9-1989

Site Location Study for an Integrated Traffic Data Collection System: Phase 1 Report

Kenneth Dueker
Portland State University

Richard Ledbetter
Portland State University

Bruce Rex
Portland State University

SITE LOCATION STUDY FOR AN INTEGRATED TRAFFIC DATA COLLECTION SYSTEM: PHASE 1 REPORT

Kenneth J. Dueker
Richard Ledbetter
Bruce Rex

July, 1989
(Revised September, 1989)

A report to the Oregon Department of Transportation

Center for Urban Studies
School of Urban and Public Affairs

PHASE I FINDINGS AND RECOMMENDATIONS

Introduction

Improved technology for counting, classifying, and weighing vehicles in the highway traffic stream is becoming increasingly available. Research has been undertaken to make effective use of these new technologies, not just to replace current data collection programs of counting, classifying, and weighing with more efficient versions. This reports on the Phase I research, which had three major objectives;

- a literature review to guide the analysis of alternative methods;
- a statement of requirements to modernize and integrate the data collection programs of traffic counting, vehicle classification and truck weights, and to identify the type and use of traffic data for design, planning maintenance, enforcement and vehicle taxation; and
- analysis of alternative site selection methods to meet the locational requirements of an integrated traffic data collection system.

The following sections report on progress toward these objectives and provide detailed recommendations for continuation of the research into Phase II.

Literature Review

The results of the Literature Review task are reported separately (see Attachment 1). The literature review became an on-going task throughout Phase I. The process of literature review guided the research and we were able to apply what was learned in various other states to Oregon, but no state was sufficiently comprehensive in their approach. Only Florida addressed what we came to feel was the most important issue - how to infer from continuous data collection locations to other points on the highway system. Yet, we were able to learn from Wisconsin, Washington, and Virgina a considerable amount about clustering, integrating, and factoring counts, respectively. We also learned from experience in Israel that elegant network approaches did not yield very good results.

The lack of a comprehensiveness and consensus in the literature forced us to perform a more empirical analysis in Phase I than was originally anticipated.

Requirements for Integrated Data Collection

Modernizing the OSHD traffic data collection equipment consists of replacing old and obsolete equipment that requires expensive maintenance and repairs. However, it may not be appropriate to merely replace new equipment for old. A more integrated data collection system is needed and care concerning siting is warranted. This research addresses these issues.

Integrated data collection is encouraged by FHWA in the procedures embodied in the HPMS and TMG. However, the purpose of the HPMS and TMG is to aggregate upward from the sample points to the universe and strata of the universe in terms of functional class, urban/rural, and volume class. Sampling strategies in the TMG are based on nested random selection of HPMS segments. A larger sample is needed for traffic counting. A subsample is selected for vehicle classification. A subsample of this subsample is selected for collecting truck weight data. This approach may be appropriate for aggregating up to make reliable statements about the universe and various substrata, but it may not serve State needs well for making inferences about traffic at point locations. Point location data are needed for design and construction zone traffic management. Continuous data collection is needed to factor less frequently collected data at Coverage Count locations.

It is recommended that OSHD implement a program of integrated traffic data collection, consisting of a set of ATRs to collect traffic data continuously. These continuously collected data should be used to factor counts, classification, and weights collected to meet the requirements of HPMS, and to meet the needs of the state for factoring data collected in the Coverage Count program and manual count locations.

This recommendation is supported by the conduct of empirical analysis, which was not originally planned for Phase I, but became necessary due to the lack of consistent findings from the literature review. We felt the need to create a database and perform analysis of data from the ATRs. Analysis of ATR data indicates the following:

- a clustering of ATRs by volume indicates that five clusters are needed to reduce the coefficient of variation and that additional clusters do not add much in reducing the measure (Attachment 2);
- a clustering of ATRs by seasonal factor, design hour volume, or percent heavy trucks does not yield logical spatial groups nor logical type of highway groups (Attachment 3);
- a clustering of ATRs by log of AADT, urban/rural type, and functional class yields five logical groups (Attachment 4); and
- use of multiple linear regression in lieu of cluster analysis is not complete, additional model specification will be done in Phase II (Attachment 5).

Site Selection Methodology

The Study Design specified that the Detailed Workplan (Task 4) would involve an Analysis of Alternative Methods (Phase I), consisting of a Literature Review (Task 1), a Statement of Requirements (Task 2), and a Site Selection Methodology (Task 3). The literature review and the analysis of requirements for the type and use of traffic data necessitated a change in the conduct of the site selection methodology task and consequently, affects the detailed workplan task. The lack of direction from the literature review led to earlier than anticipated empirical analysis to determine whether spatial clustering of seasonal, design hour and heavy truck factors exist. We did not find evidence of spatial clustering of factors. This, coupled with the finding in Israel that network analysis was not fruitful, has led to major reconsideration of relying on spatial clustering and network analysis. Consequently, the recommendations for siting integrated traffic data collection equipment differ from what was originally anticipated.

The Phase I empirical analysis leads to the following recommendations:

- Cluster ATRs into five groups based on log AADT, U/R class, and functional class;
- Use the cluster means of seasonal, design hour, growth, heavy trucks factors to apply to HPMS, Coverage Count, and other locations, of the same volume, U/R, and functional class;
- Research an alternative to the above described method. The alternative would use a

weighted mean of the three ATRs of the same cluster type that triangulate the unknown point, as illustrated in Figure 1;

- Research whether some clusters contain too many ATRs, for purpose of making reliable statements about the system, and for purpose of making point estimates of factors. This will be performed by means of pairwise regression of ATRs within clusters. Means by which to evaluate results need further analysis; and
- Research whether some clusters contain too few ATRs. This will be accomplished by analysis of errors at manual count locations, where "ground truth" data exist. Areas where error is too large may indicate locations where additional ATRs may be needed. This type of analysis needs to be done for traffic counting, vehicle classification, and truck weights.

Detailed Work Plan

The stated recommendations provide the basis for the development of a detailed work plan for Phase II. The Phase II research will provide further development of the recommendations of Phase I research. Although cluster analysis is the recommended approach, additional statistical confidence is sought by refining the estimates within clusters using Multiple Linear Regression (MLR) or triangulation.

Task 6: Expand Data Set and Refine ATR Clusters

The data set used in Phase I will be expanded by adding data for prior years that will enable estimation of growth factors and the inclusion of additional data for WIM and HPMS sites. The cluster analysis done in Phase I will be rerun and finalized by including the new locations and using Log of AADT. Phase I research determined that five clusters was optimal, but the analysis used only AADT. We will test that conclusion using log AADT, U/R class and functional class.

Task 7: Estimating Models: Cluster Means, Cluster Specific MLR, and Triangulation

Seasonal, growth, design hour, and heavy truck factors derived from cluster means will be compared to two other approaches that should refine and improve estimation. Cluster specific MLR models will be estimated with the factors as dependent variables. Triangulation of ATRs of clusters will provide another estimate of factors, which will be compared in Task 8. Phase II research will also deal with the problem of hetroscedasticity, which was encountered in Phase I. Heteroscedasticity is a phenomenon where the residuals do not have a common variance. We are examining remedies to this problem, particularly the deflation of all variables by some measure of "size", in this case dividing by AADT.

Task 8: Statistical Comparison of Cluster Means, Cluster Specific MLR and Weighted Triangulation

Inference of factors from ATRs can employ cluster means, MLR, or a weighted mean of the three ATRs of the same cluster type that triangulates the unknown point. This task will compare the three methods using "ground truth", a set of known points which will be systematically selected ATRs. Analysis of error at "ground truth" locations will provide evidence to determine which of the three methods should be employed and guidelines for use developed.

Task 9: Identify Locations Where Additional ATRs May Be Needed

Using the models to estimate factors for ATR locations will help to identify where additional ATRs may be needed. Spatial analysis of error will help to determine locations where additional ATRs may be needed.

Task 10: Extension of Approach to Vehicle Classification and Truck Weight

The analysis done for traffic counting will be modified and applied to vehicle classification and truck weighing to determine the subset of ATR locations that should be given priority for vehicle classification and WIM equipment. This will be accomplished by additional cluster analysis, using vehicle class and heavy trucks as additional criteria.

Task 11: Draft Final Report

The research will result in recommendations for an integrated traffic data collection system, which will provide guidelines for the deployment of ATR, AVC and WIM equipment.

Task 12: Review and Discussion of Recommendations

Review and discussion of results and recommendations may require additional analysis and/or interpretation.

Task 13: Prepare Final Report

The Final Report will provide guidelines and recommedations for Highway Division staff to develop an Implementation Plan.

Time Schedule

The duration for Phase II is to be 12 months. The following timeline shows the scheduling of the individual tasks.

Timeline

<u>Month</u>	0	2	4	6	8	10	12
Task 6	xxxx						
Task 7		xxxx					
Task 8		xxxxxxxxx					
Task 9			xxxx				
Task 10				xxxxxxxx			
Task 11					xxxxx		
Task 12						xxxxx	
Task 13							xx

Products

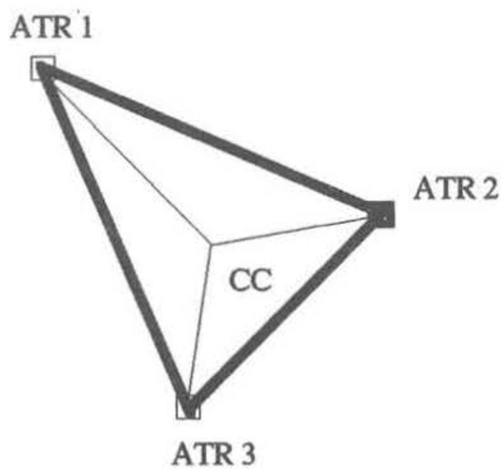
Progress Reports .... Monthly
Draft Final Report .... End of Month 9
Final Report.. End of Month 12

BUDGET

Personnel	
K. Dueker, PI, 10% 8 months, 5% 2.5 months	$5,637.00
R. Ledbetter, GRA, 30% 12 months	5,688.00
Student Hourly	1,000.00
Fringe Benefits	4,256.00
Consultant	600.00
Travel	1,000.00
Services and Supplies	1,689.00
Equipment – Computer & Software	1,000.00
Total Direct Costs	$20,870.00
Indirect Costs	3,130.00
TOTAL COST	$24,000.00

Figure 1

Triangulation



Triangulation of the point subset of ATRs for a given cluster, plus the Coverage Count location for which factors are being computed. This will yield a weighted average of factors from the three nearest relevant ATRs from the cluster that most resembles the Coverage Count location, in terms of similar functional, urban/rural, and volume classes.

A more efficient method is to use the traffic information provided by existing ATR's to predict AADT at point location on highways with similar seasonal and day of week factors, that effect traffic volumes. The underlying assumption of this approach is that ATR's can be grouped so that each group contains point locations with similar traffic characteristics. When data is collected on a statistically selected sample of ATR's in any one group, it can be used to factor raw data counts from other point locations where ATR's are not installed.

PURPOSE

In accordance with Task #1 entitled "Literature Review," this paper presents a brief discussion of the approaches undertaken by various state transportation agencies to monitor traffic and develop integrated traffic-counting programs that are statistically accurate and economically efficient. It is recognized that valuable information can be obtained by reviewing other states' experiences, especially those that are similar to Oregon in size and spatial characteristics. The states surveyed included Colorado, Florida, New Mexico, Virginia, Washington, Wisconsin, and the National State of "Israel." Each of these states has recently undergone a program evaluation or change to improve their traffic counting and monitoring efforts. The purpose of the literature review is to determine if any of the approaches utilized are statistically valid and replicatable, and would enhance Oregon's efforts to improve the efficiency of its traffic data collection system.

In four of the states (New Mexico, Washington, Wisconsin, and "Israel"), the review included a summary of their traffic counting program. The analysis of the remaining states (Colorado, Florida, and Virginia) focused on particular aspects of their program, such as new statistical procedures, and/or highway link grouping procedures to calculate seasonal factors for predicting AADT. Recommendations on applicability and relevance to Oregon are discussed. Information was also obtained from California on their traffic counting methods, however, after review of the material, and considering the extreme differences in size and spatial characteristics with Oregon, no apparent applicability to Oregon was found.

NEW MEXICO

Background

A study of traffic monitoring practices was undertaken by the New Mexico State Highway and Transportation Department in response to concerns for state data integrity, and to standardize the implementation procedures for a statewide traffic counting program recommended by the FHWA. The study reviewed current practices in relation to the Traffic Monitoring Guide (TMG), published by the Federal Highway Administration in 1985.

Findings

The traffic monitoring study determined that the current traffic monitoring practices in New Mexico produced questionable traffic monitoring data. The primary source of error in traffic monitoring was the inconsistent use of professional judgment in factoring traffic data. Professional judgment was applied to adjust base data and complete missing data on the transportation network.

It was found that traffic counts were modified up to eight separate times by individuals who were unaware of previous or subsequent changes to the data. These judgments were made during field collection, data summarization, selection and application of adjustment factors, and data use.

The duration of counts and the use of adjustment factors varied substantially among transportation agencies within the state. All traffic volume data were typically labeled "Average Daily Traffic" (ADT). This resulted in a traffic flow map with Average Weekday Traffic (AWD) on some roads, and Average Seven-Day Week Traffic on other roads--all labeled as "ADT." Some 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 factors; and non-count based estimates. All volumes ere labeled "ADT" and presented as equivalent data. Clearly, a more statistically valid and accurate method was needed to classify traffic volumes and factor traffic count data.

The study reported that data consistency is an issue nationally and that no state surveyed by New Mexico (not identified in their study) was found to have adopted statewide standards for collecting and reporting traffic summary statistics.

The basic function of an ATR is to provide data which can be analyzed and grouped on common patterns for the development of traffic volume adjustment factors. These factors can then be applied to coverage counts to estimate AADT. Previous analysis of ATR 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 used by New Mexico are based on summary statistics from ATR's on roadways with the same functional classification. The study also determined that individual ATR's effectively control volume factors on the same roadway for relatively short geographic distances. Beyond a two-mile distance, the mean statistic provides a more accurate adjustment factor than individual count data. As a result, there is a state restriction on the maximum distance, on the same highway segment, for which data from an individual ATR would be applied.

In past years, New Mexico has installed a variety of ATR's to continuously count traffic volumes. In most instances, this has not been a random process and unknown bias has been introduced into the mean statistic used for determining adjustment factors. The State now requires that a random sample of ATR's, by functional classification, be taken. The following functional classifications are used by New Mexico:

Rural	Urban
Interstate	Interstate
Principal Arterial	Principal Arterial
Minor Arterial and Major Collector	Minor Arterial
Recreational Route	Major and Minor Collector

The data from an ATR not meeting the above state standards is excluded from calculating the mean summary statistics for highway groupings based on functional classification. New Mexico separately stores such data in a research file. At the end of each year, the statistical significance of including data which has been excluded, is evaluated. If merited, the state standards may be modified to include the data. The report did not discuss the effects of such a procedure, or what additional information would be necessary to make such a decision.

Conclusions

The new traffic counting procedures proposed for New Mexico are aimed at improving the accuracy of calculating seasonal adjustment factors from existing ATR locations. The program uses the standard FHWA functional classifications to group ATR's. The mean statistic for each group is used rather than individual ATR data, for calculating the adjustment factors. New Mexico found that the mean statistic provided better accuracy than individual ATR data, at distances greater than two miles on a particular highway segment. This has significance to Oregon because of the desire to use individual ATR counts at coverage count locations to determine seasonal adjustment factors. Further investigation of the distance constraint needs to be tested in the Oregon site location study.

WASHINGTON

Background

The Washington State Department of Transportation (WSDOT) currently collects traffic count data from 80 permanent traffic recorders (PTR) at approximately 65 locations throughout the state. The data is collected using a ,recently installed, telemetry system. The conversion to this method has reduced the amount of manpower needed to collect and manipulate ATR counts.

The data collected from the ATR's is seasonally adjusted and entered into the existing traffic volume data base for future reference. Volume data already in the data base, and not replaced by a new volume count, are adjusted annually to reflect the growth in VMT in the state. The seasonal factors applied to each raw count are derived from available ATR data. A transportation data office engineer or technician determines the particular ATR to be used for calculating the factor on the basis of his/her knowledge of the road being counted, the roads that contain ATR stations, and on previous estimates of seasonal factors for various road sections in the highway system.

Based on concern about the appropriate level of resources to be allocated to data collection activities and the statistical basis for the program, WSDOT undertook research to evaluate their overall approach.

Findings

To a limited extent, the evaluation showed the existing data collection program fulfills the majority of the department's current needs. However, in spite of the program's low level of data collection activities, an increase in costs has been experienced. Data deficiencies and collection inefficiencies are attributed to the increase.

The study found that the WSDOT did little traffic counting other than at existing ATR locations. Ad hoc seasonal factors were applied manually, as opposed to statistically derived factors using an automated or computer approach. No HPMS data is collected by WSDOT off the state highway system. The lack of a statistically rigorous procedure for collecting data created doubt about the accuracy and validity of current data collection efforts.

Stephen G. Ritchie ("A Statistical Approach to Statewide Traffic County," Transportation Research Record, 1090) presents a statistical framework for volume counting and vehicle classification to estimate AADT from short-duration axle counts, at any location on a state highway system. The data used was taken from the WSDOT case study discussed above. The procedure adopted by WSDOT begins with a basic model for estimating AADT from a coverage count location as follows:

AADT = VOL(Fs)(Fa)(Fg) where

VOL = average 24-hour volume from a standard WSDOT 72-hour Tuesday-Thursday short count;

Fs = seasonal factor for the count month;

Fa = weekday axle correction factor if VOL is in axles;

Fg = growth factor if VOL is not a current year count.

The relative precision of estimated AADT was measured by the coefficient of variation (cv).

Ritchie then evaluated several alternative methods for performing seasonal factoring including the continued use of existing WSDOT Data Office procedures; cluster analysis of existing ATR's; and procedures suggested in the FHWA counting guide.

The approach selected was a revised FHWA procedure using a simple linear regression technique. The highway system was stratified by region and functional classification. For the purposes of the Washington study, seven factor groups were chosen based on 1980 through 1984 ATR data. Seasonal factors for each month of the year were then derived for each of the seven factor groups according to the following regression model:

AADT = bVol + e where

AADT and VOL are as defined above. The b is the seasonal factor to be estimated and the e is the predicted error term.

Conclusions

Ritchie cites the advantages to the regression approach used in the WSDOT study to be:

- The seasonal factors are statistically valid.
- The overall errors are equal to or smaller than the errors associated with any other seasonal factoring approach considered.
- The factoring procedure is transparent to any user of volume information and thus allows the recalculation of the raw traffic count at some later time if desired.

The study further concluded that the use of a computer routine to group ATR's could be problematic. First, the clusters computed were not consistent across years. An ATR in one group may change to another group the following year. No method was available to account for this adjustment. Next, individual road sections were not easily or accurately assigned to cluster groups, irrespective of the difficulties mentioned previously. Finally, the total error in the AADT estimates (including seasonal variation) was only marginally better than that obtained by the regression approach, before inclusion of the indeterminate error that is present as a result of the above problems.

WISCONSIN

Background

The Wisconsin Department of Transportation (WisDOT) currently maintains an extensive traffic counting program (TCP) to collect, process, summarize and report traffic flow data throughout the state. Traffic volumes are monitored periodically at 20,000 locations across the state. Approximately 6,700 of these locations are counted annually at a cost of $800,000. Vehicle miles traveled and AADT are two of the most basic measures of highway utilization monitored by the program.

Wisconsin's present program groups the continuous traffic count stations (ATR's) by seasonal traffic variation factors. Monthly factors--defined as the ratio of AADT to Monthly Average Weekday Traffic (MAWDT)--are used to measure seasonal traffic variations. After the groups are finalized, monthly group mean factors are computed for each group for use in adjusting short period traffic counts to estimates of AADT. Highway sections that do not contain continuous count stations are assigned to these groups on the basis of seasonal traffic counts. The appropriate mean factors from the short period counts are applied in arriving at estimates of AADT. Excluding sources of error due to traffic counts collected by road tubes, seasonally adjusted short counts were found to be accurate, on average (across all factor groups), within plus or minus 23%, with 95% confidence. However, there are several issues that Wisconsin found that needed to be addressed.

First, weekly factors are used as the seasonal adjustment factors yet, monthly factors are used to group ATR stations. This creates an inconsistency in the procedures. Second, the current seasonal factoring procedure requires excessive subjective judgment and time consuming manual manipulation in both the process of grouping ATR's and assigning highway sections to defined factor groups. Third, the current seasonal factoring procedure offers no guidance on the optimum number of factor groups. The current procedure results in six basic groups. Yet there is no theoretical basis to expect this number. Fourth, the current procedure requires the collection of large amounts of traffic data. Annually, eight to twelve weeks of counts are taken at 300 locations for the purpose of assigning highway sections to factor groups.

To help address the above issues, Wisconsin developed a simulation program using the Statistical Analysis System (SAS) to test the effects of different seasonal factors, different grouping procedures, and different numbers of factor groups, on the prediction of AADT and levels of statistical significance. The simulation uses the SAS computer routing "Cluster" to group ATR's using monthly seasonal factors. This grouping procedure uses a standard agglomerative hierarchical clustering algorithm (K-Means) that places each observation in a cluster by itself; merges the two closest clusters to form a new cluster; and repeats the process until all observations have been assigned to a cluster. The K-Means algorithm along with other clustering routines are discussed later under Clustering Algorithms.

Findings

Results of the new procedures indicate that using weekly (rather than monthly) factors provide no significant increase in accuracy in predicting AADT. Additionally, monthly factors capture the overwhelming proportion of seasonal variation and are subject to less variation over time, than weekly factors.

Wisconsin also found that the large number of factor groups currently used (6 groups) resulted in instability, with some ATR's changing groups each year after the grouping procedure (this finding is consistent with the WSDOT study discussed previously). Wisconsin considered this possibility a concern because it implied potential errors in peak and design hour factors. Extensive analysis of existing traffic flow data across the state of Wisconsin indicated three groups to be the optimum number for calculating seasonal adjustment factors. Wisconsin predicts that AADT estimates based on three groups should be accurate at the 95% confidence interval, within plus or minus 9%, for urban highways (Group 1), plus or minus 16% for rural highways (Group 2), and within plus or minus 27% for tourism/recreation area highways (Group 3).

Finally, the new seasonal adjustment procedures allow for the elimination of most seasonal coverage counts which results in a cost savings to the state.

Conclusions

The benefits from Wisconsin's program are reported to be:

- Ease in developing the monthly factors for seasonal adjustments.
- Increased reliability and accuracy of AADT estimates.
- Substantial cost savings due to the elimination of coverage counts for the purpose of assigning highway links to specific factor groups.

THE NATIONAL STATE OF "ISRAEL"

Background

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 a week, every one or two years, by mechanical counters. Some manual counts are also taken to determine the percent of heavy vehicles using the highway system.

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.

Yehuda J. Gur and Irit Hocherman in a recent Transportation Research Board paper , "Optimal Design of Traffic Counts," describe the results of a study which examined possible ways to incorporate the arterial urban road network into the current traffic counting procedures. The basic strategy was to reduce the frequency of counting by ,sampling the sections to be counted every year, and devising a method for estimating AADT on those sections that were not counted. The methods evaluated included: (1) estimation of traffic volumes based on road characteristics; (2) estimation of traffic volumes based on past values; (3) the use of network connectivity information as an aid in updating volume estimates.

The data used in the study included the following:

- An identification number for each section of roadway including its two nodes;
- Average weekday traffic volume in each of the years 1982-1985;
- The class of road, region, number of lanes, length and a dichotomous variable denoting proximity to a metropolitan area.

Findings

The study found a significant association between link volume and type of road, and an apparent association between volume and region. However, no systematic time trend between years, seasons, or months was found.

The first method tried in the study to estimate AADT on a given segment of highway involved a regression model of average weekday traffic (AWDT) as a function of road classification, region, number of lanes, and geographic area (1-rural, 2-metropolitan). A separate model was calibrated for each of the years 1982-1985. The "R" squared ranged between .70 and .74. However, the error in the model was very high (approximately 6500 vehicles per day). This large of error was considered unacceptable, particularly for estimating traffic volumes of individual road sections from coverage counts.

The second method for estimating traffic volumes on selected sections of highway involved a regression equation that utilizes the count data on the same section of highway from the previous year. The procedure provides a growth factor that can be applied to prior year counts to arrive at AADT. The procedure was tried after disaggregating the highway network by route type and region. The results showed that estimating a uniform growth factor for the whole system was superior to individual growth factors for each road type and region. The "R" squared was reported to be .97, and the error present in the model much lower than the regression from above which was based on the physical characteristics of the highway segment alone. The model suggested that current year link volumes could be estimated by assuming a uniform growth factor from the previous year for each region and road classification.

A more complex procedure focused on updating link volume estimates based on volume changes in connected links. Standard statistical methods treat links in the highway system as independent members of a population. In reality, links are connected into a network, and trips are distributed over a chain of connected links. The authors believed it likely that volumes are correlated and that a volume change in one link is associated with changes in related links.

The proposed model established chains consisting of links which belong to the same route. Links with updated counts were used to estimate volume changes over the entire chain. Weighting factors used included the following:

- The distance between the counted link and the link of interest;
- The volumes on the links along each chain and their stability;
- The number of chains with and without counts.

Upon calibration of the model with 1985 data, the results showed large and inconsistent variations in the extent, and even direction, of volume changes. These findings tend to support the use of clustering techniques for the purpose of determining seasonal adjustment and growth factors, rather than network analysis.

Conclusions

The authors concluded that the use of a systemwide, uniform growth factor, provided the best estimates for updating link volumes. This method (method 2 above) provided satisfactory estimates, even at a 30% sampling rate. Based on these findings, the study recommended an annual sampling rate of about 40%, with about a quarter of the coverage counts taken at new roads and areas where significant changes are suspected.

COLORADO

Background

In a draft report "Seasonality Analysis of Colorado's 1986 ATR Data, Mr. Esteve from the FWHA examined Colorado's 1986 ATR data to establish seasonality patterns and to recommend an appropriate grouping methodology. Colorado uses a computer generated clustering technique (SAS), (similar to Wisconsin) to group its 54 ATR's based on volume counts. These groups were compared with the TMG proposed groups developed by the FWHA (Interstate Rural, Interstate Urban, Interstate Recreational, Other Rural, Other Urban, and Other Recreational). The mean seasonal factors and coefficient of variation for each group were then calculated.

Findings

The findings showed that the reliability of both procedures was very similar. After a thorough review of the state report, the author could not proclaim one procedure superior to the other. However, the draft report did emphasize the simpler, functional class grouping contained in the TMG, because they make assignment of highway segments easy and require no control counts.

Based on an examination of the precision of the monthly factors calculated for the recommended TMG groups, it was concluded the existing number of ATR's (54) was more than sufficient to achieve a precision level of plus or minus 10% with 95% confidence. In addition, there was room for reducing the number of ATR locations. The same precision level was obtained with the following reductions in ATR's.

	Group	Current ATR's	Recommended
1.	Interstate Rural	6	7
2.	Other Rural	21	15
3.	Interstate Urban	5	5
4.	Other Urban	18	8
5.	Recreational	4	4
	TOTAL	54	39

Conclusions

The report concluded that in the application of the group factors, it is important to realize that these are system factors, not specific point factors for each ATR. Although the TMG recommends computing and maintaining tables of the individual factors at each ATR location for greater accuracy, the Colorado analysis showed the group factors to be appropriate for their program.

FLORIDA

Background

In order to provide as many traffic counts as possible, the Florida Department of Transportation established a 24-hour counting period to collect traffic data. Prior to 1982, seasonal 24-hour counts were taken from one to four times a year, and the mean of the counts was assumed to be the AADT for the count location. A study conducted by the Traffic Counts Office in 1981, using continuous count information, verified that AADT could be replicated accurately using one 24-hour seasonal count, factored by weekly factors generated at continuous count sites.

Findings

The Florida program operates on the assumption that continuous counters replicate the traffic patterns within contiguous areas. Based on this, these counters have been assigned to individual counties within Florida for the purpose of developing monthly or weekly factors that are used to factor a single 24-hour count to AADT. The factor development process is shown below:

- WADT/AADT = MF
- The sum of all MF/N = MFi or MFj
- (MFj-MFi)/# Weeks = AWF
- ADT/AWF = AADTn

Definitions:

WADT = weekly average daily traffic for an ATR

AADT = annual average daily traffic

MF = monthly factor

MFi = average monthly factor for a county/area for preceding month.

MFj = average monthly factor for a succeeding month

Weeks = number of weeks between mid-month to next mid-month

AWF = average weekly factor

ADT = 24 -hour traffic count

AADTn = annual average daily traffic for a non-continuous count site

Conclusions

All transportation districts in Florida now use the above procedure for estimating AADT at locations that do not have ATR's. The FHWA report on Florida did not indicate if ATR's are grouped in any particular manner other than by county. Geographic and demographic variables were not considered in determining the groupings. The assumption that contiguous areas have the same traffic patterns would have to be tested statistically before any conclusions can be drawn about the procedure.

VIRGINIA

Background

As discussed in the introduction, the primary factors commonly used for grouping highway links are the functional class as given in the HPMS and AADT. Recent evidence indicates that a highway link classification system based solely on these two factors may give clusters or groups that include links with different traffic characteristics (Garber and Bayant-Mokhtari). Differences in seasonal variation and the lack of accurate information on AADT for each link are the reasons cited. The authors describe a new classification system developed for the state of Virginia's rural highway system. They indicate the procedure can be adapted to urban roads as well.

Findings

Each highway in the rural area of Virginia was broken down into homogeneous links according based on AADT and seasonal variations in traffic volumes. Variables that had a significant effect on AADT and can be used as surrogates in the clustering system were identified from a literature search. Statistical tests, based on AADT data for 1977 through 1980, collected at 112 continuous traffic count stations, were conducted to identify the most significant variables. The resulting list included the following:

- FHWA functional class;
- Functional use in Virginia (Recreational, local service, long-distance, industrial, local commercial);
- Land use of the county in which the link is located;
- Population of the county in which the link is located;
- Type of terrain.

Groups were formed based on the above variables using MacQueen's K-Means clustering technique. Nominal variables were converted to interval variables by representing each category of the nominal variable by one (1) percent of the average AADT of the continuous count stations in that category. The order of magnitude between variables was compensated for by standardizing the data prior to clustering. Based on the calculation of the coefficient of variation (cv), the results indicated a reasonable number of clusters to be between four and nine. The test results showed that the cv's from this procedure were lower than the recommended FHWA values. One advantage to Virginia's procedure is that groups of highway links with similar traffic characteristics can be formed without the necessity of first assigning a value for the AADT of each link. However, the procedure does require that all highways being considered for grouping be divided into homogeneous links such that traffic volume characteristics on any one link do not vary along that link.

Conclusions

The conclusion from the authors was that the computerized highway link classification system for estimating AADT in Virginia was statistically accurate and would result in some level of cost savings. This conclusion was based on the finding that because the coefficients of variation of the AADT in each group would be low, the number of links required to be sampled for a given level of accuracy will be relatively small.

NCHRP HEAVY VEHICLE MONITORING PROJECT

A NCHRP study of heavy vehicle monitoring dealt with deployment of the technology on the highway system for both enforcement and planning (Grenzeback, Stowers and Boghani, 1988). For planning purposes, they conclude: "A relatively small number of sites would be used as permanent recording stations to obtain seasonal and weekly factors, and a much larger number of sites would be used for short-term classification and weight studies." They also conclude that the TMG recommended approach of randomly selecting sites for the data collection sessions from among the state's sample of highway data collection sections used in the HPMS is inefficient, because there is a high correlation in traffic characteristics from section to section over long segments of highway, particularly for the Interstate System. This suggested to them that states might easily develop a more efficient data collection program by selecting WIM data collection sites that are representative of longer segments.

CLUSTERING ALGORITHIMS

Background

A substantial amount of work in the development of clustering and classification methods has been directed towards methods that focus on the minimization of the within-group sum of squares (error). These methods have been termed minimum-variance constraint methods (Wishart, 1969) and use the squared Euclidean distance as the grouping metric. Several popular methods based on this measure of distance and described in the literature are described below.

Literature Findings

MacNaughton-Smith (1965) defines a hierarchical method, called the furthest neighbor technique. This method treats each data element as a single-point cluster. The elements are combined in steps such that the two clusters having the smallest Euclidean distance between the two most distant elements in each cluster are combined.

Ward (1963) uses the within-group sum of squares (WGSS) as the grouping metric. His criterion is also termed the error sum of squares (ESS) and is just the sum of squares of the distances from each data point to its parent cluster mean.

Sokal and Michener (1958) describe a procedure wherein the distance between two clusters is given in terms of the squared Euclidean distance between their centroids. This method also combines clusters having the minimum squared distance between data elements.

A procedure of Ball and Hall (1965) forms a desired number of clusters by first selecting a specified number of data elements at random and then assigning each of the remaining elements to the cluster center nearest it. The cluster centroids are computed and any two clusters are combined if the squared Euclidian distance is less than a pre-determined threshold value. A cluster is split if the variance within the cluster of any one variable exceeds the threshold value.

MacQueen (1966) proposed a technique similar to Ball and Hall's. His procedure chooses a desired number of data elements at random to be used as cluster centers. Each data point is assigned to the center nearest it if its distance from that cluster center does not exceed a pre-determined threshold value. If a data element is farther than the threshold value from all centers it forms a new cluster. After each allocation, the cluster centroid is recomputed and becomes the new cluster center. When the distance between two centroids becomes less than a specified threshold value, the two clusters are combined. This technique is known as K-Means clustering.

There are various objections to the minimum-variance approach for forming clusters. First, changes in scale will modify the resultant clusters. In addition, the addition of variables will change the clusters. One remedy to the scale problem is to standardize the data first, before clustering (Hartigan, 1975). This procedure centers and scales data and subsequently effects the weighting factor in various types of distance calculations. This may be advantageous at times, however, Hartigan discusses the possibility of masking true differences in clusters by using the technique.

The literature pointed to the fact that little has been done to evaluate the differences between the various methods. It is apparent that the purpose of a particular study and the type of data used is the best determination of the appropriate clustering procedure to use.

Conclusions

The existing literature on clustering techniques is in-conclusive as to whether one method is better than another. Several new additions to the literature have emerged which describe new procedures for identifying significant variables that influence AADT, cluster-segmentation methods, and highway link identification. These procedures can be used to estimate AADT from coverage count locations on selected highways where ATR's are not installed.

The evidence indicates that a systematic grouping of the ATR's does provide more accuracy in the estimation of AADT and seasonal factors. Whether the groupings should be based on a computer algorithm involving geographical, functional, and individual highway characteristics, or the standard TMG procedures, is not certain. Good results were reported by states using each method.

The use of seasonal factors to estimate AADT provides better precision than raw data counts. This result was common among the states surveyed. The one exception was the National State of "Israel," which used a uniform growth factor void of any seasonal variation among highway types.

APPLICATION OF LITERATURE REVIEW FINDINGS TO THE OREGON STUDY

This section provides a brief assessment of the applicability of selected state findings to the Oregon study.

New Mexico

New Mexico found a high degree of correlation between the functional class of the highway and the seasonal variation for individual ATR's. For this reason, they rely heavily on volume counts from ATR's on roads with the same functional classification. However, their study also found that there was a distance constraint of approximately two miles for which individual ATR counts could be used in accurately calculating seasonal adjustment factors. Beyond the two mile distance, group means provided a better estimate for calculating seasonal variation, and ultimately, AADT. If the Oregon study relied solely on volume counts at ATR locations to determine seasonal factors, the same distance constraint may be encountered. However, it is contended that the use of other factors such as functional class, urban/rural designation, volume class, and socio-economic characteristics, as surrogates for volume counts will serve to eliminate the distance problem.

Washington

Washington found that ATR's could be grouped geographically by functional class for the purpose of calculating seasonal adjustment factors. This same finding is not visible in the Oregon study as of this writing. The use of various clustering criteria including state plane coordinates, AADT, seasonal factors, design hour volumes, and the percent of heavy vehicles, did not show a consistent spatial pattern for determining seasonal variations in traffic volumes.

Wisconsin

The results of the Oregon study to date agree with Wisconsin's findings that the use of a computer clustering technique provides a more systematic and statistically significant way to group ATR's. These groupings allow for the calculation of seasonal factors that capture the majority of the seasonal variation at specific ATR locations. In addition, the coefficient of variation provides a means to statistically check the accuracy of assigning a particular ATR to a factor group. The difference in the Oregon approach is that the functional classification, urban/rural designation, and volumes from continuous ATR's are used in the grouping procedure. After grouping, the seasonal adjustment factors can be found through linear regression involving the grouping variables. The Oregon study is considering five (5) factor groups instead of just three (3), to more closely parallel the FWHA functional classifications for highways.

"Israel"

The use of a single growth factor for updating volume counts has only limited applicability to the Oregon study because only one year of data (1986) is currently being used. During Phase II of the study, the use of longitudinal data over time may show more relevance to the Israel approach. In addition, Israel does not appear to have a coverage count problem for estimating AADT because of the lack of seasonal variation in traffic volumes. In Oregon, seasonal variation plays a larger role in estimating traffic volumes and must be considered along with yearly growth factors.

Colorado

The results of Colorado's statistical analysis of their ATR data showed that the existing locations provided reliable estimates of AADT for federal reporting. In fact, Colorado concluded they could reduce the total number of ATR's without changing the level of accuracy. It is felt that the group factor approach is relevant to Oregon and the overall goals of it's Integrated Traffic Counting Program. It also seems feasible, based on the Colorado results, that a reduction in the number of ATR's used in Oregon is possible, while maintaining the desired level of accuracy and precision for federal reporting purposes. The results did not, however, address the statistical reliability of short duration coverage count locations. It is possible that additional ATR sites will be needed to estimate AADT and seasonal factors for these locations, not less.

Florida

Florida's program assigns ATR's to individual counties for the purpose of developing weekly factors for use in factoring coverage count data to AADT. They operate on the assumption that ATR volume counts replicate the traffic patterns within contiguous areas of the transportation system. The Oregon study does not make this assumption. Instead, each ATR point location is grouped by its functional class, urban/rural designation, and ADT. These groupings are used to determine seasonal adjustment factors for estimating AADT at non-continuous count locations in the state. The groupings of ATR's in Oregon have not resulted in any significant spatial orientation, whether it be county boundary or other regional delineation.

Virginia

Virginia's clustering approach for grouping highway links with similar traffic characteristics, without the necessity of first assigning a value for the AADT of each link, looks promising. The extension of this idea to point locations will be looked at in Oregon's Phase II study. In addition, their method for converting nominal level variables to interval variables will be evaluated further.

NCHRP Heavy Vehicle Monitoring Project

The WIM deployment conclusions and recommendations contained in NCHRP Report 303 are consistent with those of Phase 1. This will be developed more in Phase II. We are particularly interested to see the extent to which cluster techniques accomplish the aggregation of like short highway segments to longer segments.

Selected References

Albright, David; Blewett, Chris; and Goldstine, Rachel., "Development of State Traffic Monitoring Standards: Implementing the Traffic Monitoring Guide, in New Mexico." Transportation Research Board 68th Annual Meeting, January 22-26, 1989. Washington D.C. Paper No. 880241.

Ball, G. H. and Hall, D. J., "A clustering technique for summarizing multi-variate data," Behavioral Sciences, Vol. 12, No. 2, (March 1967) pp. 153-155.

California Department of Transportation, Traffic Manual, Ch. 2, Traffic Volumes System, July 1979.

Esteve, M., Draft Report, "Seasonality Analysis of Colorado's 1986 ATR Data," Federal Highway Administration, May 1987.

Everitt, Brian., Cluster Analysis, Halsted Press. 1974.

Federal Highway Administration, "Traffic Forecasting For Pavement Design," Report No. FHWA-TS-86-225. U.S. Department of Transportation, March 1988.

Garber, Nicholas J., and Bayat-Mokhtari, Faramarz, "A Computerized Highway Link Classification System for Traffic Volume Counts," Transportation Research Record 1090, TRB, Washington, D.C.

Grenzeback, L.R., Stowers, J.R., and Boghani,A.B., "Feasibility of a Mational Heavy-Monitoring System," NCHRP Report #303, Transportation Research Board,1988.

Gur, Yehuda J., "Optimal Design of Traffic Counts," Transportation Research Board 68th Annual Meeting, January 22-26, 1989, Washington, D.C., Paper No. 880552.

Hartigan, John A., Clustering Algorithms, John Wiley & Sons, New York, 1975.

MacNaughton-Smith, P., "The Classification of Individuals by the Possession of Attributes Associated with a Criterion," Biometrics, 19, (1963) pp. 364-366.

Ritchie, Stephen, G., "A Statistical Approach to Statewide Traffic Counting," Transportation Research Record 1090, TRB, Washington, D.C.

Ritchie, Stephen G. and Hallenbeck, Mark E., "Evaluation of a Statewide Highway Data Collection Program," Transportation Research Record 1090, TRB, Washington, D.C.

Sokal, R. R. and Michener, C. D., "A Statistical Method for Evaluating Systematic Relationships," University of Kansas Sci. Bull., (March 1958), pp. 1409- 1438.

State of Wisconsin, Wisconsin Department of Transportation, A Statewide Traffic Counting Program For The Wisconsin Department Of Transportation. Madison, Wisconsin. January, 1987.

Attachment 1

Selected Literature Survey of Statewide Traffic Counting Programs and Data Clustering Techniques

Introduction

Consistent with state and federal mandates, highway data and specifically traffic count data have been collected by state transportation agencies to support a wide range of programs and needs. These have included the use of traffic count data to develop estimates of annual average daily traffic (AADT), vehicle miles of travel (VMT) and design hour volume (DHV) for individual highway sections, and functional classifications or other regional grouping of the state highway system. In addition, the Federal Highway Administration (FHWA) has required submission of various traffic and truck data estimates for use by FHWA and other federal agencies. These have been required in order to establish national travel trends, prepare reports requested by Congress, plan for future transportation needs, and assess the overall efficiency of various programs, policies, and funding sources.

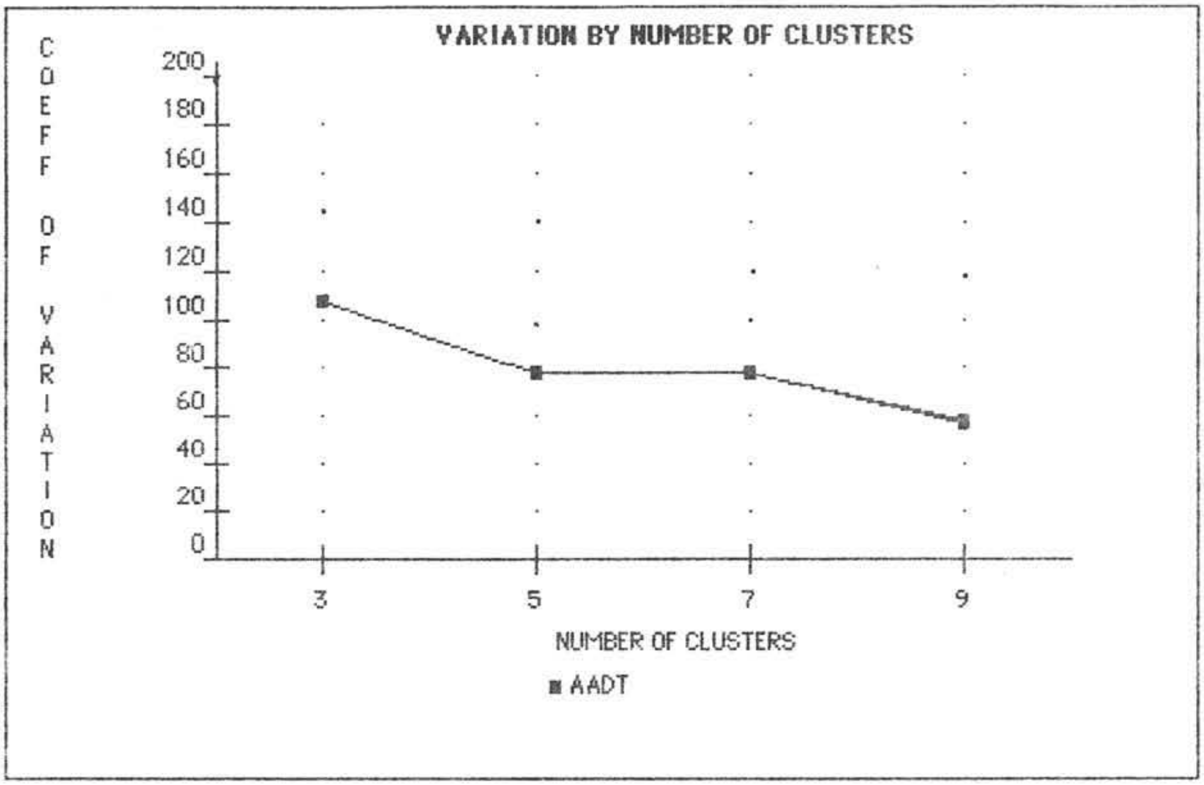
In recent years, considerable concern has existed about the appropriate level of resources to be allocated to various state data collection activities and the statistical basis for these activities (Ritchie and Hallenbeck, Transportation Research Record, 1090). As a result, many state transportation agencies are making a concerted effort to reduce their annual expenditures on traffic counting programs, while attempting to maintain the desired level of statistical accuracy.

The most accurate way to monitor changes in traffic volumes at specific locations on the highway system is to continuously count traffic for each 24 hour period throughout the year. This is best accomplished through the use of Automatic Traffic Recorders (ATR). The methodology provides a tremendous amount of information, however, it is also very costly. It is therefore, not economically feasible to install an ATR every few miles on each class of highway within a particular jurisdiction (state, city, county, etc.). A more efficient and cost effective method is to group highways of similar traffic characteristics and apply known seasonal factors from continuous count locations within the group in order to estimate AADT.

The primary factors commonly used by states for grouping highway segments are the functional class of the highway as given in the Highway Performance Monitoring System manual (HPMS) and its AADT. The use of these factors alone can be problematic (Garber and Bayat-Mokhtari, Transportation Research Record, 1090). First, the assignment of a particular highway link to a functional class may be subjective because of the difficulty of distinguishing between major and minor arterials, and collectors. In addition, highways of the same HPMS functional class may not have similar traffic characteristics (e.g., seasonal variation) if they are located in different geographical areas of a state. Second, the AADT for each link is required in order to properly assign the link to a specific group. However, the AADT is not known in many cases and links must be assigned based on past experience and personal judgement of transportation and engineering personnel. Estimates of the coefficients of variation (ration of the standard deviation to the mean) of the AADT of groups formed by the standard FHWA procedures tend to be high. This means that large sample sizes are needed to obtain the required accuracy (Garber and Mayat-Mokhtari). The apparent lack of a sound statistical basis for collecting traffic count data has resulted in increased program costs for many state transportation departments.

Attachment 2

COEFFICIENT OF VARIATION IN AADT AS A FUNCTION OF NUMBER OF CLUSTERS OF ATRs



Plot shows a reduction in the coefficient of variation when the number of clusters is increased from three to five. The reduction from seven to nine clusters is achieved by further division of high volume clusters and does not yield real improvement.

Attachment 3

GEOGRAPHIC GROUPING OF ATRs BY SEASONAL FACTOR, DESIGN HOUR VOLUME, AND PERCENT HEAVY TRUCKS

The overlapping geographic groupings of ATRs indicate there are not natural regions having similar seasonal, design hour and percent heavy truck factors.

GROUPING BY JANUARY SEASONAL FACTOR CLUSTERS

TOTAL OBSERVATIONS: 40

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	40	40	40	40	40
MINIMUM	198.000	0.948	0.733	8.400	3.400
MAXIMUM	117928.000	1.221	0.937	17.500	40.700
MEAN	30227.450	1.096	0.855	10.590	11.440
STANDARD DEV	35964.537	0.089	0.046	1.493	7.169

TOTAL OBSERVATIONS: 39

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	39	39	39	39	39
MINIMUM	208.000	1.228	0.679	9.500	6.900
MAXIMUM	22982.000	1.438	0.878	15.500	39.500
MEAN	5300.513	1.350	0.776	12.679	21.521
STANDARD DEV	5723.147	0.057	0.051	1.650	8.541

TOTAL OBSERVATIONS: 3

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	3	3	3	3	3
MINIMUM	380.000	2.239	0.576	15.600	10.400
MAXIMUM	1330.000	2.639	0.734	47.600	31.700
MEAN	743.667	2.430	0.646	27.900	18.200
STANDARD DEV	512.650	0.201	0.080	17.236	11.738

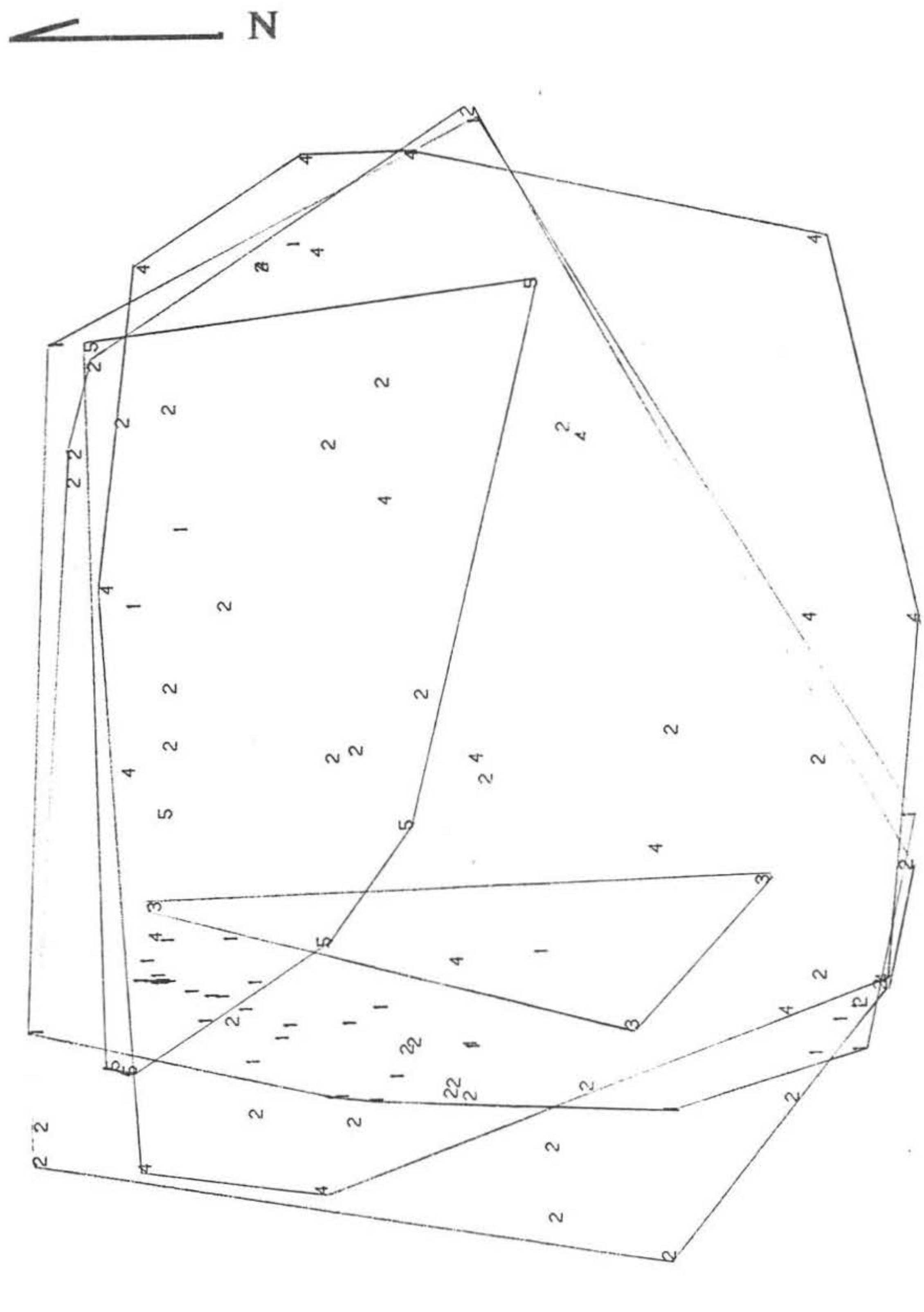
TOTAL OBSERVATIONS: 21

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	21	21	21	21	21
MINIMUM	569.000	1.474	0.629	11.000	10.600
MAXIMUM	16850.000	1.693	0.832	19.100	46.000
MEAN	3896.190	1.574	0.740	14.252	22.843
STANDARD DEV	4071.794	0.065	0.048	2.113	10.385

TOTAL OBSERVATIONS: 7

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	7	7	7	7	7
MINIMUM	739.000	1.770	0.644	13.700	9.500
MAXIMUM	4962.000	1.967	0.818	24.500	26.400
MEAN	2809.000	1.860	0.715	19.957	17.943
STANDARD DEV	1861.220	0.067	0.061	3.512	6.480

MAP OF OREGON CLUSTERED BY JANUARY SEASONAL FACTOR



DESIGN HOUR VOLUME CLUSTERS

TOTAL OBSERVATIONS: 57

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	57	57	57	57	57
MINIMUM	198.000	0.948	0.716	8.400	3.400
MAXIMUM	117928.000	1.643	0.937	12.100	46.000
MEAN	23941.649	1.193	0.831	10.572	16.258
STANDARD DEV	31731.503	0.170	0.057	0.864	11.001

TOTAL OBSERVATIONS: 1

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	1	1	1	1	1
MINIMUM	380.000	2.639	0.734	47.600	10.400
MAXIMUM	380.000	2.639	0.734	47.600	10.400
MEAN	380.000	2.639	0.734	47.600	10.400
STANDARD DEV	.	.	.	.	.

TOTAL OBSERVATIONS: 39

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	39	39	39	39	39
MINIMUM	208.000	1.144	0.629	12.400	6.900
MAXIMUM	16850.000	2.239	0.878	15.700	37.000
MEAN	2986.718	1.450	0.761	13.846	20.133
STANDARD DEV	3600.262	0.200	0.059	0.924	7.373

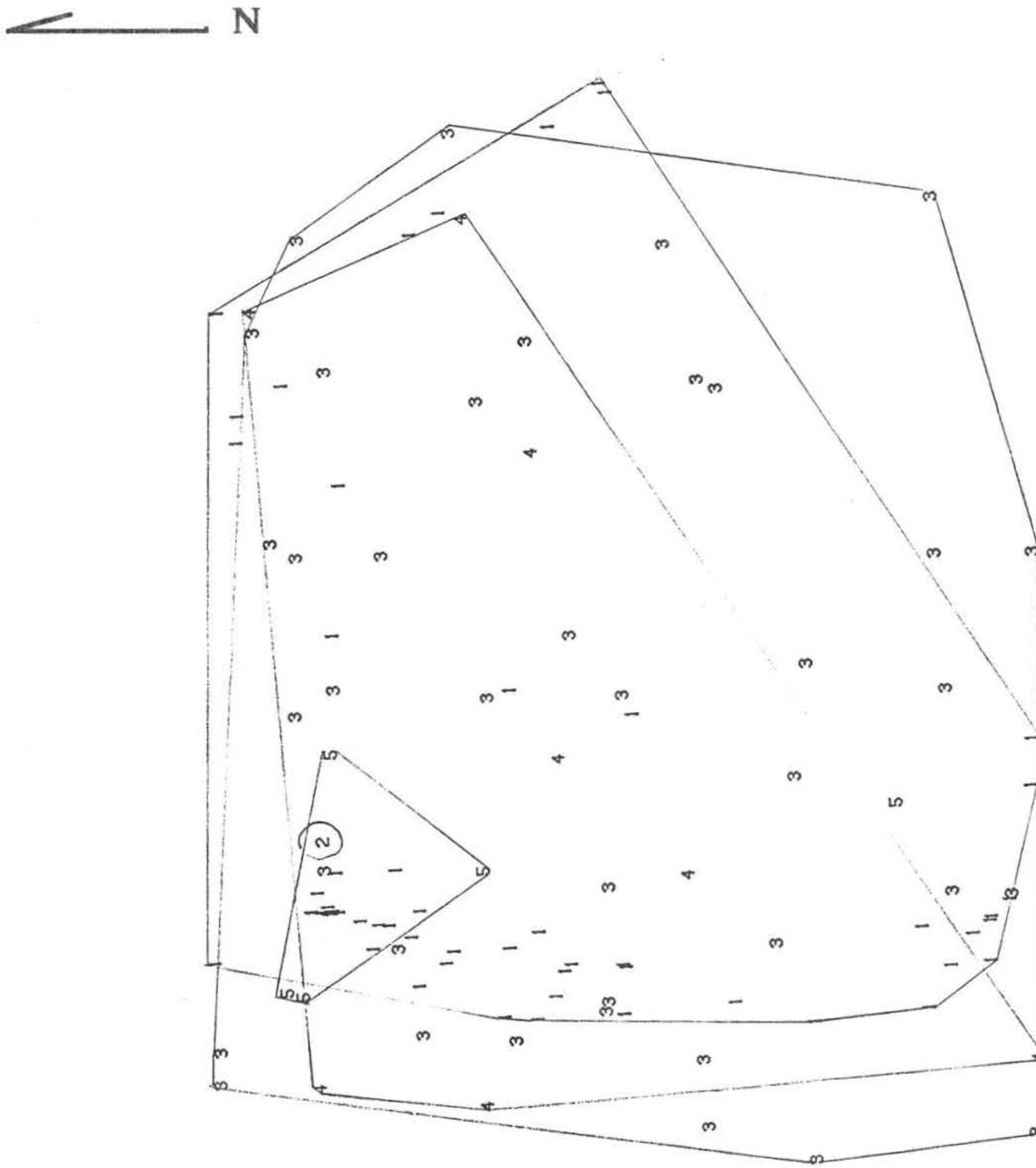
TOTAL OBSERVATIONS: 8

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	8	8	8	8	8
MINIMUM	724.000	1.221	0.629	16.300	10.600
MAXIMUM	7109.000	1.967	0.790	19.100	40.700
MEAN	3070.500	1.627	0.705	17.588	18.688
STANDARD DEV	2256.524	0.229	0.052	0.916	10.250

TOTAL OBSERVATIONS: 5

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	5	5	5	5	5
MINIMUM	521.000	1.770	0.576	20.500	9.500
MAXIMUM	4962.000	2.412	0.818	24.500	24.600
MEAN	2686.400	1.934	0.703	21.980	17.000
STANDARD DEV	1959.103	0.269	0.095	1.509	6.387

MAP OF OREGON ATRs CLUSTERED BY DESIGN HOUR VOLUME



PERCENT HEAVY VEHICLES CLUSTER

TOTAL OBSERVATIONS: 39

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	39	39	39	39	39
MINIMUM	208.000	1.048	0.576	9.500	11.600
MAXIMUM	45735.000	2.412	0.932	21.500	19.300
MEAN	6392.128	1.394	0.788	13.279	14.928
STANDARD DEV	10935.016	0.268	0.075	2.806	2.055

TOTAL OBSERVATIONS: 21

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	21	21	21	21	21
MINIMUM	349.000	1.149	0.644	9.600	20.200
MAXIMUM	21032.000	1.910	0.848	24.500	27.400
MEAN	4395.238	1.481	0.760	14.205	23.900
STANDARD DEV	5698.914	0.221	0.056	3.565	2.371

TOTAL OBSERVATIONS: 14

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	14	14	14	14	14
MINIMUM	984.000	1.298	0.629	9.500	28.300
MAXIMUM	22982.000	2.239	0.825	15.600	37.000
MEAN	7212.214	1.471	0.746	12.450	31.964
STANDARD DEV	6550.844	0.245	0.050	2.001	2.729

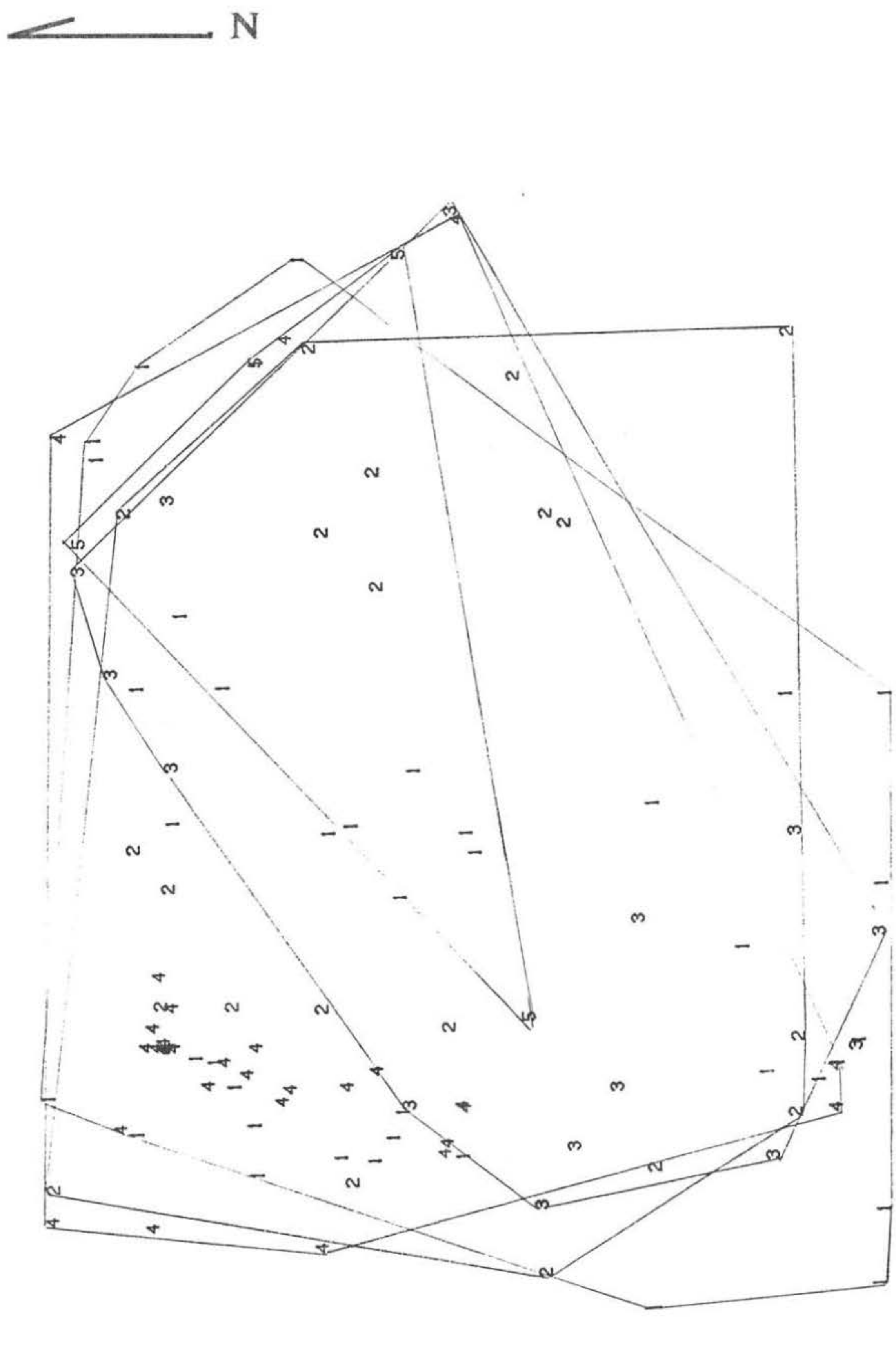
TOTAL OBSERVATIONS: 32

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	32	32	32	32	32
MINIMUM	198.000	0.948	0.663	8.400	3.400
MAXIMUM	117928.000	2.639	0.937	47.600	11.200
MEAN	33206.313	1.186	0.838	12.450	7.922
STANDARD DEV	38620.806	0.332	0.070	7.052	2.485

TOTAL OBSERVATIONS: 4

	AADT86	JANWSF	AUGWSF	HR30	PCTTRKS
N OF CASES	4	4	4	4	4
MINIMUM	1975.000	1.221	0.716	11.000	39.500
MAXIMUM	4958.000	1.643	0.790	17.500	46.000
MEAN	3591.500	1.469	0.748	12.975	42.925
STANDARD DEV	1576.067	0.188	0.031	3.036	3.305

MAP OF OREGON CLUSTERED BY PERCENT HEAVY VEHICLES



Attachment 4

CLUSTERS OF ATRs BASED ON LOG AADT, URBAN/RURAL CLASS, AND FEDERAL FUNCTIONAL CLASS

Group on URC Log AADT

TOTAL OBSERVATIONS: 110 TOTAL DATA BASE.

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	110	110	110	110	110
MINIMUM	198.000	8.400	3.400	0.948	1.000
MAXIMUM	117928.000	47.600	46.000	2.639	5.000
MEAN	13813.927	13.098	17.789	1.362	2.655
STANDARD DEV	25172.922	4.514	9.651	0.296	1.505

TOTAL OBSERVATIONS: 38 (1)

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	38	38	38	38	38
MINIMUM	2830.000	9.600	3.700	1.034	1.000
MAXIMUM	11498.000	22.000	46.000	1.967	1.000
MEAN	5689.842	13.174	18.555	1.393	1.000
STANDARD DEV	2236.765	3.426	10.597	0.232	0.000
	CV .39				

TOTAL OBSERVATIONS: 21 (2)

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	21	21	21	21	21
MINIMUM	8963.000	8.400	4.000	0.948	2.000
MAXIMUM	117928.000	13.200	34.000	1.604	2.000
MEAN	48097.143	10.381	17.371	1.184	2.000
STANDARD DEV	35323.983	1.303	10.185	0.196	0.000
	0.73				

TOTAL OBSERVATIONS: 6 (3)

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	6	6	6	6	6
MINIMUM	14511.000	9.200	3.400	0.985	3.000
MAXIMUM	102259.000	10.600	7.800	1.126	3.000
MEAN	39848.833	9.967	5.467	1.048	3.000
STANDARD DEV	33162.430	0.532	2.109	0.052	0.000
	0.83				

TOTAL OBSERVATIONS: 31 (4)

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	31	31	31	31	31
MINIMUM	198.000	10.300	5.500	0.996	4.000
MAXIMUM	2262.000	47.600	31.700	2.639	4.000
MEAN	1071.806	15.219	16.423	1.461	4.000
STANDARD DEV	612.043	6.709	6.443	0.392	0.000
	0.57				

TOTAL OBSERVATIONS: 14 (5)

	AADT86	HR30	PCTTRKS	JANWSF	CLUSTER
N OF CASES	14	14	14	14	14
MINIMUM	497.000	11.300	13.500	1.221	5.000
MAXIMUM	2504.000	17.500	40.700	1.910	5.000
MEAN	1497.071	13.614	24.643	1.464	5.000
STANDARD DEV	667.776	1.830	8.786	0.180	0.000
	0.45				

Total Coeff. of Variation = 59.1

		RECORDER TRICLASS	AADT86	HR30	PCTTRKS	JANWSF
CASE	4	1011.000	4929.000	11.600	45.500	1.643
CASE	4	151.000				
CASE	8	2007.000	4163.000	11.000	12.200	1.200
CASE	8	132.000				
CASE	13	4001.000	8611.000	13.900	6.900	1.359
CASE	13	142.000				
CASE	15	5006.000	8103.000	11.400	17.100	1.201
CASE	15	142.000				
CASE	16	6004.000	4961.000	12.600	13.900	1.253
CASE	16	142.000				
CASE	18	8005.000	5748.000	13.100	11.600	1.437
CASE	18	142.000				
CASE	19	9003.000	11498.000	10.900	13.900	1.353
CASE	19	143.000				
CASE	21	9014.000	4788.000	18.500	13.700	1.967
CASE	21	142.000				
CASE	22	10001.000	6681.000	12.700	22.200	1.286
CASE	22	142.000				
CASE	23	10003.000	3125.000	15.300	30.400	1.390
CASE	23	141.000				
CASE	25	10006.000	4964.000	9.600	21.100	1.149
CASE	25	142.000				
CASE	29	11008.000	6556.000	14.700	37.000	1.575
CASE	29	151.000				
CASE	40	15013.000	4822.000	11.900	19.100	1.505
CASE	40	133.000				
CASE	41	15014.000	6800.000	10.200	3.700	1.057
CASE	41	133.000				
CASE	45	16002.000	7953.000	11.700	16.800	1.358
CASE	45	142.000				
CASE	46	16006.000	4779.000	13.700	15.100	1.381
CASE	46	142.000				
CASE	49	18006.000	4080.000	13.200	34.300	1.693
CASE	49	141.000				
CASE	51	18019.000	2912.000	10.700	30.700	1.419
CASE	51	141.000				
CASE	52	18020.000	3606.000	11.500	14.500	1.203
CASE	52	141.000				
CASE	59	20005.000	4732.000	11.700	15.500	1.372
CASE	59	141.000				
CASE	61	20010.000	2830.000	15.700	22.500	1.567
CASE	61	141.000				
CASE	64	21002.000	7109.000	17.000	10.800	1.580
CASE	64	142.000				
CASE	65	21006.000	3123.000	12.600	23.100	1.312
CASE	65	141.000				
CASE	66	22010.000	4065.000	10.100	8.400	1.160
CASE	66	132.000				
CASE	68	22013.000	8190.000	10.500	11.200	1.132
CASE	68	142.000				
CASE	70	23006.000	4001.000	10.100	10.700	1.212
CASE	70	131.000				

CASE	74	23016.000	4958.000	11.000	46.000	1.584
CASE	74	151.000				
CASE	75	24001.000	7359.000	10.200	7.600	1.187
CASE	75	133.000				
CASE	77	24013.000	3655.000	22.000	24.600	1.814
CASE	77	141.000				
CASE	79	24016.000	4245.000	11.600	10.500	1.034
CASE	79	132.000				
CASE	94	27001.000	9923.000	15.500	12.500	1.404
CASE	94	143.000				
CASE	96	29001.000	4738.000	19.100	10.600	1.571
CASE	96	142.000				
CASE	98	30004.000	8790.000	11.000	27.400	1.311
CASE	98	151.000				
CASE	102	30021.000	10249.000	10.400	9.700	1.189
CASE	102	143.000				
CASE	103	30025.000	7297.000	10.600	36.100	1.438
CASE	103	151.000				
CASE	107	34001.000	4962.000	21.400	9.500	1.821
CASE	107	142.000				
CASE	108	34004.000	3555.000	21.500	16.100	1.770
CASE	108	132.000				
CASE	110	36005.000	3354.000	10.400	12.600	1.048
CASE	110	132.000				

		RECORDER	AADT86 (2)	HR30	PCTTRKS	TRICLASS
CASE	9	3011.000	45597.000	10.500	17.700	155.000
CASE	12	3016.000	45735.000	10.600	14.500	152.000
CASE	26	10007.000	14460.000	10.600	28.300	152.000
CASE	36	15001.000	21032.000	10.100	21.900	153.000
CASE	37	15002.000	11269.000	12.100	34.000	152.000
CASE	42	15018.000	21976.000	9.600	18.300	153.000
CASE	43	15019.000	29725.000	9.500	16.900	252.000
CASE	47	17001.000	14270.000	11.200	33.900	152.000
CASE	60	20008.000	39473.000	10.100	6.300	252.000
CASE	69	22016.000	22982.000	11.300	28.700	153.000
CASE	73	23014.000	8963.000	9.500	31.600	251.000
CASE	82	26001.000	16850.000	13.200	25.300	251.000
CASE	85	26004.000	87017.000	9.100	8.200	254.000
CASE	86	26005.000	76527.000	10.000	6.900	254.000
CASE	88	26013.000	117928.000	8.400	4.000	256.000
CASE	89	26016.000	96002.000	10.200	8.800	255.000
CASE	90	26019.000	99778.000	8.500	9.900	255.000
CASE	91	26024.000	58832.000	11.100	8.400	253.000
CASE	92	26026.000	87083.000	9.000	5.200	254.000
CASE	93	26027.000	82941.000	10.300	10.000	254.000
CASE	105	33001.000	11600.000	13.100	26.000	152.000

		RECORDER	AADT86 (3)	HR30	PCTTRKS	TRICLASS
CASE	57	20003.000	22209.000	10.100	3.400	236.000
CASE	76	24004.000	14511.000	10.600	7.800	235.000
CASE	78	24014.000	51486.000	10.000	3.700	232.000
CASE	83	26002.000	102259.000	9.200	3.600	229.000
CASE	84	26003.000	28342.000	9.500	7.500	235.000
CASE	109	36004.000	20286.000	10.400	6.800	235.000

4

		RECORDER	AADT86	HR30	PCTTRKS	TRICLASS
CASE	2	1007.000	198.000	10.600	6.500	121.000
CASE	3	1010.000	569.000	13.900	13.900	121.000
CASE	5	1012.000	1366.000	17.100	22.200	131.000
CASE	6	2003.000	2022.000	11.500	12.900	131.000
CASE	7	2005.000	813.000	11.600	13.800	131.000
CASE	10	3013.000	1928.000	11.000	10.400	131.000
CASE	11	3014.000	1808.000	11.000	26.600	131.000
CASE	14	4010.000	349.000	15.200	27.100	121.000
CASE	24	10004.000	1330.000	15.600	31.700	131.000
CASE	27	11004.000	208.000	15.400	12.500	121.000
CASE	28	11007.000	357.000	14.000	15.200	131.000
CASE	30	12003.000	724.000	16.300	24.400	131.000
CASE	32	12009.000	948.000	13.600	20.200	132.000
CASE	35	14003.000	739.000	24.500	22.300	131.000
CASE	38	15007.000	962.000	15.000	14.100	131.000
CASE	39	15011.000	2262.000	10.800	5.500	131.000
CASE	50	18017.000	984.000	14.100	29.400	131.000
CASE	53	18021.000	521.000	20.500	12.500	121.000
CASE	54	19004.000	735.000	14.100	19.300	131.000
CASE	55	19008.000	879.000	13.200	13.000	131.000
CASE	56	19010.000	580.000	14.500	17.400	131.000
CASE	58	20004.000	1675.000	13.000	9.900	121.000
CASE	63	20023.000	1156.000	14.400	10.700	131.000
CASE	67	22012.000	2158.000	10.300	14.300	131.000
CASE	80	24020.000	1612.000	13.200	17.000	131.000
CASE	81	25007.000	1344.000	11.400	15.200	131.000
CASE	87	26012.000	380.000	47.600	10.400	121.000
CASE	100	30012.000	984.000	18.100	13.000	131.000
CASE	101	30016.000	319.000	13.500	15.900	121.000
CASE	104	31005.000	1596.000	13.800	18.300	136.000
CASE	106	33005.000	1720.000	13.000	13.500	131.000

5

		RECORDER	AADT86	HR30	PCTTRKS	TRICLASS
CASE	1	1001.000	761.000	11.700	13.500	111.000
CASE	17	7001.000	2207.000	14.100	16.500	141.000
CASE	20	9005.000	1577.000	12.400	14.300	141.000
CASE	31	12006.000	497.000	14.500	20.900	141.000
CASE	33	13001.000	743.000	13.400	21.500	141.000
CASE	34	13003.000	1667.000	12.800	22.700	141.000
CASE	44	15020.000	2032.000	13.100	27.200	141.000
CASE	48	17003.000	2351.000	17.100	14.100	141.000
CASE	62	20017.000	2504.000	17.500	40.700	141.000
CASE	71	23012.000	922.000	13.100	26.300	141.000
CASE	72	23013.000	980.000	13.700	26.400	141.000
CASE	95	28001.000	1732.000	11.300	31.900	141.000
CASE	97	30002.000	1975.000	11.800	39.500	141.000
CASE	99	30007.000	1011.000	14.100	29.500	141.000

ATTACHMENT 5

PRELIMINARY MULTIPLE LINEAR REGRESSION RESULTS

Model 1 estimates January seasonal factor as a function of Urban/rural class, federal functional class, and log of AADT, and their interaction terms. Urban/rural class and the interaction of log of AADT and federal functional class are significant, but the explanatory power of the model is poor. This is due to heteroscedasticity, which means the error terms were not uniformly districted.

Model 2 reduces the heteroscedasticity problem by deflating all variables by some measure of "size". All variables were divided by log of AADT. Thus, Model 2 estimates January weekly traffic as a function of Urban/rural class divided by log AADT, federal functional class divided by log of AADT, one over log AADT and the interaction terms. Further specification of this model will be done in Phase II.

MODEL ONE

ASSUMING MIXTURE MODEL.

DEP VAR: JANWSF N: 110 MULTIPLE R: .518 SQUARED MULTIPLE R: .269
ADJUSTED SQUARED MULTIPLE R: .226 STANDARD ERROR OF ESTIMATE: 0.260

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
URC	2.155	0.643	9.009	0.0009815	3.349	0.001
FED	0.259	0.250	3.386	0.0006658	1.037	0.302
LOGADT	0.073	0.072	2.117	0.0016683	1.026	0.307
URC* FED	-0.278	0.244	-4.721	0.0004144	-1.141	0.257
URC* LOGADT	-0.197	0.047	-7.693	0.0020938	-4.177	0.000
FED* LOGADT	-0.011	0.028	-1.281	0.0006788	-0.396	0.693
URC* FED* LOGADT	0.022	0.020	3.573	0.0006589	1.088	0.279

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	2.565	6	0.428	6.302	0.000
RESIDUAL	6.988	103	0.068		