the next 4 years, assuming that the same company car is available to the household, the fair market value is reassessed and the new annual lease value is estimated (*9*).

The proportion of households with the availability of a company car is 13.5 and 10.4 percent in outer- and inner-ring suburbs, respectively; the 3.1 percent difference is statistically significant at the 91 percent level. The outcome that more households with company cars were captured in outer-ring suburbs agrees with the outcome of Bates et al., who found that company car users tend to live farther from work. Furthermore, 12.7 percent of those households that relocated between 1987 and 1989 had the availability of a company car, whereas only 11.3 percent of those that did not relocate had the availability of a company car. These may be indications that companies essentially subsidize (affordable) distant housing.

DYNAMIC ANALYSIS OF AUTOMOBILE OWNERSHIP

The objective of this part of the analysis is to explore the potential causal links between changes in household characteristics (i.e., life-cycle stage, household size, number of workers and residence location) and changes in household automobile ownership between 1987 and 1989. Several limitations apply to this part of the analysis because of the restricted breadth of the retrospective part of the questionnaire survey. Information about incomes, work locations, and use of public transportation and company-subsidized automobiles is not available for 1987; thus their changes between 1987 and 1989 cannot be estimated and analyzed. As a result, outcomes from this part of the analysis can serve only as indications of potential associations between changes in causal factors and changes in automobile ownership.

This objective was approached using analysis of variance and regression modeling. The following variables were used:

- *Dependent Variable*
  - –DCARS: automobile ownership difference = (1989 automobiles) – (1987 automobiles).
- *Independent Variables*
  - –DHHSIZE: difference in household size = (1989 size) – (1987 size);
  - –DWORKFULL: difference in number of full-time employed household workers = (1989 full-time workers) – (1987 full-time workers);
  - –DDRIVERS: difference in number of persons eligible to drive in the household; the variable takes integer values from –3 to 2.

Residence location changes between 1987 and 1989 are defined with three dummy variables:

- SUBURB: 1 = relocated from another suburb, 0 = no change;
- CITY: 1 = relocated from city of Chicago, 0 = no change; and
- FAR: 1 = relocated from some place outside metropolitan Chicago, 0 = no change.

DSTAGE is the change in household life-cycle stage. The effects of the change of life-cycle stage (DSTAGE variable) on automobile ownership are nonlinear; thus, a special procedure was devised to determine appropriate values. This procedure involved the estimation of a linear regression automobile ownership model in which each life-cycle stage was included as a dummy variable. The parameter estimates represent the effect of each life-cycle stage on automobile ownership. Then, these estimates were used in a dynamic model. Automobile ownership models with life-cycle-stage dummy variables were estimated with both 1987 and 1989 data. All parameters but one were significant in the 1987 model, whereas five parameters in the 1989 model were not significant. The 1987 model was used (*4*). Although its overall goodness of fit is mediocre ($R^2 = .34$), all parameters are significant and intuitive in sign and relevant size; therefore, the model was used with confidence.

The model includes major factors such as residence location, household size, and workers in the specification; thus, the parameters for each life-cycle stage largely express the effect of each life-cycle stage on automobile ownership. These parameters were used to create a matrix, the cells of which contain the specific effect of changing stages on automobile ownership. For example, a change from Stage 2 to Stage 3 results in a decrease of automobile ownership by 0.28 automobile. A small part of this matrix is illustrated. The 121 values contained in the full 11-×-11 matrix in Table 6 are the values for the DSTAGE variable (*4*).

The use of the DSTAGE variable relies on the assumption that the effects of life-cycle stage on automobile ownership remain the same in 1989 as in 1987. Application of the F_{CHOW} test between the 1987 and 1989 models did not reject the null hypothesis that the parameter estimates of the two models are equal (*4*); therefore, changes in tastes and needs caused by changes in culture and economy were assumed to be negligible (over the 2 years studied).

The best-fit linear regression specification for the dynamic model is presented next; all parameters are significant at 95 percent.

$$\begin{aligned} \text{DCARS} = {} & 0.087 + 0.206 * \text{DHHSIZE} + 0.139 * \text{DWORKFULL} \\ & + 0.249 * \text{DDRIVERS} + 0.378 * \text{DSTAGE} \\ & - 0.148 * \text{SUBURB} + 0.209 * \text{CITY} \\ & + 0.115 * \text{FAR} \qquad R^2 = .23,\ F = 60,\ n = 1{,}415 \end{aligned}$$

The model suggests that increases in household size and number of workers and drivers have a positive effect on household automobile ownership. The mediocre overall fit of the model may be because (a) the period covered is short, and only two time points are available; and (b) several explanatory variables could not be included because they are not known for 1987.

The parameter of the DSTAGE variable must be multiplied by the appropriate value corresponding to each particular life-cycle change (i.e., for change from Stage 1 to Stage 2 the value is 0.34) to estimate the effect on the change in automobile ownership. This effect can be positive, negative, or zero; based on the model estimate, the range of the effect of life-cycle stage change varies between –0.4 and +0.4 auto-

TABLE 6 Matrix of DSTAGE Values

CHANGE IN AUTOMOBILE OWNERSHIP		TO (1989 LIFE-CYCLE STAGE)				
		1	2	3	4	. . .
	1	0	0.343	0.061	-0.008	
FROM	2	-0.343	0	-0.282	-0.352	
	3	-0.061	0.282	0	-0.070	
(1989	4	0.009	0.352	0.069	0	
LIFE-	.					
CYCLE	.					
STAGE)	.					

mobile, that is, in several cases it may have a powerful effect on automobile ownership.

The estimates for the relocation variables indicate that relocation from Chicago (CITY) and from another state or abroad (FAR) have a positive effect on automobile ownership, whereas relocation from another suburb (SUBURB) has a small negative effect on automobile ownership, possibly because some people relocate to certain locations to take advantage of their better public transportation supply or to get closer to their workplace.

Although part of the effect of missing variables is likely to have been captured by the constant, their absence probably affects, to an unknown degree, the magnitudes of the estimated parameters. Thus, all outcomes in this section should be treated with caution. The results probably serve as good indications of associations between changes in causal factors and changes in automobile ownership, but they may not be used to quantify the effect of independent variables on the dependent variable.

CONCLUSIONS AND DISCUSSION OF RESULTS

This paper presented a multiperspective investigation aimed at identifying static and dynamic factors affecting suburban household automobile ownership. Key findings are summarized below, followed by a short discussion.

- The life-cycle stage, number of persons eligible to drive, and household income are the major socioeconomic factors affecting household automobile ownership.
- The location of residence in outer-ring, low-density growing suburbs, as well as the location of workplace in the suburbs, increase household automobile ownership.
- Automobile availability (i.e., automobiles per driver) is 9 percent higher in outer- than in inner-ring suburbs; this is partly because most of the population of outer-ring suburbs is in early life-cycle stages with high automobile availability, whereas most of the population in inner-ring suburbs are senior households with low automobile availability.
- Relocation from the central city (Chicago), another U.S. state, or abroad to a suburb has a strong positive effect on automobile ownership.
- Higher-income households own more expensive rather than more automobiles than lower-income households with the same number of drivers.
- Economies of scale with respect to automobile ownership are clear: smaller households (i.e., with up to two drivers) own a number of automobiles equal to the number of drivers in the household, whereas larger households (i.e., with three or more drivers) own approximately 20 percent fewer automobiles than the number of drivers.
- Availability of a company car increases the average household automobile ownership by 0.27 automobile. Most households (91 percent) substitute company cars for household-owned cars, which translates into substantial income benefits for households with the availability of a company car.

This study indicates that "traditional" demographic and socioeconomic factors continue to have close associations with household automobile ownership. In addition, life-cycle stage, location of residence and employment, public transportation use, and company car availability affect automobile ownership. These less traditional but important factors affecting automobile ownership are discussed in the following.

There are indications that changes in life-cycle stage affect automobile ownership substantially. The effect is clearly nonlinear: the change from single person to couple marginally decreases automobile ownership; the change from couple without children to couple with young children increases automobile ownership; the change from couple with young children to couple with children, some of whom are eligible to drive, increases automobile ownership substantially; and most changes to single parent or to senior households decrease automobile ownership.

Location of residence in outer-ring suburbs and location of the workplace in the suburbs have positive effects on automobile ownership characteristics (i.e., automobile ownership and availability are 5.6 and 8.9 percent higher in outer- than in inner-ring suburbs, respectively).

Use of public transportation has a negative effect on automobile ownership. The lower automobile ownership characteristics in inner-ring suburbs can be partly attributed to the fact that public transportation services are much better in these locations: not only is more service available, but transit can serve most work trip commuting destinations of inner-ring suburb workers (i.e., central city destinations).

Whether the higher automobile ownership in outer-ring suburbs is a cause or an effect, the implications of this finding for congestion in the developing suburban fringe are significant and serious. More automobiles are connected with more travel and more pressure on the limited roadway network.

The number of eligible drivers in a household was found to be the most significant factor determining household automobile ownership across most social groups. It is theorized that the causal link lies among the following three facts: (a) irrespective of age, people face certain needs and activity sets—they need a means of transportation to fulfill these needs and to participate in activities; (b) public transit diminishes as a transportation option in the suburbs, particularly in outer-ring suburbs; and (c) most members of today's households are "busy with their lives" (i.e., work, school, individual exercise, and social activities), which is partly an outcome of the individualism and freedom of expression in modern society (more so in the United States than in other western societies). Therefore, there is little time left to serve other members of the household and less homogeneity in tastes and activities, which limits opportunities to consolidate destinations and share automobile usage. Thus, it is beneficial (if not necessary) to have an automobile available for every person in the household who is eligible to drive.

Individuals who have a company car available travel much longer distances by automobile; the automobile also is their exclusive means of travel to work. There is some evidence that more of them tend to reside farther away (outer-ring, low-density suburbs) as well. Thus, the availability of company cars may be fueling congestion because high use and long commutes by automobile become affordable.

ACKNOWLEDGMENT

This work was supported in part by the Illinois Department of Transportation and FHWA.

REFERENCES

1. P. D. Prevedouros and J. L. Schofer. Suburban Transport Behavior as a Factor in Congestion. In *Transportation Research Record 1237*, TRB, National Research Council, Washington, D.C., 1990, pp. 47–58.
2. P. D. Prevedouros and J. L. Schofer. Social and Cultural Factors Affecting Transportation Behavior in Rapidly Developing Suburbs. *Proc., 5th WCTR*, Yokohama, Japan, Vol. 4, 1989, pp. 169–183.
3. P. D. Prevedouros and J. L. Schofer. *Demographic, Social and Cultural Factors Affecting Suburban Congestion*. Illinois Department of Transportation, Springfield, 1990.
4. P. D. Prevedouros. *Demographic, Social, Economic and Personality Factors Affecting Suburban Transport Behavior*. Ph.D. dissertation. Department of Civil Engineering, Northwestern University, Chicago, Ill., 1990.
5. *Nationwide Personal Transportation Study: Personal Travel in the U.S., 1983–1984*, Vol. 1–3. U.S. Department of Transportation, 1986.
6. A. E. Pisarski. *Commuting in America: A National Report on Commuting Patterns and Trends*. Eno Foundation for Transportation, 1987.
7. P. D. Prevedouros. Associations of Personality Characteristics with Transport Behavior and Residence Location Decisions. *Transportation Research*, in press.
8. J. Bates, M. Roberts, S. Lowe, and P. Richards. *The Factors Affecting Household Car Ownership*. Gower Publishing Company, Ltd., United Kingdom, 1981.
9. Internal Revenue Service. *Internal Revenue Code Regulations* 1.61-21d2iii, *RIA* 1989.

The findings in this paper reflect the view of the authors and do not necessarily represent the positions of the Illinois Department of Transportation or FHWA.

Publication of this paper sponsored by Committee on Transportation Planning Applications.

Proposed Method for Calibrating Weigh-in-Motion Systems and for Monitoring That Calibration Over Time

CURTIS DAHLIN

Calibration of weigh-in-motion (WIM) systems is a difficult process that has many limitations and problems. The problems encountered with using test trucks for calibration are documented. WIM systems often produce differing weights for individual test trucks. Furthermore, calibration achieved in this traditional manner addresses how the system is performing only on those days that tests are conducted. Actual system performance for all other days is unknown. An alternative method for calibrating a system and a method for monitoring the performance of the system on an ongoing basis are proposed. This ongoing monitoring uses an aggregation of specific types of weight data for five-axle tractor-semitrailers that are collected by the system. The results obtained from applying this procedure can also be adapted for use in editing weight data.

The Minnesota Department of Transportation (MnDOT) has been operating permanent weigh-in-motion (WIM) systems since 1981. Currently there are installations at 16 locations. The methods used to calibrate the systems and monitor that calibration have evolved over this period as experience was gained.

TRADITIONAL METHODS OF CALIBRATION

Initially, the department relied solely on one test truck for the calibration. It was a loaded two- or three-axle single-unit department dump truck. Repeated runs over the scales were made at highway speeds, and the WIM scales were calibrated to reflect the static gross weight of the truck. It was then assumed that the weights that the system collected were valid until the next time a test truck was used at the site. Because of a shortage of personnel, a great deal of time would often elapse between tests.

Various aggregations of volume and weight data collected by the system were examined on a weekly basis. Each site usually exhibited fairly predictable patterns. Some of the early weight patterns noted were the distribution of gross weight and front-axle weight of five-axle tractor-semitrailers. It was also observed that the calibration for an individual lane occasionally would drift off. This raised concerns because of the inability to use test trucks every time the occasion appeared to require it.

When test trucks were used, sometimes both two- and three-axle single-unit trucks were available. These test results showed that a system could and frequently did weigh the trucks differently. For example, the two-axle truck might register on the average 2 percent high, whereas the three-axle truck would be 3 percent low. The system would then be calibrated to an average of these readings, but there was no way to determine how well it was actually doing with trucks in the traffic stream.

In 1989 extensive tests of WIM systems were conducted in an effort to determine the simplest, most reliable method for calibrating a WIM system. A portable WIM system was set up at the site of MnDOT's permanent WIM system on I-94 east of St. Paul. Two- and three-axle test trucks were used as well as monitoring trucks in the traffic stream. The traffic stream trucks were weighed at the St. Croix weigh station, which is 3 mi upstream from the permanent WIM site. Identifying characteristics of the trucks were noted, and they were matched up when they passed the WIM site. The data reported in Table 1 represent results based on a calibration that correctly weighs five-axle tractor-semitrailer(s) in the traffic stream. These results clearly focused on the pitfalls of relying solely on test trucks for calibration. It appears that the dynamics of any specific truck are unique and can be very different from those of other trucks with the same axle configuration. Recorded weights of the test trucks and the trucks of the same axle configuration in the traffic stream often corresponded poorly. Also, dynamics constantly change, so varying results are obtained on the same route, even when the scales are close together. These circumstances make it difficult to rely on a test truck to calibrate a system. Test trucks may not be representative of the dynamic weights of the most critical trucks in the traffic stream.

TABLE 1 Results of Calibration That Correctly Weighs Five-Axle Tractor-Semitrailers in Traffic Stream

	Permanent WIM		Portable WIM	
Vehicle Type	# of Vehicles	% Deviation	# of Vehicles	% Deviation
Traffic Stream				
2 axle 6 tire	27	+ 1.7%	31	- 7.9%
3 axle single unit	17	- 3.7	19	+ 1.2
5 axle semi	208	0.0	243	0.0
Twin trailers	19	+ 5.8	19	- 7.3
Test Trucks				
2 axle 6 tire	19	+13.8	21	-10.3
3 axle single unit	9	-11.5	11	- 5.2

Minnesota Department of Transportation, 813 Transportation Building, 395 John Ireland Boulevard, St. Paul, Minn. 55155.

OTHER EFFORTS TO ADDRESS PROBLEM

The lack of a standardized procedure for acceptance of the various types of WIM systems and performance standards led ASTM to examine this issue. ASTM conducted a study of WIM, and in 1990 it published the results (ASTM E1318-90). Its procedure for WIM acceptance involves using test trucks and randomly selecting and statically weighing trucks from the traffic stream. ASTM specifies how many trucks of each type to weigh and then calibrates the system to the average of those weights. This is unquestionably the best method. Unfortunately, however, it is impractical in most cases. One needs either a static weigh station in the immediate vicinity or portable static scales. The former situation is a rare occurrence, and the latter is time consuming, labor-intensive, and dangerous, assuming that an agency has access to portable static scales.

PROPOSED ALTERNATIVE

MnDOT's prime concern when dealing with the issue of WIM system calibration is to ensure that the weight data collected are valid. The principal thrust of this paper is to develop a system that can be used to identify those instances in which the weights are systematically off by a significant amount, defined as 4 percent at this time. There are two reasons for using 4 percent. First, the fourth-power relationship between weight and equivalent single axle loads (ESALs) means that a 4 percent difference in weight translates into a significant difference of about 16 percent when dealing with ESALs. The second reason for choosing 4 percent is that currently it is at best difficult and perhaps even impossible to achieve and especially to maintain a calibration that has a true systematic error of 3 percent or less.

After studying the calibration issue for some time, MnDOT developed a comprehensive procedure that works well, one that is being used as a manual procedure. The techniques being used concentrate on five-axle tractor-semitrailer(s), the vehicle type that has the greatest impact in terms of ESALs on Minnesota's highways. They typically contribute 70 to 90 percent of the ESALs on many of the state's trunk highways. The goal is to collect accurate weight data on them.

For the initial system calibration, a five-axle tractor-semitrailer loaded to 75 or 80 kips is used. The loaded five-axle tractor-semitrailer(s) in the traffic stream contribute the vast majority of the ESALs of all five-axle tractor-semitrailer(s). This test truck is equipped with a leaf spring suspension, which is the most common one in use. A minimum of 25 passes over the scales at highway speeds are specified after the final calibration adjustment. These runs confirm that the calibration with that vehicle is correct.

Next, the data collected on each individual lane at a site are monitored. The weight data for five-axle tractor-semitrailers are monitored in three areas:

1. Distribution of gross weight,
2. Front-axle weights, and
3. Flexible ESAL factors.

If there is a system malfunction that is severe enough, any one of these areas can indicate that there is a problem and that the data are invalid. All three are used because each plays a strong supportive role in making this determination.

DISTRIBUTION OF GROSS WEIGHT

The first area checked to determine the status of the calibration and the validity of the weight data is the distribution of

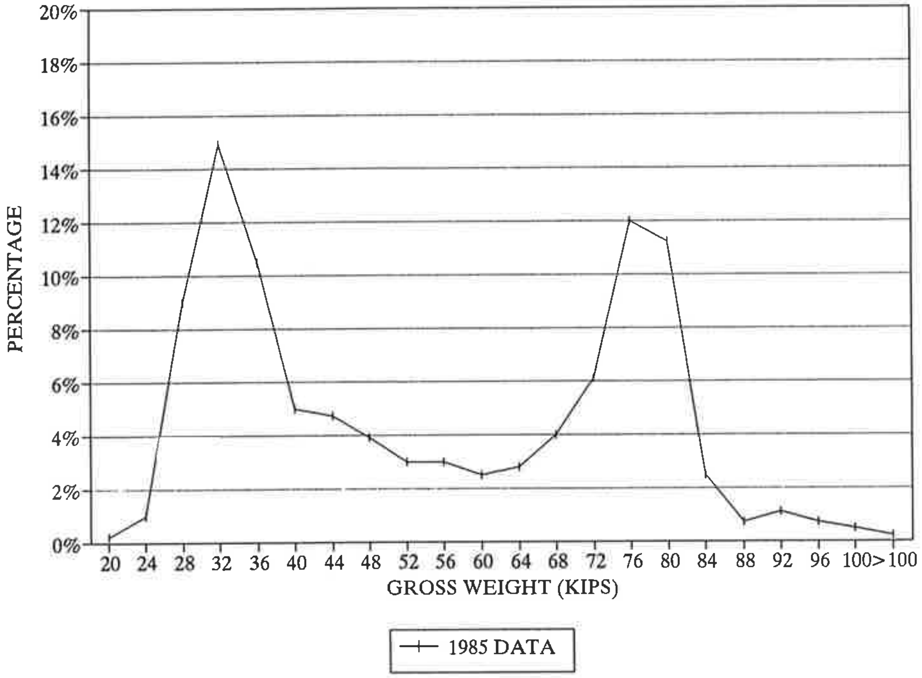


FIGURE 1 Distribution of static gross weight of five-axle tractor-semitrailers.

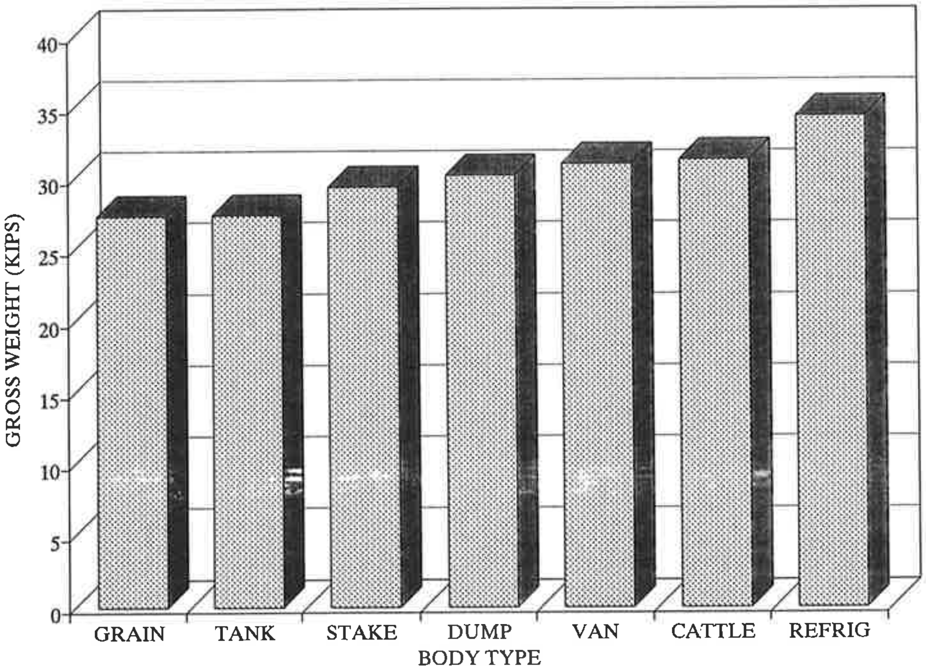


FIGURE 2 Average static weight of empty five-axle tractor-semitrailers.

the gross weight of five-axle tractor-semitrailers. It is here that the average weights of the loaded and unloaded vehicles assist in the process of separating valid and invalid weight data. This distribution should have one peak for empty vehicles at about 28 to 32 kips and another peak in the 70- to 80-kip range for loaded vehicles. This peak in the 70- to 80-kip range reflects Minnesota's gross weight limit of 80 kips. Figure 1 shows the distribution from static statewide weighing in Minnesota in 1985. These data, which represent about 3,100 vehicles, have the general pattern usually found. The gross weight figures represent the upper end of each category; for example, 32 represents those weights from 28 to 32 kips.

The type of trailers used on five-axle tractor-semitrailers varies somewhat from one area of the country to another. Figure 2 shows the static weight of empty five-axle tractor-semitrailers by body type in Minnesota. They range from a

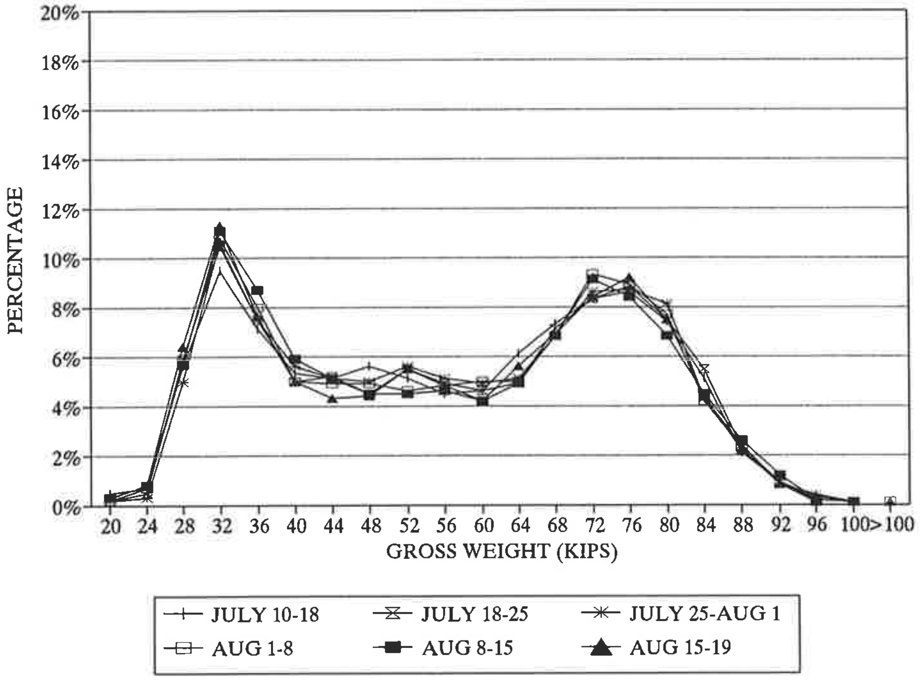


FIGURE 3 Distribution of gross weight of five-axle tractor-semitrailers on I-94: right lane, July and August.

low of 27.4 kips for trucks with grain boxes up to 34.3 kips for refrigerator trucks. The exact placement of the peak for empty trucks will depend on the type and mix of trucks that are in the traffic stream at the site. Generally, it will be close to 30 kips.

When a WIM system is functioning properly, the pattern of the distribution of the gross weight repeats from week to week. Figure 3 shows six consecutive weeks of data collected in the right lane on I-94. Each data set represents approximately 5,500 five-axle tractor-semitrailers. Note that this pattern is generally similar to that in Figure 1. They both show a distinct bimodal distribution with the peaks at approximately the same locations.

SYSTEM MALFUNCTION

Sometimes WIM systems malfunction, and incorrect weight data are recorded. Figure 4 shows an example of a valid set of weight data (October 22 through 29) compared with an invalid set (December 3 through 10). Each data set represents more than 5,000 five-axle tractor-semitrailers. Note the similar patterns between the October 22–29 data and the data in Figures 1 and 3. The data in Figure 4 are from the right lane at the I-94 site, the same location as shown in Figure 3. Note that the invalid data set in Figure 4 has the highest percentage in the lightest category; it does not have the classic bimodal distribution.

Other techniques can be used to determine the validity of the December 3–10 data in Figure 4. The first is to check the volume of five-axle tractor-semitrailers for both periods. If the volume changed significantly, the observed results could be a true reflection of the actual weights on the roadway. If the volumes did not change, the data are likely to be invalid. The volumes did not change significantly in the example used here.

Further supportive evidence that the data shown in Figure 4 are invalid is shown in Figure 5. These data are from the left lane at the same site on I-94. Figure 5 shows the same period as presented in Figure 4. The patterns of both sets of data in Figure 5 are the same. Consequently, for both sets of the data in Figure 4 to be considered valid, the patterns should be very similar. They are not. There should not be such vastly different patterns in lanes that are side by side traveling the same direction. On the basis of the analyses described, the presence of a pattern such as that shown in Figure 4 indicates the likelihood of a system malfunction and not a simple drift in calibration. It requires that a technician examine the system and correct the problem. A change in calibration will not solve the problem in this case.

CALIBRATION ERRORS

Figure 6 shows data from the WIM on I-94. Both of these data sets have the desired general pattern. The problem here is that the calibration is off by about 20 percent for the October 6–16 data set. The first peak is at about 38 kips instead of 32, and the second peak is at 90 kips instead of 74. The flexible ESAL factor for July 11 through 17 was 0.90, whereas it was 1.41 for October 6 through 16. The October 6–16 data are invalid and should not be used. The calibration should be adjusted by a factor determined from an analysis of the distribution of gross weight, front-axle weights, and flexible ESAL factors. A minimum of 7,500 five-axle tractor-semitrailers are represented in each data set here.

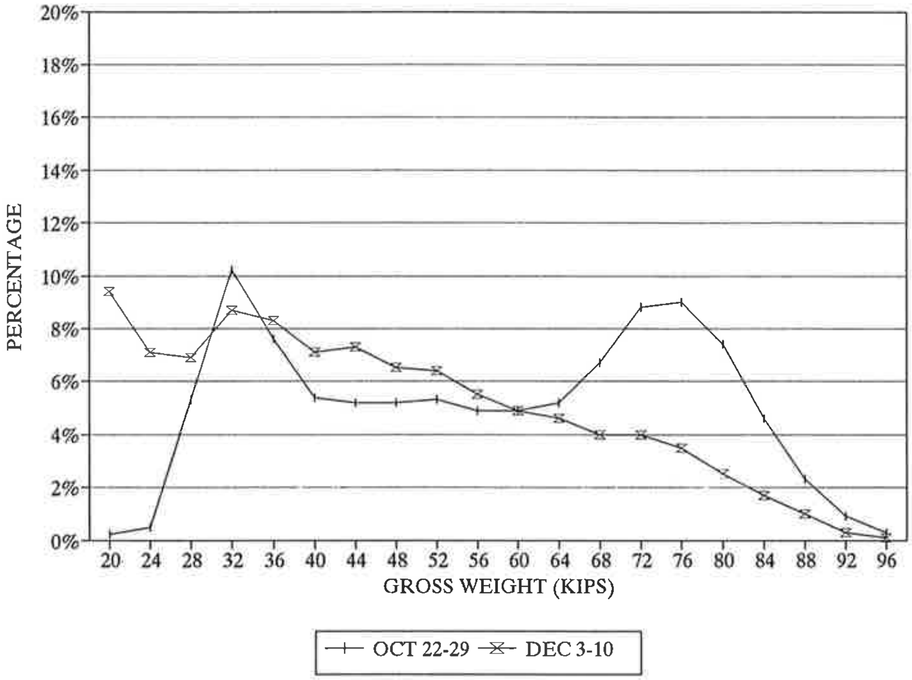


FIGURE 4 Distribution of gross weight of five-axle tractor-semitrailers on I-94: right lane, October and December.

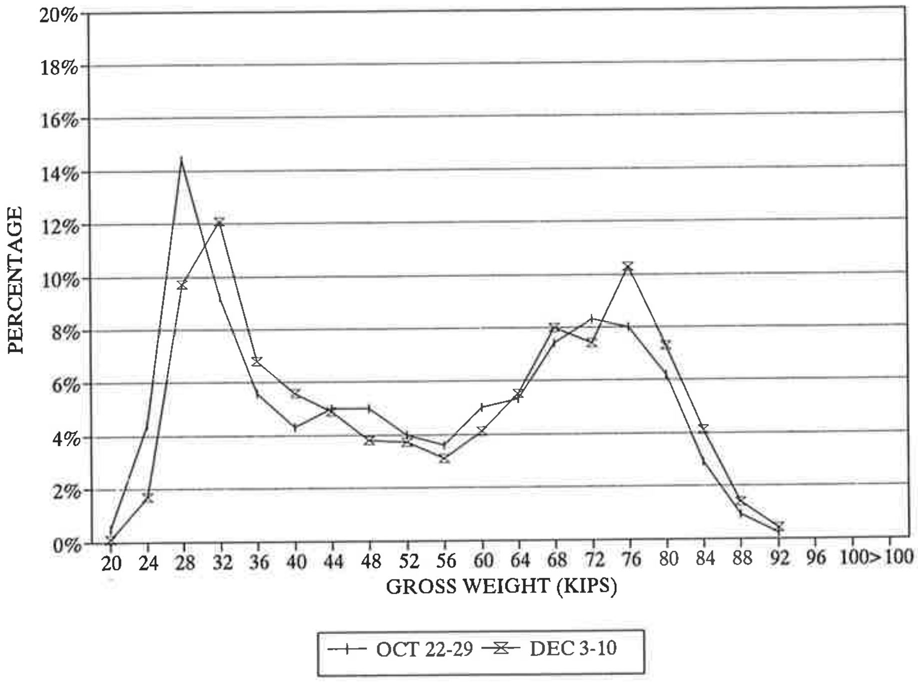


FIGURE 5 Distribution of gross weight of five-axle tractor-semitrailers on I-94: left lane, October and December.

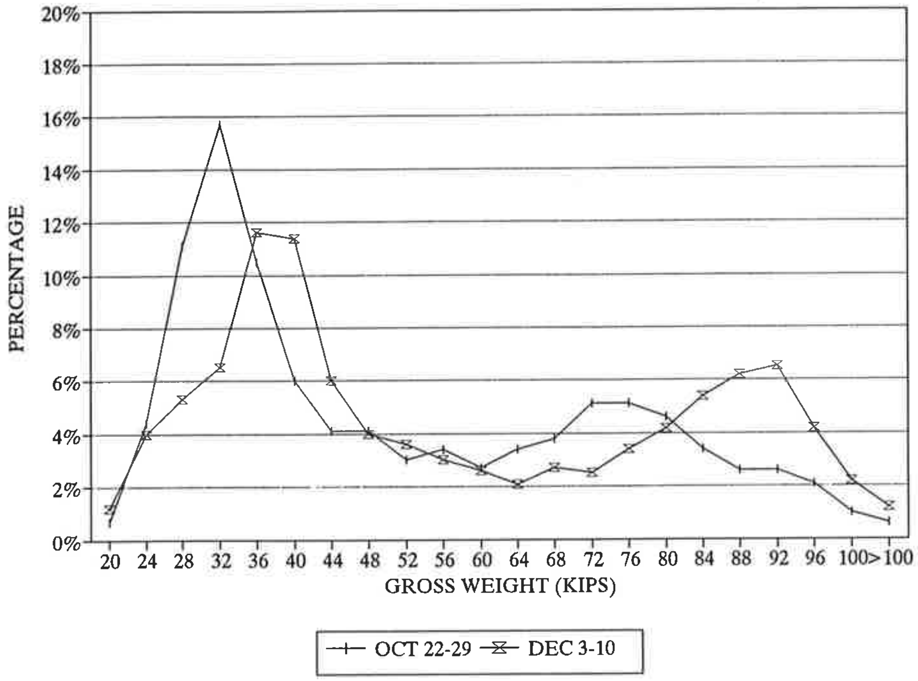


FIGURE 6 Distribution of gross weight of five-axle tractor-semitrailers on I-94: both eastbound lanes, October and December.

Also, referring to Figure 5, it appears that the calibration was off for the October 22–29 period. The peaks are slightly offset. The flexible ESAL factor for October 22 through 29 was 0.77, whereas it was 0.88 for December 3 through 10. The calibration should be adjusted by a factor determined from an analysis of the distribution of gross weight, front-axle weights, and flexible ESAL factors. Approximately 1,500 five-axle tractor-semitrailers are represented in each data set here.

FRONT-AXLE WEIGHTS

The second area that may be analyzed to determine the status of the calibration and the validity of the weight data is front-axle weights. The distribution of front-axle weights should repeat in a predictable manner at each site for a group of five-axle tractor-semitrailers. The distribution may be grouped into three categories, with the weights in kips.

Gross Weight Range	*Average Weight*
Less than 32.0	8.5
32.0–70.0	9.3
More than 70.0	10.4

These are average values that do not always apply to a specific site. However, the pattern that exists at each site should be consistent. The weights recorded for front axles vary to some degree depending on the vendor that produced the system.The important point here is not the weights that are noted but the pattern that is recorded by the system being studied. A minimum of 30 vehicles probably should be weighed in each group to be considered valid.

Figure 7 shows front-axle weights as recorded for individual weeks over an extended period. These data were also taken from the right lane on I-94. Note that each of the respective three weights were steady over an extended period. Then the weights began to show a large amount of deviation. It is interesting that the last value in January was lower than it should have been in the group of less than 32 kips and higher than it should have been in the group of more than 70 kips. The system was weighing some trucks too light and others too heavy. The weights in December indicate invalid data.

MnDOT's systems now contain an optional feature that provides an automatic system recalibration. This recalibration procedure is based on an aggregation of the observed weights of the front axles of five-axle tractor-semitrailers in the three weight ranges previously discussed. This procedure is being monitored to determine if it performs satisfactorily.

FLEXIBLE ESAL FACTORS

The third data set that may be analyzed to determine the status of the calibration and the validity of the weight data is flexible ESAL factors. Either rigid or flexible factors can be used in the evaluation. As was mentioned, these factors are sensitive to changes in the weights because of the fourth-power relationship between weight and ESALs. The approximate range for these values must be determined for each lane. These values are obtained by noting what the system produces after it has been properly calibrated.

Generally the ESAL factors for each individual day—particularly weekdays—are examined. Weekday ESALs tend to be quite stable. That is not necessarily true for weekends. Figure 8 shows weekday data from October to mid-December for the right lane on I-94. Data from this lane are shown in

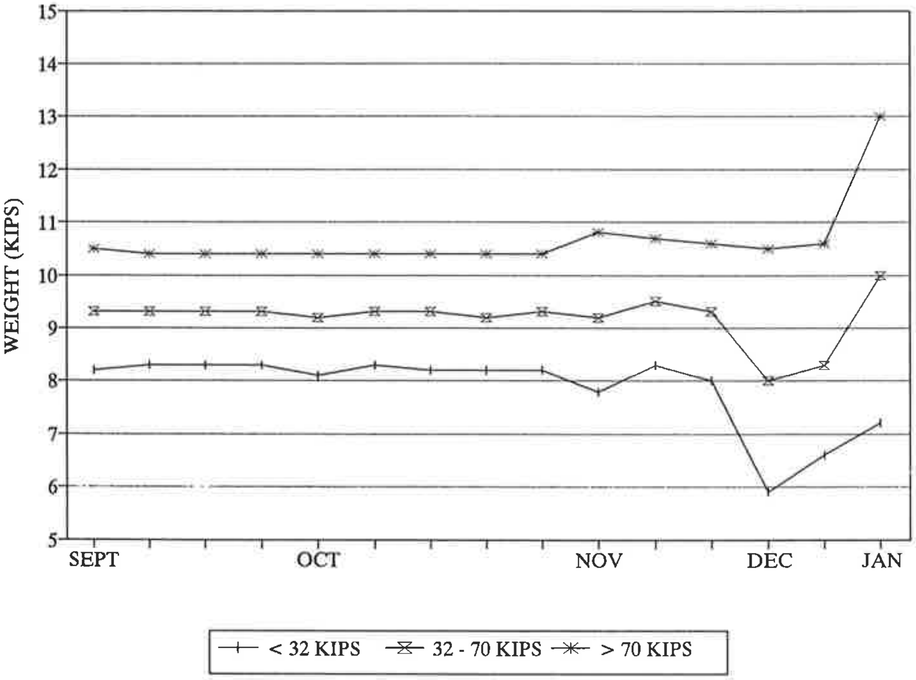


FIGURE 7 Front-axle weights of five-axle tractor-semitrailers on I-94: right lane, by gross weight range.

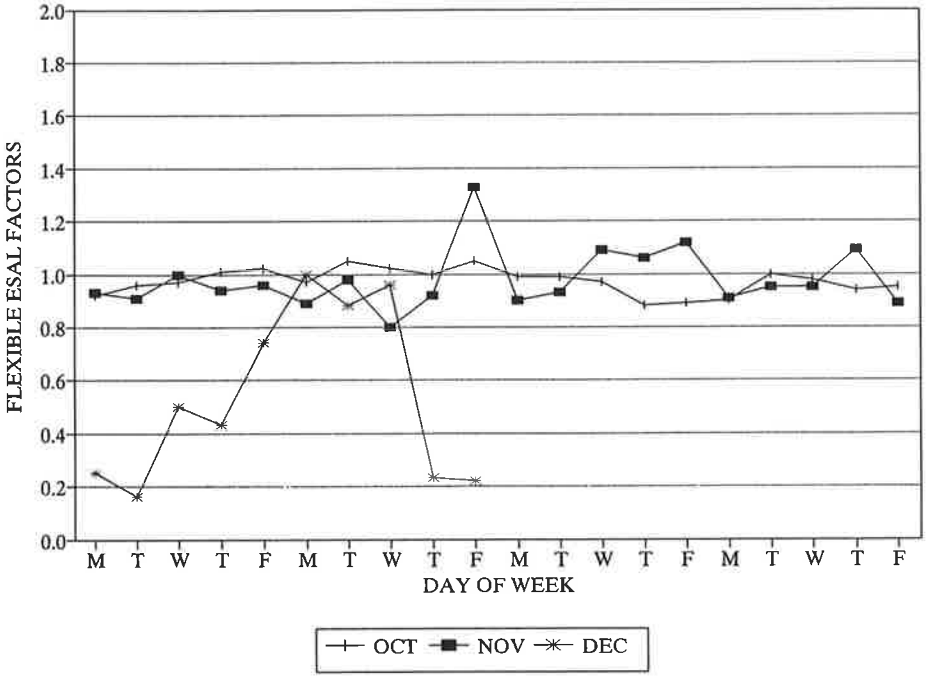


FIGURE 8 ESAL factors for five-axle tractor-semitrailers on I-94: right lane.

Figures 3, 4, and 7. It becomes evident after viewing these data that the valid factors should be in the area of 0.90 to 1.05. Approximately 900 five-axle tractor-semitrailers per day are represented here.

Figure 9 shows weekday data for the left lane on I-94. These data cover the same period as those shown in Figure 8. The pattern is consistent and generally falls in the 0.70 to 0.90 range. There are no drastic deviations as observed in Figure 8. These data are valid, as was also determined for Figure 6. Approximately 200 trucks a day are represented here.

SUMMARY OF PROCEDURE AFTER INITIAL CALIBRATION

The following steps outline how to conduct an analysis of WIM data to determine the status of the calibration and the

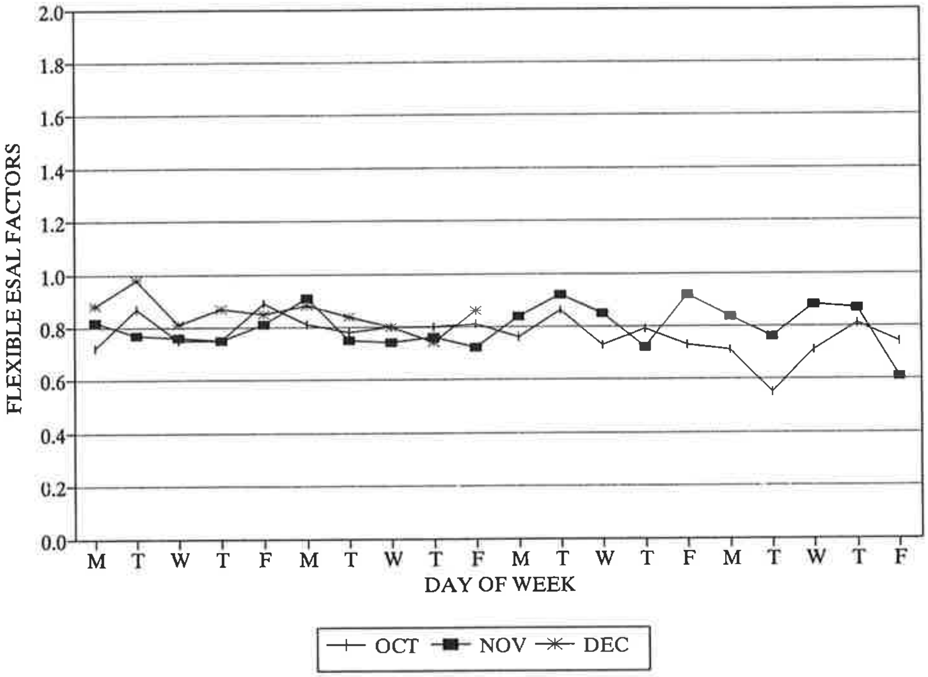


FIGURE 9 ESAL factors for five-axle tractor-semitrailers on I-94: left lane.

validity of the data. They also suggest an application in the process of editing weight data. The analysis examines the distribution of gross weight, front-axle weight, and flexible ESAL factors for five-axle tractor-semitrailers. It also contains guidelines for determining whether the weight data are valid. When an analysis indicates that the calibration is incorrect, a recalibration should be done on the basis of the percentage indicated by the outlined procedures. When there is a system malfunction, recalibration should not be done. The system should be examined and repairs should be made. These procedures follow the initial calibration of the system. The application is on an individual-lane basis.

1. Determine whether the distribution of the gross weight is logical. The first peak in the bimodal distribution should be between 28 and 32 kips, and the second peak should be between 70 and 80 kips. Many states have legal limits that are greater than 80 kips; this means that the second peak will occur at a different location, but the first peak should remain at about 30 kips. If the peaks are where they belong, the system is properly calibrated.

If a shift in the peaks has taken place, determine if the volume of trucks has changed appreciably from the number that regularly uses that lane. If the volume did change, it might explain the shift in distribution of weight. If the number has not changed, and if both peaks have shifted in the same direction and they are off by 4 percent or more, the data should be considered invalid. This indicates the need for recalibration. If both peaks are off by 4 percent or more and they have shifted in opposite directions, the data should also be considered invalid. In this second circumstance, do not consider recalibration, because such data indicate a system malfunction.

2. Determine the proper front-axle weights for each of the three gross weight categories. These are the values that the system produces, assuming that the basic calibration check described in Step 1 has been completed. Then monitor the values in each of the three categories. If a change is occurring, check the volume of vehicles passing through the system as described in Step 1 to determine if that might be the cause of the change.

If the weights in at least two of the three categories are shifting in the same direction and they pass beyond a specified point—for example, 4 percent—the system should be considered to be out of calibration and the data collected during this period invalid. This shift in weights indicates the need for recalibration. Also, if one value shifts 4 percent in one direction and another value shifts 4 percent in the other direction, the data are invalid. Do not consider recalibration because this type of action indicates a system malfunction.

3. Determine the average flexible (or rigid) ESAL factor for weekdays by using the values calculated from weight data collected by the properly calibrated system. Then set up the acceptable range of factors. Coordinate this range with what was deemed acceptable in the other two areas being monitored. For example, when 4 percent is used in those areas, approximately 16 percent should be used here. If values are observed outside the range, check to see if the volume of vehicles changed as described in Step 1. The data are invalid for any days that fall outside those values. For those stations collecting either continuous or a significant amount of weight data, assume that Saturday and Sunday are valid as long as the adjacent weekdays are valid.

4. Compile the results of the analysis of the three measures discussed. Determine whether the data are valid. If they are valid, continue to collect data with the system and use the data in the desired analysis. If they are invalid, the system requires either recalibration or repair. Recalibrate the system if at least two of the three areas indicate that this is needed, or repair the system if at least two of the three areas indicate that this is needed. These three measures all complement one another. They serve as cross checks in determining whether the weight data are valid.

5. Apply the results of the analysis to the WIM site data editing process. This step is a vital supplement to standard editing techniques that focus on the validity of individual vehicle records. Such standard editing techniques cannot detect the types of problems that have been discussed here.

CONCLUSIONS

There is a practical, effective method that can be used to calibrate a WIM system and monitor that calibration over time. This method uses predictable patterns of five-axle tractor-semitrailer weight data. The distribution of the gross weight data provides the best vantage point to determine its overall validity. This is a key part of the process. The other two measures—front-axle weight and flexible ESAL factors—are also important. Any one or two of these items can function independently, but they work best when used together. These indicators can also be used to identify and edit out invalid weight data.

ACKNOWLEDGMENT

The author gratefully acknowledges the contributions made to this report by Roberta Casey, George Cepress, Mark Flinner, and Morris Gildemeister of MnDOT.

Publication of this paper sponsored by Committee on Vehicle Counting, Classification, and Weigh-in-Motion Systems.

Applications of Weigh-in-Motion Data in Transportation Planning

JERRY J. HAJEK, GERHARD KENNEPOHL, AND JOHN R. BILLING

Participation in the Strategic Highway Research Program (SHRP) and Canadian-SHRP (C-SHRP) long-term pavement performance monitoring programs has prompted all state and provincial highway agencies in the United States and Canada to install at least one weigh-in-motion (WIM) scale. The WIM scales provide not only the traditional traffic monitoring data but also a lot of detailed data for individual vehicles, including length, speed, axle weights, and axle spacing. Because the WIM data are often associated only with pavement research, many potential users are not yet aware of (a) the capabilities of WIM technology, (b) the type of data available, and (c) the opportunities to use WIM data in their areas of expertise. Traffic monitoring data generated by typical WIM installations are described. A number of practical examples show that WIM data are useful for a wide range of transportation planning and decision-making purposes. The specific application areas discussed include (a) planning and programming of transportation facilities, (b) pavement design and rehabilitation, (c) apportionment of pavement damage, (d) compliance with vehicle weight regulations, (e) development of geometric design standards, (f) compliance and regulatory policy development of truck dimensions, (g) safety analysis, (h) traffic operation and control, and (i) analysis related to highway bridges. WIM data can be used in all these application areas. Their usefulness cuts right across the organizational structure of any highway agency. Thus, WIM data should be considered corporate data and should be managed accordingly.

During the past few years, there has been a rapid and widespread introduction of weigh-in-motion (WIM) technology and automatic vehicle classification (AVC) technology in the United States and Canada. The main impetus for this has been the participation in the Strategic Highway Research Program (SHRP) and Canadian SHRP (C-SHRP) long-term pavement performance monitoring programs, which mandate a detailed automatic monitoring of traffic loads. Already, all major state and provincial highway agencies have installed, or are in the process of installing, at least one WIM scale. The Ontario Ministry of Transportation (MTO) experimented with the earliest WIM equipment in the mid-1970s and has been operating at least one WIM scale more or less continuously since the early 1980s (*1*). MTO is currently operating four in-highway WIM scales, six WIM sorter scales at truck inspection stations (TISs), and, to comply with the SHRP and C-SHRP guidelines, is planning to install three more WIM scales in 1991.

The existing and planned WIM scales provide a large amount of detailed traffic monitoring data for individual highway vehicles, such as axle spacing and axle weights, and vehicle length and speed. This is in addition to the traditional aggregated traffic characteristics, such as hourly and daily vehicle volumes. Installing and operating WIM scales, and their associated traffic data retrieval and analysis, are neither easy nor inexpensive. It is desirable to ensure that the wealth of traffic monitoring data generated by WIM scales is used properly.

It is often assumed that WIM data are applicable only to pavement research, because of their association with the SHRP and C-SHRP pavement research effort. As a result, many potential users of WIM-type traffic monitoring data do not know

1. The data monitoring capabilities of WIM technology,
2. The type of data available, and
3. How the data can be used within their area of interest.

The objective of this paper is to show, by practical examples, that WIM data are useful for a wide range of transportation planning and decision-making purposes. The treatment of any individual area is brief, by necessity, to provide a comprehensive overview of all the main areas of application.

In this report, axle and gross vehicle weights (GVW) results are presented in kilograms, axle spacings in meters, and vehicle speeds in kilometers per hour—the units in which Ontario regulates vehicle loads and dimensions.

DESCRIPTION OF TRAFFIC MONITORING DATA PROVIDED BY WIM SCALES

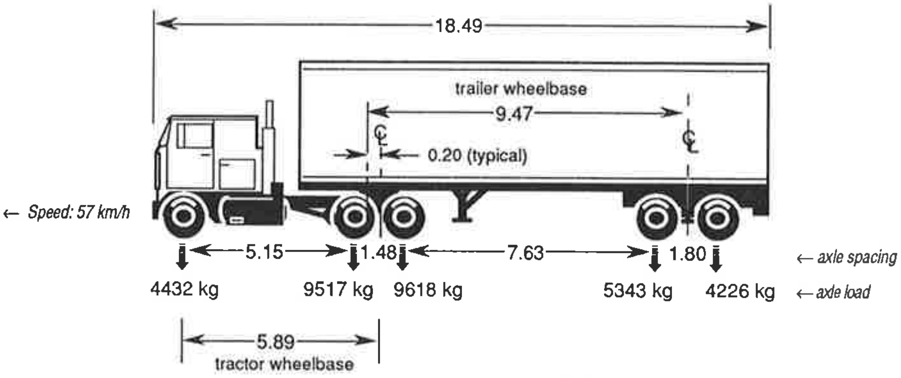
A typical WIM scale consists of magnetic loops and axle sensors embedded in the pavement and a microcomputer housed in a roadside cabinet. Magnetic loops and axle sensors respond to axles passing over the pavement by generating electric signals. The signals are processed by the computer and are transformed into engineering parameters including instantaneous vehicle speed, vehicle length, distances between consecutive axles, and axle weights. An annotated example of an individual vehicle record provided by a WIM scale is shown in Figure 1.

The majority of data used in this study were obtained from two WIM scales, one installed on Highway 7N and the other on Highway 402. The Highway 7N scale is in an eastbound truck lane of a six-lane suburban (metropolitan Toronto) arterial road and uses bending plate technology (*2*). The Highway 402 scale is in an eastbound truck lane of a four-lane rural freeway (near Sarnia) with a speed limit of 100 km/hr;

J. J. Hajek and G. Kennepohl, Pavements and Roadway Section, Research and Development Branch, and J. R. Billing, Vehicle Technology Office, Transportation Technology and Energy Branch, Ontario Ministry of Transportation, 1201 Wilson Avenue, Downsview, Ontario, Canada M3M 1J8.

```
(910)LANE £3  TYPE 7  GVW 33.1 tonnes   LENGTH 1849 cm
18-K ESAL 4.058 SPEED 57 km/h Mon Jan 15 08:54:46 1990

               UNIT       SEPARATION       WEIGHT
                             (cm)           (kg)
                1                           4432
                2             515           9517
                3             148           9618
                4             763           5343
                5             180           4226
```



tractor wheelbase: 5.15 + (1.48/2) = 5.89 trailer wheelbase: (1.48/2)+ 0.20 + 7.63 + (1.80/2)= 9.47

Note: All dimensions are in metres

FIGURE 1 Individual truck record from International Road Dynamics, Inc., WIM scale and its interpretation.

it uses piezoelectric cable technology (*3*). Compared with static conditions, the scales provide dimensions accurate within 2 to 3 percent, GVWs within about 5 percent, and axle loads within about 5 to 12 percent. Accuracy depends on vehicle dynamics (e.g., vehicle configuration and speed) and pavement roughness in the vicinity of the scale.

The WIM computer can store a record for each vehicle passing over the scale, as well as some selected summary data. These data can be verified, edited, and analyzed to meet the various application requirements.

Common to all WIM data applications is the need to classify individual vehicles into distinct categories. An example of the basic MTO vehicle classification scheme is given in Table 1. The scheme resembles that used by the FHWA *Traffic Monitoring Guide* (*4*). However, unlike the FHWA guide, it attempts to classify nonpassenger vehicles not only according to number of axles and number of connected units, but also according to axle configuration. Thus, for example, there is a specific category for 3S2 trucks (Category 7 in Figure 2).

Although the Ontario scheme recognizes 14 highway vehicle classes, some applications discussed later require a more specific or detailed classification scheme. Figure 2 shows one method for identifying six-axle tractor-semitrailers that are loaded close to their allowable gross weight. The semitrailer is of a triaxle design, having a fixed dual axle unit at the rear and a liftable "belly axle" usually at least 2.54 m ahead of it. The scheme relies on knowledge of the impact of vehicle weight and dimension regulations on truck design. This was done by constructing a "filter" that specifies limits on various truck weight and dimension parameters so that only those trucks of interest are selected.

Considering the variety and the large samples of vehicles analyzed, there must be a certain level of speculation associated with any vehicle classification scheme based on WIM data. There is no reason to believe that these uncertainties have a significant effect on the observations made from the results or on the more global conclusions presented herein.

USE OF WIM DATA

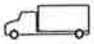
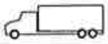
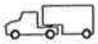
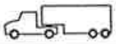
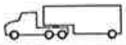
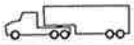
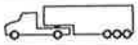
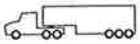
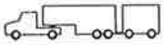
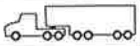
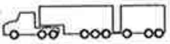
To enhance the understanding of and to clarify the various ways in which WIM data can be used for transportation planning and decision making, the following discussion is divided into nine general application areas, some of which may overlap.

Planning and Programming of Transportation Facilities

Any WIM scale installed on a highway is, basically, a continuously operating traffic counting station. It can generate not only traditional traffic volume characteristics, such as the annual average daily traffic (AADT) and the 30th highest hourly volume used for planning and programming of transportation facilities, but also a wealth of additional detailed data for individual highway vehicles.

When planning the installation of new permanent traffic counters, or the replacement of existing ones, consideration should be given to upgrading the counters into WIM scales. The cost of upgrading is mainly due to higher-quality axle sensors and a more powerful computer in the roadside cabinet. Installation and operating costs, such as highway closure during installation, provision of power and communication lines at the site, system maintenance, and data retrieval ac-

TABLE 1 Ontario Vehicle Classification Scheme

Class	Type	Description
1	PVLC 	**Passenger cars** 2-axle vehicles with spacing from 1.2 to 3.9 m
2	2 	**2-axle trucks** axle spacing from 3.9 to 15.0 m
3	3 	**3-axle trucks** with a dual axle spacing up to 2.4 m
4	UD-3 	**All other 3-axle trucks**
5	2S2 	**4-axle trucks** 2-axle tractor with one dual axle semitrailer dual axle spacing up to 2.4 m largest tractor and semitrailer axle spacing from 2.4 to 15.0 m
6	UD-4 	**All other 4-axle trucks**
7	3S2 	**5-axle trucks** 3-axle tractor with one dual axle semitrailer dual axle spacing up to 2.4 m largest tractor and semitrailer axle spacing from 2.4 to 15.0 m
8	UD-5 	**All other 5-axle trucks**
9	3S3 	**6-axle trucks** 3-axle tractor with a 3-axle semitrailer tractor dual axle spacing up to 2.4 m trailer axle spacing up to 2.6 m largest tractor and semitrailer axle spacing from 2.4 to 15.0 m
10	UD-6 	**All other 6-axle trucks**
11	3S2-2 	**7-axle trucks** 3-axle tractor with two semitrailers or with one semitrailer and one trailer; the first semitrailer has a dual axle tractor dual axle spacing up to 2.4 m largest tractor axle spacing up to 5.0 m first semitrailer dual axle spacing up to 3.0 m largest 1st and 2nd trailer axle spacing from 5.2 to 15.0 m
12	UD-7 	**All other 7-axle trucks**
13	UD-8 	**All 8-axle trucks**
14	UD-9 	**All trucks with 9 or more axles**

count for the bulk of the total cost. They are not very influenced by the type of traffic counting device. The benefits of upgrading can be significant.

Pavement Design and Rehabilitation

WIM data have been closely associated with pavement design because traffic loads constitute the basic pavement design parameter (*5*). For pavement design and evaluation purposes, the effect of traffic loads on pavement structural damage has been traditionally expressed using the concept of load equivalency factors (LEFs). The equivalency factors equate the damaging effect of any given axle, or axle combination, to that of a standard axle. For convenience, LEFs have been related to the standard axle load of 8160 kg (18,000 lb) carried on a single axle with dual tires, called the equivalent single axle load (ESAL).

Understanding of the number of ESALs associated with different truck types, in different highway corridors and during different times is now shaped mainly by WIM data. As an example of the usefulness of WIM data, Table 2 compares the average ESAL per truck obtained from WIM scales with that estimated by the OPAC truck prediction subsystem (*6*). The latter is currently used for design and was established before the advent of WIM scales. The existing prediction methodology is seen to underestimate the average ESALs per truck by a factor of 2 or more. The ESALs reported herein are related to flexible pavements with a structural number equivalent to 5 (*5*).

Whereas Table 2 shows the average ESALs per truck, a rather global measure of pavement traffic load, WIM technology enables changes in ESAL for specific vehicle types during specific periods to be evaluated. For example, Figure 3 illustrates a year-long variation in average weekly ESALs for 3S2 tractor-semitrailers on Highway 7N. (For definition of 3S2 trucks, refer to Figure 2.)

The WIM data can provide the main building block for developing procedures for predicting ESALs for pavement structural design. What is required, basically, is to obtain

Criteria for Truck Definition

(a) 6-axle trucks
(b) single (steering) - dual (tractor) - single (liftable) - dual (trailer) axle arrangement
(c) axle 2-3 spacing from 1.07 to 1.83 m
(d) axle 3-4 spacing greater than 4.00 m
(e) axle 4-5 spacing greater than 2.40 m
(f) axle 4-5 spacing greater than axle 5-6 spacing
(g) axle 5-6 spacing from 1.07 to 3.05 m
(h) gross weight within 1000 kg of allowable load

Explanation

The first item simply ensures the proper number of axles. The second is descriptive, and may be redundant. Item (c) covers the known range of drive axle spreads. Item (d) ensures that the filter captures semitrailers at least 10 m (32 ft) long, so it will exclude tractors pulling 7 m (23 ft) tridem container chassis. Items (e) and (f) ensure the liftable axle is properly separated from the trailer tandem axle, according to either Quebec, Ontario, or Michigan regulations. Item (g) covers the known range of trailer axle spreads. The final item ensures the gross weight is close to the allowable limit.

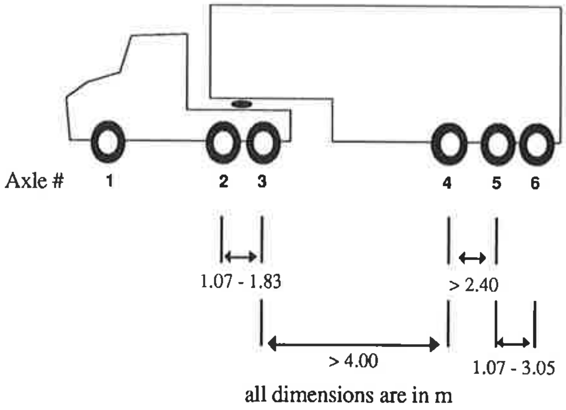


FIGURE 2 Criteria identifying six-axle tractor-semitrailers with liftable axle.

TABLE 2 Comparison of Average ESALs per Truck (5)

Location Facility type	Average ESAL per truck	
	Based on WIM Measurements	OPAC Estimate
Hwy. 7N Urban arterial road	1.40	0.40
Hwy. 402 Rural freeway	1.60	0.67

volume estimates for major truck categories and their corresponding average ESALs. Detailed step-by-step procedures are readily available.

Apportionment of Pavement Damage

In many jurisdictions there is a general tendency to set truck license fees in proportion to the highway resources (pavement damage) consumed. The WIM scales can quantify pavement

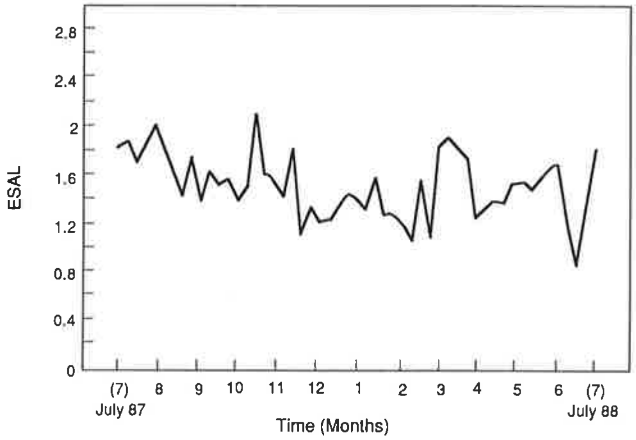


FIGURE 3 Variation in weekly ESALs for 3S2 tractor-semitrailers: Highway 7N, Lane 3.

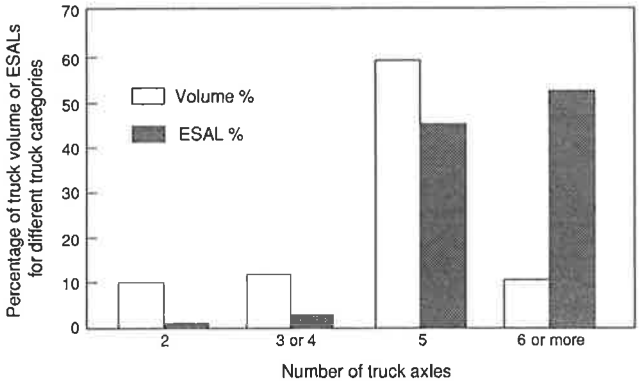


FIGURE 4 Pavement damage contribution by different truck categories: Highway 402, 1989.

damage contributions of various truck classes, as shown in Figure 4 for Highway 402. For example, trucks with six or more axles constitute less than 20 percent of the total truck volume but are seen to be responsible for more than 50 percent of pavement damage.

Compliance with Vehicle Weight Regulations

In-Highway WIM Scales

WIM scales can provide realistic and detailed insight into vehicle axle and gross overloads. The assessment is realistic because WIM scales can operate continuously and they are unobtrusive. It is detailed because by generating weight and dimension records for individual highway vehicles, they enable assessment of overloads during a specific time period and for a specific truck category. WIM scales do not directly provide a valid estimate of violation rates because of the impact of vehicle dynamics and other factors on weight estimates. Although they cannot be used at this time for vehicle weight enforcement, WIM scales are sufficiently accurate to identify substantial deviations from regulatory limits.

An example result in Figure 5 shows the compliance with weight regulations for liftable axles on six-axle tractor-semitrailers with a dual-axle unit at the rear and a liftable belly axle at least 2.54 m ahead of it. This truck type is the most common among the trucks with liftable axles and is also the most prone to axle overloads.

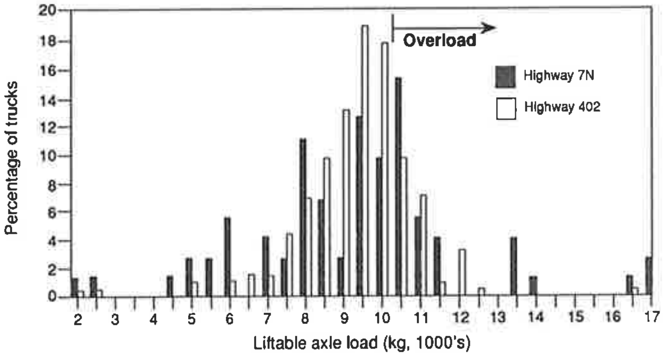


FIGURE 5 Liftable axle load distribution for selected group of six-axle tractor-semitrailers.

The trucks were identified using the filter defined in Table 1 that was applied to individual vehicle records obtained by the Highway 402 and Highway 7N WIM scales during 1 month in 1990. The same data set is used for the other examples in this paper, unless otherwise stated. The filter produced a sample of 72 trucks from the Highway 7N WIM scale and 184 trucks from the Highway 402 scale. The small numbers arise because the range of GVW of the selected trucks is very narrow (condition in filter of Table 1), ranging from about 49 000 to about 55 500 kg.

A few selected observations, related only to liftable axles and based mainly on Figure 6, follow. [A more comprehensive discussion of GVW and axle overloads for six-axle tractor-semitrailers with liftable axles is available (*7*). This literature also describes weight compliance observed for five- and six-axle semitrailers without liftable axles.]

- Loads on the liftable axles ranged from 2000 to 17 000 kg for each highway. The allowable load limit on the liftable axle, considering that its distance from the surrounding axles is more than 2.5 m, is 10 000 kg.
- There was somewhat better load compliance on Highway 402 than on Highway 7N. About 21 percent of liftable axles were overloaded on Highway 402 compared with about 37 percent on Highway 7N. This may be partly because of the proximity of the Highway 7N WIM scale to intersections, at which turning trucks would be raising the liftable axle in preparation for turning or after making a turn, and partly because the Highway 402 WIM scale is close to a TIS. The liftable axle overloads apply only to a selected group of trucks with the GVW close to the allowable limit (Table 1, Criterion h).
- Because the GVW of all trucks investigated was within 1000 kg of their allowable (total) GVW, when the loads on the liftable axles were significantly lower than their allowable load—say, less that 8000 kg—the slack caused by these underloaded liftable axles had to be compensated for by overloading the surrounding axles. About 20 percent of liftable axles were underloaded on Highway 7N.

Analyses of this type, based on statistically relevant samples obtained by WIM scales, are far removed from simplistic conclusions such as "10 percent of trucks are overloaded."

WIM Scales near TISs

WIM scales near TISs can be used to evaluate the effectiveness of weight enforcement programs and help in developing cost-effective enforcement strategies. This can be accomplished, for example, by comparing compliance rates, determined by a continuously operating WIM scale, achieved by various compliance enforcement strategies employed at the TIS (e.g., an intensive versus a spotty enforcement strategy).

WIM as Sorter Scale

To facilitate the enforcement of weight regulations, TIS on high-truck-volume highways have been equipped with WIM scales. These scales are referred to as "sorter scales" because

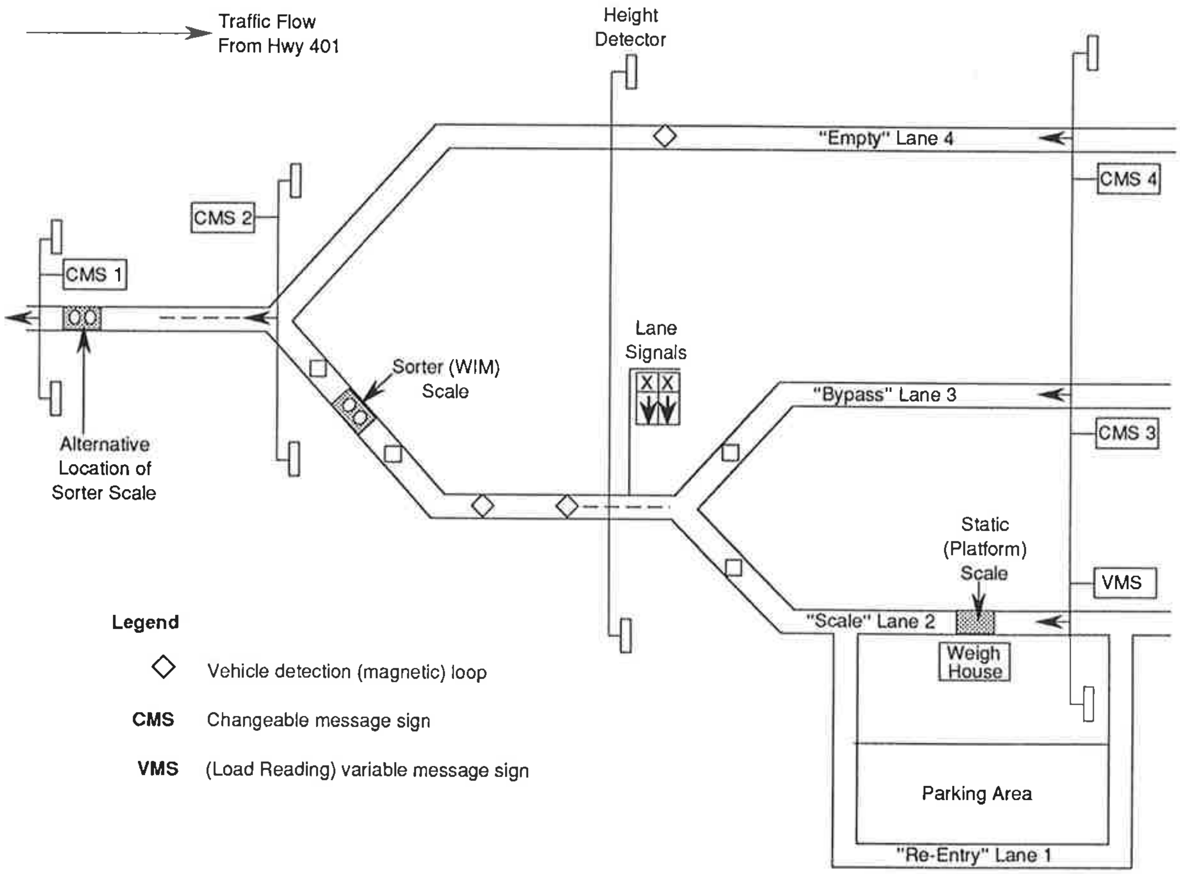


FIGURE 6 **Schematic layout of TIS with sorter scale.**

they are used to sort incoming trucks into two categories: (a) those that clearly comply with weight regulations, and (b) those that should be weighed at a static scale. Figure 6 shows a schematic layout of a TIS incorporating a sorter scale. Upon entering the TIS area, trucks with no payload can be directed to bypass both the sorter scale and the static scale. When this option is used, the weight data obtained from the sorter scale are no longer representative of the entire truck population and cannot be used for certain applications. Whenever feasible, the sorter scale should be used to weigh all trucks entering the TIS. Even in this case the data can be biased: the proclivity of trucks to avoid TIS is well known.

Development of Geometric Design Standards

The rational development of highway geometric design standards (turning radii and width of traffic lanes at intersections, dimensions of loading and unloading facilities, design of truck climbing lanes, etc.) requires the knowledge of vehicle weights and dimensions. In particular, it requires knowledge of the dimensions of the most common, and from the space requirement aspect, one of the most demanding, large truck configurations—the five-axle tractor-semitrailer. The realistic assessment of dimensions of existing vehicles requires a large truck population sample, which can easily be provided by WIM scales. The use of WIM data is illustrated for one of the key parameters governing the space requirements for turning movements of a 3S2 truck—its semitrailer wheelbase. A detailed analysis and interpretation of the observed dimensions of five-axle tractor-semitrailers on geometric standard development, in terms of the overall truck length, tractor and semitrailer wheelbase, and the drive and trailer axle spreads, is given elsewhere (*7*).

The relationship between semitrailer wheelbase (the distance from its kingpin to the center of its rear dual axle; Figure 1) and the maximum low-speed offtracking offset (the distance between the turning radii of the steering axle and the rear axle) is shown in Figure 7; Figure 8 shows the observed trailer wheelbase distribution. Some of the observations on trailer wheelbase distribution, as they apply to the development of geometric standards, are summarized as follows.

- On Highway 7N, the distribution appears to be bimodal. The first peak, about 8.8 m (29 ft), is believed to be associated primarily with a 40-ft container chassis. The second peak, about 10.7 m (35 ft), is more likely to be associated with older 13.7-m (45-ft) semitrailers now relegated to local uses. On Highway 402, the predominant wheelbase is about 11.9 m (39 ft), which is typical of a 14.7-m (48-ft) semitrailer now used for long-distance hauling.
- The maximum observed semitrailer wheelbase was about 13 m on Highway 402 (Figure 8). This wheelbase results in an offtracking of about 7 m (Figure 7). However, considering the generally shorter wheelbase distances observed for local Highway 7N traffic and the longer wheelbase distances ob-

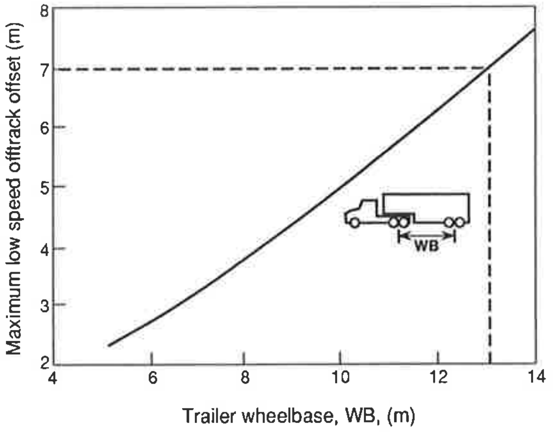


FIGURE 7 Maximum offset during low-speed turns (*8*).

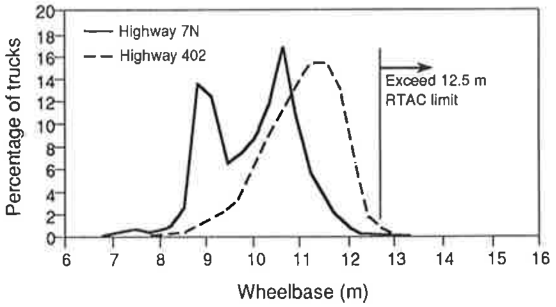


FIGURE 8 Trailer wheelbase distribution for five-axle tractor-semitrailers.

served for long-distance Highway 402 traffic, the geometric standards need not be applied uniformly to all transportation facilities.

Truck Dimensions: Compliance and Regulatory Policy Development

The knowledge of existing truck dimensions can be used to assess the consequences of regulatory changes in truck dimensions, the need to except the existing truck configurations for the present (grandfathering), or both. This issue is connected with the development of geometric design standards. However, whereas the development of geometric design standards tends to simply reflect the existing situation, exploration of the policy development issues is more proactive. For example, the following observations may be formulated when considering the results of the semitrailer wheelbase distribution given in Figure 8, from a compliance perspective:

- The Canadian *Memorandum of Understanding on Interprovincial Heavy Vehicle Weights and Dimensions* (*9*) establishes an upper limit of 12.5 m for a semitrailer wheelbase. Twenty-eight trucks on Highway 7N (0.39 percent) exceeded this limit, compared with 194 (0.49 percent) on Highway 402. Most of these exceeded the 12.5-m limit by no more than 0.6 m.
- Some of the problem trucks may be carrying a long, indivisible load and operating under permit. It would be useful to know "who" such trucks are, particularly in terms of their body style and commodity. However, the WIM data alone cannot provide such detailed information.

Safety Analysis

Safety emerges as a major issue in all debates or polemics about changes in vehicle weights and dimensions and, invariably, it is concluded that adequate information about safety implications of the proposed changes is lacking (*10*). WIM data can contribute to analysis of safety issues by providing detailed information on vehicle types using highway facilities and vehicle behavior on highway facilities.

Vehicle Types Using Highway Facilities

The knowledge of mileage traveled by various vehicle types (exposure rate) is a prerequisite for evaluating their accident rates. The accident rate (number of accidents divided by exposure) is instrumental in identifying the influence of vehicle design parameters on vehicle safety.

Trucks are often registered in several jurisdictions or in a jurisdiction other than the one in question. It is, therefore, difficult to estimate mileage traveled by different truck classes and, thus, to obtain accident rates for different truck types. WIM data can help in establishing truck exposure measures, particularly for facilities in which WIM scales have been installed. For example, referring to Figure 4, Truck Type 3S2 (defined in Table 1) composes about 55 percent of the total truck volume on Highway 402. The volume percentage of 3S2s can be directly related to the percentage of accidents involving the 3S2s on this facility.

Another descriptive parameter useful in accident studies (and provided by WIM data) is GVW. It has been observed that the accident rates of unloaded trucks are significantly higher than those for loaded trucks (*11*).

Vehicle Behavior on Highway Facilities

The unobtrusive presence of WIM scales can provide a reliable description of vehicle driving patterns and enable rational assessment of truck driving behavior. In this summary report, driving behavior is described using only simple frequency distribution functions for vehicle speed and headway. It is certainly possible to study more complex functions, such as the relationship between vehicle speed, headway, GVW, and time of day for different vehicle categories, and to provide data to develop and manage police enforcement strategies.

Vehicle Speed Distribution WIM scales routinely provide instantaneous vehicle speeds (Figure 1). Excessive vehicle speed, and particularly speed differentials between different vehicles, is considered to be a main cause of accidents. Overloaded and speeding trucks may constitute an additional safety hazard.

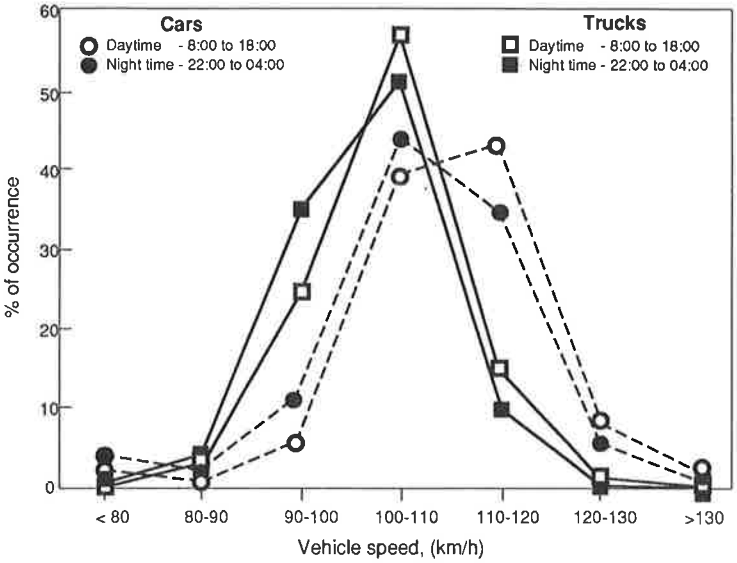


FIGURE 9 Vehicle speed distribution: Highway 402, truck lane, March 19–22, 1991.

Figure 9 shows vehicle speed distribution for cars and trucks, during the day and at night, derived from the Highway 402 WIM scale data. Data were obtained for four consecutive weekdays, without any precipitation, in March 1991. The WIM scale is in one of the truck lanes of this four-lane rural freeway, and low traffic volumes (about 300 vehicles per hour in the WIM lane during daytime and 75 vehicles per hour at night) enable a large degree of traffic operational freedom.

Overall, data in Figure 9 indicate that truck drivers are more disciplined than car drivers. Some specific observations follow:

- Most cars were speeding. The speed limit on this facility is 100 km/hr (about 60 mph). During daytime, about 53 percent of all cars exceeded 110 km/hr, whereas at night 42 percent of all cars exceeded this speed. The corresponding numbers for trucks were 16 and 10 percent, respectively.
- Compared with cars, truck speed distribution is more uniform. Looking at the extremes, during daytime, 1.3 percent of cars had speeds less than 80 km/hr compared with only 0.3 percent of trucks. At the high end, 1.3 percent of cars (in the truck lane) exceeded the speed of 130 km/hr compared with 0.1 percent of trucks.

Headway Distribution According to Ontario's *Highway Traffic Act* (*12*), maintaining "reasonable and prudent" headway is mandatory for all drivers. There is an extra stipulation for drivers of commercial vehicles who, while driving at speeds exceeding 60 km/hr, "shall not follow within 60 metres of another motor vehicle."

WIM scales routinely provide a time stamp, truncated to the nearest second, for all individual vehicle records. Only this routine time measurement precision was available for data used in this exploratory study. A more focused study of headway distribution would require software modification that gives time measurements in tenths of a second.

Figure 10 compares the difference in headway distributions of cars and trucks. The figure uses the same data set as that used for Figure 9. The greater discipline of truck drivers, indicated by the speed distribution, is also indicated by the headway distribution. Some observations follow:

- During daytime, 7 percent of all cars followed other cars with a 1-sec headway, whereas only 2.5 percent of trucks did so. Nevertheless, considering an average truck speed of 100 km/hr or 27.8 m/sec, more than 2.5 percent of all trucks appear to be in violation of the Ontario Highway Traffic Act headway requirement.
- A total of 3.5 percent of all trucks were following other trucks with a headway of 1 sec, whereas only 2.5 percent of trucks were following cars with this headway. The difference in the headway distribution for these two cases was found to be statistically significant.

Traffic Operation and Control

Several computerized analytical models are used for analysis, optimization, and control of traffic operations on highway facilities. To derive full benefits from the more sophisticated analytical models, it is necessary to divide the traffic flow into vehicle categories according to vehicle weight and length and frequency of occurrence. For example, the model TRARR (*13*), used for analysis of undivided highways, can accommodate up to 16 vehicle types, and FOMIS (*14*), a simulation model for freeway sections, can accommodate up to 100 vehicle types. Undoubtedly, WIM-supplied data can take advantage of the models' options and can increase their usefulness and accuracy.

Analysis Related to Highway Bridges

Two levels of WIM data usage can be envisaged: to determine load levels used for reviewing the design and maintenance

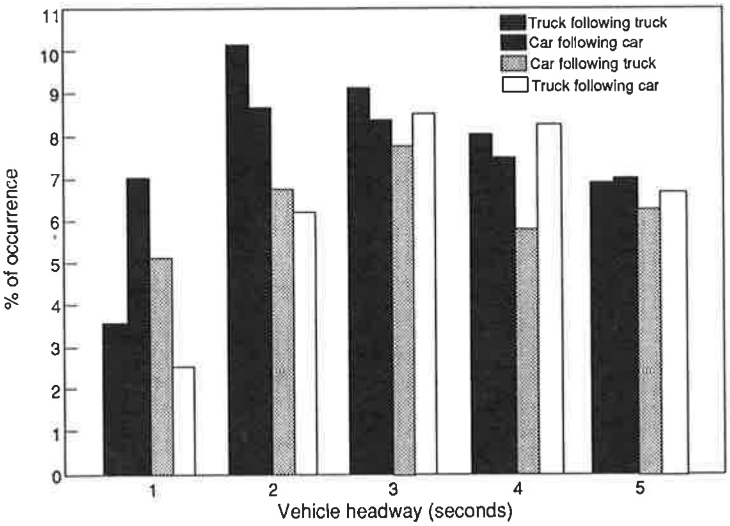


FIGURE 10 Daytime vehicle headway distribution: Highway 402, truck lane, March 19–22, 1991.

standards for bridges, and to determine load levels for a specific bridge structure.

Review of Loading Requirements for Design and Maintenance

Periodic vehicle weight surveys are essential to determine loading standards for the design and maintenance of bridges (*15*). The most recent load survey for Ontario, conducted in 1989, was done manually, by measuring (with a tape) and weighing (with static scales) a selected sample of about 2,000 trucks in 17 locations across the province. This type of survey could become obsolete as the number and accuracy of WIM scales increase. Some preliminary work on using WIM data, obtained by an instrumented highway bridge, was done by Agarwal and Bakht (*15*). At any rate, WIM scale results can supplement manual surveys and have the advantage of realistically capturing overloaded trucks, which are usually one of the main reasons for conducting such surveys.

Load Levels for Specific Bridge Structures

The loading and safety concerns about a specific bridge structure can be addressed by temporary installation of a portable WIM scale. The WIM scale can provide the loading information to the bridge owner and can also provide a warning to drivers of trucks that are overloaded for the specific bridge structure.

DISCUSSION OF RESULTS

The discussion on using WIM scale data and extracting specific information of interest in various application areas is not exhaustive. It simply illustrates possible usage in the traditional application areas. Other application areas may include modeling of goods movement and macroeconomic applications (input-output models). For example, the Pennsylvania Department of Transportation uses WIM scale data to determine total cargo weight transported by the highway on a long-term basis (*16*). WIM scales, combined with automatic vehicle identification systems, may be used for fleet management and management of truck traffic through long-distance corridors (Crescent project). On a somewhat unrelated topic, WIM scales can be installed on remote airports to document aircraft movements (e.g., type of aircraft, how loaded, and time of operation).

CONCLUSIONS

1. This paper has demonstrated, through examples, that WIM data can provide insights not previously available to a wide range of issues that cut across the organization of a highway agency.
2. WIM technology can provide statistically reliable samples that should supplant older labor-intensive manual survey methods.
3. WIM scales, because of their unobtrusiveness and continuous operation, can provide truly unbiased data yielding a realistic long-term picture of highway usage and driver behavior.
4. Currently available WIM data do not satisfy all requirements connected with analyzing the effects of regulatory policies on truck weights and dimensions.
5. When planning the installation of new permanent traffic counters, or the replacement of the existing ones, consideration should be given to upgrading the counters into WIM scales.
6. The demonstrated usefulness of WIM data requires that they be treated as corporate data. Highway agencies should consider establishing dedicated (formal) WIM data banks to facilitate data storage and retrieval as a service to potential users. WIM scale operation and WIM data storage should be integrated with other traffic data gathering and storage processes as the corporate resource.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support received from the members of the SHRP Expert Task Group on Traffic Data Collection and Analysis, in particular from David Albright, Mark Hallenbeck, and Curtis Dahlin. The contribution of Alexander Ugge on the use of WIM data for traffic operation and control and the contribution of Akhilesh Agarwal on the use of WIM data for bridge-related analysis are also gratefully appreciated.

REFERENCES

1. W. A. Phang. A Single Transducer Weigh-in-Motion Scale for Unattended Highway Traffic Data Collection, An RTAC Project. *International Colloquium, Full Scale Pavement Tests*, Zurich, Switzerland, May 1982.
2. *IRD High Speed WIM-5000*. International Road Dynamics, Inc., Saskatoon, Saskatchewan, Canada, 1989.
3. *IRD Weigh-in-Motion Piezoelectric 1060P, User's Manual*. International Road Dynamics, Inc., Saskatoon, Saskatchewan, Canada, 1990.
4. *Traffic Monitoring Guide*. Office of Highway Planning, FHWA, U.S. Department of Transportation, June 1985.
5. *AASHTO Guide for Design of Pavement Structures, 1986*. AASHTO, Washington, D.C., 1986.
6. *Pavement Design and Rehabilitation Manual*. Report SDO-90-01. Ontario Ministry of Transportation, Downsview, Canada, Jan. 1990.
7. J. R. Billing and J. J. Hajek. *Weigh-in-Motion Scale Data as a Basis for Geometric Design Standards Development, Weight Compliance Assessment, and Regulatory Policy Development*. Ontario Ministry of Transportation, Downsview, Canada, Sept. 1990.
8. B. G. Hutchinson. Large-Truck Properties and Highway Design Criteria. *Journal of Transportation Engineering*, Vol. 116, No. 1, Jan. 1990, pp. 1–22.
9. *Memorandum of Understanding on Interprovincial Heavy Vehicle Weights and Dimensions*. Roads and Transportation Association of Canada, Ottawa, Ontario, Canada, Feb. 1988.
10. H. W. McGee. Accident Data Needs for Truck Safety Issues. In *Transportation Research Record 1052*, TRB, National Research Council, Washington, D.C., 1986, pp. 146–150.
11. D. J. Parker and B. G. Hutchinson. *Large Truck Braking at Signalized Intersections*. Report TDS-88-01, Ontario Ministry of Transportation, Downsview, Canada, Dec. 1988.
12. *The Highway Traffic Act* (as amended to May 1989), Government of Ontario, Toronto, Canada, May 1989.
13. C. J Hoban, G. J. Fawcett, and G. K. Robinson. *A Model for Simulating Traffic on Two-Lane Rural Roads: User Guide and Manual for TRARR, Version 3.0*. Technical Manual ATM 10A. Australian Road Research Board, Victoria, 1985.
14. B. Hellinga. *FOMIS—A Microcomputer Freeway Traffic Simulation Program*. Transport Group, University of Waterloo, Ontario, Canada, Sept. 1987.
15. A. C. Agarwal and B. Bakht. Error Analysis of Weighing-in-Motion by Highway Bridges. *Transportation Research Forum*. Vol. 3-3, pp. 31–40, Dec. 1986.
16. B. M. Hains, T. TenEyck, and A. J. Pietropola. Pennsylvania Department of Transportation's Weigh-in-Motion Program. Presented at 70th Annual Meeting of the Transportation Research Board, Washington, D.C., 1991.

Publication of this paper sponsored by Committee on Vehicle Counting, Classification, and Weigh-in-Motion Systems.

Trial Implementation of a Photoelectric Sensor System To Classify Vehicles

J. L. GATTIS AND CLYDE E. LEE

The development and field testing of an automatic vehicle classification (AVC) system in 1990 and 1991 are described. The system employs an inductance loop detector and a pair of photoelectric sensors emitting modulated infrared-spectrum light beams, which are reflected off raised pavement markers. The presence of a vehicle activates the loop detector, prompting the system to classify the passing vehicle. The pattern of the infrared light beam interruptions enables the system to infer the vehicle's speed, number of axles, and number of tires. The number of axles and tires defines the vehicle's class for assessing the proper toll. Signals from the activation go through a special microprocessor into a personal computer for interpretation, display, and recording. The system was tested initially at a high-speed (50 to 65 mph) Oklahoma Turnpike Authority site. Evaluation of the computer records and written field notes obtained during a 5-wk study indicated that the first-generation AVC system correctly classified about 95 percent of the 1,736 vehicles observed. After a review of the data and the number and type of errors found therein, it was concluded that the system software could be modified to correctly classify about 97 percent of the vehicles. The AVC system was subsequently modified and installed in four automatic vehicle identification toll-gate lanes on the Turner and Will Rogers turnpikes in Oklahoma, where vehicles with a special device on board are allowed to debit their toll account as they pass through without stopping. Evaluation of the AVC system operating in this environment is continuing.

Several different types of traffic data are needed for planning, designing, and operating street and highway systems. One such type of data are vehicle classification data, which report the number of various types of vehicles in the traffic stream. Some classification methods aggregate vehicle types into groups or classes according to the number of axles and the tire arrangement, whereas others use different criteria.

In the past, human observers have been the predominant means by which to obtain vehicle classification data. Development of mechanical or automated systems to do this job may lower data-collection costs while improving data accuracy and reliability.

This paper describes the conceptual development and field evaluation of an automatic vehicle classification (AVC) system in 1990 and 1991. The Oklahoma Turnpike Authority (OTA), wanting to deploy an AVC system to enhance their toll collection and auditing operations, contacted the Center for Transportation Research at The University of Texas at Austin and the Oklahoma Transportation and Infrastructure Center at The University of Oklahoma concerning the development of such a system. The recommended system employs an inductance loop detector and a pair of photoelectric sensors emitting infrared-spectrum light beams, which are reflected off raised pavement markers (RPMs) of the type commonly used to mark traffic lanes. The RPMs are positioned near the middle of the lane, so the tires of passing vehicles will straddle the RPMs. The presence of a vehicle activates the loop detector and prompts the system to classify the passing vehicle into one of eight classes in the OTA toll schedule. When a vehicle straddles the RPMs, each of the tires on one side will for an instant interrupt the pair of infrared beams. The pattern of the interruption of the infrared light beams enables the system to infer the vehicle's speed, number of axles, and number of tires; thus, its class. Signals from the activation go through a special microprocessor into a personal computer for interpretation, display, and recording.

The initial charge to the development team was to recommend an AVC system for application in a manually operated toll gate situation in which vehicles stop—sometimes more than once—and then pass through the gate area at slow speeds, probably while accelerating. Later, the objective was shifted to developing an AVC system that would function in a toll collection lane in which vehicles pass through the toll gate—perhaps at medium to high speeds—without stopping. This system was envisioned for possible application in toll collection lanes designated for vehicles equipped with automatic vehicle identification (AVI) devices.

The higher-speed AVC system was built and field tested during the summer of 1990 on the Turner Turnpike east of Oklahoma City, where traffic speeds were about 50 to 65 mph. These tests indicated promising performance from the system and suggested software modifications that would improve the accuracy of vehicle classification.

In late 1990 and early 1991, specially designed hardware and software were installed at four AVI-equipped toll collection lanes on the Turner and Will Rogers Turnpikes between Oklahoma City and the Missouri state line. Evaluation of these systems, operating in an environment in which vehicle speeds are generally about 25 to 40 mph, is progressing. Continuing observations have identified problems not encountered during the initial tests. Means of addressing these new problems are being considered and developed.

BACKGROUND

OTA has operated six toll roads in the state of Oklahoma for a number of years; four new roads are now opening. There are about 100 toll gates. The amount of the toll charged to a

---

J. L. Gattis, Oklahoma Transportation and Infrastructure Center, University of Oklahoma, 202 West Boyd, Room 334, Norman, Okla. 73019. C. E. Lee, Department of Civil Engineering, University of Texas, ECJ Hall 4.2, Austin, Tex. 78712.

vehicle is determined by the classification of that vehicle and the distance traveled. Classification is determined by the number of axles and tires per vehicle. There are eight vehicle classes:

Class 1. Automobile, motorcycle, or two-axle, four-tired truck;
Class 2. Class 1 vehicle towing one-axle trailer;
Class 3. Class 1 vehicle towing two-axle trailer;
Class 4. Two-axle bus, or two-axle, six-tired truck;
Class 5. Three-axle bus, or single/combination three-axle truck;
Class 6. Four-axle combination truck;
Class 7. Five-axle combination truck; and
Class 8. Six-or-more-axle combination truck.

The Need

Until 1991, OTA's toll collection system relied on the toll booth attendant to correctly classify vehicles and assess the proper toll. A treadle in each toll lane provided an axle count to verify the attendants' totals. Manual auditing of this system was costly and time consuming. The treadles occasionally required maintenance.

OTA wanted to take advantage of current technology and upgrade the established methods of vehicle detection, vehicle classification, assessing tolls, collecting tolls, and auditing toll collection. A number of considerations prompted the decision to investigate new AVC technology. Because public expectations of governmental agency accountability are continually rising, OTA felt a need to upgrade the means of auditing toll collections. In short, OTA had decided to aggressively pursue new technology to make its operation more efficient.

AVI Toll Collection System

OTA introduced systemwide automatic toll collection on January 1, 1991. The automatic toll collection system was based on AVI system technology. Motorists who wish to participate purchase a debit account. When they do, they get a transponder tag the size of a deck of playing cards. This tag is attached with removable adhesive strips to the inside-front windshield of the vehicle.

The AVI system employs a transmitter and receiver at each toll collection station. The signal sent by the transmitter is altered by the transponder tag inside each subscribing vehicle and reflected to the receiver. Subscribing vehicles do not stop to pay tolls at booths with AVI equipment. Instead, the account of the participating vehicle is automatically debited as the vehicle passes through the toll-assessing area (i.e., the toll gate).

Each AVI account is "classification specific." That is, the user of an AVI tag is supposed to use that tag only in a vehicle of a certain classification. For instance, an AVI tag purchased for an automobile (Class 1) is not supposed to be used with an automobile pulling a trailer (Class 2) or in a tractor-semitrailer truck (Class 7), because the toll for a Class 1 differs at a given toll booth from the toll for a Class 2 or a Class 7.

While finalizing plans for installation of the AVI system, OTA decided that an AVC system should complement the automatic toll collection system. Used in the toll collection environment, automatic vehicle classification data can

1. Verify each manual vehicle classification made by toll collector;
2. Audit the automatic vehicle classification of each tagged vehicle, to determine whether the proper tolls were being assessed; and
3. Override the AVI tag signal to correctly charge a vehicle with an incorrect tag in the windshield.

Thus, AVC was needed to complement both manual and automatic toll collection operations.

Initially Recommended AVC System

In early 1990, before actually implementing the AVI system, OTA contracted with the Center for Transportation Research at The University of Texas at Austin and the Oklahoma Transportation and Infrastructure Center at The University of Oklahoma to conduct a feasibility study and suggest a method that could automatically classify vehicles at manually operated turnpike toll gates. The study (*1*) reported some methods that were currently under development or in use on other toll systems, such as those using ultrasonic detectors. The study recommended an AVC system using a pair of photoelectric sensors and an inductance loop vehicle presence sensor; this system was then under development at The University of Texas at Austin (*2*).

Later in 1990, the Turnpike Authority modified its original intent and requested the demonstration of an AVC system for use in a nonstop, higher-speed environment. The researchers recommended the same basic sensor array for this application and proceeded to develop hardware and software for demonstration at a high-speed sight.

INITIAL DEVELOPMENT AND TESTING

The recommended system was a "first-generation" concept and therefore needed additional testing and modification before being ready for full-scale implementation. The system was tested from mid-August to mid-September 1990 on the west end of the Turner Turnpike (I-44), northeast of Oklahoma City. The test was conducted in the right, or outside, lane only; vehicles in the left lane were not classified.

Procedure

The researchers conducted seven tests over a period of about 5 wk. Each test lasted no more than a few hours during daylight. The tests were conducted in warm to hot weather. There was no rain during any of the tests.

The system hardware included the modulated infrared light beam sensors, a loop detector, raised pavement markers, a special microprocessor, a microcomputer, and wiring. In addition, housing is needed to protect the sensors and the computers from the elements. Two adjacent steel pillbox-like cylinders with lids served as housings for the infrared sensors. The pillboxes were positioned beside the shoulder, shielded by an existing guardrail. The pillboxes were less than a foot high and about a foot in diameter. Both pillboxes had portholes with visors, to allow the infrared beam to be sent and received. Figure 1 shows a schematic layout of parts of the initial system.

The main goal of the test was to evaluate the system's ability to correctly classify passing vehicles according to the OTA toll classification scheme. To accomplish this, human observers classified passing vehicles just before the vehicle entered the location where automatic classification was to occur. The human observer orally communicated his or her classification to a person at a microcomputer keyboard, who in turn entered the observer's classification as the automatic classification was taking place. Researchers decided to employ two persons (observer and keyboarder) instead of one so the observer would not have to look away from the road and because physical constraints at the sight (such as electric power availability and the limited view of lanes from a narrow embankment) made it undesirable to try to classify from the spot where the computer was located. The recording procedure created a computer file with both the observer's record and the automatic record stored together. Later, these records were compared to determine the number and nature of differences between the observer and the AVC system. Another person at the test site recorded any observer or keyboard input errors, so when the record was reviewed later, it would be easier to separate human errors from automatic system errors.

Problems and Solutions

There were some computer and software problems during the first few tests. The problems were identified and solutions implemented, so the system functioned well during the later tests. There were also human-input errors, especially when a number of closely spaced vehicles came by. With less than 2-sec headways, it was difficult to correctly key in the vehicle class before a following vehicle passed by the loop and sensors.

In addition to computer problems, the researchers had to find an ideal spacing between the sensors. The system sometimes confused wide single tires with narrow dual tires. During the latter data-taking sessions, the AVC system occasionally indicated, in a seemingly random pattern, dual tires on the front axle. For a dual tire to be indicated, both infrared light beams must have been interrupted simultaneously. The cause of this problem was not obvious. Cleaning the lenses did not alleviate the problem. Eventually, researchers decided that the dimensions of some wide single tires and some narrow dual tires were too similar.

A few vehicles did not straddle the RPMs but instead passed by entirely to the right or to the left of them. All four tires of the few passenger cars that traveled to the extreme right side of the instrumented lane interrupted the infrared light beams. Vehicles positioned in this manner registered "extra" axles because four tires interrupted the infrared beams, not two. The AVC system had difficulty classifying motorcycles. Some motorcycles came by in pairs, pulled small trailers, or passed by to the left of the reflectors.

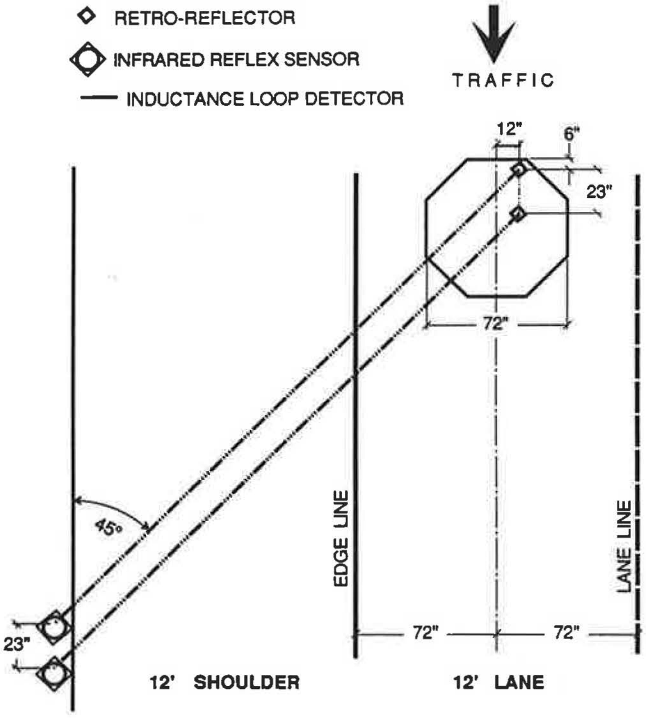


FIGURE 1 AVC sensor arrangement for initial high-speed tests.

Results

By reviewing the computer record, researchers were able to classify errors as caused by either human input, AVC-system error, or "unsure." The automatic vehicle classification system correctly classified about 95 percent of the 1,736 passing vehicles.

Some errors occurred when following vehicles were closely spaced; close headways create a need for a fast-executing program. The AVC system did not accurately classify motorcycles, and motorcycle detection problems accounted for 0.4 percent of the system errors.

The nature of the problems experienced was such that it was expected that some of the problems with the first-generation system were correctable, and it was estimated that the AVC's accuracy could be improved to about 97 percent. The testing of the first-generation AVC system led to the identification of issues that needed to be addressed during future research and development of the AVC system.

DEVELOPING AVC SYSTEM FOR IMPLEMENTATION

The Turnpike Authority notified the researchers in late 1990 that they had decided to implement the tested system on a small scale at four of the highest-volume toll plazas and wanted the system installed by early 1991. These sites are at each end of the Turner and the Will Rogers Turnpikes, which together extend from Oklahoma City northeast to the Missouri state line near Joplin, Missouri.

Each of the four plazas has one AVI lane, in which motorists with the proper AVI tags can proceed without stopping to manually pay a toll. Vehicle speeds in the AVI lanes generally range from 25 to 40 mph. Infrequently, a motorist without an AVI tag, apparently unfamiliar with the system and not comprehending the meaning of the numerous signs, gets into the through-AVI lane and then comes to a stop, trying to find someone to whom to give a toll.

Designing the System

The researchers had only a few weeks to prepare for installation. Experiences from the initial system testing prompted some changes in the hardware and software installed at the four sites. The design configuration was also constrained by the existing layout of the toll plaza areas.

The toll plaza areas consisted of a number of islands in parallel, to channel all vehicles alongside one of the toll booths where collections were made. The newly created through-AVI lanes were located on the left or inside. To the left of the AVI lane is a median barrier curb, about 30 in. high. To the right of the AVI lane is an island and a tollhouse serving the lane to the right of the AVI lane. The tollhouses are enclosed, except for an open window through which the driver passes cash to the attendant. A crash attenuator is located in advance of the island and tollhouse.

An initial challenge was identifying a suitable location for the sensor installations, out of traffic. There was a space of about 4 to 5 ft between the backside of the crash attenuators and the front nose of the islands, with guardrail at the face of the lane. This location offered just enough space for a roadside sensor installation, so none of the island had to be jackhammered out.

The researchers chose not to use two pillboxes to house the infrared sensors as was done in the initial test. Instead, a single low-profile aluminum box was fabricated to house both sensors in one unit. This allowed the proper spacing between the sensor pairs to be set in the factory, as opposed to trying to correctly position them during a field installation. Based on previous experience, a 24-in. spacing between sensors was used. Visors projected inward into the aluminum box, so nothing protruded outside the box.

The widths of the AVI lanes varied among the four sites. The two sites on the Will Rogers had lanes approximately 10.5 ft wide; Figure 2 shows a schematic layout used on the Will Rogers sites.

The lanes on the Turner Turnpike were approximately 14 ft wide. The 14-ft widths posed a problem: that is, vehicles have more latitude than normal in the path taken as they pass through the loop and sensor area. The researchers needed to find a simple configuration that would detect vehicles anywhere within the lane, no matter how far to the left or to the right the vehicle was. The vehicle detection has to be made before the front tire interrupts the leading infrared beam, and the detection must continue until the rear tire has completely crossed the path of the trailing infrared beam. One constraint was keeping the loop relatively small so as to retain the ability to continually sense the presence of high-body trailers. In addition, researchers wanted to minimize the number of loops to simplify installation.

The researchers addressed this problem by employing a diagonal "figure 8" detector loop. The relative positions of the sensors, the reflectors, and the loop permit a vehicle in the right side of the lane to pass over the loop in time to activate the "call for detection" before the right front tire interrupts the leading infrared light beam. A vehicle at the extreme left side of the lane will still be over the other end of the figure 8 until after all tires have cut the trailing light beam. Figure 3 shows this layout. The reflectors were placed

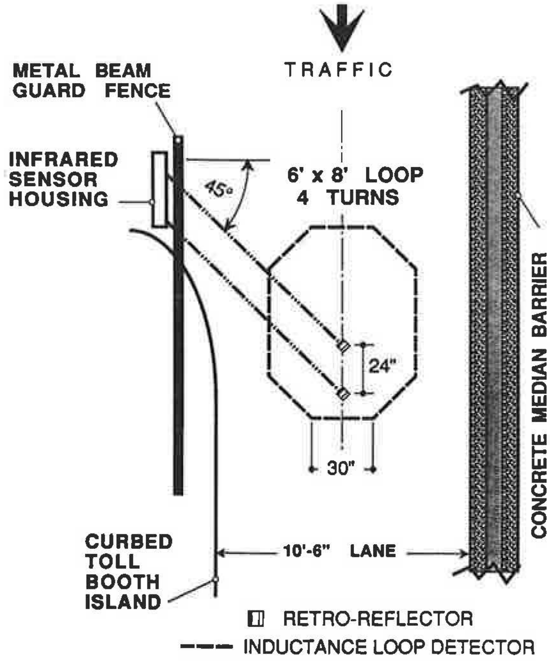


FIGURE 2 AVC sensor arrangement for narrow AVI lane.

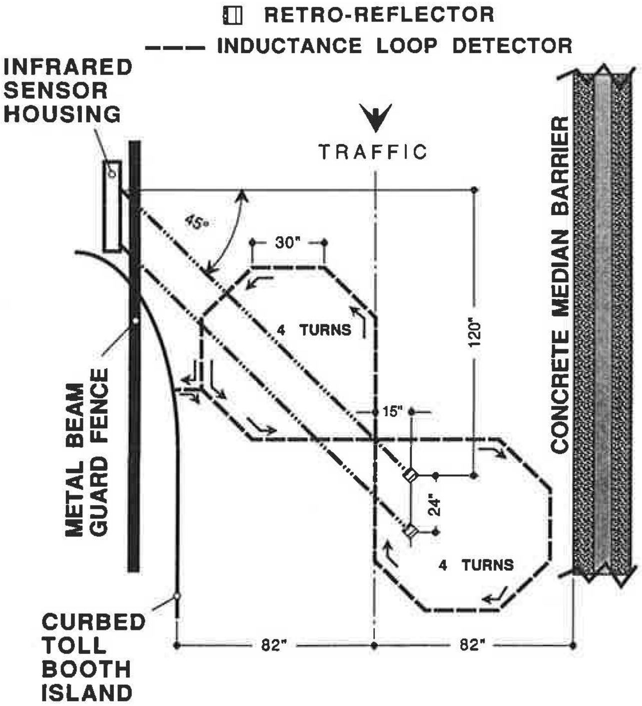


FIGURE 3 AVC sensor arrangement for wide AVI lane.

to the left of the lane centerline, in the hope that most motorcycles would pass to the right side of the reflectors, so the tires would interrupt the infrared light beams and be classified.

The researchers modified the original software so data collection could continue for a number of days before the data had to be downloaded. They also addressed the problem encountered during the initial test of confusing wide single tires with narrow dual tires. The solution was assuming that all front tires were single tires and then referencing following tires to the front tires. If the following tires were wider than 1.2 times the front tire width, the following tires were said to be dual tires. Width was inferred from the duration of infrared beam interruption.

Installing the System

The system was installed during the winter, with subfreezing temperatures often encountered. The cold weather influenced the researchers to use a poured asphalt adhesive to install the raised pavement marker reflectors in the lanes, as well as to attach the aluminum sensor housing boxes to the concrete surface. Wiring extended from the sensors, through previously installed conduit, into the nearest tollhouse, in which the microprocessor and microcomputer were housed.

Figure 4 shows a pair of sensors installed in the aluminum box; the box cover has been removed. Figure 5 shows the box in the background and the reflectors and loop in the foreground.

The loop and reflectors were located slightly upstream of the tollhouses. Any passenger car or other short vehicle in the AVI lane that erroneously stops alongside the tollhouse to pay a toll will have already left the detector loop before coming to rest. All four systems were turned over to the Turnpike Authority in working order.

MONITORING AND FURTHER DEVELOPMENT

OTA planned for the toll collection staff to check the AVC system performance. The other regularly assigned tasks prevented toll personnel from checking the system as often as needed, so monitoring was irregular. From available data, it was apparent the system was having problems, but there was not enough constant monitoring to identify the causes. It seemed that the problems could be arising from a combination of human input error and system malfunction.

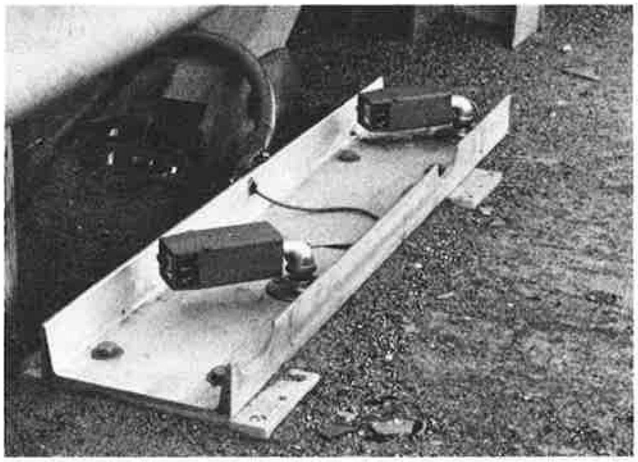


FIGURE 4 AVC sensors attached to base of box.

FIGURE 5 Loop, reflectors, and sensor housing box.

In summer 1991, two students employed by OTA for the summer assumed the task of monitoring the system. With regular field monitoring, and some intensive field checking, researchers think they have identified a number of problem causes.

A snowplow accidently removed the raised pavement markers during a late winter storm at the state line site. This site was the most time consuming to check, because for the person coming from Oklahoma City, it required another 3 hr of driving time past the third site. OTA decided not to reinstall the markers and to use the remaining three sites for evaluation and development.

Symptoms of Problems Encountered

There were many system problem symptoms. An AVC installation would appear to correctly classify for a few days, then have a few hours of obviously incorrect data. This period sometimes would be followed by apparently correct operation. Incorrect data took the form of a "Class 0" (i.e., AVC could not recognize a vehicle type) or unreasonably high speeds (i.e., 9,530 mph), or both. Sometimes the system would prematurely cease operation, and some data files contained more than one day's data (a mistake). There also had been some visible physical damage to the system wiring at one site.

Most of the observed classifications have been correct. In an exception, researchers observed the effects of a truck trailer having rear mud flaps that were dragging the ground. The flap touching the pavement surface interrupted the infrared beam, creating the impression of an extra axle. This caused the system to incorrectly classify a Class 7 as a Class 8 vehicle.

Attempts To Correct Problems

Many of the malfunctions happened with no one present to document a cause. From conjecture, researchers created a list of potential problem causes. They were

- Infrared beam misalignment;
- Dirty reflectors;
- Pavement vibration;
- Human operator errors when downloading data and restarting system;
- Computer hardware problems; or
- Computer software problems.

A number of separate actions have been implemented in an attempt to improve the AVC system operation.

To eliminate operator input error as a source of problems, the computer keyboards were removed from the toll plazas. The keyboards are taken to the sites only by central office personnel. This action seemed to eliminate the problem of one file containing more than one day's data.

Researchers traced the root of some problems to equipment malfunction. Specifically, the microprocessor experienced electrical troubles. The cause could have been improper handling of the unit. The dirty environment is suspected as the cause of another computer malfunction, the failure of one floppy disk drive to copy. Computer covers had been installed over the front of the units, but the researchers making field visits often found them removed.

In the summer of 1991, one of the infrared sensor housing boxes came loose from the concrete pavement. The field monitors noted that they could reposition it and align the sensors, and then the system would correctly work until one or two heavy trucks passed by. The truck vibration caused the infrared beams to go out of alignment and no longer be reflected back to the unit for processing.

Suspected Chief Problem

The chief problem seems to be one of infrared beam reflection. After anywhere from 18 hr to a few days, the reflector surfaces of the raised pavement markers become coated and need a quick wiping. It was found that a product that "cuts" oil is better for this task than a window cleaning product. This problem was not encountered during the 5-wk test of the initial system. One difference is that the initial test site was at an unconfined location. At the toll plazas, there are structures, the median barrier curb, and highly channelized traffic. There may be fewer deposits and more natural cleansing at unconfined sites.

There also have been instances of the beam losing proper alignment. When this happens, the photoelectric emitting/receiving device has to be realigned or retargeted onto the RPM for the AVC system to resume operation.

If the trailing reflector is dirty or the trailing infrared beam is misaligned, then an unreasonably high-speed reading results. The series of events during this malfunction is as follows: vehicle interrupts leading beam at time *t*; trailing beam is not being successfully reflected, so computer reads this beam as being interrupted also at time *t*—the distance between the infrared beams traversed in an infinitesimally short time yields a high speed.

The current suspicion is that premature operation cessation resulted from an overloaded computer buffer. A series AVC system readings giving unreasonably high speeds and numbers of axles could overload the buffer. On encountering an operation cessation, the field monitors have sometimes found that simply cleaning the RPM reflective surface will allow the system to commence operation without any computer input. At other times, the computer system has to be restarted.

It seems that after a few months on the road, the reflectors are not as effective; the cause of this is not certain. At one site, the reflector had deteriorated after 5 months until the beam aimed at the marker could not be reliably received for more than a few minutes. The installation of new raised pavement markers in August 1991 seemed to address the problem. A second site is showing characteristics of the same problem. The computer software was modified to check for loss of infrared signal reflection, so malfunctions resulting from this cause could be identified with certainty.

AVC Operation During Rain

By chance, it rained during an August 1991 field check when the researchers were present. This gave the researchers an opportunity to assess the system operation during moderate precipitation. It appeared that spray from vehicle tires has the effect of interrupting the infrared light beam, which can distort the classification. For instance, the AVC system may incorrectly evaluate the infrared beam interruption pattern generated by a Class 7 truck and semitrailer as a Class 1 vehicle because the spray from the front tire blocked the beam until the first pair of dual-axle tires crossed the path of the beams. This extended interruption also created a very large front tire dimension in the computer memory, so the following dual tires appeared to be single tires. The distance between the front dual-tandem axle and the rear dual-tandem axle was such that the beam interruption ceased before the rear dual-tandem tires crossed the beam, although spray caused the rear tandem tires to be viewed as only one axle. The system would identify a slower-speed Class 7 as a Class 5 vehicle. The slower vehicle produces less spray, so the classification distortion is not as great.

Researchers observed one occurrence of the unaware driver stopping in the AVI lane to pay a toll, with another vehicle following closely. The impatient following vehicle was so close that the loop detector apparently did not differentiate between the two vehicles and recorded them as one Class 4 vehicle.

Additional Actions To Address Problems

The recent field monitoring has led to identification of other actions that may solve the observed problems. The following actions are now being considered.

1. Designing a custom raised pavement marker. Most of the marker body would be solid, but a slight recess would be formed in the side toward the sensor. The reflector would reside in the recess, which would permit a reflector surface to be covered from the top but exposed for reflection from the side.

2. Creating a larger RPM reflective surface. This could help improve reliability. A larger vertical dimension would allow the infrared beam to avoid losing alignment.

3. Modifying the computer logic to overcome the spray problem in wet weather. The vehicle speed is determined when the front tire interrupts the two infrared beams. Once the speed is known, then a duration for which a passenger car would occupy the loop can be estimated. If a loop occupancy duration exceeded the maximum for a car at that speed, then the vehicle classification would be "other than a Class 1." Because present experience shows that most users of the AVI system, indeed most vehicles on the turnpikes, are either Class 1 or Class 7, it is of some benefit to differentiate between Class 1 and "others," although it is still desirable to correctly identify all eight classes.

CONCLUSION

Evaluation of this AVC system using an inductance loop detector and a pair of photoelectric sensors emitting infrared light beams reflected off raised pavement markers is continuing. Recent observations have identified problems not encountered during the initial tests. Means of addressing these new problems are being considered and developed.

Recent testing has revealed that some vehicle classifications may be distorted when a layer of water covers the pavement surface. The apparent effects of tire spray blocking the light beams constitute a significant system limitation. Computer programming techniques may partially overcome this problem. In many areas rainfall occurs only a small fraction of the time, so this limitation may be tolerable for many AVC system applications.

If system modifications can eliminate other problems and lead to creation of an AVC system that will operate reliably with minimal maintenance, then the next step would be to link the AVC input with elements of the AVI system and other OTA toll collection operations. This AVC system may have other practical classification applications as well.

ACKNOWLEDGMENT

The authors thank the Oklahoma Turnpike Authority for the support of this research.

REFERENCES

1. C. E. Lee and J. L. Gattis. *Recommended System to Classify Vehicles for Toll Collection.* Oklahoma Turnpike Authority, Oklahoma City, 1990.
2. J. E. Garner, C. E. Lee, and L. Huang. Photoelectric Sensors for Counting and Classifying Vehicles. In *Transportation Research Report 1311,* TRB, National Research Council, Washington, D.C., 1991.
3. C. E. Lee and J. L. Gattis. *Test of a System to Classify Vehicles for Toll Collection.* Oklahoma Turnpike Authority, Oklahoma City, 1990.

Publication of this paper sponsored by Committee on Vehicle Counting, Classification, and Weigh-in-Motion Systems.

Calibration and Adjustment of Weigh-in-Motion Data

RALPH GILLMANN

Several methods have been used for calibrating weigh-in-motion (WIM) systems. One difference between the methods is in the formula used to make the calibration. A systematic approach is taken here in deriving and comparing four calibration methods. Adjustments of WIM data beyond calibration may be used to more closely emulate the properties of static weight data, such as in equivalent axle load calculations. Several adjustments of WIM data are also derived. Examples are given of the various methods of calibrating and adjusting WIM data.

The accuracy of weigh-in-motion (WIM) data depends on many factors, including calibration of the WIM system. The many reports received at FHWA show a variety of approaches to WIM system calibration. One way in which they differ is in the formula used to make the calibration. It does not seem to be appreciated how calibration affects the accuracy of the WIM data. The underlying criteria for four different calibration methods are examined and guidelines in their application are given. One of the main points is that the way in which WIM is evaluated should guide the selection of a calibration method.

WIM data are often used in place of static weight data because of their relatively low cost and convenience. What is then desired is to minimize the differences between WIM and static weight data. In particular, ways in which WIM data can be made to approximate static data in the determination of equivalent axle loads are examined here. Although other approaches are briefly mentioned, the main approach taken is to adjust the WIM data. The results of calibrating and adjusting three data sets are shown as examples of the methods.

ASTM standard specification E 1318-90 states that WIM is "the process of measuring the dynamic tire forces of a moving vehicle and estimating the corresponding tire loads of the static vehicle" (*1*). Note that WIM measures dynamic weights and estimates static weights. All too often these two aspects have been combined or confused. One reason for this is that static weights are the reference standard for WIM data. Another reason is that WIM weights are often taken as a proxy for static weights.

The purpose of WIM data collection determines whether static or dynamic weights are desired. For weight enforcement purposes, WIM is only a proxy for static weights. Weight regulations are aimed at the static weight of individual axles, axle groups, and vehicles. For most other purposes, such as pavement design and management, however, dynamic weights may be acceptable, if not preferred. The actual forces of moving vehicles that affect the roadway are more meaningful than the static weights.

CALIBRATION METHODS

A WIM sensor produces a signal whose value depends on the instantaneous dynamic wheel loads of a moving vehicle. When the output for the sensor is properly calibrated, a dynamic load measurement is produced. This dynamic load may then be used to estimate a static weight. Calibration is the process of adjusting the outputs of a WIM sensor to match the measurements of a static scale.

Some devices can function as both static scales and WIM systems, and thus can be calibrated statically. However, WIM systems usually have to be calibrated with moving vehicles. If a test vehicle could be instrumented in such a way as to record the dynamic wheel forces as it moves, than a WIM system could be calibrated to these dynamic weights. But in the absence of any dynamic weight measurements to compare with, the static weights derived from a nearby scale are the only standard for comparison.

One calibration method involves passing the same vehicle or group of vehicles over the WIM sensor repeatedly. Then an average WIM measurement for each vehicle can be calculated and related to the static weight. The problem with this approach has been that it is costly to get appropriate test vehicles, and the calibration is keyed to a small group of vehicles that may not be representative of the general traffic stream.

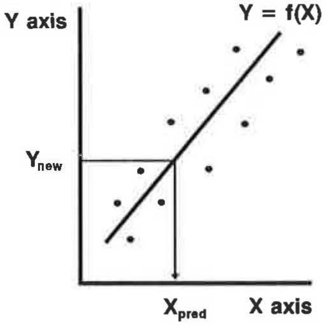
The calibration methods examined here use the gross vehicle weight from a sample of vehicles that have been weighed with both a static scale and a WIM system. The vehicles may include test vehicles as well as the general traffic stream (*1*) but should be representative of the heavy vehicle traffic at the WIM site. The vehicles should also cover the range of weights expected to avoid extrapolating the calibration.

INVERSE PREDICTION

Regression analysis determines a functional relationship between independent and dependent variables that expresses their statistical relationship. A particular regression function is chosen according to given criteria that evaluate the difference between the dependent variable and its estimate by the regression function. Once it is determined, the regression function may be used to predict values of the dependent variable for new values of the independent variable.

For calibration, inverse prediction is needed (*2,3*). The inverse of the regression function is used to predict values of the independent variable given new values of the dependent variable. This is done because the independent variable repre-

FHWA, 400 Seventh Street, S.W., Washington, D.C. 20590.



$X_{pred} = f^{-1}(Y_{new})$

FIGURE 1 Inverse prediction.

sents the reference values that need to be estimated from the dependent variable that is being calibrated.

Let X_i be the reference values, Y_i the measured values to be calibrated, and $f(X)$ the regression function, where the index i ranges from 1 to n, the size of the data set (this same index will be used throughout). Then fit $Y = f(X)$ according to the regression criteria (see Figure 1). Finally, use the inverse of the regression function for the inverse prediction of X given a new value of Y:

$$X_{\text{pred}} = f^{-1}(Y_{\text{new}}) \tag{1}$$

For WIM calibration, static weights are the reference values X_i, and WIM weights are the measured values Y_i. Through inverse prediction, the static weight corresponding to a WIM weight may be estimated. The WIM system may then be adjusted to automatically produce these static weight estimates.

REGRESSION MODELS

A regression model specifies the type of function to be fitted and the criteria to be used to select the particular regression function. The functional form of the regression function may be determined by examining the way in which it is used. WIM calibration involves transforming the original WIM values of a sample data set, Y_i, into the calibrated values in a way that depends on the relation of WIM data to static weight data as defined by the regression model.

The gross vehicle weight is normally used for calibration because it varies less than the dynamic axle or wheel loads. The calibration is then applied to the wheels, axles, and axle groups. Thus the transformation that arises from calibration must be the same whether it is applied to the gross vehicle weight or the individual wheels, axles, or axle groups. That means the regression function must satisfy the functional equation

$$g(y_1) + g(y_2) = g(y_1 + y_2) \tag{2}$$

where y_1 and y_2 represent, for example, WIM axle weights, and $(y_1 + y_2)$ represents the gross weight. The general solution of this equation for nonnegative functions is

$$g(y) = k \cdot y \tag{3}$$

where the coefficient k is a nonnegative constant (*4*, p. 8). This constant is called the calibration factor (CF). This calibration function is the inverse of the regression function:

$$g(y) = \text{CF} \cdot y = f^{-1}(y) \tag{4}$$

so that

$$f(x) = \frac{x}{\text{CF}} = A \cdot x \tag{5}$$

where A is the coefficient derived from the regression analysis. The CF is thus the inverse of the regression coefficient A. We seek to minimize the error of the equation

$$Y_i = A \cdot X_i \tag{6}$$

by an appropriate choice of A. The errors are supposed to be in the WIM weights, not the static weights, so the regression of WIM on static data is used.

The CF is applied to the original WIM values Y_i to yield calibrated values Z_i:

$$Z_i = \text{CF} \cdot Y_i \tag{7}$$

which approximate the static weights X_i for each vehicle i in the data set.

To derive the CF, there must be some criteria for determining the particular regression function. The regression criteria should match the criteria that will be used to evaluate the accuracy of the WIM system. The type of calibration will be determined by the regression criteria employed. If the calibration criteria are inconsistent with the evaluation criteria, then the calibration will be inaccurate.

A particular regression function is chosen by optimizing a property of the residuals, which are the differences between the uncalibrated WIM and the regression estimate:

$$Y_i - A \cdot X_i \tag{8}$$

We will consider two approaches here: making the sum of the residuals equal zero, and minimizing the sum of the squares of the residuals. We will also apply these two approaches to the relative residuals, which are the residuals divided by the regression estimates:

$$\frac{Y_i - A \cdot X_i}{A \cdot X_i} \tag{9}$$

The mean of the differences between calibrated WIM and static weights is called the systematic difference and the standard deviation is called the random difference. This difference may be an absolute or relative difference as we shall see.

Evaluating a WIM system for accuracy and ensuring that it is within tolerance are two approaches to WIM evaluation. For accuracy the systematic difference is the most important consideration, but for tolerance the random difference may be more important. This is because accuracy is usually measured by an average value, whereas tolerances are concerned with keeping within a certain range of values.

A WIM regression model implicitly answers the question whether WIM deviations from static weights are considered measurement errors or dynamic differences. That is, does WIM measure static weight or dynamic weight? If WIM weights are expected to match static weights, then any differences are errors that should not be allowed to cancel one another. The first calibration approaches the residuals as errors.

LEAST-SQUARES CALIBRATION

A least-squares (LS) regression model focuses on the squares of the residuals and seeks to minimize their sum by an appropriate choice of the regression coefficient A:

$$\mathrm{LS}_i = (Y_i - A \cdot X_i)^2 \tag{10}$$

All errors in the LS model are nonnegative, so negative errors cannot cancel out positive errors. To derive the LS calibration factor, take the derivative with respect to A of the total LS error and set it equal to zero:

$$\frac{d}{dA} \sum (Y_i - A \cdot X_i)^2 = 0 \tag{11}$$

or

$$\sum X_i Y_i = A \cdot \sum X_i^2 \tag{12}$$

so that A is (*3,5*)

$$A = \frac{\sum X_i Y_i}{\sum X_i^2} \tag{13}$$

Since the coefficient A is the inverse of the CF,

$$\mathrm{CF} = \frac{1}{A} = \frac{\sum X_i^2}{\sum X_i Y_i} \tag{14}$$

In other words, the LS calibration factor is the sum of the squared static weights divided by the sum of the product of static and WIM weights.

When the errors are assumed to be uncorrelated and normally distributed with mean zero, the least-squares method gives the maximum likelihood estimator (*3*). However, several of the desirable properties of the least-squares estimator with intercept do not apply when there is no intercept. The sum and mean of the residuals do not equal zero. The sum of the observed values does not equal the sum of the fitted values. The regression line does not go through the center of the data points. LS calibration minimizes the variance, not the mean, of the residuals. When the sum or the mean of the residuals is made to equal zero, a different calibration method results.

ABSOLUTE DIFFERENCE CALIBRATION

If WIM is considered as a measurement of dynamic weight, then differences with static weights will be expected. WIM values will vary about the corresponding static values, but WIM should vary as much above static weights as below static weights. Thus, an average of the differences would be expected to equal zero as positive and negative differences cancel each other. Two models result from this approach, depending on which type of difference is used.

An absolute difference (AD) model evaluates WIM using the difference between the WIM and static weights (*6,7*):

$$\mathrm{AD}_i = Y_i - X_i \tag{15}$$

This means that the regression line uses the residuals, not their squares. The sum of residuals is made to equal zero, which means that the sum of WIM and WIM estimates are equated. Thus

$$\sum Y_i = \sum A \cdot X_i \tag{16}$$

so that

$$A = \frac{\sum Y_i}{\sum X_i} \tag{17}$$

and the CF is the ratio

$$\mathrm{CF} = \frac{1}{A} = \frac{\sum X_i}{\sum Y_i} \tag{18}$$

The AD calibration factor is the ratio of the sum of the static weights and the sum of the WIM weights. Geometrically, the AD calibration line goes through the origin and the center of the data points. The mean of AD-calibrated WIM equals the mean static weight, and the mean AD is zero. This calibration has the properties mentioned above that least squares without an intercept lacks.

PERCENT DIFFERENCE CALIBRATION

Instead of the absolute difference, the relative difference (RD) between WIM and static weights may be more important (*8*):

$$\mathrm{RD}_i = \frac{Y_i - X_i}{X_i} \tag{19}$$

The relative difference is usually expressed as a percent difference (PD) (*1,6,7*):

$$\mathrm{PD}_i = 100 \cdot \mathrm{RD}_i \tag{20}$$

The RD and PD are related to the impact factor (IF), which is the ratio of corresponding WIM and static weights (*6,9*):

$$\mathrm{IF}_i = \frac{Y_i}{X_i} = \mathrm{RD}_i + 1 = \frac{\mathrm{PD}_i}{100} + 1 \tag{21}$$

In this case, the regression line uses the relative residuals. The sum of the relative residuals is made to equal zero

$$0 = \sum \frac{Y_i - A \cdot X_i}{A \cdot X_i} = \frac{1}{A} \sum \mathrm{IF}_i - n \qquad (22)$$

which implies that A equals the mean impact factor. Since the CF is the inverse of A,

$$\mathrm{CF} = \frac{1}{A} = \left(\frac{1}{n} \sum \frac{Y_i}{X_i}\right)^{-1} = (\overline{\mathrm{IF}})^{-1} \qquad (23)$$

the CF is the inverse of the mean impact factor (*8*). The mean of the impact factor for PD-calibrated WIM equals 1 and the mean PD is zero.

RELATIVE LEAST-SQUARES CALIBRATION

The relative least-squares (RLS) calibration is derived in a manner similar to LS calibration, except that the relative residuals are used instead of the residuals. First, the derivative of the sum of the squares of the relative residuals is set equal to zero:

$$\frac{d}{dA} \sum \left(\frac{Y_i - A \cdot X_i}{A \cdot X_i}\right)^2 = 0 \qquad (24)$$

or

$$\sum \frac{Y_i}{X_i} = \frac{1}{A} \cdot \sum \frac{Y_i^2}{X_i^2} \qquad (25)$$

Therefore,

$$\mathrm{CF} = \frac{1}{A} = \frac{\sum Y_i/X_i}{\sum Y_i^2/X_i^2} = \frac{\sum \mathrm{IF}_i}{\sum \mathrm{IF}_i^2} \qquad (26)$$

That is, the RLS calibration factor is the sum of the impact factors over the sum of the squares of the impact factors. This calibration method appears to be new.

EXAMPLES

Examples of these calibration methods are given in Table 1, which is based on data sets provided to FHWA from Kansas and Utah plus a test data set that we generated. Only five-axle tractor trailers (3S2s) were selected from the data. There are 81 trucks in the resulting Kansas data set, 327 in the Utah data set, and 21 in the test data set. The tolerance levels used in Table 1 are plus or minus 5,000 lb and 10 percent for the test data and Kansas data and plus or minus 15,000 lb and 25 percent for the Utah data.

TABLE 1 Examples of Weigh-in-Motion Calibrations

Type of Calibration	Data Set	Mean Absolute Difference	Mean Percent Difference	Percent Outside Absolute Tolerance	Percent Outside Percent Tolerance
None	Kansas	-0.26	-1.17	3.70	4.94
	Utah	6.06	10.24	18.04	16.82
	Test	-0.23	4.81	0.00	19.05
Absolute Difference	Kansas	0.00	-0.68	4.94	4.94
	Utah	0.00	0.42	6.73	8.26
	Test	0.00	5.23	4.76	23.81
Percent Difference	Kansas	0.36	0.00	4.94	6.17
	Utah	-0.26	0.00	6.42	8.26
	Test	-2.98	0.00	38.10	14.29
Least Squares	Kansas	-0.29	-1.22	3.70	4.94
	Utah	0.21	0.76	6.73	8.26
	Test	1.20	-7.33	4.76	23.81
Relative Least Squares	Kansas	0.24	-0.24	4.94	4.94
	Utah	-1.59	-2.14	7.65	7.03
	Test	-3.85	-1.52	42.86	23.81

The good news is that the accuracy of various calibrations with the two "real world" data sets varied less than 2 percent. The bad news is that it is not hard to generate a data set in which the difference is 5 to 7 percent as the test data set illustrates. Table 1 also shows that LS calibration does the best for two of the three data sets with an absolute tolerance. Similarly RLS calibration does the best for two of the three data sets with a percent tolerance.

Supposing that the best approach for calibration with a tolerance level is the direct approach, find the CF with the most points within tolerance. The problem is that there may be more than one CF that does this. An algorithm that picked one would be arbitrary. The particular tolerance level chosen would also have a significant effect on the CF.

Another approach is to minimize the variance of the residuals. For an AD tolerance level this leads to LS calibration. For PD tolerance levels this leads to RLS calibration. However, minimizing the variance without minimizing the mean of the residuals can produce skewed results. The systematic difference should be minimized before the random differences are addressed. That is why AD calibration is best for absolute tolerances and PD calibration is best for percent tolerances.

By choosing an inappropriate calibration method, accuracy may be needlessly lost. The calibration should at least work right on the calibration data set. It is the WIM regression model that determines what "right" means. This in turn depends on what criteria are used to make the evaluation of accuracy. For slow-speed WIM where high accuracy is expected, choosing the right calibration method is critical.

CALIBRATION STANDARDS

ASTM E 1318-90 specifies the evaluation of a WIM system in terms of tolerance levels for a test data set (*1*). A given percentage (95 percent) of the data must have an error within the tolerance interval. For most cases, the tolerance level is defined as plus or minus a certain percentage difference. For some cases, the tolerance is plus or minus an absolute value.

Standard E 1318-90 has much detail about the calibration sample and tolerance level but little about calibration calculations, simply (*1*),

> Make the necessary adjustments to the WIM-system settings which will make the mean of the respective differences for each basic measurement equal zero.

Earlier in E 1318-90 "difference" is defined as percent difference. This would imply that the recommended calibration is based on percent differences. However, for absolute tolerances a calibration based on absolute differences makes more sense.

ADJUSTMENTS WITH INTERCEPT

If wheels or axles are calibrated separately, then Equation 2 does not apply. In that case the intercept term need not equal zero. Although it is just possible that the dynamic weight of an axle could be zero, a WIM system reading of zero should mean that no axle is present. But by allowing a nonzero intercept within a certain range, another condition may be included with the calibration criteria. In this way, both the systematic difference and the random difference may be minimized.

Let Z_i be the adjusted data, AF the adjustment factor, and AI the adjustment intercept. Adjustments may be derived by inverting the regression equation:

$$f(x) = A \cdot x + B \qquad (27)$$

(compare Equation 5) so that

$$\mathrm{AF} = \frac{1}{A} \qquad (28)$$

and

$$\mathrm{AI} = \frac{B}{A} \qquad (29)$$

If the mean absolute difference is made to equal zero and the standard deviation of the AD distribution is minimized, the result is the standard least-squares regression line with intercept. Because of the intercept term, the adjustment is applied either to the individual axles or to the gross weights but not to both. If the mean percent difference is made to equal zero and the standard deviation of the PD distribution is minimized, a different adjustment results.

There are other approaches to reducing the random difference. One may try to control the factors that give rise to the random difference in the first place. These are factors that make the dynamic weights differ from static weights, such as pavement condition and vehicle speed. If these can be avoided or limited at the WIM site, then the systematic difference should be reduced.

A list of factors causing dynamic weight to differ from static weight has been compiled by Lee (*10*). These include the vehicle factors of gross vehicle load, distribution of gross vehicle weight, suspension, tires, and aerodynamic characteristics. However, control of these variables and, hence, of the random difference, is limited.

Other approaches to reducing random differences involve the postprocessing of WIM data. We briefly look at the use of multiple calibration factors and then examine adjustments to WIM data.

MULTIPLE CALIBRATION FACTORS

Another approach to minimizing both systematic and random differences is the use of multiple calibration factors. Known sources of random difference can be accounted for by separate calibration factors. Different vehicle types, for example, might have their own calibration factor. Since WIM is normally used in conjunction with automatic vehicle classification (AVC) devices, the vehicle type should be known.

Standard E 1318-90 notes that some WIM systems allow various calibration factors for each wheel, axle, or axle group on a vehicle (*1*). For example, the steering axle usually weighs light with WIM systems because of the torque associated with the drive train (*11*, p. 340). A calibration that includes the steering axle weights will therefore be too high, making the other axles weigh heavier than they should. Having a separate CF for the steering axle would compensate for this. The CF derived from the other axles' weights would then be lower. There is also evidence that drive tandems and trailer tandems have different dynamic properties (*12*) and so might have separate calibration factors. In any case, more research needs to be done to separate out the sources of random difference.

ADJUSTMENT FOR ESALs

The calculation of equivalent single axle loads (ESALs) from WIM data requires special treatment. Unadjusted WIM data in ESAL calculations usually will overestimate ESALs because the standard deviation of WIM data is usually greater than that of the corresponding static weight data, and ESAL calculations are related approximately as a fourth-power function of the static axle weights (*13*). (This is also addressed in an unpublished paper by Tony Esteve of FHWA entitled *An Analysis of WIM versus Static Truck Weight Data*.) Minimizing both systematic and random differences should help alleviate this problem. Multiple calibration factors would be better than an adjustment with intercept because various adjustments would be needed for the axles and axle groups used in ESAL calculations because of the adjustment intercept. A more direct approach is to use ESALs calculated from static and WIM data instead of gross vehicle weights in a manner similar to calibration. Then the adjustment factor is defined in a manner similar to the calibration factor and is applied to the WIM ESALs instead of the weights:

$$\mathrm{ESAL}(Z_i) = \mathrm{AF} \cdot \mathrm{ESAL}(Y_i) \qquad (30)$$

(compare with Equation 7). Since ESALs are added together, it would seem appropriate that the AD criteria be used. In that case the AD ESAL adjustment factor is similar to the AD calibration factor:

$$\mathrm{AF} = \frac{\sum \mathrm{ESAL}(X_i)}{\sum \mathrm{ESAL}(Y_i)} \qquad (31)$$

(compare with Equation 18). If one uses the PD criteria, then the PD ESAL adjustment factor is similar to the PD calibration factor:

$$\mathrm{AF} = \left[\frac{1}{n}\sum \frac{\mathrm{ESAL}(Y_i)}{\mathrm{ESAL}(X_i)}\right]^{-1} \qquad (32)$$

(compare with Equation 23). Since the ratio of ESALs is approximately as the fourth power of the ratio of the corresponding axle weights (*14*), a PD adjustment factor may be derived that is applied to the axle weights:

$$1 = \frac{1}{n}\sum\left(\frac{Z_i}{X_i}\right)^4 = \frac{1}{n}\sum\left(\frac{\mathrm{AF}\cdot Y_i}{X_i}\right)^4 \qquad (33)$$

TABLE 2 Examples of Weigh-in-Motion Adjustments

Type of Adjustment	Data Set	Equivalent Single Axle Load Mean	 Mean Absolute Difference	 Mean Percent Difference
Unadjusted Static Weight	Kansas	0.929	—[a]	—[a]
	Utah	1.426	—[a]	—[a]
	Test	1.979	—[a]	—[a]
Unadjusted WIM	Kansas	0.984	0.054	-0.86
	Utah	2.328	0.902	59.66
	Test	3.366	1.380	85.65
AD ESAL Adjusted WIM	Kansas	0.929	0.000	-6.35
	Utah	1.425	0.000	-2.22
	Test	1.979	0.000	9.37
PD ESAL Adjusted WIM	Kansas	0.992	0.063	0.00
	Utah	1.460	0.032	0.00
	Test	1.809	-0.170	0.00
PD Axle Adjusted WIM	Kansas	0.949	0.020	-4.33
	Utah	1.339	-0.086	-8.00
	Test	1.504	-0.475	-13.17

[a] Not applicable

so that

$$AF = \left[\frac{1}{n}\sum\left(\frac{Y_i}{X_i}\right)^4\right]^{-1/4} \tag{34}$$

which is the inverse of the fourth power mean of the impact factor. The adjusted axle weights would then be used to estimate the ESALs.

The data set used for deriving the ESAL adjustment factors that are applied to the WIM ESALs should be representative of the ESALs experienced at the WIM site. Because making adjustments is difficult to do in practice, it may be preferable to adjust the WIM axle weights instead of the WIM ESALs.

Table 2 gives examples of WIM adjusted for ESALs for the same data sets used in Table 1. This was done using a sensitivity index of 2.5 and a structural number of 3 on flexible pavement. The second and third axles as well as the fourth and fifth axles were combined into tandems since only 3S2s were included in the data set.

The 60 percent average percent difference in the unadjusted Utah data set shows the need for adjusting WIM ESALs. The form of the PD adjustment applied to the WIM axle weights is less accurate than the other adjustments. Further research is needed in this area. The calculation of ESALs directly from dynamic, rather than static, loads may be the best solution.

CONCLUSIONS

- The WIM calibration method should correspond to the following evaluation criteria:
  - –AD calibration should be used for WIM when the evaluation is based on absolute differences;
  - –PD calibration should be used for WIM when the evaluation is based on percent differences;
  - –AD calibration is preferable for slow-speed WIM with absolute tolerances;
  - –PD calibration is preferable for WIM with percent tolerances; and
  - –LS and RLS calibrations should be used with caution.
- Adjustment of calibrated WIM data may be useful for the following special purposes:
  - –To minimize both systematic and random differences; and
  - –To approximate static ESALs.
- Further research is needed
  - –To quantify the factors that make dynamic weights differ from static weights;
  - –To determine the advantages of using multiple calibration factors; and
  - –To use dynamic weights directly in ESAL calculations.

REFERENCES

1. American Society for Testing and Materials. *E 1318-90: Standard Specification for Highway Weigh-in-Motion (WIM) Systems with User Requirements and Test Method.* ASTM, Philadelphia, Pa., 1990.
2. B. Izadmehr and C. E. Lee. Accuracy and Tolerances of Weigh-in-Motion Systems. In *Transportation Research Record 1123*, TRB, National Research Council, Washington, D.C., 1987, pp. 127–135.
3. J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Statistical Models.* Irwin, Homewood, Il., 1990, pp. 167–176.
4. J. Aczél. *A Short Course in Functional Equations.* D. Reidel, Boston, Mass., 1987.
5. R. Gillingham and D. Heien. Regression thru the Origin. *American Statistician*, Vol. 25, No. 1, pp. 54–55.
6. P. Davies and F. Sommerville. Calibration and Accuracy Testing of Weigh-in-Motion Systems. In *Transportation Research Record 1123*, TRB, National Research Council, Washington, D.C., 1987, pp. 122–126.
7. W. D. Cunagin. WIM System Accuracy. Presented at 3rd National Conference on Weigh-in-Motion, St. Paul, Minn., 1988.
8. B. Izadmehr and C. E. Lee. On-Site Calibration of Weigh-in-Motion Systems. In *Transportation Research Record 1123*, TRB, National Research Council, Washington, D.C., 1987, pp. 136–144.
9. R. C. Moore. High Speed Axle Load Measurement: A European Perspective. Presented at 2nd National Conference on Weigh-in-Motion, Atlanta, Ga., 1985.
10. C. E. Lee. Factors That Affect the Accuracy of WIM Systems. Presented at 3rd National Conference on Weigh-in-Motion, St. Paul, Minn., 1988.
11. L. Henion, N. Ali, A. T. Bergan, and M. Krukar. Evaluation of Multi-Sensor Piezoelectric Weigh in Motion Systems in Oregon. *Proc., National Traffic Data Acquisition Technologies Conference and Expo*, Austin, Tex., 1990.
12. H. F. Southgate. Dynamic Forces Versus Static Weights—Highway Design. Presented at 3rd National Conference on Weigh-in-Motion, St. Paul, Minn., 1988.
13. M. Hallenbeck and T. C. Kuo. Use of Weigh-in-Motion for Pavement Design: A Preliminary Examination of ESAL Calculation Using WIM Data. Presented at 3rd National Conference on Weigh-in-Motion, St. Paul, Minn., 1988.
14. E. J. Yodel and M. W. Witczak. *Principles of Pavement Design*, 2nd ed. Wiley, New York, N.Y., 1975.

Publication of this paper sponsored by Committee on Vehicle Counting, Classification, and Weigh-in-Motion Systems.

Challenges Confronting Metropolitan Portland's Transportation Decision System

SY ADLER AND SHELDON EDNER

The structure and dynamics of the regional transport regime in Portland, Oregon, are discussed and its consensus decision-making process through a series of challenges to the regime is explored. These challenges include (a) cultivating new sources of project finance as the federal government reduces its contribution, (b) integrating transport projects with regional and local land use plans designed to manage urban growth, and (c) intensifying competition between business centers within the region as rapidly growing suburban business centers seek transport projects that will facilitate locally oriented economic growth. Initially, the institutional and normative elements of the regional consensus process are discussed. Then a set of case studies that illuminate the challenges confronting the regime and the nature of regime responses is developed. To conclude, the future stability of the regime is considered. The challenges call into question the stability of the informal institutional character of Portland's regime that has prevailed since it was created. Clearly, an informal, consensus-based decision process functions well when compensatory resources are available to mitigate zero-sum consequences of decisions. Whether the process can sustain itself without being able to remediate the consequences of internal regime resource allocations is the true test of an institutionalized versus a noninstitutionalized regime.

A regional transport regime—a set of urban institutional structures and related behavioral norms, rules, principles, and decision-making procedures governing relations between public and private political actors—evolved in the late 1970s in the Portland metropolitan area (*1–4*). This evolution took place around the process of withdrawing Mount Hood Interstate Freeway funds and transferring these funds to a variety of other highway and transit projects in the region. The most important product of this regime conversion process has been and continues to be a regional consensus regarding project priorities. The structure and dynamics of this consensus process are discussed through an exploration of a series of challenges to the regime. These challenges include (a) cultivating new sources of project finance; (b) integrating transport projects with regional and local land use plans designed to manage growth; and (c) intensifying competition between central business district (CBD) suburban business centers within the region. The initial discussion focuses on the institutional and normative elements of the regional consensus process. Subsequently, an analysis of contemporary issues illuminates the challenges confronting the regime and the nature of regime responses. The conclusion considers the potential long-term stability of the regime.

NATURE OF THE CONSENSUS

In the 1970s the Interstate withdrawal and transfer process produced a consensus that a light-rail transit (LRT) line linking the City of Portland CBD with the commercial center of Gresham, a suburban city in eastern Multnomah County (which also includes Portland), should be the region's top priority project. In addition to this LRT line, regime decision makers assembled a package of about 140 other highway and transit projects that were to be undertaken using withdrawal funds. The process demonstrated the crucial importance of a regional consensus in the effort to persuade higher-level governmental decision makers to accept this package. Since state and national officials are reluctant to make choices regarding local projects when local officials are themselves in conflict regarding priorities, presenting a united front plays a key role in facilitating financial commitments by higher levels of government. The regional consensus enabled Portland-area representatives in the state legislature to secure necessary state matching commitments and assisted Oregon congressional representatives in securing funding at the federal level (*5*).

The Joint Policy Advisory Committee on Transportation (JPACT) and the Transportation Policy Advisory Committee (TPAC) are the two most important institutional products of the 1970s. The dominant "cultural" feature of these committees is the value they attach to sustaining the 1970s consensus-building process in support of a continuing vision of metropolitan development. This sustaining effort has been greatly facilitated by the relative stability of regime personnel and their commitment to professionalism.

Emergence of METRO

During the 1970s an effort to create a regional government was also under way. Authorized by state legislation and a citizen vote in early 1978, METRO came into existence in January 1979 (Figure 1). The new regional governmental authority was created by merging two existing entities, the Columbia Region Association of Governments (CRAG) and the Metropolitan Service District (MSD). The new entity had an independent board of councilors elected from 12 electoral districts and provided solid waste, drainage, zoo, and planning services in portions of Multnomah, Clackamas, and Washington counties. The effect of creating this new entity was to shift the regional transportation planning process out of the hands of local government officials who had dominated CRAG—representing the respective jurisdictions—to a popularly elected governing body representing districts within the

---

S. Adler, Department of Urban Studies and Planning, and S. Edner, Department of Public Administration, Portland State University, P.O. Box 751, Portland, Oreg. 97207.

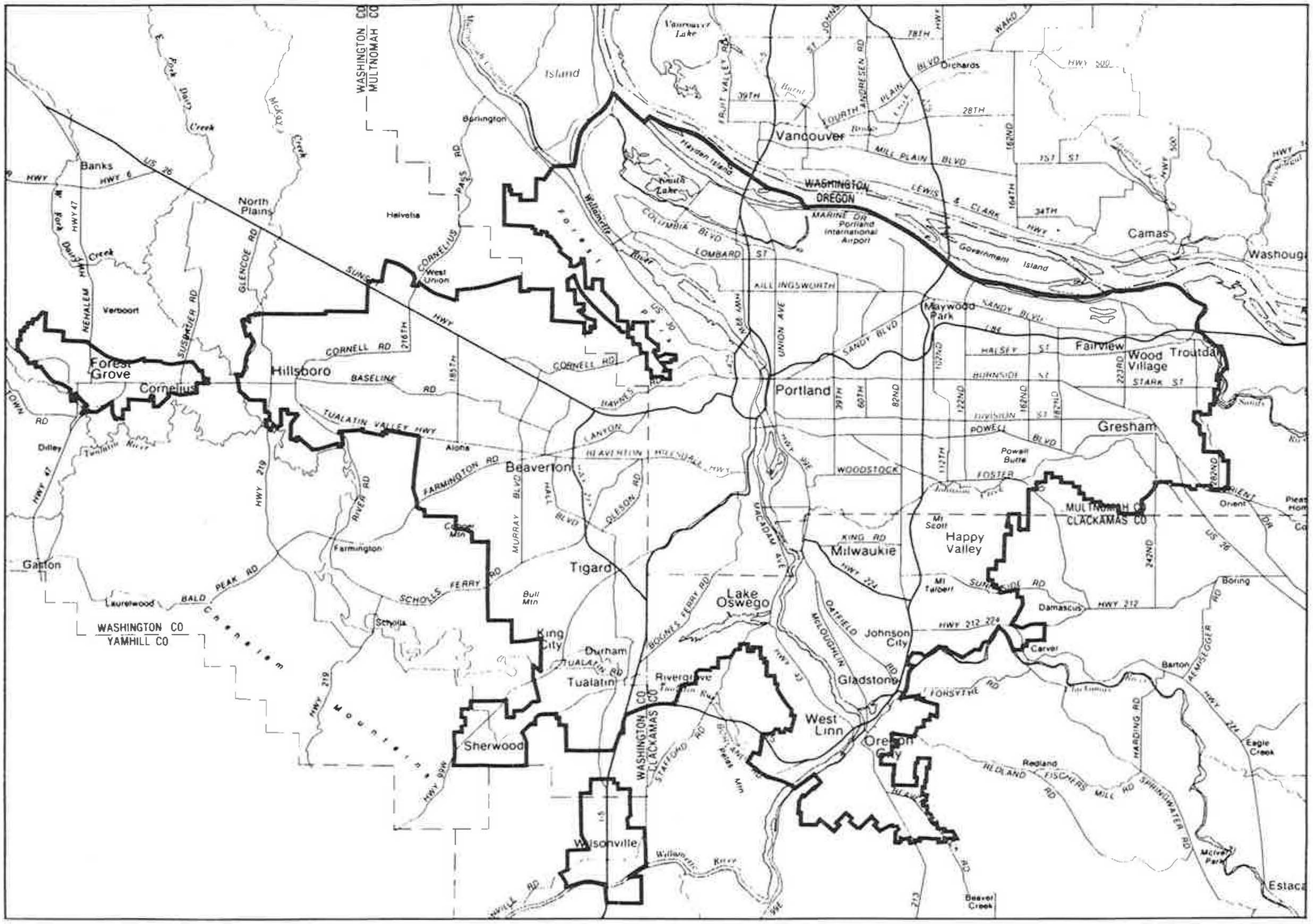


FIGURE 1 METRO boundary.

region. METRO staff assumed responsibility for the development of regional transportation plans, and for support to state and local planning staffs working on transport and land use issues.

Led by its Chairman Neil Goldschmidt, CRAG, in one of its last acts (December 1978), adopted a resolution specifying the projects that would be funded with Interstate withdrawal funds in an effort to set the transport agenda for METRO. METRO attempted to adopt this resolution as the region's federally required transportation plan, but this idea was rejected by the federal government. The U.S. Department of Transportation (DOT) did not view METRO as either a general-purpose government or as composed of representatives of general-purpose governments [a requirement for Metropolitan Planning Organization (MPO) designation]. This was confirmed by a state attorney general's opinion and led directly to METRO's creation of the JPACT in 1979. At the same time, the TPAC was formed to provide technical input to JPACT. Together these supplemental bodies allowed METRO to meet the federal MPO requirements.

JPACT and TPAC

The composition of JPACT has not changed since it was created. Members include elected officials representing the four counties in the bi-state metropolitan area, elected officials representing the cities of Portland, Vancouver, and other cities within the three Oregon counties, the Oregon and Washington departments of transportation, Tri-Met, the Port of Portland, the Oregon Department of Environmental Quality, and the METRO Council. TPAC consists of technical representatives of FHWA, FAA, and UMTA [now the Federal Transit Administration (FTA)], the Intergovernmental Resource Center of Clark County—the designated MPO for that part of Washington state—and six citizen representatives appointed by the METRO Council.

JPACT does not have permanent funding. Local government dues, Highway Planning and Research funds, UMTA planning monies, Tri-Met and Portland grants, and special project funds compose the overall budget of $900,000 (excluding special project funds).

The organizational relationship of JPACT to METRO was never formally defined. There was a split between those involved in creating JPACT regarding which agency would actually make decisions. An unwritten agreement evolved that allowed METRO to approve JPACT decisions, although it was clearly understood that such approval would be a formality. As a result, the regional consensus process could be maintained in a way not available through METRO alone since it does not represent Washington and Oregon state agencies, or other implementing agencies. Thus, JPACT became the defacto MPO body making recommendations to the federal government through METRO.

JPACT deals with issues in an atmosphere that presumes a continuing cooperative process. Projects are debated on their merits, and each project proposal is given a full and fair hearing. Participants describe the process as a "professional game" in which anyone (citizen, professional, or politician) who plays by the rules can have access. This professional atmosphere contributes to the maintenance of high levels of commitment and trust among those involved. The continuity of JPACT's formal membership—and of particular committee members—contributes another important element to the stability of the regime. The political trade-offs and compromises involved in assembling packages of projects and in determining regional project priorities have been easily retained within institutional memory. JPACT has, therefore, been able to integrate new project proposals in a manner that is least disruptive to the underlying consensus.

Another crucial aspect of the culture that sustains the regime is the relationship between technical staff and policymakers, and among the technical officials themselves. METRO's Director of Transportation chairs TPAC and serves as Executive Director of JPACT. Reflecting earlier commitments to upgrade METRO's technical capacity, the METRO staff is large and technically sophisticated and has significant resources available to it for planning and interacting with regional, state, and federal governments. METRO staff works closely with JPACT, aided by TPAC. TPAC has been particularly important in providing the technical expertise necessary to clarify the complexities of project planning and implementation resulting from myriad state and federal funding requirements, especially regarding the sequencing of events. There are also close working relationships between METRO staff and staff in local government planning agencies around the region. On occasion, professional staff from TPAC member agencies are loaned to METRO to assist in completing projects.

Structural relationships also contribute to a high degree of interagency cooperation. When METRO was designated as the MPO, a formal agreement was worked out placing responsibility for socioeconomic forecasting with METRO staff. Therefore, local government planners depend on METRO for population, employment, and transportation forecasts. Although local jurisdictions may suggest and test the consequences of alternative growth and development scenarios, the regional data base used for these projections is maintained by METRO and any changes must be adopted by METRO. The high level of interaction among technical people and between technicians and political officials has tended to produce a substantial level of agreement on transport policy issues, as well as on project priorities over time. This is consistent with the findings of Schmitt et al. (*6*), who found that MPO effectiveness in coordinating transportation planning at the local level was more a function of history and the attitudes of the people involved than of any MPO structural characteristics. The stability of the participants has further reinforced this tendency.

In addition to these institutional and political culture factors, the great weight of the Portland CBD in the regional political economy undergirds the transport regime. During the period of the withdrawal of funds, downtown Portland was just beginning to confront suburban pressures for office and commercial investment. The choice of a downtown/radial LRT line—and the reconstruction of a downtown/radial freeway sharing the same corridor—reflected the weight of the Portland CBD in the consensus-building process. Several suburban business centers have, however, grown rapidly in recent years, and the now-familiar pattern of suburban gridlock is evident in these outlying areas. Nevertheless, downtown Portland maintains its dominant position in the region, even as it faces increasingly intense competition for investment in the activities historically concentrated there. This dominant role is widely, if grudgingly, acknowledged, and the recognition serves as a counterbalance to pressures that would fragment the consensus.

Finally, leading members of the Oregon congressional delegation emerged as effective advocates of the regional consensus with the national government during the withdrawal period and have continued to play this role. The absence of any competing major metropolitan areas in the state has enhanced the willingness and the capacity of the delegation to serve as an effective advocate.

NATURE OF THE CHALLENGES

The Portland regime now confronts a set of challenges that generate conflicting pressures on the consensus process. The first challenge is financial. Regional transport activists anticipate a continuing reduction in the level of federal funding for projects, necessitating an increase in state, local, and private-sector contributions to offset the decline (*7*). On the one hand, reductions in the level of federal investment generate pressure to maintain the regional consensus to effectively compete with other metropolitan areas for a share of a shrinking federal pie. On the other hand, the prospect of lowered funding levels increases the intensity with which project sponsors advocate a higher priority on behalf of particular projects and, therefore, the level of conflict that the consensus process must contain. Moreover, the necessity of including private-sector financial participation may exacerbate these conflicts, because places with poor private developmental prospects will be less able to secure commitments to projects.

A second challenge is the integration of transport projects with comprehensive land use plans. Planners and planning processes were greatly strengthened in Oregon by the environmental movement that surfaced in the 1970s. The new environmentally based planning challenges the relatively narrow concerns of the transportation regime. In the face of this challenge, regime supporters may close ranks to defend against a broadening of perspective that might compromise the effectiveness of transport projects as facilitators of economic growth. At the same time, effective environmentally based planning is widely seen as a crucial competitive advantage enjoyed by metropolitan Portland, as the region struggles to attract investment from Seattle and California urban centers that are portrayed as choking on unplanned urban sprawl. The desire to maintain this widely appreciated competitive advantage creates pressure to restructure the regime as a way of tipping the balance of power between transport projects and land use plans toward the latter.

Generally, the most basic challenge to any regional consensus process is competition between subregions to maintain and attract investment. Competition causes each business center to sponsor transport projects that will enhance its locational attractiveness. Projects become weapons that are deployed to gain competitive advantage. Those projects that concentrate benefits in space, as freeways and rail transit lines do around interchanges and stations, generate the most controversy because they will give a great advantage some places and a disadvantage to many others (*8,9*). The package of projects assembled during the withdrawal process enabled the Portland regime to avoid the political stalemates that have characterized other areas.

As noted above, however, the Portland CBD faces increasingly intense competition from suburban business centers, though it continues to maintain its dominant position in the regional economy. The consensus process now includes more assertive suburban activists seeking projects that will enhance the autonomy of their centers. Designing individual projects and assembling packages of projects that will serve the development aspirations of the many competing regions represented in this regime becomes more problematic. At the same time, the continued dominance of downtown Portland encourages CBD activists to insist on the priority of projects that will maintain the privileged position of their region.

The consensus that was constructed during the 1970s was built using federal funds, so the challenge of making up for recent reductions in these contributions has only recently surfaced. Spreading 140 projects around the region effectively addressed the challenge of competition in the 1980s. The LRT line and the reconstructed freeway primarily benefited downtown Portland; however, the location of LRT stations enabled CBD activists to form an alliance with their counterparts in Gresham at the end of the line. Including in the package the myriad highway and transit projects throughout the metropolitan area induced activists from suburban Washington and Clackamas counties to support the consensus. Finally, the 1970s coincided with the start-up phase of Oregon's statewide land use planning program. The jurisdictions in the region, including METRO, were busily mapping urban growth boundaries and preparing a comprehensive land use plan that would accommodate a substantial amount of development. The projects included in the original consensus did not engender any significant conflicts with emerging land use goals. How well a transport regime born in an era of plenty can hold up to these internal and external challenges is examined in the case studies presented below.

WESTSIDE BYPASS

The proposed Westside Bypass (Figure 2) presents a more serious land-use-based challenge to the regime for two reasons: (a) many jurisdictions are directly affected by the bypass, including several suburban cities and Washington County at the western edge of the region, and (b) part of the proposed route would lie outside of the METRO-adopted regional urban growth boundary. The bypass will be a north-south arc traversing the western edge of the metropolitan area and has raised the specter of traffic-induced sprawl. The bypass has been Washington County's top-priority project for some time, especially for the western part of the county, which has experienced an electronics-industry-related employment boom. County activists have been pointing out that traffic increasingly circulates within the county instead of between the county and Portland and that the bypass was crucially important to serve the growing economic autonomy of the area.

Following JPACT and METRO Council decisions to add the bypass to the regional consensus priority list, *The Oregonian* lectured Washington County officials on their responsibilities. A lecture was seen as necessary because of the per-

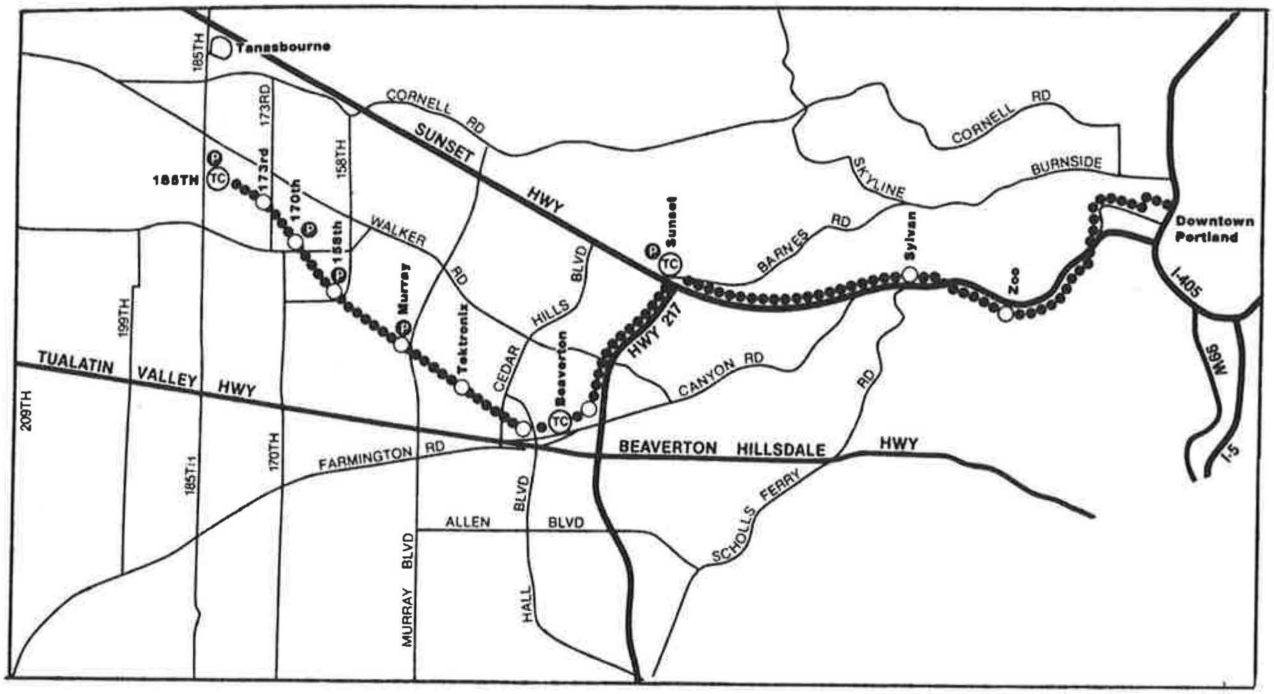


FIGURE 2 Westside Bypass.

ception that Washington County had historically been insufficiently attentive to regional concerns, had displayed a "go-it-alone," independent attitude, and was prone to pursue narrow, growth-related objectives.

The Oregonian (*10*) editorialized two challenges to county officials:

> One is to show that participation in a metropolitan partnership is genuine and not just lip service masking insular inclinations. The other is to show enough backbone in controlling development to give the rest of the region confidence that the bypass would not lay open land supposedly protected by the urban growth boundary.

Washington County's Director of Land Use and Transportation Planning, a TPAC member, had already promised to work out a program immediately to resolve land use issues with METRO. JPACT approved funds for Washington County to begin preliminary engineering. At that point, the bypass project was conceived as a long-term effort that would be built in stages, the first stage of which would not involve any growth-management-related concerns. The sections of the bypass that would cross over the urban growth boundary were several years away.

The land use brew thickened when the Washington State legislature expressed interest in a third bridge over the Columbia River, several miles west of the existing Interstate bridge on the freeway connecting downtown Portland and Vancouver, Washington. This project was especially exciting to Clark County officials as one that would encourage the Oregon members of the regional transport regime to extend the Westside Bypass in a northeasterly direction all the way to the new river crossing, providing a much more direct connection between Clark County and industrial development zones in western Washington County. The bypass extension would, however, cross Portland's Forest Park, one of the largest urban parks in the United States, much of which is forested. Intense environmental opposition soon surfaced in Portland and in Clark County because of the routing of the connecting highway. The Intergovernmental Resource Center, the Clark County MPO, strongly supported the idea and was studying it at the request of the Washington state legislature. The legislature recommended a much more elaborate third bridge study to be jointly funded by Washington and Oregon agencies. Clark County officials readied a presentation to JPACT to secure a regime commitment. A Washington County transportation planner approved the logic of extending the bypass. However, a city of Portland transportation planner and TPAC member sounded a note of caution, arguing that before funds were committed to a third bridge study, a policy question ought to be resolved: Was opening up new lands for economic growth the objective, or was relieving traffic congestion on an existing bridge the main concern? This planner thought the former goal was uppermost in the minds of Clark County officials. TPAC's chair, who was also METRO's Director of Transportation Planning, thought so as well. The Portland regime's agenda, however, was concerned with the latter.

When the Clark County members of JPACT presented their bistate study idea to TPAC, the technical officials clearly indicated that a third bridge would be a low-priority item on the regional transport project list. They recommended against regime participation in the study. TPAC's chair recommended instead that METRO and Clark County's Intergovernmental Relations Center cooperate in a study of congestion on existing routes and the possibilities of extending light rail transit into the county. A Clark County official asked that a vote on this recommendation be delayed while further discussions were pursued. JPACT later agreed to delay a decision.

When JPACT did consider Clark County's invitation to a study, the Clark County officials' presentation of their request markedly changed. They now stressed traffic congestion in general and the third bridge as one possible answer to corridor congestion concerns. JPACT later decided to reject any examination of a third bridge, but to support a bistate study along the lines suggested by the TPAC chair. This was to involve an earlier and more concentrated look at possible LRT lines linking downtown Portland and the city of Vancouver than JPACT had previously contemplated. The Washington State JPACT delegation was pleased, having decided to give priority to LRT in the face of the third bridge idea. JPACT member Earl Blumenauer, a leading LRT supporter, noted that the city of Portland would likely participate financially in such LRT studies.

The resolution of this issue illustrates the way in which the regime was responsive to Clark County concerns. The regime responded in a manner that maintained the integrity of the existing consensus regarding project priorities, as well as recognized the dominant role of downtown Portland—and its congestion concerns—in the regional political economy.

Although JPACT dealt with Clark County issues, the level of conflict on the Westside Bypass dramatically increased. 1000 Friends of Oregon, the state's land use law watchdog, legally challenged the status of the bypass in Washington County's transportation plan on the grounds that land use issues had not been addressed before the county committed to the freeway in its plan. The state Department of Land Conservation and Development (DLCD), which administers Oregon's land use laws, filed a separate appeal, arguing that the county's plan lacked an explicit agreement that the bypass would be dropped if state land use requirements were not met. There had been an understanding on this point between the county and the state. The state had permitted the county to include the bypass to facilitate the search for financial support through the consensus process. However, the official plan failed to make the understanding explicit. Both 1000 Friends and DLCD were concerned about the sprawl-inducing effects of those portions of the bypass located outside the existing urban growth boundary.

The Washington County Board of Commissioners quickly moved to include the explicit agreement sought by DLCD and to do a study that would determine whether the bypass should be excused from land use laws protecting agricultural land outside urban growth boundaries. In addition, the county committed itself to study alternatives to the bypass in case the project could not be exempted under state land use laws. DLCD then dropped its appeal. However, 1000 Friends refused to follow DLCD's lead, intending to press their appeal unless Washington County dropped all reference to the bypass in its transportation plan and addressed the land use question of whether a freeway was needed in that corridor. The attorney representing 1000 Friends said, "We think this is going to be used to bust the urban growth boundary" (*11*). This constituted a direct challenge to the transportation regime, which had approved the bypass in JPACT's and METRO's transport plan.

The regime countered with a METRO Council decision to intervene in the case being heard by the Land Use Board of Appeals (LUBA):

> METRO should be "at the table," the Council's presiding officer said, if 1000 Friends and Washington County wind up negotiating some sort of settlement that could affect METRO's system for reaching consensus on transportation problems. (*12*)

LUBA decided in favor of 1000 Friends' position ordering Washington County to redo its transportation plan so that it did not appear that the county was already committed to the project. Land use analyses aimed at determining whether a freeway was needed in the corridor had to come first, according to LUBA. 1000 Friends hailed the LUBA decision as a "major rerouting of the mindset" regarding the project. Washington County and METRO representatives disagreed, claiming that LUBA had not altered the transportation planning process they had been following. County commissioners were saddened, however, that the decision would set back construction of the crucial first stage of the project for 2 years while the need for the entire freeway was restudied.

The Oregonian saw the LUBA decision as a victory for comprehensive planning, which would consider land use and transportation issues together. In the wake of the bypass ruling, a DLCD official pointed out the implications of the decision for projects being planned elsewhere in the region, particularly in Clackamas County, where similar urban growth boundary concerns were likely to emerge. State land use and transportation officials began meeting to clarify the rules regarding when the transportation planning process would have to take state land use goals into account, particularly those concerning defense against urban sprawl, and when particular projects would be exempt. The TPAC chair pointed out that current rules were unclear and welcomed the effort at clarification.

The problem of integrating land use plans and transport projects has been noted in analysis of the state land use planning program (*13*). Integration is clearly more than a technical challenge for the regime. 1000 Friends of Oregon is an extremely well-organized, attentive, and politically and technically sophisticated organization. The organization has prospered in its watchdog role and it stands ready and willing to oppose the regime on land use and environmentally related transport matters. The group members will closely monitor the formulation and implementation of the rules that emerge from the process of clarification now under way. The question is, Will the regime accommodate plans, or will the logic of the project continue to drive the urban development process?

LIGHT-RAIL TRANSIT PLANNING

After the federal government committed funds to build the eastside LRT line in place of the Mount Hood Freeway in 1981, the regime turned its attention to the next LRT project. A westside line, linking downtown Portland and the eastern Washington County city of Beaverton, emerged as the top-priority project. From the viewpoint of Portland CBD leaders, this route would add commuter capacity in the increasingly congested Sunset Highway Corridor. The proposed line would not, however, extend significantly into western Washington County.

JPACT had endorsed a particular westside LRT alignment in 1983; however, the Tri-County Metropolitan Transit District (Tri-Met) hesitated to apply for funds from the federal government to begin preliminary engineering on this project because it lacked the financial resources to apply for a matching grant. In addition, however, opposition to the project had surfaced in western Washington County. At the same time, Clackamas County transportation and economic development planners had joined with planners at the Port of Portland to propose an LRT line running south from Portland International Airport through a connection with the eastside LRT line to a major regional shopping center in Clackamas County. This line was to share a corridor with the recently constructed eastside bypass freeway, I-205, for most of its length.

Washington County transport activists, particularly those in the west, were shifting their focus away from a westside LRT line to the bypass freeway. Applauding this shift, *The Argus*, one of Washington County's leading newspapers based in the western county seat of Hillsboro, editorialized, "We should far rather see . . . money . . . spent to facilitate north-south traffic in Washington County—where it is urgently needed—rather than boosting travel from our County into downtown Portland" (*14*). The fact that the JPACT-approved alignment did not reach into the western portion of the county—even though all Washington County business firms paid a payroll tax to Tri-Met—underlined what many in the west believed was yet another effort by downtown Portland to use transit to capture Washington County office workers and shoppers.

Clackamas County officials stressed the complementary nature of their LRT proposal, arguing that the airport link would boost patronage on the eastside LRT line. However, they also pointed out that their LRT line would cost a great deal less than a westside line, since virtually all the necessary right-of-way was in governmental hands, and that it could be constructed much more quickly than a westside counterpart. JPACT responded by authorizing a search for possible funding sources.

Tri-Met began preliminary engineering on the westside line in 1987 (Figure 3). During the course of this work, Tri-Met was reluctantly forced to conclude that the 1983 JPACT-approved alignment was no longer acceptable. During the intervening years, parts of the approved route had been built over, and a great deal more employment and residential development had taken place in the western portion of Washington County than had been anticipated. The new alignments to be considered would reach all the way to Hillsboro in the west. What was most disturbing about the alignment discovery was the prospect of having the regime go through the regional consensus process once again to reach agreement on a new route.

While westside LRT proponents contemplated major changes in their project, Clackamas County government and technical officials continued to press their LRT line forward as a high-priority endeavor. These officials called for building both westside and I-205 LRT lines as part of a regional package, ushering in a new era of cooperation between Clackamas and Washington counties, which historically had found themselves competitors for limited transport funds. Clackamas County's assertiveness troubled Washington County officials, however, particularly when the former attempted to compare the I-205 line with the westside line in a manner that reflected less favorably on the latter's regional priority. A Washington County commissioner and JPACT member said she thought Clackamas County officials no longer supported westside LRT as the region's top priority. She argued that adding more lines to the regional list would jeopardize funding for Number 1: "You can't do a little part of everything, or nothing gets done" (*15*). A Clackamas County commissioner, who was also a JPACT member, countered that he was "shocked and appalled" by the accusation that his county was trying to disrupt the regional consensus (*16*). Clackamas officials stressed their allegiance to the regime and to the westside LRT as they continued to argue that there would not be any competition between the two projects because various funding sources likely would be involved.

Extending westside LRT to downtown Hillsboro would permit that city to join an alliance with downtown Portland and downtown Beaverton. An appropriately chosen alignment would enable the three business centers to coexist peacefully, each enjoying an enhanced capacity to reach potential workers and shoppers locating in the rapidly growing parts of Washington County. The Hillsboro city council strongly supported efforts to bring LRT to town, as did the city council of Forest Grove, a small city beyond Hillsboro that hoped to secure a place on the line for the future.

However, two obstacles were present. The first was the concern that adopting a new alignment all the way to Hillsboro would necessitate a restudy of the entire project, delaying efforts for up to 2 years. The federal government had approved preliminary engineering funds solely for the 1983 alignment. In an effort to dispel any thoughts about foregoing the extension, a Washington County newspaper (*17*) editorialized the crucial significance of the Hillsboro connection:

> A project that ends (at the point designated by JPACT in 1983) will primarily serve commuters travelling between Beaverton and downtown Portland. A project that extends to Hillsboro will not only serve a greater number of commuters, but it will also provide a new route for intracounty travel. . . . Indeed, the support of many Washington County residents and officials is contingent upon seeing the line extended to Hillsboro. Proceeding with plans to build the line only to (the 1983 end point) could fracture consensus that is starting to grow for the light-rail project.

Washington County's Director of Land Use and Transportation Planning, a TPAC member, voiced the same concern.

The other obstacle involved the politics of route selection. An association representing land development and industrial firms in the corridor adjacent to the Sunset Highway argued for an alignment that would follow the highway all the way from downtown Portland to Hillsboro, relocating that portion of the LRT route in the central portion of the county between Hillsboro and Beaverton. Beaverton's planning director opposed this alignment, arguing that the Sunset Highway Corridor route would not support the land use plans of many Washington County cities, including Beaverton. A citizens advisory committee to the Tri-Met Board of Directors recommended that the Hillsboro-Beaverton route, instead of the Sunset Highway Corridor, be chosen. This route had also been approved by the local governments in the county.

While westside LRT work went forward, Clackamas County officials became increasingly concerned that their LRT project was languishing. The County's Director of Transportation Planning, who was a TPAC member, thought that "Portland and Washington County have formed a pretty formidable coalition. What they want, they get. Unfortunately, they do not speak for Clackamas County" (*18*). Earl Blumenauer bri-

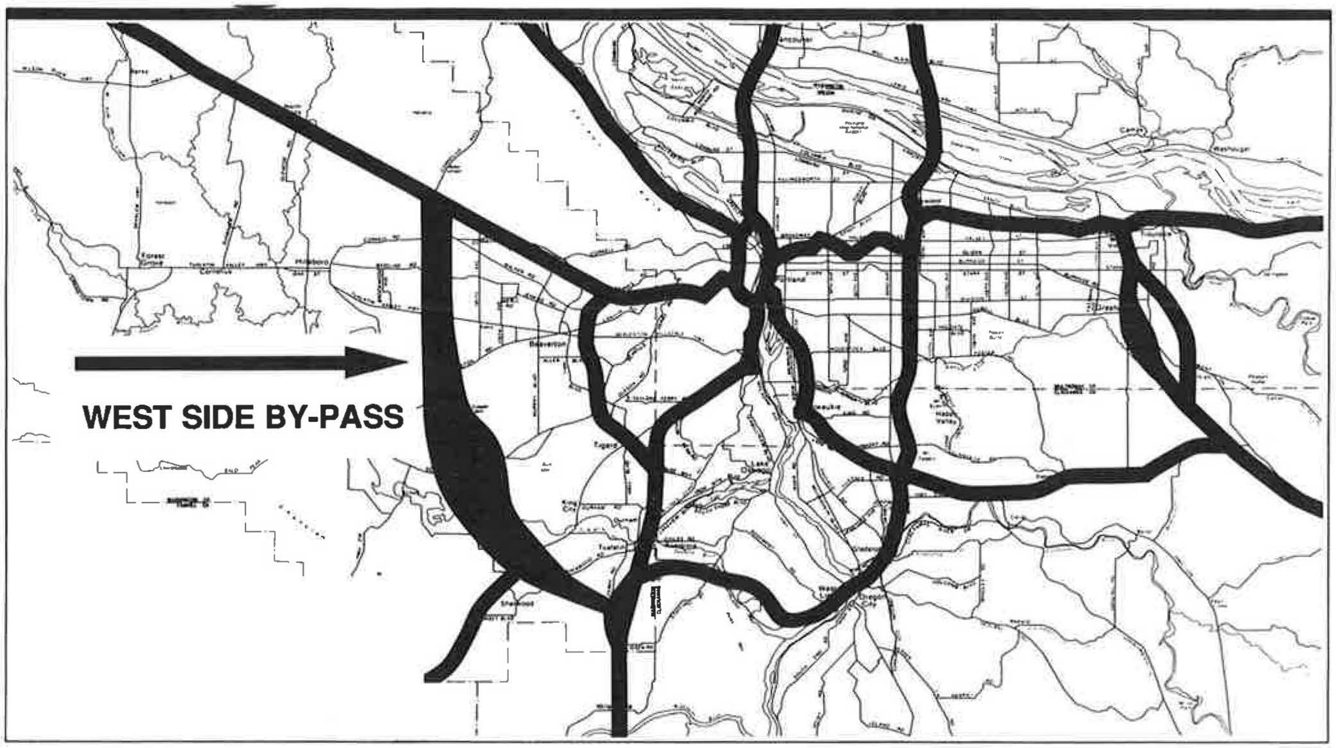


FIGURE 3 Alignment of proposed Westside light-rail project (heavy lines indicate major highway corridors) (Source: Tri-Met).

dled at the accusation that political power tactics were keeping LRT out of Clackamas County. He defended his cooperative efforts on behalf of county interests and his support for bringing LRT to the county.

This was clearly a critical period for the regime. Competition for LRT funds was intensifying. Bringing western Washington County into the LRT alliance was problematic. At the same time, the Westside Bypass project was under attack, and Clark County officials were concerned about the capacity of the regime to meet their needs. A leading member of the Oregon congressional delegation, whose district includes Washington County, announced that he intended to lead an effort to seek additional federal funds to study the Hillsboro westside LRT extension and to permit this study to be done without having to restudy the entire route, thus saving 2 years. The significance of this plan is that the project would be eligible for 75 percent federal funding, instead of the 50 percent contribution that the federal government intended to institute when its funding formula expired in 1991.

The Oregon delegation was successful. House and Senate transportation bills incorporated language directing UMTA to permit the Hillsboro extension to be conducted without delay and providing funds for the work. Moreover, the Oregon delegation's efforts stimulated a similar effort by their Washington state counterparts, who lined up funding for LRT studies in a broad corridor including Clackamas County, Portland, and Clark County. Two possible Clackamas County–Clark County alignments, one via I-205 and another via downtown Portland, would be eligible for study within this broad corridor notion. The LRT program for the entire region was boosted, greatly enhancing the stability of the regime.

The nature of rail transit projects—the fact that they concentrate access around a small number of station locations—greatly complicates the process of regional consensus building. The Portland transport regime illustrates one possible solution to the problem posed by the spatial concentration of benefits: a "coverage" strategy, including alliances between leading business centers based on transit network design. The success of the congressional delegations, in turn based on the underlying stability of the regime, enabled the regime to pursue several rail transit possibilities simultaneously, thereby covering the region with potential rail transit benefits. The regime has also been sensitive to issues of alliance-sustaining network design. The coverage strategy requires large amounts of funding, however, to make it work (*19*). Finding the funds to construct and operate these projects is the remaining challenge to the regime that we discuss.

FUTURE OF REGIME

The case studies illustrate the ways in which the Portland transport regime has responded to changes in its environment. The financial challenge to JPACT includes greater dependence on state and local sources for planning monies, which might disturb the consensus process. If one or a small number of regional or local jurisdictions ends up shouldering a much larger share of the burden of financing JPACT studies—Tri-Met, for example—the regime's consensus building process will be more complicated than when federal funds were shared around the metropolitan area. The challenge of competition is also reflected in the concern felt by smaller cities in the region that they are under-represented on JPACT. As these smaller cities grow they will become increasingly interested in projects that will enhance their position in the metropolitan economy, and in articulating their interest through JPACT.

Pressures on JPACT to respond to this concern will likewise complicate the consensus-building process.

Integrating transport projects and land use plans is a different sort of challenge from dealing with funding shortfalls and an increasing number of project sponsors. These latter sorts of difficulties produce conflicts, but these conflicts are internal to the regime. Often led by its technical members, the regime has been able to learn to deal with them without undermining the culture of consensus. The fact of downtown Portland's continuing dominance in the regional hierarchy and the absence of rapid economic growth pressures should moderate disruptive tendencies, permitting the regime to continue to deal with these conflicts in the ways that have produced the region's past successes. Conflicts between transport project and land use plans, however, arise from sources external to the regime; the package assembly approach to maintaining consensus does not apply.

The integration conflict represented by the Westside Bypass challenges the regime in two profound ways. One issue is the possibility of policy learning across belief systems; another involves the relationship between JPACT and METRO. Learning within a policy domain, such as urban transportation, most readily takes place when the core values adhered to by activists within the domain are not challenged. The funding and competition conflicts do not directly call into question the core values of the transport regime, which involve assembling packages of growth-facilitating transport projects. The regime has learned and continues to learn how to deal with these. However, integrating transport projects and environmentally based land use plans, as represented by the urban growth boundary, necessitates learning across systems. 1000 Friends of Oregon and the other citizen activist groups opposing the bypass hold to a different set of core values, including the priority of environmental goals and the subordination of growth-facilitating transport projects to these goals.

Sabatier hypothesizes (*20*) that policy-oriented learning across belief systems is most likely to occur when there is an intermediate level of conflict between advocates and when experts of the respective coalitions are forced to confront each other in a forum that is dominated by professional norms. Given the legal and political resource of its environmental challengers, the question is whether METRO, JPACT, and TPAC will attempt to defeat their opposition in an effort to maintain the integrity of the existing consensus on projects or whether the professionalism that has characterized JPACT and TPAC deliberations can be extended to include the advocates of land use and environmental planning.

Finally, the integration challenge calls into question the nature of the informal relationship between METRO and JPACT that has prevailed since JPACT was created. Since METRO has jurisdiction over the urban growth boundary for the region as well as the regional transportation plan, METRO may have to take a more distant, critical view of JPACT's project recommendations to address the land use concerns that are also within the agency's domain. Perhaps the conflicts surrounding integration will produce a forum that will enable METRO transportation and land use professionals to produce the policy-oriented learning necessary to transcend the existing limits of Portland's transport regime.

REFERENCES

1. C. Stone. *Regime Politics: Governing Atlanta 1946–1988*. University of Kansas Press, Lawrence, 1989.
2. S. Krasner. *Structural Conflict*. University of California Press, Berkeley, 1985.
3. R. Keohane. *After Hegemony*. Princeton University Press, Princeton, N.J., 1984.
4. O. Young. *Resource Regimes*. University of California Press, Berkeley, 1981.
5. S. Edner and G. B. Arrington, Jr. *Urban Decision Making for Transportation Investments: Portland's Light Rail Transit Line*. U.S. Department of Transportation, 1985.
6. R. Schmitt, E. Beimborn, and J. Weitman. Local Organization of Transportation Planning Activities in the United States. *Transportation Planning and Technology*, Vol. 9, 1984, pp. 225–236.
7. E. Weiner. Redefinition of Roles and Responsibilities in U.S. Transportation. *Transportation*, Vol. 12, No. 3, 1984, pp. 211–224.
8. S. Adler. Infrastructure Politics: The Dynamics of Crossing San Francisco Bay. *The Public Historian*, Vol. 10, No. 4, 1988, pp. 19–41.
9. S. Adler. Why BART but no LART? The politics of Rail Rapid Transit Planning in the Los Angeles and San Francisco Metropolitan Areas, 1945–57. *Planning Perspectives*, Vol. 2, No. 2, 1987, pp. 149–174.
10. Challenges of a Bypass. *The Oregonian*, Portland, Oct. 20, 1987.
11. A. Green. 1000 Friends Say It Won't Drop Freeway Appeal. *The Oregonian*, Portland, Feb. 2, 1989.
12. METRO Set to Enter LUBA Case. *The Oregonian*, Portland, Feb. 27, 1989.
13. J. DeGrove. *Land, Growth and Politics*. American Planning Association, Chicago, 1984.
14. Editorial. *Argus*, Hillsboro, Oreg., March 12, 1985.
15. A. Green. Counties Each Push Transit Plan. *The Oregonian*, Portland, Oct. 6, 1988.
16. V. Kohler. Lindquist Denies County Light Rail Effort Would Derail Westside Bid. *The Oregonian*, Portland, Oct. 10, 1988.
17. See That Light Rail Gets to Hillsboro. *The Valley Times*, Washington, County, Oreg., Feb. 23–March 1, 1989.
18. M. Gerdes. Will Light Rail Come to Clackamas County? *The Review*, Clackamas County, Oreg., March 10, 1989.
19. S. Adler and S. Edner. Governing and Managing Multi-Modal Regional Transit Agencies in a Multi-Centric Era. In *Managing Public Transportation*. (G. Guess, ed.), Greenwood Press, Westport, Conn., 1990.
20. P. Sabatier. An Advocacy Coalition Framework of Policy Change and the Role of Policy-Oriented Learning Therein. *Policy Sciences*, Vol. 21, 1988, pp. 129–68.

Publication of this paper sponsored by Committee on Intergovernmental Relations and Policy Processes.