1236

TRANSPORTATION RESEARCH RECORD

Transportation Data 1989

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1989

Transportation Research Record 1236
Price: $12.00

mode
1 highway transportation

subject areas
12 planning
13 forecasting

TRB Publications Staff

Director of Publications: Nancy A. Ackerman
Senior Editor: Edythe T. Crump
Associate Editors: Naomi C. Kassabian
Ruth S. Pitt
Alison G. Tobias
Production Editor: Kieran P. O'Leary
Graphics Coordinator: Karen L. White
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 data, 1989.
p. cm.—(Transportation research record, ISSN 0361-1981 ; 1236)
ISBN 0-309-04958-X
1. Traffic surveys. 2. Traffic estimation. 3. Traffic flow.
I. National Research Council (U.S.). Transportation Research Board.
II. Series.
TE7.H5 no. 1236
[HE369]
388 s—dc20
[388.3 14] 90-33749
CIP

Sponsorship of Transportation Research Record 1236

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Chairman: Ronald F. Kirby, Metropolitan Washington Council of Governments

Transportation Data, Economics, and Forecasting Section
Chairman: Edward Weiner, U.S. Department of Transportation

Committee on Transportation Data and Information Systems
Chairman: Alan E. Pisarski, Falls Church, Virginia
Secretary: Rolf R. Schmitt, Federal Highway Administration, U.S. Department of Transportation
W. Bruce Allen, Kenneth J. Dueker, Douglas C. Frechtling, Philip N. Fulton, Larry R. Goode, David Lloyd Greene, David T. Hartgen, Brad H. Johnson, Don C. Kirkpatrick, Santo J. LaTores, James J. McDonnell, Wayne K. Minger, Gordon A. Shunk, Arthur B. Sosslau, Marcus Ramsay Wigan

Committee on Traveler Behavior and Values
Chairman: Mary Lynn Tischer, Federal Highway Administration, U.S. Department of Transportation
Elizabeth Deakin, James C. Echols, David T. Hartgen, David A. Hensher, Joel L. Horowitz, R. Ian Kingham, Ryuichi Kitamura, Lidia P. Kostyniuk, Stephen C. Lockwood, Hani S. Mahmassani, Patricia L. Mokhtarian, Eric Ivan Pas, Richard H. Pratt, Earl R. Ruiter, Peter R. Stopher, Antti Talvitie, Carina Van Knippenberg, Edward Weiner, William Young

James A. Scott, 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, 1988.

NOTICE: The Transportation Research Board does not endorse products or manufacturers. Trade and manufacturers' names appear in this Record because they are considered essential to its object.

Transportation Research Board publications are available by ordering directly from TRB. They may also be obtained on a regular basis through organizational or individual affiliation with TRB; affiliates or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Research Council, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.

Transportation Research Record 1236

Contents

Foreword

In the paper "Simulation Results of the Highway Performance Monitoring System," Memmott examines the adequacy of the Highway Performance Monitoring System (HPMS) sample for making need estimates at the district level (Texas) and recommends increases in the current sample based on the results of the examination. A simulation model is developed to calculate errors of a given sample size. Also, the simulation model is used to test the usefulness of stratifying the functional classes by volume group.

In the paper "Development of State Traffic Monitoring Standards: Implementing the *Traffic Monitoring Guide* in New Mexico," Albright et al. review current practices as published in the *Traffic Monitoring Guide* (FHWA) and conclude that state traffic data requirements could best be met, and the *Traffic Monitoring Guide* could be best implemented, by developing statewide traffic monitoring standards. Standards were drafted, refined, and adopted. All traffic data are required to be in conformity with the state standards.

In his paper "Geographic Information Systems: An Important Technology for Transportation Planning and Operations," Simkowitz defines Geographic Information Systems (GIS), describes how a number of transportation agencies are using GIS, and further examines the usefulness of spatially integrated data to transportation and clarifies the distinction between GIS and other data base systems that use spatial data.

Stokes and Chira-Chavala in their paper "Design and Implementation of Intercity Origin-Destination Surveys" discuss and summarize the results of a postcard origin-destination (O-D) survey conducted to estimate intercity travel patterns in the I-38 corridor between Austin and San Antonio. The results indicate that the postcard O-D survey method can be used to develop representative, reliable estimates of intercity corridor travel patterns.

Gur and Hocherman, in their paper "Optimal Design of Traffic Counts," describe the control traffic counting program in Israel whereby week-long mechanical counts are performed at each rural highway link every 1 to 2 years. They discuss possible ways to increase the program's efficiency by counting only a sample of links every year.

In the paper "On-Line Generation of Synthetic Origin-Destination Counts for Application in Freeway Corridor Traffic Control," Van Aerde et al. discuss the need to provide O-D matrices for each time slice during the peak period to be analyzed. The authors propose a fully automated approach. The approach relies on existing algorithms for formulating synthetic O-D counts from observed link flows. However, it uses a special relationship that exists between O-D matrices for consecutive time slices to carry out these computations more efficiently and often also with greater accuracy.

Walker, in his paper "Method to Synthesize a Full Matrix of Interdistrict Highway Travel Times from Census Journey-to-Work Data," discusses the incompleteness of census data in that most cells in the interdistrict matrix have no census travel time observation. The paper explores the accuracy and usefulness of census travel time data versus survey- and network-based methods and airline distance approaches at the Minor Civil Division (MCD) level of analysis.

Goulias and Kitamura, in their paper "Recursive Model System for Trip Generation and Trip Chaining," discuss a model system developed to describe both trip generation and trip chaining in a coherent manner. The model systems described offer theoretically consistent coefficient values and quantify the relationship between the number of trips and the number of trip chains and can be used in the conventional forecasting procedures in place of home-based and non–home-based trip generation models.

Simulation Results of the Highway Performance Monitoring System

JEFFERY L. MEMMOTT

This paper covers a detailed examination of the adequacy of the Highway Performance Monitoring System (HPMS) sample for making needs estimates at the district level in Texas and recommends increases in the current sample based on the results of the examination. To test the accuracy of sample sizes, a simulation model was developed to calculate the errors of a given sample size. Since the FHWA procedure uses average annual daily traffic (AADT) to calculate the required sample size, the simulation model was used to compare the accuracy of needs estimates with the assumed accuracy using AADT. It was found that in general, the errors were larger than the assumed error range at the functional class level, but the errors decreased substantially as functional classes were aggregated. The simulation model was also used to test the usefulness of stratifying the functional classes by volume group. It was found that in most cases stratifying did improve the accuracy of the sample estimates. Further, it was found that in most cases calculating the sample at the functional class level and distributing the sample to volume groups proportionately by mileage performed better than calculating the sample at the volume group level.

The Highway Performance Monitoring System (HPMS) (*1*) was developed by the FHWA to collect data on a large sample of highway sections throughout the United States and to make estimates on the current condition of the highway system and future needs, including effects of different funding levels. For that purpose, each state was required to select a stratified random sample of highway sections and to collect several items of information on each section. They are further required to maintain and update that information with annual submittals to FHWA. The sample covers all public roads above the local functional class.

The FHWA also developed a package of computer programs to summarize and analyze the data submitted by the states. The programs provide an analysis of the current or existing condition of the highway system and a number of options that consider future needs as well as impacts of different funding limitations. The basic procedure the analysis package uses is first to estimate the current condition of the sample highway sections. Those conditions are then compared to minimum tolerable conditions. For those sections that have values falling below those minimum values, an improvement is simulated. Both the type of improvement needed and the construction cost are estimated internally within the program. If a funding limitation is imposed, then the program selects the highest ranked needed improvements until the next funding period.

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

The FHWA also provided to the states a version of the analysis package for use at the state level. Since the states must collect the HPMS information anyway, it would be advantageous to make use of that data if it can successfully be adapted to the needs within Texas. That is the purpose of this study, to examine the sample and analysis package for possible adaptation and use in Texas.

One potential use of the HPMS data and analysis package for Texas is to provide information on the current condition of the highway system in Texas and estimates of future needs, similar to what is done at the federal level. Currently the Texas Department of Highways and Public Transportation (TSDHPT) compiles a document called the Strategic Mobility Plan (SMP) (*2*). This gives estimates of 20 year needs of the department and is updated every 2 years. The estimates are based on a combination of the projects submitted by the districts, which cover the anticipated needs over the next 20 years, and a computer program, which estimates aggregate rehabilitation and maintenance needs over the same 20-year period.

The HPMS has the potential to be used in combination with, or as a substitute for, the current procedure. Use of HPMS sample data would eliminate the need for submitting individual projects for estimating needs, and estimating several categories of needs together in one analysis eliminates the double counting in the current system. For example, the same highway section could have an added-capacity project submitted by the district and a pavement rehabilitation estimated with the computer program. The HPMS would eliminate that type of double counting if the improvements are needed in the same time frame, usually 5 years (which can be varied).

A disadvantage of the HPMS system is that it does not cover all construction areas TSDHPT is interested in. These include new location projects, bridges, interchanges, and routine maintenance. These would have to be handled outside the HPMS framework.

Overall, however, HPMS does seem to have a potential for providing consistent needs estimates over time. Another significant advantage is the ability to quickly and easily make estimates of the effects of changes in funding levels, something that cannot be done with the present system. This could be very beneficial when making policy and when working with the legislature in determining the required funding for TSDHPT.

This paper documents some of the work that has been done in examining the HPMS sample and analysis package for use in Texas. Extensive work has been done in examining the adequacy of the sample size and in making recommendations for increasing the sample to make needs estimates at the district level.

ADEQUACY OF THE HPMS SAMPLE

Current Sampling Procedure

The FHWA has recommended to the states the procedures that should be followed in calculating the size of the sample needed for each state, the selection of the samples, and the criteria for selecting additional samples as needed over time. The highway system is first stratified into rural, small urban, and urbanized categories. The urbanized areas can either be handled collectively or as individual areas. Currently, there are 30 designated urbanized areas in Texas that are sampled separately. Each area is broken down by functional class (excluding the local functional class) and further stratified by volume group (up to 13) within each functional class using average annual daily traffic (AADT).

The FHWA provides a formula for calculating the required sample size for the highway sections in each volume group. The formula is given below and is from Appendix G of the HPMS Field Manual (*3*). The minimum sample size is three unless there are three or fewer sections in the volume group; in such case all sections are sampled.

$$n = F/[1 + 1/N(F - 1)] \text{ with } n \geq 3 \qquad (1)$$

where:

n = required sample size
F = $[(Z_\alpha)(c)/d]$
Z_α = value of the standard normal statistic for (α) confidence level (two-sided)
c = AADT coefficient of variation
d = desired precision rate
N = universe or population stratum size

The FHWA has recommended values of both Z_α and d, based on functional class, with generally higher precision and confidence levels for higher functional classes. The Texas Transportation Institute (TTI) Research Report 480-1 (*4*) shows that the current procedure can cause a problem in calculating the required sample size in some circumstances. The problem results from the use of AADT as both the variable for stratifying the highway sections and in calculating the required sample size within each of those stratified volume groups.

An example from a TRB paper on the same subject should illustrate the potential problem (*5*). The example looks at the minor arterial functional class from the Houston urbanized area. In volume group 1, from 0 to 2,499 AADT, there are 40 sections. Using Formula 1, n = 11.34. However, in volume group 5, from 15,000 to 19,999, with 154 sections, Formula 1 gives n = 0.29, which would use the minimum of 3. In the paper the reasons that this can happen are described and a procedure is recommended for eliminating the problem by calculating the required sample size at the functional class level using Formula 1 and then distributing the sample to the volume groups.

The problem with the FHWA procedure does not seem to be serious in most cases because there are not usually a large number of sections in higher volume groups where the problem is most significant. In addition, as shown in the simulation results, aggregating tends to reduce the error introduced in under sampling some volume groups. It should not affect estimates at the national level, but it could have some effect at the state level and substate level. For that reason, the sampling simulation in the next section uses the changes in the recommended procedure for testing and determining the required sample size for district level HPMS estimates in Texas.

A recent General Accounting Office (GAO) study (*6*) found the current HPMS sampling procedure to be adequate for making national needs estimates.

HPMS Sample Simulation

The objective and goal of taking a sample is that if chosen properly it can be used to represent the population being sampled subject to a known margin of error. In the case of HPMS, the sample is being used to represent the entire highway network. One of the main concerns is how well the sample represents the overall network in terms of the estimated needs over time. However, those needs are not known before the sample is selected and data are collected and analyzed.

As described previously, the size of HPMS sample is calculated using AADT, a commonly available data item on most highway sections. Since the sample size calculations and the margin of error is based on AADT, the accuracy of estimating needs is not known. In a sense, AADT is being used as a proxy for sampling purposes for other unknown items, such as needs. Although AADT may be a good proxy for some needs, it does not cover all possibilities over time. For example, AADT and 20-year needs tend to be correlated, but that correlation varies considerably, with higher correlations in rural areas than in urban areas. Therefore, it became necessary as part of this study to devise a method to determine the accuracy of various sampling rates for estimating needs.

There have been some studies that examine the accuracy and reliability of the HPMS sample in representing the highway network (*6,7*). There has also been considerable interest in the area of pavement management, to determine the sample coverage required to estimate pavement condition and rehabilitation costs (*8*). It was determined that a similar analysis could be performed on the HPMS sample in Texas.

A simulation model was developed to test several sampling rates and methods using the Texas HPMS sample data as the base of comparison. Only those samples on the state maintained highway system were used, because estimates of needs are required by TSDHPT for those highways. The sample was treated as the universe, like a district, and samples were taken from that universe. The sample sizes were calculated for this universe using the formula and procedure described in the previous section, and samples were selected randomly. The accuracy of the sample was calculated by comparing the needs estimate of the sample with the needs estimate of the universe. This procedure was repeated several times to generate an error distribution curve, giving the probability of any margin of error occurring, which could then be compared to the assumed accuracy, based on the AADT used to calculate the sample size. This was done for both 5- and 20-year HPMS needs estimates.

Three different groups of precision rates were tested in the simulation model. The precision rates recommended by FHWA for statewide sampling and for individual urbanized areas were used along with a lower precision rate I developed for testing. The three groups of precision rates are presented in Table 1. The precision rates specify the probability that the sample mean will fall within a specified range. For example, if a 90-5 precision is specified, it means that the sample mean AADT will be within ± 5 percent of the universe mean AADT 90 percent of the time. If a sample were drawn 100 times from the universe, the sample mean AADT would be expected to be within 5 percent of the universe mean 90 times.

A sample size was calculated for each precision rate on all the functional classes; Formula 1 was used to calculate the required sample size. These numbers are presented in Table 2. The estimated needs for all the HPMS sections were used as the basis of comparison in calculating the sample errors. For example, in rural class 1, there is a total of 176 HPMS sections. For the statewide precision rate, 142 of those sections need to be sampled. Sections were randomly selected, without replacement, until 142 of the 176 total had been selected. Selection without replacement was used so that the same section could not be in the sample more than once.

After the samples were selected, the 5- and 20-year costs were summed and then expanded to represent all the sections, using the ratio of the universe mileage to sample mileage as the expansion factor. That is the same process used in the HPMS analytical package. The samples were not stratified by volume group during this part of the simulation; stratification is tested in the next section. The error then can easily be calculated by taking the difference between the expanded sample cost and the universe cost and dividing by the universe cost. This was repeated 350 times to give a distribution of the errors. The number of replications, 350, was chosen because the distribution seemed to stabilize at about that point in testing of the simulation model and the distribution changed very little with higher replications. Some examples of the simulation results are shown in Figures 1 to 6. The complete set of graphs is contained in TTI Research Report 480-2F (*9*).

Each of the figures depicts the accuracy range along the horizontal axis, which is based on the calculated percent sample error previously described. The vertical axis gives the percentage of the 350 samples that fall within that range of accuracy. For example, in Figure 1, rural functional class 1, the top line represents the error distribution with the statewide precision rates. About 82 percent of the samples were within 5 percent of the 20-year needs estimate for all the sections in the universe and about 98 percent were within 10 percent of the universe amount. None of the samples was more than 15 percent off.

TABLE 1 PRECISION RATES USED IN SAMPLE SIMULATION

Functional Class	FHWA Statewide Precision Rates[1]	FHWA Individual Urbanized Area Rates[1]	Lower Precision Rates
Rural			
1 - Interstate	90 - 5	80 - 10	70 - 10
2 - Principal Arterial	90 - 5	80 - 10	70 - 10
3 - Minor Arterial	90 - 10	80 - 15[2]	70 - 15
4 - Major Collector	80 - 10	70 - 15	60 - 15
5 - Minor Collector	80 - 10	70 - 15	60 - 15
Small Urban and Urbanized			
1 - Interstate	90 - 5	80 - 10	70 - 10
2 - Other Freeway	90 - 5	80 - 10	70 - 10
3 - Principal Arterial	90 - 5	80 - 10	70 - 10
4 - Minor Arterial	90 - 10	80 - 15[2]	70 - 15
5 - Collector	80 - 10	70 - 15	60 - 15

[1] FHWA rates taken from Appendix F, HPMS Field Manual (3). The first number in each entry of the table is the confidence interval and the second number is the precision or range of error for that confidence level.

[2] Precision rates changed from those recommended in HPMS Manual to make them consistent with statewide rates. The recommended precision rate is 70 - 15.

TABLE 2 SUMMARY OF SAMPLE SIZES USED IN HPMS SIMULATIONS

Functional Class	Total Number of HPMS Sections	Sample Size Simulation: Statewide Precision Rate	Individual Urbanized Rate	Lower Precision Rate	Stratifying by Volume Group Simulation
Rural					
1 - Interstate	176	142	69	52	92
2 - Principal Arterial	437	355	174	132	171
3 - Minor Arterial	165	112	60	45	53
4 - Major Collector	173	139	94	76	74
5 - Minor Collector	162	111	63	48	66
All	1113	859	460	353	456
Small Urban					
1 - Interstate	42	38	27	23	31
2 - Other Freeway	25	23	18	18	20
3 - Principal Arterial	264	184	69	49	98
4 - Minor Arterial	57	49	36	30	36
5 - Collector	8	7	6	6	8
All	396	301	156	124	193
Urbanized					
1 - Interstate	148	112	48	35	60
2 - Other Freeway	167	137	69	52	86
3 - Principal Arterial	442	291	100	71	97
4 - Minor Arterial	147	97	51	37	48
5 - Collector	15	13	11	10	14
All	919	650	279	205	305
Statewide	2428	1810	895	682	954

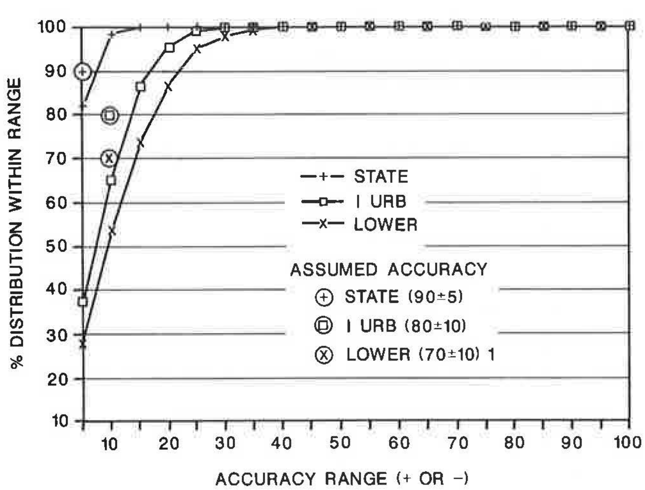


FIGURE 1 Results of HPMS simulation, testing sample size—20-year needs, rural functional class 1. ***Note*****: State, statewide precision rate; I URB, individualized rate; Lower, lower precision rate.**

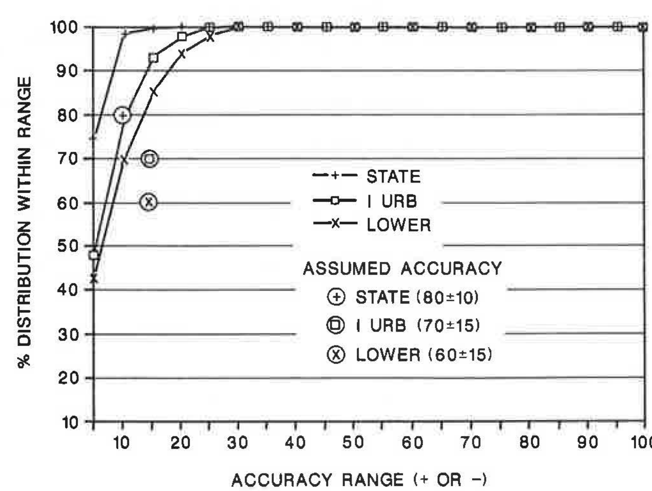


FIGURE 2 Results of HPMS simulation, testing sample size—20-year needs, rural functional class 4.

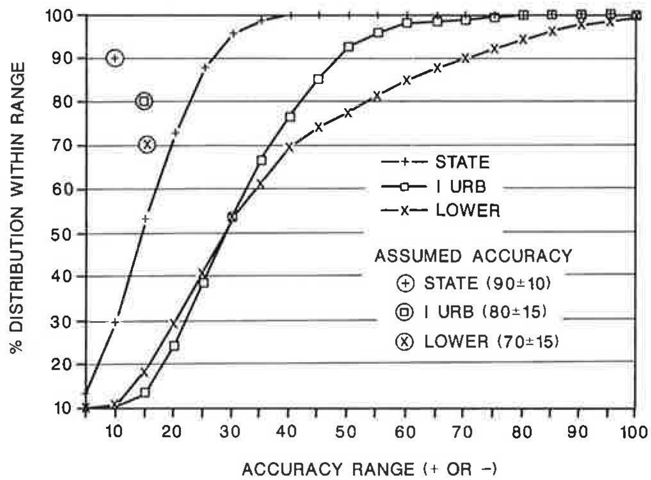


FIGURE 3 Results of HPMS simulation, testing sample size—20-year needs, urbanized functional class 4.

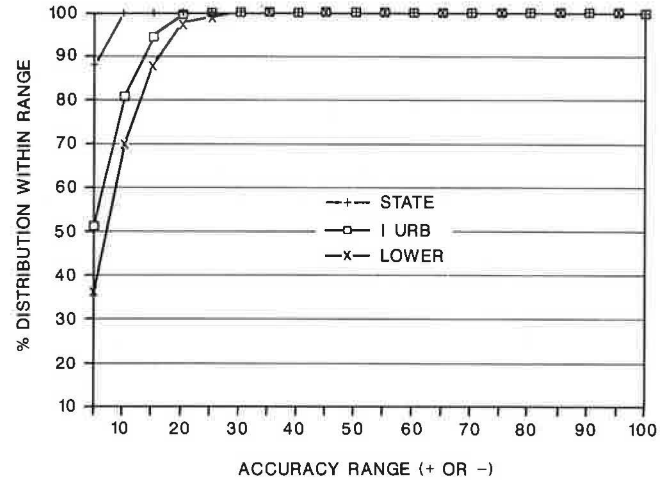


FIGURE 5 Results of HPMS simulation, testing sample size—20-year needs, all urbanized functional classes.

The assumed precision rates are also shown on each figure when applicable. The cross with a circle around it at 90-5 in Figure 1 is the statewide precision rate used to calculate the sample size. The simulated error is somewhat below the assumed precision rate, but not by much. The other precision rates miss by more. For example, the lower precision rate assumes 70 percent of the samples will be within 10 percent of the actual amount, but only 54 percent are in that range in the simulation. The rate does go over 70 percent at the 20 percent range, but doesn't reach 100 percent of the distribution until about a 50 percent range of accuracy.

Much better results are shown in Figure 2, rural functional class 4, in terms of the accuracy of the simulation versus the assumed precision. In each case, the simulated accuracy is much higher than the assumed precision. For example, the assumed statewide precision is 80-10, whereas the simulated results give about 98 percent of the distribution within 10 percent of the actual value.

At the other extreme, the results of urbanized functional class 4 are shown in Figure 3. Here the errors are much larger than was assumed in calculations of the sample sizes. For example, the statewide precision rate assumes 90 percent of the distribution will be within 10 percent of the actual amount, but only about 30 percent of the distribution is within that range. At the lowest precision rate there are some samples that have an error greater than 100 percent.

Fortunately, aggregating substantially reduces the errors for individual functional classes. The combined functional classes in the rural area are depicted in Figure 4. The assumed precision rates are not shown because they vary by functional class. The statewide precision rate gives about 96 percent of the samples within 5 percent and 100 percent within 10 percent. Even the lower precision rate has 90 percent of the samples within 10 percent and 100 percent within 20 percent.

As would be expected, the combined urbanized area, shown in Figure 5, is not as high as the combined rural area, but the

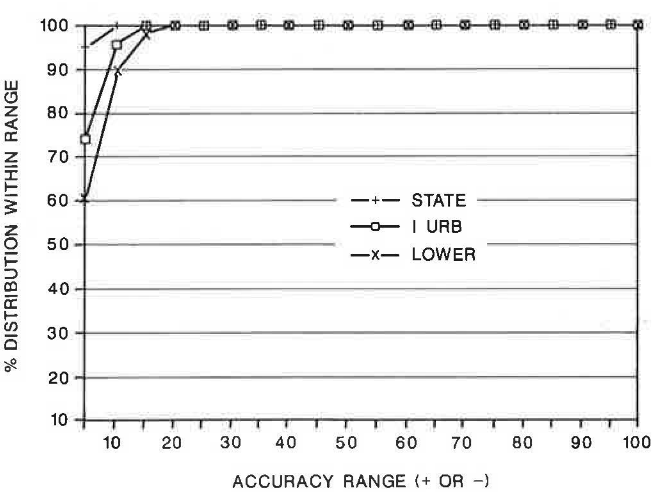


FIGURE 4 Results of HPMS simulation, testing sample size—20-year needs, all rural functional classes.

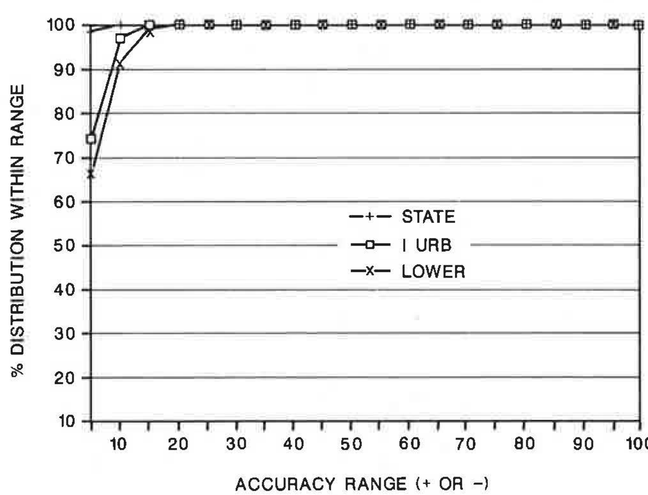


FIGURE 6 Results of HPMS simulation, testing sample size—20-year needs, all statewide functional classes.

improvements are substantial. The statewide precision rate gives about 88 percent of the samples within 10 percent and 100 percent within 10 percent. The lower precision rate gives about 88 percent within 15 percent and 100 percent within 30 percent.

The combined statewide distributions are shown in Figure 6. The statewide precision rate is very high, with 100 percent of the distribution within 10 percent. The other two are somewhat lower, but still very high, with both above 90 percent at the 10 percent accuracy level.

There appears to be considerable increases in the accuracy levels when functional classes are combined, even if the accuracy of individual functional classes is not very high.

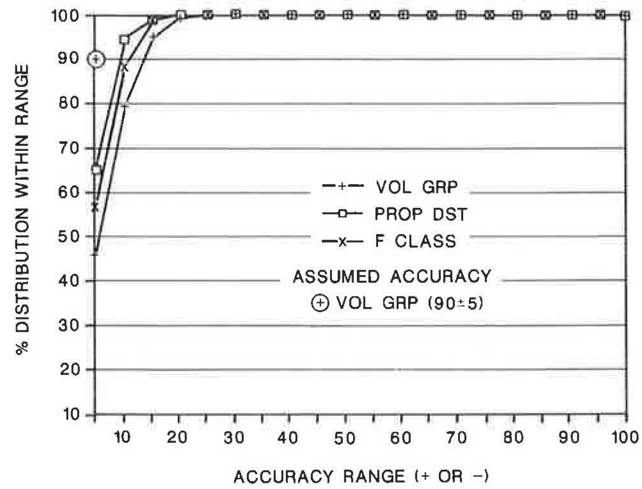


FIGURE 7 Results of HPMS simulation, testing volume group stratification—20-year needs, rural functional class 1.

Stratification by Volume Group

One common way to improve the accuracy of the sample is to stratify it into more homogenous groups. The HPMS sampling procedure attempts to do this by stratifying functional classes by volume group. The objective is to reduce the required number of samples for a given precision rate. One reason for that was to reduce the data-collection burdens on the states when the HPMS sample data was originally collected. Stratification does not necessarily improve the accuracy of the sample and it can actually make it worse, though it generally helps.

In an effort to determine the usefulness of stratifying the HPMS sample by volume group using AADT, the simulation model was used to test two stratification strategies. The first is the current technique recommended by FHWA. The HPMS sections are stratified by AADT volume group and then the required sample size for each volume group is calculated using Formula 1. In the second method the same AADT volume group stratification is used, but the sample is distributed proportionately by mileage to the volume groups. The functional class level of sampling, presented in the previous section, is also used in this stratification simulation as the basis of comparison. The functional class sample is not stratified at all, so it can be compared with the previous graphs.

A summary of the sample sizes used in the volume group simulation is given in Table 2. A complete set of sampling rates by volume group used in the simulation is contained in TTI Research Report 480-2F (*9*).

For these HPMS sections, the sample size calculated at the volume group level using the statewide precision rate gives roughly the same sample size as using the individual urbanized precision rate at the functional class level. For example, the statewide total in Table 2, the sample for individual urbanized rate is 895, compared to 954 for the volume group stratification. Of course, the sample sizes for some functional classes vary considerably, influencing the accuracy of individual functional classes.

Some examples of the simulation results of stratification are shown in Figures 7 to 11. Rural functional class 1 is presented in Figure 7. All three lines are below the assumed precision of 90-5, with the proportional distribution performing the best. Surprisingly, the unstratified functional class distribution is higher than the volume group calculated distribution. In this case, the calculation of sample size by volume group actually was worse than if no stratification at all had been done.

A somewhat different result for rural, functional Class 3 is shown in Figure 8. In this class the highest accuracy is the volume group calculation with the stratified proportional distribution higher at most levels of accuracy than the unstratified distribution.

As was the case with the simulation results of the previous section, the errors tend to be much larger in the urban areas (Figure 9), but decline substantially as functional classes are combined (Figures 10 and 11).

In general, stratification improved the overall accuracy of a given sample. This can be seen in the combined statewide distributions in Figure 11. Both stratification strategies give a substantial improvement at the 5 percent and a lesser improvement at the 10 percent level. In most cases, the proportional stratification improved the sample accuracy as compared to the volume group calculated accuracy, although that

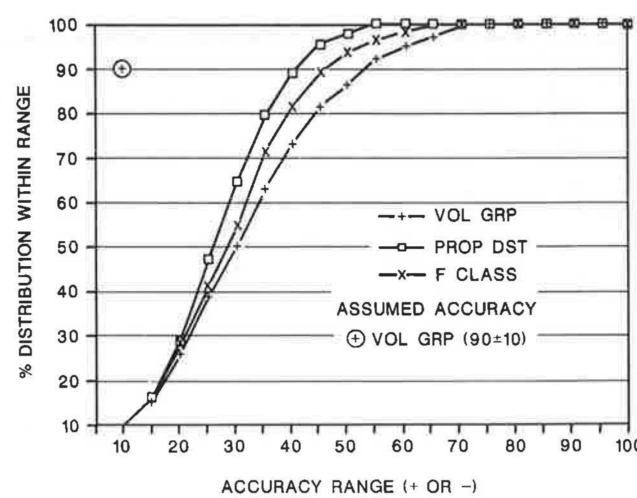


FIGURE 8 Results of HPMS simulation, testing volume group stratification—20-year needs, urbanized functional class 4.

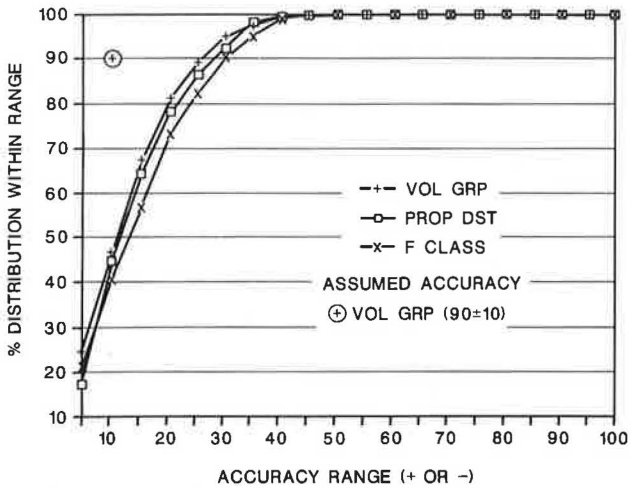


FIGURE 9 Results of HPMS simulation, testing volume group stratification—20-year needs, rural functional class 3.

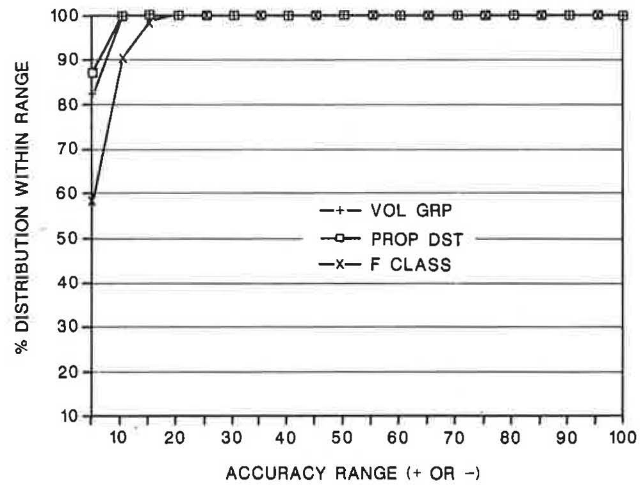


FIGURE 11 Results of HPMS simulation, testing volume group stratification—20-year needs, all statewide functional classes.

is not always the case, and many times the difference is very small. However, the benefits of stratification that FHWA anticipated when stratifying the HPMS sample by volume group are very small. The same sort of accuracy can be obtained by a simple proportional stratification.

HPMS Sample Size Recommendation for Texas

As a result of the simulation results, and taking into account the requirement to estimate needs at the district level, the TSDHPT advisory committee for this project decided to substantially increase the HPMS sample size in Texas. It was decided that estimates were not required at the functional class level within districts, so the errors estimated with the simulation model at the area level were acceptable. The recommended sample represents about a 133 percent increase in the on-system HPMS sample in Texas. The recommended

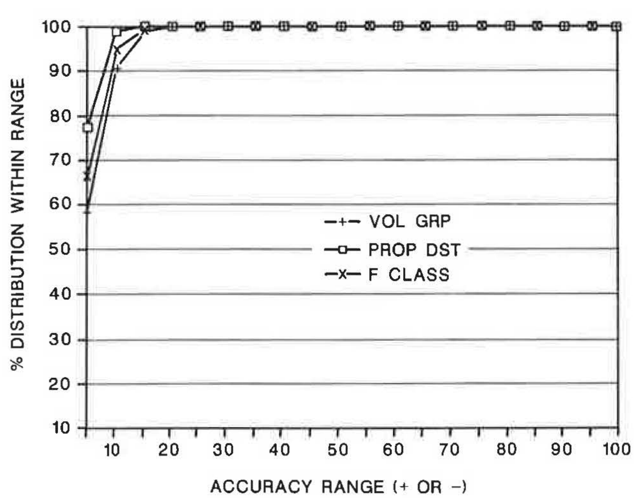


FIGURE 10 Results of HPMS simulation, testing volume group stratification—20-year needs, all urbanized functional classes.

sample sizes were calculated at the functional class level by district, with proportional distribution of samples within volume groups. The individual urbanized precision rates were used for the urban area and the lower precision rates, presented in Table 1, were used for the rural and small urban areas. The lower precision rates were used for the rural area because the simulation results indicated satisfactory accuracy levels with that precision rate. The small urban areas constitute a very small proportion of the estimated statewide needs, so the lower precision rates were also used for that category. There was also a concern to keep the increase in sample sizes as low as possible because it would entail a significant data-collection burden for the districts, and the small increase in the accuracy of the estimates would not justify the increased data-collection costs.

CONCLUSION

The HPMS sample data and analytical package offer an opportunity for Texas to make estimates of future needs in a consistent and comprehensive fashion that is not available at the present time. In addition, it provides a tool for estimating the effects of different funding levels on the condition of the highway system and the motorists using the highways. This method should be very valuable in the future.

One of the biggest areas of concern with the HPMS is the sample. Anytime a sample is used to represent a larger population, in this case the highway network, there is a legitimate concern that the sample may not accurately represent the population for estimating those unknown elements from the population. In the case of HPMS, the sample is based on AADT, a commonly used and widely available data item for highways. However, one of the principal items of interest is not the input AADT, by itself, but how it affects, along with the other data items, the estimated needs in the output. For that reason a simulation model was developed to determine how good a sample, based on AADT accuracy, is in estimating needs.

The results of the simulation showed that in general the needs accuracy is not as high as assumed when the sample size for individual functional classes is calculated. But when aggregating over functional classes, the accuracy significantly improves, which suggests that for highly aggregated needs estimates the sample is probably not introducing much error. In other words, the sample is accurately representing the overall highway network. However, more caution should be exercised when making estimates at lower levels of aggregation.

The recommended increases in the HPMS sample for use in Texas represent the results from the simulation model, as well as the need to stratify the sample to the district level. The increased sample will provide adequate coverage for district level needs estimates, and because it is far above the minimum required sample from FHWA, should pose no problems for reporting purposes to the FHWA.

ACKNOWLEDGMENT

The research findings presented in this paper were developed as part of a Research Study sponsored by the Texas Department of Highways and Public Transportation and FHWA.

REFERENCES

1. *Highway Performance Monitoring System Analytical Process. Volume 2, Version 2.0: Technical Manual.* Office of Highway Planning, FHWA, U.S. Department of Transportation, Jan. 1986.
2. *Strategic Mobility Plan.* Texas State Department of Highways and Public Transportation, Austin, Aug. 1984.
3. *Highway Performance Monitoring System Field Manual.* Office of Highway Planning, FHWA, U.S. Department of Transportation, Jan. 1984.
4. J. L. Memmott. *Sample Size and Accuracy of Highway Performance Monitoring System.* Research Report 480-1, Texas Transportation Institute, The Texas A&M University System, College Station, Tex., May 1986.
5. J. L. Memmott. Adequacy of the Sample Size and Accuracy of the Highway Performance Monitoring System for Use in Texas. In *Transportation Research Record 1134*, TRB, National Research Council, Washington, D.C., 1987, pp. 32–39.
6. *Highway Needs, An Evaluation of DOT's Process for Assessing the Nation's Highway Needs.* U.S. General Accounting Office, Washington, D.C., Aug. 1987.
7. D. T. Hartgen. How Good is the Highway Performance Monitoring System? A Comparison with State Results. In *Transportation Research Record 1060*, TRB, National Research Council, Washington, D.C., 1986.
8. C. J. Templeton and R. L. Lytton. *Estimating Pavement Condition and Rehabilitation Costs Using Statistical Sampling Techniques.* Research Report 239-5, Texas Transportation Institute, The Texas A&M University System, College Station, Tex., June 1984.
9. J. L. Memmott. *Results of the Evaluation of the Highway Performance Monitoring System for Use in Texas.* Research Report 480-2F, Texas Transportation Institute, The Texas A&M University System, College Station, Tex., June 1988.

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

Development of State Traffic Monitoring Standards: Implementing *Traffic Monitoring Guide* in New Mexico

DAVID ALBRIGHT, CHRIS BLEWETT, AND RACHEL GOLDSTINE

A study of traffic monitoring practices was undertaken by the New Mexico State Highway and Transportation Department. The study reviewed current practice in relation to the *Traffic Monitoring Guide*, published by the FHWA. The study concluded that state traffic data requirements could best be met, and the *Traffic Monitoring Guide* could best be implemented, by developing statewide traffic monitoring standards. Standards were drafted, refined, and adopted. All traffic data are required to be in conformity with the state standards. The immediate benefit is standardization of traffic data. Data are equivalent and directly comparable, enhancing project selection and design. This process establishes the basis for testing and evaluating traffic monitoring statistics and procedures.

A study of traffic monitoring practices was undertaken by the New Mexico State Highway and Transportation Department. The study reviewed data collection, summarization, and analysis practices. Practices were reviewed within the department, among metropolitan planning agencies and other governmental units in New Mexico, and departments of transportation nationally. The study was conducted in response to concerns for state data integrity, and in relation to implementation of procedures recommended in the *Traffic Monitoring Guide*, published by the FHWA in 1985 (*1*).

METHODOLOGY

A study design was drafted, reviewed, and adopted. The study methodology included review of literature, interviews concerning current traffic monitoring practices in New Mexico, interviews concerning current traffic monitoring practices nationally, drafting state traffic monitoring standards, and standards implementation and evaluation.

The study began with a comprehensive review of traffic monitoring literature. Through this review, the *Traffic Monitoring Guide* was identified as providing the basis for improving the accuracy and efficiency of traffic monitoring.

After the review of literature, interviews were conducted with all departmental personnel involved in traffic monitoring. A series of three separate interviews was conducted with each individual. The interviews addressed operational procedures, statistical methods, and current and potential computerization of work. Current practices among the three state metropolitan planning organizations were also reviewed.

On the completion of the in-state interviews, all state departments of transportation were surveyed. Two separate telephone surveys were conducted. One survey concerned statistical methods used in traffic monitoring. The second identified computerization of traffic monitoring among the state departments.

On the basis of the literature review, interviews, and surveys, a process was proposed to develop state traffic monitoring standards. Standards were drafted through a consultative process and were refined through state and federal review. After formal adoption, training was provided to implement the state standards.

The final element of the study design was to maintain the state standards as an open-ended process. An annual review and refinement of the standards was established, as well as verification of standard compliance.

NEED FOR STATE TRAFFIC MONITORING STANDARDS

The traffic monitoring study determined that the current traffic monitoring practices in New Mexico produced questionable traffic monitoring data. Most practices resulted in data for which a confidence level and interval could be calculated neither on a system- nor site-specific level.

A primary source of error in traffic monitoring was inconsistent use of professional judgment in factoring traffic data. Professional judgment was applied to adjust base data and complete missing data. The judgment of professionals resulted in multiple modifications of the same data set. Traffic counts were modified up to eight separate times by individuals unaware of previous or subsequent modifications. These judgments were made during field collection, data summarization, selection and application of adjustment factors, and data use. The adjustments modified the initial data in conflicting ways. In most instances existing practice prevented identification of original base data after modifications had been made.

The study also identified inconsistency in data-collection procedures. Among agencies in New Mexico, and within some

D. Albright and R. Goldstine, New Mexico State Highway and Transportation Department, P.O. Box 1149, Santa Fe, N. Mex. 87504-1149. C. Blewett, Middle Rio Grande Council of Government, 317 Commercial N.E., Suite 300, Albuquerque, N. Mex. 87102.

agencies, the length of count and use of adjustment factors varied substantially. All traffic volume data were typically labeled average daily traffic (ADT). This resulted, for example, in a traffic-flow map with average weekday traffic on some roads and average 7-day week traffic on other roads all being labeled as ADT. Some same-map traffic volumes were mixed data, including average annual daily traffic (AADT) from permanent traffic recording devices, 24-hour unadjusted count data, 48-hour count data with axle correction and seasonal adjustment factor, and non-count-based estimates. All volumes were labeled ADT and presented as equivalent data. Of more concern, all data were analyzed as equivalent data in highway project simulation, selection, and design.

The error in professional judgment must be assessed in relation to data use. One use of traffic volume data is for assessment of alternative routes for urban areas. Nonequivalent data, as described above, were used in a gravity-model computer simulation of Taos, New Mexico. From these data a bypass was indicated. Equivalent traffic data as potentially (and eventually) required under state standards were then collected, and one-way street changes were indicated rather than a bypass.

Review of the department's highway project selection process indicated the dominant role of traffic volume. Potential projects with similar roadway distress were being selected on the basis of ADT. Comparison of highway projects in the project selection process will be enhanced, and some selected projects will change, as traffic data are equivalent.

Similar changes have been calculated in relation to pavement design. Pavement thickness was determined by the department as sensitive to differences in traffic-volume data summarization. Inaccurate current-year data typically establish future trends that, when forecast to design year, result in under-building or over-building a facility (*2*). Improving current-year traffic volume data as required under the state standards typically resulted in a design change of one-half in. in pavement thickness.

Improvement in vehicle classification data will have additional design impact. Rural Interstate data from then current vehicle classification practices were compared with data collected as potentially specified under state standards. Under then current practice there was a characteristic underestimate of heavy commercial vehicles by 6 percent in the total traffic volume data.

In network simulation, highway planning, project selection, and pavement design, traffic data practices were inadequate to data use. The problem with individual practices is accentuated when comparing data from various locations and from various sources. Data consistency is an issue nationally. No state surveyed was found to have adopted statewide standards for traffic summary statistics. No models were available from other states for such standards. Practices then common in New Mexico were typical of other states. This would suggest that national traffic data bases contain nonequivalent data. Future efforts at national traffic data bases should examine base data through specified collection procedures, summarization techniques, adjustment factors, and factor sources applied to the base data.

At every governmental level in New Mexico, the study determined that traffic monitoring summary statistics were inadequately reported. Proper understanding and use of these data would require definition of the summary statistic and method. Except for cursory descriptions of count factors and methods, information was unavailable that would allow summary statistics to be appropriately used. There was a critical need for adequate reporting procedures in transmittal of traffic monitoring summaries.

As a result of the study, it was concluded that the traffic monitoring practice within the department must be redesigned. It was also concluded that the improvement should be statewide and required for all public and private traffic monitoring practices related to state and federally funded road projects. This would provide the opportunity for the department to compare data, to avoid mixed data, to avoid politically inspired data (traffic data collected in a manner to show a desired conclusion) in the state data base, and to use these data appropriately in transportation planning and engineering.

PROCESS OF DEFINING AND IMPLEMENTING THE STATE STANDARDS

Nine traffic monitoring technicians in New Mexico were invited to participate in drafting state traffic monitoring standards. The department, each metropolitan planning organization, and other interested agencies were represented. Copies of traffic monitoring technical documents were distributed to the participants. Preliminary discussions of traffic monitoring issues were conducted. All participants were then gathered for an intensive 2-day period. The traffic monitoring standards were drafted through this consultation.

The procedure used during the consultation was based on specific types of traffic data. Each type of data (or data element) was addressed, from collection through summarization and analysis. After data elements were defined, discussion proceeded on what practices should be followed. Although data elements were considered individually, the principle of nesting was adhered to, as defined in the *Traffic Monitoring Guide*. The intention was to help ensure the most efficient use of collection activities. After the consultants had reached consensus on separate data element standards, an overall review was conducted. This resulted in simplification of the proposed standards, and a final draft was adopted by the participants.

The draft standards were reviewed and refined at the state and federal level. Final standards were prepared and adopted under departmental administrative memorandum. The memorandum was signed in May, 1988, and the standards went into effect October 1, 1988. This allowed a period of standards training before implementation. The standards are required for all traffic monitoring data on any road in New Mexico for which state or federal moneys are used or are proposed to be used.

OVERVIEW OF THE STATE TRAFFIC MONITORING STANDARDS

The New Mexico state traffic monitoring standards contain 89 separate specifications for traffic data collection, summarization, and analysis. The structure of the state standards follows that of the *Traffic Monitoring Guide* (*1*). The standards begin with characteristics required of all traffic monitoring in New Mexico and then detail practices required by type of monitoring: automatic traffic recorders, coverage counts, vehicle classification counts, weigh-in-motion (WIM) counts, manual counts, and other concerns such as monitoring equipment.

In the early stages of the New Mexico study it was decided that any modifications to current practice should implement the *Traffic Monitoring Guide*. In implementing the *Guide*, standards had to be derived to bridge the gap between specific monitoring practice required to meet state data needs and the general recommendations advanced by the FHWA. It was also necessary to produce standards that identified preferred methods where the *Guide* provided more than one alternative.

The state standards begin with general requirements for all traffic monitoring practices. Of these standards, the critical points adopted were restrictions on data manipulation, statement of confidence level and interval, unique road segment basis for traffic volumes, and establishing annual review and updating of standards. Each point is described below.

The state standards stipulate that missing or inaccurate data may not be completed, filled-in, or replaced for any type of traffic count, at any location, under any circumstance. This preserves the integrity of base data. Partial or incomplete data will be separately stored by type of data, and will be analyzed to quantify the errors that are associated with current imputation procedures.

The state standards require that all transmitted traffic counts and volume estimates must include a confidence level and interval. The specified system level confidence is 95 percent and the specified confidence interval is ± 10 percent. These statistics are calculated, as in the *Traffic Monitoring Guide*, from system level variability of data.

The standards identify that traffic summary statistics must be reported on the basis of "unique road segments." If two values of the same traffic volume summary statistic within a road segment have a volume difference that exceeds their combined confidence interval, the road segment is not unique. It must be divided into separate, unique segments and separate traffic summary statistics must be provided. This division prevents averaging highly variable traffic volumes, for example, as representative of a road segment, which results in the arithmetic mean being an inadequate measure of the central tendency of traffic on the roadway.

In New Mexico there are roads that primarily serve extractive industries. On these roads not only the volume difference but the load difference is important in discerning what is a unique road segment. For this reason, the standards identify that if two equivalent single-axle loading (ESAL) summary statistics within a road segment have an equivalent loading difference that exceeds their combined confidence interval, the road segment is not unique. In this instance it must be divided into separate, unique segments. This is an important element of the standards. It identifies, for example, that if separate segments are indicated, separate design solutions may be indicated.

Standards revision was made part of the general traffic monitoring standards. There will be an annual review of standards and their implementation. The same consultative process will be used as served the drafting of the standards. A standards compliance and revision mechanism, based on a government audit, was designed and distributed to all agencies involved in traffic data collection. The audit includes a record of implementation for each state standard and an equipment maintenance record. The standard audit will provide a check for data integrity. The equipment maintenance record will enable estimation of additional data collection attributable to anticipated equipment failure.

After the general standards are stated, standards related specifically to continuous automatic traffic recorders (ATRs) are discussed. The basic function of ATRs is to provide data that can be analyzed and grouped on common patterns for the development of adjustment factors. The adjustment factors may then be applied to coverage counts. Previous analysis of ATRs monthly, or seasonal, patterns in New Mexico identified that functional classification and seasonal variation are highly correlated. For this reason the annual and monthly adjustment factors are based on summary statistics from ATRs on the same roadway functional classification. In the same study of traffic statistical variability by functional classification, it was determined that individual ATRs effectively control volume factors on the same roadway for relatively short geographic distances. Beyond a 2-mile distance on the same roadway, except for lengthy, rural unique road segments, the mean statistic from ATRs on the same functional classification provides a more adequate count adjustment factor. This led to state standard restrictions on the maximum distance, on the same route, for which data from an individual ATR would be applied.

The state standards require an adequate sample of ATRs by functional classification. As importantly, it is required the sites be randomly selected within each functional classification. Over the years New Mexico has installed a variety of ATRs. In most instances this has not been a random process, but rather what is referred to as a "pseudorandomization" process: where there were road construction projects and an ATR was considered important, one was installed.

This practice has potentially introduced unknown bias into the mean statistics by functional classification. Under the state standards new permanent counters are required. The old counter data will also be collected. After the first year of enforcement of the standards, there will be an opportunity to measure the impact of pseudo-randomization of counter location by functional classification and volume grouping of roadways.

There is another opportunity to measure the impact of traffic monitoring practices other than the practices specified in the state standards. Under the general standards it was noted that incomplete data cannot be imputed. Incomplete permanent counter data cannot be used in computing average day of the week, month, or year in calculating AADT. The standards specify that monthly traffic summary data must be based on a representative sample of the days within the month, which must include a minimum of 2 days for each day of the week. Complete, standard data are stored in the primary file for calculating summary statistics.

The data from an ATR not meeting state standards, including complete data, will be excluded from calculating the mean summary statistics of their functional classification. These data will be separately stored in a research file. At the end of the first year of the state standards, and each subsequent year, it will be possible to evaluate the statistical significance of including data that had been excluded. If merited, the standards may then be appropriately modified.

The principle adopted in the New Mexico state standards is clear: questionable data are excluded from traffic monitoring statistics until it is shown that their inclusion adds to rather than diminishes the quality of the analysis. All data are retained and stored for use in the primary file for computing traffic monitoring statistics or for use in the research file for data analysis and possible standards modification.

In keeping with this principle, other quality controls are identified within the state standards. Among these are

- When the same recorded traffic volumes, other than zero, occur at an ATR for four successive hours, an error message will be displayed and the day's data will be excluded from calculation. They will be separately stored in th research file for analysis.
- When 8 hours of successive zeros occur at an ATR, an error message will be displayed and the day's data will be excluded from calculation. They will be stored in the research file.
- If the daily directional total of an ATR is in the range of 60 to 79 percent of the total traffic for the same day, a data and ATR review message will be displayed. The data will be excluded from initial calculation and stored in the primary file only after operator review and acceptance.
- To provide the operator an indication of ATR correct monitoring, when the daily traffic volume for a given day of the week exceeds ± 2 standard deviations from the annual average day of the week for the same day, a warning message will be issued.

The intention of these standards is to ensure that suspect data are not automatically loaded into the ATR primary data base. Further, excluded data will be reviewed through the research activities of the department to help determine whether or not exclusion is appropriate.

The third section of the New Mexico state standards relates to volume coverage counts. Forty-eight consecutive hour counts, with direct computer interface fo the traffic monitoring data base, are specified. An adequate sample to attain the specified confidence requirements is calculated for the following functional classifications:

- Rural
  - Interstate
  - Principal arterial
  - Minor arterial and major collector
  - Recreational route

- Urban
  - Interstate
  - Principal arterial
  - Minor arterial
  - Major and minor collectors

Following the principle of excluding suspect data, and the standard prohibiting data imputation, if there are less than 48 hours of data, the count will not be included in the coverage sample. Incomplete data sets will be separately saved. At the end of the first year, and each subsequent year, it will then be possible to measure the impact of including various hours of imputed data, by alternative imputation techniques.

Vehicle classification counts are a subset of the volume coverage count program. The coverage count sites are randomly selected from all unique road segments. The number of counts is based on the variability of data by functional classification. Vehicle classification count sites are randomly selected from the coverage counts. This makes the count program more efficient: the same effort used for collecting classifications also collects vehicle volume for these sites. Two required counts are taken with one operation.

The classification counts are used in estimating the loop correction factor. The mean loop correction factor by functional classification is used. The loop correction factor from the classification subset will be compared with the factors from the ATRs on the same road classification. This will allow testing the assumption that the coefficient of variation for coverage count loop correction factor (based on shorter term, more extensive number of counts) will be lower than that of the ATR loop correction factor (based on longer term, fewer count locations).

In addition to nesting classification counts within the coverage count program, the New Mexico state standards nest the speed count program. Speed compliance monitoring will be based on 48-hour intervals, in both directions. Both speed and volume will be recorded so that volume can be used to either satisfy a required sample in the coverage count program, or to provide more current traffic count data for a segment in the data base. Noncoverage count site data provide a basis for evaluating noncount year functional classification growth factors.

The third nested count program is WIM. The WIM program is a subset of the classification count. As specified in the *Traffic Monitoring Guide*, a minimum of 90 sites are sampled for 48-consecutive hours, over a 3-year cycle. Thirty of the sample sites will be selected from the Interstate vehicle classification sample. Volume and vehicle classification data from the WIM program will be used to satisfy part of the vehicle classification and coverage count requirement.

After the count programs have been completed, traffic-flow maps are a primary source of data distribution in New Mexico. The New Mexico state standards clarify the way in which data are to be published. Traffic-flow maps will not use smoothing techniques. All volumes represented must, beginning with the second year of standards implementation, be based on count data.

During the phase-in of the standards, there will be some data that are not in compliance. The first year after adoption of the state standards, all traffic-flow maps may designate volumes in three ways:

1. Traffic volume with general confidence level and interval, which denotes the volume method, is in compliance with state standards.
2. Traffic volume in parentheses, which denotes the volume is based on count data for the segment, is not in compliance with state standards.
3. Traffic volume in brackets, which denotes that the volume estimate is not based on count data for the road segment, is not in compliance with state standards.

In the second year of implementation of the state standards, no estimated volume data may appear on the traffic-flow maps. In the third year of implementing the state standards, and all subsequent years, the only volumes designated may be those in compliance with the state standards.

The New Mexico State traffic monitoring standards include requirements for manual counts, other specific project counts, and traffic monitoring equipment. As a whole, the standards create the condition for equivalent data use statewide, for both the governmental and private sectors. The standards create the condition for understanding traffic data and their appropriate use.

BENEFITS FROM IMPLEMENTING STATE STANDARDS

There are three primary benefits from adoption of the state traffic monitoring standards. The first benefit, which has accrued through the development of state standards, is traffic data efficiency. Whether the data are collected by the state, city, or a consultant employed by a governmental agency, the data are standard and may be electronically collected and transmitted to the state traffic data base. This saves staff time and error in manual transcription of data, and provides a single traffic data resource for New Mexico.

The second benefit is systematic traffic data summarization and analysis. No matter how efficiently the data were transmitted, if mixed data were gathered, the traffic data base would be of marginal planning and engineering support. The systematic data collection and summarization practices provide equivalent traffic data. It is compelling to compare some current practices (which result in typical traffic data with no known accuracy, or accurate 90 percent of the time ± 100 percent), with standard-based practices. The present study and standards provide appropriate traffic data for highway planning and engineering activities.

The third benefit is ongoing traffic data research and development. Over the longer term, benefit is anticipated from data comparisons. With the standards implemented, on an annual basis the impact of random and pseudorandom ATRs can be calculated. The impact of excluding partial data can be evaluated. The impact of alternative imputation techniques, with various hours and days of data estimated, will be compared. The change in project selection, pavement design, and urban traffic network simulation may be evaluated with each proposed change in the state standards.

Each of these changes and each of these benefits are provided through statewide traffic monitoring standards. The potential exists for refinement of traffic monitoring practice through the development of consistent, comprehensive statewide traffic monitoring standards.

REFERENCES

1. *Traffic Monitoring Guide*, U.S. Department of Transportation, FHWA, Washington, D.C., 1985.
2. D. Albright. A Quick Cluster Control Method: Permanent Control Station Cluster Analysis in Average Daily Traffic Calculations. In *Transportation Research Record 1134*, TRB, National Research Council, Washington, D.C., 1987, p. 57.

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

Geographic Information Systems: An Important Technology for Transportation Planning and Operations

HOWARD J. SIMKOWITZ

A Geographic Information System (GIS) is a computerized data base management system for the capture, storage, retrieval, analysis, and display of spatial (i.e., locationally defined) data. The purpose of this paper is to explain why GIS technology is important to transportation professionals, describe how a number of transporation agencies are using GIS, and provide insight on how to participate in this technology. Transportation agencies are still in their infancy with respect to exploiting the power and possibilities offered by GIS technology. The usefulness of spatially integrated data to transportation is examined and the distinction is made between GIS and other data base systems that use spatial data. The benefits of GIS are summarized, and examples of GIS activities at the FHWA and state highway agencies are described. Sources for digital geocoded data, including U.S. Geological Survey digital line graphs and Bureau of the Census topologically integrated geographic encoding and referencing files, are discussed.

A Geographic Information System (GIS) is a computerized data base management system for the capture, storage, retrieval, analysis, and display of spatial (i.e., locationally defined) data. The purpose of this paper is to explain why GIS technology is important to the transportation professional, describe how a number of transporation agencies are using GIS, and provide insight on how to participate in this technology.

THE ROLE OF SPATIALLY INTEGRATED DATA IN TRANSPORTATION

Spatial considerations are fundamental to most transportation activities. A transportation system consists of nodes, links, and entities distributed in two- or three-dimensional space. Events happen within this system at a point (an accident or a signal location), along a segment (vehicle volumes or pavement deficiencies), or within a geographical area (the number of people living within two blocks of a bus stop or working in an industrial park).

The collection of highway-related data involves a wide variety of activities: traffic counting, sign inventories, skid-resistance measurements, photologging, accident investigation, recording of construction and maintenance projects and funding, right-of-way surveys, inventories of signs and roadside obstacles, bridge inspection, rail-highway crossing inventories, speed monitoring, pavement condition surveys, geometric design inventories, and other data-collection and maintenance activities. Unfortunately, these various data files within the same transportation agency are typically unrelated to each other, duplicative, and inconsistent.

An integrated highway information system, on the other hand, is a system for collection and storage of highway-related data in such a way that data from different sources that apply to the same point on, or section of, a highway are correlated or linked (*1*).

Because transportation-related data always have a spatial component, the most natural way to associate elements from different data sets is through a consistent spatial referencing system. GIS technology can provide the core of a framework for an integrated highway information system (*2*). As far back as the Roman Empire, highway referencing systems have been used to locate roadway features. Popular systems include milepost, reference post, and paper document methods (*3*). More recently, states have begun incorporating georeferencing systems such as state plane or latitude and longitude coordinates into their data bases.

The following examples illustrate the importance of using a consistent georeferencing system:

- In addition to the data on the drivers involved in an accident, accident analysis requires the correlation of a number of explanatory roadway and environmental variables, such as pavement condition, roadway geometrics, weather conditions, traffic volumes, signage, signalization, and lighting. It is important to observe these variables at both high and low accident locations to determine causality. An easy way to associate these variables is through a common geographical referencing system.
- Transportation-demand modeling makes use of population and employment statistics that have been collected for small, homogeneous areas such as census tracts or traffic analysis zones (groups of census blocks). Trip productions and attractions are developed for these analysis zones, and vehicles or person-trips are assigned to each link in the transportation network. The results of the analysis might be summarized by link, transportation corridor, analysis zone, city, county, or region.
- A pavement management system records the construc-

Caliper Corporation, 4819 Cumberland Avenue, Chevy Chase, Md. 20815.

tion and maintenance history as well as the condition of roadway segments and assists decision makers in selecting appropriate treatments for deficiencies. Typically, the roadway condition and treatments are printed in tabular form, and it is up to the highway engineer to transfer this information to a base map by hand as a first step in developing actual projects. A map-based interface and geocoded data would greatly facilitate the process of data entry and project development and scheduling. It could also be used to produce any number of statistical reports (such as the number of deficient miles in each jurisdiction) and graphical outputs indicating the condition of the network and the scheduling of projects. These products could be easily comprehended by the engineers, politicians, and citizen groups. Projects could be scheduled to have minimal impact on flow through the network if demographics and journey-to-work data were included in the system. User costs resulting from delays could be determined.

One would assume that agency-wide spatially related data bases would have become commonplace. Unfortunately, this is not the case. Although all state Departments of Transportation (DOTs) use some form of referencing system, the system may not be consistent throughout the organization. Files for different data groups may have been created independently of one another, using different referencing systems or computer formats. In the worst case, some of the data required for analysis may not be spatially referenced at all. Relating the files to create an agency-wide data base has proven to be difficult and costly. Yet, in order to perform meaningful analyses of a transportation system, it is imperative that these different data elements be linked by a common reference framework.

CHARACTERISTICS OF A GIS

Although the integrated data base systems being implemented by the state DOTs represent major improvements in data base design, in most cases the data are not referenced to a geographical coordinate system (e.g., latitude and longitude or state plane coordinates). Rather, integrated data base systems are tied to the traditional milepoint or reference point systems. Therefore, these systems do not provide the data necessary for most spatial analyses.

By contrast, the sophisticated data base engine in a GIS has the ability to associate and manipulate diverse sets of spatially referenced data that have been geocoded to a common referencing system. To permit this it might be necessary, for example, to use software that transforms state plane coordinates and milepoint data to latitude and longitude. A GIS is capable of topological operations; that is, it understands how elements contained in the data base are related to each other spatially and it can perform spatial manipulations on these elements.

A GIS contains two broad classifications of information, geocoded spatial data and attribute data. Geocoded spatial data define objects that have an orientation and relationship in two- or three-dimensional space. Each object is classified as either a point, a line, or a polygon and is tied to a geographic coordinate system. These objects have precise definitions and are clearly related to each other according to the rules of mathematical topology.

In addition to the topological information, a GIS contains the same attribute data that is found in traditional data bases. Attributes associated with a street segment might include its width, number of lanes, construction history, pavement condition, and traffic volumes. An accident record could contain fields for vehicle types, weather conditions, contributing circumstances, and injuries. What distinguishes a GIS from a traditional data base is that this attribute data is associated with a topologic object (point, line, or polygon) that has a position somewhere on the surface of the earth.

BENEFITS OF A GIS

A GIS can lead to new ways of thinking about and dealing with old problems. Because the data is tied to a common referencing system, it is easy to use the same data across applications as well as to associate diverse data sets previously unavailable for joint analysis. Topology permits new questions to be asked and encourages a new style of analysis that is in many cases fundamentally better than those used traditionally.

Because it permits the use of spatial relationships, a GIS adds a degree of intelligence and sophistication to a transportation data base that has previously been unknown. For a segment (a line) on a road network, a GIS system knows what routes (other lines) cross it and whether there is an actual physical intersection. It knows the position of roadside features (points) along the segment; and can tell which census blocks (polygons) are to the right and the left of the segment or within any distance of it.

Rather than being limited to textual queries, it is now possible to perform geographic queries in a straightforward, intuitive fashion. For example, a GIS with the appropriate algorithm and data can easily compute and display the route that will result in the minimum population exposure to a shipment of hazardous materials. With the route drawn on the computer screen, the analyst can see immediately how the logic of the model has bypassed certain population centers. The analyst can create a detour by pointing to a road segment and "deleting" it from the network and then watch the algorithm redraw the path. Similarly, the analyst can ask a series of geobased questions and obtain the answers quickly in an easy-to-understand color-coded display on the screen, hard copy, or disk file. Two examples are the number of low-income residents living within two blocks of a bus stop and the total employment within each traffic analysis zone.

COMPUTER-AIDED DESIGN AND DRAFTING, MAPPING, AND GIS

There are a number of very useful tools available to the transportation professional that, at first blush, might be mistaken for a GIS. However, there are fundamental differences that should be understood. For example, a transportation demand model that has for an interface a graphical representation of a network is not a GIS. Rather than being built on a flexible data base engine, these models use a rigid file structure suitable only for the task at hand. The network, although displaying a graphical abstraction of the nodes and the links, is lacking in topology (spatial relationships).

Computer-aided design and drafting (CADD) systems have revolutionized the cartographic process within most state DOTs. They are used to produce map products nearly as elegant and sophisticated as those drawn by the most experienced cartographer. Line widths, shadings, colors, and symbology can be changed electronically in an instant, and productivity has been greatly increased [for a complete discussion of computer assisted cartography, see (*4*)]. However, at nearly every state transportation agency, this CADD mapping activity is not producing a GIS.

Dueker (*5*) makes clear the distinction between GIS and computer-aided mapping (CAM). He points out that the two terms are often used interchangeably, but they should not be. A CADD mapping system lacks at least one essential ingredient, topology. Each theme (e.g., interstate highways, state routes, or county boundaries) is drawn on its own electronic layer. This is physically equivalent to creating a set of transparent mylar overlays. The observer can see where an interstate and a state route cross, and might even be able to ascertain that there is an interchange at this point, but the computer system has no knowledge about this intersection. Most CADD systems probably can't even specify a route from point A to point B on the same road. The geometry is present, but the topology and network connectivity is missing.

Without topological intelligence, spatial analysis is difficult, if not impossible. For example, determining the number of dwelling units or the length of a roadway within a bounded area like a transportation analysis zone requires both geometric manipulation and appropriately structured data, which most CAM systems do not support.

GIS is a tool whose applications are limited only by the sophistication of the hardware and software, the quality of the data, and the imagination of the users. The following two sections summarize some of the GIS activities on-going at the FHWA and the state DOTs that illustrate the usefulness of this technology.

GIS ACTIVITIES AT THE FEDERAL HIGHWAY ADMINISTRATION

The FHWA Office of Planning has worked closely with the Bureau of the Census, the U.S. Geological Survey (USGS), the City of Columbia, and the Caliper Corporation in a project to assess the applicability of the Census topologically integrated geographic encoding and referencing (TIGER) file for transportation planning and analysis. A number of data sources are used in the demonstration: TIGER/Line file for Boone County, 1980 census data, Urban Transportation Planning Package (UTPP), and accident records and sign inventory compiled by the Columbia Department of Public Works.

The effort clearly illustrates the usefulness of the TIGER/Line file to solve transportation problems. A number of transportation-specific applications of GIS are demonstrated:

- A connected transportation network is built from TIGER/Line;
- The shortest path and traffic assignment are computed for this transportation network;
- Demographic characteristics are displayed and analyzed by census tract and traffic analysis zone; and
- An accident analysis system prototype making use of spatial relationship between accidents and traffic signs is developed.

TransCAD GIS software, developed by the Caliper Corporation, was used for the project. Some results are illustrated in Figures 1 to 5. Figure 1 is a portion of the road network from the TIGER/Line file. In Figure 2, population data from Census tape STF2 has been associated with tracts built from the TIGER/Line file. The tracts are then grouped by population, color coded, and labeled. In Figure 3, accidents and signs have been geocoded by latitude and longitude and are displayed as a layer over the TIGER/Line road network. As with all other objects in the data base, the accidents and signs have a large number of attributes associated with them such as drivers' ages and type of sign. TransCAD permits traditional SQL (Standard Query Language, developed by IBM) and boolean queries such as, "find all accidents within two miles of a high school involving male teenagers who have been drinking." Because it is a GIS, it is also possible to perform spatial queries such as finding all signs within 50 ft of an intersection that is going to be rebuilt or all accidents along a certain stretch of highway.

In Figure 4, a portion of the TIGER/Line road network has been converted to a connected transportation network and the shortest path computed between two points. The origin is assumed to be a retirement home and the destination a shelter. The purpose of the computation is to develop an evacuation route in case of a chemical spill at a factory near the retirement home.

The result of performing a traffic assignment as part of the four-step transportation planning process is shown in Figure 5. The assignments (link volumes) are shown as double bandwidths along the links in the regional network. The bandwidths are proportional to the actual vehicle loadings.

The FHWA Office of Planning has sponsored the development of the Geographic Roadway Information Display System (GRIDS). GRIDS is an interactive microcomputer program that accesses and displays data about the U.S. Interstate highway system. GRIDS produces a map of a state using data associated with state and county boundaries, city location and population, and interstate highway road sections. The GRIDS database is a subset of the data contained in the Highway Performance Monitoring System (HPMS) data base. This information can be displayed in various formats. It is possible to produce statistical summaries, pie and bar charts, and highlight highway sections with certain characteristics in distinctive colors. This concept could be expanded to meet many of the State's transportation data analysis needs.

The FHWA Office of Policy has developed a Highway Traffic Forecasting System (HTFS). The HTFS is a specialized decision support system that is designed primarily to analyze truck size and weight issues. It can also be used to analyze a wide variety of other policy issues that affect the highway network. The current model uses a simplified representation of the national Interstate highway network. Routines generate freight traffic and assign it to the network. The network uses Bureau of Economic Analysis regions as nodes with single link inter-

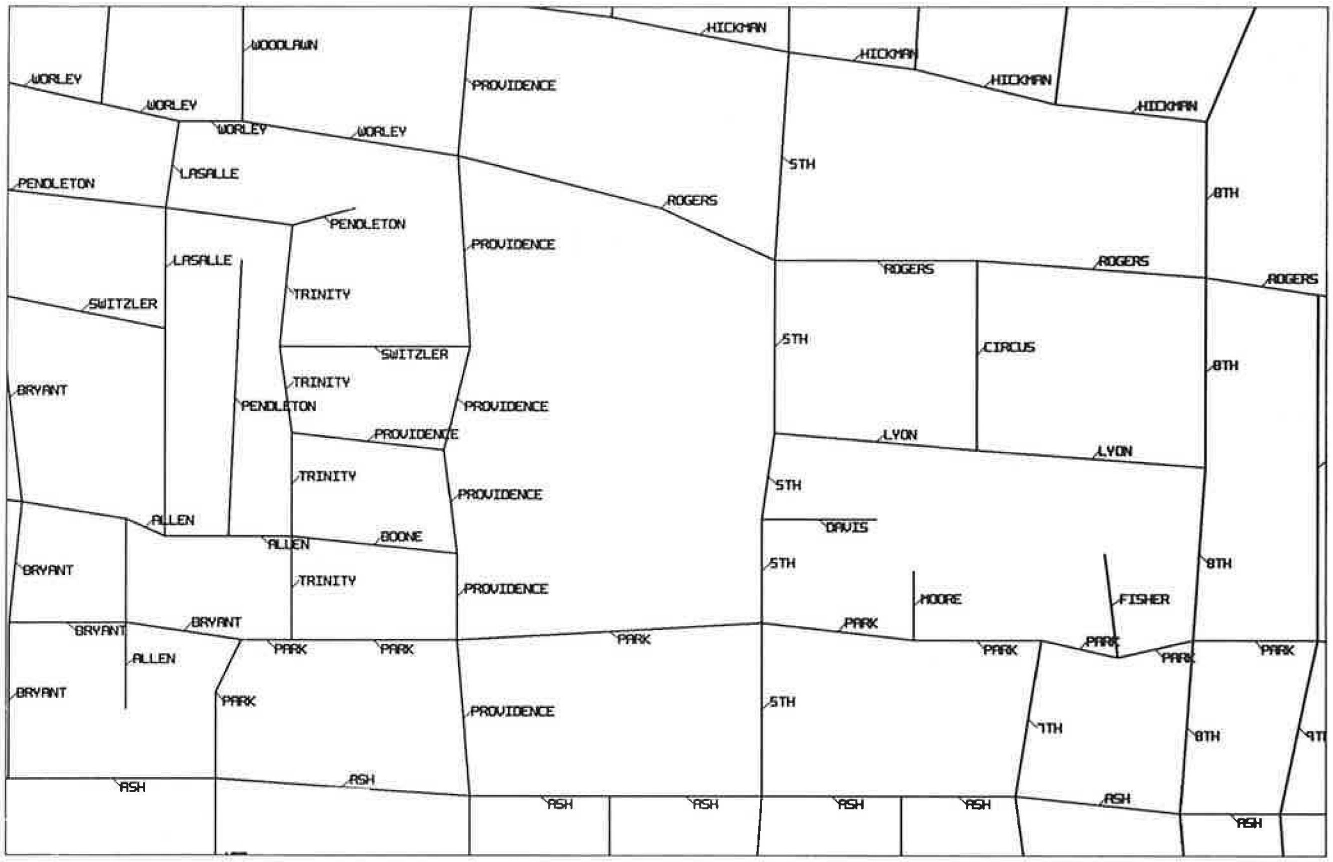


FIGURE 1 Boone County TIGER file—road network.

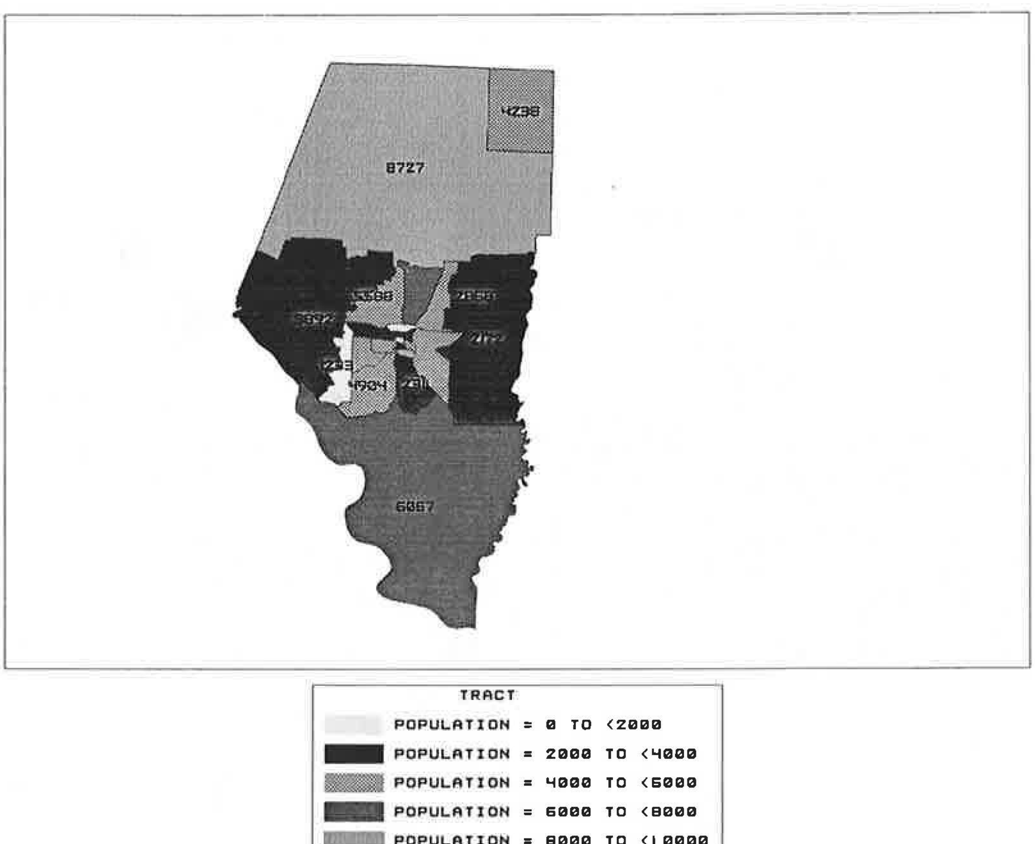


FIGURE 2 Census tracts grouped by population.

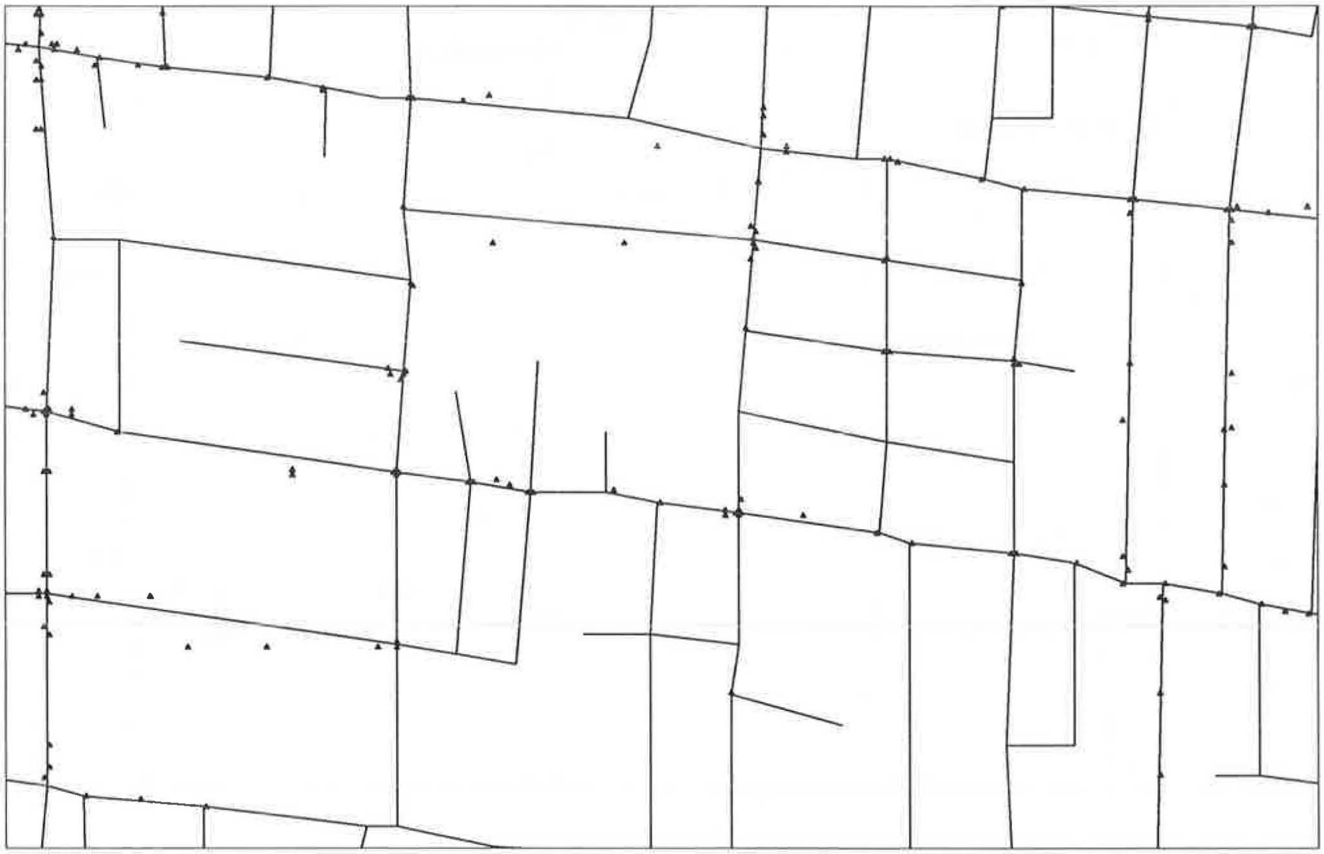


FIGURE 3 Accidents and signs geocoded to TIGER file.

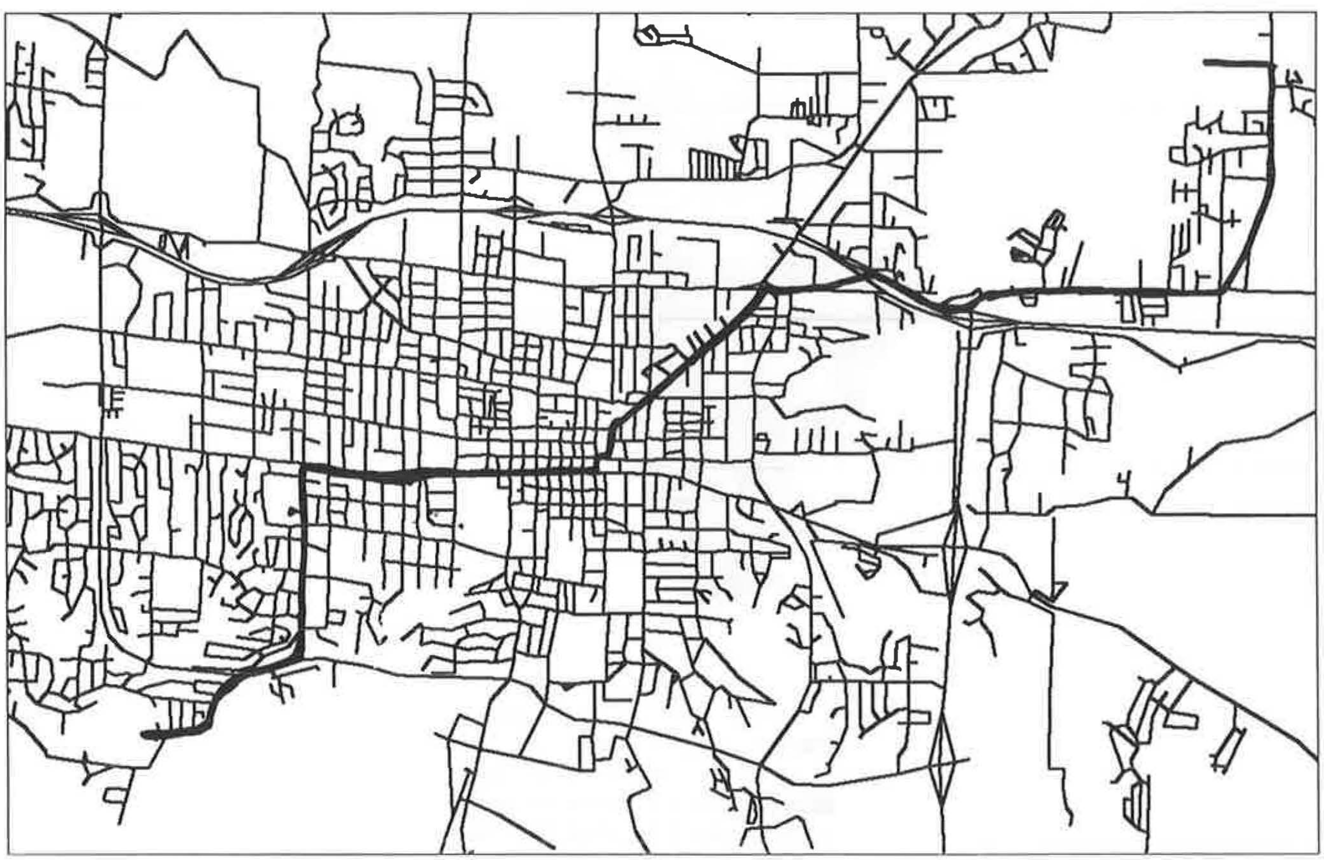


FIGURE 4 Shortest-path evacuation route.

FIGURE 5 Traffic assignments shown as dual bandwidths.

connections. The current network model generates a relative distribution of Interstate (long-haul) freight traffic among 10 geographic regions. When truck size and weight regulations change, the distribution of traffic among vehicle types and regions will also be changed by the model. The model is being converted to run on a microcomputer using Caliper Corporation's TransCAD GIS software (*6*).

The FHWA Office of Engineering and Highway Operations Research and Development is developing a nationwide databank of pavement research results (*7*). The mainframe computer model permits other geographically based nonhighway data, such as weather records, to be accessed for statistical and other analysis needs. The data base is designed to maintain compatibility with other current and proposed nationwide systems. Plans are to move the model to a microcomputer.

GIS ACTIVITIES AT STATE TRANSPORTATION AGENCIES

Several state transportation agencies are doing interesting research and development in the general area of GIS. This section summarizes a number of these activities.

Alaska—Integrated Highway Information System

Alaska's Highway Analysis System (HAS) is a mainframe data base of highway-related information and a collection of computer programs to process and analyze the data and to produce reports (*8*). The data base contains a number of files that had in the past resided on paper, personal computers, or in isolation on the mainframe. The HAS contains highway inventory, traffic, accidents, project history, HPMS, pavement conditions, bridges, signs, permit locations, right-of-way location information, and railroad crossings. HAS also contains information to integrate any of the above data to photolog and cartographic data bases. One interesting feature is that the user is able to choose the method most comfortable for "viewing" the data. A user can locate events on the roadway according to a route name and a milepoint from the beginning of the road, by a displacement (feet or miles) from a point (e.g., an intersection or a reference marker), by engineering station, or by state plane coordinate. This is made possible by using a common x,y,z coordinate georeferencing system.

Idaho—GIS Coordination Among State Agencies and the USGS

The State of Idaho has formulated a policy to encourage the use of GISs when such use enhances the overall cost effectiveness of administrative functions or improves productivity (*9*). It is the state's policy to acquire and support GISs through well-planned implementation strategies that include: (a) development and maintenance of data standards for base category data, statewide exchange data, and project data; and

(b) development and maintenance of contracts for state agency use covering the purchase of GIS software and hardware. To implement this policy, Idaho has established a statewide Geographic Information Advisory Committee that has responsibility for developing data standards for GISs and for overseeing GIS activities, including the purchase of GIS hardware and software. Idaho also has established a work share agreement with the USGS for the creation of the 7.5-min digital line graphs for the State.

Virginia and Tennessee—Feasibility of Using Global Positioning System in a Mobile Unit to Collect Roadway Data

The Virginia DOT and the Virginia Transportation Research Council have instituted a study to determine the feasibility of using a mobile unit at speeds of from 30 to 35 mph to collect selected roadway data on secondary, primary, and Interstate roads in Virginia using Global Positioning System (GPS). The study is designed to answer a number of questions related to the feasibility of using GPS: (a) the rate of travel in the vehicle on the various types of roads as data are collected; (b) the accuracy of the collection of the roadway items (number collected versus actual number in place); and (c) the cost per mile to collect data on the various types of roads (*10*). The Tennessee Department of Transportation has conducted a similar study (*11*).

Wisconsin DOT—Integrating Photolog Data into a GIS

The Wisconsin Department of Transportation (WisDOT) has been using photolog to assist in the planning, design, and operation of the state highway system since 1975. The system creates 35-mm camera images every 0.01 mile (52.8 ft). It also collects horizontal and vertical vector displacements of the vehicle's bearing and slope and calibrates these numbers to the photographs. The photo images are transferred directly from the negatives to a video disk. Each disk can hold up to 24,000 frames or 240 miles of highway. A microcomputer controller allows interactive random access by either frame number or highway and log-mile. Search and scan capabilities are available, including the ability to advance the frames at simulated speeds. Any frame on the disc can be randomly accessed in less than 1 sec.

WisDOT has integrated the photolog with its GIS (*12*). It is no longer necessary for the analyst to know the appropriate log-mile. One need merely point to a map location, and the system will retrieve the correct image. This permits access to all data about a specific location in a coherent, logically integrated manner.

ADDITIONAL APPLICATIONS OF GIS TO TRANSPORTATION

The last two sections have discussed a number of ways that the FHWA and state transportation agencies are using GIS. This section presents some additional applications.

Bridge Maintenance and Management

The application of GIS to bridge maintenance and management provides opportunities similar to those for pavement management (discussed in the section on spatially related data). In addition, detailed drawings and inspection records of each bridge could be maintained on laser disk and accessed through the GIS.

Traffic Engineering

The FHWA has developed a number of sophisticated traffic optimization and simulation models. A GIS could be used as a front-end processor to a collection of these models. The GIS would handle all data entry, editing, and display, including the real-time display of traffic patterns.

Highway Safety

The applicability of GIS for accident analysis has previously been discussed. The technology currently exists to take GIS into the field. Police officers could enter accident reports directly into a lap-top computer. On country roads it is often difficult to pinpoint the exact location of an accident. If the officer could display a detailed map of the area, including sufficient shape points, on the computer screen, he or she could more easily identify the exact milepoint location and use the cursor keys or a mouse to record it on disk. The officer could also enter the side of the road, lane, or other locational information. In the future, it may be economically feasible for each patrol car to have its own GPS receiver for automatic geographic coordinate determination.

Relations with Public Interest Groups and the Legislature

A large part of the work performed by the planning office in a transportation agency involves preparing materials for presentation to the public or elected officials. A GIS that integrates the agency's data base will make this process considerably easier and less time consuming and will provide better organized information for these groups.

Transportation-Demand Modeling

Transportation-demand modeling has been discussed, in part, in the section on spatially referenced data. A few additional points should be made. Historically, transportation-demand modeling has required the tedious building of abstract travel networks involving large numbers of nodes, links, and population and employment centroids. This entire process, including the data creation and editing, algorithms to solve the travel demand estimation, assignment to links, mode choice, and display of results can now be accomplished within a GIS. A GIS will make it easier to perform subarea analysis and to modify the traffic analysis zone structure as required by the focus of each study.

Routing of Hazardous Materials

The shipment of hazardous materials involves a tradeoff between cost of shipment and exposure to population centers. A GIS with the appropriate optimization algorithms can be used to find the "shortest path" between two points, where "shortest" is some function of cost, population exposure, and risk. The GIS can then print out the route and suggest departure times to minimize the potential for traffic accidents. A GIS that combines road, rail, air, and water networks could provide even better solutions to this growing problem.

SOURCES OF DIGITAL CARTOGRAPHIC INFORMATION

A GIS requires digital data with geographic coordinates. Dildine (*13*) lists a number of methods to obtain geocoded data:

- *Use of existing digital data.* There are a number of files available from federal agencies and private providers. These include the USGS digital line graphs (DLGs) and the Bureau of the Census TIGER file. These files are described in detail in the next section.
- *Scanning.* Scanning is the process of electronically reading a map and converting it into a raster (point) image. Since a raster format can not represent a connected network, it must be edited, converted into a vector format of links and nodes through the process of vectorization, and edited again. This process is most successful when the map being scanned is drawn clearly and contains only one map feature.
- *GPS.* The Navigation Satellite Timing and Ranging System (NAVSTAR), more commonly known as Global Positioning System (GPS), is a satellite system being developed by the Department of Defense. Presently, six satellites providing positioning information are in orbit. This six-satellite constellation can be used for measurements only during a limited time each day. An 18-satellite constellation providing 24-hour coverage is expected to be fully operational between 1990 and 1992. These satellites will then provide very precise three-dimensional information on a continuous basis (*14*). Today the time window for collecting coordinates is limited to 4 hours per day because of an insufficient number of satellites. A further complication is that this window changes daily and may not occur during daylight hours. However, it is safe to assume that GPS will soon have a major role in surveying and mapping. This technology potentially offers at least a 6:1 increase in speed over hand digitizing.
- *Image Processing to Update Map Features.* In this process, a recent aerial photograph is superimposed over a GIS basemap. The operator is able to identify new features while the software identifies features that have changed. The operator then enters all modifications into the GIS.
- *Photogrammetry.* Photogrammetric techniques provide a high degree of accuracy, but they are not necessarily any more economical than hand digitizing. This technology could be used effectively for updating base maps (*13*).
- *Hand Digitizing.* Hand digitizing is the slowest method of inputting cartographic base data. It is an appropriate method when no potential for automation exists, usually because of poor source documents or when the project is small in scale.

DIGITAL DATA BASES FOR TRANSPORTATION

There are a number of excellent sources of digital data available to the transportation professional.

Digital Line Graphs

The USGS National Mapping Program produces a number of digital map products at large, intermediate, and small scales for the United States. The principal data source is the National Digital Cartographic Data Base (NDCDB). Currently, the NDCDB contains base categories of digital cartographic data that include: geographic and other coordinate reference systems, hydrography, hypsography, transportation, boundaries, miscellaneous culture, geodetic control, and vegetative and nonvegetative cover. In addition, geographic names and the land use and land cover and associated mapped categories of census tracts, political boundaries, hydrologic units, and federal and state land ownership are included in the NDCDB (*15*).

There are a number of digital products produced and sold by USGS. The DLG files are digitized layers for transportation and hydrography at 1:24,000 (1 in. equals 2000 ft), 1:100,000 (1 in. equals 1.5 miles), and 1:2,000,000 (1 in. equals 30 miles) scales. The DLG transportation data are composed of topologically structured link and node records depicting the roads and trails found on the source material. The 1:100,000 series is complete for the continental United States. The 1:2,000,000 series maps are complete for the the entire United States. The 1:24,000 scale maps are available for some locations. The 1:100,000 DLGs were prepared from source materials originally drawn at 1:24,000 scale. They contain attributes that describe the road classes, but character fields (such as street names) are not supported. A major use of the 100,000 DLG is in the Bureau of the Census' TIGER file (*16*). Other products include digital elevation models at 1:24,000 and 1:250,000 scales and land use and land cover at 1:100,000 and 1:250,000 scales.

The 1:2,000,000 scale is too small for most state transportation purposes. By contrast, the 1:100,000 and 1:24,000 DLGs offer sufficient detail for most applications. However, to use the DLGs, a number of complex operations must be performed. First, the individual quads must be edge matched to join corresponding links. Next, routes must be identified and names assigned to the links. Finally, attribute data in the agency's data base, which is probably referenced by milepoint, need to be attached to the DLGs that are geocoded in latitude and longitude. Because of the large amounts of data involved, these steps should be automated using sophisticated software.

Since the 1:100,000 DLGs are available for the 48 contiguous states, they provide a quick and inexpensive way to build a GIS. As the 1:24,000 DLGs become available, they could be substituted for the corresponding smaller-scale data.

Census TIGER Files

As part of its preparation for the 1990 Census, the Bureau of the Census, in cooperation with the USGS, has been building a topological file containing every street and block face in the

United States. The availability of TIGER could be important in how transportation agencies, particularly those with a local or regional focus, establish GISs. [For a more detailed discussion on using TIGER for transportation purposes, see (*17*)]. The GBF/DIME files used in the 1980 census have been edited to remove topological errors (e.g., block boundaries are now guaranteed to achieve closure). These have been merged with USGS 1:100,000 scale DLGs (topological map files) for all areas of the country not covered by GBF/DIME. The result is a seamless, nationwide data base that will provide the foundation for the 1990 decennial census (*18*).

The transportation data in TIGER is derived from the following three sources:

1. GBF/DIME—within urban areas previously covered by DIME data; not quadrangle based;
2. Extended DIME—new files created by the Census Bureau to extend their DIME data to the nearest 7.5-min quadrangle border; and
3. 100,000 DLG—all remaining areas (*16*).

The first product of use to the transportation community is the TIGER/Line file, an extract of selected information from the TIGER file data base organized as a topologically consistent line network. Each street or block side (including selected nonstreet features) is coded as a separate record in the TIGER/Line file. Associated with each record are the census geographic area codes found to the left and right of the feature segment represented by the record, the name of the feature (including the relevant census feature class code identifying the segment by category) and, for selected areas, the address range and associated ZIP code of the street segment.

The TIGER/Line file contains record types that identify segment records, shape records, selected 1980 geographic area codes, and additional feature name and address information applicable to the segment. The shape records provide coordinate values that describe the shape of the segment. Generally, shape records will be available in areas outside the area covered by the 1980 GBF/DIME files. The Census will make TIGER/Line files available on a county basis (*19*).

CONCLUSIONS

Transportation agencies are still in their infancy with respect to exploiting the power and possibilities offered by GIS technology. I have attempted to explain GIS from the point of view of the needs of the transportation professional. Examples of GIS undertakings by FHWA and the states have been described along with a summary of sources for digital geocoded data. These examples show there is no shortage of opportunities to apply GIS technology to transportation problem solving.

However, there are a number of difficult questions that will have to be answered along the way. How does the GIS relate to the agency's mapping activities? Can an existing cartographic data base, such as the DLGs, be used? Is there an easy way to link milepoint-referenced attribute data to the coordinate-based GIS? How is this enormous quantity and diversity of data kept up to date? How is the issue of network overlay dealt with so as to minimize data redundancy but not lose necessary detail? What should be the structure of the computer environment?

With the attention that GIS is currently receiving by the transportation community, it seems certain that these questions will soon find answers. As computer hardware continues to become less expensive and more powerful, and as the software continues to grow more sophisticated, cost-effective GIS-based solutions to traditional transportation problems will become more and more common within the transportation community.

REFERENCES

1. D. W. Briggs and B. V. Chatfield. Integrated Highway Information Systems. *National Cooperative Highway Research Program Synthesis of Highway Practice No. 133.* TRB, National Research Council, Washington, D.C., 1987.
2. T. L. Nyerges and K. J. Dueker. *Geographic Information Systems in Transportation.* Prepared for the FHWA, Office of Planning, HPN-22, Washington, D.C., 1988.
3. W. T. Baker et al. Highway Location Reference Methods, National Cooperative Highway Research Program. *Synthesis of Highway Practice No. 21.* TRB, National Research Council, Washington, D.C., 1974.
4. *Regional Computer Assisted Cartography Conferences Summary.* Cosponsored by the Pennsylvania Department of Transportation, the Washington State Department of Transportation, and the FHWA, Office of Planning, Washington, D.C., 1988.
5. K. J. Dueker. Geographic Information Systems and Computer-Aided Mapping. *American Planning Association Journal*, Summer, 1987.
6. R. R. Schmitt. *Geographic Information Systems and the Transportation Studies Division.* Internal memo, FHWA Washington, D.C., 1988.
7. D. M. Freund and C. G. Conley. *The Pavements Databank: A Research Application of a Geographic Information System.* Presented at the 25th Annual URISA Conference, Los Angeles, Calif., August 1988.
8. K. R. Fox. *Description of Alaska's Highway Analysis System.* Boeing Computer Services, Seattle, WA, 1988.
9. M. Sety. *Policy Statement for Geographic Information Systems for the State of Idaho.* Idaho Geographic Information Advisory Committee, Boise, Idaho, 1988.
10. D. C. Wyant. *Feasibility of Using Global Position System in a Mobile Unit to Collect Roadway Data.* Virginia Transportation Research Council, Charlottesville, VA, 1988.
11. W. Aldridge. FHWA, Tennesse Division Office, series of memos and reports.
12. D. Fletcher. *Integrating Photolog Data into a Geographic Information System.* Wisconsin Department of Transportation, Madison, WI, 1987.
13. G. Dildine. *Geographic Information Systems Task Force, Feasibility Report.* North Carolina Department of Transportation, Division of Highways, Raleigh, N.C., May, 1988.
14. NCHRP, *Kinematic Differential GPS Satellite Surveying.* Project No. 20-23, National Cooperative Highway Research Program, TRB, National Research Council, Washington, D.C., 1988.
15. U.S. Geological Survey, "Guidelines for Acceptance of Digital Cartographic Data into the National Digital Cartographic Data Base," Digital Cartography Program Office, Reston, Va., 1987.
16. L. Batten. *Geographic Information Systems Research Utilizing 1:100,000 Digital Line Graph Data.* U.S. Geological Survey, Reston, Va. 22092, 1988.
17. H. J. Simkowitz. *Using the Census TIGER/Line Files for Transportation Planning.* Presented at the National Conference on Effective Planning Techniques for Small and Medium-Sized Areas, Phoenix, Arizona, December 6–8, 1988.
18. R. A. LaMacchia, S. G. Tomasi, and S. K. Piepenburg. *The Tiger File: Proposed Products.* Geography Division, Bureau of the Census, Washington, D.C., 1987.
19. G. Boudriault. *Topology in the TIGER File.* Geography Division, U.S. Bureau of the Census, Washington, D.C., 20233.

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

Design and Implementation of Intercity Origin-Destination Surveys

ROBERT W. STOKES AND T. CHIRA-CHAVALA

This paper summarizes the results of a postcard origin-destination (O-D) survey conducted to estimate intercity travel patterns in the Interstate 35 corridor between Austin and San Antonio, Texas. The paper discusses the study design, data-collection and data-processing procedures, methods used to check the representativeness of the sampled data, expansion of the sampled data to represent total corridor travel, and study costs. The results of the study indicate that the postcard O-D survey method can be used to develop representative, reliable estimates of intercity corridor travel patterns. Additionally, the results of this study show that with a good traffic control plan and a well-trained survey crew, this survey method can be safely implemented without causing any substantial traffic delays, even on high-volume intercity and interstate freeways.

As a result of current and projected growth in the I-35 corridor between Austin and San Antonio (Figure 1), the Texas State Department of Highways and Public Transportation (SDHPT) is undertaking an analysis of alternative corridor improvements. The possibility of an alternate route to the east of I-35 (Figure 1) has received considerable attention in recent years. However, other alternatives have not been eliminated from consideration at this date.

Current and statistically reliable information concerning interurban origin-destination (O-D) travel patterns in the Austin–San Antonio study area was needed to conduct the analysis of alternative corridor improvements. This paper summarizes the O-D survey that was conducted to identify current travel patterns in the study corridor. The paper describes the O-D survey study design, field data collection and data processing procedures, checking the representativeness of the sampled data, expansion of the sampled data to represent total corridor travel, and the study costs. The details of the study are presented elsewhere (*1–3*).

STUDY DESIGN

This section describes the alternative survey methods evaluated for possible use in the study, criteria for selecting survey stations, scheduling of the survey, and the sample-size determination.

Survey Method

Five traditional O-D survey methods were evaluated for possible use in the study corridor (*1*). They were roadside interview, postcard survey, license plate "trace" survey, license plate "mail-out" survey, and tag-on-vehicle/lights-on survey. The advantages and disadvantages, manpower requirements, typical response rates, and approximate sample sizes for these methods are summarized in Table 1, with the methods listed in descending order of cost and accuracy.

Neither the license plate trace method nor the tag-on-vehicle/lights-on surveys are applicable to a large intercity traffic corridor, such as the Austin–San Antonio corridor, because of the extreme difficulties in planning and implementing the survey. The manpower requirements to implement either one of these methods on a corridor of this size would be unrealistic and the analysis of the field data would be extremely cumbersome.

The license plate mail-out survey has a number of short-

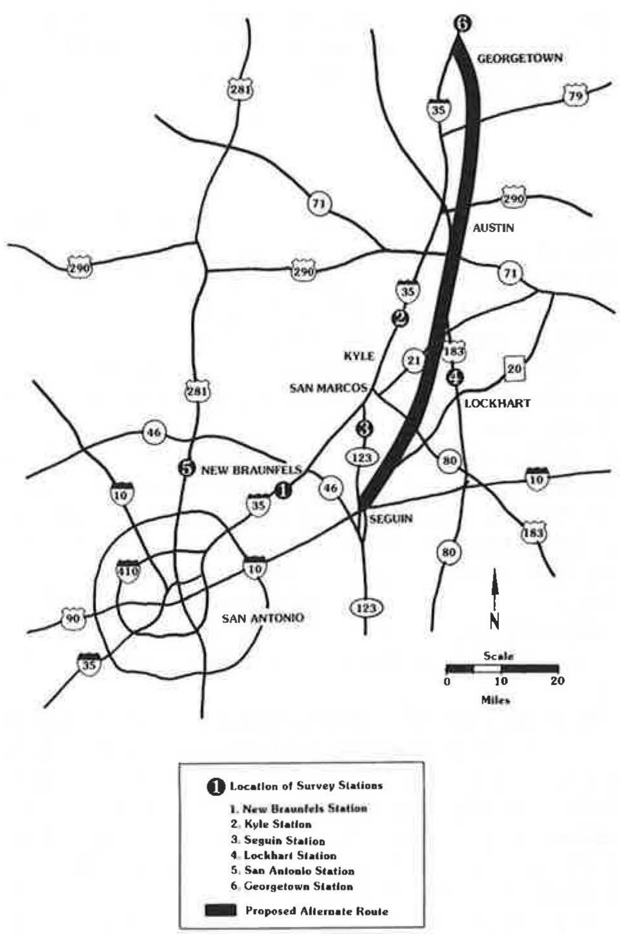


FIGURE 1 Austin/San Antonio study corridor.

R. W. Stokes and T. Chira-Chavala, Texas Transportation Institute, The Texas A&M University System, 333 West Loop North, Suite 116, Houston, Texas 77024. Current affiliation: T. Chira-Chavala, Transportation Research Board, Washington, D.C. 20418.

TABLE 1 SUMMARY OF O-D SURVEY METHODS

Survey Method	Advantages	Disadvantages	Manpower Requirements per Survey Station	Recommended Sample Size[b]	Typical Response Rates
1. Roadside Interview	• Complete information • High Response Rate • Better Sampling Control	• Relatively expensive • Traffic delays • Hazardous	• 10-20 persons/ station[a] • 2-4 police officers	20%-50%	100%
2. Postcard Surveys	• Can be completed quickly • Less traffic delay • Relatively inexpensive • Good population coverage	• Possible bias due to better response by some drivers • Low response by thru and out-of-state traffic • Requires stopping traffic • No provision for follow-up of non-responses	• 5-9 persons/ station • 1-2 police officers	60%-80%	25%-35%
3. License Plate Surveys					
a) "Trace" Method	• Simplicity of field organization • No interference with traffic • Unbiased Sample	• Data Analysis is difficult • Large number of stations required • Possible recording errors • Survey stations must operate simultaneously	• 2-3 persons/ station	35%-50%	60%[c]
b) "Mail-out"	• Similar to Method No. 2, except followup of non-responses is possible • Stations need not operate simultaneously	• Same as Method No. 2, except does not require stopping traffic • Requires access to vehicle registration files	• 2-3 persons/ station	60%-80%[d]	20%-35%[d]
4. Tag-on-vehicle/ Lights-on Surveys	• Same as Method 3a, except may result in minor traffic delays	• Same as Method 3a, except less recording errors	• 2-3 persons/ station	100%	-

[a]Number of interviewers varies with traffic volume but on the average is about 3-4 times the number of persons required to hand-out postcards. The above estimate is for relatively low hourly traffic volumes.
[b]Sample sizes have been adjusted for typical response rates to insure at least 20% sample.
[c]Response rate is estimate of percentage of license plates which can be traced.
[d]Response rate can be increased by follow-up of non-responses.

comings if applied to this study. The most notable problem is that after the vehicles passing a station are selected and their license-plate numbers read, it would be difficult to send questionnaires to drivers of trucks or out-of-state vehicles and it would be almost impossible to reach drivers of leased vehicles. This survey method might therefore result in noncoverage of many subgroups within the population that cannot be easily corrected for.

Postcard surveys may be based on the "controlled" or "roadside-distribution" method. The former utilizes vehicle-ownership or licensed-driver records, whereas the latter involves distributing the postcards to vehicles passing the survey stations. The controlled postcard method suffers the same shortcomings as the license plate mail-out method in its inability to effectively survey trucks and leased and out-of-state vehicles. This method was therefore considered unsuitable for this study.

The roadside-interview and the postcard-distribution methods are similar in providing good coverage of the vehicle population and in the amount of information that can be effectively sought from the drivers. In terms of costs and manpower requirements, the roadside-interview method, on the average, requires 3 to 4 times more field personnel than the postcard-distribution method, and this estimate can be much higher for very high-volume facilities. A trained interviewer can complete about 30 to 40 interviews in an hour, while postcards can be handed out to drivers every 4 to 5 sec. The response rate of the roadside-interview method, however, may be up to 3 times higher than the roadside-distribution postcard method. Despite its higher response rate, the interview crew would need to work at least as long as the postcard-distribution crew in order to obtain a sufficient number of responses. One disadvantage of the postcard-distribution method is that the nonresponses may introduce biases. Therefore, a survey based on postcard distribution must include a mechanism for checking the nonresponses to ensure that they are not substantially different from those who respond to the survey.

In terms of adaptability, the postcard-distribution method is more desirable in terms of traffic delays, station set-up, traffic control plans, survey management, and safety to the survey crew and motorists. On a high-volume facility, such as Interstate 35, it would not be practical to stop traffic to complete interviews with drivers on-site because traffic congestion and delays could become excessive, even with a large interview crew. Furthermore, as the number of interviewers increases, so does the complexity of setting up the site and managing the survey in order to maintain safety and to minimize traffic delays and confusion. Previous experiences (*4–6*) with the roadside-distribution postcard method have shown that with a good traffic control plan, well-trained survey personnel, and the use of an appropriate vehicle selection technique, this survey method can be safely implemented without causing any substantial delays to traffic.

A combination roadside-distribution postcard and roadside interview survey was also considered. This approach would involve distributing postcards to drivers during high-volume time periods and conducting on-site interviews during low-volume time periods. Past experience suggests that such a combination would not enhance the amount of information obtainable, nor would it improve the quality of the survey results.

On the basis of these considerations, the roadside-distribution postcard survey method was selected. A sample of the postcard questionnaire used in the study is shown in Figure 2. The survey form was designed to solicit information concerning vehicle type, trip purpose, trip origin and destination, vehicle occupancy, and trip frequency. The survey form requested street address, city, and zip code of the trip origin and destination. This information made it possible to code O-D zones with sufficient detail to evaluate the range of improvements being considered for the corridor. The questionnaire portion of the form was printed on the back of a prepaid, preaddressed postcard. Also, each questionnaire was individually numbered to facilitate recording the time and location of distribution.

Selecting Survey Stations

The following six survey stations were selected (Figure 1): (a) I-35, between San Marcos and San Antonio, south of SH-46 (New Braunfels Station); (b) I-35, between Austin and San Marcos (Kyle Station); (c) SH-123, between I-35 and I-10 (Seguin Station); (d) US-183, between SH-21 and I-10 (Lockhart Station); (e) US-281, north of San Antonio between FM 1604 and SH-46 (San Antonio Station); and (f) I-35, north of Georgetown (Georgetown Station).

The I-35 stations between Austin and San Antonio, and the stations on SH-123 and US-183 were chosen to provide samples of intercity and through-traffic, as well as traffic with

AUSTIN/SAN ANTONIO ORIGIN-DESTINATION STUDY

Survey Station:
Northbound US 281 Near San Antonio № 95217

1. Type of vehicle?
   Passenger Car ☐ Pickup ☐ Van ☐ Other Truck ☐
2. Purpose of trip today?
   Work ☐ School ☐ Shopping ☐ Recreation ☐ Other ☐
3. Where were you coming from when you received this questionnaire?
   Street Address (or nearest intersection) City Zip Code
4. Where were you going when you received this questionnaire?
   Street Address (or nearest intersection) City Zip Code
5. How many people in vehicle (including driver)? ____
6. How many days per week do you make this trip?
   1 ☐ 2 ☐ more than 2 ☐ Other (please specify) ____
7. Any additional information on your trip that you think might be helpful to us would be appreciated. ____

FIGURE 2 Sample postcard questionnaire.

O-Ds at key intermediate points. These stations were considered to be particularly important in terms of assessing the potential feasibility of an alternate Austin–San Antonio route to the east of I-35. The US-281 station was selected to sample potential traffic for an alternate route between Austin and San Antonio to the west of I-35. The I-35 station north of Georgetown was identified to obtain a sample of traffic that might use an I-35 Austin bypass.

The following criteria were used to identify precise survey station locations.

- *Sight-distance.* The primary consideration in selecting survey stations was safety. Survey stations were located on flat, straight roadway sections that were clear of structures or other obstructions that could reduce sight-distances. Level and straight sections of highways with an unrestricted sight distance of 800 ft or more in each direction from the station were sought (*7*). Stations at or near intersections were avoided.
- *Roadway cross-section.* Wherever possible, survey stations were located where roadway width was at its maximum. On I-35, survey stations were located on sections with inside and outside shoulders. By using the freeway shoulders it would be possible to set up four-channel service areas for postcard distribution. On non-interstate roadways, survey stations were established on four-lane sections.
- *Traffic catchment area.* Survey stations were located to intercept a representative sample of intercity traffic. As a general guide, survey stations were located near the midpoints of the roadway links surveyed.

Scheduling the Survey

The following issues were considered in scheduling the O-D survey.

- *Month and day-of-week considerations.* The choice of the month and day-of-week of the survey depended on whether "typical" or "peak" O-D data were desired. An examination of monthly, daily, and seasonal traffic volumes as a percent of average annual daily traffic (AADT) from several permanent traffic recorders in the corridor revealed that the summer months of June through August generally account for the highest percentages of AADT. The fall months of September through November, on the other hand, appeared to be more representative of the AADT. In terms of average variations in the AADT, Mondays through Thursdays appeared to be "typical" days. Fridays, with their high percentages of pre-weekend traffic, tended to be higher-than-average traffic days. The objective of the study was to estimate the peak demand within the corridor. Therefore, the O-D survey was conducted on Tuesday through Thursday during the month of July.
- *Time-of-day considerations.* The O-D survey may be conducted over a 24-hour period, or more typically, during daylight hours. Given the hazards associated with nighttime operations, survey operations were restricted to daylight hours.
- *One-directional versus two-directional station operations.* In scheduling the survey and estimating manpower needs, the issue of whether each direction of travel was to be surveyed separately or simultaneously needed to be resolved. The FHWA (*7*) guidelines on conducting O-D surveys state

". . . two-directional surveying is necessary if hourly data describing origins and destinations by direction are needed. It is generally assumed that although inbound traffic patterns are similar to outbound traffic patterns for a 24-hour period, the differences are significant enough on an hourly basis to warrant two-directional surveys. Some serious problems could arise in the analysis of the data if two-directional data are not available. Where sufficient personnel are available, it is desirable to survey traffic in both directions simultaneously."

Harmelink (*8*) suggests that one-directional surveys would produce larger errors than would two-directional surveys. Hajek (*9*) found from actual O-D data that the errors for a 50 percent two-directional survey were very similar to the errors for a 100 percent one-directional survey. Hajek attributed this similarity in errors to the daily variation in traffic that might have obscured the expected difference between the two-directional and the one-directional surveys.

Miller et al. (*10*) reported that inbound and outbound frequencies of O-D trips were not exact mirror images of one another and that some differences between the two directions existed. The percent differences were likely to be higher for small trip interchange volumes than for larger trip interchange volumes.

On the basis of findings from these past studies, and to maximize the usefulness of the resulting O-D data, two-directional surveys were conducted.

Sample Sizes

A minimum sample size required at a given survey station is the number of vehicles sampled at the station whose drivers successfully complete the postcards or the interview. A minimum sample size required for an O-D survey of vehicles passing through a survey station is usually expressed as a sampling rate (i.e., a ratio of the number of vehicles sampled to the total number of vehicles passing through). The sampling rate is a function of the following: (a) p (proportion of total traffic volume at the survey station that has a particular O-D); (b) w [desired accuracy (percent error) of p]; (c) N (traffic volume at the survey station); and (d) Z (normal variate that is associated with a specified level of confidence in estimating the O-D interchange volume). The sample size formula (*9*) is given by

$$r = (Z^2pq)/(N - 1)w^2 + (Z^2pq)$$

where r is the required sampling rate, and q is $(1 - p)$.

To apply the sample size formula, some estimate of N must be known. A desired accuracy of p must be specified, as must a level of confidence in estimating p. The proportion of traffic volume at the survey station with a particular O-D must be specified. This proportion is usually not known during sample size determination. What must be specified, instead, is a minimum O-D trip interchange volume to be obtained from the survey with the desired accuracy level. In the context of this study, this minimum O-D trip interchange volume was assumed to be in the range of 2 to 10 percent of the traffic volume at the survey site.

Table 2 presents approximate sampling rates (r) for a range of average daily traffic (ADT) (N) and accuracy levels (error rates) from ± 5 to ± 15 percent. All calculations assume 95 percent confidence interval. Lower confidence intervals would, of course, result in lower sampling rates for a given ADT and accuracy level. The sampling rates shown in Table 2 assume a 100 percent response and must be adjusted for nonresponses as follows: number of vehicles sampled = (sampling rate × traffic volume)/response rate.

Table 3 summarizes recommended sample sizes for each of the survey stations in the study corridor. The sample sizes are given in terms of the number of postcards to be distributed at each station. The sample sizes were estimated from rates given in Table 2 and have been adjusted on the basis of an assumed postcard response rate of 30 percent (*1*). These recommended sample sizes were based on a 95 percent confidence limit and desired accuracy of ± 15 percent.

FIELD DATA COLLECTION

This section describes methods for setting up the survey stations, implementing traffic control plans, distributing the

TABLE 2 APPROXIMATE SAMPLING RATES FOR ERRORS WITHIN ±5%, ±10%, AND ±15% AT 95% CONFIDENCE

N	p = 0.03			p = 0.05			p = .10		
	±5%	±10%	±15%	±5%	±10%	±15%	±5%	±10%	±15%
3,000	.94	.81	.65	.91	.71	.52	.82	.54	.34
5,000	.91	.72	.53	.86	.59	.40	.74	.41	.24
10,000	.84	.56	.36	.75	.42	.25	.58	.26	.14
20,000	.71	.39	.22	.59	.27	.14	.41	.15	.07
30,000	.63	.30	.16	.50	.20	.10	.32	.11	.05
40,000	.56	.24	.12	.42	.16	.08	.26	.08	.04
50,000	.50	.20	.10	.37	.13	.06	.22	.07	.03
60,000	.45	.17	.09	.33	.11	.05	.19	.06	.02
70,000	.42	.16	-	.30	.10	-	.17	.05	-
100,000	.33	.11	-	.23	.07	-	.12	.04	-

Notes: N = Traffic Volume at Survey Station; p = Minimum O-D trip interchange volume to be estimated from the survey with the desired accuracy level (expressed as proportion of N). Sampling rates assume 100% response and must be adjusted for non-responses as follows: Number of Vehicles Sampled = (Sampling Rate X Traffic Volume)/Response Rate.

TABLE 3 RECOMMENDED SAMPLE SIZES FOR AUSTIN–SAN ANTONIO O-D SURVEY

Survey Station and Direction[a]	1985 ADT[b]	n[c]	p[d]
1. New Braunfels (I-35)			
NB	19,000	9,500	0.05
SB	19,000	9,500	0.05
2. Kyle (I-35)			
NB	20,000	10,000	0.05
SB	20,000	10,000	0.05
3. Seguin (SH 123)			
NB	4,000	4,000	0.10
SB	4,000	4,000	0.10
4. Lockhart (US 183)			
NB	3,300	3,300	0.10
SB	3,300	3,300	0.10
5. San Antonio (US 281)			
NB	9,650	5,800	$0.05<p<0.10$
SB	9,650	5,800	$0.05<p<0.10$
6. Georgetown (I-35)			
NB	13,500	8,100	$0.05<p<0.10$
SB	13,500	8,100	$0.05<p<0.10$

[a]NB = Northbound, SB = Southbound.
[b]Directional ADT assumes 50/50 split. Source: District Highway Maps, Texas State Department of Highways and Public Transportation.
[c]n = No. of postcards to be distributed.
[d]p = minimum O-D trip interchange volume which can be estimated from survey results with ±15% accuracy (expressed as proportion of ADT).

questionnaires, and collecting data to check the representativeness of the sampled data.

Survey Station Set-up and Traffic Control

With the high traffic volumes encountered on many of the roadways surveyed, great care was taken to ensure that the surveys were conducted in a safe, efficient, and professional manner. The actual distribution of the postcard questionnaires did not result in any substantial delay to individual motorists. The overall efficiency of the survey stations, therefore, was determined by the vehicle entry and exit set up at the survey station (i.e., the physical layout of the survey stations). Figures 3 and 4 show the basic setups used at the Interstate and non-Interstate survey stations, respectively. All survey stations had law enforcement officers on duty to ensure safety and to enhance motorist cooperation.

The survey stations were in operation from 7:00 a.m. to 8:00 p.m. each day. However, survey operations were occasionally suspended in order to minimize motorist delays. As a general rule, if traffic queues extended to the advance signing of the survey stations, survey operations were temporarily suspended until the queue was reduced.

Questionnaire Distribution and Data Collection

Four persons per Interstate site and two persons per non-Interstate site were required to distribute the postcard questionnaires. The questionnaire forms were bundled according to the 15-min time period during which they were to be distributed. The number of questionnaires per bundle was based on the sample sizes shown in Table 3. Additionally, postcard questionnaire identification numbers were recorded at the beginning and end of each 15-min survey period to ensure that the time and location of distribution could be identified when tabulating the survey responses.

In addition to distributing postcards, the survey crews also conducted manual counts of traffic volumes, vehicle classifications, and vehicle occupancies. At the Kyle Station, a nighttime vehicle classification study was conducted. Survey crews also recorded samples of vehicle license plate numbers at each of the survey stations. At the Kyle Station, postcard survey form numbers were recorded along with the license plate numbers of a sample of the vehicles surveyed.

The volume counts were used to expand the sample data to represent the entire vehicle population for the corridor, and the license plate data were collected to evaluate the representativeness of the sample data. The use of these data is discussed below.

DATA PROCESSING

To facilitate data analysis, the survey results and the volume classification and license plate data were coded for computer processing. The following accuracy checks were performed on the survey data.

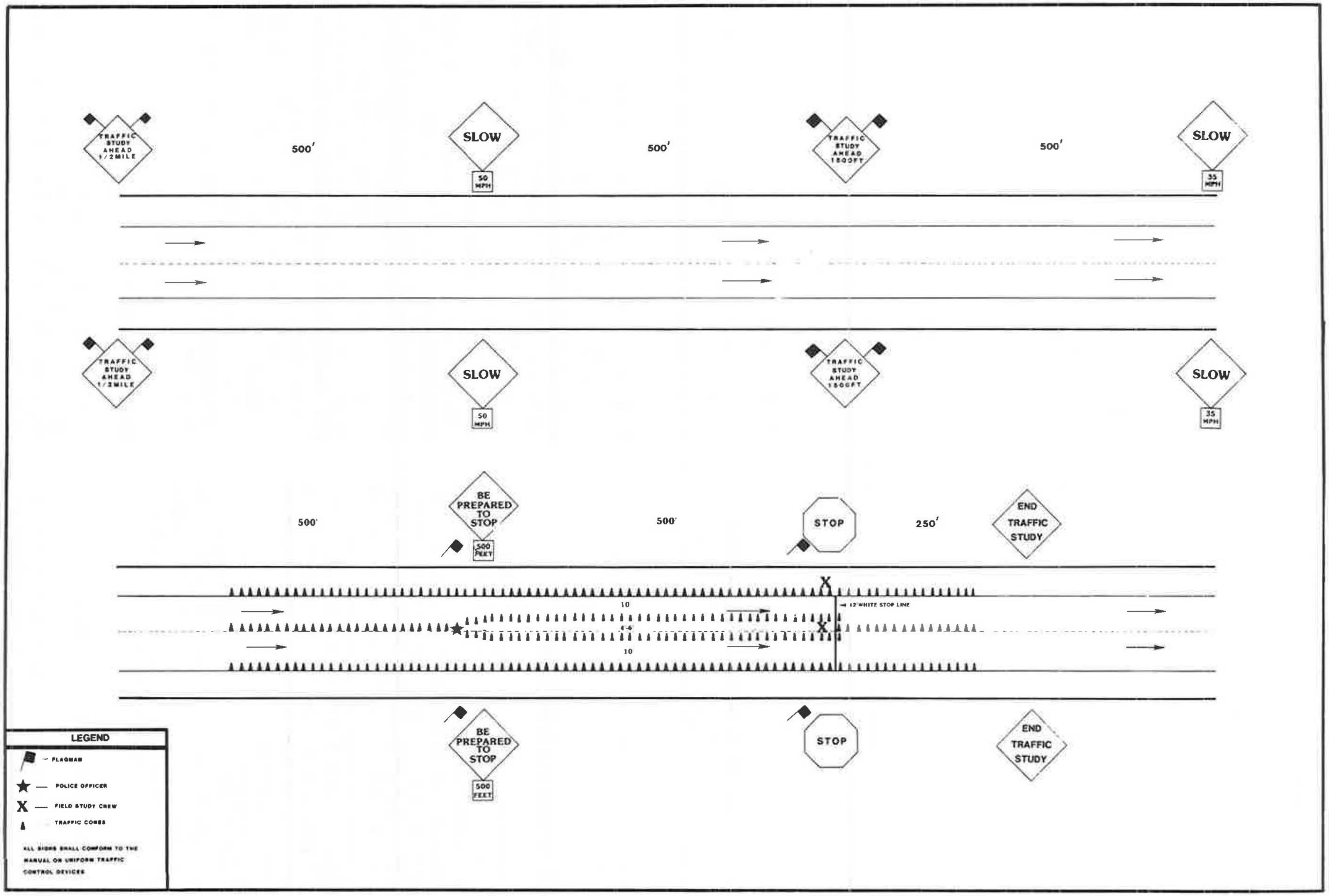


FIGURE 3 Interstate highway traffic control plan.

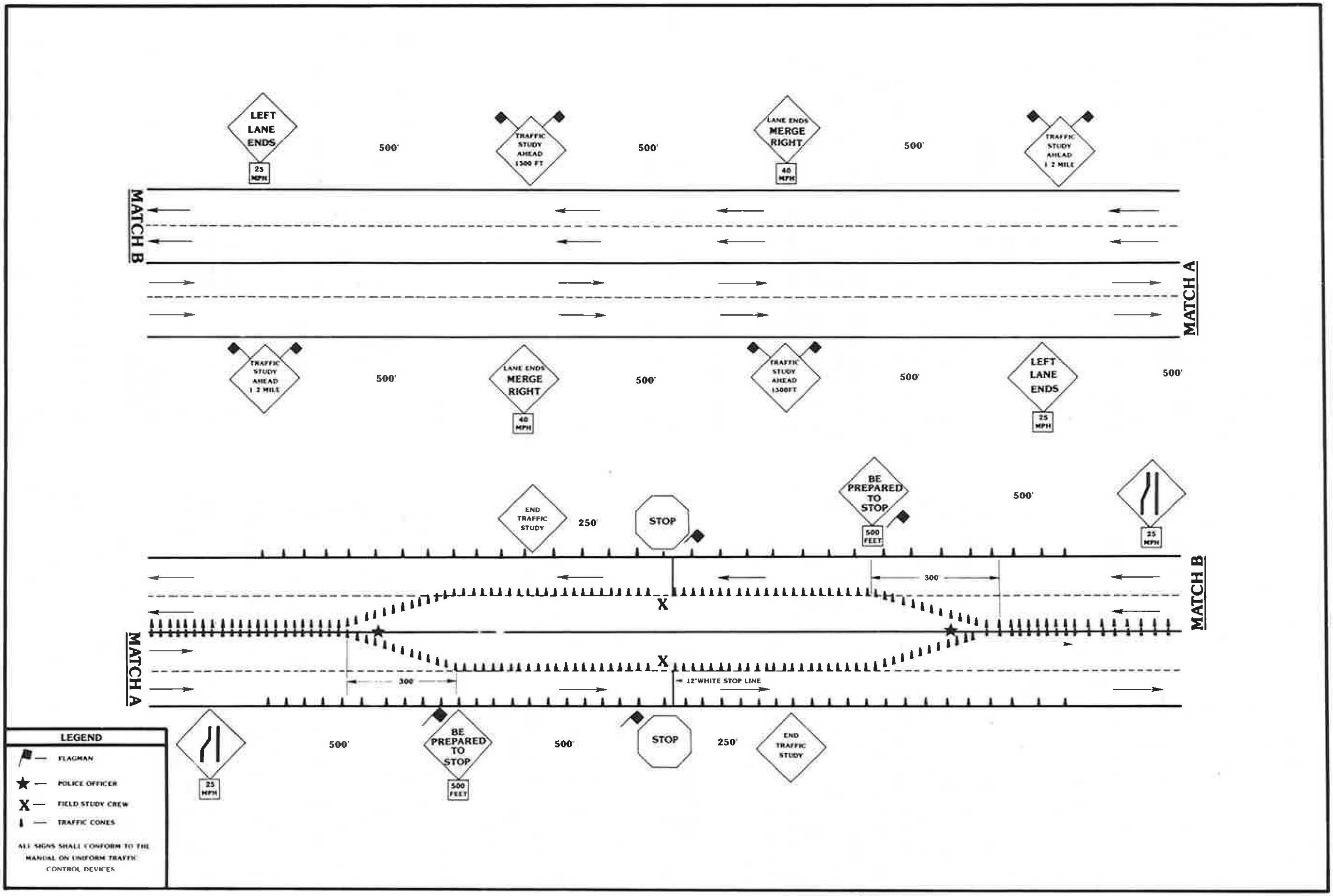


FIGURE 4 Non-Interstate highway traffic control plan.

- *Key-punch errors.* Tight quality control procedures were established for the data processing phases of the study. However, given the enormous amount of data that needed to be processed, it was recognized that coding and input (key-punching) errors would be unavoidable. In order to assess the magnitude and nature of these errors, approximately 1000 of the survey responses from the Kyle Station were processed a second time. These 1000 responses were manually checked to ensure they had been inputted correctly. Once this data set was "clean," it was merged with the initial entries and any mismatches were identified and evaluated. The results of this accuracy check indicated that the error in computer processing of the survey data was about 4 percent. However, the majority of the errors were for information not directly related to the primary objectives of the study (e.g., errors or inconsistencies in categorizing and coding comments or trip frequency).
- *Zip code reporting errors.* A zip code atlas and street address information provided by the respondents were used to compare the actual and reported zip codes of origins and destinations for 10 percent of the responses received from the Kyle Station. Approximately 5 percent of the responses examined were found to have errors in the zip codes reported for the origins or destinations. However, the errors were predominantly in the last two digits of the zip code. Because the zip code data were aggregated into large zones in the final data tabulations, these reporting errors had little effect on the overall accuracy of the results.

CHECKING REPRESENTATIVENESS OF SAMPLED DATA

The sample data were checked for possible biases caused by nonresponses, in order to ensure the representativeness of the population surveyed. Specifically, the data collected at the Kyle Station were used to perform the following analysis. The results of these checks indicate that the sample of travel patterns in the corridor obtained from the survey was representative of the population surveyed.

- *Geographic distribution of responses.* A comparison of the geographic areas (zip codes) of vehicle registrations for respondents and nonrespondents was performed to identify any bias in the survey results caused by the over- or under-representation of one or more geographic areas in the responses. This evaluation was performed using data from the Kyle Station, where it was possible to identify respondents and nonrespondents from the subset of vehicles whose license plate numbers had been matched with survey postcard numbers. The analyses revealed no significant geographic bias in the survey results.
- *Travel patterns of nonrespondents.* In an effort to assess whether the travel patterns of the survey respondents represented the travel patterns of all travelers in the corridor, a follow-up survey of nonrespondents was conducted. Approximately 80 nonrespondents, as identified from the subset of vehicles at the Kyle Station, were interviewed in a telephone survey. Although the sample size was too small to draw any definite conclusions, the analyses indicate that there was no substantial differences in the travel patterns of respondents and nonrespondents.

EXPANDING THE SAMPLE

A summary of the O-D sample by survey station is presented in Table 4, which shows nearly 83,000 survey forms were distributed during a 3-day survey period. Over 28,000 (35 percent) of the postcard questionnaires were returned. This response rate represents over one-fourth of the total traffic observed during the survey period. That is, more than one in four (29 percent) of the vehicles observed responded to the survey. The aggregate summary in Table 4 shows that roughly 90 percent of the vehicles observed were passenger vehicles. Trucks and other commercial vehicles accounted for the remaining 10 percent.

Once the O-D survey data were tabulated and checked, the sample results were expanded to obtain estimates of O-D volumes for the entire vehicle population of the study corridor. The observed traffic volumes were used to expand the sample data.

The sample data were expanded by survey station and direction for each of the following three time periods: (a) morning (7:00 a.m. to 11:00 a.m.); (b) midday (11:00 a.m. to 3:00 p.m.); and (c) afternoon (3:00 p.m. to 8:00 p.m.). The data were expanded by time period to account for possible differences in the travel patterns by time of day. Aggregate estimates of O-D volumes for the vehicle population were then obtained by simply summing over site and direction of travel.

Expansion Formula

The basic formulas used to obtain the estimates of the population O-D volumes and their standard errors are as follows:

$$p = t/n$$

$$T = pN$$

$$S_p = [p(1-p)/n]^{\frac{1}{2}}$$

$$S_T = N[p(1-p)/n]^{\frac{1}{2}}$$

where:

p = proportion of the reported trips having a particular O-D (for each site and direction),
t = number of trips reported for a particular O-D (for each site and direction),
n = total number of trips reported for each site and direction,
T = estimate of O-D volumes for the entire vehicle population,
N = observed traffic volume for each site and direction;
S_p = standard error of p, and
S_T = standard error of T.

A discussion of the relative efficiencies of alternative expansion procedures can be found elsewhere (*11*).

Total Corridor Travel

The estimated 1987 vehicle trip interchanges for the major O-D zones in the corridor are summarized in Table 5. Also

TABLE 4 SUMMARY OF AUSTIN–SAN ANTONIO O-D SAMPLE

SURVEY STATION	OBSERVED TRAFFIC VOLUME (7:00 A.M. - 8:00 P.M.)						SURVEY DISTRIBUTION		SURVEY RESPONSE		
		Commercial Vehicles									
	Passenger Vehicles	Single Unit	Combination	Tractor Only	Buses	Total Vehicles	Number Distributed	% Traffic Surveyed	Number	Return Rate	% Tot. Veh. Responding
1. NB	12322	612	1130	40	25	14129	12009	85%	4152	35%	29%
SB	12335	704	1116	20	18	14193	12484	88	4560	36	32
Total	24657	1316	2246	60	43	28322	24494	86	8712	36	31
2. NB	12498	396	939	19	19	13871	12461	90	4128	33	30
SB	12931	566	1025	8	23	14553	12583	86	4119	33	28
Total	25429	962	1964	27	42	28424	25044	88	8247	33	29
3. NB	1933	108	81	3	2	2127	1914	90	698	36	33
SB	2098	116	97	4	1	2316	1919	83	638	33	28
Total	4031	224	178	7	3	4443	3833	86	1336	35	30
4. NB	2014	303	74	5	5	2401	2178	91	778	36	32
SB	2559	99	89	3	3	2753	1898	70	822	43	30
Total	4573	402	163	8	8	5154	4076	79	1600	39	31
5. NB	4485	207	59	1	3	4755	3858	81	1617	42	34
SB	4252	165	71	1	2	4491	3335	74	1481	44	33
Total	8737	372	130	2	5	9246	7193	78	3098	43	34
6. NB	8198	500	956	18	13	9685	9000	93	2510	28	26
SB	8608	430	898	13	19	9969	9000	90	2561	28	26
Total	16806	930	1855	31	32	19654	18000	92	5071	28	26
Total	84,233	4,206	6,536	135	133	95,243	82,639	87	28,064	34	29

shown in the Table are the cell percentages and the standard errors of the estimates.

As shown in Table 5, the Austin, San Antonio, and San Marcos–New Braunfels areas account for over 75 percent of the O-Ds in the corridor. The relatively high percentage of O-Ds observed for the San Marcos area (23 percent) is particularly significant in terms of the need for an alternate route in the corridor. Since nearly one-quarter of the trips in the corridor have origins and destinations on I-35 between Austin and San Antonio, it seems unlikely that a substantial percentage of these trips would find an alternate route east of I-35 particularly attractive.

The diagonal of the trip table represents round-trips in the corridor. Since the survey questionnaire (Figure 2) requested information concerning origins and destinations on a directional basis (i.e., one-way trip information), the information in the diagonal of the trip table probably stems from "reporting errors." However, the diagonal elements account for only about 6 percent of the total vehicle trips (Table 5) and the resulting error is not considered to be substantial. Any bias resulting from the nonzero values in the diagonal would be in the form of slightly over-estimating long trips. This possible over-estimation of long trips could slightly increase the attractiveness of an alternate route in the corridor.

STUDY COSTS

A summary of estimated study costs is presented in Table 6, which shows the total study cost was $87,500, or approximately $3 per response. The data-collection phase of the study was the most costly, accounting for 57 percent of the total cost and nearly 50 percent of the person-hours expended. The data-collection costs are based on the following field crew manpower requirements.

- *Interstate Highway Survey Crews*
  - 3 Traffic data recorders (volume, classification, occupancy counts, and license plate readings)
  - 5 Survey form distributors
  - 2 Police officers
  - 1 Crew chief

  11 Persons/Crew/Survey Station

- *Non-Interstate Highway Survey Crews*
  - 2 Traffic data recorders
  - 3 Survey form distributors
  - 2 Police officers
  - 1 Crew chief

  8 Persons/Crew/Survey Station

The traffic recorder and questionnaire distribution personnel estimates each include provisions for one substitute crew member for use when rests become necessary, or in case of an emergency. The data collection costs do not include costs incurred by the SDHPT in implementing the survey station traffic control plans (Table 6).

Student workers were used extensively in the data-collection and keypunching phases of the study. As a result, the

TABLE 5 ESTIMATED 1987 VEHICLE TRIPS BY MAJOR O-D ZONES (7:00 A.M.–8:00 P.M.): ALL VEHICLES

Origins	Destinations						
	M1	M2	M3	M4	M5	M6	Total
M1	586 44.9 0.6	8686 152.6 9.1	10768 138.6 11.3	698 48.9 0.7	76 16.4 0.1	3483 107.7 3.7	24297 242.4 25.5
M2	8867 153.6 9.3	2304 90.0 2.4	5611 120.6 5.9	3774 82.8 4.0	767 51.3 0.8	5554 105.9 5.8	26877 258.7 28.2
M3	11448 136.9 12.0	5630 117.9 5.9	2356 88.6 2.5	1362 60.3 1.4	266 30.4 0.3	760 53.0 0.8	21822 218.8 22.9
M4	671 46.0 0.7	3625 80.3 3.8	1426 58.3 1.5	567 43.3 0.6	51 13.4 0.1	939 57.8 1.0	7278 131.8 7.6
M5	81 16.9 0.1	678 48.8 0.7	229 28.9 0.2	25 9.6 0.0	27 10.1 0.0	963 59.5 1.0	2002 85.1 2.1
M6	3881 115.1 4.1	58855 108.7 6.2	956 60.6 1.0	1082 63.6 1.1	974 60.2 1.0	189 27.6 0.2	12967 192.8 13.6
TOTAL	25533 245.0 26.8	26808 256.8 28.1	21346 222.5 22.4	7509 137.5 7.9	2161 87.9 2.3	11887 182.4 12.5	95243 485.1 100.0

Major Interchange Zones (See Figure 1):
M1 = San Antonio
M2 = Austin
M3 = New Braunfels/San Marcos
M4 = Seguin/Lockhart
M5 = South of San Antonio
M6 = North of Austin

Legend: XXX = Vehicle Trips
XX.X = Standard Error (vehicles)
XX.X = Cell Percent

TABLE 6 STUDY COSTS

Study Phase	Person Hours	Cost[a]
Design/Planning	440	$12,000
Data Collection	2200	50,000[b]
Data Processing		
"Keypunching"[c]	1280	17,000[d]
Editing	160	1,300
Data Analysis[e]	250	3,200
Documentation	200	4,000
Total	4530	$87,500
Cost/Response[f]	0.16	$3.12

[a]Costs include travel-related expenses and fringe benefits.
[b]Includes $14,000 in questionnaire printing and postage expenses. Does not include costs incurred by SDHPT for implementation of traffic control plans.
[c]Data entry for questionnaires, manual volume/classification count data, and license plate data.
[d]Includes $3000 for computer time.
[e]Includes costs incurred in checking the representativeness of the sample.
[f]Based on 28,064 responses (see Table 4).

costs shown in Table 6 for these activities should be viewed as conservative estimates.

CONCLUSIONS

The results of this study indicate that the postcard O-D survey method can be used to develop representative, reliable estimates of intercity travel patterns. The results show that with a good traffic control plan and a well-trained survey crew, this survey method can be safely implemented without causing any substantial traffic delays, even when conducted on high-volume intercity freeways.

ACKNOWLEDGMENTS

This paper is based on research conducted as part of a project entitled "Austin/San Antonio Origin-Destination Study Concepts," sponsored by the Texas SDHPT. The views, interpretations, and conclusions expressed in this paper are those of the authors, and not necessarily those of the Texas SDHPT. The contributions of Danny Morris, Gene Hawkins, and Wanda Hinshaw of the Texas Transportation Institute were of great value to this effort.

REFERENCES

1. R. W. Stokes and T. Chira-Chavala. *Study Design for Austin/San Antonio Origin-Destination Study*. Research Report 1186-1 (Draft), Texas Transportation Institute, College Station, Tex., March 1987.
2. R. W. Stokes, D. E. Morris, G. Hawkins, and W. M. Hinshaw. *Austin/San Antonio Origin-Destination Study, Vol. I: Summary Report*. Research Report 1186-2 (I), Texas Transportation Institute, College Station, Tex., Feb. 1988.
3. R. W. Stokes, D. E. Morris, G. Hawkins, and W. M. Hinshaw. *Austin/San Antonio Origin-Destination Study, Vol. II: Technical Appendices*. Research Report 1186-2 (II), Texas Transportation Institute, College Station, Tex., April 1988.
4. J. M. Mounce. *Southwest Freeway (US-59S) Origin-Destination Study, Houston, Texas*. Texas Transportation Institute, College Station, Tex., Jan. 1981.
5. R. W. Stokes, D. E. Morris, and J. M. Mounce. *Katy Freeway (I-10W) Origin-Destination Study, Houston, Texas*. Texas Transportation Institute, College Station, Tex., Jan. 1982.
6. A. J. Ballard, D. E. Morris, and S. M. Smith. *Loop 410 (I-410) Origin-Destination Study, San Antonio, Texas*. Texas Transportation Institute, College Station, Tex., Feb. 1985.
7. FHWA. *Urban O-D Surveys*. U.S. Department of Transportation, Washington D.C., July 1975.
8. M. D. Harmelink. *A Study of Rural Roadside Interview Sampling Techniques*. Report RR 142, Ontario Ministry of Transportation and Communication, Downsview, Ontario, Feb. 1969.
9. J. J. Hajek. *Optimal Sample Size of Roadside Interview O-D Surveys*. Ontario Ministry of Transportation and Communication, Downsview, Ontario, Jan. 1977.
10. I. Miller et al. Sampling Methods for Roadside Interviewing. In *Highway Research Board Bulletin 76*, Washington, D.C., (1953) pp. 31–52.
11. M. Kuwahara and E. C. Sullivan. Estimating O-D Matrices from Roadside Survey Data. *Transportation Research*, Vol. 21B, No. 3, 1987, pp. 233–248.
12. District Highway Maps, Texas State Department of Highways and Public Transportation, Austin, Tex.

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

Optimal Design of Traffic Counts

Yehuda J. Gur and Irit Hocherman

The central traffic counting program in Israel covers the major rural highway network. Week-long mechanical counts are performed on each link every 1 to 2 years. This paper describes a study of possible ways to increase the program's efficiency by counting only a sample of links every year. Two possible methods for updating link volume estimates based on a sample of counts are described: (a) using network connectivity information and (b) using older counts for estimating the rate of change of volumes over time. The methods have been evaluated using actual count data.

In almost every country there is an agency that operates a system for the collection and processing of traffic-count data on a permanent basis. The main purpose of such systems is to provide an overall, reliable picture of the patterns of motor-vehicle travel and traffic load on the road network, and their changes in time.

There are two main uses to traffic-count data:

- On the demand side, they provide information for the description of the amount of vehicular travel, including growth trends and distribution by type of road and location. These data are used for transportation policy decisions and strategic planning. Traffic counts may also assist in the estimation of origin-destination tables.
- From the supply side, estimates of volumes on specific road sections serve as a basis for maintenance planning, for measuring exposure (in accident analysis), and sometimes as input for detailed planning.

The permanent count program also provides adjustment factors that are used for the conversion of traffic counts into volume estimates. Examples include such factors as corrections for seasonal effects, and the conversion of average annual daily volume (AADT) into design hourly volume (DHV).

In Israel, the Central Bureau of Statistics (CBS) performs the systematic traffic counts on the rural road network. About 800 road sections are counted for 1 week every 1 or 2 years by mechanical counters. Some manual classified counts are also performed to determine the percent of trucks in traffic. In addition, three fixed counting points are located on major arterial roads. The counting program is run quite efficiently by two full-time field crews and a small office staff.

Urban road sections are not included in the program. Various local counting programs are maintained by individual municipalities, but they do not provide a satisfactory basis for a comprehensive picture of traffic trends.

The use of the CBS count data is limited to the supply aspects. The demand side uses are hampered by the absence of counts on urban roads. There is no information on the split of travel between urban and rural roads in the different regions and the way it changes with time. Thus it is not possible to estimate the total amount of travel in the system or the geographic distribution of traffic loads. Moreover, the definition of a road as urban or interurban is sometimes arbitrary and certainly subject to change over time.

In a separate study, Gur et al. (*1*) investigated the feasibility of performing traffic counts on the urban network, and concluded that it was technically feasible. The study recommended that urban arterial roads be incorporated into the CBS counting program. It also recommended that the additional road sections would be added to the system without an increase in the permanent traffic counting staff, that is, without a significant change in the number of road sections that are counted every year.

The purpose of the study that is described here is to examine ways to incorporate the arterial urban road network in the count scheme, assuming that the total number of sections counted is fixed and given. The basic strategy used is to reduce the frequency of counting by sampling the sections to be counted every year and by devising a method for estimating the traffic volumes on those sections that were not counted in a specific year.

On the basis of the findings reported by Gur et al. (*1*), it can be assumed that in Israel there are no systematic seasonality effects. That is, the weekly traffic counts are random drawings from a distribution with expectation $VOL(l,n)$, where VOL is the average annual weekly volume of section l during year n.

The major thrust of this work is to develop a method for estimating AADT based on sample counts and an optimal sampling scheme. In other words, we search for an efficient method for estimating the current traffic volume in a specific road section, given its location and physical characteristics, its own traffic counts in previous years, and current traffic counts of other sections.

THE DATA

Three data files from three different sources were used for the study:

- Traffic count data base constructed by the CBS, which included all the counts for the years 1983 to 1986.
- The physical characteristics file, which was constructed as part of the study on the basis of information from the Department of Public Works.

Transportation Research Institute, Technion–Israel Institute of Technology, Haifa, Israel.

- The interurban network file, which described the statewide arterial roads as a connected network, compiled by the Israel Institute for Transportation Planning.

The three files were edited for errors and inconsistencies and concatenated. The resulting file included the following information for each road section: The identification numbers of the section and its two nodes, average weekday traffic volume in each of the years 1982 to 1985 where the section had been counted, and various attributes such as class of road, region, number of lanes, length, and a dichotomous variable denoting proximity to a metropolitan area. This data file was used to develop and test the various methods that were examined in the project.

ESTIMATION OF TRAFFIC VOLUMES BASED ON ROAD CHARACTERISTICS

We have examined the link volume data for significant trends. Most striking is the association between a link's volume and its type of road. There is also an apparent association between the volume and the region. We found no apparent systematic time trend between years, seasons, or months.

Our first attempt at estimating the traffic volume of a section is based on the mentioned associations by a regression model of the annual weekday traffic volumes (AWDT) as a function of the various sections' attributes. The independent attributes are

- Class of road (major arterial, minor arterial, collector, or local*)
- Region (north*, Haifa, center, Jerusalem, south, or Golan)
- Number of lanes (1 or 2)
- Area (1, rural; 2, metropolitan).

Each nonstarred value of the first two attributes is represented by a dummy variable. For example, the dummy variable for the value "major arterial" is 1 if the section belongs to a major arterial road, and 0 otherwise, and it measures the additional volume that is ascribed to the section being a major arterial, relative to a local road. This regression model is basically identical to an analysis of variance (ANOVA) model.

A separate model was calibrated for each of the years 1982 to 1985; the model for 1985 is displayed in Table 1. The coefficients of the models are quite similar for all years; the proportion of variance explained, R^2, ranges between 70 and 74 percent, which may be considered a relatively high explanatory power. At the same time, the models' estimation errors are rather high, with a root mean square error (RMSE) of around 6500 vehicles per day. Such large errors make the model estimates unacceptable, particularly for estimating traffic volumes of individual road sections.

ESTIMATION OF TRAFFIC VOLUMES BASED ON PAST VALUES

Our next attempt at estimating the traffic volume of a section uses the count information on the same section from the prior year. The regression equation for 1985 is as follows:

$$U_n = -0.174 + 1.075 * C_{n-1}$$

where U is the traffic volume for year n and C is the number of cars counted ($R^2 = 0.97$; RMSE = 2.08).

This model is much better than the one based on the physical characteristics; it has a much higher explanatory power and the RMSE is much smaller. Ignoring the small constant, the model suggests that 1985 link volumes can be estimated by assuming a uniform 7.5 percent growth in 1984 volumes.

The good fit of the model (or the high correlation coefficient) may result from the huge dispersion of counts in different types of road sections. To test this we have calculated separate correlation coefficients for each combination of region and type of road. All of the correlation coefficients were greater than 0.9, showing that the correlation between counts of the same link in two consecutive years is indeed high.

Correlation coefficients of similar magnitudes are observed also between the counts of 1983 and 1984, and even with a 2-year gap, between 1983 and 1985.

This finding is meaningful. It suggests that a promising method for estimating current link volumes should be based on updating older counts. However, the model's coefficient (i.e., the annual growth rate) might differ between years and must be estimated anew for every year. This issue is discussed in the following sections.

TABLE 1 REGRESSION MODEL FOR AVERAGE TRAFFIC VOLUME IN A ROAD SECTION BY ITS PHYSICAL CHARACTERISTICS, FOR 1985 (1,000 VEH.)

Variable	Beta	S.D.	T-Value	P-Value
Constant	-24.03	1.67	-14.4	0.0001
Area	10.52	1.01	10.4	0.0001
No of Lanes	13.57	1.38	9.8	0.0001
Type of Road:				
Major Arterial	11.95	1.32	9.0	0.0001
Minor Arterial	4.99	0.79	6.4	0.0001
Collector	2.42	0.84	2.9	0.004
Region:				
Haifa	-0.30	0.98	-0.3	0.76
Center	4.16	0.89	4.7	0.001
Jerusalem	-2.52	1.33	-1.9	0.059
South	-1.35	0.94	-1.4	0.155
Golan	-2.68	1.30	-2.1	0.040

$R2 = 0.70$; $RMSE = 6.5$;

CONNECTIVITY INFORMATION AS AN AID IN UPDATING VOLUME ESTIMATES

The Problem

Standard statistical methods treat links in the system as independent members of a population. In reality, links are connected into a network, and many trips use not one, but a chain of connected links. Hence, it is likely that volumes are correlated—a volume change in one link is associated with similar changes in related links. We have tried to formulate a procedure that uses network connectivity information to improve the accuracy of link volume estimates. Assume that outdated (say, 1-year-old) volume estimates on all links in the network are given, as well as updated volume estimates (say, based on counts) for some of the links. Our problem is to provide estimates of the updated volumes on all links.

At the outset, it is tempting to try a comprehensive approach—the network can be described as a connected graph, and volumes can be estimated using assignment-type network analysis models. For example, procedures for estimating trip tables based on link volumes (*2*) can use volume estimates on some links to produce consistent flow estimates on other links. However, in closer examination it becomes clear that this approach is not compatible with the problem at hand; such an approach is dependent on information about land use and travel patterns. It might be suitable for planning, but not for providing traffic volume data, which is considered basic data and must be mostly self contained.

Principles of the Approach

A method for updating link volume estimates that is aided by connectivity information has been developed. Basic definitions are presented in Figure 1. Each link is treated separately; consider the analysis of link O, which we will term the link of interest (LOI). Each one of its nodes (1 and 2 in Figure 1) connect to a number of link chains; for example, node 1 connects to chains I and II.

Each chain consists of links that belong to the same route (as defined by the highway authority). Advancing along a chain for a set number of links, we might hit a link with an updated count; for it, we know the relative and absolute volume changes. This information may be used to estimate the volume's change in link O. The process is repeated independently for each of the chains connecting to the two nodes.

We have examined a wide range of procedures for weighting and summing the data on volume changes on each of the chains, in order to get an estimate of the expected changes on the LOI. Factors that we have accounted for include:

- Whether a chain is along the same route as the LOI;
- The distance (measured by the number of intersections) between the counted link and the LOI;
- The volumes on the links along each chain and their stability; and
- The number of chains with and without counts.

It is beyond the scope of this paper to describe in full the various methods.

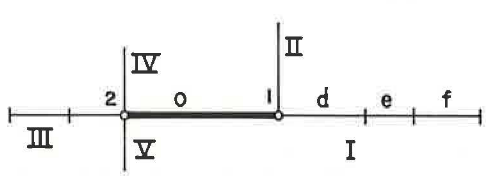


FIGURE 1 Elements of the model.

Testing

The most promising versions of the model are evaluated using the CBS data. The model is used to estimate 1985 link volumes based on 1984 counts for most links and on 1985 counts for a random sample of links. A range of sampling rates is tested. The model's estimates are compared to the available 1985 counts.

The model's estimates of 1985 volumes are also compared to estimates that are obtained by factoring the 1984 volumes by estimated aggregate growth rates in the amount of vehicle miles traveled (VMT). The following growth factors are used:

- FT: A uniform factor for the whole system,
- FC: Different factors by class of road,
- FD: Different factors by district, and
- FCD: Factors by class and district.

Typical results are presented in Table 2. Trends, which have appeared repeatedly in the different experiments, include:

- The model's performance is not superior, and in most cases mostly inferior, to simple factoring.
- Factoring using one universal factor for the whole system has performed best in most cases.
- The worst results have been obtained by use of detailed factoring by both route type and district.

In detailed examination, we have found that the model behaves as expected; that is, the estimated volume change is a reasonable resultant of the changes in the neighboring links. However, the actual counts do not behave this way—large and inconsistent variations in the extent and even direction of the volume changes are apparent.

Conclusions

The attempt to formulate a method for improving the estimates of link volume changes, by using connectivity information has not succeeded—simple estimates of the change, using a gross, system-wide growth factor, prove superior to the estimates of a rather involved model.

The findings in this section reinforce our previous findings, namely that a system-wide uniform growth factor is a powerful and promising means to update link volume estimates.

The methods for the inclusion of the connectivity information are somewhat arbitrary; possibly, there can be found more successful methods. However, considering the fact that we have tried a rather wide range of versions, revolutionary improvements seem to us unlikely.

TABLE 2 NETWORK-BASED ESTIMATION—TEST RESULTS

Nominal		Estimate's RMS Error (K veh.)				
Rate (%)	Counts	Model	FT	FC	FD	FCD
20	49	3.12	2.98	3.52	3.02	3.45
50	126	3.24	2.76	2.85	2.75	2.96
80	201	2.34	1.76	1.87	1.79	2.15

It is possible to speculate on possible reasons for these results:

- In a rather small country like Israel, factors that cause system-wide changes in travel are dominant, compared to local variations. Such factors might include vehicle fleet growth, the economic situation, or amount of tourism.
- Because of the large data errors caused by technical problems, as well as the natural variability in weekly volumes during a year, the counts provide only a rough estimate of the AWDT, with a large random error. It is possible that the benefits incurred from using a large body of data to obtain a reliable estimate of one system-wide growth factor more than compensate for ignoring rather minor detailed trends within the network.

It will be more instructional to refer to the experiment described here as another example that shows that elaboration does not necessarily lead to added accuracy. In some cases, the tendency to complicate models by adding detail (and stratifications) does not improve, and might harm, the model's performance. Even if there is a "behavioral" justification for adding details, their ultimate inclusion must be decided according to the model's performance. The elaboration and precision of a model should be tempered by our limited ability to define and measure relevant relationships.

SAMPLING

The Problem

On the basis of previous findings, we know that the use of one universal systemwide change (growth) factor provides a promising way for updating link volumes over time. We also know that there is a high correlation between counts of the same link in different years. Considering these findings, we examine possible ways to design a traffic counting program that is based on an annual count of a sample of links and uses a system-wide expansion factor.

There are two interrelated design problems: How to estimate the annual volume change and how to select the sample of links each year. In selecting the sample, it is desirable to count each link as frequently as possible; more recent counts provide a better base for volume updates. Frequent counts also correct for irregular, large volume changes that may occur in individual links (e.g., because of increased local activity). In order to achieve the highest possible count frequency for all links, it is necessary to change the links in the sample every year.

The procedure should also provide as reliable an estimate as possible of the annual growth rate. The growth rate can best be estimated by counting the same links in consecutive years. In such a procedure, the estimation error of the growth factor will not be affected by the random variations among the links in the different samples; that is, for efficient estimation of the annual growth rate, it is best to count the same links in consecutive years. (We have not examined other promising methods for estimating aggregate growth, such as using fuel sales or similar external data.)

In this section we examine different ways for estimating the growth factor. If we find that it is indeed necessary to rely on counts of the same links in consecutive years, we will have to find a way to select an optimal rate of replacement of links in the sample. For simplicity of the argument, and considering the counting techniques in Israel, we (a) ignore the existence and role of permanent counting stations, (b) assume a fixed budget and a uniform cost per count, and (c) examine annual sampling rates in the range of 30 to 70 percent.

Link Volume Updating

We consider estimation methods where the estimates of individual link volumes are consistent with the estimates of the amount of travel. The population of links in the system is well defined and known. Volume estimates are to be provided annually and simultaneously for all the links in the system. The proposed estimation procedure is applied as follows.

At the start of any year n, there exists, for each link l in the system, a volume estimate for the last year, $U_{n-1,l}$. During the year a subset of links, C_n is counted; the counts provide current volume estimates, $C_{n,l}$ for these links. On the basis of available information, U and C, a growth factor for the whole system F_n, is calculated. We discuss later how F_n is estimated. Volume estimates for each link are calculated as follows:

$$U_{n,l} = \begin{cases} C_{n,l} & \text{if } l \in C_n \\ U_{n-1,l} * F_n & \text{if } l \notin C_n \end{cases} \tag{1}$$

This method provides volume estimates $U_{n,*}$ for all the links in the system. These estimates are used as a base for estimating the volumes for year $n + 1$, and so on. They also provide the information for calculating the amount of travel in any subset of links by:

$$VMT_{n,S} = \sum_{l \in S} U_{n,l} * L_l \tag{2}$$

where L_l is the length of link l, and S is any selected subset of links.

Estimating the Growth Factor, F_n

The growth factor, F_n, is defined as the annual rate of change in the amount of travel:

$$F_n = VMT_n/VMT_{n-1}$$

or

$$F_n = \sum_l U_{n,l} \Big/ \sum_l U_{n-1,l} \tag{3}$$

Three different methods for estimating F_n have been examined:

1. On the basis of independent VMT estimations. With the counts at year n, the total system's VMT can be estimated by

$$VMT_n = \frac{1}{W_n} * \sum_{l \in C_n} (C_{n,l} * L_l) \tag{4}$$

where

$$W_n = \sum_{l \in C} L_l \Big/ \sum_l L_l \qquad (5)$$

W_n is the fraction (by length) of the counted links in year n out of the total link population. From here the estimate of the growth rate by the first method, F_n^a is obtained by Equation 3. In this method, different links are used to estimate VMT_n and VMT_{n-1}. Also, the method is consistent; the VMT estimates using Equations 2 and 4 are identical.

2. On the basis of all $l \in C$. Here, the growth rate is estimated by comparing the current year counts to the previous year volume estimates, for all the counted links:

$$F_n^b = \frac{W_{n-1}}{W_n} * \frac{\sum_{l \in C_n} C_{n,l} * L_l}{\sum_{l \in C_n} U_{n-1,l} * L_l} \qquad (6)$$

3. On the basis of only links that are counted in both years. This method is very similar to the previous one, in that F_n^c is estimated by

$$F_n^c = \frac{\sum_{l \in C_{n,n-1}} C_{n,l} * L_l}{\sum_{l \in C_{n,n-1}} C_{n-1,l} * L_l} \qquad (7)$$

where $C_{n,n-1}$ contains only links that are counted in both years.

Comparative Evaluation of the Three Estimation Methods

Qualitatively, the three methods differ in the amount and type of the information they use. The first method uses all counts made in the last 2 years. VMT_n and VMT_{n-1} are estimated independently, based on two different sets of links.

The second and third methods provide a direct estimate of F_n. Both use the volumes in consecutive years of one set of links for estimating the rate of volume change, thus exploiting the correlation between consecutive counts of the same link. The second method uses all the counts of year n. It uses the information from counts in earlier years embedded in the estimated volumes U_{n-1}. However, it uses only indirectly the counts in year $n - 1$.

The third method is similar to the second, but it uses only a subset of the links, those that were counted on both years; it also does not take advantage of counts of earlier years. Compared to the second method, its major advantage is that it relies only on actual counts, rather than on volume estimates that are subject to estimation errors.

The three methods differ also in their use of the information. The third method provides direct estimate of the volume growth rate, by examining the change in individual links. In the first method, growth is estimated indirectly, without taking advantage of the correlation in volumes of the same links.

Cochran (*3*) discusses these two estimation methods for the case where the sample sizes in both cases are equal; he shows that the third method is preferable when:

$$R_{n,n-1} > 0.5 * F_n \qquad (8)$$

where $R_{n,n-1}$ is the correlation in volumes in the 2 years. In our case, the correlation coefficients are much higher. At the same time, the sample size in the third method is likely to be much smaller. It is equal to the size in the first method only if the same links are counted in both years—an unfeasible case.

The major deficiency in the second method is that it does not consider the quality of the estimate, $U_{l,n-1}$; in particular, the age of the last count. However, if links are counted in a reasonable frequency, and care is taken to count a link whenever a major change in its volume is suspected, then it is likely that the added information will make this method superior in many instances.

The three methods were compared quantitatively, using the count data for the years 1983 to 1985. The 1985 volumes were estimated based on the 1983 volumes for all links, and random samples of counts for 1984 and 1985. Sampling rates of 0.3, 0.5, and 0.7 were tested. Estimated growth rates by the three methods were compared to the 1985 counts. The procedure was repeated a few times, in order to control random effects. Typical results are summarized in Table 3.

The test results show

- The three methods perform well in the high sampling rates.
- The first method is inferior in the medium and low sampling rates.
- The second method, which uses all the link counts, performs best in all sampling rates.
- The third method performs surprisingly well, considering the small number of links in the sample with counts in consecutive years with only about 20 in the 30 percent samples.

Similar results have been obtained in comparing the standard estimation errors of individual links.

Conclusion

Both the qualitative and quantitative analysis indicate that the most effective method for estimating the annual rate of change is by estimating the average VMT change in all the counted links (Equation 6). The method provides satisfactory estimates even at a 30 percent sampling rate. The method was tested using data for two consecutive annual count samples. In actual applications, the expected gap between counts is somewhat longer. However, this difference is not likely to have a critical effect.

For the problem at hand there is no particular need to count the same links in consecutive years. This result is valid only

TABLE 3 ESTIMATED VMT CHANGE RATE (1983–1985)

sampling rate method	30%	50%	70%	true
1	.99- 1.09	1.18-1.28	1.08-1.12	1.13
2	1.09-1.10	1.06-1.11	1.09-1.12	
3	1.06	1.08-1.12	1.08-1.09	

for relatively high sampling rates where each link is counted every 2 to 4 years.

The qualitative analysis has indicated that links should be counted whenever an irregular change in their volumes is suspected.

SUMMARY AND CONCLUSIONS

We have examined a number of issues that are related to the use of samples of links in an on-going traffic counting program. We have found trends in link volumes, and in particular a high correlation of volumes on a link between different years. By exploiting these trends, it is possible to get reliable estimates of the link volumes for years when they are not counted.

On the basis of the findings of the study, a gross annual sampling rate of about 40 percent is recommended, of which about a quarter of the count is devoted to special cases, such as new roads and areas where significant changes are suspected.

REFERENCES

1. Y. J. Gur, A. S. Hakkert, and I. Hocherman. *Traffic Counting in Israel—Analysis and Evaluation*. Publication no. 185-85. Transportation Research Institute, Technion, Israel, 1985.
2. M. Turnquist and Y. J. Gur. Estimating Link Volumes Based on Link Volume Counts. *Transportation Research Record 750*, TRB, National Research Council, Washington, D.C., 1979.
3. W. G. Cochran. *Sampling Techniques*, John Wiley and Sons, Inc., New York, 1977, Chapter 8.

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

On-Line Generation of Synthetic Origin-Destination Counts for Application in Freeway Corridor Traffic Control

M. VAN AERDE, J. VOSS, AND G. NOXON

A need exists during the application of freeway corridor control models to determine the prevailing origin-destination (O-D) matrices for each time slice during the peak period to be analyzed. As it is virtually impossible to obtain these matrices directly by survey, a fully automated approach is proposed that would use Freeway Traffic Management System (FTMS) data already being collected. This approach relies on existing algorithms for formulating synthetic O-D counts from observed link flows, but uses a special relationship that exists between O-D matrices for consecutive time slices to carry out these computations more efficiently and often also with greater accuracy. The general background to the problem and the general solution approach that has been proposed are discussed. Subsequently, several different analysis runs using the proposed approach are described that were performed with data for the Burlington Skyway FTMS system in Ontario. The results of these runs illustrate the details of the technique and demonstrate the main reasons for the improved efficiencies and accuracy. The paper is concluded with a discussion of how the procedure can be further refined and implemented in both its off-line and on-line modes within existing FTMS installations.

During the past decade a number of techniques were developed for estimating synthetic origin-destination (O-D) demands from observed link flow counts. Such techniques proved to be efficient and cost-effective in generating the demand data required for transportation planning studies, when either direct survey methods were impractical or too expensive. In freeway corridor problems, all assignment-based control models require that the traffic demands are also expressed as O-D flow rates for the freeway corridor (*1*). However, because of the operational rather than planning character of the analysis, a sequence of O-D matrices is required to express the changes in traffic demand during the peak period that is analyzed. Consequently, a number of O-D matrices must be derived, rather than just one single matrix.

Such a sequence of O-D data is difficult and expensive to obtain for use with off-line simulation models. Furthermore, at present it is virtually impossible to obtain such O-Ds on-line for use with real-time traffic control or diversion models. We propose a technique that can efficiently generate this sequence of O-D matrices on-line using real-time data. The technique is based on a special relationship that exists between O-D matrices for consecutive time slices. The objective of this paper is to show the feasibility of the technique, to illustrate its results and limitations, and to outline how the technique could be used in practice.

Department of Civil Engineering, Queen's University, Kingston, Ontario, Canada, K7L-3N6.

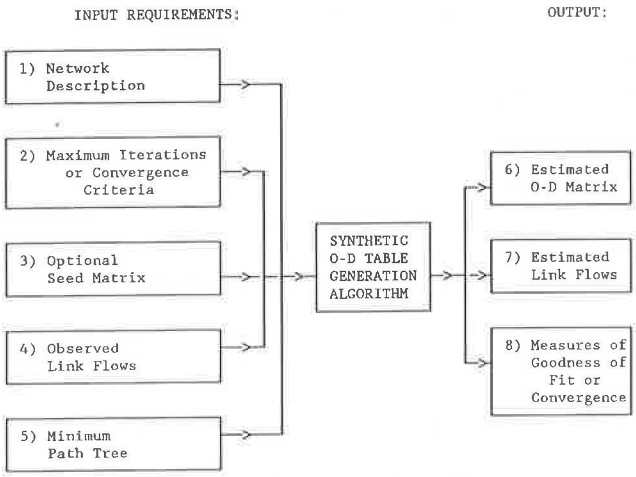
PROCEDURE FOR GENERATING AN O-D MATRIX FOR ONE TIME SLICE

The procedure for generating synthetic O-Ds was developed based on an existing algorithm by Van Zuylen and Willumsen (*2*).

Synthetic O-Ds in Transportation Planning

Many techniques exist for developing synthetic O-D data from link flows. Examples of these, which have been applied in a transportation planning context, include a generalized least-squares estimator (*3*), Bayesian statistical approach (*4*), constrained regression (*5*) and information minimization–entropy maximization (*2*,*6*).

The general procedure involved in applying these methods is illustrated in Figure 1. As shown, the inputs to the analysis consist of a network description file, a set of convergence criteria, a series of minimum path trees, a list of observed link flows, and an optional seed matrix. Within the analysis, the minimum path trees are used to determine which O-D pairs contribute to which link flows, and with the simpler algorithms only one path is allowed between each O-D pair. As there are many more variables than constraints to this problem, there are numerous different mathematical solutions possible. The derivation of a mathematical solution that closely matches the "correct" matrix is therefore assisted considerably if the algorithm is provided with a priori knowledge of the general travel pattern structure in the form of a seed matrix. This seed matrix is used to initiate the solution search and it reduces the number of required algorithm iterations. As it will also impart its general structure onto the final solution matrix, this seed matrix consistently results in a much improved final O-D matrix estimate. The quality of the generated matrix is ideally determined by measuring the deviation between the predicted O-D cell values and the actual O-D counts. However, as the technique is intended to be used



1. Network Description:	link/node structure of the network
2. Convergence Criteria:	a criterion used to terminate the algorithm
3. Optional Seed Matrix:	old or approximate matrix to initiate search
4. Observed Link Flows:	measured link traffic flows
5. Minimum Path Tree:	listing of links on most efficient routes
6. Estimated O-D Matrix:	O-D flows estimated by algorithm
7. Estimate Link Flows:	link flow counts estimated by algorithm
8. Measures of Convergence:	indicators of the algorithm's convergence

FIGURE 1 Basic procedure for generating synthetic O-Ds.

when actual O-D counts are not available, the next-best quality indicator is the ability of the matrix to reproduce the original link flows.

Selection of a Suitable Method to Generate an O-D Matrix

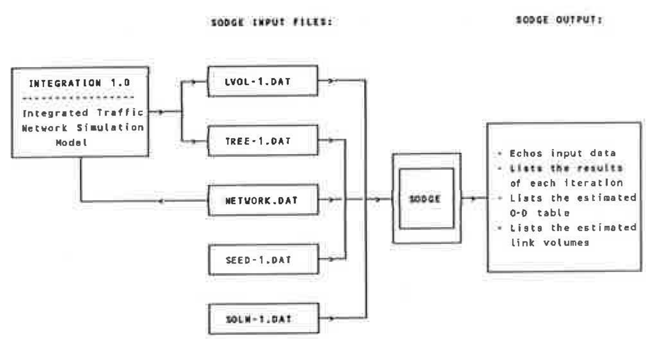
Of the available synthetic O-D techniques, the information minimization–entropy maximization algorithm by Van Zuylen and Willumsen (*2*), revised by Van Zuylen (*6*), was considered most suitable for the intended use. The principles and steps of this algorithm are well documented, and a version of the algorithm implemented in the ME2 model has been satisfactorily validated by Van Vliet and Willumsen (*7*).

The above algorithm determines the most likely O-D matrix by solving for the matrix that minimizes the information contained in the final O-D matrix. The actual solution algorithm is derived by formulating a linear equation that considers each link flow to be a result of a series of trips between all O-D pairs that have routes utilizing that particular link. For O-D cells that use multiple paths, the appropriate proportions utilizing the link along each path are expressed as probabilistic fractions. In the case of an "all-or-nothing" assignment, these probabilities end up being either zero or one, depending on whether a given link was used or not. The entire problem formulation is then a series of linear equations, including an objective function that maximizes the entropy measure of the trip matrix and a number of constraints arising from the observed link volumes.

Computer Implementation of the Selected Algorithm

The revised information-minimization algorithm (*6*) was implemented as a new computer program by Noxon (*8*) to allow model inputs that are compatible with the INTEGRATION (*9*) simulation model. The resulting program, called SODGE (synthetic origin-destination generator), requires three essential input files. The first is a network description file, which lists the network links. The second file contains the link traffic flows, whereas the third contains a minimum path tree matrix for the given network conditions. Two other input files are optional, namely a seed O-D matrix and the actual O-D matrix, if available. The former assists in initiating the search among the range of feasible solutions, whereas the latter, if known, allows the user to check the accuracy with which a true matrix can be recreated.

As SODGE was developed to provide O-D counts to the INTEGRATION simulation model, the network description file for both models was formulated for dual compatibility. In addition, the SODGE link volume and minimum path tree input files were configured such that they could automatically be generated using INTEGRATION. Consequently, a given network could be analyzed using INTEGRATION to determine minimum path trees and link flows, and with these SODGE could be run to retroactively determine the most likely O-D matrix governing the network's operation (Figure 2). This procedure was first used to determine the reliability and accuracy with which SODGE could reproduce a known matrix, supplied only with link flows and minimum path trees. However, in practice SODGE would be used to generate an estimate of the unknown O-D matrix for use within the INTEGRATION model, which in turn would evaluate different network control strategies.



LVOL-1.DAT:	Contains the link volumes generated by INTEGRATION 1.0 for the time interval being analyzed.
TREE-1.DAT:	Contains the minimum path trees generated by INTEGRATION 1.0 for the time interval being analyzed.
NETWORK.DAT:	Describes the network used both by SODGE and INTEGRATION 1.0.
SEED-1.DAT:	Contains the seed matrix (user-defined or default) for the time interval being analyzed.
SOLN-1.DAT:	Contains the actual solution matrix (if known) for the time interval being analyzed.

FIGURE 2 Synthetic O-D generation procedure utilized by SODGE.

SODGE Measures of Convergence

To monitor and evaluate the convergence of the SODGE algorithm, the program calculates three statistical indicators at the conclusion of each iteration. The first measure is the root mean square difference (RMSD) between the cell values of successive iterated solution tables. This value decreases as successive iterations produce more similar solution tables such that deviations between cell values for successive iterations are minimized. The second measure is the root mean square error (RMSE) between the observed link flows and the link flows that are produced by feeding the iterated trip table back into the network. These two statistical measures are produced for each iteration in all cases. The third indicator is the RMSE (in percentage form) between the cell values of the current trip table estimate and those of the actual trip table, if the latter is provided by the user.

If no solution matrix is available, convergence to a solution may be indicated by a stabilization of the trip cell RMSD figure between consecutive iterations. However, the algorithm tends to give a series of RMSD stabilizations at different levels of actual convergence. It is thus better to judge convergence using the link flow RMSE, which will consistently stabilize at the optimal convergence. Convergence of link flows indicates that the matrix, although perhaps not the exact one, can reproduce the observed link flows at a desired level of accuracy. In practice, the actual matrix is of course always unknown, as it is the object of the search. However, during the testing of SODGE, the search for a known O-D matrix was performed to determine the likely range of errors and problems associated with searches in practice where the true matrix is unknown.

APPROACH FOR SEQUENTIAL GENERATION OF ON-LINE O-D COUNTS

The previous section indicated how the SODGE implementation of Van Zuylen's (*6*) synthetic O-D generation technique could be used to automatically interact with INTEGRATION to derive one O-D matrix at a time. In this section we discuss how the same procedure can be used to derive a series of consecutive O-D matrices for an entire peak period.

Methodology

As the traffic demands within a peak period are not necessarily uniform, no single O-D table can accurately represent the demand pattern over the whole period. Therefore, the entire peak period to be analyzed is broken down into a series of consecutive time slices, each time slice having its own separate O-D table. As a first step, one could generate the O-D matrices for an entire peak period by simply running SODGE for each time slice by itself, without accounting for any interactions.

However, as the generation of an O-D matrix from scratch involves a large number of iterations, and as the quality of an O-D matrix generated without some reliable a priori knowledge is usually not very high, further significant performance enhancements are possible. These involve the use of the previous time slice's O-D matrix as the seed for the derivation of the next time slice's matrix. This approach is illustrated in Figure 3 for a sample sequence of two consecutive time slices.

Consequences

If there is any relationship between the true O-D matrices of consecutive time slices, the O-D estimate of the first period will make a much better seed for the second stage than would a random or uniform seed. The first consequence of this is that fewer iterations would be required to estimate the next matrix. This is an important efficiency if these O-D estimates are to be made based on on-line traffic counts in real time. More importantly, if the O-D matrix estimated for the previous time slice was a good fit, its use as a seed should also considerably improve the accuracy of the O-D prediction for the next time slice.

If there is a consistent nontrivial relationship among a time-series of O-D matrices, this technique would efficiently retain the general structure of the O-D pattern over the entire period. However, it would also selectively scale the entire matrix or selective entries in the matrix, in view of any changes in observed link flow counts. The result would be more accurate on-line O-D estimates that are responsive to real-time traffic flow counts provided by FTMS detectors.

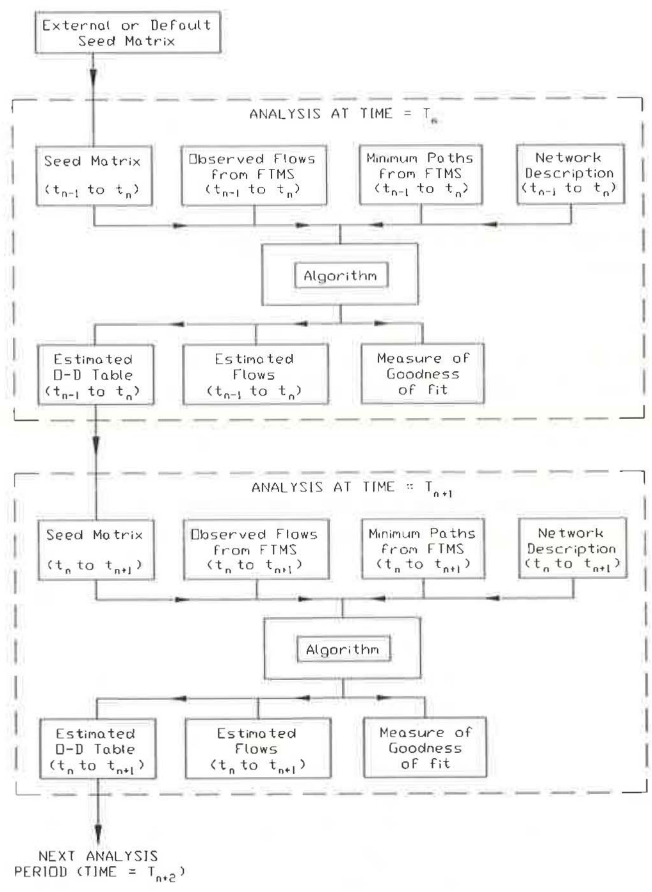


FIGURE 3 Proposed on-line synthetic O-D generation procedure.

In an off-line situation, the above features permit an efficient and economical estimation of a sequence of O-D matrices that reproduce the link flows that are observed during the peak period. In addition, when the technique is applied using on-line data, it will estimate in real time the unique O-D matrices for a particular day's peak period in view of the unique traffic flows that are observed on that particular day.

Description of the Test Network and its FTMS

To illustrate the potential of the on-line O-D generation technique, some sample test runs were performed using data for the Burlington Skyway FTMS on the Queen Elizabeth Way (QEW) between Toronto and Niagara Falls, Ontario. The general location of the Burlington FTMS system is shown in Figure 4.

The QEW is a major provincial highway between Toronto and Niagara Falls and cuts across the Hamilton Harbor at the west end of Lake Ontario. As the freeway crosses the Burlington Canal via the Burlington Skyway Bridge, its final configuration will provide three fully detectorized lanes in each direction. In addition, a four-lane arterial highway parallel to the QEW provides a second route, which acts as a diversion alternative in case of an incident on the bridge (*10*). This diversion route is signalized and fully detectorized.

At the time of this study, only the detectors for the southbound portion of the system were fully operational, as the northbound portion of the system was still under construction. Consequently, the test runs on the Burlington Skyway only considered the southbound traffic network and traffic demands.

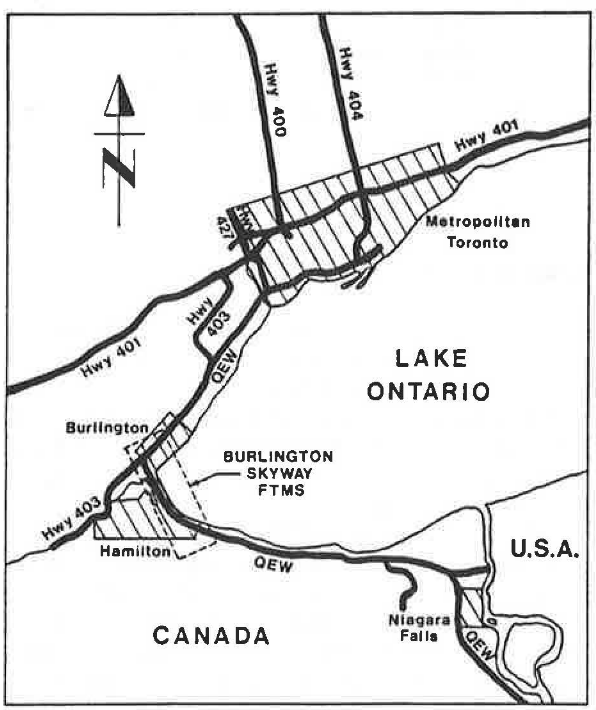


FIGURE 4 General location of the QEW and the Burlington Skyway FTMS.

The southbound Burlington Skyway network was coded and digitized in March, 1988, using 100 links and 73 nodes, of which 10 were also zone centroids (Figure 5).

Details

During the trial evaluation of this procedure, a number of details had to be resolved, such as the duration of the time slices and the updating rate of the minimum path trees.

A time slice duration of 15 min was considered to be short enough to allow the capturing of dynamic changes in the O-D pattern, while also providing sufficiently stable link flow counts to SODGE. The whole 2-hour peak period analysis was thus based on an eight-slice simulation of a typical peak period. The objective was to determine if the sequential estimation procedure could successfully back-calculate prevailing O-D patterns and their changes.

Because the current version of SODGE is set to evaluate flows for all-or-nothing route assignments, INTEGRATION was used to assign traffic flows to all-or-nothing assignments for 15 min at a time. In addition, because INTEGRATION starts its evaluations with networks that are initially empty, a state of equilibrium was allowed to develop during the first 15 min before any link summaries were computed. This time allowed all O-D patterns to fully propagate through the network, since the maximum trip length in the network was about 8 min. Finally, to provide an analysis of equilibrium conditions, all signal timings were held constant during the entire 2-hour peak period.

RESULTS OF ON-LINE O-D GENERATION TESTS

The potential of the proposed approach was assessed using a systematic evaluation of four sets of related experiments. A description of these experiments is provided below, and the results are summarized in Tables 1, 2, and 3.

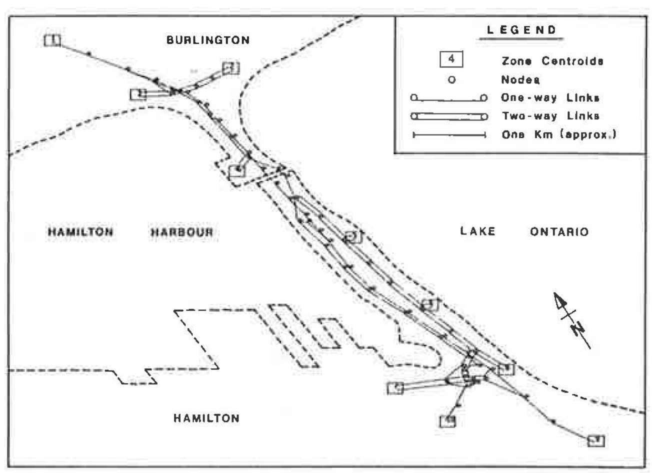


FIGURE 5 Network representation of southbound Burlington Skyway FTMS.

TABLE 1 CONSTANT O-D DEMAND PATTERNS FOR THE ENTIRE 2-HR PERIOD

	RUN A O-D Demand constant Initial seed = 100 Epsilon = 0.1			RUN B O-D Demand constant Init. seed = Act. soln. Epsilon = 0.1		
TIME PERIOD	# OF ITER	FLOW RMSE	OD TABLE RMSE %	# OF ITER	FLOW RMSE	OD TABLE RMSE %
1						
2	10	5.65	46.73	2	5.74	2.95
3	2	9.73	47.41	2	10.26	4.82
4	3	8.81	43.59	3	8.80	4.97
5	4	10.19	47.26	4	10.54	6.53
6	4	10.33	45.19	4	10.66	7.02
7	4	8.73	46.17	4	8.39	4.82
8	4	8.53	46.00	4	8.43	5.62

TABLE 2 INCREASED O-D DEMAND PATTERN FROM TIME 1.0 HRS TO 1.5 HRS

	RUN G O-D Demand varies Constant seed = 100 Epsilon = 0.1			RUN D O-D Demand varies Initial seed = 100 Epsilon = 0.1			RUN F O-D Demand varies Init. seed = Act. soln. Epsilon = 0.1			RUN H O-D Demand varies Each seed = Act. sol Epsilon = 0.1		
TIME PERIOD	# OF ITER	FLOW RMSE	OD TABLE RMSE %	# OF ITER	FLOW RMSE	OD TABLE RMSE %	# OF ITER	FLOW RMSE	OD TABLE RMSE %	# OF ITER	FLOW RMSE	OD TABLE RMSE %
1												
2	10	5.65	46.73	10	5.65	46.73	2	5.74	2.95	2	5.74	2.95
3	10	8.35	47.22	2	9.73	47.41	2	10.26	4.82	2	9.11	4.57
4	12	10.96	43.83	3	8.81	43.59	3	8.80	4.97	2	10.37	5.02
5	9	115.03	66.48	8	121.37	67.57	8	121.97	47.85	10	109.86	43.77
6	13	15.07	48.28	8	21.40	48.35	8	21.27	22.34	4	11.62	6.21
7	9	110.32	65.54	10	115.76	65.68	10	116.86	64.60	10	108.70	61.95
8	10	11.78	45.61	8	17.03	44.53	10	12.82	5.54	2	10.24	5.33

TABLE 3 INCREASED O-D DEMAND PATTERN WITH EPSILON = 0.01 RATHER THAN 0.1

	RUN E O-D Demand varies Initial seed = 100 Epsilon = 0.01			RUN I O-D Demand varies Each seed = Act. sol Epsilon = 0.01		
TIME PERIOD	# OF ITER	FLOW RMSE	OD TABLE RMSE %	# OF ITER	FLOW RMSE	OD TABLE RMSE %
1						
2	18	6.70	46.93	9	7.05	3.54
3	9	7.94	47.54	10	8.06	4.16
4	9	9.62	43.92	9	9.50	4.60
5	16	113.15	66.26	17	113.27	44.60
6	16	13.92	48.31	12	13.21	5.73
7	17	111.71	65.17	16	112.25	63.11
8	16	10.45	45.76	10	9.64	4.90

Experimental Trial I

The first experiment consisted of runs A and B, where the INTEGRATION simulation model was used to simulate traffic conditions on the Burlington Skyway for a constant demand O-D matrix for 2 hours. Link flow and minimum path tree estimates were generated by the simulation model for each time slice, after the initial 15-min start up, and these files were analyzed using SODGE.

Run A used a uniform seed O-D matrix for the first time slice, whereas run B used the actual O-D matrix as the seed. The results, which are shown in Table 1a, indicate three important facts. First, the analysis of the first time slice with a uniform seed matrix requires many more iterations (10) than any of the subsequent time slice analyses (2–4). This indicates the efficiencies that are achieved if the solution matrix for a previous time slice is used as a seed matrix for a subsequent time slice. Second, even though run B was provided with the actual solution matrix, the O-D table it estimated was not exactly the same as the one that was used in the simulation model. The main reason for this difference is the presence of traffic signals in the network, which cause link arrival and departure rates to be nonuniform. This discrepancy is also shown by the lack of a perfect convergence of the link flows.

Finally, even though the link flows in runs A and B converged to roughly the same link flow error level, the deviation from the actual true matrix was much larger for the analysis seeded with a uniform matrix (43 to 47 percent) than for the analysis that was seeded with the true matrix (2 to 7 percent). This indicates that two solutions can have comparable link flow convergences but still differ substantially in their agreement with the actual matrix.

Experimental Trial II

Although runs A and B were based on the simulation model outputs for a constant traffic demand pattern, runs G, D, F, and H considered a traffic demand pattern that was constant for 1 hr, increased for certain O-D pairs for 30 min, and then returned to its original state for the final 30 min. The original and the changed O-D matrix are presented in Tables 4a and 5, whereas the consequent statistics for each of these runs are illustrated in Table 2.

After 1 hr, the INTEGRATION simulation model increased the vehicle departure rates at the respective origins immediately, but all relevant link flows did not increase until these vehicles reached those links downstream. Consequently, there

TABLE 4 ORIGINAL O-D MATRIX FOR TIME 0–1.0 HRS AND 1.5–2.0 HRS

Orig down \ Destinations across	1	2	3	4	5	6	7	8	9	10	SUMS
1	0	180	180	60	60	60	180	60	900	180	1860
2	0	0	180	30	30	30	90	30	90	90	570
3	0	180	0	30	30	30	30	30	90	30	450
4	0	0	0	0	60	30	30	30	30	30	210
5	0	0	0	0	0	0	60	0	60	60	180
6	0	0	0	0	0	0	60	60	60	60	240
7	0	0	0	0	0	60	0	60	180	120	420
8	0	0	0	0	0	60	60	0	0	60	180
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0
SUMS	0	360	360	120	180	270	510	270	1410	630	

TABLE 5 MODIFIED O-D MATRIX FOR TIME 1.0–1.5 HRS

Orig down \ Destinations across	1	2	3	4	5	6	7	8	9	10	SUMS
1	0	240	240	60	60	60	360	180	1200	300	2700
2	0	0	240	30	30	30	90	30	90	180	720
3	0	240	0	30	30	30	30	30	90	60	540
4	0	0	0	0	60	60	30	30	60	60	300
5	0	0	0	0	0	0	60	0	60	60	180
6	0	0	0	0	0	0	60	60	60	60	240
7	0	0	0	0	0	60	0	120	240	240	660
8	0	0	0	0	0	60	90	0	0	120	270
9	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0
SUMS	0	480	480	120	180	300	720	450	1800	1080	

is a lag in the response of the link flow rates, which at worst is equal to the travel time between the two most separated O-D pairs, or approximately 8 min. This lag implies that the link flows for these time slices would be a weighted average of the previous time slice's O-D rate and the new O-D flow rate. This so-called "transient" effect should disappear in the second time slice after the change.

Run G (Table 2) provides the results for an analysis where the SODGE routine was seeded with a matrix of uniform cell values (equal to 100) for each time slice. As the analysis starts from scratch at the start of each time slice, the number of iterations stays relatively high (9–13). Also, because of the presence of transients, the link flows of time slices 5 and 7 are shown to converge very poorly, and the estimates of the O-D matrix in each case are also less accurate. These results should be compared to run D, where SODGE was seeded with a uniform matrix, but was allowed to use its predicted O-D matrix from each time slice as a seed for the next time slice. Both the flow and the O-D cell errors are shown to be the same as for run G, but the number of iterations required to find this comparable solution is usually reduced.

Run F analyzes a situation where the seed provided to the first time slice is the actual solution matrix, whereas the O-D of each subsequent time slice is then calculated using the previous slice's solution as its seed. This results in a reduced number of iterations for the first time slice analysis, and a consistent number of iterations for all the subsequent time slices. Similarly, the degree of convergence of link flows is roughly the same, as is the increase in the link flow error during the transition periods. However, the agreement with the actual O-D table is much better for run F, even after the two major changes in the traffic pattern. The importance of a good seed matrix to initiate the analysis of the first time slice is emphasized in run F. The knowledge of the underlying pattern appears to be of considerable assistance, even after the matrix has undergone a change.

Finally, in run H SODGE was provided with the correct seed matrix for each time slice. The results are the best of any of the runs, but some error still remains, especially during the transient periods. When the SODGE performance in run F is compared against that in run H, one finds that relatively good results can be achieved by seeding only the first time slice with the actual matrix. An overview of the relationships among the results for runs G, D, F, and H (all in Table 2), is provided in Figure 6. The results for runs A and B (Table 1) are compared to the results for runs D and F (Table 2) in Figure 7.

Experimental Trial III

The third experiment involved an analysis of the implications of utilizing a different epsilon value for identifying the onset of convergence. Runs E and I have results very similar to runs D and H, respectively, as shown in Figure 8. Dropping the epsilon from 0.1 to 0.01 significantly increased the number of iterations, marginally increased the degree of flow convergence, and has only a negligible effect on the accuracy of the estimated O-D table.

Consequently, the use of an epsilon value of no smaller than 0.1 seemed appropriate for the tests carried out using

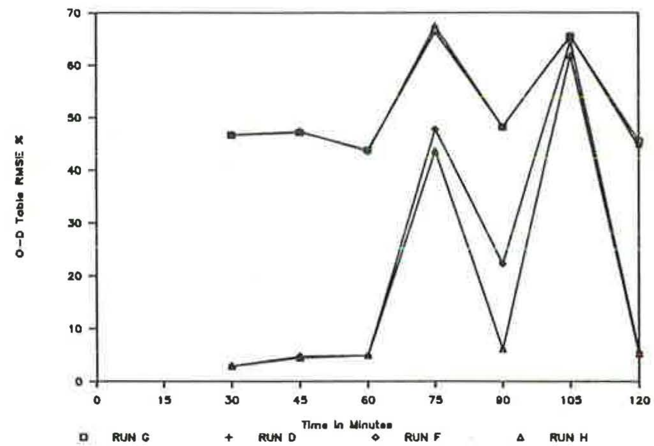


FIGURE 6 Comparison of fit for different seed matrices.

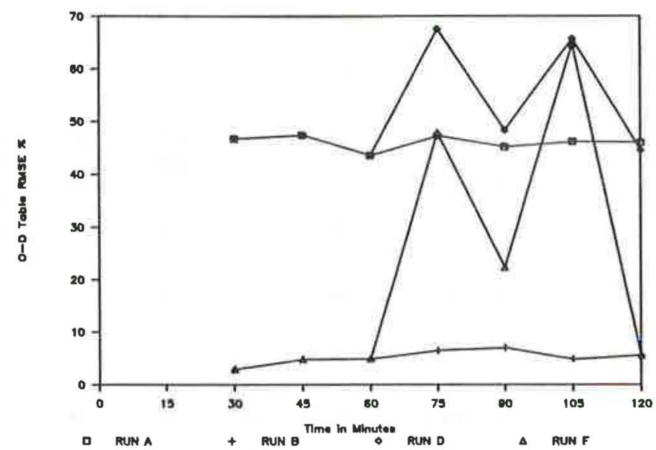


FIGURE 7 Comparison of fit for constant versus changing demands.

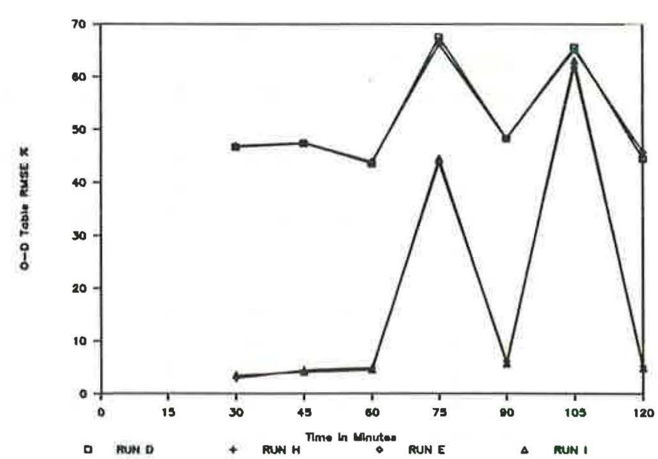


FIGURE 8 Comparison of fit for two different stopping criteria.

this network. However, this value is likely to be different for different networks and perhaps even for different traffic conditions. Epsilon should therefore be nondimensionalized and expressed in a format that is unbiased by the scale of the network and its traffic pattern.

Experimental Trial IV

The fourth and final experiment involved an analysis of the implications of using a 5-min versus a 15-min analysis interval. This analysis allowed for a more microscopic look at the dynamic behavior of the traffic within the network, and the consequent implications for the on-line generation of O-Ds. The traffic flow rates on four different links in the network are shown in Figure 9 (top).

When one compares the increase in traffic volume on the freeway links (L8 and L10), to the increases on the surface streets (L6 and L23), it is apparent that the traffic increase on the freeway is much more pronounced than on the surface streets. This is to be expected, as the change in the O-D pattern affects primarily the freeway trips.

When one compares the first freeway link (L8) to the last freeway link (L10), one can also detect the lag in the transient effect for the downstream link (L10). This lag is approximately equal to the travel time from zone 1 to zone 9. It significantly complicates the synthetic O-D analysis, as the effect of a change in O-D pattern cannot be detected on the final downstream link until about 10 min after the O-D pattern has changed. Synthetic O-D solutions based on traffic flows during this initial 10-min transient would therefore incorrectly allocate the increased flows, on the first few links, to the same origin but a nearer destination. Of course, after a decrease in traffic flow for an O-D pair, the reverse effect would take place.

The results for the iterative SODGE application every 5 min are illustrated in Figure 9 (bottom) and Table 6. This analysis was initiated with the correct seed matrix after 20 min, and all subsequent applications of SODGE were then seeded with the O-D solution matrix from the previous 5-min period.

As shown, during the first 60 min, the number of iterations, the link flow error, and the O-D matrix error vary somewhat about a relatively stable average value. Then, after the increase in the O-D matrix at time 60 min, the dramatic transient results in a considerable error in the link flow convergence and the O-D matrix estimation during the next two 5-min time periods. Subsequently, the solutions generated by the algorithm again stabilize, until the O-D demand is dropped after 90 min.

Although the link flow error peaks for the time slice that is 5 to 10 min after the change in O-D flows occurs, the O-D matrix error has already begun to decrease significantly. Consequently, during the time slice that lasts from 10 to 15 min after the O-D flow increase or decrease, the O-D matrix provides a good fit once again.

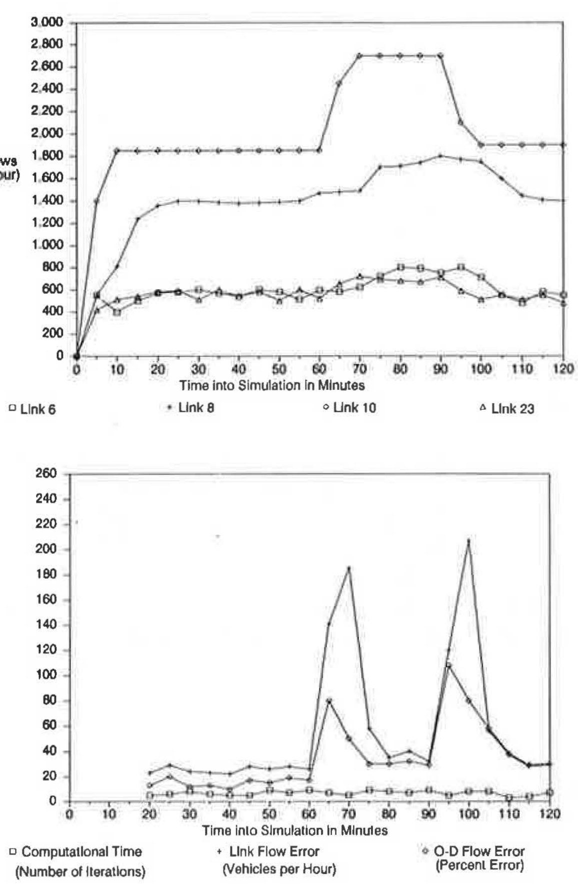


FIGURE 9 Change in 5-min flow rates on links 6, 8, 10, and 23 (top) and results for 5-min intervals and correct initial seed (bottom).

TABLE 6 RESULTS FOR 5-MIN INTERVALS AND CORRECT INITIAL SEED

TIME SLICE	TIME (Min) FROM	TO	NUMBER OF ITERATIONS	LINK FLOW ERRORS	O-D TABLE ERROR(%)
1	0	5			
2	5	10			
3	10	15			
4	15	20	5	23.0	14.3
5	20	25	6	28.8	20.9
6	25	30	9	24.7	13.3
7	30	35	6	24.0	14.3
8	35	40	5	23.0	10.3
9	40	45	5	28.0	17.8
10	45	50	9	25.2	14.7
11	50	55	6	28.6	18.4
12	55	60	9	24.2	13.4
13	60	65	7	141.3	78.1
14	65	70	6	184.8	49.2
15	70	75	10	55.8	26.3
16	75	80	9	37.4	26.4
17	80	85	8	43.7	28.0
18	85	90	11	34.4	24.4
19	90	95	6	119.0	105.9
20	95	100	10	206.4	76.9
21	100	105	10	54.9	22.8
22	105	110	4	33.8	11.5
23	110	115	5	23.6	16.8
24	115	120	9	25.5	14.9

DISCUSSION OF THE RESULTS AND THEIR SIGNIFICANCE

In most network-oriented traffic management studies it is important to determine the prevailing O-D traffic demands during a peak period in order to estimate the indirect diversion impact of any proposed control strategies or to determine the direct diversion impacts of control strategies that are deliberately intended to result in a rerouting of traffic. The need to evaluate the growth and decay of traffic patterns during a peak period requires the analyst to consider the time-varying O-D demands of the network explicitly. As it is either impractical or uneconomical to obtain time-varying O-Ds by direct survey methods, it becomes necessary to use indirect methods. These synthetic O-D methods, although subject to some error and various other technical problems, provide the next-best solution, as often no alternative approach to obtaining these data is available.

Summary of Initial Findings

The results presented here indicate that it is feasible to automatically generate a sequence of O-D matrices for a series of time slices during a peak period. Such matrices can be used in conjunction with freeway control or diversion models for off-line preevaluation of different strategies using historical traffic counts or for on-line generation of diversion strategies using real-time traffic-flow measurements.

The use of a time slice's solution O-D matrix as the seed for the subsequent slice appears to reduce the number of iterations required to achieve a certain level of accuracy, as compared to using a blank or uniform seed matrix for each new time slice. This efficiency in computation time is especially useful in real time control applications.

The use of a sound initial seed matrix at the start of a control period has benefits throughout the entire control period, even if the flow rates for a number of O-D pairs may change significantly. The structure of the initial seed is usually retained throughout the analysis period and assists considerably in selecting the most appropriate O-D matrix from among the often numerous possibilities.

Problems Unique to Synthetic O-Ds for Freeway Control

Traffic demands within congested freeway networks are never in full equilibrium. Instead, they appear to be always in some state of dynamic flux or transition. Consequently, problems arise because of the transients that occur when traffic patterns change, as these changes require a finite period of time to manifest themselves on all the downstream links that will be used by the given O-D.

Similar problems may also arise when queues cause two different link flow rates along a given path, one on the upstream end of the bottleneck and one on the downstream end. In this situation, the synthetic O-D program may mistake the resulting traffic flow observations as being indicative of two short trips rather than one long one.

It is expected that in practice the transient effects discussed here may not be as drastic or dramatic as the nearly instantaneous 50 percent increase in traffic flows that we used in this report. In actual networks it is more likely that traffic demands will change gradually during a 15- to 30-min period, in which case the problems generated because of transients will be lessened, and perhaps become insignificant.

Recommended Further Work

The proposed procedure should be implemented using real data from an FTMS and urban traffic control system (UTCS) and compared to the O-Ds as estimated from a detailed driver survey. However, it is likely that in practice the true O-D matrices will never be be available for use in assessing the quality of the solutions obtained. In this case, the relative magnitudes of the errors, as estimated in the sample runs discussed here, could be a guide as to the likely margin of error that would be present in situations where the true matrix is unknown.

To address some of the transient problems, it may be helpful to use the algorithm with fractional "link use" probabilities. In this case, the path that is taken by vehicles between a given O-D pair may be assigned decreasing-use probabilities along its length to reflect the decreased likelihood that the effects of the shift in that O-D demand have propagated a certain distance away from its origin. Similarly, different probabilities may be assigned to links before and after a bottleneck location to indicate the fraction of drivers from a particular O-D pair that are likely to be stuck in the queue.

Finally, it is proposed that the analysis discussed here be repeated for a wider corridor in which more alternate routes are available. Although it would be much more difficult to trace the causes and effects of any transients, this type of network application would be more representative of corridors in which one freeway can be avoided by traveling on any one of up to three or four alternate routes.

ACKNOWLEDGMENTS

The research reported in this paper was sponsored in part by the R&D Branch of the Ministry of Transportation Ontario (MTO), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Queen's University.

The authors would like to thank A. Ugge and E. R. Case of MTO's R&D Branch for their helpful comments during this study, and the staff of MTO's Freeway Traffic Management Office for their assistance in providing data on the Burlington Skyway FTMS for this study.

REFERENCES

1. M. Van Aerde, S. Yagar, A. Ugge, and E. R. Case. *A Review of Candidate Freeway/Arterial Corridor Traffic Models*. Annual Transportation Research Board Meeting, Washington D.C., 1987.
2. H. J. Van Zuylen and L. G. Willumsen. The Most Likely Trip Matrix Estimated from Traffic Counts. In *Transportation Research*, Vol. 14B, pp. 281–293.

3. E. Cascetta. Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Squares Estimator. *Transportation Research Record 18B*, TRB, National Research Council, Washington, D.C., 1984, pp. 289–299.
4. M. J. Maher. Inferences on Trip Matrices from Observations on Link Volumes: A Bayesian Statistical Approach. *Transportation Research*, Vol. 17B, pp. 435–447.
5. C. Hendrickson and S. McNeil. Estimation of Origin-Destination Matrices with Constrained Regression. *Transportation Research Record 976*. TRB, National Research Council, Washington, D.C., 1984, pp. 25–32.
6. H. J. Van Zuylen. *Some Improvements in the Estimation of an O-D Matrix from Traffic Counts*. Eighth International Symposium on Transportation and Traffic Theory, University of Toronto, Canada, pp. 183–190.
7. D. Van Vliet and L. G. Willumsen. *Validation of the ME2 Model for Estimating Trip Matrices from Traffic Counts*. Eighth International Symposium on Transportation and Traffic Theory, University of Toronto, Canada, pp. 191–198.
8. G. Noxon. Development and Testing of a Model to Estimate Origin-Destination Matrices from Link Counts. Undergraduate Technical Report (unpublished), Department of Civil Engineering, Queen's University, Kingston, Canada.
9. M. Van Aerde, J. Voss, A. Ugge, and E. R. Case. An Integrated Approach to Managing Traffic Congestion in Combined Freeway and Traffic Signal Networks. *ITE Journal*, Vol. 59, No. 2.
10. M. J. Delsey and S. E. Stewart. Queen Elizabeth Way—Burlington Skyway Traffic Management System. *Transportation Forum*, Vol. 2-2, pp. 68–72.

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

Method to Synthesize a Full Matrix of Interdistrict Highway Travel Times from Census Journey-to-Work Data

W. THOMAS WALKER

The 1980 census Urban Transportation Planning Package (UTPP) is a large-scale source of observed travel times derived from the journey-to-work data. Plans for the 1990 census indicate that work trip times will again be collected. These data are incomplete, however, in that most cells in the interdistrict matrix have no census travel time observation. This paper explores the accuracy and usefulness of census travel time data versus survey- and network-based methods and airline distance approaches at the minor civil division (MCD) level of analysis. A promising technique to extract and edit the census data and to synthesize the missing observations is developed.

The need for accurate estimates of a matrix of highway travel times between a fine-grained system of zones arises in many transportation activities, including market analyses for retail business locations, vehicle routing and scheduling operations, highway needs studies and cost-benefit analyses, transit service planning, highway accessibility studies, and validating the travel time data input to travel forecasting models. The large number of cells contained in realistic travel time matrices generally precludes exhaustive interdistrict estimation with floating car survey methods because of high costs. Travel time matrices are usually generated by building minimum paths through computerized networks, where the links may contain measured speeds, but more often, because of survey costs, contain generalized averages of measured speeds, perhaps stratified into a number of route types or area type and time-of-day categories. The large geographical area associated with commercial census data bases and many highway-related activities, such as scheduling long-haul truck deliveries often precludes network methods because of the large number of links required. In such cases, travel time estimation based on some form of airline distance analysis is the norm.

The 1980 census Urban Transportation Planning Package (UTPP) is a large-scale source of observed travel times derived from the journey-to-work data. Plans for the 1990 census indicate that work trip times will again be collected. These data are incomplete, however, in that most cells in the interdistrict matrix have no census travel time observation. This paper explores the accuracy and usefulness of census travel time data versus survey- and network-based methods and airline distance approaches at the Minor Civil Division (MCD) level of analysis. A promising technique to extract and edit the census data and to synthesize the missing observations is explored. Comparisons between census and network-based travel time estimates are made for a four-county portion of New Jersey that was covered by a recent travel time survey, as are comparisons between census and airline distance estimation methods. As a case study, a census-based travel time matrix is generated at the MCD level for the portion of the Northeast corridor shown in Figure 1, including southeastern Pennsylvania and the entire state of New Jersey (816 districts).

Delaware Valley Regional Planning Commission, The Bourse Building, 21 South 5th Street, Philadelphia, Pa. 19106.

BACKGROUND

Over the last quarter century there has been a great deal of work on estimating highway speeds and travel times. The relationship between highway speed, vehicular volume, and roadway capacity on individual links was studied intensively in the early 1960s under the auspices of the FHWA and the Highway Research Board. This work culminated in the 1965 Highway Capacity Manual (*1*). Because of the importance of speed in determining roadway service levels for design purposes, interest in this subject has continued to the present day (*2*). This work, however, is focused on individual highway links with little attempt to estimate interdistrict travel times.

Travel-demand forecasters, however, have considerable interest in travel times because they are used as a principal measure of interzonal highway service levels for purpose of trip distribution and modal split. These analyses, however, tend to be focused on the impact of errors in various network methods on the accuracy of the travel demands generated for specific transit and highway facilities (*3*,*4*). In this work, accuracy in trip travel time estimates is often sacrificed to obtain more accurate demand estimates or network assignments, viewing the resulting time as an impedance or composite measure including other factors. To date, there has been little attempt to use or enhance the census journey-to-work travel times except perhaps to validate network-based travel time results, because direct census observations fill only a small portion of the highway travel time matrix.

METHODS FOR ESTIMATING INTERDISTRICT TRAVEL TIMES

This section explores three methods for synthesizing travel times between a moderately dense system of analysis districts; synthesized census work-trip times, minimum path network

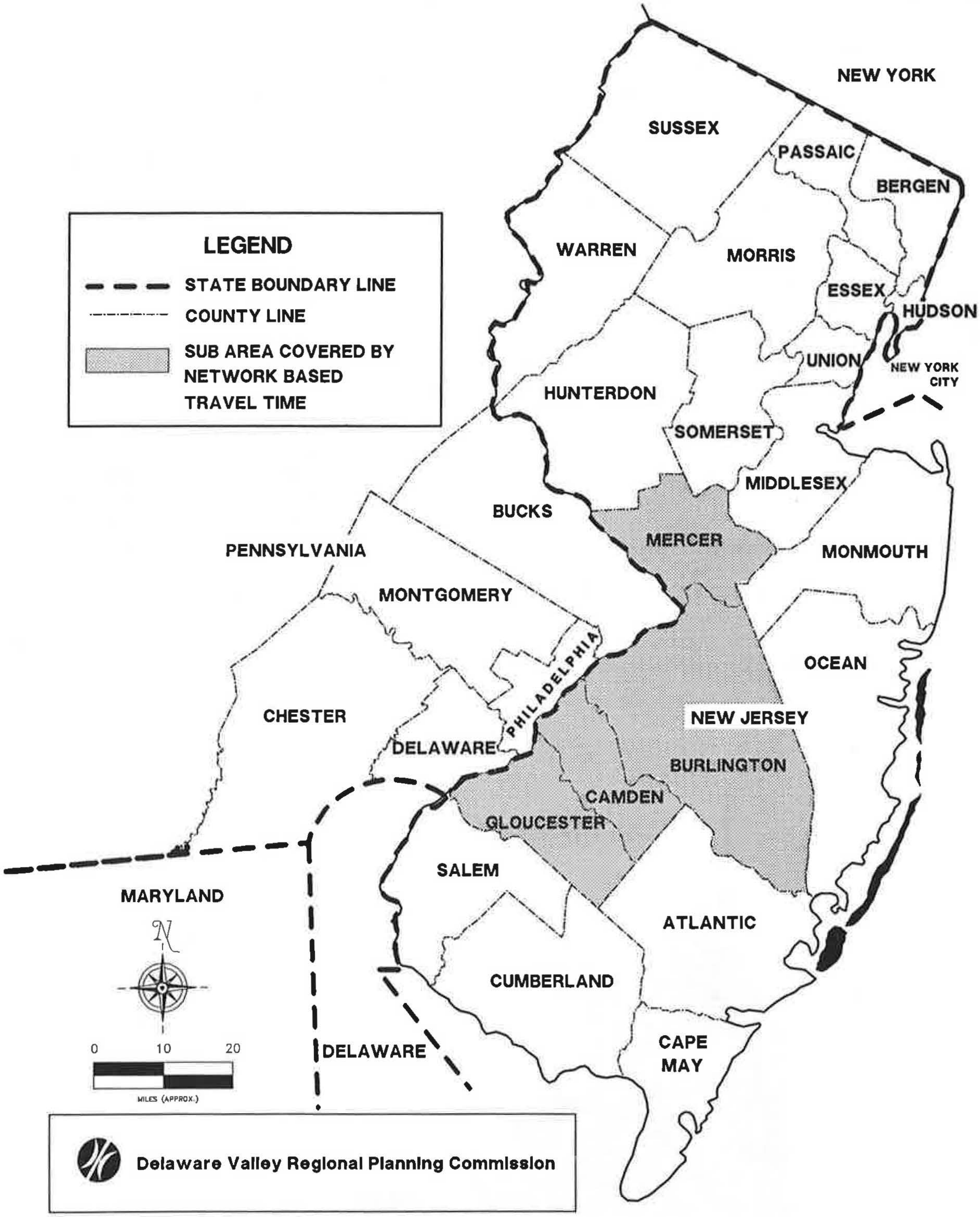


FIGURE 1 Case-study area for travel time estimation.

methods, and airline distance-based time calculations. The district system used is based on MCD within southeastern Pennsylvania and New Jersey. In total, these comprise some 816 districts with an average area of about 12 square miles, although there is considerable variation in individual district sizes ranging from small boroughs to large rural townships. The City of Philadelphia is technically one MCD. Philadelphia was broken up into 12 subdistricts in this analysis.

Questions 23 and 24 of the 1980 Census long form asked each employed resident to indicate where he worked, the work-trip time, and what travel mode was used. These data were processed, tabulated, and averaged at the census tract level for just under eight percent of all employed residents and distributed in Part IV of the UTPP (*5*). Both the Delaware Valley Regional Planning Commission and the New Jersey Department of Transportation purchased the 1980 UTPP. These data files provide a very large number of observations of interdistrict travel time. The drive-alone travel mode was chosen because it gives the most direct estimate of over the peak-period road-travel times. Carpooling travel times were rejected because they included an additional travel time associated with picking up passengers. In total, almost 160,000 census tract level observations of drive-alone travel times were recorded in the combined UTPPs, some 15,000 of which were duplicates caused by overlap of the UTPP areas. When averaged to the MCD or district level of analysis, these observations cover almost 40,000, or six percent, of the 666,000 possible interdistrict interchanges. On average, there were about 4.2 census tract travel time observations per MCD interchange covered by the sample before editing. These observations were not evenly distributed, however. Some major centers, such as the Philadelphia central business district, had hundreds of observations, while small MCDs may have had 10 or fewer observations. On average, there were almost 47 travel time observations per MCD.

This UTPP travel time data is also available at the census tract and county levels of aggregation. Although more precise in terms of areal definition, census tracts are too numerous for the resulting travel time matrix to be adequately filled (3.2% full) with the UTPP sample. The large number of tracts involved makes the methods used to synthesize missing observations less reliable and more costly in terms of computational effort. Conversely, the county level of aggregation is too coarse to adequately support most transportation applications requiring travel time estimates. The MCDs provide a good compromise between travel time data need and data availability. Although of some use as received from the Census, this MCD data set must be expanded to full rank before it is appropriate for most scheduling and accessibility applications.

Synthesizing Missing Travel Times from Census Data

The basic motivation for the method used to synthesize missing travel times derives from the fact that census travel times are not randomly dispersed through the interdistrict travel time matrix. Rather, these observations follow commutation patterns as expressed by the manifest behavior of the employed resident who filled out the 1980 Census. Generally, these persons work in employment centers, and these centers are served by the highway system with workers following a sort of minimum path from home to work. It is in the nature of work trip length frequencies for observed work trips to be clustered in the shorter travel time ranges (Figure 2); there-

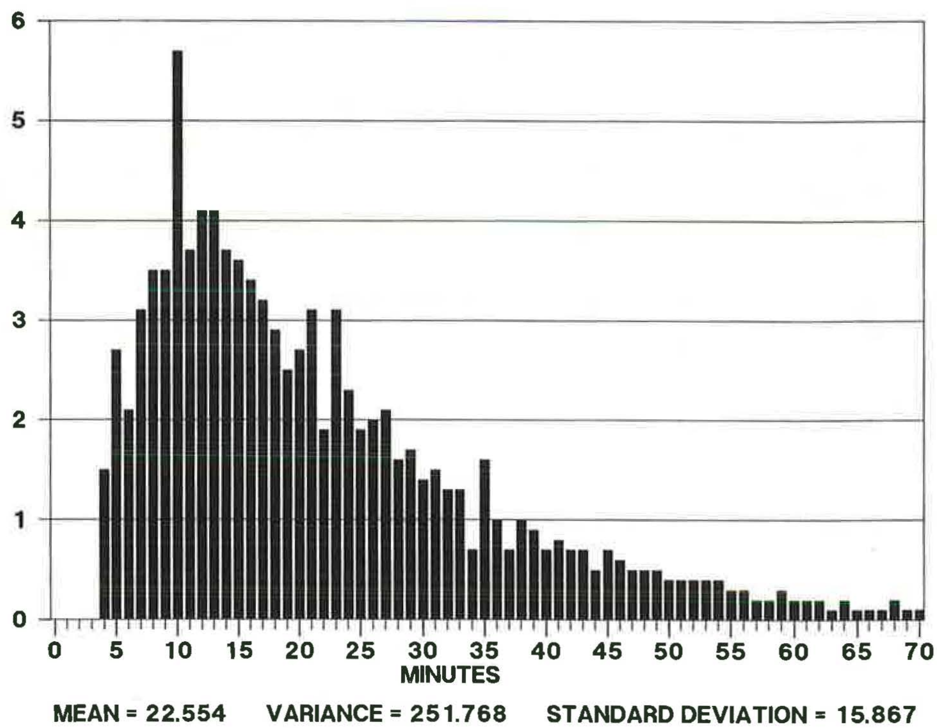


FIGURE 2 Trip-length frequency distribution of home-based work trips.

fore, it is the longer interdistrict interchanges that tend to be missing from the census data.

However, these missing observations may be synthesized from the census data by the creation of a network using each census observation as arc or one-way link. These arcs are then chained and summed into paths using a minimum path algorithm, thereby creating a synthetic time for most, if not all, missing observations. This process is similar to the spider network applications that were popular in travel analysis some years ago.

Figure 3 presents a visualization of this process. The Census has provided a drive-alone travel time estimated from MCD A to MCD B, and from MCD B to MCD C, but not from A to C. However, by chaining and adding observations A-B to B-C, one can synthesize a travel time estimate from A to C. This method, however, imposes special accuracy requirements on the census data in that spurious responses must be removed before synthesizing the missing travel times. Standard statistic techniques involving averaging individual responses do not apply in this application.

Editing the Census Travel Time Responses

The principal concern in editing the raw census travel time responses is to remove unrealistically small travel times from the file. Such a response, given the minimum path techniques used to synthesize missing travel times, could cause whole regions of the resulting matrix to have unrealistically small time estimates. For example, one employed resident from northern New Jersey reported that he made it to work in Honolulu, Hawaii, in 30 min. Even this extreme example is not necessarily a spurious response. Generally speaking, the census responses were contentious and quite good. This person may maintain a temporary or secondary residence in Hawaii 30 min from work. However, a 10-min estimate of travel time from, for example, Newark to Atlantic City, New Jersey, will have a disastrous impact on a synthetic matrix generated by minimum path methods.

The method used to edit the UTPP drive-alone travel time observations involved calculating an effective speed based on the airline distance between the home and work MCD. Those observed times that fall outside of an acceptable speed range were rejected. The airline distances were generated from grid coordinates of an arbitrary point located within each MCD, generally the geographic centroid. The edit procedure took each intercensus tract observation, calculated an effective airline speed based on the MCD of origin and destination, and rejected those observations that had an effective speed greater than 50 airline miles per hour (mph), or less than 4.9 mph. This corresponds to an over-the-road speed of 55 to 60 mph and 5 to 6 mph respectively, depending on the degree of circuity involved. The lower limit could be raised somewhat. However, many short commutes of 5 miles or less involve rather low speeds because much of the trip takes place on local and collector streets and relatively little on higher speed highway facilities. In any case, a minimum path algorithm can build paths around arcs with too low an effective speed. The

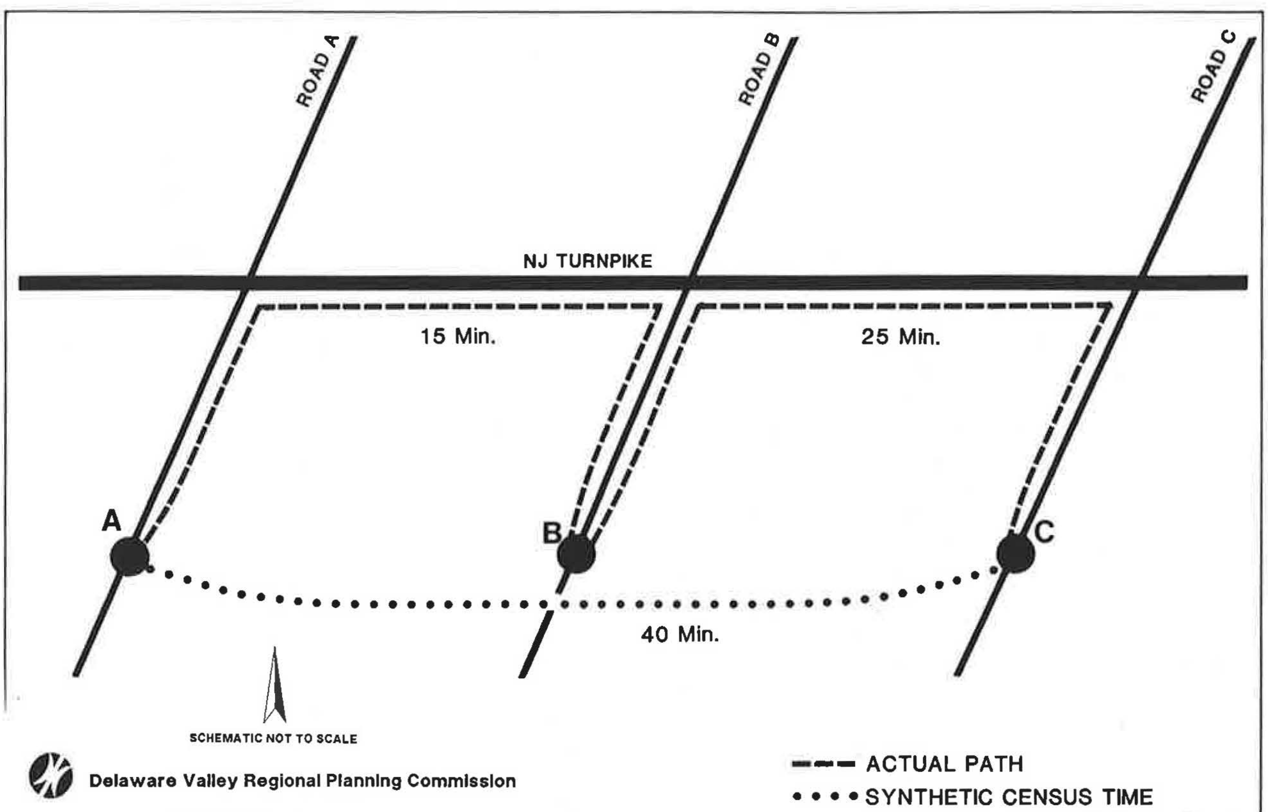


FIGURE 3 Hypothetical example of synthesized census travel time.

critical part of the edit is to remove census observations with excessive operating speeds. After the edit at the census tract level, the surviving travel time observations were compressed to the MCD level of detail, averaging the times in cases where multiple observations between two given MCDs existed. This edit procedure rejected 8,559 of 158,476 tract level travel time observations, or about 5.4 percent of the census data. After averaging to the MCD level of aggregation, 37,855 inter-MCD travel times were available to the path building algorithm.

A critic of this approach might contend that the synthetic travel times generated will tend to be over estimated because implicit local street access and terminal times will be duplicated for each arc added to the chain spanning the origin and destination MCD. In order to minimize chain length, all long-distance as well as short-distance travel time observations must be included in the calculation. Individual MCDs may have tens or even hundreds of travel time observations emanating from them. Through these long arcs, the average number of arcs per chain in the synthetic matrix was held to 3.75. The price we pay for this, however, is that the resulting network is nonsparse. Network sparseness refers to the percentage of the maximum number of one-way arcs (nodes × nodes) included in the network or the percentage of the node-link incidence matrix that is nonzero. Most packaged computer programs restrict the number of arcs out of a node to 4 or 8 to maintain sparseness and are therefore useless for synthesizing missing census travel times from UTPP data. By way of comparison, a traditional FHWA highway network may be about 0.02 to 0.3 percent full, whereas the spider network corresponding to the census travel time observations is 5.69 percent full. This manyfold increase in the fullness of the network has significant implications for the computer science techniques used to implement the Moore minimum path algorithm.

Minimum Path Building in Nonsparse Networks

The theorems defining the mathematical properties of the Moore algorithm apply to all levels of sparseness. The number of arcs per mode principally impacts the techniques used to implement this algorithm. It may be useful to consult the work of others on related problems, although most of the operations research literature on path building deals with sparse networks that have had much more computational success.

The Moore algorithm uses dynamic programming techniques to construct a tree or ordered sequence of arcs that connect the MCD of origin with all destination MCDs via a minimum time path. This tree is built one arc at a time, starting with the origin, always adding the minimum time arc emanating from the subset of active nodes. A node is considered active when it has been reached by the tree, but no arc emanating from the node has been added. Nodes not yet added to the tree and nodes that already have an outbound arc in the tree are considered inactive and therefore ineligible for link selection. This list of eligible arcs is often called a sequencing table and the problems caused by nonsparseness relate to the computational effort needed to maintain and update this table, which can grow quite large during the path building process.

Generally, these problems involve optimizing the tradeoffs between the computational savings derived from minimizing the size of the sequencing table and the computational costs required to edit its contents. The use of largely standard edit and merge techniques to maintain the sequencing table produced a method that is computationally practical, albeit much more expensive than tree building through sparse networks. The 816 district case study took a time equivalent to about 1.5 hours on an IBM 4381 computer to build minimum paths and skim the times and number of arcs over each path.

Variations on the methods described here to construct a spider network from census travel time observations are possible. For instance, a two-way link could be constructed from each home-to-work observation by attributing the same time to the reverse movement. This alternate approach would reduce the tendency to overestimate travel times because of path circuity, but would almost double the fullness of the resulting network, thereby exacerbating the tree-building problems discussed previously. In any case, there is no evidence in the following corroborative data sets that synthetic census times generated by the recommended approach are overestimated.

Network Method to Estimate MCD Travel Times

Travel times between MCDs may also be estimated by traditional network methods. Each MCD is given a centroid designation and a series of network links are defined, each representing a given street or highway segment with additional nodes placed at street intersections. Centroids are then connected to this street system via a series of approach links whose travel times are representative of approach travel over local roads (Figure 4). Travel speeds over the street links should in theory be estimated by floating car survey techniques. This technique is commonly used in traffic studies throughout the United States and Europe.

The principal problem associated with the travel time estimation with network methods vis-a-vis the census work trip times is the high cost of coding the network and measuring the speeds. The grain or density of network links should be at least equal to that of the district system boundaries. For MCDs, this implies that all freeway and major arterial facilities must be included with minor arterials and collector roads added as needed. For the 816 district example given above, this network would be large and expensive to survey.

Network and survey methods can be efficiently applied for smaller study areas, however. In 1986, such a travel time survey and network was prepared by the Delaware Valley Regional Planning Commission (DVRPC) for the portion of southern New Jersey, including Mercer, Burlington, Camden, and Gloucester counties (Figure 1). Although only a small portion of the 816 district system is covered by the census data, this network provides alternate MCD travel time estimates for comparison purposes. The network had the following characteristics: 107 districts, 800 nodes, and 1107 two-way links.

These districts were for the most part MCDs, although some small MCDs were missing, and cities such as Trenton and Camden were subdivided into smaller zones. For comparison with census data, zonal times were averaged to an MCD value and missing MCDs were dropped from the calculations. The minimum paths through this network produced travel-time

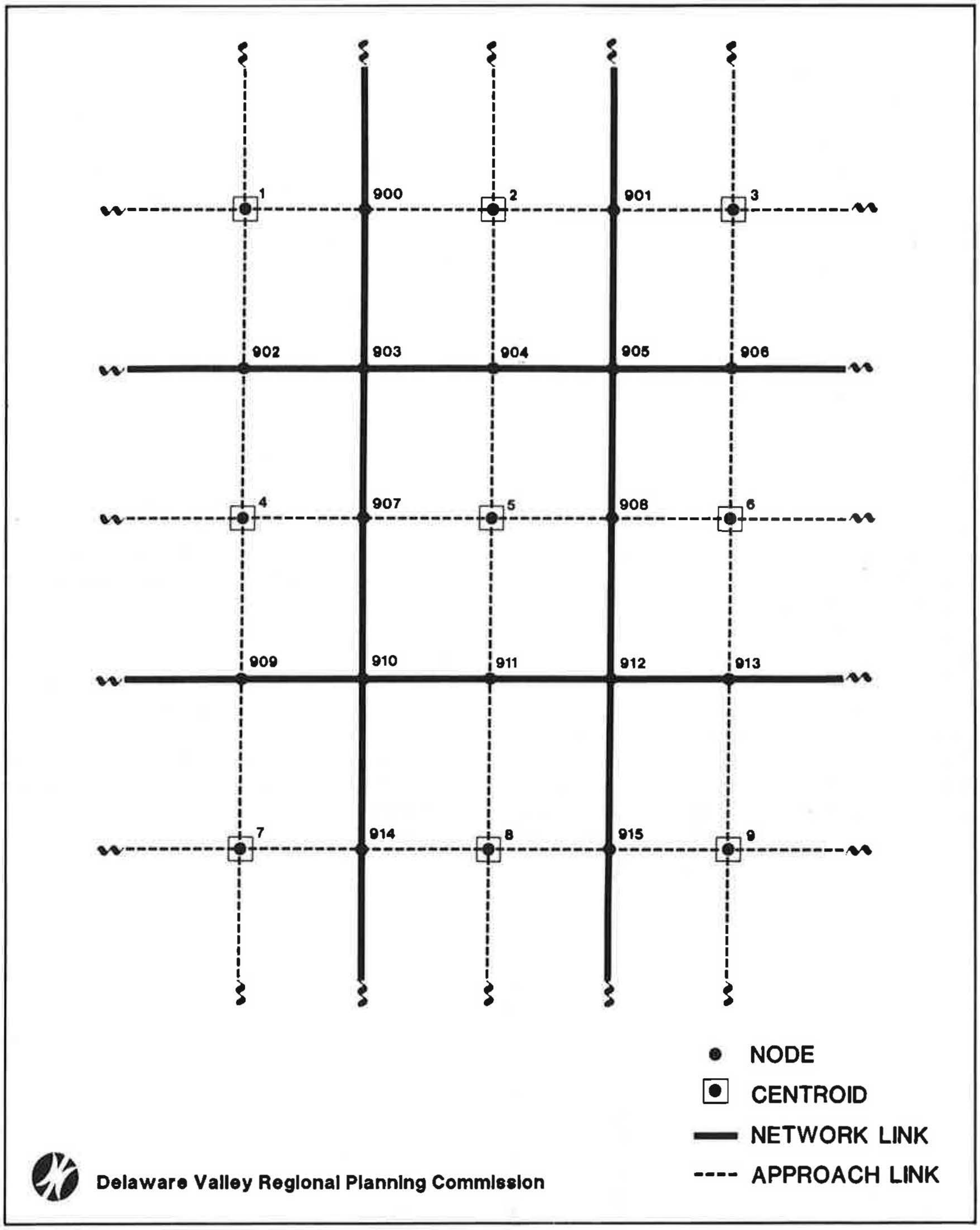


FIGURE 4 Traditional highway network.

estimates between 94 of the 112 southern New Jersey MCDs. In total, 30 percent of these links were provided with a measured speed. Unsurveyed links were given an average speed by route and area type from the survey (*6*). Peak and off-peak speeds in this survey were almost the same for most roads. The average peak and off-peak speeds were 27 and 28 mph, respectively. Some suburban highways now have slower speeds in the off peak.

Airline Distance–Based Method for Estimating MCD Travel Times

The most common method to estimate travel times in use in vehicle scheduling and market analysis studies is to use airline distance as a proxy for travel time. These distances can be calculated easily from grid coordinates by triangulation. With this technique, distances can be made specific, for example, to the individual warehouses or fast food restaurants. This specificity is in general not possible in the census or network methods because of sample requirements and network size. However, we pay a price for this specificity in the accuracy of the times obtained. It is obvious that travel speeds are not the same in all directions from a given point.

For purposes of comparison with census times, the average speed (27 mph) from the network-based travel time survey was used after adjusting for the average circuity of the street system. A factor of 1.09 calculated by dividing total over-the-road network by total airline distance reduced this speed to 24.8 mph for the area covered by the network and travel time survey.

VALIDATION OF THE SYNTHETIC CENSUS RESULTS

Although it is possible to synthesize easily a complete travel time matrix from the drive-alone census journey-to-work travel times with the methods outlined previously, the accuracy of this synthetic data is unknown. In determining this accuracy, one cannot follow the usual approach of comparing synthesized with actual times because these two subsets of the travel times for the most part do not overlap. Synthetic census times exist for interchanges, with real census observations only in cases where a faster chain has been built around the census time. Rather, an indirect validation method must be used.

We will first compare the observed census travel times with the network-based and airline distance–based estimates discussed previously. The statistical indicators calculated from these comparisons will then form a bench mark for comparing the synthesized census matrix with these alternate travel time estimates. If the synthetic census times have comparison statistics similar to the observed census observations, then this is indirect evidence that the synthesized travel times are also reflective of actual conditions. When interpreting these comparisons it is important that only the observed census time data resulted from a random sample and therefore in theory has known statistical properties. The network and airline distance travel times are also synthetic. For this reason the comparisons between the observed census, network, and airline distance estimates are of interest in their own right.

The Bureau of the Census recommends the following variation on the standard error equation for use in estimating errors in census means

$$\sigma_{\bar{x}} = C\sqrt{\frac{\sigma^2}{n}}$$

where C is an adjustment factor that compensates for processing errors in coding and tabulating the census responses. C is not given for travel times; however, the value of this parameter varies from 0.8 to 3.4 for other items in the census. Another problem is that many travel time responses were previously averaged to census tracts, perhaps reducing the variation in the sample. For these reasons it is not possible to calculate the standard error for the census data.

A histogram of intertract census drive-alone travel time responses by minutes of total time is shown in Figure 5. These census times have been averaged in cases where multiple observations between tracts existed in the raw census data. Since these multiple observations tend to be concentrated in the lower time intervals, this averaging tends to shift the center of gravity of the histogram to the right compared to the network time–based trip length distribution shown in Figure 2. Although averaged, there are pronounced peaks in this

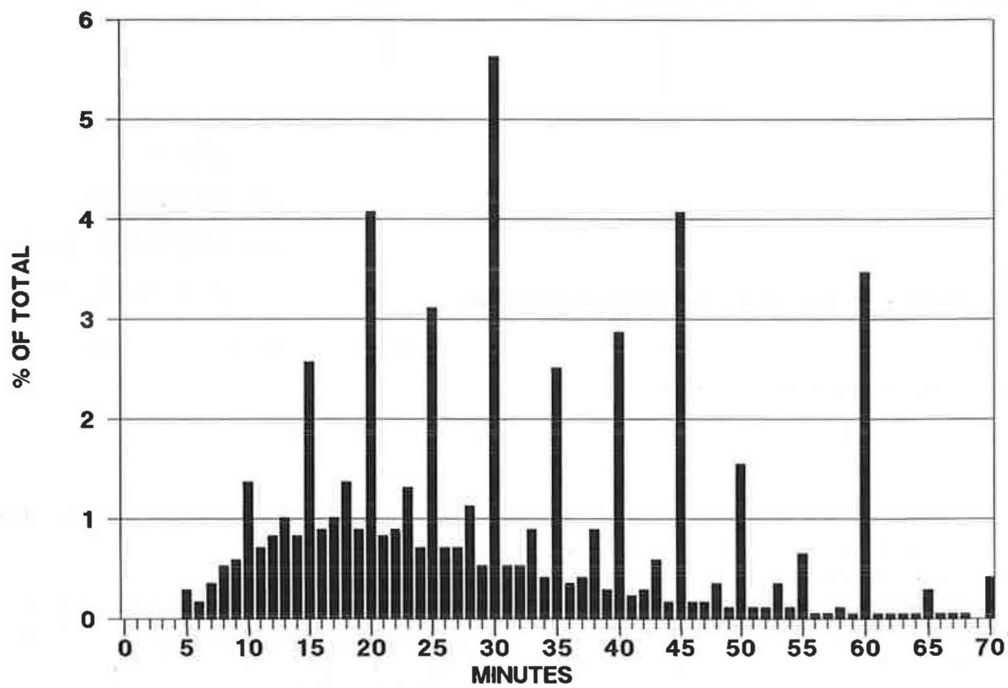


FIGURE 5 Observed census drive alone travel times.

data at even 5-min intervals and somewhat larger peaks at 15-min intervals. Respondents tended to round their journey-to-work times to the nearest 5 min and, to a greater degree, 15 min. This behavior is understandable given the probable daily variations in travel times. The journey-to-work time is known accurately by most people, given the importance of the work trip and the need to arrive at work on time, although individual variation in driving habits may be expected. The aggregation to the MCD level is also a source of travel time variation, especially in rural areas where MCDs are large. Approximately 10 percent of the travel time observations represent workers allocated to the geographic work place by the census because of incomplete address. Since reported travel time was a primary basis for determining allocation peer groups, this procedure will not introduce significant errors into the observed census travel time data (*7*). Rather, allocated observations will be associated with interchanges that already have similar times in the census data.

A statistical comparison between the observed census travel times and corresponding results of the network and airline distance methods are presented in Table 1. The R^2 measure of each is similar (0.67 and 0.70) and there is little difference between the census, network, and airline distance-based mean values. There is a substantial difference between the standard deviation of the airline distance and census travel times that is not found in the network/census comparisons. The airline distance method with its constant average speeds tends to overestimate the travel time for long trips and underestimate shorter trip times. The network-based times had an almost 8 min (30%) root mean square (RMS) difference. The corresponding RMS difference for airline distance times was somewhat larger—9 min or 35 percent.

Also included in Table 1 are three useful decomposition measures of mean square difference proposed by Theil (*8*). For the Theil tests, UM measures to the fraction of mean square error attributable to differences in the means, US refers to the fraction of this error resulting from difference in standard deviation, and UC estimates the corresponding fraction resulting from incomplete covariation (scatter). The Theil tests in Table 1 confirm that 98 percent of the observed difference between network and census times is attributable to scatter. The airline distance-based times had some 16 percent of the difference attributable to the relative standard deviations and 84 percent to scatter.

Figure 6 compares the prediction-realization diagrams for the observed census versus network times with the synthetic census versus network. From these diagrams it is apparent that the synthetic part of the census data tends to fill in the longer travel times of the matrix. As previously discussed, one might argue that the chaining method used to synthesize the missing values would overestimate the travel times for long trips by counting the terminal and local times for each link internal to the chain. The slight upward tilt of the data in Figure 4 shows that this has not happened. On average, long synthetic census times are slightly less than their network counterparts. This phenomenon is also reflected in the slight

TABLE 1 OBSERVED CENSUS VERSUS NETWORK AND AIRLINE DISTANCE-BASED TRAVEL TIMES

	Observed Census Versus	
	Highway Network	Airline Distance
Number of Observations	2,565	3,225
R^2	0.67	0.70
Difference between -		
Means (minutes)	- 1.07 (- 4.2%)	- 0.13 (0.5%)
Standard Dev. (minutes)	0.44 (3.4%)	3.58 (28.0%)
RMS Difference (minutes)	7.89 (30.7%)	9.00 (35.4%)
Theil Tests -		
UM	0.02	0.00
US	0.00	0.16
UC	0.98	0.84

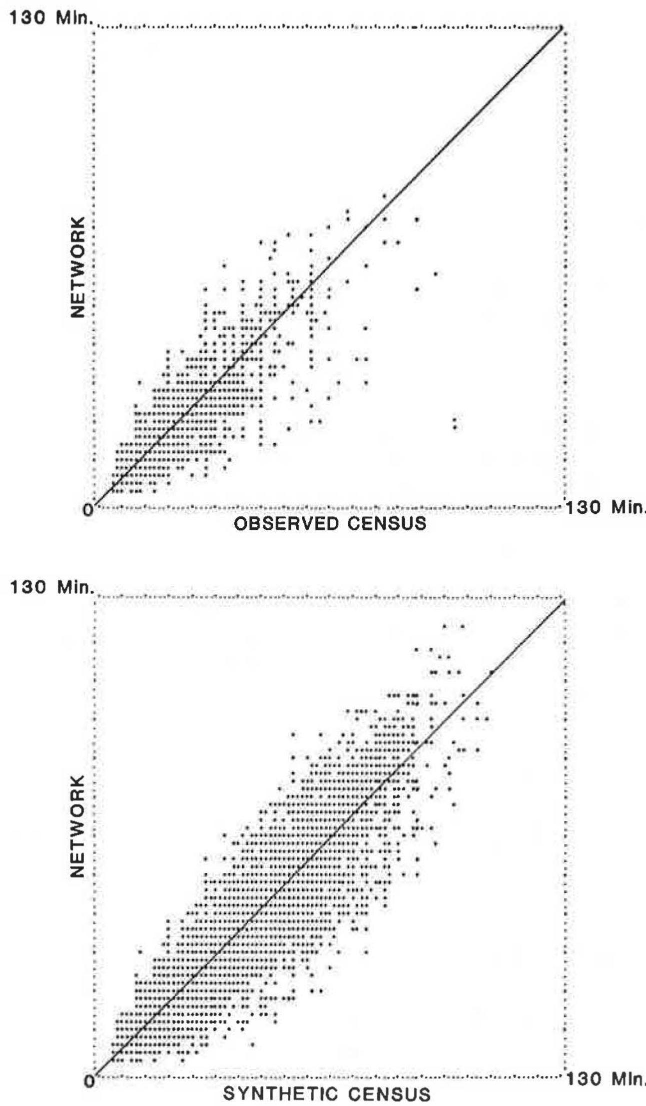


FIGURE 6 Observed census and synthetic census versus network-based travel times.

TABLE 2 SYNTHETIC CENSUS VERSUS NETWORK AND AIRLINE DISTANCE-BASED TRAVEL TIMES

	Synthetic Census Versus			
	Highway Network		Airline Distance	
Number of Observations	8,742		12,320	
R^2	0.84		0.85	
Difference between -				
Means (minutes)	- 0.64	(- 1.6%)	6.38	(15.5%)
Standard Dev. (minutes)	1.97	(10.4%)	9.61	(50.8%)
RMS Difference (minutes)	8.49	(20.9%)	14.65	(35.7%)
Theil Test -				
UM	0.01		0.19	
US	0.05		0.43	
UC	0.94		0.39	

increase in the US Theil statistic given in Table 2. This apparent overestimation of travel times by the network may result from a tendency to survey links in developed areas where network grain is finer and links more numerous, thereby imparting a small downward bias in the speeds shown in the table for unsurveyed links.

Except for the standard deviation, the statistical measures of the difference between the synthetic census and the network are essentially the same as the comparisons with the observed census data. This is strong indirect evidence that the accuracy of the synthetic part of the census matrix is comparable to the observed census travel times.

The comparisons between airline distance and synthetic census show a significant increase in the difference measures principally as a result of the strong tendency of the airline technique to overestimate long-distance travel times. A constant effective airline speed is inadequate. Effective airline speeds should be higher for long-distance travel and lower for shorter trips. Another use for the synthetic census data might be to estimate more accurate effective airline speeds. Such a speed could be calculated for the appropriate MCD interchange from the census time and airline distance matrices. This speed might then be applied to the precise airline distance in areas where MCDs are too large to supply the required geographical precision.

CONCLUSIONS

With the chain-building techniques explored here, one can synthesize a full matrix of MCD travel times from the 5 percent of these interchanges provided by the 1980 Census in the detailed journey-to-work data. Although incompatible with most transportation-planning software batteries because of the density of the spider network representing observed census times, missing census times can be synthesized through minimum path-building techniques in a cost-effective and accurate manner.

Comparisons between synthetic census and traditional transportation network travel time survey methods indicate that the synthetic census results are generally reasonable and representative of over-the-road travel times. Airline distance–based travel time estimates were found to be inaccurate because of a constant airline speed used to convert distance to over-the-road time. Overall speeds are slower for short movements and faster for longer trips, and vary significantly by direction of travel.

The process explored in the paper is capable of producing accurate, fine-grained travel time estimates from census data for large areas of the United States without the expense and difficulty associated with coding realistic highway networks and estimating link speeds through floating car surveys. This process, however, requires that adequate interdistrict connectivity be available in the census travel time observations. For the roughly 8 percent sample from the 1980 Census, work times must be aggregated to the MCD level (12 square miles on average) to achieve this. The sample size and the geographical coverage of the 1990 Census journey-to-work data will determine the grain and extent of the 1990 travel time matrix that can be generated with these methods.

REFERENCES

1. Highway Capacity Manual. *Highway Research Board Special Report 87*, Washington, D.C., 1965.
2. Highway Capacity Manual. *TRB Special Report 209*. TRB, National Research Council, Washington, D.C., 1985.
3. A. Talvitie and Y. Delghani. Comparison of Observed and Coded Network and Travel Time Measurements. *Transportation Research Record 723*, TRB, National Research Council, Washington, D.C., 1979, pp. 46–51.
4. D. E. Boyce and R. W. Eash. The Effect on Equilibrium Trip Assignment of Different Link Congestion Functions. *Transportation Research, Part A: General*, Vol. 15a, No. 3, pp. 223–232.
5. *Urban Transportation Planning Package: 1980 Census Technical Documentation for Summary Tape*, U.S. Bureau of the Census, Washington, D.C., 1983, p. H-4.
6. 1971–73 to 1986 Highway Travel Time and Speed Comparisons in the New Jersey Portion of the Delaware Valley Region. *Transportation Data Bulletin No. 9*, Delaware Valley Regional Planning Commission, Philadelphia, Pa. 1987, p. 10.
7. P. N. Fulton. Procedure Used for Allocating Incomplete Place of Work Responses in the 1980 Census Urban Transportation Planning Package. *Transportation Research Record 981*, TRB, National Research Council, Washington, D.C., 1984, pp. 15–20.
8. H. Theil. *Applied Economic Forecasting*. Rand McNally, Chicago, Ill., 1966, pp. 29–36.

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

Recursive Model System for Trip Generation and Trip Chaining

KONSTADINOS G. GOULIAS AND RYUICHI KITAMURA

A model system is developed to describe both trip generation and trip chaining in a coherent manner. A recursive structure is adopted to represent the generation of trips for different purposes, and the number of trip chains is expressed as a function of the numbers of trips by purpose. The model system offers theoretically consistent coefficient values and quantifies the relationship between the number of trips and the number of trip chains, and can be used in the conventional forecasting procedure in place of home-based and non-home-based trip generation models. This model is applied to examine how trip chaining patterns vary across sample subgroups. The results indicate no significant variations in trip chaining behavior across car ownership subgroups. It is inferred that car ownership influences household trip generation, but given the number of trips generated, the number of trip chains is not influenced by car ownership.

The importance of trip chaining—linking of trips to visit more than one destination after leaving home—has been discussed extensively in the travel behavior literature (*1–3*). The spatial distribution of trip ends and trip timing, as well as the total number of trips, vary substantially depending on the way trips are linked to each other. Empirical evidence indicates that the destination of a non–home-based trip is heavily influenced by the location of the home base (*4*). Because consolidating trips is one of the schemes that can be used to reduce travel time and other resources expended to pursue out-of-home activities (*5*), it is likely that urban residents' trip chaining behavior will change over time as travel cost, congestion, land use patterns, and other contributing factors change.

Only limited knowledge exists on how a set of trips made by a trip-maker on a given day will be combined into trip chains. Obviously this is a complex process that involves many objectives, alternatives, and constraints. Minimizing travel distance is perhaps just one of the objectives that a trip-maker attempts to achieve. Many constraints are often not identifiable from typical home interview travel survey results, and feasible alternatives that are considered by the trip-maker are usually unknown.

The approach taken in the conventional four-step procedure is to estimate the number of home-based trips and non–home-based trips separately, then distribute the two types of trips as entirely unrelated entities. This approach unfortunately does not lead to a causal model of trip chaining behavior through which future travel patterns can be inferred under alternative scenarios, such as intensive suburban land use development, resulting in suburban congestion, or increasing gasoline prices. Several models have been proposed that describe trip chaining behavior in simplified contexts (*6–10*). Although these models draw on certain causal structures and capture salient behavioral relationships, their application to demand forecasting has been difficult because they either apply only to limited and simplified cases, or require excessive effort to generate supporting data bases.

Department of Civil Engineering, University of California at Davis, Davis, Calif. 95616.

In this study, an attempt is made to develop a model system that describes trip generation and trip chaining at the household level. The analysis is conceived as an initial step toward a practical forecasting procedure that explicitly incorporates trip chaining. A model structure that may be termed "recursive" (*11*) is developed for the generation of trips for different purposes and formation of trip chains. The components of the model system are estimated using empirical data. Here we show the interrelationship among the number of trips by purpose and the number of trip chains, then use the model system to determine how trip chaining behavior varies across sample subgroups.

A short note is due on the use of the term, "recursive." In the terminology of simultaneous equations systems, a recursive system is that special case of structural equations that can be arranged into a system that involves only unidirectional cause-effect relations (*12,13*). Such a system presents a hierarchical, or sequential, structure in which predetermined variables determine the first endogenous variable, then the first endogenous variable and possibly other predetermined variables determine the second endogenous variable, and so on (*12*). This structure is applied in this study to trip generation by purpose and formation of trip chains.

This paper is organized as follows: The modeling approach of the study is discussed in the next section, and then the data set used is described. A set of recursive trip generation models by purpose is presented, as well as a model of trip chain generation. Also presented are the results of model application we used to examine the variation in trip chaining behavior across sample subgroups, followed by a summary of the study.

MODELING APPROACH

The concept of trip generation is an important element of the model system of this study because it links the proposed model of trip chaining to the conventional demand forecasting procedure. The introduction of the concept also reflects the viewpoint that needs to engage in activities motivate the members of a household to make a set of out-of-home stops and that trip chains are formed given the set of stops to be made.

The model system of this study adopts the structure in which the number of stops made for a given type of activities is expressed as a function of household characteristics, and these stops are then combined into trip chains. The resulting model system consists of a set of trip generation models by purpose, and a trip chaining model that translates the number of trips into the number of trip chains.

The relationship among trips made for different purposes is emphasized in the model development. It is assumed that certain trips are compulsory while others are discretionary, depending on the types of activities for which they are made. In the analysis, work and school trips are assumed to be compulsory, and personal business, shopping, and social trips are considered to be discretionary. (Trips in the latter category may not be entirely discretionary. For example, consider a grocery shopping trip that must be made to prepare a meal. However, unlike compulsory trips, a large degree of flexibility is often associated with the frequency, timing, and destination locations of these trips.)

Given this dichotomy, we assume that if these two classes of trips are interrelated at all, then the presence and number of compulsory trips influence the presence and number of discretionary trips. It is also assumed that the number of trip chains is a function of the numbers of trips by purpose. This viewpoint leads to the model structure shown in Figure 1.

A unique feature of the system of trip generation models of this study is the recursive relationship assumed among them; the number of discretionary trips is expressed as a function of the number of compulsory trips. This contrasts sharply with the conventional approach in which trips of different purposes are estimated independently without assuming any internal relationship. The approach of this study is consistent with the notion of time budget (*14*); those who expend a substantial amount of time for mandatory activities such as work and school have less discretionary time available, therefore are likely to make fewer discretionary trips. The validity of this conjecture can be statistically tested by estimating the model system.

The number of trip chains is modeled as a linear function of the number of trips by purpose. The coefficient of each trip variable in the model then indicates the average number of trip chains that is generated per trip. Estimated coefficient values point to which activities tend to be linked together with other activities into multistop chains.

For example, consider the hypothetical example in which a trip-maker makes only one one-stop chain per day that involves a work stop, that is, home-work-home. Then the number of trip chains (which equals the number of home trips) is identical to the number of work trips, and if we let

$$(\text{No. of trip chains}) = \beta_1 (\text{No. of work trips}) \qquad (1)$$

then $\beta_1 = 1$. Now suppose the trip-maker makes another trip chain that involves two shopping stops, that is, home-shopping-shopping-home, in addition to the home-work-home chain. Then,

$$(\text{No. of trip chains}) = \beta_1 (\text{No. of work trips}) + \beta_2 (\text{No. of shopping trips}) \qquad (2)$$

which is satisfied with $\beta_1 = 1$ as before and $\beta_2 = 0.5$. In general, trips that tend to be combined into multistop chains will have a smaller coefficient, and trips that tend to be pursued in one-stop chains will have a coefficient closer to 1 (which is the theoretical upper bound).

These coefficients also serve as indicators of the propensity to form multistop trip chains that is shared by a group of households; households that tend to consolidate trips into fewer trip chains will have smaller coefficients, whereas those that tend to make one-stop chains will have coefficients closer to 1. In the following sections, the model system is estimated and the propensity to chain trips is evaluated for household subgroups.

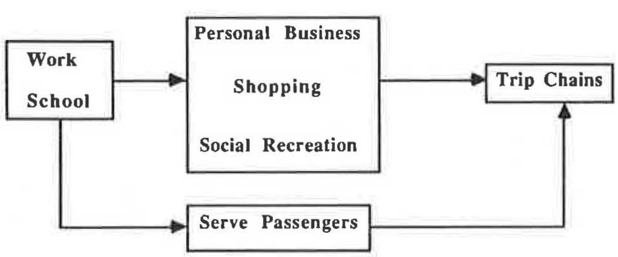


FIGURE 1 Hypothesized recursive structure involving trip generation by purpose and trip chaining.

SAMPLE FOR THE STUDY

A sample from the ongoing Dutch National Mobility Panel Survey data set is used in this study. There are two reasons for the use of this particular data set. First, weekly trip records, filled out by household members of 12 years old and over, are available from the survey. This offers reliable measures of trip generation and trip chaining that are less influenced by day-to-day variations in travel patterns. Second, use of the panel data set allows later extension of the analysis to dynamic analysis of trip chaining behavior. The analysis reported here, however, is strictly cross-sectional.

The households in the panel data set, which was intended to represent the Dutch population, were selected using a stratified sampling method based on household life cycle stage and income. The resulting sample households are scattered throughout the nation in 20 municipalities of various sizes.

Records from the first panel survey, conducted in April 1984, are used in the analysis here. All households are included in the analysis except those with missing variable values. [Discussions of the background of this panel survey and data characteristics can be found elsewhere (*14–17*).]

The set of explanatory variables used in the model development is shown in Table 1. These variables are divided into four groups. The first group consists of variables that represent household structure, which is believed to importantly influence household trip generation (*18*). In addition to the household size (HHLDSIZE), variables are included to represent the number of household members by age and sex. Also included in this group is the number of household members who filled out the weekly travel diaries (NDIARIES), which in most cases equals to the number of household members of 12 years old and over.

TABLE 1 VARIABLES USED IN MODEL FORMULATIONS

VARIABLE	DEFINITION
Household Demographics	
HHLDSIZE	Number of persons in the household
NDIARIES	Number of household members who filled out the diary
NCHILDREN	Total number of children living in the household
NCHILD:0-6	Number of children of 0 to 6 years old
NCHILD:7-11	Number of children of 7 to 11 years old
NCHILD:12-17	Number of children of 12 to 17 years old
NCHILD:18-	Number of children of 18 years old and over
NADULTS	Number of adults (> 18 years old) in the household
NMEN	Number of adult male household members
NWOMEN	Number of adult female household members
Household Socioeconomics	
NWORKERS	Number of employed persons in the household
HIEDUCATN	1 if the household member with the highest level of education has a college degree
LOEDUCATN	1 if the household member with the highest level of education have completed only elementary school
√INCOME	Square-root of annual gross household income (Dfl) divided by
√MAXPINCOME	Square-root of the annual gross income (Dfl) of the major breadwinner, divided by 100
Car Availability	
NCARS	Number of cars available to the household
ONECAR	1 if exactly one car is available to the household
TWOCARS	1 if two or more cars are available to the household
NDRIVERS	Number of licensed drivers in the household
NONDRIVERS	Number of household members (>12 years) who are not licensed to drive
Residence City Class	
LARGECITY	1 if the household resides in a large metropolitan area with highly developed multi-mode transit systems
RURALAREA	1 if the household resides in a community that is not served by rail

The variables in the second group, socioeconomic attributes, include the number of workers (NWORKERS), income measures (√INCOME and √MAXPINCOME) and levels of education (HIEDUCATN and LOEDUCATN). The use of the square root of income is based on the results of previous studies using the same data set (*17, 19*). The education level of the household member with the highest education is used to define these education variables. Car availability is represented by the number of cars available to the household (NCARS) and two dummy variables (ONECAR and TWOCARS) to account for nonlinear effects. The number of licensed drivers and nondrivers (NDRIVERS and NONDRIVERS) is also included to represent the use of, and competition for, family cars. The two variables in the last group (LARGECITY and RURALAREA) are measures of the residence city size and an indicator of transit service levels.

TRIP GENERATION MODELS

A set of trip generation models is developed following the recursive structure of Figure 1 and using the variables shown in Table 1. Their dependent variables are the weekly total numbers of trips in the respective purpose categories reported by the diary-keepers in the household.

The general form of the models of this study is

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i = \beta' X_i + \varepsilon_i$$

where Y_i is the number of trips reported by household i, the β's are coefficients, the X's are explanatory variables, $\beta' = (\beta_0, \beta_1, \ldots, \beta_k)'$, $X_i = (1, X_{i1}, X_{i2}, \ldots, X_{ik})'$, and ε_i is a random error term. All models are estimated using weighted least squares regression, with the weight formulated as $\theta(1 + |Y_i|)^\tau$, where Y_i is the predicted number of trips for household i, and θ and τ are estimated by regressing the squared residual on the predicted number of trips.

The set of trip purpose categories used in this analysis includes: work, school, shopping, social, and serving passengers. Shopping trips include personal business, and social trips include recreational and trips for meals. Work-related business trips are not included in the analysis of this study. The results of model estimation are presented by purpose in the next section.

Work Trip-Generation Model

The weekly work trip generation model shows a good fit, explaining 56 percent of the total variation in weekly household work trip generation (Table 2). The most significant variable is the number of workers in the household (NWORKERS). The frequency of work trips is also influenced by the household size and gender composition (HHLDSIZE, NWOMEN, and NMEN).

The education and income variables (HIEDUCATN, LOEDUCATN, and √INCOME) indicate that households with higher education or higher income tend to make (or report) more work trips. These variables are multiplied by the number of workers (NWORKERS). Therefore their coefficients represent the impact of education or income on the number of work trips per worker. The city size variables (LARGECITY and RURALAREA) indicate that, other things being equal, households in large cities or in rural areas tend to generate fewer work trips.

However, the coefficients of the variables representing education, income, and residence area are statistically not significant. This model development effort indicates that work trip generation is primarily determined by the number of workers and other household demographic attributes and is not significantly influenced by education, income, and residence area, despite their likely correlation with the type of employment.

School Trip-Generation Model

The number of school trips generated by a household is determined primarily by the number of children (NCHILD:12–17 and NCHILD:18–) whose coefficients are extremely significant (Table 3). The coefficient of the number of children between 12 and 17 years old (NCHILD:12–17), who are practically all students, is very close to the number of school days in a week. Recall that household members below 12 years old were not requested to fill out the diary; otherwise the number of younger children would also have entered the model.

TABLE 2 WORK TRIP-GENERATION MODEL

Variable	β	t
NWORKERS	3.933	9.12
NWOMEN	.416	2.49
NMEN	.918	6.30
HHLDSIZE	.282	3.30
NONDRIVERS	-.457	-4.73
HIEDUCATN*NWORKERS	.234	1.14
LOEDUCATN*NWORKERS	-.349	-1.31
√INCOME*NWORKERS	.225	1.09
LARGECITY*NWORKERS	-.106	-.34
RURALAREA*NWORKERS	-.431	-1.32
Constant	-.355	
R^2	.557	
F	216.8	
df	(10,1728)	
N	1739	

β = Estimated model coefficient
t = t-statistic

TABLE 3 SCHOOL TRIP-GENERATION MODEL

Variable	β	t
NWORKERS	-.444	-6.27
NCHILD:12-17	5.378	34.16
NCHILD:18-	1.439	11.40
√INCOME*NDIARIES	.083	2.35
LOEDUCATN*NDIARIES	-.219	-4.23
HIEDUCATN*NDIARIES	.533	9.95
Constant	.566	
R^2	.546	
F	347.3	
df	(6,1732)	
N	1739	

β = Estimated model coefficient
t = t-statistic

Unlike the case of the work trip-generation model, the education and income variables have significant coefficients with anticipated signs. The education coefficients indicate that the school trip rate differs by 0.752 [0.533 − (−0.219)] trip per diary-keeper between households with the highest and lowest education levels. The significance of the education variables, however, may be in part due to the fact that the household education level may be determined by that of a child; in certain cases a higher household education level is attributed to the presence of children pursuing higher degrees. The possibility exists that the coefficients of the education variables reflect differences in trip reporting as well as actual trip making; it is likely that members of households with higher education levels tend to complete trip diaries more accurately without leaving out trips that were actually made (*16*).

Shopping Trip-Generation Model

The number of adult household members is among the significant variables that contribute to shopping trip generation (Table 4). The coefficient of the number of female adults

TABLE 4 SHOPPING TRIP-GENERATION MODEL

Variable	β	t
NWOMEN	2.376	8.57
NMEN	1.949	7.21
HHLDSIZE	.307	2.82
√INCOME*NDIARIES	.247	2.78
NONDRIVERS	-.567	-4.10
LOEDUCATN*NDIARIES	-.274	-2.35
HIEDUCATN*NDIARIES	.644	6.78
RURALAREA*NDIARIES	-.577	-3.61
Y(work)	-.257	-6.71
Constant	2.066	
R^2	.281	
F	1291.6	
df	(9,1729)	
N	1739	

β = Estimated model coefficient
t = t-statistic

(NWOMEN) is substantially larger than that of the number of male adults (NMEN). The difference, which is statistically significant at $\alpha = 0.05$, indicates that an adult woman of a household tends to make 0.4 more shopping trip per week than does an adult man. The coefficient of the number of nondriving diary-keepers is negative and significant, suggesting that shopping trip generation is correlated with automotive mobility.

As anticipated, the income coefficient indicates that shopping trip generation increases with income. The coefficients of the education variables are again significant and suggest that a diary-keeper from a household with the highest education level tends to make 0.91 more shopping trip per week than does his or her counterpart from a household with the lowest education level. As in the case of work and school trips, however, this difference reflects different degrees of reporting accuracy as well as genuine differences in trip rates. This shopping trip–generation model includes the number of work trips as one of the explanatory variables. The coefficient of this variable was estimated using the predicted number of work trips obtained from the work trip–generation model as an instrument. The estimated coefficient implies negative correlation between work trip generation and shopping trip generation. This is consistent with the conjecture postulated earlier; work trip generation implies that a considerable amount of time is spent by household members for mandatory work activities, leaving less discretionary time available and leading to a lower level of discretionary trip generation. The estimated coefficient value of -0.26 suggests that approximately 1.3 less shopping trips per week will be generated by the household for each household member who is employed (the number of school trips is not included as an explanatory variable of the model because school trips are made primarily by the children of a household, who tend not to make shopping trips).

Social Trip–Generation Model

The most significant variable in the social trip generation model is the number of diary-keepers (NDIARIES). The number of children (NCHILD:12–17 and NCHILD:18-) also has a positive contribution (Table 5). The education variables (LOEDUCATN and HIEDUCATN) indicate similar effects as before. The table shows that, like shopping trip generation, social trip generation is negatively correlated with work trip generation.

The variables associated with car availability and use (ONECAR, TWOCARS, and NONDRIVERS) are all significant and indicate that social trip generation is influenced by automotive mobility. No car ownership variables were used in the work, school, and shopping trip–generation models because they were insignificant.

This significance of the car ownership variables in the social trip–generation model is presumably because social trips are least mandatory among these trip purpose categories. Generation of mandatory trips will be determined by external factors and their frequency will be independent of the relative ease of trip making. Alternatively, generation of discretionary trips is regulated by the household to a larger extent and therefore is influenced more significantly by the ease of travel and the mobility resources available to the household. The results of this study support this conjecture.

TABLE 5 SOCIAL TRIP–GENERATION MODEL

Variable	β	t
NDIARIES	2.376	8.57
NCHILD:12-17	.855	1.99
NCHILD:18-	.698	1.82
NONDRIVERS	-.785	-3.15
ONECAR*NDIARIES	.597	3.02
TWOCARS*NDIARIES	.687	2.72
LOEDUCATN*NDIARIES	-.275	-1.69
HIEDUCATN*NDIARIES	.549	4.23
RURALAREA*NDIARIES	-.131	-.59
Y(work)	-.249	-5.07
Constant	1.466	
R^2		.324
F		83.0
df		(10,1728)
N		1739

β = Estimated model coefficient
t = t-statistic

Serve-Passenger Trip–Generation Model

The structure of this trip generation model is substantially different from those of the models presented previously, as it reflects the unique nature of serve-passenger trips, that is, they are made to fulfill activity needs of other individuals, quite often other household members. Obviously very few, if at all, serve-passenger trips will be generated by households to which no automobile is available or that have no drivers. It can also be inferred that, given automobiles are available, fewer serve-passenger trips will be generated by single-person households, whereas households with nondrivers or children will on average generate more trips of this type. The model presented in Table 6 reflects these considerations.

Most of the explanatory variables in this model are mul-

TABLE 6 SERVE-PASSENGER TRIP–GENERATION MODEL

Variable	β	t
NDRIVERS*MPHH	1.21	4.73
NCHILDREN	1.24	13.51
LOEDUCATN*NDRIVERS*MPHH	-.28	-1.78
√INCOME*NDRIVERS*MPHH	-.04	-.37
Y(work)*ONECAR*MPHH	.06	1.80
Y(work)*TWOCARS*MPHH	-.15	-2.37
Y(shop+social)*ONECAR*MPHH	-.08	-4.78
Y(shop+social)*TWOCARS*MPHH	-.05	-1.68
Constant	.48	
R^2		.17
F		44.39
df		(8,1730)
N		1739

MPHH = 1 if HHLDSIZE > 1
β = Estimated model coefficient
t = t-statistic

tiplied by a dummy variable (MPHH), which takes on a value of one for multiperson households. This specification has been chosen over several alternative model specifications in which effects of the explanatory variables are not differentiated between single- and multiperson households. In addition to MPHH, the education and income variables (LOEDUCATN and $\sqrt{}$INCOME) are multiplied by the number of drivers (NDRIVERS), and predicted numbers of trips [Y(work) and Y(shop + social)] by car ownership dummy variables (ONECAR and TWOCARS).

The number of children (NCHILDREN) is highly significant and positively contributes to the number of serve-passenger trips. This result supports the conjecture that serve-passenger trips are often made for children in the household. The table also shows the anticipated result that the number of drivers (NDRIVERS*MPHH) positively contributes to the number of serve-passenger trips. As in the previous models, households with lower levels of education tend to make (or report) fewer trips in this category, but income is not significantly associated with serve-passenger trips.

The work trip instrument variable in one-car households [Y(work)*ONECAR*MPHH] has a positive coefficient, while that in multicar households [Y(work)*TWOCARS*MPHH] has a significant negative coefficient. The result suggests that fewer serve-passenger trips are made in connection with work trips by multicar households. The shopping and social trip instrument variable [Y(shop + social)] has a negative coefficient regardless of the car ownership level. Overall, the estimation results have shown that the number of children is the most significant variable influencing serve-passenger trip generation.

TRIP CHAIN MODEL

The trip chain model of this study represents the relationship between the number of trips by purpose and the number of trip chains made by household members. The model, shown in Table 7, is simple, consisting of three instrument variables [Y(work), Y(school), and Y(shop + social)]. No constant term is included because no trip chains can be made when no trips are made.

The coefficients of the trip generation instrument variables are measures of how trips of the respective categories are formed into trip chains. The coefficient of Y(work), 0.897, implies that a work trip on average generates 0.897 trip chain. The smaller coefficient (0.800) of Y(school), on the other hand, suggests that school trips are linked to other trips into multistop chains more frequently. The coefficient of Y(shop + social), which is very close to one, suggests that practically all trips of this category are made in one-stop chains.

TABLE 7 TRIP CHAIN MODEL

Variable	β	t
Y(work)	.897	15.15
Y(school)	.800	10.49
Y(shop+social)	.994	32.76
R^2	.887	
F	4560.9	
df	(3,1736)	
N	1739	

β = Estimated model coefficient
t = t-statistic

The theoretical upper bound of these coefficients is 1.0, because a trip cannot generate more than one trip chain. The estimation result is logical with none of the coefficient estimates exceeding this theoretical maximum, and all coefficients are positive and significant.

This trip chain model is now used to examine possible variations in trip chaining behavior across sample subgroups. The focus of the analysis is on the tendency in consolidating trips for various purposes into trip chains, and how the tendency varies across subgroups. Subgroups defined in terms of income, car ownership, and city size are examined in the analysis. The results are summarized in Table 8 in terms of the estimated coefficients for the respective subgroups. Also presented in the table are F-statistics to test the significance of the variation in the coefficient values across the subgroups.

The F-statistics indicate that the coefficient vectors are significantly different across income subgroups and city-size subgroups. The coefficients of the three income subgroups indicate that the high-income group tends to make work trips and discretionary (shopping and social) trips in multistop chains more frequently, and school trips by themselves in single-stop chains. No differences are appreciable in the coefficient estimates between the low and medium income subgroups.

The result that high-income households tend to consolidate work and discretionary trips into multistop chains is consistent with previous findings (*3*). However, this tendency does not uniformly apply to all trip purposes. The result of this analysis implies that patterns of trip chaining must be examined by trip purpose.

The coefficients of the trip chain models are also significantly different across the city-size subgroups. The estimation results indicate that residents of large cities tend to make discretionary trips in multistop chains. The coefficients of the work trip instrument are stable across the subgroups, whereas those of the school trip instrument suggest that school trips are linked to others more often in both small and large urban areas.

The coefficient of the school trip instrument is extremely small for the no-car subgroup. This, however, did not lead to a significant F-statistic for the car ownership subgroups. In fact all of the other coefficients are within 8 percent of the coefficient estimates for the entire sample shown in Table 7. This result suggests that trip chaining behavior may not be as much influenced by car ownership as it has been believed to be; the effects of car ownership on trip chaining observed previously may have been caused by the effect of car ownership on trip generation rather than on trip chaining. The validity of this conjecture must be tested in further investigation.

The estimated coefficient values shown in Table 8 are again theoretically supportable, lying between 0 and 1 for most cases. A few estimates are above the upper bound of 1.0, but only by small amounts, and in no cases do they exceed the bound with statistical significance. This consistency in the coefficient estimates is extremely encouraging, especially in

TABLE 8 TRIP CHAIN MODELS BY INCOME, CITY SIZE, AND CAR OWNERSHIP SUBGROUPS

Income	Low (≥$21,000)		Middle		High (≤$37,000)		
Variable	β	t	β	t	β	t	Difference
Y(work)	1.027	8.73	1.076	8.84	.861	8.11	
Y(school)	.736	6.64	.770	5.48	1.060	6.53	
Y(shop+social)	.997	25.39	.984	15.85	.900	12.99	
R^2	.857		.888		.906		
F	1396.8		1465.7		1547.3		3.081*
df	(3,697)		(3,553)		(3,480)		(6,1730)
N	700		556		483		1739
City Size	Small		Medium		Large		
Variable	β	t	β	t	β	t	Difference
Y(work)	.879	2.88	.902	13.86	.858	6.07	
Y(school)	.626	1.78	.802	9.66	.665	3.14	
Y(shop+social)	1.092	6.35	1.004	30.31	.839	11.82	
R^2	.857		.891		.886		
F	234.5		3924.7		457.0		3.246*
df	(3,117)		(3,1437)		(3,176)		(6,1730)
N	120		1440		179		1739
Car Ownership	No Car		One Car		Two Cars		
Variable	β	t	β	t	β	t	Difference
Y(work)	.885	7.99	.862	10.76	.933	5.15	
Y(school)	.393	2.48	.864	8.96	.758	3.07	
Y(shop+social)	.982	21.39	1.011	25.06	.978	8.42	
R^2	.821		.889		.918		
F	690.3		1465.7		737.4		1.582
df	(3,452)		(3,1080)		(3,198)		(6,1730)
N	455		1083		201		1739

β = Estimated model coefficient
t = t-statistic
Note: None of the estimated coefficients β is significantly larger than the theoretical upper bound of 1.0.
* Significant at α=0.05.

light of the small sample sizes of some of the subgroups. The study results also offer an encouraging indication that this simple trip chain model can be used to examine the characteristics of trip chaining behavior by population subgroups.

CONCLUSION

An attempt has been made in this study to develop a model system that describes both trip generation and trip chaining in a coherent manner. The model system adopts a recursive structure in representing the generation of trips for different purposes. The number of trip chains is expressed as a function of the numbers of trips by purpose. The estimated coefficients of the recursive models have indicated the presence of negative correlation between mandatory and discretionary trips, suggesting the influence of time budgets. The trip chain model has offered theoretically consistent coefficient values and quantified the relationship between the number of trips by purpose and the number of trip chains. This model has been applied to examine how trip chaining behavior varies across sample subgroups. The results indicate that variations in trip chaining exist among income subgroups and city-size subgroups, but not among car ownership subgroups.

The finding that higher income households tend to consolidate trips into multistop chains is consistent with earlier findings. However, the analysis of this study has shown that not all trips are consolidated at the same rate; it has been shown that school trips are combined with other trips less frequently by the high-income group. Such associations between trip chaining and trip purposes must be properly reflected in models of trip chaining behavior.

No significant variations in trip chaining behavior have been found across car ownership subgroups in this study. An important conjecture stems from this: Car ownership influences household trip generation, but given the number of trips generated, the number of trip chains is not influenced by car ownership. This conjecture is worthy of further examination in the future.

This study has shown that a model system can be successfully developed integrating trip generation and trip chaining. The theoretically supportable coefficient estimates of the trip chain model are extremely encouraging. The proposed model system can be incorporated almost immediately into the conventional travel-demand forecasting procedure. The trip generation models indicate the total number of trips made by trip purpose, including both home-based and non-home-based trips. A procedure can be developed to classify a predicted number of trips into these two types based on the coefficients of the trip chain model. The model system will then offer the same forecasts as does the conventional set of home-based and non-home-based trip generation models, in a logically coherent manner. However, the model system presented here is in its initial stage of development. Its validation and careful refinement remain as a future task.

ACKNOWLEDGMENTS

This research was supported by a National Science Foundation grant to the University of California at Davis. The data set used was obtained from the Dutch National Mobility Panel Survey results. The survey was sponsored by the Projectbureau for Integrated Traffic and Transportation Studies and the Directorate General of Transport, Dutch Ministry of Transport and Public Works.

REFERENCES

1. S. Hanson. Urban-Travel Linkages: A Review. In *Behavioral Travel Modelling* (D. Hensher and P. Stopher, eds.), Croom Helm, London, 1979, pp. 81–100.
2. J.-C. Thill and I. Thomas. Towards Conceptualising Trip Chaining Behaviour: A Review. *Geographical Analysis*, Vol. 19, 1987, pp. 1–17.
3. R. Kitamura. Recent Developments in Trip Chaining Analysis. Presented at the 15th PTRC Transport and Planning Summer Annual Meeting, University of Bath, England, September 1987.
4. R. Kitamura and M. Kermanshah. A Sequential Model of Interdependent Activity and Destination Choice. *Transportation Research Record 987*, TRB, National Research Council, Washington, D.C., 1984, pp. 81–89.
5. S. P. Pfifer, A. J. Neveu, and D. T. Hartgen. Family Reactions to Energy Constraints. *Preliminary Research Report 161*, New York State Department of Transportation, Albany, NY, 1979.
6. T. Adler and M. Ben-Akiva. A Theoretical and Empirical Model of Trip Chaining Behavior. *Transportation Research B*, Vol. 13B, 1979, pp. 243–57.
7. J. Horowitz. A Utility-Maximizing Model of the Demand for Multidestination Non-work Travel. *Transportation Research B*, 1980, Vol. 14B, pp. 369–86.
8. D. Damm. Parameters of Activity Behavior for Use in Travel Analysis. *Transportation Research A*, Vol. 16A, 1982, pp. 135–48.
9. K. Kondo and R. Kitamura. Time-Space Constraints and the Formation of Trip Chains. *Regional Science and Urban Economics*, Vol. 17, pp. 49–65.
10. K. Nishii, K. Kondo, and R. Kitamura. An Empirical Analysis of Trip Chaining Behavior. *Transportation Research Record* (forthcoming).
11. H. O. A. Wold. *Econometric Model Building, Essays on the Causal Chain Approach*. North-Holland, Amsterdam, 1964.
12. J. L. Murphy. *Introductory Econometrics*. Irwin, Homewood, Ill., 1973.
13. H. Theil. *Principles of Econometrics*. John Wiley and Sons, New York, N.Y., 1971.
14. P. M. Allaman, T. J. Tardiff, and F. C. Dunbar. *New Approaches to Understanding Travel Behavior*. National Cooperative Highway Research Program Report 250, TRB, National Research Council, Washington, D.C., 1982.
15. J. Golob, L. Schreurs, and J. Smit. The Design and Policy Applications of a Panel for Studying Changes in Mobility Over Time. In *Behavioral Research for Transport Policy*, VNU Science Press, Utrecht, The Netherlands, 1985, pp. 81–95.
16. T. F. Golob and H. Meurs. Biases in Response Over Time in a Seven-Day Travel Diary. *Transportation*, Vol. 13, 1986, pp. 163–81.
17. R. Kitamura and P. H. L. Bovy. Analysis of Attrition Biases and Trip Reporting Errors for Panel Data. *Transportation Research A*, Vol. 21A, 1987, pp. 287–302.
18. R. Kitamura and T. van der Hoorn. Regularity and Irreversibility of Weekly Travel Behavior. *Transportation*, Vol. 14, 1987, pp. 227–51.
19. R. Kitamura. Lifestyle and Travel Demand. In *A Look Ahead: Year 2020, Special Report 220*, TRB, National Research Council, Washington, D.C., 1988, pp. 149–189.

Publication of this paper sponsored by Committee on Traveler Behavior and Values.