NATIONAL TRAFFIC DATA ACQUISITION CONFERENCE (NATDAC '96), PROCEEDINGS (VOL. II)

NOTE TO READER:

THIS IS A LARGE DOCUMENT

Due to its large size, this document has been segmented into multiple files. All files separate from this main document file are accessible from links (blue type) in the table of contents or the body of the document.

PROCEEDINGS

National Traffic Data Acquisition Conference

Albuquerque
New Mexico

May 5-9, 1996

N A T D A C

National Traffic Data Acquisition Conference

The New Mexico State Highway and Transportation Department

Alliance For Transportation Research

The Federal Highway Administration

NATIONAL TRAFFIC DATA ACQUISITION CONFERENCE (NATDAC '96)

PROCEEDINGS
Volume II

Albuquerque, New Mexico
May 5-9, 1996

Hosted By:

Alliance For Transportation Research (ATR)
New Mexico State Highway and Transportation Department (NMSHTD)
Federal Highway Administration (FHWA)

Sponsored By:

American Association of State Highway and Transportation Officials (AASHTO)
Transportation Research Board (TRB)
Institute of Transportation Engineers (ITE)

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NM-NATDAC-96		NATDAC-96

4. Title and Subtitle

National Traffic Data Acquisition Conference (NATDAC '96) Proceedings, Volume II, Albuquerque, New Mexico, May 5-9, 1996

5. Report Date

6. Performing Organization Code

CN9981

7. Author

Geri Knoebel

8. Performing Organization Report

NATDAC-96

9. Performing Organization Name and Address

Alliance for Transporation Research
1001 University Blvd, SE, Ste 103
Albuquerque, NM 87106

10. Work Unit No.

11. Contract No.

SPR-PL-RD-0002(33)NM

12. Sponsoring Agency Name and Address

New Mexico State Highway & Transportation Department
PO Box 1149
Santa Fe, NM 87504

13. Report Type and Period Covered

Final Report
May 5-9, 1996

14. Sponsor Code

CN9981

15. Supplementary Notes

This project was co-sponsored by the Alliance for Transportation Research, New Mexico State Highway and Transportation Department, and the Federal Highway Administration.

16. Abstract

The National Traffic Data Acquisition Conference 1996 (NATDAC '96) was held in Albuquerque, New Mexico on May 5-9 1996. A broad range of topics was covered during the conference, including: National Travel Trends, Travel Time Data Collection Using GPS, Congestion Management System Data Issues, Detector Technologies, and Metropolitan Travel Data. Concurrent sessions included Technology and Data Collection, Traffic Monitoring Systems, Weigh-In-Motion, and Enforcement Issues. There were 119 presenters, panelists, and moderators involved in the program. Conference participants were transported to three field sites for demonstrations of data collection equipment from 26 vendors. Thirty four vendors maintained displays and exhibits at the conference facility.

The 378 registered participants represented 48 states, the federal government, and industry. Twenty international participants attended.

The papers within this publication are a compilatation of those presented at NATDAC '96. It is presented in two volumes. Volume I contains all General Sessions papers as well as Track A Concurrent Sessions papers. Volume II contains Track B and Track C Concurrent Sessions papers.

17. Key Words	18. Distribution Statement
Global Positioning Systems Congestion Management System Traffic Detection Technologies Weigh-In-Motion	No restrictions. This document is available to the public through the National Technical Information Service. Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	700	

Form DOT F 1700.7

PREFACE

The National Traffic Data Acquisition Conference, 1996 (NATDAC '96) was held in Albuquerque, New Mexico on May 5-9, 1996. The Conference was co-sponsored by the Federal Highway Administration, the Alliance for Transportation Research, and the New Mexico State Highway and Transportation Department.

Presentations covered a broad range of topics related to traffic monitoring. General sessions included presentations covering policies, perspectives, and innovations of interest to a wide audience. Three concurrent session tracks were more specific in nature. Track "A" sessions focused on traffic monitoring in urban areas - congestion and other performance monitoring activities, including travel time and the unique challenges of data collection in and urban environment. Track "B" sessions were oriented toward traffic monitoring program management and data analysis with particular attention to vehicle classification data. Track C" sessions focused on weigh-in-motion equipment and truck weight issues including Long-Term Pavement Performance (LTPP) project and data collection for weight enforcement.

There were 378 registered participants representing 48 states, the federal government, academia, and private industry. Twenty international participants attended. Thirty-four vendors maintained exhibition booths. One day was dedicated to field demonstrations where participation vendors demonstrated devices on the roadway for delegates to view. Altogether, NATDAC '96 had 119 presenters, panelists, and moderators involved in the program.

The papers within this publication are a compilation of those presented at NATDAC '96. It is presented in two volumes. Volume I has General Sessions I, II, III, and Track "A". Volume II contains Track "B" and Track "C".

Thank you to all the speakers, participants, and exhibitors who made the conference an outstanding success.

The contents of this report reflect the views of the authors of the papers and abstracts, who are responsible for the facts and accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the New Mexico State Highway and Transportation Department or the Federal Highway Administration. This report does not constitute a standard, specification or regulation.

Neither the United States government, nor the State of New Mexico endorse products or manufacturers. Trade or manufacturer names appear herein only because they are considered essential to the objec of this document.

TABLE OF CONTENTS -VOLUME II

TRACK "B"
CONCURRENT SESSION IB - DATA COLLECTION ISSUES

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

CLASSIFICATION ALGORITHMS/VEHICLE CLASSIFIER ACCURACY

Speaker: Bruce A. Harvey
Georgia Institute of Technology
Authors: Bruce A. Harvey, et al.
Georgia Institute of Technology

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Classification Algorithms / Vehicle Classifier Accuracy

Bruce A. Harvey, Ph.D. and Glenn H. Champion
ITL/CND/Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, Georgia 30332-0821
Phone: 404-894-5597
FAX: 404-894-3906

ABSTRACT:

This paper describes the results of field tests of a number of common vehicle classification devices currently in use throughout the United States. The objectives of these tests were to determine the accuracy of vehicle counting devices, the adequacy of equipment to correctly sort vehicles into the 13 FHWA vehicle classes, the accuracy of automatic measurement of overall vehicle length, the effects of vehicle and axle sensor technology on the accuracy of the vehicle classification, and the effects of vehicle repetitions and weather on axle sensors. These tests were performed by the Georgia Tech Research Institute (GTRI) and the Georgia Department of Transportation (GDOT). Funding for this project was provided by the Federal Highway Administration (FHWA) Office of Highway Information Management.

The tests were conducted on I-20 a high-volume interstate in the metro Atlanta area where there was a good mix of vehicles. All FHWA vehicle classes were recorded during these tests except triple trailers which are illegal in the state of Georgia. All the devices were installed in a single lane and testing was conducted simultaneously on the same traffic stream to ensure a fair comparison of the devices. The tests included two continuous 48-hour tests for detailed vehicle-by-vehicle analysis, a 7-day test to determine long term effects on the vehicle classification devices, and a brief test after pavement overlay at the test site.

Data reduction for these tests included vehicle-by-vehicle comparisons of vendor classification to ground truth data. Ground truth data was collected using computer-assisted analysis of each vehicle recorded on video by a camera mounted above and to the side of the traffic lane. Markings on the roadway were used to derive vehicle length and axle spacing on each of the vehicles.

The test setup, analysis methods and results of the tests as well as the problems encountered are presented in this paper. This paper will focus on the classification accuracy of the devices measured during the tests and the classification errors common to many of the devices tested.

BACKGROUND & PURPOSE OF TESTS

Traffic monitoring equipment and especially automatic vehicle classification (AVC) equipment are a very important part of an overall traffic management system. Many devices are currently being marketed which are designed to not only monitor traffic density and speed, but also to classify vehicle into the 13 FHWA vehicle classes. The accuracy of the classification function of these devices should be independently verified under normal operating conditions. The tests described in this paper were used to evaluate nine (9) vendor's equipments including 14 sensor sets of these AVC devices in a side-by-side test under normal traffic conditions.

The tests performed under this program were designed to update the results from the testing of vehicle classification devices by the State of Maine in 1984 for the FHWA. (I) New technologies included in equipment for the current tests included new types of sensors (such as piezoelectric axle sensors), and programmable classifiers that allow the user to specify the dimensional thresholds for various vehicle types and to retain individual vehicle information.

The devices being tested in this project were commercially available, off-the shelf devices representing the current state-of-the-application classification devices as of September 1992, when the project began. Participating vendors of vehicle classification devices loaned their classification equipment, and most participated in the installation, check-out and calibration of their equipment at the test site. The sensors used for the tests were specified by the vendor and purchased by the project.

The objectives of the test were to:

- Determine the accuracy of vehicle counting devices.
- Determine the accuracy of various types of equipment to correctly sort vehicles into the 13 FHWA vehicle classes (as identified by the FHWA Traffic Monitoring Guide).
- Determine the accuracy of automatic measurement of axle spacings, wheel base length, and overall vehicle length.
- Determine how vehicle and axle sensor technology affects the accuracy of the vehicle classification.
- Determine the effects of vehicle repetitions, heavy axle loadings, and weather on pneumatic tube axle sensors.

The final report for this project has been published and submitted to the FHWA. (2) A supplementary report has been published summarizing the result of the pavement overlay test. (3) A paper was presented at NATDAC '94 containing a description of the tests and presenting preliminary general results. (4) This paper summarizes the fmal results from the project. Individual vendor results are not provided in this paper, but are available in the final report.

TEST SET-UP & INSTALLATION

The tests were conducted on the "slow" or outside lane of westbound Interstate 20 near Covington Georgia, 30 miles east of the metro Atlanta area. This particular location was chosen for several practical reasons including available shoulder space for installing an equipment trailer, available power, and a conveniently located overpass. This section of the roadway met the conditions for the test including radius of curvature greater than 1740 meters (5700 feet), a longitudinal gradient of less than 2 percent, and a cross-slope of less then 2 percent with 46 meters (150 feet) available for sensor locations. The width of the roadway lanes were between 3 and 3.7 meters (10 and 12 feet), the pavement surface was relatively smooth, and the normal traffic flow was in the range of 45 and 105 kilometers per hour (30 to 65 miles per hour). The test site also had a good mix of vehicles for all FHWA vehicle classes except triple trailers (FHWA Class 13) which are illegal in the state of Georgia. Average daily traffic for the test lane was approximately 10,000 vehicles.

Between December 1992 and April 1993, the sensors required by each of the classification device vendors were installed by GDOT, mostly under the supervision of vendor representatives or with provided instructions. The vendors participating in the test used only piezoelectric axle sensors, or some combination of piezoelectric axle sensors and loop sensors. Table 1 contains a list of the vendors participating in the tests and the sensor configuration used.

Table 1. Equipment Vendors and Configurations

VENDOR	MODELS	CONFIGURATION
Peek Traffic, Inc.	TrafiCOMP III GK-6000	P-L-P P-L-P L-P-L
Mikros Systems	TEL-2CM	L-P-L
PAT Equipment Corporation, Inc.	AVC-100 AVC-100	P-L-P L-P-L
Diamond Traffic Products	TT-2001	P-L-P P-L-P
International Road Dynamics, Inc.	TC/C 530-4D/4P/4L	PR-L-PR P-L-P
Mitron Systems Corp.	MSC-3000 DCP	P-P
Golden River Traffic, Ltd.	Marksman 660	P-L-P
Electronic Control Measure	HESTIA	P-L-P
TimeMark, Inc.	Delta II	P-P

The sensor configuration is depicted using "P" to indicate a piezoelectric axle sensor (voltage output), an "L" to depict a inductive loop and "PR" to denote a piezoelectric resistive device. Therefore, a "P-L-P" configuration uses two piezoelectric sensors with a loop in between.

To monitor the traffic at the test site, two video cameras were installed. The first camera was installed on an overpass at the end of the test site to monitor vehicles changing lanes within the test area. The second was mounted on a utility pole that was set in place by GDOT off the side of the road pointing down at the test lane at a 30" angle. Two specially mounted street lights were used to illuminate the test lane at night. The pole camera was used to record the individual vehicles in the traffic stream as they passed through the test site. The video tapes of the traffic stream were used during the data reduction to form the ground truth (or reference) data which listed vehicle class, axle spacings and overall vehicle length for each vehicle which passed through the test site.

Figure 1 depicts the test set-up used for this project. A GDOT mobile trailer was used to house computer equipment, the video recording equipment for the pole camera, and the test personnel. The pole camera was mounted approximately 13.4 meters (44 feet) above the surface of the road and 21 meters (69 feet) from the side of the road. The classification equipment was installed along the right-hand lane of the highway with spacing sufficient to insure no interference between vendor equipment (60 feet between loop sensors of different vendor systems). Some of the vendor's equipment had to be modified (primarily software modifications) to provide individual vehicle records. In some cases, portable computers were required to provide the storage capacity needed to store individual vehicle records.

A camera was mounted on an overpass near the end of the test site to provide a record of vehicle changing lanes within the test site. This data was used to eliminate vehicles from the test which enter or leave the right lane within the test site. The resulting ground truth data was a file listing the vehicles which passed over all of the classification sensors and could be classified by the video from the pole camera.

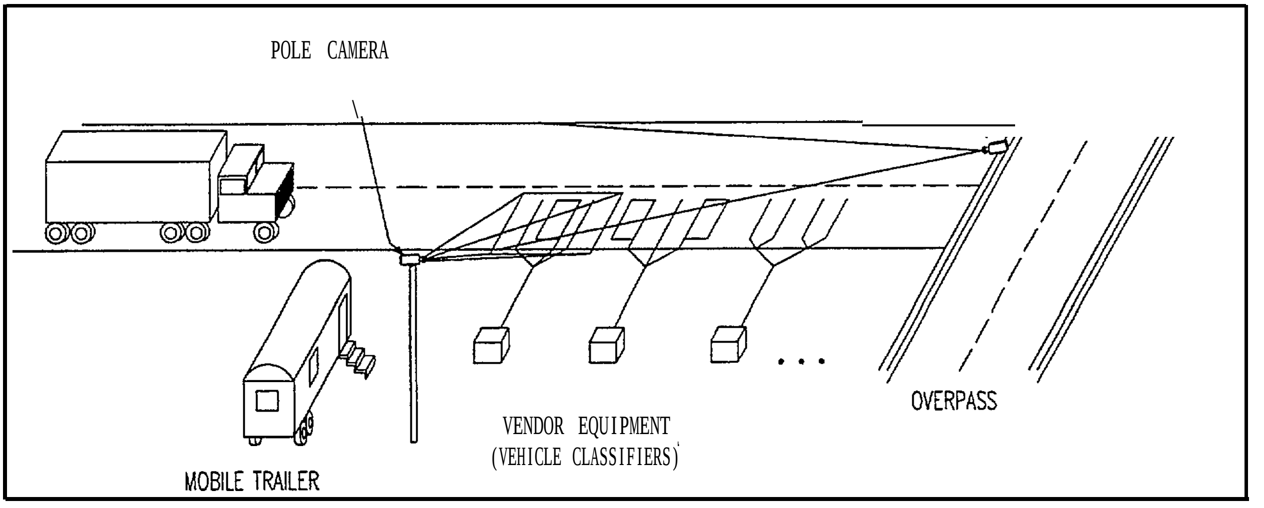


Figure 1. Vehicle Classification Test Site

DESCRIPTION OF TESTS

The required tests for the project included two 48-hour tests for detailed vehicle-by-vehicle analysis and a 7-day test to obtain longer term accuracy statistics. The classifier data collected during the 48-hour tests was compared on a vehicle-by-vehicle basis with the ground truth data set derived from the pole camera and the lane-change camera. During the 7-day tests, the classifiers were programmed to collect vehicle counts for each vehicle class over 15 minute intervals. This binned data was then compared to similarly binned ground truth data to obtain longer term statistics for each of the classification systems.

The first 4%hour test was conducted on May 5 - 7, 1993 and the second 48-hour test was conducted on September 9-11, 1993. Temperatures were moderate to warm for each of these tests, and no measurable precipitation occurred during either test.

The augmented pneumatic tube test was conducted in parallel with the second 48-hour test. Two pairs of pneumatic tubes attached to a single classifier were used to monitor the traffic in both westbound lanes through the test site. One pair was used to monitor the outside (slow) lane, and one pair was placed across both lanes. The purpose of the test was to determine if the traffic in multiple individual lanes could be accurately monitored using the pneumatic tubes. Technical problems with the data collection prevented the successful analysis of the data from the pneumatic tube tests The Georgia DOT repeated this test after the completion of this project and the data is not available for this paper.

The 7-day test was conducted in conjunction with the second 48-hour test. Binned data on 15 minute intervals were recorded for five days following the second 48 hour test. The data from each vendor's equipment for the first 4%hours of the test were converted to binned data to give a full seven days of binned data. This data was then used to compile long term statistics on the performance of each classifier.

A pavement overlay test was performed following improvements to the highway at the test site. The improvements included repaving the test lane. The repaving included a leveling course of pavement (thickness ranging 0.5 to 1.4 inches), a 1 l/2 inch layer of dense asphaltic concrete (B-mix) and a 3/4 inch layer of open-graded friction course (D-mix). Axle sensor output voltages were measured and classifier performance was assessed after the leveling course and after the layer of E-mix was applied and before D-mix was applied.

DATA REDUCTION

The data reduction required for this project consisted of two phases for each test. The first phase was the reduction of the video data to a baseline set of vehicles with classifications, length and axle spacings. The baseline set of data is referred to as the ground truth data for the test. The second phase involves comparing the ground truth data to the data collected by each of the vendor's classification equipment.

Video data reduction from the pole camera was accomplished using the Computer Vehicle Classification/Reduction System (CVCRS) developed by GTRI for this project. The CVCRS consisted of a PC/486 computer with a video capture and processing card installed, a second VGA monitor, and a video cassette recorder (VCR). Custom software was written to allow a user to measure overall vehicle length, measure axle spacings, and classify each vehicle in an efficient manner.

Calibration marks in the test lane allowed the user to calibrate the CVCRS measurements and to correct for the viewing angle of the pole camera. The pole camera video tape of the traffic stream included a time stamp which was read by the CVCRS and stored with each vehicle record. The accuracy of the axle spacing measurements made by the CVCRS was assessed using a specially marked and manually measured test vehicle provided by the GDOT. The test vehicle was driven through the test site several times at varying speeds during the first 4%hour test. The tests revealed that the mean error of the CVCRS measurements was 6.1 cm (2.4 inches) and the standard deviation was 3.3 cm (1.3 inches).

Video data reduction from the lane change camera consisted of identifying vehicles which entered or exited the test lane within the length of the test site. The vehicles which did change lanes were removed from the ground truth data record for the 48-hour tests. The resulting data set included only vehicles which passed completely through the test site without changing lanes.

The second phase of the analysis began with the conversion of data from the specific formats of the individual members to a single, standard format for analysis. The results of the individual classifiers were then compared against the ground truth data and a statistical analysis performed.

The binned data from the 7-day test was used to determine the long term accuracy of the classifiers tested. The total vehicles counted in each class and the number of axles counted were compared against the ground truth data for the 7-day test. Total count volumes and percent differences were compared to determine the counting accuracy of the classifiers. Also, statistics on the accuracy of the classifiers during the first day of the test were compared against the accuracies during the last day of the test to determine if the accuracy was affected by the length of time in operation of the classifiers.

The peak outputs of the piezoelectric axle sensors were measured during initial installation of the equipment using an oscilloscope. After the leveling course and E-mix layers of pavement were completed during highway improvements, these measurements were repeated for comparison. A manual assessment of the classification performance of the working systems was performed after the leveling course. After the E-mix layer was applied, 4-hour binned tests were performed on some the working classifiers in order to assess classifier accuracy.

CLASSIFICATION ACCURACY RESULTS FROM TESTS

The following results summarize the performance of the vehicle classification systems tested under this program. To avoid the appearance of a competitive evaluation, the names of the companies and equipments have been removed from the results. Each individual classifier is identified by a letter (A, B, . ..) which remains constant throughout the presentation of the results. Also, all piezoelectric axle sensors are indicated by the letter 'P', including the resistive sensors. The specific results of each individual classifier is presented in this project's final report. (2)

Overall Classification Accuracy

The classification accuracy for each system was summarized two ways. First, a classification matrix was generated which compared the actual classification of vehicles (ground truth) to the classification determined by the system under test. Then, the overall classification accuracy was calculated (ii percent) for each system.

Figure 2 shows a typical classification matrix derived for one of the vehicle classifiers tested. The matrix is divided into two sub-matrices. The top sub-matrix displays the actual vehicle count while the bottom sub-matrix displays the same data as a percent of the total vehicles of each class (from ground truth). The first column in each sub-matrix indicates the actual (ground truth) classification of the vehicles. The top row of each sub-matrix indicates the vehicle classification assigned by the system under test. Therefore, the number located in row '2', column '3' of the top sub-matrix is the number of class 2 (passenger cars) which were classified as class 3 (pickups, vans) by the system being tested. In the bottom sub-matrix, the number is the percent of class 2 vehicles classified as class 3.

Two general results were derived from these classification matrices. First, the most common classification errors were between class 2 and class 3 vehicles. This error is understandable since both of these classes consist of Z-axle, 4-tire vehicles with often identical wheelbases. Second, the most accurately classified vehicle classes were the large trucks (classes 8 - 12).

Table 2 summarizes the overall classification accuracy of the systems tested. Included in this table is the percentage of axle sensor errors (vehicles with miscounted axles) and the percentage of correctly classified vehicles recorded for each system. Classification accuracy of the systems if classes 2 and 3 were combined (assumed to be a single class of vehicles) are included in parentheses. As expected from the classification matrices, the classification accuracy is significantly higher when classes 2 and 3 are combined into a single class of vehicles.

Classification Accuracy Versus Axle Sensor Errors

A strong correlation between the accuracy of a classifier and the reliability of the axle sensor used was identified during this test. To quantify this correlation, the classification accuracy of each configuration tested was plotted versus the percentage of axle sensor errors (defmed as miscounted axles on a vehicle). Figure 3 depicts the classification accuracy as a function of percent axle sensor errors with and without combining vehicle classes 2 and 3.. With the exception of one or two systems, this figure shows that axle sensor reliability has a very strong effect on the performance of the vehicle classifiers. It appears that if axle sensor error were eliminated, most of the classifiers would be nearly 98% accurate with classes 2 and 3 combined. This result indicates that the axle sensor installation is critical to the performance of the vehicle classification systems.

Classes Determined by AVC Equipment

COUNT

Actual (Ground Truth) Vehicle Classes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	5501	174	0	6	1	2	2	10	1	1	0	0	0	0
3	0	2080	1541	4	219	2	2	19	4	0	0	0	0	0	0
4	0	1	1	26	3	6	0	0	0	0	0	0	0	0	0
5	0	5	45	73	176	0	2	17	3	0	0	0	0	0	0
6	0	2	1	6	0	168	0	1	7	0	0	0	0	0	0
7	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
8	0	2	6	1	0	1	3	231	0	0	0	0	0	0	0
9	0	6	4	0	1	0	2	3	2575	0	2	0	0	0	0
10	0	0	1	0	0	0	0	0	3	17	0	1	0	0	0
11	0	0	0	0	0	0	1	0	1	0	63	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0
13	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0

PERCENT

Actual (Ground Truth) Vehicle Classes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0	66.7	33.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	96.5	3.1	0.0	0.1	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	53.7	39.8	0.1	5.7	0.1	0.1	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	2.7	2.7	70.3	8.1	16.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	1.6	14.0	22.7	54.8	0.0	0.6	5.3	0.9	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	1.1	0.5	3.2	0.0	90.8	0.0	0.5	3.8	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.8	2.5	0.4	0.0	0.4	1.2	94.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.2	0.2	0.0	0.0	0.0	0.1	0.1	99.3	0.0	0.1	0.0	0.0	0.0	0.0
10	0.0	0.0	4.5	0.0	0.0	0.0	0.0	0.0	13.6	77.3	0.0	4.5	0.0	0.0	0.0
11	0.0	0.0	0.0	0.0	0.0	0.0	1.5	0.0	1.5	0.0	96.9	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
14	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	50.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0

Figure 2. Example Classification Matrix

Classification Accuracy Versus Percentage of Trucks in Traffic Stream

The classification matrices (Figure 2, columns 8-12) indicated that trucks are more accurately classified than are smaller vehicles. The effects of this accuracy can be shown by comparing overall vehicle classification accuracy to the percentage of vehicles with greater than two axles (trucks) in the traffic stream. The average vehicle classification accuracy versus the percentage of trucks in the traffic stream (determined on 15 minute intervals) is depicted in Figure 4. It can be seen from this figure that there is little difference in the classification accuracy if classes 2 and 3 are combined. Therefore, the classification accuracy noted for trucks is not due to the weight or size of the vehicles, but rather is due to the distinct separation in the number and spacings of axles as defined by the FHWA vehicle classification guide. All of the classifiers tested used axle spacing, wheelbase and number of axles to determine vehicle classes. Classes 2 and 3 are simply not "distinct" enough in terms of these parameters to accurately differentiate between many vehicles in these classes. Other parameters, such as weight or height, may be useful in differentiating between vehicle classes 2 and 3.

Table 2. Overall Classification Accuracy

Vendor	Sensor Config	48-Hour Test #	% Sensor Errors	% Correctly Classified
A	L-P-L	1	13.91	70.3 (82.5)
		2	4.81	63.5 (78.8)
B	P-L-P	1	3.71	75.3 (93.0)
		2	3.53	74.8 (93.2)
C	P-P	1	6.29	73.7 (90.1)
		2	1.19	79.0 (96.2)
D	P-P	1	4.02	77.1 (92.4)
		2	1.07	79.1 (96.2)
E	P-L-P	1	N/A	N/A
		2	N/A	N/A
F	L-P-L	1	1.49	76.6 (95.0)
		2	0.51	76.6 (95.1)
G	P-P	1	N/A	N/A
		2	N/A	N/A
H	P-L-P	1	10.05	67.5 (86.4)
		2	6.70	72.0 (94.3)
J	P-P	1	2.83	78.9 (94.4)
		2	1.97	77.3 (94.6)
K	P-L-P	1	8.89	69.3 (88.4)
		2	5.35	72.6 (92.9)
L	P-L-P	1	7.25	70.8 (89.7)
		2	4.37	73.8 (93.9)
M	P-L-P	1	N/A	N/A
		2	7.91	63.9 (82.3)
N	P-L-P	1	3.17	73.1 (92.0)
		2	11.65	*
O	P-L-P	1	3.47	75.5 (93.9)
		2	3.4	*

* - Software Bug During Test 2
N/A - Not Available for This Test
(#) - Percent Correctly Classified with Classes 2 & 3 Combined

Accuracy of Axle Spacing Measurements

Another objective of these tests was to determine accuracy of the axle spacing measurements made by the classification systems tested. The average axle spacing measurement error and standard deviation of the errors for each system tested was calculated by comparison with ground truth data. Note that vehicles for which the axle count was incorrect were necessarily removed from consideration in this analysis. With very few exceptions, the errors and standard deviation calculated were on the same order as those for the overall length of the test vehicle used to validate the data reduction. The average axle spacing measurement error and standard deviation for the test vehicle measured by the CVCRS was 6.1 cm and 3.3 cm.

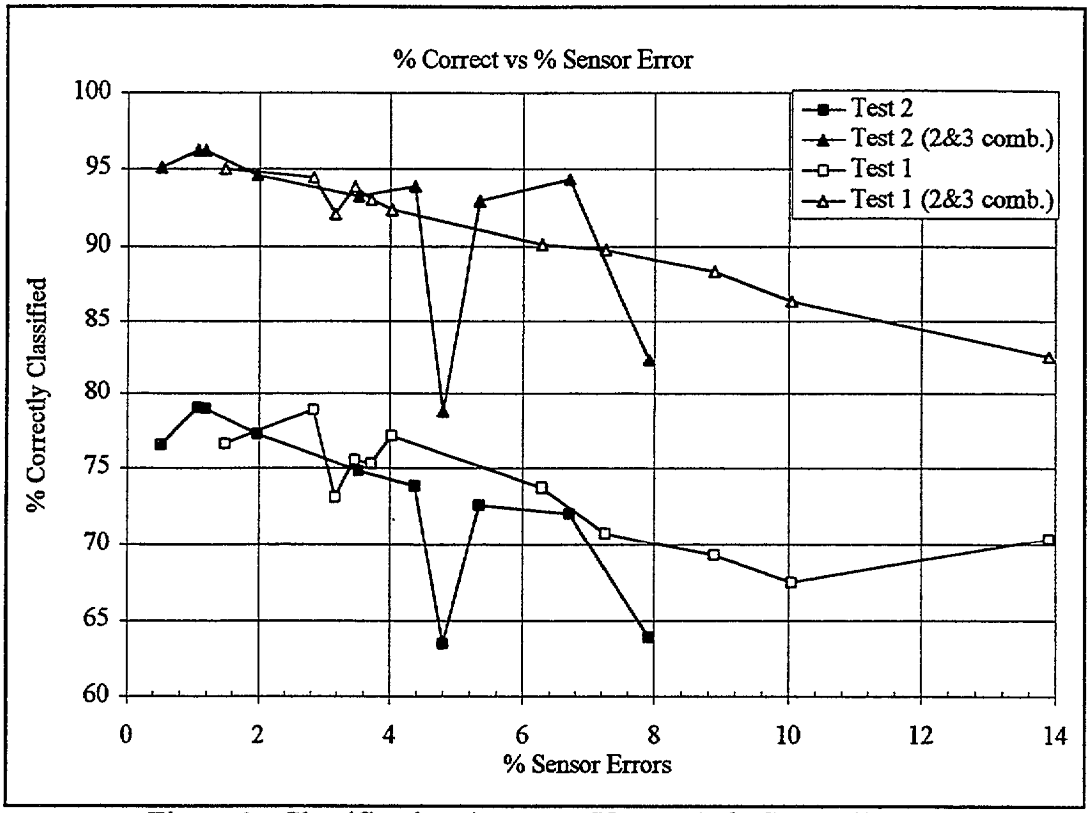


Figure 3. Classification Accuracy Versus Axle Sensor Errors

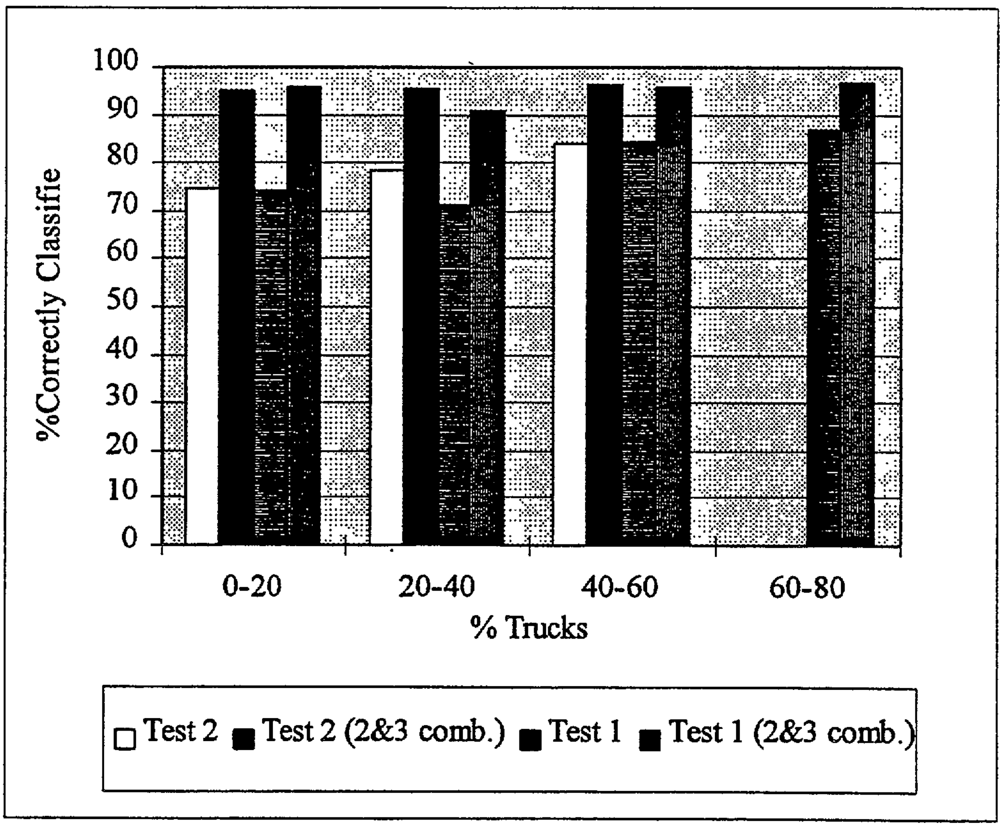


Figure 4. Classification Accuracy vs. % Vehicles With Greater Than 2 Axles

Accuracy of Overall Length Measurements

Some of the devices tested measure overall, bumper-to-bumper length of the vehicles. The accuracy of these measurements (average error and standard deviation) were calculated as part of this project. Table 3 lists the results of these calculations.

The accuracy of the classifiers in measuring overall length was considerably worse than the accuracy in measuring axle spacings. This is logical given that the overall length measurements must be made using the outputs of the inductive loops. The axle spacing measurements rely primarily on the output from precise tire contact with the narrow piezoelectric axle sensors.

Table 3. Overall Length Measurement Errors
(All Measurements in Feet)

Vendor	Sensor Config.	48 Hour Test #	Mean Error	Standard Deviation
A	L-P-L	1	2.093	2.549
		2	-2.167	4.031
H	P-L-P	1	0.598	2.157
		2	2.232	3.073
J	P-L-P	1	-1.321	11.302
		2	-1.739	10.952
K	P-L-P	1	-1.154	11.423
		2	-2.152	10.493
N	P-L-P	1	-1.271	a.747
		2	-3.717	4.613
0	P-L-P	1	-1.196	5.845
		2	-2.382	1.600

Long-Term Count/Classification Accuracy Results

The performance of the classifiers as a function of time after installation is of great importance for traffic planning and monitoring. The accuracy of the classifiers as a function of time was assessed in this performance by comparing the accuracies of the classifiers during first 48-hour test to the accuracies during second 48-hour test, and by reviewing the accuracies of the first and last days of the 7-day test.

The classification accuracy of most systems improved from the first to the second 48-hour test. This was primarily due to rehabilitation of the piezoelectric axle sensors for the second test. Asphaltic tape was placed on the piezoelectric sensors on locations where the sensor height was near or below that of the roadway. This improved the outputs of several of the piezoelectric sensors. Also, several broken piezoelectric sensors were replaced between the tests. There was no significant trend in accuracy for the classifiers for which no rehabilitation or changes were made to the axle sensors.

Comparing the performance of the classifiers on the first and last days of the 7-day test showed no general trend in accuracy versus time. Some of the vendor systems improved over the seven days, while other had declining performance. Most of the time the change was nearly insignificant. It is concluded, therefore, that the accuracy of the classifier/sensor combination is dependent on the particular system and installation. No general trends were evident.

RESULTS OF THE OVERLAY TESTS

The overlay tests were performed to determine how well the vehicle classifiers performed after the sensors (loop and piezoelectric) were overlaid with new pavement. The highway improvements at the test site afforded a unique opportunity to study the effects of pavement overlay on a number of sensors simultaneously.

The analysis of the piezoelectric axle sensor effects due to pavement overlay was conducted by measuring the output voltage of the sensors at installation, after the leveling course of pavement was installed and again after the E-mix layer was completed. Such voltage measurements were made and recorded on only 6 of the piezoelectric axle sensors during original installation. Typical output voltages were up to 400 mV and 2000 mV for class 2 and class 9 vehicles, respectively. After the leveling course was applied the maximum recorded output voltages were 250 mV and 700 mV for vehicle classes 2 and 9, respectively. After the E-mix pavement layer was completed, the maximum output voltages measured from the piezoelectric axle sensors were 100 mV and 400 mV for vehicle classes 2 and 9, respectively.

The ability of the classifiers to accurately classify vehicles was, for all practical purposes, negated by the addition of pavement layers. After the leveling course, a quick assessment of the performance of the classifiers was conducted by manually observing the classification outputs. The results were discussed with the vendors and adjustments were made to sensitivities and sensor gains where possible. A 4-hour classification test was performed on some of the working systems after the E-mix layer was applied and after all adjustments were made.

No classifier tested could achieve adequate classification accuracies after all adjustments were made following the E-mix pavement laydown. Some classifiers detected no vehicles, others detected only a few heavy truck axles, and some detected multiple extra axles (apparently due to noise detected after sensitivities were lowered). Only two systems performed well enough to calculate classification accuracy performance. Both of these classifiers over counted vehicles by more the 20%, over counted axles by more than 10% and had very poor classification performance.

The overall conclusion from the overlay tests is that there is insufficient adjustments and filtering in the systems tested to compensate for the reduced output voltages of the piezoelectric axles sensors after pavement overlay. With these systems, new axle sensor would have to be installed if the traffic monitoring site was overlaid with new pavement.

OTHER GENERAL RESULTS

A few other general results can be derived from the installation of the classifiers and the conduct of the tests. Some of these conclusions are as follows:

- Experiences from the installation and check-out of the sensors and classifiers resulted in several conclusions of interest. First, while the inductive loops caused no problems, roughly l/3 of piezoelectric axle sensors installed failed and had to be replaced. Also, the in-pavement piezoelectric sensors were very sensitive to installation depth and the rigid sensors were difficult to install with even very slight rutting in the lane. These problems are not entirely unexpected and have been addressed in an earlier report (5) Often the classifier sensitivities had to be adjusted to obtain proper response for all types and sizes of vehicles.
- Most of the classifiers tested worked without any or with very minor problems when they were first installed. Most, however, required some calibration or adjustment in order

to optiiize their performance. In some case, this adjustment required the devices to be opened to access internal adjustments while in other cases, the adjustments were made using software. This showed that many of the devices are not simply turnkey devices, but instead would require adjustments after installation. This would only be a minor problem if the devices were permanently installed or used for long periods of time, but multiple adjustments may be required.

- One problem that was anticipated prior to the test was the misclassification of pole trucks, which are in Class 9 (tractor trailers), carrying utility poles or logs. The trailers on these units consist of a long metal beam between the trailer axles and the front of the trailer. It was anticipated that the classifiers would sometimes mistake the beam or pole for a gap between vehicles and classified the truck-trailer combination as two separate vehicles. Where a Class 9 vehicle (tractor trailer) was expected, the devices might sometimes classified it as a Class 6 (3-axle truck) and a Class 2 (car) vehicle. This problem turned out to occur only infrequently and did not significantly impact the results of the tests.
- The classification accuracy of the units tested was also assessed as a function of the air and pavement temperature at the test site. Since the winter test was not conducted, the range of temperatures experienced was rather limited. The data did, however, tend to show a slight reduction in classification accuracy as temperature increased.

CONCLUSIONS

This study quantified the ranges of expected accuracies for the state-of-the-application vehicle classification devices (as of September 1992). It also provided information on axle spacing and overall length measurement accuracies for the classifiers. The installation and operation of the devices also provided valuable results. The overall results and conclusions for this project include:

- Piezoelectric axle sensor failures were a significant problem during and immediately after installation. A greater percentage of the flexible piezoelectric sensors failed than did rigid sensors.
- Piezoelectric axle sensors performance was the key factor in the overall classification accuracy of the devices tested. Even slight rutting in the lane caused significant problems in installation and operation of most axle sensors.
- All of the classifiers required some adjustments after installation; some even required disassembly of the cases to adjust internal elements. None were turnkey devices.
- Classification errors were most common between Classes 2 (passenger cars) and 3 (pickups and minivans).
- The most accurately classified vehicle was the tractor trailers (Class 9), probably due to heavy axle loadings and unique axle configuration.
- The measurement of axle spacings was very accurate for most devices.
- Overall vehicle length measurement accuracies *were* less accurate due to the use of inductive loops to make the measurements.
- The augmented pneumatic tube test should be repeated due to a setup error.

For more information on the performance of individual classifiers tested, refer to the project final report published in June 1995 (2).

ACKNOWLEDGMENTS

This project is funded and directed by the Federal Highway Administration Office of Highway Information Management under the title "Accuracy of Traffic Monitoring Equipment." The FI-IWA funded the State of Georgia Department of Transportation to perform the study. The Georgia DOT contracted the Georgia Tech Research Institute to conduct the tests and analyze the results.

REFERENCES

(l) Wyman, J. H., Braley, G. A., and Stevens, R. I., *Field Evaluation of FHWA Vehicle Classifiation Categories,* Technical Paper 84-5, Maine Department of Transportation, Final Report, December, 1984.

(2) Harvey, B. A., Champion, G. H., Ritchie, S. M., and Ruby, C. D., *Accuracy of Traffic Monitoring Equipment,* Final Report, Project # A-929 1, Georgia Tech Research Institute, June 1995.

(3) Harvey, B. A., Champion, G. H., Ritchie, S. M., and Ruby, C. D., *Accuracy of Traffic Monitoring Equipment - Pavement Overlay Tests,* Final Report, Project # A-9684, Georgia Tech Research Institute, July 1995.

(4) Harvey, B. A., Champion, G. H., and Deaver, R., *Accuracy of Traffic Monitoring Equipment,* Proceedings of the National Traffic Data Acquisition Conference 1994 (NATDAC '94), Rocky Hill, Connecticut, September 18-22, 1994, Vol. 2, pp. 111-122.

(5) Majdi, S. and Middleton, D., *Installation of Sites for Automatic Vehicle Classification.* Research Report 1901-1F. Texas Transportation Institute, November 1992.

BEYOND SCHEME F

Speaker: C. James Elliott
Los Alamos National Laboratory
Authors: C. James Elliot, et al.
Los Alamos National Laboratory

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Beyond Scheme F

C. James Elliott, Henry Fisher, and Jason Pepin

Los Alamos National Laboratory

and

Ralph Gillmann, FHWA

Abstract

Traffic classification techniques were evaluated using data from a 1993 investigation of the traffic flow patterns on I-20 in Georgia. First we improved the data by sifting through the data base, checking against the original video for questionable events and removing and/or repairing questionable events. We used this data base to critique the performance quantitatively of a classification method known as Scheme F. As a context for improving the approach, we show in this paper that scheme F can be represented as a McCullogh-Pitts neural network, or as an equivalent decomposition of the plane. We found that Scheme F, among other things, severely misrepresents the number of vehicles in Class 3 by labeling them as Class 2. After discussing the basic classification problem in terms of what is measured, and what is the desired prediction goal, we set forth desirable characteristics of the classification scheme and describe a recurrent neural network system that partitions the high dimensional space up into bins for each axle separation. The collection of bin numbers, one for each of the axle separations, specifies a region in the axle space called a hyper-bin. All the vehicles counted that have the same set of bin numbers are in the same hyper-bin. The probability of the occurrence of a particular class in that hyper-bin is the relative frequency with which that class occurs in that set of bin numbers. This type of algorithm produces classification results that are much more balanced and uniform with respect to Classes 2 and 3 and Class 10. In particular, the cancellation of errors of classification that occurs is for many applications the ideal classification scenario. The neural network results are presented in the form of a primary classification network and a reclassification network, the performance matrices for which are presented. Cancellation of errors is site dependent and the question of generalization to many sites suggests another measurement of goodness for a classification scheme that takes into account the need to operate the network under a variety of operating conditions.

Introduction

An unofficial standard for the classification of axle-spacing-based highway data into the thirteen classes of the Federal Highway Administration was developed by the Maine Highway Department a decade ago and is known as Scheme F. This scheme is not a FHWA standard and presents a highly inaccurate count of Class 2 and Class 3 vehicles in the Georgia study. We show that the assumptions built into this method can be overcome by adaptive means or by using probabilistic neural networks. Furthermore, these more general methods of characterizing partitioned axle parameter space, yield higher precision results.

Regions of parameter space requiring special attention are those regions where two or more of the 13 FHWA classes can not be uniquely separated. The Scheme F approach rigidly split parameter space into fixed compartments that were declared by fiat to be one and only one of the 13 classes. It was recognized that this approach was not strictly correct, but one would hope that the boundaries could be carefully modified so that the errors incurred in one region of parameter space would cancel those in the other regions. For this to be the case, a number of assumptions have to hold. One of these assumptions is that the distribution of passenger cars (Class 2) and pickup-up trucks and vans (Class 3) remains of invariant shape from place to place. Another is that the relative numbers of Class 2 and Class 3 vehicles conforms to a fixed ratio. Common sense opposes these assumptions as does extensive ground truth data. We utilize the 1993 study of traffic on 120 near Covington Georgia to demonstrate that the assumption of cancelling errors under scheme F does not hold, even remotely.

The 120 data suggest new approaches to the axle-based vehicle classification process. One of these is application of a probabilistic neural network that assumes static distributions. This network is tested with two data sets both of which have good cancellation properties. We also discuss adaptive classification procedures, the simplest of which is the moving demarcation line model. This model moves the linear boundary between two classes depending on observations. This method is highly non-linear in that it involves ascertaining the characteristics of the flow from all of the vehicles before deciding how to partition up the ambiguous regions of parameter space. Although the subclass decomposition is often well behaved, we point out instances where the concept of axle spectra has overwhelming experimental difficulties preventing its realization. Practical data-reduction difficulties that occurred in acquiring the Georgia data are also discussed.

Utilizing the Georgia data, we describe a new neural net approach, reclassification of the rejected vehicles by a specially designed network for that purpose. Combined with the adaptive approach described above for the overlap regions, this approach promises greatly enhanced generalizability that tailors itself to particular geographic sites and/or functional categories.

Scheme F results

Scheme F is a method developed by the Maine Department of Transportation [Wyman 1984] for the classification of vehicles. The basic regions for Scheme F are shown in Table 1 for each of the 13 Classes of the FHWA classification. Scheme F has not been approved by the FHWA as a method of classification nor has it been based upon a systematic analysis of an underlying database that supports it as an optimal data classification scheme. We hold that any modem classification scheme must meet the criterion of conforming well to an established database. The results of classification using Scheme F of the Georgia I20 database are presented in Tables 2 and 3. These results closely track the relative performance matrices reported by Harvey for manufactured equipment in his study. That is, it appears that many vendors are basing their classification scheme on something that closely resembles Scheme F. The lack of symmetry of errors can be seen by examination of the vehicles misclassified as Class 2 while they are really Class3. These are compared to the Class 2 vehicles that have been classified as Class 3. The latter are in the hundreds whereas the former are in the thousands. This type of asymmetry is characteristic of Scheme F shown in Tables 1 and 2. Another problem is the misclassification of Class 10 vehicles as Class 13. The definition of Scheme F that we used is shown in Table 3.

Next we point out that Scheme F can be formulated as a McCulloch-Pitts neural network [McCulloch, 1943]; an example of this formulation is shown in Figure 1. Here we look at all the 3 axle vehicles. They can fall into Classes 2, 3, 4, 6 or 8. The McCulloch-Pitts neural network has two input nodes; one for the separation between axles 1 and 2 and the other for axle separations 2 and 3. We give an example and trace how the neural net responds to inputs of 12 and 14 respectively.

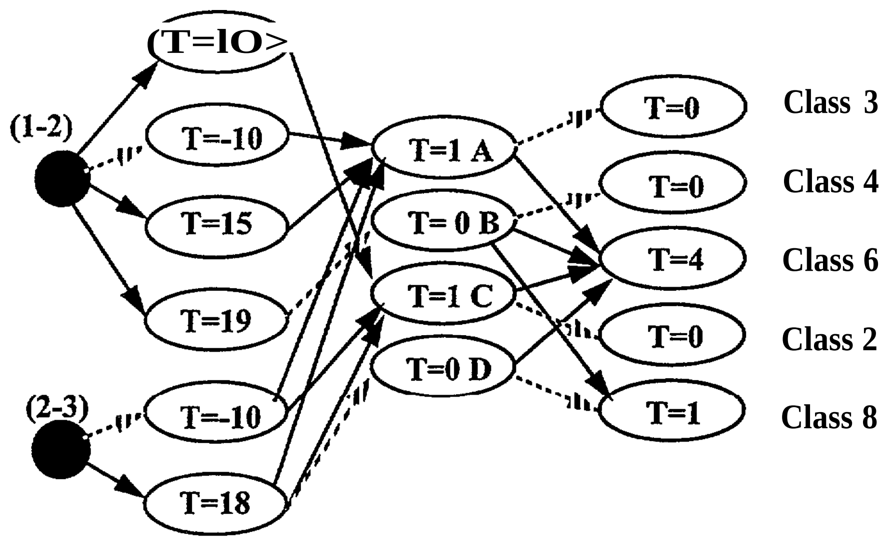


Figure 1: Scheme F for three axles as a McCulloch-Pitts neural net

With 12 for the first axle separation (l-2), only those nodes that are above threshold fire. The dashed line indicates a negative weight of - 1 . Thus 12 becomes -12 going into the t=-1 0 node. Because -12 is below -10, the second node in the first layer does not fire. The first node in the first layer has a threshold of 10 which is below the value of 12, and thus it fires in the above threshold

Table 1: M_{ij} for Scheme F for the May 5-7, 1993 Georgia test

Truth\Scheme F	1	2	3	4	5	6	7	8	9	10	11	12	13
1	18	2				1							
2		8316	243			1		5					
3		3018	2443		4	3		87			5		
4			3	60	6			1					
5		22	361	93	235	1	9	22			16		
6			3	20		461		1	14	1		1	
7													
8			3					319					
9									3671		5		
10										31			12
11								3	1		101		
12												29	
13													1

Table 2: M_{ij} for Scheme F for the September 9-11, 1993 Georgia test

Truth\Scheme F	1	2	3	4	5	6	7	8	9	10	11	12	13
1	30	3											
2		10128	234			2		9					
3		3700	2495		44	3	1	79			3		
4			2	44	5	4							
5		7	145	62	149		6	18		1	7		
6			2	10	1	191		3	5				
7							1	2					
8		2	6			1	1	278					
9									3152	1	11	1	
10										27			9
11											93		
12												28	
13													

Table 3: The Scheme F assumptions

				Axle	Space	In	Feet
Class	FHWA Vehicle Type	No of Axles	Axle 1 to 2	Axle 2 to 3	Axle 3 to 4	Axle 4 to 5	Axle 5 to 6
1	motorcycle	2	0-5.8				
2	car	2	5.8-10				
	car/1 axle trailer	3	0-10	10-18			
	car/2 axle trailer	4	0-10		<3.5		
3	pickup	2	10-15				
	pickup/1 axle trailer	3	10-15	10-18			
	pickup/2 axle trailer	4	10-15		<3.5		
	pickup/3 axle trailer	5	9.9-15			<3.5	
4	bus	2	>20				
	bus	3	>19				
5	single unit truck/dual rear axle	2	15-20				
6	single unit truck	3		<18			
7	single unit truck	4					
8	2 axle tractor with 1 axle trailer	3		>18			
	3 axle tractor with 1 axle trailer	4		<=5	>10		
	2 axle tractor with 2 axle trailer	4		>5	>3.5		
9	3 axle tractor with 2 axle trailer	5					
	2 axle tractor with 3 axle trailer	5		<6.1		3.5-8	
10	any single tractor/trailer comb. with 6 or more axles	6 or more			3.5-5		
11	any tractor/double trailer unit with 5 or less axles	5		>6			
12	tractor/double trailer unit	6					>10
13	any tractor/double trailer unit with 7 or more axles	7 or more					

condition. It is the only node that fires in the first layer that is connected to the first axle separation value. The second axle separation has 14 as the input. For the first of two nodes to which it is connected, the weight is -1 as indicated by the dashed arrow; thus the value going into the t=- 10 node is -14, a value that is below threshold. Thus the first of the two nodes in layer one attached to (2-3) does not fire. Neither does the second node with t= 18 because 14 is below threshold. For these values, the only node that fires in the first layer is the first node. When a node fires it produces a value of unity; otherwise it produces zero.

We now examine our special case for the second layer neurons that are denoted by A,B,C, and D. Of these, B,C, and D fire but A does not. The neuron C fires because it is connected to the firing neuron of layer 1. Neurons B and D fire because they have thresholds of 0 and have 0 inputs. Neuron A does not fire because none of its inputs are active. The interpretation of these nodes may be helpful in understanding the remaining part of the net: Node A fires whenever Class 3 is not active. Node B fires whenever the second axle separation is less than 19. Node C fires whenever Class 2 is not active. Node D fires' whenever the second separation is less than 18.

The third layer is the classification layer. The reader may establish a few facts: Class 2 occurs whenever node C does not fire. Class 3 occurs whenever A does not fire. Class 4 occurs whenever B does not fire. The Class 6 and Class 8 nodes rely on the fact that the neural inputs sum the signals entering the node. Each of the lettered nodes emit either a zero or unity pulse; only if all four inputs leading to Class 6 fire will the threshold of 4 be achieved. Because of the negative weight of - 1 for the D node leading to Class 8, its edge can contribute only zero or - 1 to the sum, whereas the edge from the B node can contribute either zero or unity depending on its firing condition. For this case only, if B fires and D does not, will the Class 8 node occur.

Although the language of how a neural network operates is more difficult for most humans to follow in detail, its description, as given above, is nearly in the form required for a computer to absorb. Humans have a highly developed visual cortex and pictures are much more to their liking than the set of 15 rules computers like so well. In Figure 2, we show the same three axle system as a picture with the shaded areas representing regions where the Scheme F classification occurs. The axle separation (l-2) is plotted as the ordinate and the separation (2-3) is plotted as the abscissa. The scale can be established by a comparison with Table 1. Figure 2 contains the same information as Figure 1, but it is in a much more digestible form.

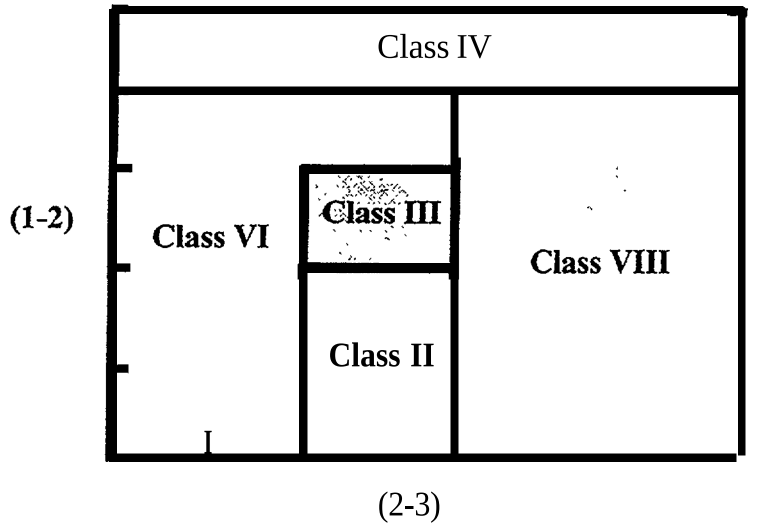


Figure 2: Scheme F for three axle vehicles visually

We can see that a neural network partitions a plane up into black and white regions that correspond to whether a particular node fires or does not fire. Although this is not quite in the form required for a classification scheme, it is very close. The scheme we describe later involves dividing each axle spacing up into **bins,** thereby dividing the plane up into a set of rectangular (hyper-bin) regions. In Scheme F, if properly defined, the hyper-bins fall into one - and only one - of the 13 classes. In the general scheme described later, known as a probabilistic neural network, each of the hyper-bins has a mixture of the classes. For a given data set, a particular hyper-bin could contain, for instance, 10 vehicles in Class 2 and 34 vehicles in Class 3.

The Neural Network Approach

The basic information that we measure on the highway can be cast in terms of the hyper-bin in which the axle separations of the vehicle falls. If two vehicles fall into the same hyper-bin of bins with bin numbers (4,3,5,0,0,0) their axle separations are considered to be equal within the context of our model. We use the convention of bin 0 for the i'th axle separation of a vehicle when that vehicle has too few axles to generate that separation. Thus, the example vehicle described is a four axle vehicle with three axle separations. All vehicles having these bins for its axles will be compared against each other. If a hyper-bin of bins has only one class of vehicle, it can be assigned to that class uniquely. The most general situation is for a hyper-bin of bins to have more than one type of vehicle. Accumulation of the number of each type of vehicle that falls into a given hyper-bin of bins occurs during the training phase of the neural network. These numbers are then converted into a probability (more correctly a likelihood) of finding the vehicle in that hyper-bin of bins. In the testing phase, a vehicle that falls into a given hyper-bin of bins is equally considered to be any of the vehicles that have fallen into that hyper-bin during the training phase. Thus, a probability for each class of the 13 classes of the FHWA is assigned to the test vehicle according

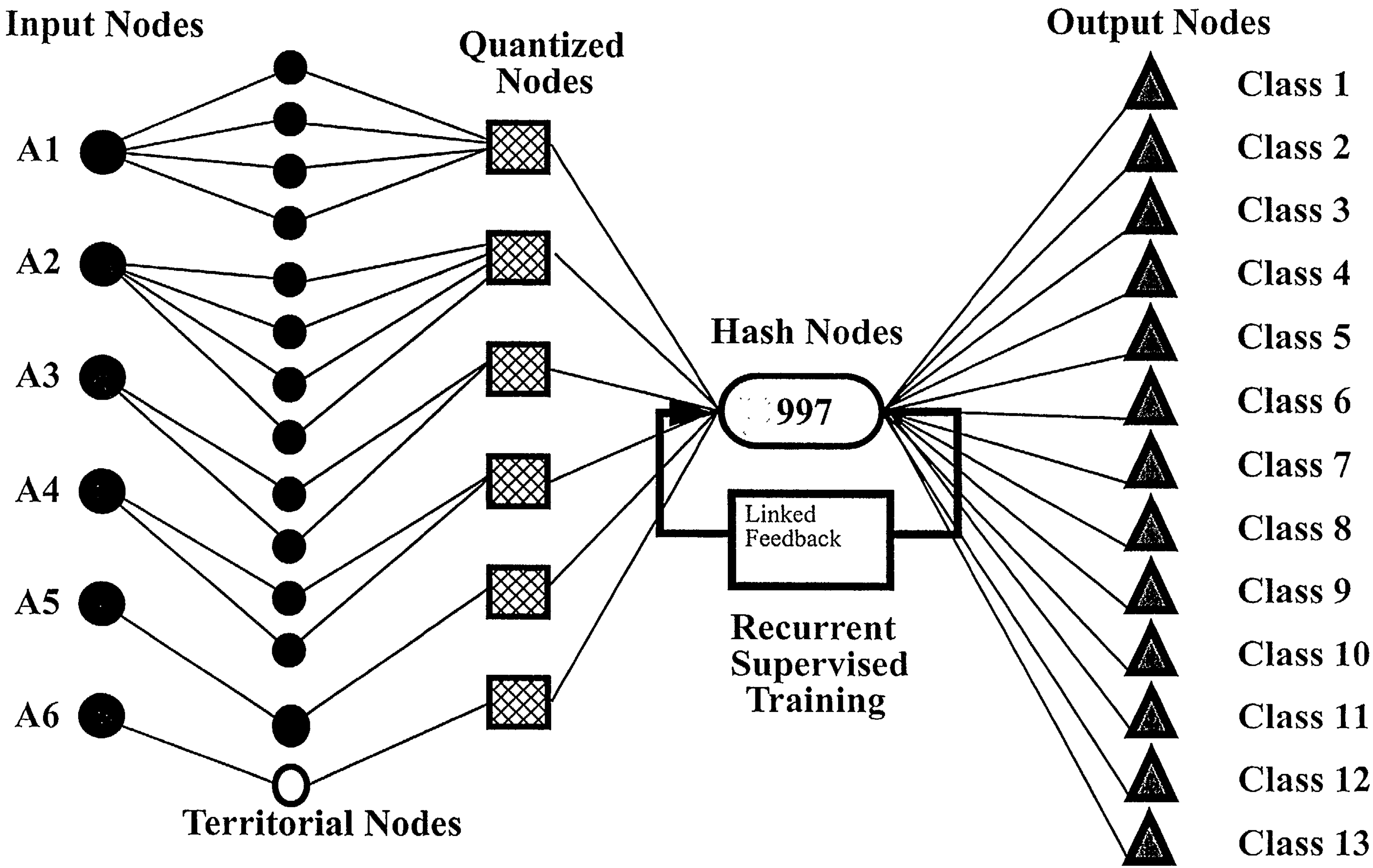


Figure 3: The Probabilistic Neural Network Architecture

to these numbers. In the course of determining the performance matrix during the testing phase, the performance matrix element is incremented according to the probability.

The neural network has six axle separations as the inputs. Each axle separation has a range of continuous values whose lower range is greater than zero. The value of zero is used to indicate that the vehicle is missing such a separation. The continuous range of separations is broken up into a user determined number of bins. A technical problem in achieving such a system is the potentially large number of bins that can occur. For the example below, the number of bin combinations is 3.2 million. The computer science technique of a hash table [Sedgewick, 94] is used to condense the number of possibilities to only those hyper-bins that actually have vehicles. This is approximately a thousand such hyper-bins, a savings of about a factor of 3 000 in the number of storage elements that could be accessed as the storage bins. The net classification time is less than the input and output time - less than 6 seconds total on a SUN SPARC 10.

Table 4: Axle Spacings in 30.48 cm units

axle spacing	min space	max space	number of bins
1	4.0	30.5	40
2	2.7	63.0	40
3	1.8	52	20
4	2.2	43	20
5	2.2	31	10
6	3.5	18	1

Table 5: Axle Spacings for reject classification

axle spacing	min space	max space	number of bins
1	4.0	30.5	10
2	2.7	63.0	1
3	1.8	52	1
4	2.2	43	1
5	2.2	31	5
6	3.5	18	3

In this technique, hyper-bins of parameter space are occupied during training and then accessed during testing. The data is divided into three parts by random choice: org.tm, org.tes, and org.tst, the training and two testing sets. Of the approximately 13,5 10 vehicles in file org.tes and 13,649

Table 6: org.tes data of georgia 1993 study probabilistic neural net 40 40 20 20 10 1 bins

truth\guess	1	2	3	4	5	6	7	8	9	10	11	12	rej
1	17	0.5	0.5										0
2		4703	1548		7								3
3	0.7	1579	2214	3	113	1		6					55
4				23	16	6							3
5		7	127	11	203			3					34
6				2		208			2				12
7									1				1
8			7		2	1		143					45
9						3			2146				146
10										4			18
11											49		11
12												12	4
13													0

Table 7: org.tst data of Georgia 1993 study

truth\guess	1	2	3	4	5	6	7	8	9	10	11	12	rej
1	18	1	1										1
2		4794	1611		7								4
3	2	1584	2153	3	116	1		7					42
4			2	13	14	3							2
5		7	118	13	206			2					31
6				2		221			5				18
7													1
8			10		3			134					45
9						2			2164				165
10										1			26
11											54		14
12												14	6
13													0

Table 8: Truth Guess Reclassification with network applied to org.tst with spacings from Table 5. If all the vehicles listed as rejected in Table 7 are reclassified with this net, less than one vehicle(0.4) is rejected. The 1 in truth item 10 under rej column is the count in Class 13.

t\g	1	2	3	4	5	6	7	8	9	10	11	12	rej
1	3.4	11.6	3.0										0
2	0.6	4458	1778	0.81	18.7	1.5	0.	3.16			.03		0
3	0.6	2040	1762	1.9	81	17.4	.4	67.7	.9		.7		0
4			0.6	23	17.5	7.8		.07					0
5		30.8	130	13.6	175	4.1	0.	23.4	8.6		.56		0
6		0.8	18.7	5	4.9	175		14	5.8		.13		0
7			1										0
8		3.3	62	.65	16.7	14.1	.5	100			1.1		0
9			3		4.7	6			2212		69.4		0
10					.5					20.1			(.4) 1
11			.5		.2	.1		.5	55.6		3.1		0
12						1.3						14.6	0
13													0

vehicles in the org.tst, 282 and 355 vehicles, respectively, are rejected as not being in any of the bins established during training. A methodology that we plan to adopt in the near future is to examine each of the rejected vehicles and determine how close they are to the established data set. Then to incorporate them if appropriate. The reclassification neural network classification is shown in Table 8. Of the combined nets, only one vehicle is unclassified. The rejected Class 10 and Class 12 vehicles are shown in Table 10. These numbers are consistent with Tables 5 and 8.

The classification results are shown in Tables 6 and 7. These are seen to provide a much cleaner division in Class 2 and 3 and Class 3 and 5 than provided by Scheme F. Here the errors of classification of Class 2 as Class 3 and Class 3 as Class 2 balance out, giving the correct overall numbers in each class. Likewise, the errors in identifying Class 3 as Class 5 and Class 5 as Class 3 cancel out in a similar fashion.

TABLE IO. Hyper-Bins of Rejected vehicles in Class 10 and Class 12

as1	*as2*	*as3*	*as4*	*as5*	*as6*
class 10					
24	2	5	2	1	0
9	1	1	6	1	1
13	2	12	2	1	0
12	2	3	1	1	0
9	2	9	2	1	0
13	2	9	2	1	0
19	2	13	2	2	1
9	1	1	6	1	1
13	13	4	7	1	0
9	25	1	1	1	0
9	26	1	1	1	0
9	25	1	1	1	0
16	2	2	20	1	0
15	2	3	1	10	1
class 12					
11	2	6	7	5	0
21	2	8	4	8	0

3. Utility Theory

When there are many applications for a given set of data, utility theory can help sort out the important aspects of the data. The details of a particular application govern the utility function. One can imagine many scenarios of usage of traffic data: the most common involves estimation of a parameter at a particular site that depends linearly on the number of vehicles in each class, or a derived quantity such as the average speed times the number of vehicles, or perhaps, the equivalent single axle loadings of all vehicles. An enforcement utility function, on the other hand, might involve specification of the characteristics of a single stolen vehicle. The best classification scheme depends on the characteristics of the associated utility function. For estimation the following utility function is appropriate for repeated samples with the same distribution at a given site:

$$Q = -\sum_{ij} W_i (M_{ij} - M_{ji})^2$$

where W_i indicates the importance of a particular class and M_{ij} is the performance matrix element. This indicates that cancellation of errors is useful,; thus, if Class 2 vehicles are misidentified as Class 3 in the same number as Class 3 vehicles are misidentified as Class 2, the overall counts will be preserved and to the extent that cancellation happens, the prediction with cancellation will be as good as that without.

The problem is that most classification schemes are not geared for one site. Even if they were, the variations in traffic flow at a single site can change with weather, degree of urbanization, technological factors, and the mission of society. In the language of neural networks, this phenomena is called the problem of generalization. The network is trained under one set of circumstances and is used for another; how does its performance change? In the case that the performance matrix showed perfect performance, this would be expected to hold at all sites. When cancellation of errors occurs, the cancellation effect depends on the volume of traffic in the two cancelling classes. This is shown in the simple model moving demarcation line model, shown if Figure 4.

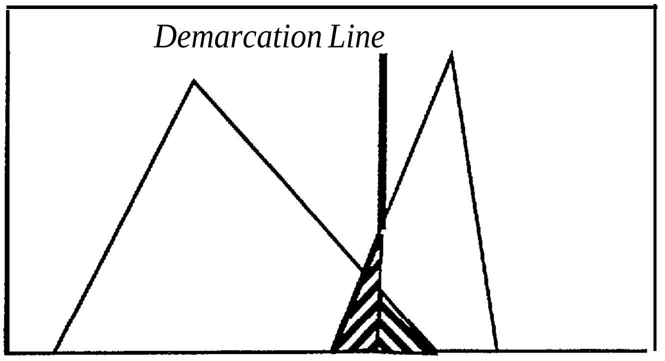


Figure 4: Moving demarcation line for intermediate case

The two distributions of vehicles are assumed to maintain the same shape, but have variable amplitudes. If there are almost no vehicles in the second class, the height of the second triangle will be very small. There will be almost no overlap area under it. This means that the demarcation line moves to the right until the area under the first triangle, to the right of the line, is almost zero. Contrariwise, if there are almost no vehicles in the first class, but lots of vehicles in the second class, the first triangle will have nearly zero height while the second triangle remains tall. Then the demarcation line moves to the left, until the area to the left of the demarcation line of the second triangle is nearly zero. An intermediate case is shown in Figure 4. Other models that vary with vehicle concentrations also show these effects and could be more important than the moving demarcation line model in practice. One such model would be to utilize a least square fit to the observed profile with variable coefficients multiplying two basis distributions. The conclusion is the same: namely a model's generalization characteristics need to be examined in detail.

The generalization characteristics drives the importance of the diagonal matrix elements. The two effects of cancellation of errors and large diagonal matrix elements in the performance matrix suggest the following form of utility function to decide on the best classification scheme with data derived from one site to be applied at other sites for the purpose of parameter estimation.

$$Q = \sum_i W_i \left[M_{ii}^2 - c \sum_i (M_{ij} - M_{ji})^2 \right]$$

Here W_i indicates the relative importance of class i. The parameter c is chosen large if the same distributions of vehicles occur at all sites and small if the variations at sites are totally unpredictable.

Conclusion

The undesirable characteristics of Scheme F have been described qualitatively and more generally with utility theory. A recurrent probabilistic neural network with a reclassification network for the rejected vehicles has been constructed and is a major remedy for the difficulties of Scheme F. The framework of utility theory can be the basis for determination of a best algorithm that applies to a variety of sites. With this and a study of the underlying phenomena at various sites, the details of best algorithm implementation can be determined.

References:

John H. Wyman, Gary A. Braley, Robert I. Stevens, Field Evaluation of FHWA Vehicle Classification Categories - MDOT, 1984, Maine Department of Transportation, Bangor, Maine, 04401.

Bruce A. Harvey, Glenn H. Champion, Steven M. Richie, Craig D. Ruby, Accuracy of Traffic Monitoring Equipment, GDOT 92 10, Technical Report GTRI Project A-9291, June 1995.

W. S. McCulloch and W. A. Pitts, A Logical Calculus of the Ideas Immanent in Nervous Activity, Bull. Math. Biophysics 5 115-133, 1943.

Robert Sedgewick, Algorithms in C++, Addison-Wesley Publishing Company, Reading Massachusetts, 1992.

Acknowledgments:

The authors acknowledge the support of the FHWA, the helpful assistance of the Alliance for Transportation Research, and the assistance of Larry Blair at Los Alamos National Laboratory in securing funding under the auspices of the Department of Energy. The helpfulness of the management and support personnel of groups XCM and EES-5 at Los Alamos National Laboratory facilitated the accomplishment of this work. The technical content was influenced by helpful discussions with Drs. Rajendra B. Patil and Azim Eskandarian.

TRAFFIC ANALYSIS EXPERT SYSTEM: CONSIDERATIONS FOR IMPLEMENTATION

Mark Flinner
Minnesota Department of Transportation

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

For further information contact:
Mark Flinner
Minnesota Department of Transportation
MS 450 MnDOT
St. Paul, MN 55155
Ph: (612) 297-1466

TRAFFIC ANALYSIS EXPERT SYSTEM:

CONSIDERATIONS FOR IMPLEMENTATION

MARK FLINNER

MINNESOTA DEPARTMENT OF TRANSPORTATION

Using the state of Minnesota's experience during 1995 as an example, this presentation will share tips and advice on the use of the recently developed expert system for Automatic Traffic Recorder (ATR) data screening and editing. Beginning first as a proof of concept project for short duration as well as continuous traffic data, the Traffic Analysis Expert System (TAES) evolved into a production oriented system that was jointly developed by Colorado, Kansas, Minnesota and the FHWA. Last year's NATDAC featured a presentation on TAES which outlined the theoretical concepts and practical features of the system.

Following initial testing of the continuous traffic data component of TAES Minnesota began to use TAES exclusively in its efforts to monitor the performance of its 140 continuous traffic monitoring sites. Implementation considerations that will be addressed briefly include the following:

- Assessing the effects of missing ATR data on a traffic monitoring and estimating program and the role of the "Truth In Data" principle,
- Polling and raw data integrity checks that must still take place outside of the TAES system,
- Specially developed programs for using past years' data files to create a historically-based ATR specific target value and tolerance files essential to TAES,
- What to do if one is adding a new site where no historical data is available,
- Minnesota's experience in "fine tuning" the rule base sensitivity in order to balance the benefits of automated traffic data screening with the costs of manual, hourly data inspection,
- Reporting against the fully edited TAES traffic record structure (expanded Card 3),
- TAES work process limitations and benefits, and
- Proposed changes and enhancements to TAES.

“TRAFFIC EDIT PROCEDURES” POOLED FUND - PROJECT DIRECTION AND STATUS REPORT

Mark Flinner
Minnesota Department of Transportation

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

For further information contact:
Mark Flinner
Minnesota Department of Transportation
MS 450 MnDOT
St. Paul, MN 55155
Ph: (612) 297-l 466

"TRAFFIC EDIT PROCEDURES" POOLED FUND - PROJECT DIRECTION AND STATUS REPORT

MARK FLINNER

MINNESOTA DEPARTMENT OF TRANSPORTATION

This presentation will describe the pooled fund activities as planned by sixteen states and the FHWA Participating states, ranging from California to New York, Montana to Florida, responded to the 1994 solicitation to contribute to the development of computerized traffic data editing routines that either supplement existing routines or replace significant manual efforts. Recognizing very early that there is much to be learned and shared, the pooled fund advisory committee intends to build upon existing levels of understanding and expertise through:

- facilitated exchanges of and challenges to existing wisdom,
- analysis and synthesis of existing and proposed traffic data screening and editing mechanisms,
- joint development of platform and operating system 'independent' data screening and editing routines, and potentially,
- enhancement 'of existing FHWA or AASHTO software so that the benefits of this effort can be shared in all states and agencies as they may desire.

The year 1996 marks the beginning of significant pooled fund activities. Consequently, conference attendees will be encouraged to comment and ask questions regarding specific actions and plans encompassed in this project. States that are considering greater involvement in this effort will be asked to contact the 'lead state' (Minnesota) with their intent.

CONCURRENT SESSION 2B - TRAFFIC MONITORING SYSTEMS PANEL

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

TRAFFIC MONITORING SYSTEMS PANEL SUMMARY

Tony Esteve
Federal Highway Administration

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

TRAFFIC MONITORING SYSTEMS PANEL

Moderator: Mr. Tony Esteve, Federal Highway Administration, HPM-3 0

Members: Mr. Rick Reel, Florida Department of Transportation
Mr. Bill Hughes, Kansas Department of Transportation
Mr. Curt Dahlin, Minnesota Department of Transportation
Allan Heckman, Missouri Highway and Transportation Department

Purpose: Discuss the status of traffic monitoring activities at the Federal and State levels.

Presentations:

Federal Requirements for Traffic Monitoring Systems
Tony Esteve briefly discussed the history of traffic monitoring requirements introduced by the Inter-modal Surface Transportation Efficiency Act of 1991 and the resulting Federal Highway Administration (FHWA) and Federal Transit Administration regulations. As required, each State provided a plan which was reviewed and approved by the FHWA. Current regulations were superseded by the National Highway System Designation of 1995, which relaxed the requirements for management systems. The traffic monitoring systems remain a requirement with wide flexibility provided to meet State differences. The traffic monitoring requirements are being reevaluated to allow State/MP0 to meet traffic data needs in a cooperative manner free of penalties or federal reporting requirements.

Florida's Traffic Monitoring Program
Rick Reel summarized the current traffic monitoring system in Florida. The State has about 113,000 miles of roads and about 11,000 miles within the State system. The State has a large permanent traffic volume, vehicle classification, and WIM program consisting of about 250 permanent counters, many capable of classification, and about 17 permanent WIM sites. The traflic volume coverage program is designed to completely cover the State road system annually. Rick briefly discussed the many real-life difficulties surmounted during the short-term installation of several permanent counters in the Florida Turnpike and Interstate 95. Installation of permanent sites remains a costly and labor intensive task. The State of Florida has made large investments in traffic data collection and processing and-continues to work with its District Offices and urbanized areas to meet customer traffic data needs.

Kansas' Traffic Monitoring Program
Bill Hughes discussed the implementation of the Kansas program of traffic volume, vehicle classification, and truck WIM data following the guidelines of the Traffic Monitoring Guide. The success of the program made the conversion to the TMS regulations very simple. The State program-meets current needs for traffic volume, vehicle classification,

and WIM data. Due to new requirements, the location of permanent sites is being reevaluated. An outreach program is underway to coordinate the traffic data programs of cities and counties to increase data interchange. The ORACLE database system is being used to implement a coordinated data processing operation to manage and utilize all the available data from all sources. The data processing will also be interfaced with a geographic information system.

Minnesota's Traffic Monitoring Program

Curt Dahlin discussed the Minnesota program, which is designed to provide comprehensive road system coverage, account for the variability caused by volume levels, and reduce count duplication.The expertise and skill of traffic data collectors is critical to reduce the need for recounts. The State is sponsoring a pooled-fimd study to develop data editing procedures and software to increase the efficiency of traffic data processing and insure the quality of traffic information. Curt highlighted the need for adequate resources to meet traffic data needs, and the importance of better marketing of traffic data and services to high level managers as a means of competing for scarce resources. He suggested the need to explore institutional and organizational support from AASHTO and other traffic engineering or highway organizations as a means to increase the visibility of traffic data products.

Missouri's Traffic Monitoring Program

Allan Heckman discussed the implementation of Missouri's traffic data program following the guidelines of the Traffic Monitoring Guide. The program consists of about 108 telemetry permanent volume sites, 5,000 annual coverage counts, 4,000 annual special volume counts, 300 portable classification counts over a 3-year cycle, and WIM data collected at both portable and permanent sites. All short counts are of a 48-hour duration. Missouri plans to continue its development of management systems and to fully integrate the management and monitoring systems. The traffic monitoring system will be expected to meet the traffic data needs of the management systems.

Open Forum

Each of the speakers entertained questions and comments from the audience on issues related to traffic monitoring. Participants asked questions and raised several interesting issues. A lively discussion ensued, and a variety of opinions were expressed.

TRAFFIC MONITORING SYSTEMS PANEL SUMMARY

Rick Reel
Florida Department of Transportation

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9,1996

Florida's Traffic Monitoring System

Florida public highway system include 112,804 miles of roadway. Of these, 89,099 miles are local roads that are excluded from the state traffic monitoring system. The remainder comprises 23,705 miles of federal aid eligible public roads, of which 11,881 miles are federal aid eligible local jurisdiction roads and 11,824 miles are federal aid eligible state highways. Of the state highways, 4,188 miles are on the National Highway System (NHS).

FDOT has designed, implemented, and operated a traffic monitoring system for the 11,824 miles of federal aid eligible state highways. Application of the Federal Highway Administration's (FHWA's) Traffic Monitoring Guide to Florida's highway system indicated that 32 automatic traffic recorder (ATR) sites were needed to meet minimum federal reporting requirements. FDOT currently has a network of 250 ATR's.

The FHWA's Highway Performance Monitoring System (HPMS) mandates a coverage count program of 1,981 sample sections. The FDOT program includes these locations among a total of 5,891 sites. Vehicle classification is required under the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 for 217 NHS to NHS connectors. FDOT has adopted a procedure that at least one vehicle classification survey will be taken on each highway section, of which there are 1,061. These vehicle classification data collection sessions are needed to derive truck percentages and axle adjustment factors for all roads on the state highway system. During the past year a total of 1,593 different locations were counted using vehicle classifiers.

The past decade has seen major improvements to the statewide FDOT traffic monitoring system. In 1991, FDOT began a major expansion of its ATR system to meet demand for accurate countywide seasonal adjustment, growth, K (30th highest hour), and directional distribution factors. It is FDOT's goal to install one ATR for each highway functional classification group for which there is a minimum of ten miles in each county. This approach will require a total of 275 ATR's. When the ATR expansion plan was approved, there were a total of 86 ATR's statewide. Twenty-five of the state's sixty-&en counties had no ATR's within their boundaries. Approximately half of the new ATR sites were included as part of resurfacing/ reconstruction projects. The remainder were installed under two ATR construction contracts.

Within the coverage count program, the FDOT Districts have gone mostly to portable vehicle classification surveys rather than simple traffic volume counts. The portable vehicle classifiers provide more useful data than simple volume counting such as truck volumes, peak hour volumes and factors, directional distribution factors, etc. Many of the coverage count locations have been equipped with permanently installed sensors. While this feature was provided principally for the safety of FDOT traffic data collection personnel and the traveling public, it has also contributed to the accuracy and reliability of the data collection process.

Analysis of the 1991 weigh-in-motion (WIM) data from 13 continuous and 21 seasonal Strategic Highway Research Program (SHRP) Long Term Pavement Performance (LTPP) monitoring sites suggested that there may be some geographical differences in the 18 Kip equivalent single axle load (ESAL) factors. In an effort to substantiate this hypothesis, a pilot program was begun in District 3 to install eight new WIM sites on different functional classification/level of service highways. A contract was issued in March 1996 to permanently install single wheelpath bending plates in the roadway at four sites. These should be operational by March 1997. The remaining four sites will be constructed as soon as road resurfacing permits. The eight WIM sites in the study will be operated as ATR vehicle classifiers year round. They will be programmed to collect truck weight data for one week each quarter. This approach will provide the WIM data needed for analysis without the high costs associated with transmitting daily WIM files back to the polling computer and the effort required to process large quantities of data.

The importance of the traffic monitoring program was brought into focus last fall after several accidents occurred involving cars and trucks that closed I-95 for six or more hours at a time. In an effort to entice some of the large trucks away from I-95 and onto the nearby parallel route on Florida's Turnpike, the Secretary of FDOT decided to reduce the tolls for trucks with five or more axles on Florida's Turnpike to their 1989 levels. This rollback is now in effect and is costing the Turnpike $100,000 per month in lost revenues. This has, in turn, had a negative impact on the Turnpike's improvement program and bond rating.

On September 20, the Statistics Office was assigned the task of installing all the ATR's required to determine if the toll reduction strategy was having the desired effect. The cost of this project was secondary to the need to have the entire data collection infrastructure operational by October 1,1995.

ATR sites were installed on both I-95 and the Turnpike. The Turnpike wanted one ATR between each interchange between Ft. Pierce and the Golden Glades Interchange in Miami. This section contains 21 highway segments of which one was instrumented with a WIM site using piezoelectric axle weight sensors. I-95 has five ATR's within this area of which only two were providing acceptable vehicle classification data when the project was initiated.

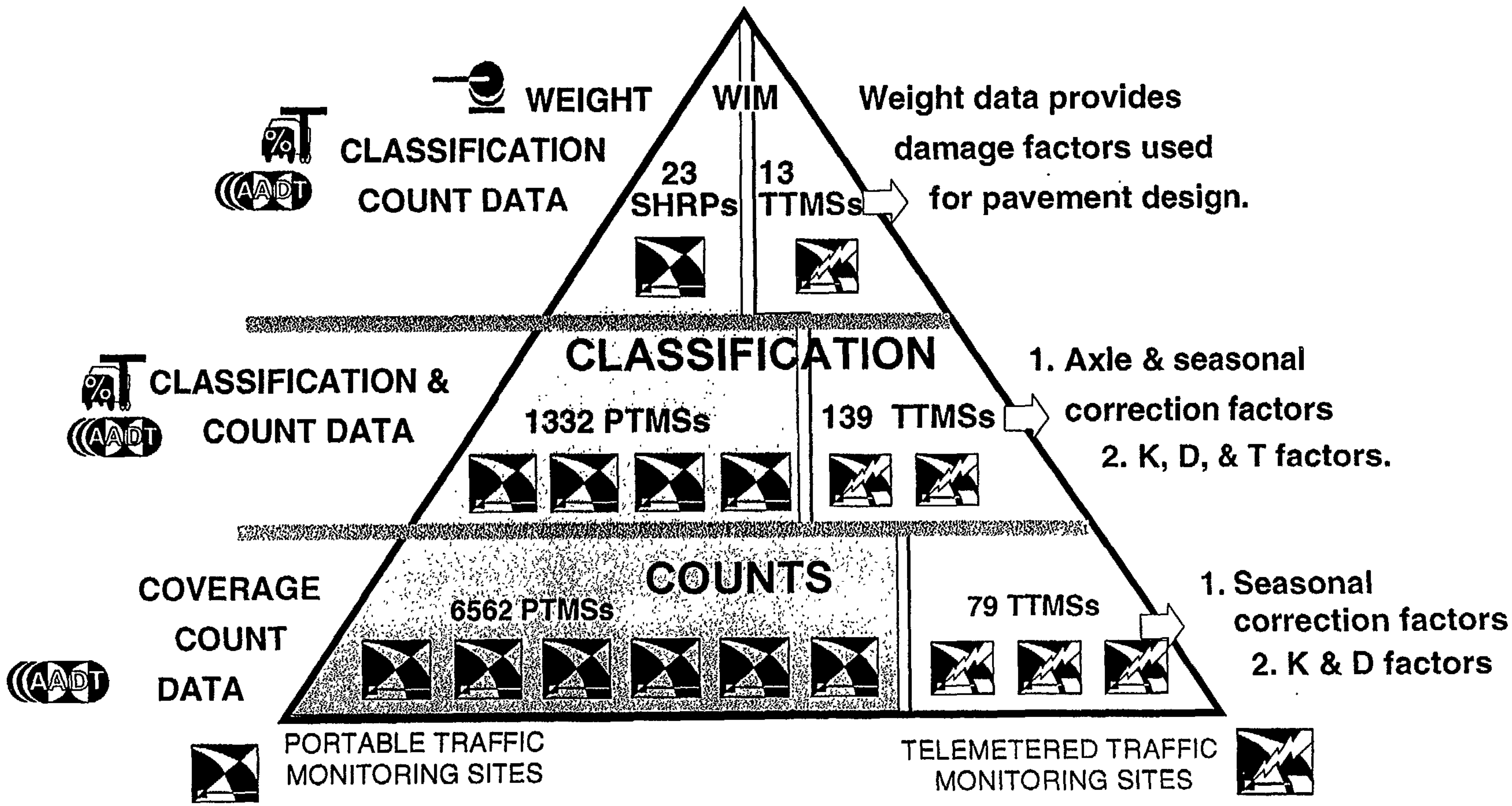
Two inspection crews spent two days surveying the highways to provide more accurate cost estimates. When the cost estimates were repotted to FDOT management, cost then became a consideration. The project was scaled back to seven screenlines, utilizing all existing ATR's. This **required** that six new ATR's be constructed on the Turnpike, two new ATR's be built on I-95, and four existing ATR's on I-95 be reconstructed.

The extremely short time line associated with this project resulted in several problems that were overcome. First, it required three to four weeks to identify a legal and rational method for hiring a construction contractor in a manner that would meet the need of the truck diversion study. Second, it was required to coordinate among two FDOT Districts, three FDOT maintenance yards, and six permits offices. The construction effort on the Turnpike was complicated by the presence of an AT&T fiber optic trunk line for south Florida in the median under a concrete barrier wall. Additionally, there is at least one fiber optic telephone line for local service on each shoulder. There are also two ten inch Florida Gas Transmission Company high pressure gas pipelines on the eastern right-of-way line.

Another complicating issue was the limitation of construction on I-95 to night hours and then only after the FDOT District Public Information Office had given the local news media a week's advance notice. The shortened construction schedule also caused difficulties in the procurement of roadside cabinets, concrete utility poles, loop wire, and sealant. Weather conditions were not optimal during the construction. There was ten inches of rainfall within six hours during the second day that the contractor had his entire crew on the job. Obtaining phone service required a significant effort.

Despite the above difficulties, the project resulted in the installation of 74 lanes of inductive loops and piezoelectric axle sensors, thirteen cabinets, eighteen ATR's, eleven new phone lines, and two cellular telephone units by November 1, 1995. Data from this infrastructure is now being used to evaluate the diversion effects of the truck toll schedule.

As indicated in the attached figure, FDOT has structured its traffic monitoring program to collect the types, quantity, and distribution of data required to support the Department's work efforts. Federal reporting of traffic data is an offshoot of FDOT's program, and willingly provided to FHWA.


WEIGHT
CLASSIFICATION
COUNT DATA
WIM
23
SHRPs
13
TTMSs
Weight data provides
damage factors used
for pavement design.
CLASSIFICATION &
COUNT DATA
CLASSIFICATION
1332 PTMSs
139 TTMSs
1. Axle & seasonal
correction factors
2. K, D, & T factors.
COVERAGE
COUNT
DATA
COUNTS
6562 PTMSs
79 TTMSs
1. Seasonal
correction factors
2. K & D factors
PORTABLE TRAFFIC
MONITORING SITES
TELEMETERED TRAFFIC
MONITORING SITES
AADT
%

CONCURRENT SESSION 3B -VEHICLE CLASSIFICATION DATA

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

USES OF VEHICLE CLASSIFICATION DATA

Speaker: Nikiforos Stamatiadis
University of Kentucky
Authors: Nikiforos Stamatiadis, et al.
University of Kentucky

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

INTRODUCTION

The increased use of equipment with an Automatic Vehicle Classification (AVC) capability and Weigh In Motion (WIM) devices produces a large amount of new data that can provide some insights on understanding traffic patterns more efficiently. Accurate estimates of annual average daily traffic (AADT) as well as vehicle-miles of travel (VMT) by vehicle type are required for a variety of traffic studies. Such data can be used in estimating truck traffic for highway cost allocation studies, in predicting traffic volumes for roadways, and in estimating accident rates for various vehicle categories.

The use of traffic counts and vehicle classification data to develop AADT and VMT measures is a common practice throughout the USA. In the past, limitations in manpower and equipment have prevented vehicle classification data from being obtained in all seasons of the year. Even though currently improved technologies allow for collection of continuous vehicle classification data for longer periods, data are not collected for all seasons due to budgetary constraints and equipment shortage. Therefore, it is important to understand the relationship between the data collected in one season or month of the year to the entire year. This relationship, if known, could be used to "expand" or extrapolate short-term traffic data to describe traffic characteristics for the entire year. Therefore, seasonal, weekly, and hourly adjustment factors become important. Traffic information developed from these adjustment factors can be used to develop accurate estimates of traffic for any given period and can be used for a variety of purposes and studies.

Historically, the Kentucky Transportation Cabinet (KYTC) has applied seasonal adjustment factors to factor routine short-term volume counts. However, a procedure to apply seasonal as well as weekly and hourly adjustment factors to short-term vehicle classification counts has not yet been developed. Previous research has shown that there are indeed seasonal and weekly variations among the various vehicle classes (1,2, 3). Therefore, a research effort was initiated to develop such adjustment factors to produce more accurate estimates of AADT by vehicle type. The first step in this study was to conduct a survey of current uses of vehicle classification data and identify methods used for data collection throughout the United States. At the same time, continuous vehicle classification data collected in Kentucky over a two year period was analyzed to develop seasonal, weekly, and hourly factors.

RESEARCH FINDINGS

Survey Results

The objective of the survey was to determine the current use of vehicle classification data by each state and to identify potential methodologies on developing and using adjustment factors for predicting AADT by vehicle type. The survey included questions concerning the collected classification data, methods for developing adjustment factors, use of any adjustment factors, current equipment usage, and other related topics. Thirty-two of the fifty states surveyed returned the questionnaire. With a response rate of 64%, a clear scope of the current vehicle classification practices across the country can be obtained.

Since vehicle classification data can be used for a variety of reasons and purposes, it was deemed essential to determine these uses from the outset of the survey. Thus, the first question identifies the uses of vehicle classification. The predominant use of these data is for operational purposes, which include roadway design, pavement maintenance, and development of AADT estimates, as it was indicated by fifteen states (47 percent). Current ISTEA and HPMS requirements as well as ESAL/EAL estimates were the primary reasons for collecting these data by another thirteen states (41 percent). Three states, 9 percent of the those surveyed, use the classification data for user fee and cost allocation studies while another three states (9 percent) consider these data important for SHRP test sections. Another five states (16 percent) report that these data are currently not used for any specific purpose. Please note that the total percentages do not add to 100 percent because a number of states indicated multi-use of these data.

The next question asked aims to determine whether seasonal adjustment factors are used and how they are applied. The practice of using seasonal adjustment factors to correct short-term counts is not a widely spread practice among the states. Only nine states (28 percent) currently use some adjustment factors to address seasonal variability in traffic counts. Four of those states develop adjustment factors and use them to adjust. short-term counts to obtain AADT while another state developed factors to be used only for truck AADT estimates. The remaining states did not specify the use of these factors and thus, no further comments can be made. Moreover, there is a great variability among these nine states on the development of these adjustment factors. The use of automatic traffic counters (ATR's), WIM., and AVC devices are prominent methods for developing seasonal factors. In addition to these devices, use of toll receipts to adjust truck data only, use of monthly average daily traffic as seasonal indicators, and sampling in various seasons are other approaches identified in the survey. An additional question, addressed to those states which do not currently use any seasonal factors, aimed to determine whether any such factors will be developed in the future. Among the 23 states that do not currently use any seasonal adjustment factors only 10 (3 1 percent of all) are considering any future development and use of such adjustment factors.

The AVC data collectors classify vehicles into a vehicle type using the length and weight of the vehicle. However, a number of vehicles cannot be properly identified as an exiting type for a variety of reasons and thus, categorized as unclassified. Obviously a large percentage of unclassified vehicles is not desirable, since it will produce unreliable estimates of traffic. The appropriate treatment of these unclassified vehicles--elimination or reallocation-is an additional issue that requires attention . Moreover, identifyng the type of vehicle for these unclassified vehicles is usually a difficult task and, most likely, will affect the reliability of the traffic estimates. The absence of any standards for acceptable range as well as treatment of unidentified vehicles lead to an investigation of existing practices among the states. Only one state, Wyoming, reported a zero tolerance for unidentified vehicles when using Scheme F. Five states indicated that their acceptable level is less than 10 percent (three in the 1 to 5 percent range and two in the 6 to 9 percent range). The majority of the surveyed states--twenty states, 71 percent--indicated 10 percent as the acceptable level for unclassiied vehicles. Finally, only two states have acceptable ranges higher than 10 percent for unclassified vehicles.

The treatment of the unclassified vehicles is also of concern. Since unclassified vehicles do not have a classification category of their own, these vehicles may be either ignored or allocated to another vehicle category. The surveyed states provided five diierent treatments. The majority of the surveyed agencies (19 states, 63 percent) completely ignore any vehicles which are not classified properly during the classification counts. The remaining states use some reallocation of the unclassified vehicles. Two states allocate them to Class 15 (heavy vehicles) and three other states allocate these vehicles in the passenger vehicle category. Four more states indicate that they allocate the unclassified vehicles but no specific category was identified. Finally, two states, North Carolina and Minnesota, either allocate or ignore the vehicles depending on the surrounding circumstances. In Minnesota, the vehicles are either allocated or ignored depending on the amount of information known about the vehicles. If the vehicle length or axle spacing is available, the vehicle class is estimated by length. Ifneither the vehicle length nor spacing is known, the vehicle is ignored. In North Carolina, treatment of the vehicles is dependent on the intent of the data. Also, the unidentified vehicles are ignored or allocated based on volume distribution percentages.

An aspect that is of interest is the relationship between the acceptable percentage of unclassified vehicles and the treatment' of these vehicles. This relationship is shown n Table 1. The examination of the data shown in Table 1 does not reveal any relationship between the variables of concern. For states with a tolerance level less than 10 percent, the number of states that elect to ignore these vehicles is equal to the number of those which reallocate them. The same is true for those states with the 10 percent level of acceptance. Moreover, there is no specific pattern among those states that reallocate the unclassified vehicles, since almost equal numbers allocate them in all possible categories. Finally, the two states with the above 10 percent tolerance indicate that they ignore all unclassified vehicles. Based on these observations, one may conclude that the 10 percent seems to be the common threshold for accepting unclassified vehicles. However, the treatment of these vehicles is a decision that an agency needs to make based upon the needs of the data, the type of data available, and the effort willing to put forth to correctly classify these vehicles.

An additional question addressed the use of other than AVC means for collecting vehicle classification data and especially the use of length-based classification. The vast majority of those surveyed (72 percent) do not have any experience in using length-based classification nor are planning to se any such systems. Seven states (22 percent) have some experience with length-based systems but do not use these systems to produce EAL's. Only two states, Idaho and Kansas, indicated that they do use length-based classification and do currently utilize such systems to produce EAL's The length-based data is used to determine the percent of commercial traffic in Idaho, where an average ESAL per commercial vehicle is then used to estimate ESAL's by vehicle type. The length-based data are used in Kansas only when a car/truck differential is required. Otherwise, the 13-class system is used for classification.

Since part of the research effort is to also develop lane distribution factors for use in calculating equivalent axle-loads, the use of such factors among the various state agencies was also surveyed. Most of the states responded (22 states, 69 percent) do not currently use any lane distribution factors. The remaining ten states indicated that use some type of lane distibution procedures are applied. These methods include detailed calculation of lane distribution factors (3 states), default values of percent of heavy trucks using the right lane (4 states), use of the values indicated in the Highway Capacity Manual (1 state), using estimates from SHRP sites (1 state), and using the factors from the AASHTO Interim Guide (1 state).

The last item in the survey asked the responding agencies to voice any comments on the current practices of vehicle classification counts as well as to indicate any suggestions to future developments in vehicle classification technology. A number of states expressed a concern for the difficulty of developing seasonal adjustment factors as well as the reliability of these factors over time. The desire to move to a length-based classification was another comment made by several states. The amount of data required to develop adjustment factors is probably the single biggest deterrent for developing such factors. The required number of permanent classification stations is another issue that needs to be addressed. Given the fact that unclassified vehicles will always be present, the treatment of these vehicles has a significant impact on the development of any procedures. Finally, an issue that is of concern for a number of state agencies is the reliability of the equipment; a concern that was captured by the unanswered question of a state official "Is there a product available that classifies accurately?"

Seasonal Factors

The next step in this study consisted of developing adjustment factors that can be applied to short-term counts for determining the AADT estimates. This research effort aimed to develop two sets of adjustment factors to account for seasonal and daily variations.

Data were collected from automatic traffic recorders (ATR) from rural Interstates and rural Non-Interstates for developing these adjustment factors, which applied to AADT could then be used in calculating the annual average daily (AADT) traffic volumes by vehicle type. These adjustment factors were developed following the recommended practices in the "Standard Practice for Highway-Traffic Monitoring" (4) but were defined slightly different to accommodate existing data in Kentucky. The adjustment factors were determined for four different classes of roadways--rural interstates and parkways, urban interstates and parkways, rural non-interstates and non-parkways, and urban non-interstates and non-parkways. Moreover, 84 adjustment factors were developed for each day and month of the year (7 days x 12 months) and a separate set was developed for each vehicle type (15 vehicle types).

To proceed with the development of these factors, it was first necessary to calculate the AADT for each vehicle type. To accomplish this task, a three step process was taken. First, the seven averages for each day of the week, Monday through Sunday, in a given month were estimated for each vehicle type. Then, these Monthly Average Days of Week (MADW) traffic were averaged for each month and multiplied by the number of days of the month to estimate the Total Monthly Volume (TMV_i) for each month i. Finally, the AADT was calculated as the division of the total annual volume ($\sum(TMV_i)$, for i =l to 12) by the 365 days of the year. The adjustment factors for each vehicle type k, $(f_{ij})_k$, are then defined for each day i and month j as $(f_{ij})_k = (MADW_{ij})_k / (AADT)_k$

For the development of these factors, six ATR stations were used for interstate calculations and four stations were used for non-interstate calculations. All stations are permanent locations and collect continuous data throughout the year. Two years of data, 1993 and 1994, were used for the calculation of the seasonal adjustment factors. An example of the adjustment factors for rural interstate and parkways for vehicle types 2 (automobiles) and 9 (5-axle, semi-trailer) is shown in Table 2 while similar factors for rural non-interstate and non-parkways are given in Table 3. To obtain an AADT estimate, one needs to divide the short-term count by the corresponding day and month factor. For example, if a count 4,000 automobiles was taken on a rural interstate on a Wednesday in July, the estimated automobile AADT for this road would be (4,000/l .036) = 3,86 1 automobiles.

Figures 1 and 2 illustrate the monthly and daily variation in traffic reflected in through the adjustment factors for the two vehicle types listed above--type 2, automobiles, and type 9, 5-axle semi-trailers. Figure 1 indicates a sharp seasonal variation for automobiles and very little variation for 5-axle, semi-trailers. Figure 2 also indicates a strong relationship between the day of the week and the adjustment factors. Moreover, the adjustment factors for these two vehicle types follow different trends. Traffic counts for semi-trailers during weekdays will show volumes higher than their AADT while weekend counts will show volumes lower than their AADT. The opposite trends are noted for automobile counts--lower volumes in weekdays and higher volumes during weekends.

Figures 3 and 4 demonstrate the monthly and daily variations in traffic for the rural non-interstate and non-parkway data for vehicle types 2 and 9. The monthly variation is still present in Figure 3 but there is less variability for automobiles and higher seasonal variability for the 5-axle semi-trailers. This is an expected phenomenon, since seasonal patterns do affect the travel trends of heavy vehicles on non-interstate roads; contrary to a more constant pattern on the interstate system. Figure 4 illustrates similar trends as those observed for interstate roads (Figure 2) with the exception of more stable trends for heavy vehicles on weekdays and significant reduction of travel in weekends. Moreover, automobiles show less variation between weekday and weekend travel indicating more stable trends for the non-interstate automobile travel.

To demonstrate the usefulness of distinguishing among the various vehicle types, similar adjustment factors were developed for all traffic without any distinction for the vehicle type. These trends are also shown in all figures. An examination of the three trend lines indicates that the general trend of all vehicles follow closely the automobile trends and by no means can be used to approximate the adjustment factors for the 5-axle semi-trailers--an observation that holds true for both intestate and non-interstate system. Therefore, the development of such factors for each vehicle type is essential for accurate estimates of the AADT by different types of vehicles. However, if the desired estimate is the AADT without any distinction for vehicle types then, the adjustment factors shown for all vehicles can be used to adjust the short-term count.

Therefore, it appears that adjustments in traffic volume estimates that account for day of the week and month of the year in which the classification count was obtained could provide better estimates of AADT by vehicle type and thus, can provide a more accurate basis from which other traffic estimates, like ESAL and VMT, can be calculated.

Currently, the validation of these adjustment factors is under way. The validation consists of using these adjustment factors to predict AADT estimates by vehicle type for roadways where ATR counts were taken and then comparing the estimates to the actual counts. Results of this effort should be completed in the near future.

SUMMARY

The questionnaire completed indicates that there is a wide use of the vehicle classification data collected among the various states, with predominant use for traffic operations. Even though data are collected over various times of the year and for different durations, the practice of using seasonal adjustment factors to correct short-term counts is not a widely spread practice among the states. The research effort presented here focuses on the development of such factors for Kentucky. The factoring procedures presented here are developed with the idea of generating accurate estimates for diierent vehicle types and thus, improving estimates of AADT.

A number of issues related to vehicle classification are of concern--improper vehicle classification, unclassified vehicles and their treatment, and so forth--and may pose problems in the development of seasonal adjustment factors. However, the development of adjustment factors by vehicle type is a procedure that can definitely improve AADT estimates and assist in the development of more accurate estimates of VMT by vehicle type.

REFERENCES

1. Pigman, J. G., Allen, D. L., Harison, J., Tollner, N., and Cain, D. H. "Equivalent Single Axle-load Computer Program Enhancements," Research Report KTC-95-7, Kentucky Transportation Center, University of Kentucky, Lexington, KY. 1995

2. US. Department of Transportation, Federal Highway Administration. "Traffic Monitoring Guide, Third Edition," Washington, D.C. 1995

3. Weinblatt, H. "Using Seasonal and Day-of-Week Factors to Improve Estimates of Truck VMT," Transportation Research Board Paper 96053 1, Washington, D.C. 1996

4. American Society for Testing and Materials. "Standard Practice for Highway-Monitoring," ASTM E 1442-92a, Philadelphia, PA. 1993

Table 1. Percentage of unclassified vehicles and their treatment

Percent tolerance	Treatment approach				
	Ignore	Reallocate-- Not specified	Reallocate-- Passenger cars	Reallocate-- Class 15	Ignore or Reallocate
0	1				
<10	2	2			
10	10	3	3	2	2
10>	2				

Table 2. Examples of adjustment factors for rural interstate and parkways

Vehicle type 2--automobiles												
Day	Month											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
SUIT	0.734	0.815	0.975	1.113	1.155	1.323	1.307	1.433	1.229	1.252	1.261	0.910
MON	0.651	0.691	0.792	0.894	0.970	1.012	1.142	1.069	1.011	0.940	0.900	0.897
TUE	0.566	0.631	7 5 0	0.800	0.818	0.921	1.042	0.936	0.881	0.840	0.885	0.841
WED	0.612	0.680	0.731	0.830	0.845	0.959	1.036	0.965	0.867	0.868	1.038	0.898
THU	0.632	0.655	0.868	0.939	0.973	1.076	1.173	1.082	0.933	1.017	1.041	0.987
FRI	0.764	0.815	1.165	1.219	1.283	1.385	1.493	1.440	1.309	1.345	1.194	1.079
SAT	0.767	0.783	1.005	1.066	1.138	1.293	1.394	1.353	1.217	1.189	1.193	0.934

Vehicle type 9--5-axle semi-trailer												
Day	Month											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
SUN	0.601	0.718	0.607	0.563	0.554	0.662	0.535	0.656	0.556	0.671	0.675	0.546
MON	1.050	0.983	0.991	0.992	0.886	0.995	0.826	0.988	0.923	1.089	1.016	0.944
TUE	1.267	1.188	1.284	1.278	1.202	1.253	1.222	1.274	1.232	1.356	1.381	1.168
WED	1.310	1.278	1.299	1.345	1.268	1.309	1.219	1.238	1.266	1.345	1.325	1.282
THU	1.234	1.201	1.261	1.297	1.267	1.274	1.198	1.284	1.271	1.359	1.109	1.204
FRI	1.044	1.014	1.033	1.037	1.011	1.070	0.976	1.125	1.045	1.105	0.917	0.904
SAT	0.583	0.641	0.587	0.554	0.595	0.609	0.623	0.639	0.609	0.703	0.568	0.560

Table 3. Examples of adjustment factors for rural non-interstate and non-parkways

Vehicle type 2--automobiles												
Day	Month											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
SUN	0.696	0.879	0.897	1.087	1.154	1.147	1.102	1.091	1.057	0.936	0.772	0.690
MON	0.819	0.906	0.971	1.072	1.104	1.090	1.036	1.064	0.941	0.851	0.809	0.722
TUE	0.811	0.901	0.980	1.078	1.089	1.093	1.036	1.046	0.983	0.814	0.734	0.764
WED	0.805	0.912	0.896	1.047	1.071	1.069	1.045	1.065	0.941	0.821	0.738	0.786
THU	0.844	0.881	1.004	1.128	1.133	1.107	1.144	1.079	0.997	0.915	0.799	0.836
FRI	0.920	0.969	1.154	1.261	1.330	1.287	1.302	1.232	1.218	1.096	0.965	0.922
SAT	0.866	0.917	1.022	1.200	1.254	1.251	1.228	1.173	1.126	0.956	0.839	0.787

Vehicle type 9--5-axle semi-trailer												
Day	Month											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
SUN	0.317	0.539	0.328	0.356	0.333	0.445	0.375	0.438	0.332	0.339	0.412	0.316
MON	1.000	0.989	1.201	1.309	1.259	1.560	1.157	1.288	1.082	0.972	1.210	1.090
TUE	0.998	1.083	1.248	1.415	1.523	1.565	1.403	1.230	1.296	1.118	1.290	1.193
WED	0.988	1.042	1.225	1.405	1.464	1.569	1.472	1.259	1.356	1.300	1.374	1.184
THU	1.029	0.940	1.135	1.445	1.493	1.595	1.441	1.253	1.276	1.299	1.112	1.139
FRI	1.024	0.765	1.140	1.182	1.535	1.530	1.370	1.289	1.269	1.158	0.984	0.985
SAT	0.342	0.378	0.412	0.472	0.564	0.571	0.518	0.468	0.516	0.391	0.391	0.381

Figure 1. Monthly variation for rural interstate and parkways

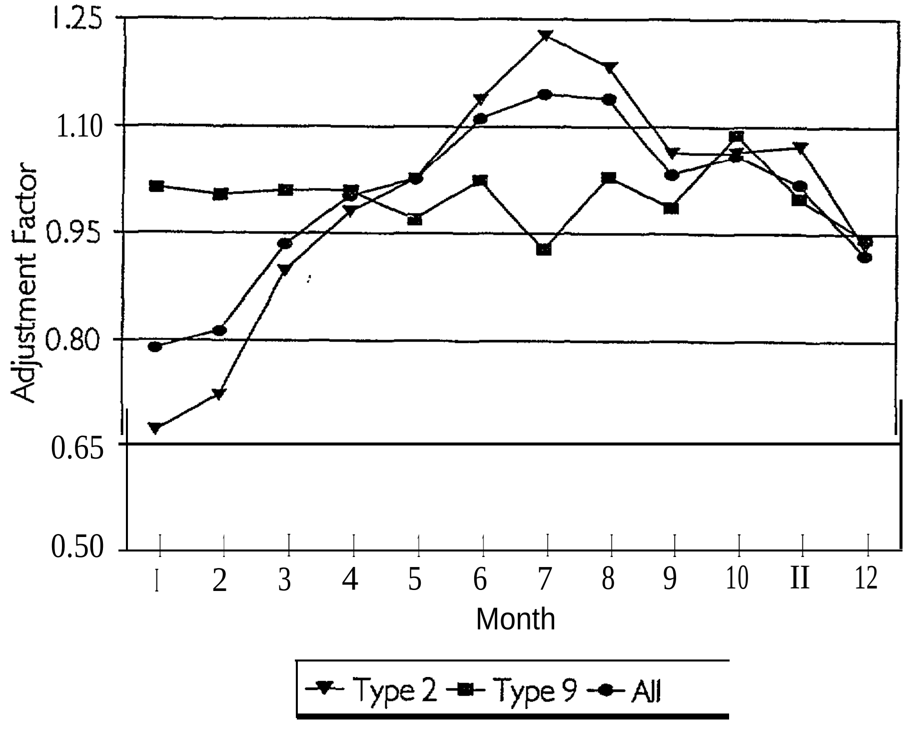


Figure 2. Daily variation for rural interstate and parkways

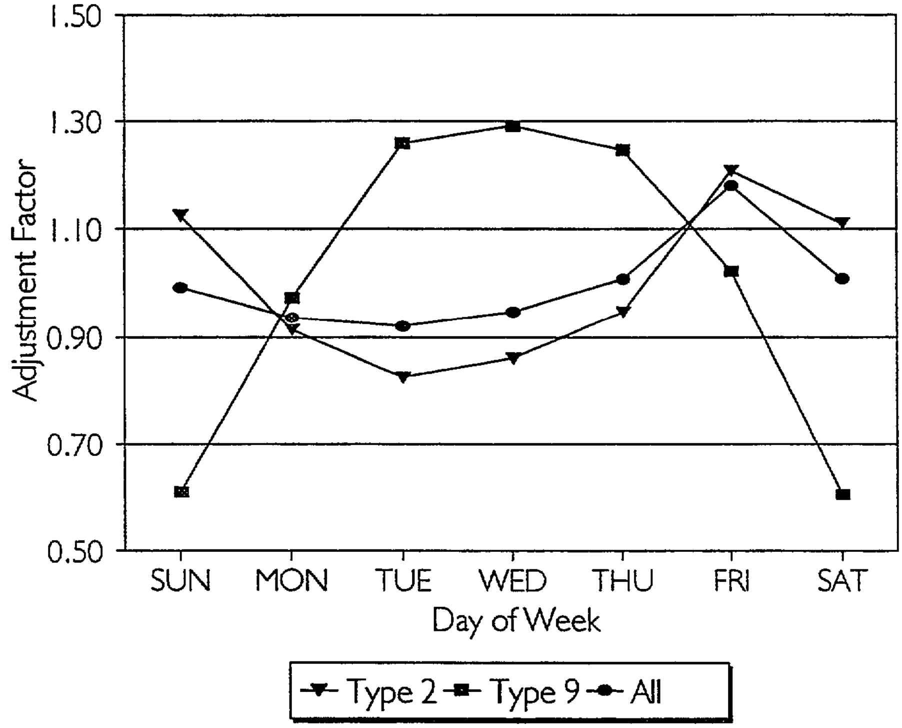


Figure 3. Monthly variation for rural non-interstate and non-parkways

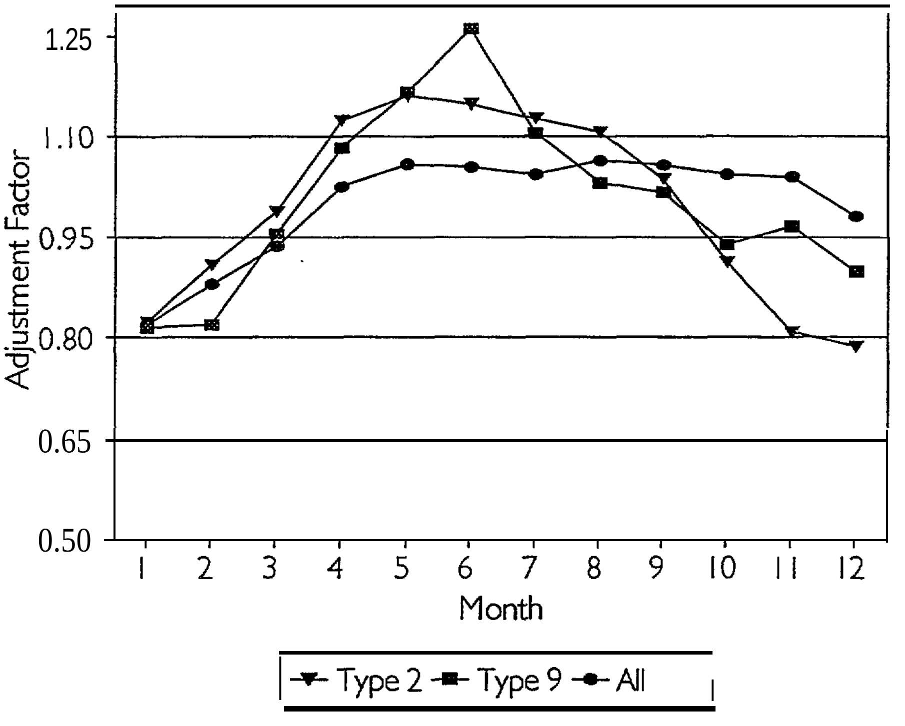
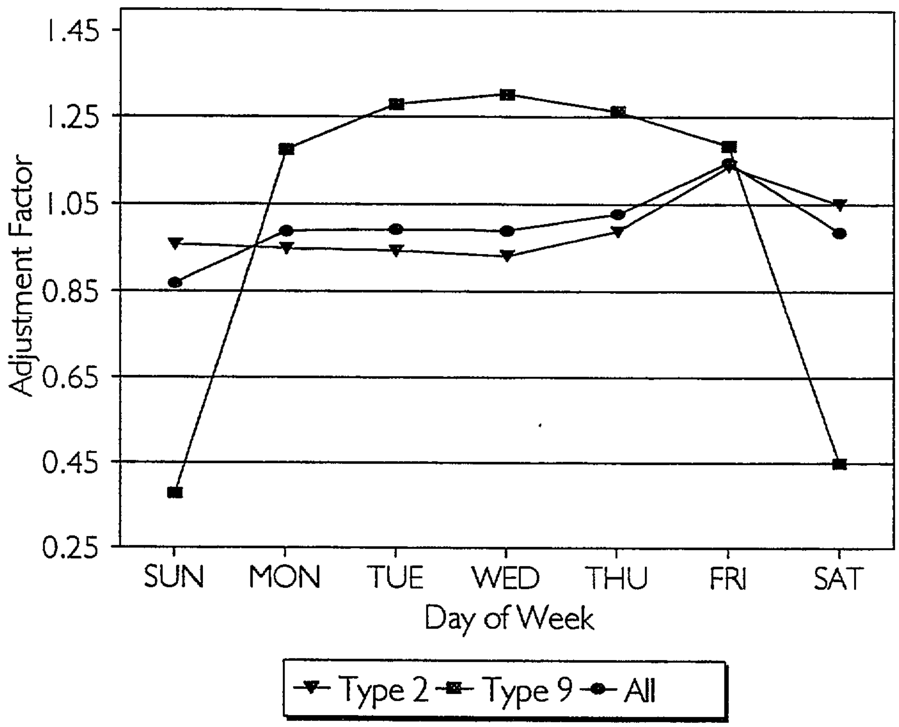


Figure 4. Daily variation for rural non-interstate and non-parkways



FACTORING OF SHORT-DURATION CLASSIFICATION COUNTS

Herbert Weinblatt
Cambridge Systematics, Inc.

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Factoring of Short-Duration Classification Counts

ABSTRACT

Procedures developed by FHWA for "factoring" short-duration traffic counts for seasonal and day-of-week variations in traffic volumes are capable of producing estimates of annual average daily traffic (MDT) that are quite good. Moreover, there is virtually no bias in these estimates, so AADT estimates for a set of sections can be used to produce unbiased estimates of total vehicle-miles traveled (VMT) for systems of roads. Unfortunately, corresponding procedures are not generally used for estimating AADT by vehicle class, and the less sophisticated procedures that are commonly used contribute to substantial overestimates of truck AADT and VMT. Current procedures apparently overestimate VMT by 25 to 40 percent for combination trucks.

This paper presents modified versions of the FHWA factoring procedure that are capable of producing substantially unproved estimates of truck VMT and of AADT of combination trucks. These procedures use seasonal and day-of-week factoring to reduce the errors in truck AADT estimates and to eliminate the upward bias in truck VMT estimates that result from the use of unfactored weekday classification counts.

Keywords: VMT
AADT
Trucks
Seasonal Factoring

Factoring of Short-Duration Classification Counts

Accurate estimates of both vehicle-miles of travel (VMT) by vehicle class and annual average daily traffic (AADT) by vehicle class are required for a variety of highway-related planning and policy analyses. Estimates of VMT by vehicle class are required to derive accident rates by vehicle class and to compare accident rates across classes and also to allocate highway costs across vehicle classes. Estimates of truck VMT also are required in order to understand the importance of trucks to the nation's economy and to evaluate the costs and benefits of potential changes in truck regulation. Estimates of truck AADT by class on individual sections of road are required to design pavements that are appropriate for the truck volumes to be carried.

Although a number of ways exist to estimate national VMT, estimates of AADT on individual highway sections and estimates of VMT on individual systems of roads generally must be developed from traffic counts. However, the development of these estimates from traffic counts poses some nontrivial analytic challenges. Resource limitations usually make it impractical to count traffic on all sections of interest, and most of the sections that are counted can be counted only at infrequent intervals (e.g., once every three years) for a relatively short time period (e.g., 24 or 48 hours). In order to produce reliable estimates from relatively limited data, FHWA has developed "factoring" procedures for adjusting short-duration traffic counts for seasonal and day-of-week variations in traffic volumes (as well as for annual growth in traffic) (1,2).
The FHWA recommended seasonal and day-of-week factoring procedures have been found to produce estimates of total AADT that are quite good. The mean absolute error is only 4 to 7 percent for urban sections and 5 to 10 percent for nonrecreational rural sections (3, Vol. I, p. 23). Moreover, there is virtually no bias in these estimates, so AADT estimates for a set of sections can be used to produce unbiased estimates of VMT for entire systems of roads.

Unfortunately, the procedures currently used by most states for estimating AADT by vehicle class are less sophisticated and produce substantially less satisfactory results. Most typically, "classification" counts are collected for a 48-hour period on weekdays (excluding Friday evenings). These counts are then used, without any seasonal or day-of-week adjustment, as the basis for distributing estimated AADT across vehicle classes. On most roads, truck traffic drops on weekends and, outside of urban areas, automobile traffic generally rises. Accordingly, the use of weekday distributions for apportioning AADT across vehicle classes contributes to substantial overestimates of truck AADT and VMT; current procedures apparently overestimate VMT by 25 to 40 percent for combination trucks, and possibly by more for single-unit trucks with six or more tires.

Substantially better estimates of VMT for all classes of trucks and of AADT for combination trucks can be developed by using modified versions of--the *Traffic Monitoring Guide* (TMG) *(2)* seasonal and day-of-week factoring procedure. The

California Department of Transportation (CalTrans) has used such a procedure since 1993 (4); and the Virginia Department of Transportation (VDOT) is introducing one in 1996 (5 - 7).

The first section of this paper presents several comparisons of the truck VMT estimates derived using traditional count-based estimation techniques with corresponding estimates from other sources. The second section presents a generalized version of the procedure developed by Cambridge Systematics for VDOT that could enable all states to produce improved estimates of truck VMT and of AADT for combination trucks. The concluding section of the paper contains: a recommendation that VMT estimates from the Census Bureau's Truck Inventory *and Use Survey* (TIUS) (8) be used when national truck VMT estimates are needed; and a brief discussion of possible changes to the FHWA system for classifying vehicles that would improve the usefulness of the estimates of VMT and AADT by vehicle class.

CURRENT ESTIMATES OF TRUCK VMT

The Mingo/Wolff Study

FHWA publishes annual estimates of truck VMT reported by the states in the VM-1 Table of *Highway* Statistics (9). A particularly thorough evaluation of these estimates of truck VMT was performed recently by Roger Mingo and Holly Wolff (10).

As an initial step in their evaluation, Mingo and Wolff compared the VM-1 estimates for 1987 to TIUS estimates developed from a quinquennial survey of trucks. In 1987, approximately 135,000 trucks were surveyed (including light trucks). To improve comparability with the VM-1 estimates, Mingo and Wolff used the TIUS public-use file to adjust the TIUS estimates for off-road travel, travel by combination power units without trailers, and travel by government vehicles. Mingo and Wolff also made a separate adjustment to the TIUS estimates to reflect the FHWA practice of classifying trucks with utility trailers as combinations (though this classification is undesirable from the standpoint of most users of the data). However, they apparently did not adjust for recreational vehicles, which are excluded from the TIUS data but frequently included in the VM-1 estimates of trucks with six or more tires.

After all adjustment&, the Mingo and Wolff data indicate that the VM-1 estimates are 0.4 percent lower than the 1987 TIUS estimates for four-tire trucks, but that they are 57 percent higher for other single-unit trucks and 39 percent higher for combination trucks

l Mingo and Wolff actually focused on the results obtained before the adjustment for utility trailers. Before this last adjustment, the VM-1 estimates are 36 percent higher than TIUS for single-unit trucks and 51 percent higher for combinations.

Mingo and Wolff considered various potential sources of the differences in the VM-1 and TIUS estimates. The possibility of systematic underreporting of annual vehicle mileage by truck operators is at variance with a finding by the University of Michigan Transportation Research Institute (UMTRI) that truck operators tend to overestimate annual mileage (by 38 percent for single units and 28 percent for combinations) (10). Also, the TIUS VMT estimates were found to be consistent with FHWA data on diesel fuel consumption. Mingo and Wolff concluded that the principal sources of the differences appeared to be:

1. The derivation of truck VMT estimates primarily or exclusively from weekday classification counts; and

2. Misclassification of light vehicles as trucks.

The significance of the first source of error can be seen from the data in Table 1. This table compares estimates of the average percent trucks obtained from 6, 10, 14, and 24-hour classification counts collected on Tuesdays through Thursdays with the percent trucks obtained using data for an entire year. The data indicates that use of 24 or 48-hour counts collected on Tuesdays through Thursdays increase the estimates of average daily traffic by 19 percent and that the increase probably is greater for combinations than for single-unit trucks. Use of Monday and Friday counts decreases the overestimates somewhat, while collecting counts only during May through September increases the overestimates.

The factoring procedures presented in the second part of this paper address only the first source of error.

Other Comparisons

Table 2 shows five comparisons of estimates of truck VMT or AADT developed using conventional procedures for analyzing data from classification counts with corresponding estimates obtained from alternative sources.

The first comparison in Table 2 is the comparison between 1987 TIUS and VM-1 data discussed above, and the second is a similar comparison between 1992 TIUS data (with off-road VMT excluded) and VM-1 data (11). The comparison of more recent data shows the same difference in the estimates for combination trucks as the earlier comparison but a very sharply reduced difference in the estimates for single-unit trucks.

The remainder of Table 2 compares estimates of truck VMT and AADT obtained using conventional count-based procedures with corresponding estimates developed using three other procedures. Each of these comparisons is described briefly in the subsections that follow. The first of these comparisons is made to estimates of truck VMT derived from odometer readings, while the remaining comparisons are made to count-based estimates that adjust for the use of weekday classification counts but not for misclassification.

All the Table 2 comparisons indicate that conventional count-based procedures produce higher estimates than the estimates produced by the alternative procedures. For combination trucks, the differences between pairs of results are fairly consistent, with the conventional count-based procedures producing estimates that are 27 to 40 percent higher than those produced by the alternative procedures; while, for single-unit trucks, there is substantially more variation in the extent of the differences.

Although we have not performed a detailed study of the reasons for the variation in the differences for single-unit trucks, there is some evidence that the variation on lines 1-3[2] may be due in part to the increasing use over time of automatic vehicle classifiers (AVCs) and to weaknesses in the "Scheme F" classification algorithm used by nearly all AVC vendors. Indeed, data developed in the course of a recent test of AVCs (12, Tables IV, VII, XIII, XIV, XVII, XXI, XXIV, XXVII, XXX, XXXII, XXXV and XXXVIII) indicate that, of twelve AVCs, eleven undercount Class 5 (six-tire) vehicles by over 20 percent and eight undercount these vehicles by more than 50 percent.

The National Truck Trip Information Survey

The National Truck Trip Information Survey (NTTIS) (13,14), conducted by the University of Michigan Transportation Research Institute (UMTRI), consisted of an initial telephone interview covering 6305 medium and heavy trucks conducted in early 1985 and four follow-up interviews conducted over the course of a year. The NTTIS VMT estimates were developed from odometer data obtained in the course of these interviews from 58.6 percent of the respondents. These estimates are compared in Table 2 to VM-1 estimates for 1985 (14).[3]

---

2 The results shown on Lines 4 and 5 are not affected by the classification procedure used, though they are affected by the way short-duration classification counts are used to estimate AADT.

3 The NTTIS also obtained respondent estimates of annual mileage provided during the initial interview and developed a separate set of estimates by mapping descriptions of truck usage obtained for the day preceding each of the follow-up interviews. The respondent estimates were found to be appreciably higher thau the odometer data (even after adjusting for the normal decline in usage over time), while the data for the four survey days produced estimates that were about one-third lower than the odometer readings different than those used in the VDOT procedure.

Oregon

A recent study of the Oregon Weight-Mile Tax (15) required reliable estimates of VMT by vehicle class as one component of a comprehensive analysis of evasion of this tax. In particular, the VMT estimates were required to estimate the extent to which carriers subject to this tax might be underreporting their mileage.

Oregon, like most states, estimates VMT by vehicle class primarily by using weekday classification counts to distribute AADT across vehicle classes. In order to adjust for the resulting overestimates of truck VMT, during November 1994 through January 1995, additional classification data was collected for one or two week periods at eight rural sites and two urban sites. Sydec, Inc., then used these data to estimate how weekday truck volumes compare to volumes on weekends, full holidays, and semi-holidays. Separate correction factors were developed, by highway functional class, for each of several vehicle classes; and these factors were further adjusted for seasonal effects using single-lane classification data collected for the Strategic Highway Research Program (SHRP).

Line 3 of Table 2 compares the resulting adjusted estimates of truck VMT to unadjusted estimates. The unadjusted estimates are appreciably larger than the adjusted ones (by 21 percent for single-unit trucks and 27 percent for combinations), though the differences are smaller than those obtained in the earlier comparisons. The smaller differences result, at least in part, because the analysis focused on adjusting the VMT estimates for day-of-week variations in traffic and it did not consider other issues such as misclassification.

Virginia

The Virginia data shown on Line 4 of Table 2 were obtained from classification counts collected at two multi-lane SHRP sites on I-64 in the Norfolk area (3, Vol. II, pp. 46-49). The first two columns show estimates of actual AADT derived from classification counts obtained for 118 days in 1993 using the standard AASHTO procedure (16) and averaged across the two sites. The next two columns show averages of AADT estimates derived in two different ways from 48-hour weekday counts.

The "unfactored" Virginia estimates were derived by extracting from the data 30 sets of 48-hour weekday classification counts obtained at each of the two sites. Each set of counts was obtained for a 48-hour period starting at noon on a Monday, Tuesday, or Wednesday. Each of these sets of counts was treated as a set of "unfactored" estimates of AADT by vehicle class. The "unfactored" estimates shown in Table 2 were obtained as the averages of all estimates for single-unit trucks and for combinations. These estimates exceed the estimates of actual AADT by 9 percent for single-unit trucks and by 28 percent for combinations.

The "distributed" AADT estimates were derived by using each set of unfactored 48-hour classification counts as the basis for distributing total AADT across vehicle classes. This is the most common procedure for estimating AADT by vehicle class from short-duration classification counts. The last line of Table 2 shows the results of averaging all estimates obtained in this way. These "distributed" estimates are appreciably larger than the "unfactored" estimates, exceeding the estimates of actual AADT by 21 percent for single-unit trucks and 40 percent for combinations. At these two sites, the distribution procedure that is in common use produces higher overestimates of truck AADT than does the simple use of unfactored classification counts. This result will occur wherever total daily traffic is higher on weekends (including Friday evenings and Monday mornings) than it is during weekdays (and the distribution procedure will produce better estimates of truck AADT wherever the reverse is the case).

The "percentage difference" figures shown for the Virginia data represent estimates of the average error introduced by two different procedures for estimating truck AADT from 48-hour classification counts. Like the percentage differences shown for Oregon, the Virginia values reflect the effects of errors that result from the use of unadjusted or improperly adjusted short-duration classification counts, and they do not reflect the effects of any other error sources such as misclassification.

A SEASONAL AND DAY-OF-WEEK FACTORING PROCEDURE

This section presents a generalized version of the procedure for estimating truck AADT and VMT developed by Cambridge Systematics for VDOT (5 - 7). This procedure uses seasonal and day-of-week factoring to reduce the errors in truck AADT estimates and to eliminate the upward bias in truck VMT estimates due to the use of unfactored weekday classification counts (but not the bias due to misclassification). We believe that most states will find some variant of this procedure appropriate for their use.

The overall procedure distinguishes four categories of highway sections and uses different procedures for estimating truck AADT and VMT on sections in each category. The four categories are:

1. Sections that contain permanent automatic vehicle classifiers (AVCs);
2. Sections on which short-duration classification counts are collected periodically;
3. "Nearby" sections on the same road as Category 1 or 2 sections; and
4. All other sections of road.

The first subsection below discusses each of these four categories of highway sections; the second subsection describes the factoring procedure that we recommend for application to the short-duration classification counts collected on sections in Category

2; and a third subsection presents time-of-day factoring procedures that we recommend using for manual classification counts that are collected for periods of less than 24 hours and when permanent AVC counts for a few hours of a day are missing or unreliable.

The Four Categories of Highway Sections

1. *Sections Containing Permanent AVCs*

For each section containing permanent AVCs, AADT by vehicle class (AADTVC) should be estimated by applying the standard AASHTO three-step AADT averaging process to the daily counts for each vehicle class (17, p.52). In the first step of this process, for each vehicle class, seven averages, corresponding to the seven days of the week, are obtained for each month of the year. These "Monthly Average Days of the Week" (MADW) traffic volumes are then averaged across all twelve months to produce a single set of seven "Annual Average Days of the Week" (AADW) traffic volumes for each vehicle class. Finally, the seven AADW values are averaged to produce estimated AADT for the vehicle class.

In order to develop the seasonal and day-of-week factors to be applied to classification counts obtained on Category 2 sections, we recommend that the highway system be divided into at least three "factor groups" (urban, rural Interstate, and rural other) and that permanent AVCs be established on a representative sample of five to eight sections in each factor group. Sections selected for this purpose should have seasonal and day-of-week patterns of truck volume that are reasonably typical of the patterns obtained on other sections in the factor group. Sections with idiosyncratic patterns of usage (such as those occurring in the vicinity of a quarry) should be avoided.

In order to maximize the quality of the seasonal and day-of-week factors, we **recommend** that permanent AVC classification counts be collected continuously (rather than only for one or two weeks each month, as is the practice in many states).

Also, in order to make maximum use of all data that is collected, we do not recommend that counts **for** an entire day be discarded whenever counts for one or two hours in the day are missing or rejected as being unreliable. Instead, when **part of** a day's count is missing or rejected, we recommend that the missing hourly counts be imputed, *provided that* the imputed counts for each vehicle class represent no more than 25 percent of the day's count for that vehicle class. An appropriate imputation procedure is presented in a subsequent subsection of this paper.

2. *Sections Used for Short-Duration Classification Counfs*

The second category of highway sections consists of sections on which short-duration classification counts are collected on a regular basis. These sections should be distributed across all functional systems (with the possible exceptions of the local functional systems and rural minor collectors).

The TMG recommends that each state collect classification counts on at least 300 sections (including sections that are monitored continuously) and that these sections be distributed across functional systems and traffic volume groups on the basis of total VMT in the functional-system / volume-group strata. Short-duration traffic counts obtained with AVCs should be collected for at least one 48-hour period at least once every three years. At locations where AVCs cannot be used reliably because of nonuniform speeds, classification counts may be collected manually for part of a day. A procedure for using partial-day classification counts to estimate daily classification counts is presented in a subsequent subsection.

As discussed earlier in this paper, the raw 48-hour classification counts do not provide very good estimates of AADTVC and, if the counts are taken on weekdays, the estimates for trucks are likely to be quite high. To improve the estimates, AADTVC estimates should be derived by applying seasonal and day-of-week factors to the raw counts and then scaling the results so that the AADTVC estimates for any section add to estimated AADT for that section. The development of the factors is presented in the next major subsection of this paper.

Limited testing of the use of factored counts indicates that factoring generally is quite successful in eliminating the upward bias, and that factoring also tends to improve the accuracy of AADTVC estimates of combination trucks. However, even for combination trucks, the factored estimates for AADTVC are not as reliable as those for AADT, and they are likely to be quite unreliable for sites with unusual day-of-week traffic-volume patterns.

The results of a very limited test of a variant of our proposed factoring procedure are presented in Table 3. This table compares the errors produced using factored 48-hour weekday classification counts to the corresponding errors produced when unfactored versions of the same counts are used for distributing AADT across vehicle classes. The results indicate that, for buses and for all classes of trucks, conventional procedures produce average errors that are large and positive, resulting in substantial overestimates of VMT for these vehicle classes (and underestimates of VMT for four-tire vehicles). However, except for single-unit trucks with three or more axles, the average errors are virtually eliminated by factoring.

Factoring also appears to produce an appreciable improvement in the quality of the estimates of AADT of combination trucks for individual sections (as measured by the mean absolute error). The estimates for multi-trailer combinations (Classes 11-13) are not as good as those for single-trailer combinations (Classes 8-l0), primarily because of the low daily volumes of vehicles in these classes. Similar effects might have been observed for Classes 8 and 10 if we had analyzed AADT estimates for these classes. separately; and very poor estimates (in percentage terms) would be expected for Classes 12 and 13, two very low volume vehicle classes.

Factoring was less successful in reducing the mean absolute error of AADT for buses and single-unit trucks. The high errors for these vehicle classes appear to be due to substantial site-to-site variation in the relative weekday and weekend volumes of vehicles in these classes. The poorest results were obtained for single-unit trucks with three or more axles, a category of vehicles subject to relatively unpredictable variations in traffic volume.

3. *Sections That Are "Near" a Category 1 or 2 Section*

Consider a section that is on the same road as a Category 1 or Category 2 section but several miles away from it. If the "character" of the traffic on the road does not change significantly between the two sections, then the two sections are likely to have very similar distributions of traffic across vehicle classes. On the other hand, if there is an intervening intersection or interchange with a major truck route, or if there is a significant change in the volume of commuter traffic as the road changes from urban to rural, then the two distributions are likely to be less similar.

Category 3 sections consist of all sections that are relatively close to a Category 1 or 2 section and that are likely to have traffic distributed across vehicle classes in the same way as the Category 1 or 2 section. For such sections, we recommend that AADTVC be estimated by assuming that total AADT be distributed across vehicle classes in accordance with the distribution on the "nearby" Category 1 or 2 section (or, in some cases, in accordance with an appropriately weighted average of the distributions on two such nearby sections). When the comparison section is a Category 1 section, this procedure is likely to produce **substantially better estimates of AADTVC than can be** obtained using factored short-duration classification counts.[4]

The determination of whether a section is sufficiently close to a Category 1 or 2 section **requires some judgment. In urban areas, the sections usually should be** no more than a few miles apart However, on rural arterials, 50-mile separations frequently may be appropriate; and, in Nevada, all sections of the 235~mile stretch of I-80 between Wendover and Winnemucca might be treated as being "neaf each other (with the possible exception of a one-mile section adjoining Wells, which might have an elevated volume of local four-wheel vehicles).

The quality of AADTVC estimates produced for a Category 3 section will depend upon whether the comparison section is a Category 1 or 2 section and the degree of similarity of the character of the traffic on the two sections. However, even when the similarity of

[4] **The VDOT system uses approximately** *150* **continuous AVCs on the National Highway System (NHS) so that all sections of the Interstate System and most other sections of the NHS are Category 1** or **3 sections for which high quality estimates of AADTVC are produced.**

character is relatively weak, the **quality** of AADTVC estimates is likely to be at least as good as can be obtained by treating the section as a Category 4 section.

4. All Other Sections

Highway sections in each functional system **should** be grouped on the basis of their traffic volumes using volume groupings such as those recommended in the *Highway Per-romance Monitoring System Field Manual* (17, Appendix F). For each functional system and each of the corresponding traffic volume groups, a set of distribution factors should be developed by:

1. **aggregating the AADTVC estimates obtained for the Category 1 and 2 sections in the functional-system/volume-group stratum; and**
2. **dividing these results by total AADT for these sections.**

As observed previously, the distributions obtained for Category 2 sections are likely to be appreciably less reliable than those obtained for Category 1 sections. Accordingly, for strata that contain sections belonging to both categories, we would be inclined to weight the distributions for the Category 1 sections more heavily - perhaps by developing separate sets of distribution factors for the Category 1 sections and the Category 2 sections, and then obtaining a weighted average of these two sets of distributions. Alternatively, it may be desirable to use data from Category 2 sections only when developing distributions for strata that do not contain any Category 1 sections[5]

Estimates of AADTVC for sets of Category 4 sections in a given functional system and volume group should be developed by applying the distribution factors obtained for that functional-system / volume-group stratum to estimated average AADT for all sections in the set. The resulting estimates of AADTVC are likely to be relatively unbiased, and so they can be used in the development of unbiased estimates of VMT by vehicle class. However, the AADTVC estimates may not be particularly accurate for individual sections in the set, and they should not be used when accurate AADTVC estimates are required for individual sections.

---

5 As described in the text, if there is more than one Category 1 section (resp., Category 2 section) in a given stratum, the distribution used would reflect an AADT-weighted average of the distributions for all the Category 1 (resp., Category 2) sections. If the distribution of AADT across all Category 1 and 2 sections in the stratum is reasonably representative of the distribution across all sections in the stratum, such a weighted average is appropriate. However, if the Category 1 and 2 sections tend to have higher AADTs than other sections in the stratum, the procedure in the test will overrepresent the distributions of relatively high volume sections. If this is the case, it probably would be preferable to develop unweighted averages of the distributions for all Category 1 sections and, separately, of those for all Category 2 sections (and then obtaining a weighted average of these two sets of distributions).

Seasonal and Day-of-Week Factors

Seasonal and day-of-week factors for factoring classification counts should be developed separately for several groups of vehicle classes for each of several highway factor groups. For this purpose, we recommend that at least three highway factor groups be distinguished: urban, rural Interstate, and rural other. A separate factor group for non-Interstate rural roads on the National Highway System might also be considered.

We suggest that five groups of vehicle classes be distinguished:

a) Four-tire vehicles (Classes 2 and 3);

b) Buses (Class 4);

c) Other six-tire, two-axle vehicles (Class 5);

d) Other single-unit vehicles with three or more axles (Classes 6 and 7); and

e) Combination trucks (Classes 8-13).

The development of factors for Groups (a) and (e) are discussed in the first subsection below. The development of factors for the remaining three groups pose some additional complications and are discussed in the second subsection.

Four-Tire Vehicles and Combination Trucks

For Vehicle-Class Groups (a) and (e), we recommend that separate sets of 84 Combined Month and Day-of-Week (CMDW) factors be developed for each Category 1 section and for each factor group.

For a given Category 1 section, a set of Group (a) CMDW factors is obtained by using daily counts for four-tire vehicles with the AASHTO procedure for deriving MADW traffic volumes and AADT (16), and then dividing the 84 MADW values for **four-tire** vehicles by AADT for these vehicles. In this process, we suggest the MADW values be computed twice: once using all available daily data; and once excluding any weekdays for which it is known that routine short-duration traffic counts were not collected (e.g., holidays, Thanksgiving Friday, and Christmas / New Year's week). If two sets of MADW values are computed, the first set should be used for deriving AADT and the second set should be used as the numerators of the MADW / AADT ratios.

For each factor group, a set of Group (a) CMDW factors is obtained as an unweighted average of the factors obtained for **all** Category 1 sections in the factor **group.**

Sets of 84 Group (e) CMDW factors for each Category 1 section and for each factor group are obtained by applying the same procedure to daily counts for combination trucks.

Factoring works best when the factors applied to a given count are developed from data for a 12-month period that includes the days on which that count was taken (3, Vol. I, pp. 12-15). Hence, for the purpose of estimating VMT for a given year, we prefer that final computations not be performed until the year is over and a complete set of factors for that year can be developed and applied. If AADT or VMT estimates are required for individual sections or counties on a more timely basis, we prefer that they be developed using factors developed for a rolling 12-month period that ends in the month when the section was counted or when counting in the county was completed. Use of factors developed entirely from counts obtained in the preceding year is a less desirable, but also less computer intensive, alternative.

Other Single-Unif Vehicles

Daily traffic volumes for Vehicle-Class Groups (b) and (d) and, to a lesser extent, for Group (c), are appreciably lower than they are for Groups (a) and (e). For this reason, the above procedure may produce some relatively unreliable factors for individual Category 1 sections, particularly for rural non-Interstate sections. Accordingly, some modification to the above procedure probably would be desirable for Groups (b) and (d), and perhaps for (c) as well.

We have tested the possibility of combining two or three of these groups and found that these combinations reduced the effectiveness of the factoring procedures, primarily because of differences in the three groups' day-of-week patterns of traffic volume. Accordingly, we do not recommend that separate factors be developed for each of these three groups. Instead, we suggest that consideration be given to using a modified version of the Group (a) and (e) procedure for Groups (b) and (d), and perhaps for Group (c) as well. Combinations of the three following modifications would appear to be appropriate:

1. Instead of developing 84 CMDW factors, develop seven day-of-week factors and 12 separate monthly factors, applying these factors in pairs. The day-of-week factors are obtained by dividing AADW traffic volumes by AADT, and the monthly factors are obtained by dividing Monthly Average Daily Traffic (MADT) by AADT. When used to estimate total AADT, we have found that these "Separate Month and Day-of-Week' (SMDW) factors increase the mean absolute error by about 0.1 percentage points and the root-mean-square error by about 0.25 percentage points. . However, since the numerators of these factors (AADW and AADT) are derived using a more extensive set of traffic counts than are used for MADW (the numerator of the CMDW factors), the SMDW factors should

be appreciably less sensitive to random variation in daily bus and truck volumes.

2. Instead of obtaining factors for a factor group as an unweighted average of the factors for the Category 1 sections in the factor group, obtain them as a *weighted* average, using as weights AADT for the relevant Vehicle-Class Group. Using weighted averages for a particular Vehicle-Class Group diminishes the effect of the factors computed for the sections that have relatively low counts of vehicles in the group.

3. Reduce the number of highway factor groups used.[6]

Time-of-Day Factors

If manual classification counts collected during part of a day are used at some sites, time-of-day factors should be used to convert these counts to estimates of total traffic by vehicle class for that day. Also, we recommend that time-of-day factors be used to impute daily counts at Category 1 classification sites when some hourly counts are missing or unreliable.

Procedures for handling missing hourly counts and for converting partial-day counts to estimated daily counts are presented in the two subsections below.

Missing Hourly Counts

We present below a procedure for imputing daily classification counts when some hourly classification counts are missing or unreliable. An obvious analog of the procedure can be used for imputing total daily volume when some hourly volume counts are missing or unreliable.

For the purpose of developing the imputation factors, we suggest that data for vehicle classes with similar time-of-day usage patterns be grouped, and that all uncommon vehicle classes be grouped with a more common class. Possible groupings are: Classes 2 and 3; 6 and 7; 9 and 10; and 11 and 12, or 11-13.

For *the* purpose of imputing daily *counts, we* define a set of *holiday-affected* days (HADs) to consist of all holidays, holiday weekends (including Thanksgiving Friday), and the weekdays immediately preceding and following any holiday weekend. The imputation

6 For Vehicle-Class Groups (b) - (d), the VDOT system uses just two factor groups (one urban and one rural) and uses weighted averages to develop factors for each factor group:.

procedure presented below works best on non-HADs. This procedure involves the following steps:

1. For each continuous AVC site and each day of the week, consider the set of daily and hourly counts available for those non-HADs for which reliable classification counts have been obtained for every hour of the day.

2. For each continuous AVC site and each hour of each day in the set developed in Step 1, obtain an hourly *fraction* for each group of vehicle classes as the ratio of the hourly classification counts for the group to the corresponding daily classification counts.

3. For each continuous AVC site, each hour of the day, each day of the week, and each group of vehicle classes, obtain an *hourly fraction* by day of the week (HFDW) by averaging the corresponding Step 2 hourly fractions (e.g., obtain a Monday l-2 A.M. HFDW for buses by averaging all the Step 2 l-2 A.M. hourly fractions for buses for non-HAD Mondays).

4. For each continuous AVC site, any single day, and any vehicle class for which some (but not all) hourly counts are missing or unreliable, obtain the sum of the HFDWs for the missing hours for the corresponding group of vehicle classes. This sum is the estimated missing *fraction of the* daily *count* (MFDC) for that group of vehicle classes.

5. If the MFDCs are acceptably small (as discussed below), impute a set of daily classification counts for the site by dividing the sum of the reliable hourly counts for the day for each vehicle class by the corresponding value of l-MFDC.

The reliability of the resulting imputed daily classification counts varies inversely with the values of the MFDCs and is appreciably greater for non-HADs than for HADs For non-HADs, we suggest using imputed classification counts for a given day when none of the MFDCs calculated in Step 4 exceeds 0.25; and, for HADs, we suggest using imputed counts when none of the MFDCs exceeds 0.1. When higher MFDCs are obtained, we are inclined to reject the imputed counts as not being sufficiently reliable.

We suggest recomputing the HFDWs at least once every three years using classification counts for at least a 12 month period.

Partial-Day Classification Counts

If manual classification counts are collected at some sites during part of a day (e.g., at urban sites where AVCs cannot be used reliably), a variant of the above procedure should be applied to these counts in order to estimate daily classification counts for these sites. For this purpose, for each highway factor group and each vehicle-class grouping, a set of HFDWs should be developed by averaging the corresponding HFDWs obtained for the continuous AVC sites in the factor group.

For each partial-day classification count, the HFDWs for the appropriate factor group should be used to compute MFDCs for each vehicle class (as in Step 4 of the above procedure). These MFDCs should then be applied to the partial-day classification counts to produce a preliminary set of estimated daily classification counts (as in Step 5). Finally, we recommend that a separate count of total traffic volume be obtained at the site with an automatic traffic recorder (ATR) for a 24-hour period that includes the period during which classification counting was performed, and that the preliminary set of classification counts be scaled to be consistent with this count of total traffic.

RECOMMENDATIONS

Sources of Truck VMT Estimates

The factoring procedures presented in the preceding section are designed to eliminate the upward bias in estimates of truck AADT and VMT that are obtained primarily or exclusively from weekday classification counts. Accordingly, implementation of these procedures should make it possible to obtain substantially improved estimates of truck AADT and **VMT.** However, as stated earlier, these procedures do not address the biases that result from consistent misclassification of trucks – a bias that should be addressed by the development of better classification algorithms for AVCs.

We believe that the estimates of truck VMT produced using the above factoring procedures will be substantially **better than the** state and local truck VMT estimates currently produced in nearly all states. However, because of the inherent difficulty of developing VMT estimates from traffic counts and the current limitations of automatic vehicle classifiers, the resulting estimates of national truck VMT are unlikely to be better than those already available from the Bureau of Census' Truck *Inventory and Use Survey* (TIUS). The TIUS estimates of national truck VMT appear to be the best estimates currently available, and they are likely to remain so for the foreseeable future. Accordingly, we recommend that all studies requiring estimates of national truck VMT use the TIUS estimates.

The Definition of Combination Vehicles

Mingo and Wolff (9) make two recommendations relating to the VM-1 definition of combination vehicles:

1. Produce separate estimates of travel by truck-trailer combinations and tractor-trailer combinations; and

2. Exclude light trailers and light trucks from both categories of combinations.

We view the second recommendation as being particularly desirable. Most studies that require the use of vehicle classification data are concerned primarily with the numbers and activity of heavy trucks and heavy trailers. Adoption of the second recommendation would greatly enhance the value of VM-1 data to policy analysts and would reduce opportunities for misinterpretation of the data.

Implementation of the second recommendation can best be accomplished by expanding the number of vehicle categories distinguished by pertnanent AVCs with weight-sensing capability.[7] For data-collection purposes, it probably will be necessary to obtain separate counts of light vehicles pulling light trailers, heavy vehicles pulling light trailers, and heavy vehicles pulling heavy trailers. Counts obtained using portable classifiers without weight-sensing capability will not be able to make these distinctions, but the overall counts of vehicles with a given axle configuration can be allocated into the three categories using data from the permanent classifiers. The data developed in this way can then be aggregated into a more appropriate set of publication categories.

GLOSSARY

AADT	Annual average daily traffic
AADTVC	Annual average daily traffic by vehicle class
AADW	Annual average day of the week
ATR	Automatic traffic recorder
AVC	Automatic vehicle classifier
CMDW	Combined month and day-of-week
HAD	Holiday-affected day
HFDW	Hourly fraction by day of the week
MADW	Monthly average day of the week
MFDC	Missing fraction of the daily count
SMDW	Separate month and day-of-week
TIUS	Truck Inventory and Use Survey
VMT	Vehicle-miles of travel

7 We also recommend that consideration be given to establishing triple-trailer combinations as a separate vehicle category and to combining the categories of five and six-axle multi-trailer combination trucks (Classes 11 and 12).

ACKNOWEDGMENT

This paper is based primarily on studies sponsored by the Virginia DOT and FHWA. The analyses reported were performed by Carolyn Frank of Information Systems and Services, Inc., and Richard Margiotta of Science Applications International Corporation. Mark Hallenback, of the Washington State Transportation Center, provided a substantial amount of insight into practical considerations in the collection, interpretation and factoring of traffic data; and additional guidance and insight were provided by our project monitors: James B. Robinson and Eugene Martin of VDOT, and Tony Esteve of FHWA.

REFERENCES

1 M. E. Hallenbeck and L. A. Bowman. *Development of a Statewide Traffic Counting Program Based on the Highway Performance Monitoring System*. FHWA, U.S. Department of Transportation, 1984.

2. *Traffic Monitoring Guide.* FHWA, U.S. Department of Transportation, Third Edition, 1995.

3. Cambridge Systematics, Inc., Science Applications International Corporation, and Washington State Transportation Center. *Use of Data from Continuous Monitoring Sites.* FHWA, U.S. Department of Transportation, Two Volumes, 1994.

4. J. Avis, California Department of Transportation, personal communication, 1995.

5. Cambridge Systematics, Inc., Gorove/Slade Associates, Inc., and Information Systems and Services, Inc. *Traffic Monitoring Systems* Development Study, Task 2: *Review and Evaluation of the Current TMS and Recommendations for an Improved TMS.* Virginia Department of Transportation, Richmond, VA, 1994.

6. Cambridge Systematics, Inc., Gorove/Slade Associates, Inc., and Information Systems and Services, Inc. *Virginia State Traffic Monitoring Standards.* Virginia Department of Transportation, Richmond, VA, 1995.

7. Cambridge Systematics, Inc., Gorove/Slade Associates, Inc., and Information Systems and Services, *Inc. TMS Algorithms.* Virginia Department of Transportation, Richmond, VA, 1995.

8. U.S. Bureau of the *Census. 1992 Census of Transportation:* Truck Inventory and Use Survey: United States. TC92-T-52. U.S. Department of Commerce, 1995.

9. Federal Highway Administration. *Highway Statistics. Annual.*

10. R. D. Mingo and H. K. Wolff. Improving National Travel Estimates for Combination Vehicles. *In Transportation Research Record* 1511, TRB, National Research Council, Washington, DC, 1996, pp. 42-46.

11. R. D. Mingo, R. D. Mingo and Associates, personal communication, 1996.

12. B. A. Harvey, *et-al.* *Accuracy of Traffic Monitoring Equipment*. Georgia Tech Research Institute, Atlanta, GA, 1995.

13. D. L. Massie, K. L. Campbell and D. F. Blower. Large Truck Travel Estimates from the National Truck Trip Information Survey. In *Transportation Research Record 1407,* TRB, National Research Council, Washington, DC, 1993, pp. 42-49.

14. D. L. Massie, K. L. Campbell and D. F. Blower. Comparison of Large-Truck Travel Estimates from Three Data Sources. In *Transportation Research Record* 1407, TRB, National Research Council, Washington, DC, 1993, pp. 50-57.

15. Cambridge Systematics, Inc., and Sydec, Inc. *Evasion and Enforcement of Oregon's Weight-Mile* Tax. Oregon Legislative Revenue Office, Public Utilities Commission, and Department of Transportation, Salem, Ore., 1995.

16. *AASHTO Guidelines for Traffic Data Programs.* American Association of State Highway and Transportation Officials, Washington, DC, 1992.

17. *Highway Performance Monitoring System Field Manual.* FHWA, U.S. Department of Transportation, 1993.

TABLE 1 Truck Overcount Ratios for Tuesday-to-Thursday Counts

<table>
<tr><th>Sample Time Period</th><th>Single Units</th><th>Combinations</th></tr>
<tr><td>Midnight to Midnight</td><td colspan="2">1.19</td></tr>
<tr><td>6 AM to 8 PM[b]</td><td>1.09</td><td>1.18</td></tr>
<tr><td>8 AM to 6 PM[b]</td><td>1.15</td><td>1.28</td></tr>
<tr><td>10 AM to 4 PM[b]</td><td>1.43</td><td>1.56</td></tr>
</table>

a Reference 4. Data is for all trucks combined from a single site in California.

b Reference 9. Data is for six sites in California.

TABLE 2 Comparisons of Truck VMT and AADT Estimates from Alternative Sources

Source	Estimates from Source		Estimates from Conventional Truck counts		Percentage Difference	
	Single Units	Combs.	Single Units	Combs.	Single Units	Combs.
1. 1987 TIUS VMT[a]	31.6 B	62.2 B	49.6 B	86.3 B	57%	39%
2. 1992 TIUS VMT (II)	51.2B	71.5 B	53.7 B	99.1 B	5	39
3. NTTIS VMT (14)	29.5 B	62.0 B	45.8 B	79.1 B	55	28
4. Oregon VMT (15)	748 M	1214 M	904M	1546 M	21	27
5. Virginia AADT[a]	1366	2200				
Unfactored			1493	2824	9	28
Distributed			1654	3078	21	40

a See Text

TABLE 3 Quality of Estimates of AADT by Vehicle Class

			Conventional Estimates		Factored Estimates	
Vehicle Class		Actual AADT	Average Error	Mean Absolute Error	Average Error	Mean Absolute Error
4.	Buses	156	+27	31	+0.5%	23%
5.	2-axle, single-unit trucks	1090	+16	23	-0.9	31
6-7.	3+ axle, single-unit trucks	403	+36	44	+28.9	52
S-10.	Single trailer combinations	2742	+41	41	+1.0	5
11-13.	Multiple trailer combinations	82	+38	38	+0.4	9

VEHICLE CLASSIFICATION: FLORIDA'S PERSPECTIVE

Rick Reel
Florida Department of Transportation

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Vehicle Classification: Florida's Perspective

The classification of vehicles by type meets several needs of the Florida Department of Transportation (FDOT) as well as other public and private entities. First, projections of the number of vehicles by type assists pavement design engineers in determining desired pavement strength and composition. Second, these data are used by planners to assess a highway's Level of Service. Next, business persons use vehicle classification data to identify and evaluate sites for placement of businesses. The data are also used by reporters to prepare their stories. Finally, vehicle classification data are used by traffic engineers to develop axle adjustment factors that are use to compensate for the over-counting with pneumatic tube counters caused by vehicles with more than two axles.

Vehicle classification data are obtained using a variety of methods, including manual counts and the use of portable and continuous automated equipment. FDOT purchased its first portable automated vehicle classifiers in 1983. They were integrated progressively into the vehicle classification program until 1990, when manual vehicle classification was discontinued. Presently, manual classification is used only for turning movement studies at intersections. Traditionally, vehicle classifiers were installed using two pneumatic tubes. Recently, permanent sensors have been installed in the pavement. The sensor array consists of two inductive loops and one piezoelectric axle sensor. Vehicle classification surveys are conducted from one to four times per year with a duration of 24 hours in urban areas and 48 hours in rural areas.

Since 1988, all new continuous traffic monitoring sites that have been constructed are capable of vehicle classification. In 1995, 162 continuous classifiers were operational. The conversion of the program to continuous classifiers was undertaken provide a base of data that could be used in the future to develop a method of factoring portable classification survey data into more accurate annual representations.

FDOT reports vehicle classification in four ways. These include the Annual Vehicle Classification Report, axle correction factors, the traffic flow map, and the allocation of a truck percentage to every segment of state highway on the Department's Roadway Characteristics

Inventory (RCI) database. The Annual Vehicle Classification Report includes classification percentages averaged by station on an annual basis, irrespective of the number of surveys. The annual classification percentages are multiplied by the stations' AADT to provide annual average vehicles per class. The axle correction factor report allows the Districts to determine which factors should be applied to axle count data taken on each roadway section. The traffic flow map shows graphically on a statewide map the approximate volumes of trucks on each major section of Florida highway. The truck percentages are assigned to each highway segment in the database using software developed by FDOT for that purpose.

There are several problems associated with the FDOT vehicle classification program. The first is the difficulty of assessing the validity of the data. Another issue is that permanently installed piezoelectric axle sensors are not as durable as is desired when they are installed in flexible pavements. It is also difficult to obtain accurate vehicle classification data on multilane undivided highways. The classification of vehicles in a center paved median (two way left turn lane) is an unresolved issue. Finally, automated vehicle classification does not work under stop and go traffic conditions. These situations most frequently occur during the morning and evening rush hours in urban areas and is often a result of queuing at traffic signals.

FDOT experience indicates that short term vehicle classification surveys are not feasible. As previously indicated, prior to 1983 all vehicle classification surveys in Florida were performed manually. The last procedure for manual surveys called for a traffic technician to manually tally vehicles from I 1:00 AM to 5:00 PM. The six hour truck percentage was then converted to a 24 hour percentage by applying adjustment factors. These adjustment factors were based on a single study performed by a staff statistician using a limited 24 hour sample. The factors developed from that study were not verified to determine whether they were applicable to any other survey. The application of adjustment factors to new sites was problematic. When the first portable vehicle classifiers were distributed to the FDOT Districts, one of the District statistics engineers conducted a study to compare a 24 hour manual classification to the data acquired using the automated vehicle classifier. He reported that there were a surprising number of trucks traveling on the highways at night. The manual surveys and factors did not accurately depict the truck volumes.

Although FDOT is collecting more vehicle classification data than ever before, the users of those data have grown accustomed to more and better traffic characteristics data, and demand even more. Consequently, FDOT is planning to collect more vehicle classification survey data in the future, as opposed to simple volume counts. One of the FDOT Districts has already taken this step by specifying that all traffic count surveys will be vehicle classification surveys.

TEMPORAL DISTRIBUTION OF VOLUME BY VEHICLE CLASSIFICATION

Mark Hallenbeck
Washington State Transportation Center (TRAC)

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Temporal Distribution of Volume By Vehicle Classification
bY

Mark Hallenbeck
Washington State Transportation Center (TRAC)

Analysis is based on data

from145 sites

in 22 states

stored in the LTPP traffic database

The sites are predominately

rural (78%)

inter-states and principal arterials (85%)

The results presented show national trends

Individual sites may have different characteristics caused by local conditions

---

There are four basic time-of-day patterns for all vehicle classes

commuter cars
weekend or rural cars
business day trucks
through trucks

The "commuter" pattern is the classic double humped, urban, AM/PM distribution

It contains FHWA Vehicle Classes 1,2,3, and 4 on weekdays in all urban functional classes

---

It also appears on weekdays in vehicle classes 1,2, and 3 in some rural functional classes

In rural areas, the morning peak tends to contain a lower proportion of daily travel than occurs in urban areas

---

The "weekend" or "rural" pattern contains these same classes on weekends in both urban and rural areas, and on weekdays in some rural areas

---

In almost all cases, for both the Business Trucking and Car patterns

- 75 percent of travel occurs between 6 AM and 6 PM
- less than 60 percent of travel occurs during this period for the through truck pattern

---

Truck percentage also has very stable patterns over time.

These patterns are essentially equal for all functional classes,

Although the absolute truck percentage changes from functional class to functional class

---

Weekend truck percentages tend to be lower than weekday truck percentages

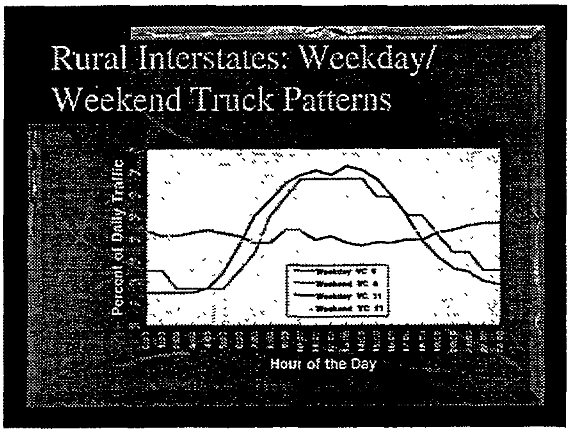
Truck percents are highest in the middle of the night

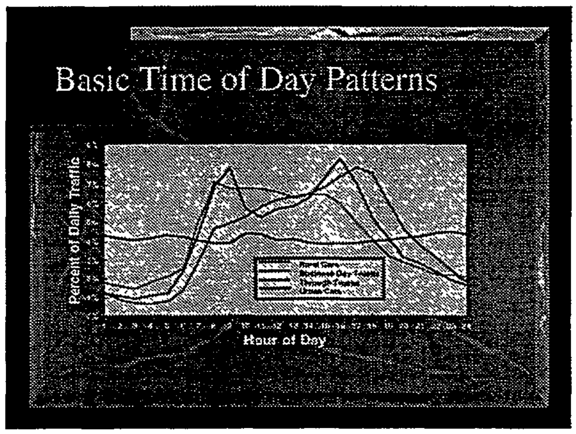
However,

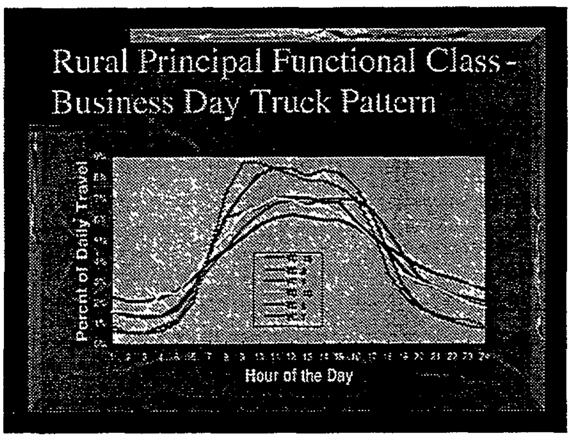
the total volume of trucks and the percentage of daily travel by trucks traveling late at night is quite low

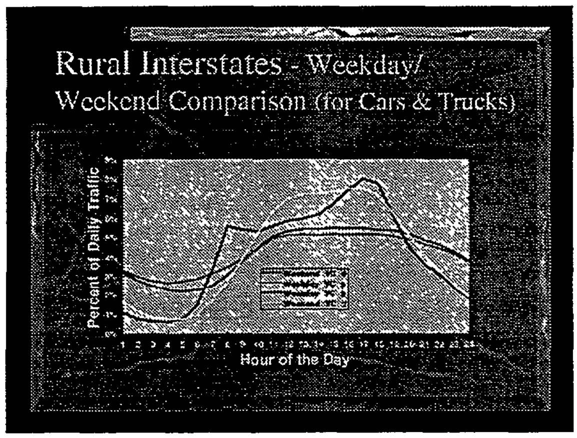
---

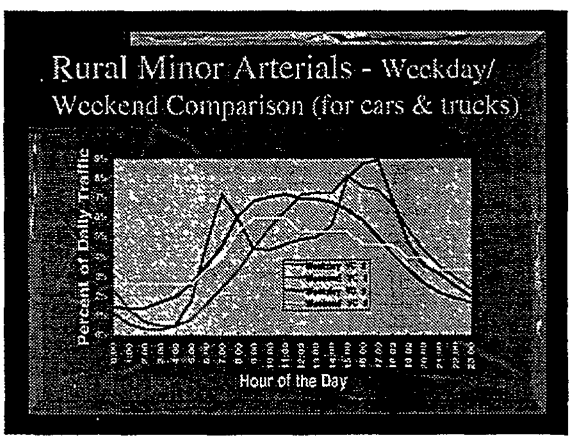
Functional Class	Number of Sites
1	35
2	58
6	15
7	4
8	1
11	19
12	4
14	8
16	1

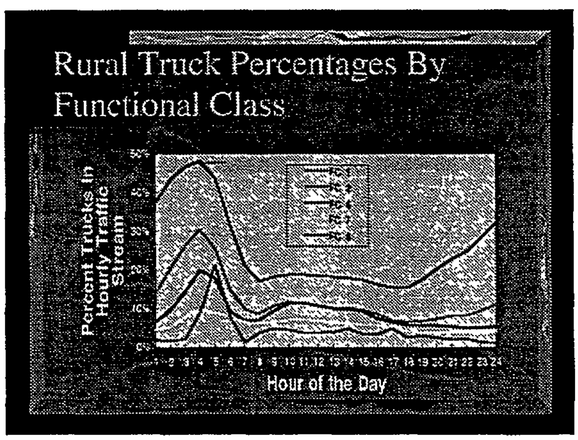

Rural Interstates: Weekday/
Weekend Truck Patterns
Percent of Daily Traffic
Hour of the Day

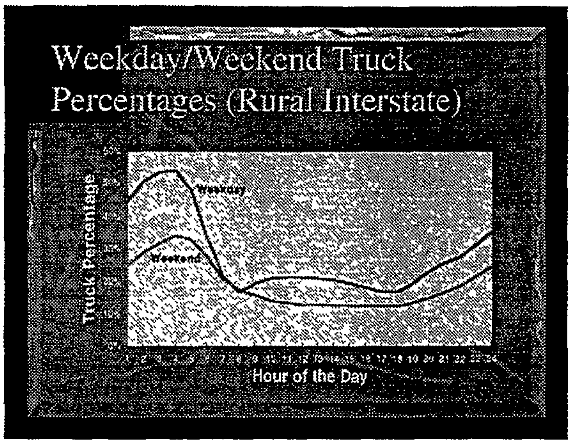

Basic Time of Day Patterns
Percent of Daily Traffic
Hour of Day

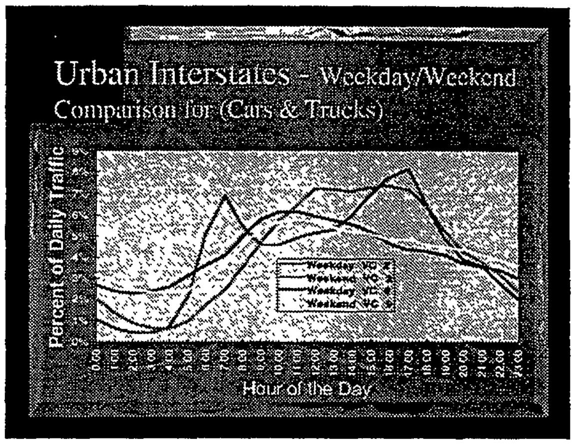

Rural Principal Functional Class -
Business Day Truck Pattern
Percent of Daily Travel
Hour of the Day

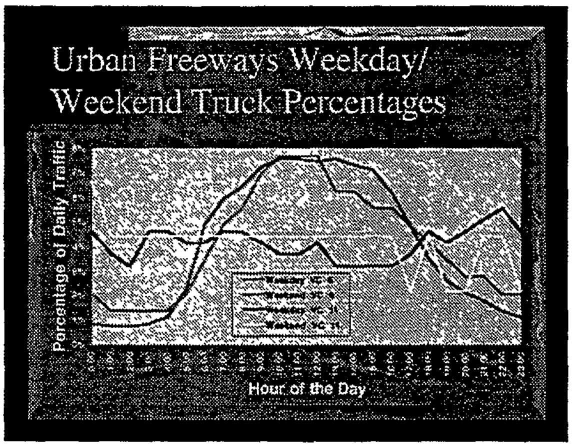

Rural Interstates - Weekday/
Weekend Comparison (for Cars & Trucks)
Percent of Daily Traffic
Hour of the Day

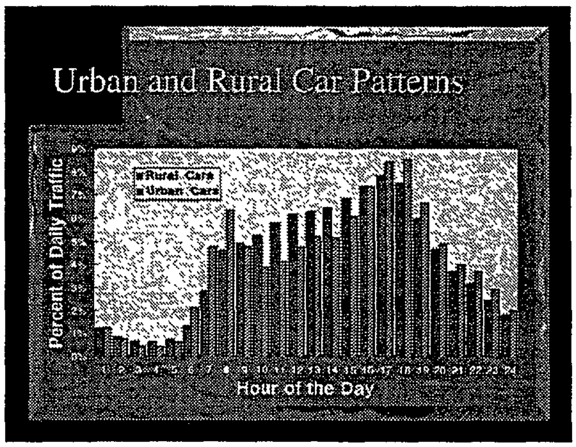

Rural Minor Arterials - Weekday/
Weekend Comparison (for cars & trucks)
Percent of Daily Traffic
Hour of the Day

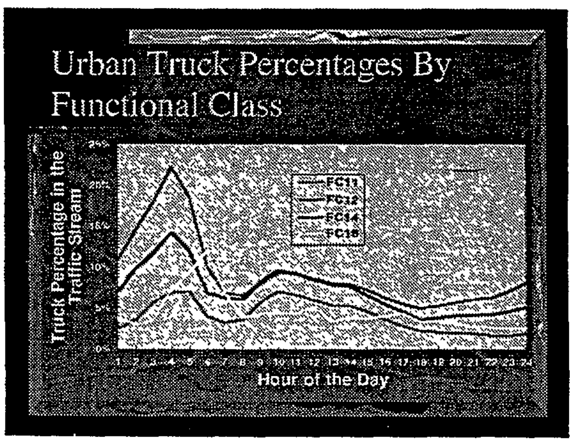

Rural Truck Percentages By
Functional Class
Percent Trucks in Hourly Traffic Stream
Hour of the Day

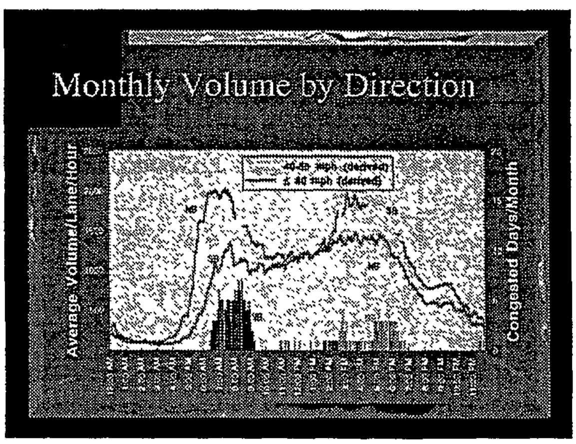

Weekday/Weekend Truck Percentages (Rural Interstate)
Truck Percentage
Weekday
Weekend
Hour of the Day

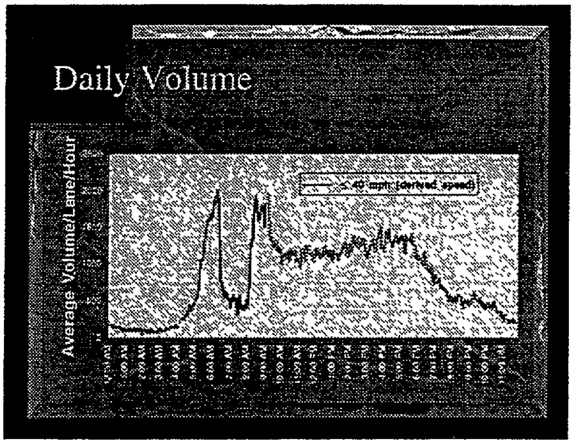

Urban Interstates - Weekday/Weekend Comparison for (Cars & Trucks)
Percent of Daily Traffic
Hour of the Day

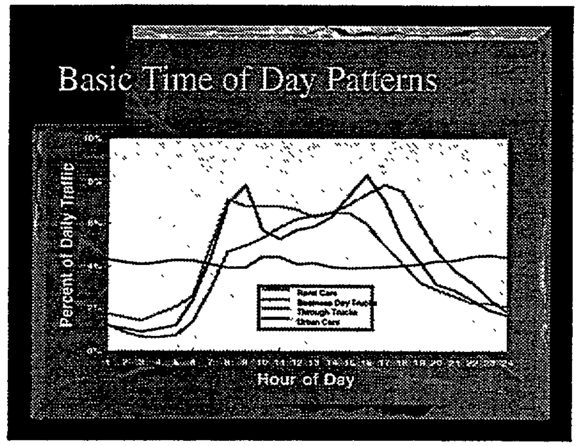

Urban Freeways Weekday/ Weekend Truck Percentages
Percentage of Daily Traffic
Hour of the Day


Urban and Rural Car Patterns
Percent of Daily Traffic
Rural Cars
Urban Cars
Hour of the Day


Urban Truck Percentages By Functional Class
Truck Percentage in the Traffic Stream
FC11
FC12
FC14
FC16
Hour of the Day


Monthly Volume by Direction
Average Volume/Lane/Hour
Congested Days/Month


Daily Volume
Average Volume/Lane/Hour


Basic Time of Day Patterns
Percent of Daily Traffic
Hour of Day

CONCURRENT SESSION 4B - SAMPLING AND PRECISION

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

PRECISION ESTIMATION APPLIED TO TRAFFIC MONITORING DATA

Antonio E. Esteve
Federal Highway Administration

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

Precision Estimation Applied to Traffic Monitoring Data

Introduction

Traffic monitoring data and information needs continue to expand and the use of the information to guide decisions is becoming more common. Once decision-making based on information becomes the norm, the quality of the information becomes a critical concern. Decision makers begin to question the value of the information on which the decisions are based.As a result, the quality of the data is increased and the data itself subject to a higher level of scrutiny. The perspective of data quality quickly changes from a technical concept applied by technicians to a high-level managerial concern dear to the operation and critical to mission success. Traffic monitoring has been approaching this stage for the past few years. Heavy emphasis has been placed on automation and on improving the capabilities of data collection procedures. The 1991 ISTEA placed additional emphasis on information management and introduced requirements resulting in the development of Traffic Monitoring Systems for Highways (TMS/H) in each State to coordinate, stabilize, standardize, and improve the quality of traffic information.

The use of statistical principles in traffic data collection and management has also increased over the last few years. The precision of estimates is a central concept of the statistics profession and probability sampling, where it is used as a means of meeting quality objectives in a cost-effective manner. Traffic monitoring data collection and processing is growing and is expected to grow further as new intelligent transportation initiatives advance. The amount of traffic monitoring data resulting from continuous installations is growing prodigiously and creating situations where we are becoming overwhelmed by the quantities of data available.. Yet, in the midst of this enormous data wealth, we still seem unable to quantify data quality.

The purpose of this paper is to highlight the use of statistical precision as a means of controlling the quality of the traffic monitoring estimates produced, to further the discussion of precision and error estimation, and to mainstream the use of statistical precision concepts to improve the cost-effectiveness of traffic estimation.

Precision in Statistical Sampling

The concept of precision is basic to the development of probability sampling and relates directly to the theory of sample size estimation and confidence intervals. Statistical theory and practice recognize the direct link between accuracy and cost and use the precision concept to provide alternatives between the two. Statistical sampling allows survey professionals to determine the highest accuracy/precision that can be achieved for a fixed cost or alternatively to minimize the cost of meeting a desired precision.

The two terms accuracy and precision are often used interchangeably. Webster's New World Dictionary defines accuracy as "the quality or state of being accurate or exact; precision; exactness." Likewise, precision is defined as "the quality of being precise; exactness; accuracy; definiteness." These very circular definitions clearly indicate that in general language usage, the terms precision and accuracy have the same meaning.

Statistical sampling, on the other hand, makes a clear distinction between accuracy and precision by defining statistical accuracy as the sum of two components: bias and precision. Bias refers to all the error introduced by issues outside-of the statistical sampling design and it is often termed non-sampling error. Statistical precision is defined as the error introduced by the use of a probability sample to estimate the. parameter in question and is usually termed sampling error.

The dictionary definition of both accuracy and precision also mentions exactness. Statistical accuracy implies approximation to an objective based on the concept of estimation. The statistics profession explicitly recognizes that perfection is only achievable by mathematical theory, not by statistical data collection programs. Statistics is an art as much as a science and as a profession should never be confused with accounting.

Statistical precision is usually reported in studies based on statistical sample designs and is used as a measure of how closely a specific estimate is from the 'true" but unknown value the sample was designed to measure. Examples abound in the literature and in studies based on statistical sampling. Specific transportation examples of sampling studies include the Nationwide Personal Transportation Study (NPTS), a household survey sponsored by the USDOT, the Highway Performance Monitoring System (HPMS), and the Truck Inventory and Use Survey (TIUS), a trucking firm survey conducted by the Bureau of the Census.

The major difficulty involving accuracy in statistical studies is the issue of statistical bias or non-sampling error caused by problems with the sample design's ability to represent the universe, coverage of the sampling frame from which the sample is selected, non-response error, or measurement and processing errors introduced in the administration of the survey.

The major difficulties in producing sampling error estimates in statistical studies are the large number of estimates produced and the complexity of the sampling designs employed, which usually necessitate extensive and complex computations. The major theoretical issue with sampling error relates to the dependence on the normal distribution model theory as the source of the established methodology.

Traffic monitoring has its own set of problems in defining accuracy due to the error **introduced by the data collection equipment, the critical inability to apply rigorous statistical** sample designs, the non-random selection of the sample sites, the human difficulties with data collection, and the factoring procedures used to develop the traffic **estimates.** Traffic monitoring is subject not only to the variability of traffic and to the statistical (or data collection) design, but also to equipment error and the human factors involved in traffic data collection. In fact, it is the equipment and human factors that dominate traffic monitoring programs.

Statistical accuracy in traffic monitoring means controlling equipment error. managing the human component. and achieving subjective precision targets within realistic cost constraints. Cost is the limiting or controlling factor in traffic monitoring.' Given unlimited funding; random data collection methods, measurement equipment, and controlled procedures could be implemented to meet any desired accuracy.

The Simple Random Sample Model

The simple random sample equation for determining the sample size needed to achieve a desired precision level derives directly from the confidence interval computation and is shown below:

$$n=z^2 s^2/D^2$$

where

n gives the sample size needed,

z is the value of the normal distribution given a specific significance level,

s is the standard deviation of the characteristic under study, and

D is the numerical precision desired.

In this equation, z and s are basically fixed while n and D are inversely related to one another. Further, the inverse relationship is to the second power (squared) which means-that a reduction in the precision interval by one half (doubling of precision) causes a fourth fold increase in the number of samples needed.

Although the equation is a simplification, it is not a difficult equation to understand or apply, and it clearly explains the trade-offs between sample accuracy and cost, which **indicates that doubling precision quadruples the cost. The equation presented applies only** to estimation of the mean (one of many possible estimates) and based on the **normal** distribution (the most basic statistical distribution). The equation quickly increases in complexity for more involved sampling designs and depends on the statistic for which precision is desired.

Another form of the same equation is computed by dividing the standard deviation and the numerical precision desired by the mean, thereby converting it to the coefficient of variation and the percent deviation from the mean. The revised equation takes the form:

$$n=z^2 \ C^2/d^2$$

where

C is the coefficient of variation, and

d is the desired precision as a proportion from the mean.

The second equation allows simpler assessments of precision. For example, to achieve an AADT estimate with 95 percent confidence and plus or minus 5 percent precision [95-5], the equation becomes: $n = 4 C^2 / (.05)^2 = 1600 C^2$. For the example in appendix B, with a C of .27 and day as the sampling unit, a simple random sample of 117 days would be needed to estimate the AADT with a precision of plus or minus 5 percent with 95 percent confidence. If the precision statement were relaxed to [95-10], the sample needed reduces to 30 days. This is much better, but still a long way off from the 48-hour or 24-hour monitoring periods which are the standard short counts on which most estimates of AADT are based.

The previous example makes it obvious that the traditional sampling methods are not directly applicable. A rigorous application of statistical sampling precision methods is not directly applicable because the traffic data are not collected based on a rigid statistical sampling method and most important because of the periodicity in the traffic data that distorts the normal distribution assumptions of the random sampling model.

Traffic data collected by continuous counters has no sampling limitations or sampling error since all possible data (with the exception of equipment shutdowns) is being collected. Simply stated, a 100 percent sample has no sampling error. Since the data is produced by continuous monitoring, the statistical parameters needed to establish sample size or to measure precision or variability can be easily computed and applied as shown in the previous examples. However, since continuous traffic monitoring data -represent a 100 percent sample and are subject to periodic variation, it may be more appropriately analyzed by using time-series methods.

Sweden's example

The application of rigorous statistical sampling precision statements to traffic volume data has been questioned due to the lack of a random sampling approach and the periodicity prevalent in traffic data which affects the assumption of normality in the sampling model.

In the review that preceded this paper, I have not found a single example of the actual development of statistical precision statements to short count traffic volume data in State programs. However, an example of the use of precision statements in traffic volume flow maps is shown by the flow maps produced by Sweden (sample attached in Appendix A) and described in a 1993 report'.

These maps show the traffic flow estimate for each highway link followed by a plus or minus percentage. The volume estimate represents Average Daily Traffic (ADT) and the percentage is the 95 percent confidence interval. The confidence interval is based on sets of short counts taken several times a year over several years.

In Sweden's 1990 traffic flow map the precision values range from a high of 51 percent to a low of 11 percent. As expected the lower variability is found in the higher volume links. The precision estimates can remain fairly constant along a route or vary by link, which seems to indicate that the computations were performed individually by link. The information provided by these precision statements allows users to generalize the quality of the estimates presented, to conduct sensitivity analysis assessing the effect of changes in volume due to traffic variability, or to determine statistical significance. The use of the precision statement also serves to alert the casual user to the dynamic nature of the traffic information presented.

The development of precision estimates based on short counts in the Sweden example leaves many concerns. The effect of variation caused by temporal variation (seasonal, monthly, daily, and hourly), weather, construction, maintenance, special events, traffic growth, or the many other factors which affect traffic are not addressed. The precision estimates presented do not apply to the specific year for which the map was developed, but rather are a general estimate from data collected over several years.

A comparison of the sample of days needed to achieve the precision listed in Sweden's flow maps and the random sample methods shown in the previous section can be used to estimate the samples used to compute the estimates.

The continuous counter in Appendix A has a C of 27 percent. Using the equation in the previous section, a random sample of two days would produce an estimate of plus or minus 50 percent with 95 percent confidence [95-50]. Likewise, an estimate of [95-10] requires a random sample of 30 days.

The following table presents a summary of the example and also the estimates if the value of C were lower (15 percent) as is usually the case in urban sites:

C	Days	Precision
.27	2	95-50
.27	8	95-20
.27	30	95-10
.27	117	95-5
.15	1	95-30
.15	3	95-20
.15	9	95-10
.15	36	95-5

Sweden's report does not provide the complete details of AADT estimation, whether temporal factors are used, whether continuous counters are used, or how their counting program is established. If Sweden's traffic counting program were based on a random number of daily counts each year, then the AADT and C at each site could be easily determined using the standard methods shown. However, it remains unclear how the effects of day-of-week and monthly variation are accounted for to compensate for the temporal variation in the AADT estimate.

Current traffic data precision auidelines

Traffic monitoring data consists of three basic components: traffic volume, vehicle classification, and truck weight. The estimation procedures and continuous data available for traffic volume are far more defined than for vehicle classification or truck weight. The discussion and application in this paper is limited to traffic volume.

The most common estimate derived from traffic volume data is Annual Average Daily Traffic (AADT). But, many other estimates are also developed from volume data. The Traffic Monitoring Guide[2] (TMG) identified AADT as the principal objective of traffic volume data collection and established the AADT target precision criteria as [95-10], plus or minus 10 percent with 95 percent confidence. This general precision criteria was based on historical or traditional criteria and it was used to support the Highway Performance Monitoring System (HPMS), which itself established sample size criteria for estimation of Vehicle Miles of Travel (VMT).

The reasoning behind the establishment of the AADT criteria was somewhat arbitrary and not derived from rigorous sampling theory analysis. It was based on an extensive deterministic analysis of the data', the data collection equipment, the data collect&

procedures and the cost. During the development of the TMG, it was quickly recognized that very tight precision levels could only be achieved by using a continuous counter. Cost is a primary consideration in traffic monitoring programs. Given the cost of continuous counters and the inability to place more than a few in each State, the traditional process of using short counts to estimate AADT remained the only feasible alternative.

Estimating precision levels tighter than the 10 percent criteria would require repetitive short counts at the same site, which involves a high cost. The result was the cost-effective recommendation of a traditional short count adjusted for seasonality based on a well defined temporal adjustment. process derived from the permanent traffic counters. The recommended short count period, 48 hours, was based on a comparison of the precision, equipment, and installation cost of several alternatives.

The application of the TMG-recommended 95-l 0 statistical precision criteria is difficult due to the natural variability inherent in traffic volume, the effect of random but frequent traffic events, the different traffic patterns caused by levels of urbanization, the extreme patterns encountered at many locations, and the many equipment related issues. It is generally accepted that a plus or minus 10 percent precision can be achieved using a seasonally adjusted 48-hour count in higher volume urban traffic [assuming that the appropriate equipment can be safely installed]. Indeed, several States-have found that the target criteria is achievable with 24-hour counts (the 1992 AASHTO Guidelines[3] recommended a 24-hour minimum in urban areas based on that finding as well as the equipment difficulties encountered in urban areas). The 95-10 precision criteria is more difficult to meet [based on 48-hour counts] in lower volume urban locations or in rural areas [although equipment installation is usually not a problem there], and may well be Impossible to meet in recreational locations due to the extreme variation patterns.

In the final analysis, the target precision criteria representtargets that identify the precision objectives, relate precision to cost, and allow future program assessment. Whether precision criteria are specified on a comprehensive basis [the TMG example] or separate precision targets are identified depending on other characteristics or needs, the purpose of specifying criteria remains the same, to provide a general indication of the quality of the estimate. As with all targets, the expectation of scoring a bull's eye each time is unrealistic. Specific cases where the precision criteria is not met do not indicate a failure of the data collection method, a lack of validity of the statistical measures, or reduce the value of the resulting information. They simply reflect the reality of traffic monitoring and the inability of any statistical tool to achieve success 100 percent of the time.

The first aspect of quality improvement, the input phase, involves defining precision targets to meet specified objectives before the data are collected. The second aspect, the output phase, is measuring precision after the data have been collected to determine whether the objectives of the data collection were achieved. It is the measurement of precision that allows objective evaluation of the data collection program. The evaluation aspect becomes critical because it indicates whether the objectives have been met and/or allows program improvement as techniques and equipment improve over time.

A simplified method for traffic monitoring precision estimation

Establishing precision criteria can be difficult in real-life situations. In theory, sampling precision criteria can be based on the desired objective irrespective of cost, or alternatively to maximize the precision achievable based on a fixed cost. Neither of these sampling theory options offers a clear solution for traffic monitoring data collection.

The definition of traffic monitoring objectives can get complicated due to the many uses of the data and the types of information needed. Costs can be quite intractable to specify due to complex budgeting mechanisms, the equipment involved, the operational difficulties with traffic monitoring, resource allocations based on need, personnel, organizational responsibilities, and the wide variety of ways in which costs are assigned or allocated by different organizations.

Estimates of AADT are computed directly from continuous counters or estimated based on short counts adjusted by factors. Estimation from continuous counter is fairly straightforward. Recent unpublished work[4] by Oak Ridge National Laboratory and also presented in this conference show that computational method or missing data have negligible effect on the AADT estimates and their variability. The finding also extends to vehicle classification estimates. This is a very important finding that will greatly simplify the computation of AADT and variability estimates. These findings also increase the credibility of traffic estimates by highlighting their statistical robustness.

As previously indicated, estimates of daily traffic variability can be directly computed at continuous counter sites. The estimate of AADT from a continuous counter has no sampling error since it was based on all available data. Of course, data collection equipment, data processing, and missing data still remain sources of non-sampling error.

At short count sites where data is collected for a 24-hour or 48-hour period, the direct estimation of variability is not possible due to the fact that a random sample of short counts is not used to estimate AADT at the short count site. At portable sites, a single short count is used to estimate AADT based on the application of seasonal factors (day and month) that compensate for the periodic variability that affects traffic. The AADT estimation method, using daily and monthly factors to compensate for the temporal variability, provides a highly precise method of estimation whose statistical precision is not adequately reflected in the standard simple random sampling approach. The computation and application of seasonal factors to estimate AADT is detailed in the TMG and the AASHTO Guidelines.

The proposed approach to estimate the precision, or more precisely the error, of the estimate of AADT makes inferences based on comparisons from the continuous counters. This is done by using the same inference base from which the temporal adjustment factors for the estimation of AADT are derived. The final result is an empirical precision estimate for the AADT estimate derived from the short count and the applicable seasonal factors based on a comparison at the continuous counter base from which the factors were derived.

The proposed approach makes a critical assumption, that the same continuous sites used to develop the seasonal factors and determined to be appropriate to estimate the AADT are also appropriate to estimate the precision of the AADT estimate. This is not a farfetched assumption. It is well established that the variability in traffic is compensated for by the use of day and monthly factors and that the variability patterns can be applied consistently across highway systems or other groups of continuous counters- All that the proposed approach does is to extend the use of the information extracted from the continuous counters to estimate not only the AADT but also the precision of the AADT estimate.

The procedure and the results are demonstrated by example. **Appendix B** provides a list of daily traffic at a continuous traffic volume station in 1993. The table provides the date, day of week, daily volume, and identifies the national holidays.

Appendix C consists of two pages. The first page provides a list of statistics derived directly from the data in Appendix B. The statistics include the AADT (16,729), the daily standard deviation (4.479), the coefficient of variation (27 percent), a table showing the twelve monthly factors, and a table showing the day of week factors for each month.

The AADT **is** computed as the straight average of the 365 days. The standard deviation is the simple random sample statistic. The coefficient of variation is the ratio of the standard deviation to the mean (AADT) shown as a percentage. The monthly factors are computed using standard TMG procedures as the ratio of the AADT to the monthly average daily traffic (MADT) The daily factors are the ratio of the MADT to the average of the specific day of week within the month. The computations include all days of the week including holidays and influence days.

The second page in Appendix C shows the computations to estimate the error of the

AADT estimate derived from a 24-hour count on May 5, and the AADT estimate from a 48-hour count on May 5 and May 6. The tables were computed using Lotus 123 Version 5.

Single-day error estimation

Using the data in Appendices B and C, an AADT estimate and its empirical error (which I refer to as estimated precision) will be computed using a 24-hour monitoring period and the standard factorization approach based on day of week and monthly factors to compensate for temporal variation. The selected day is May 5, which in 1993 was on a Wednesday. The following table presents the AADT estimating procedure and the error of the estimate.

Date	Day	Volume	Day Factor	Month Factor	AADT Estimate	True AADT	Error %
May 5	Wednesday	13,661	1.08	1.05	15,492	16,729	-7.4

This information, extracted from the continuous counter in Appendix B, shows that the AADT estimate computed using the standard factoring procedure from the May 5 count underestimates AADT by 7.4 percent. This is not a precision estimate. This is the actual error based on complete information at this site. Given complete continuous data, there-is no need for confidence intervals to compensate for non-existent sampling error or for the theoretical assumptions that support sampling methods.

The next step is how to infer the information from this actual process and station to estimates based on short counts taken at other sites. For this we need a generalized inferential process since we lack temporal information at the short count sites.

The generalized error inferential process recommended in this paper is to use the highest absolute error identified, based on the same day of the week and month as the short count, at the continuous station or stations from which the day and month factors used to make the AADT estimate were taken. If the factors used to expand the short count came from a single site, then the process is shown in Appendix C. If the factors came from a group of sites then, the Appendix C method would be computed at each of the group sites and the average of the highest absolute errors at each site used as the generalized estimate.

The procedure will be described in detail by example. If a 24-hour count is taken on Wednesday, May 5 and the factors from the station in Appendix B are used to expand the short count to AADT, then the errors for each Wednesday in May are computed in as shown in the second page of Appendix C. The highest error identified is an underestimate of 7.4 percent, which can be rounded up to 8 percent. Interestingly, the highest error detected for the Wednesdays in May occurred on May 5, the date of the count. The error of the AADT estimated from the 24-hour count is the absolute value of the highest error found and, in this case, is plus or minus 8 percent.

Any short counts taken on Wednesdays in May 1993 where the AADT was estimated using the factors from the continuous counter in Appendix B will have an estimate of error of plus or minus 8 percent. Although no analysis of year effects has been done, it is recommended that the proposed method be applied only to current year counts using current year factors computed after analysis of the continuous data for the full year.

As can be seen from the table in Appendix C, proposed method results in a very conservative estimate, the highest absolute error found. The AADT estimates from other Wednesdays in May show much lower error. Other less conservative estimates could be developed. flowever, the most conservative estimate is recommended because the inference is made from the continuous counter to a completely different site where the short count is taken and where no information on temporal variation is available. As other studies[2, 6] have found, the correct application of factors is one of the most important aspects of traffic estimation, far more important than the length of the count.

If instead of a single counter, the factors came from a group of IO continuous stations, the procedure in Appendix C would be applied to each of the IO continuous sites using the group factors. Then the average of the absolute errors at each of the 10 sites would be used as the empirical error estimate equivalent to a precision estimate. The average absolute highest error would then become the error applied to any AADT estimate based on the group factors for the appropriate day and month..

Since the individual site and group factors are known (and used frequently to estimate AADT at the short counts), tables of estimates can be easily computed using any spreadsheet package for each day of the week for each month for each continuous counter or group of counters. The error tables would be similar to the day of week factor table shown in Appendix B, providing an easy to apply error estimate for each day of the week for each month. A decision to apply factors to estimate AADT based on a short count would immediately point to a similar table providing the empirical error of the AADT estimate. The process provides both easy computation of error estimates from continuous counters and simple direct application to the short counts.

Two-day error estimation

The second page of Appendix C also presents a similar procedure for a 48-hour monitoring period. The only difference being that the process consists of two days and the AADT is estimated as the average of the two days. The error estimate is based on exactly the same procedure once the 48-hour AADT is estimated.

The procedure is shown by an actual example using Wednesday, May 5 and Thursday, May 6 and also result in an empiric error estimate of plus or minus 8 percent. The procedure is essentially the same except that instead of using 4 daily periods to estimate AADT for the single-day example, we now have 8 periods for the two-day example. Once the 48-hour AADT estimates are produced, the procedure remains the same.

The empiric error estimate for the May 5 & 6 (48-hour) count using the procedure advanced in this paper is plus or minus 8 percent estimated as the largest absolute difference detected rounded up to the next percent. If the seasonal factors came from a group of stations, then the estimate would be computed in the same manner as indicated for the single-day period as the average of the highest absolute error at the group sites.

The application of empirical error for a 48-hour count would be based on the day of week of the first day of the count for the specific month. In this example, 48-hour counts starting on a Wednesday in May 1993 where the factors came from the counter in Appendix B would show an error of plus or minus 8 percent. As before, if group factors were used, the procedure would be applied to each of the continuous counters in the group and the average absolute highest error would become the empirical error estimate.

By comparing the one and two day procedures, an interesting result is highlighted. The error estimate for the 48-hour count consisting of May 5 and 6 is -7.84 percent, higher than for the May 5 count, - 7.4 percent. Since both are rounded to the next highest percent, both AADT estimates result in errors of plus or minus 8 percent.

This finding begs discussion and explanation. The equal error result is not a mistake, it simply reflects the reality of the comparison between periods of 48 and 24 hours using a single day approach. As can be seen in Appendix C, the AADT estimate from May 5 is 15,492. The same estimate from May 6 is 15,344. These estimates are very close. The estimate from the 48-hour period is the average of the two or 15,418. By definition, the average of two numbers is going to be between the two, lower than one and higher than the other. If we had compared the 48-hour estimate to May 6 rather than May 5, then the 48-hour error estimate would have been slightly lower.

On the average and over large samples, a 48-hour count will always'be more precise than a 24-hour count, but not by a significant amount. This is supported by sampling theory, since a larger sample is, on the average, more precise than a smaller sample. In the application of the real-life AADT temporal estimation procedure, the real gain of the 48-hour period occurs in cases where the separate estimates from each of the two days err in opposite directions.

For example, if the first 24 hour period produces a 5% underestimate and the second day produces a 5% percent overestimate, then the average of the two days, the 48-hour period AADT, results in a perfect estimate with zero percent error and is far superior to both 24-hour periods. This is, of course, a contrived example which will not occur often. In general, when the average error of 48-hour AADT based on the average of two days AADT is compared with a single 24-hour count, the 48-hour will show less error of a magnitude of 1 to 2 percent. If the AADT estimates from the two days err in the same direction then the AADT estimate from the 48-hour period will always be better than one day and worse than the other. if the two days estimates err in opposite direction, then the 48-hour estimate may produce an estimate which is better than each of the two days.

A reflection of the fact that 48-hour counts will be more precise than 24-hour counts is shown by comparing the average error of the single versus two-day periods from the table in Appendix C. In this actual example, the single day average error is 0.26 while the two-day average error is 0.06. For a while, I contemplated using the average error as the empirical estimate of precision. However, as can be seen from the examples presented, the average errors get very small and reflect the averaging effect in opposite directions.

The average error presents a very misleading implication of very low error from an AADT estimate. Experienced data collectors and users of traffic data recognize the many sources of error that often affect traffic estimates. However, the low average errors are also a reflection of the high accuracy that can be achieved with the temporal factorization based on day and monthly factors- Given the extreme variability of traffic, I opted for the conservative approach using the absolute highest error detected, which I believe is far more realistic.

A comparison of the proposed methodology to Sweden3 example can not be shown-until the proposed approach is applied and tested in the real world. However, by examining the weaknesses of Sweden's approach and the errors presented, it seems apparent that the AADT temporal factorization method as recommended in the TMG will result in more precise AADT estimates.

<u>Summary and conclusions</u>

Estimating the precision of traffic data is another step to improve the quality of traffic data and enhance its decision-making potential. This paper has examined the theory of simple random sampling and precision and highlighted its lack of applicability to real traffic monitoring data as exemplified by the AADT estimation process based 'on temporal factors. The theory has been followed by an actual application from Sweden's 1990 traffic flow map. Finally, an empirical procedure for estimating the error of AADT estimates as it applies to real traffic data is proposed and demonstrated. Use of the simple procedure described will provide an adequate estimation of AADT error and highlight the quality of the information presented. It will also serve to increase the comfort level of both data collectors and users of traffic data.

EFFECT OF TRAFFIC DETECTION ERRORS ON THE ACCURACY OF VOLUME REPORTS

Speaker: Dario Hidalgo
Ohio State University

Authors: Dario Hidalgo, et al.
Ohio State University

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9,1996

ABSTRACT

We consider an approach to modeling traffic detection errors that has impacts on procedures used to adjust for systematic bias and on design of tests to estimate correction parameters for these procedures. In this approach, we consider three kinds of detection errors: not counting existing vehicles, counting existing vehicles more than once, and counting non-existing vehicles. We then develop a stochastic model of the detection process and derive a formula to adjust observed traffic for systematic bias. The formula uses correction parameters estimated through a test in which the observed traffic is compared to the actual traffic in a disaggregated (i.e., vehicle-by-vehicle) fashion.

We conducted a simulated, numerical study that allows a comparison of the accuracy of this procedure with that of the currently recommended procedure. We compare the two procedures as *a* function of: traffic volumes used to estimate parameters of the adjustment formulas; traffic volumes that are to be monitored; and levels of the various error probabilities. The results show that the approach developed in this paper provides better average estimates of traffic volumes in almost all cases. The current adjustment procedure provides equal average estimates only when the traffic volumes used to estimate the parameters of the adjustment formula are the same as the traffic volumes to be monitored, an unlikely situation in practice.

The results imply that an approach that considers different types of errors separately might be preferred to the currently recommended procedure. This type of approach would require dissagregate tests, however, which are more difficult and expensive to perform than the aggregated tests currently recommended. The advantages of the improved accuracy, compared to the additional cost and complexity of the tests, would need to be evaluated.

INTRODUCTION

Highway traffic data provides basic information for design, maintenance, and management of highway systems (I). Hence, errors in this data may have important impacts on the final uses. Despite the fact that little is known about the size of these impacts (2), there is a perceived demand for quality.

Traffic monitoring data is prone to various sources of error, namely, sampling, parameter estimation, special conditions and equipment errors (I). Sampling errors are the result of the faulty selection of the number, location and duration of traffic counts. Parameter estimation errors follow incorrect determination of factors used to expand short term counts to estimate annual average daily traffic. Special conditions errors are the result of certain events, such as bad weather or construction, during the traffic studies. Finally, equipment errors are either the consequence of equipment malfunction or detection errors. In the Traffic Monitoring Guide (I), the federal government recommends procedures to minimize the sampling and parameter estimation errors while keeping costs at tolerable levels. This guide also suggests that practitioners exert judgment upon the occurrence of special conditions or equipment malfunction.

In addition, the guide emphasizes the need to correct traffic counts for systematic bias resulting from equipment detection errors and suggests an adjustment procedure. The adjustment procedure is based on the assumption that the ratio of number of vehicles "observed" by the equipment to the "actual" traffic volume is independent of the actual traffic volume. The guide recommends that this ratio be obtained using a test in which the number of vehicles observed by the equipment and the actual number of vehicles are recorded. To obtain the actual number of vehicles, or "ground-truth", the guide recommends using manual counts or a calibrated traffic monitoring device. The ratio is then calculated by dividing the number of observed vehicles by the ground-truth. The resulting ratio is used to estimate actual traffic volumes in monitoring studies. The estimated actual traffic volume in a given monitoring study is equal to the number of vehicles observed by the equipment divided by the ratio obtained in the test.

In this paper, we approach equipment detection errors differently. We propose that there are different types of detection errors. We also consider that these errors are probabilistic. This approach has impacts on the procedures used to adjust for bias and on the design of equipment accuracy tests. We develop a simple model based on this view in the next section. In section three, we derive a correction formula based on the model and present the formula for the currently recommended adjustment procedure. In section four, we conduct a simulated, numerical comparison of the accuracy of the estimates of traffic using the correction formulas. The results of this comparison indicate that distinguishing among different types of errors would provide better estimates of actual traffic volumes. We discuss our results and their implications in the final section.

DETECTION PROCESS MODEL

To model the vehicle detection process we distinguish among three types of errors. We call them "error of omission", "error of commission type A", and "error of commission type B". An error of omission is an error such that a vehicle that is present is not counted. An error of commission type A is an error such that a vehicle that is present is counted more than once. We assume in this paper that it is counted as two vehicles. An error of commission type B is an error such that a vehicle count is obtained when no vehicle is present.

With this framework, the number of vehicles (N_O) that would be observed by the detection device for a given time period would be:

$$N_O=N_R-N_M+N_A+N_B, \qquad (1)$$

where N_R is the actual (real) traffic volume during the time period; N_M is the number of errors of omission committed during the time period; N_A is the number of errors of commission type A committed during the time period; and N_B is the number of errors of commission type B committed during the time period.

We model the detection of a vehicle that is present as a discrete process with probability p of not counting the vehicle, probability q of counting the vehicle twice, and probability (1 -p-q) of counting the vehicle once. Hence, the number of errors of omission (N_M) would be a Bernoulli random variable with expectation and variance given by

$$E[N_M] = N_R*p\ ; \qquad (2)$$
$$Var[N_M]=N_R*p*(1-p)\ . \qquad (3)$$

Similarly, the number of errors of commission type A (N_A) would be a Bernoulli random variable with expectation and variance given by

$$E[N_A] = N_R*q\ ; \qquad (4)$$
$$Var[N_A]=N_R*q*(1-q) \qquad (5)$$

As a first approximation, we model the number of errors of commission type B as a Poisson variable with parameter λ. This approach assumes that the number of errors of commission type B are independent of the actual traffic volume and dependent on the length of the observation period only. In an actual situation this type of errors might be a function of the traffic in other lanes (3) or objects different than vehicles (4). The expected value of the number of false alarms is given by

$$E[N_B] = \lambda\ . \qquad (6)$$

CORRECTION FORMULAS

This approach to the problem leads to a different formula for estimating actual traffic from observed traffic from that used in the currently recommended adjustment procedure. We first derive a correction formula based on the model discussed in the previous section. Then, we present the formula for the currently recommended adjustment procedure.

We derive the correction formula based on the expected number of observed vehicles ($E[N_O]$). Taking the expectation of(l), $E[N_O]$ is found to be

$$E[N_O] = N_R - E[N_M] + E[N_A] + E[N_B] \tag{7}$$

Replacing the corresponding expected values on the right hand side by equations (2), (4) and (6), respectively, we write

$$E[N_O] = N_R - N_R{*}p + N_R{*}q + \lambda\ . \tag{8}$$

Solving for N_R, we have

$$N_R = \frac{E[N_O] - \lambda}{(1-p+q)} \tag{9}$$

The parameters p, q and λ are not observable. However, they could be estimated from a test. Unbiased estimators of p, q and λ would be

$$\hat{p} = \frac{N_M^e}{N_R^e} \tag{10}$$

$$\hat{q} = \frac{N_A^e}{N_R^e} \tag{11}$$

$$\hat{\lambda} = N_B^e \tag{12}$$

where N_M^e is the number of errors of omission in the test, N_R^e is the actual traffic volume in the test (ground-truth), N_D^e is the number of error of commission type A in the test, and N_B^e is the number of errors of commission type B in the test. The number of errors of omission, commission type A and commission type B can be obtained by vehicle-by-vehicle comparison of the observations using the equipment being tested and the ground-truth.

To estimate the actual traffic volume in a monitoring study from an observed number of vehicles (N_O^a) we replace p, q and λ by the estimators obtained through equations (10), (11), and (12), respectively, and $E[N_O]$ by N_O^a in equation (9). This leads to

$$\hat{N}_{R1} = \frac{N_O^a - \hat{\lambda}}{(1-\hat{p}+\hat{q})} . \qquad (13)$$

Given that (13) is based on a disaggregate (i.e. vehicle-by-vehicle) test, we refer to it as “dissagregate-based formula”.

We want to compare the results of adjusting the traffic using the dissagregate-based formula with the adjustment procedure suggested in the Traffic Monitoring Guide (I). In the suggested procedure, the estimated traffic, which we shall call $\hat{N}_{R2}$, is the result of dividing the number of vehicles observed by the equipment in a traffic monitoring study (N_O^a) by a correction factor ($\hat{F}$):

$$\hat{N}_{R2} = \frac{N_O^a}{\hat{F}} , \qquad (14)$$

where $\hat{F}$ is the ratio of the number of observed vehicles in the test (N_O^e) to the actual traffic volume in the test (N_R^e), as described in section 1. Given that this formula is based on the total actual and observed traffic volumes, we call this an “aggregate-based formula”.

NUMERICAL COMPARISON

To compare the performance of the correction formulas presented in the previous section, we consider the difference between the estimated traffic volume ($\hat{N}_{R1}$ or $\hat{N}_{R2}$) and the actual traffic volume in a monitoring study (N_R^a), divided by the actual traffic volume in the study. For the dissagregate-based formula the performance measure, or relative error (E_1), is given by

$$E_1 = \frac{\hat{N}_{R1} - N_R^a}{N_R^a} . \qquad \mathbf{(15)}$$

Similarly, for the aggregate-based formula, the performance measure, or relative error (E_2), is given by

$$E_2 = \frac{\hat{N}_{R2} - N_R^a}{N_R^a} \qquad (16)$$

We are also interested in comparing the magnitude of these relative errors. We do this by calculating the ratio of the size of the two relative errors (RE)

$$RE=\frac{E_2}{E_1}. \qquad (17)$$

The relative errors (E_1 and E_2) and the ratio of the size of the relative errors (RE) are random variables. Their probability distributions are difficult to obtain. Therefore, we approximate these distributions using Monte Carlo simulation. We calculate the mean and the standard deviation of the relative errors (E_1 and E_2) and the mean (RE) We perform the simulations for different sets of initial values, namely, different values of p, q and λ, and different levels of the actual traffic volume in a test (N_R^e) and actual traffic volume in a monitoring study (N_R^a).

The simulation procedure follows the actual process for estimating correction parameters using a test and adjusting the number of observed vehicles in a traffic monitoring study. The procedure consists of the following steps: i) simulate the outcomes of an accuracy test (i.e., N_M^e, N_A^e, N_B^e, and N_O^e); ii) estimate the correction parameters for each adjustment formula ($\hat{p}$, $\hat{q}$, and $\hat{\lambda}$ for the dissagregate-based formula, and $\hat{F}$ for the aggregate-based formula); iii) simulate the outcome of a traffic monitoring study (N_O^a); iv) we estimate the actual traffic using each adjustment formula ($\hat{N}_{R1}$ and $\hat{N}_{R2}$) with the parameters obtained in step ii); and v) calculate performance measures (E_1, E_2, and RE) using $\hat{N}_{R1}$ and $\hat{N}_{R2}$ from step iv).

To simulate the outcomes of an accuracy test in step i), we randomly sample a number of errors of omission (N_M^e), a number of errors of commission type A (N_A^e), and a number of errors of commission type B (N_B^e) from the probability distributions of each type of error. We then use the sampled values to calculate a number of observed vehicles for the test (N_O^e) using equation (1). To obtain the correction parameters $\hat{p}$, $\hat{q}$, and $\hat{\lambda}$ in step ii), we use equations (l0), (11) and (12), respectively. To get ($\hat{F}$) we divide N_O^e by N_R^e. We repeat the sampling procedure described in step i) to obtain the components of equation (1) and calculate the outcome of a traffic monitoring study (N_O^a), in step iii). In step iv) we apply the adjustment formulas --equations (13) and (14) for disaggregate and aggregate adjustment, respectively. Finally, in step v) we apply equations (15), (16) and (17) to calculate the relative errors E_1 and E_2 and the ratio RE.

We replicate the five simulation steps 3,000 times. In each replication the numbers N_M^e, N_A^e, N_B^e, N_M^a, N_A^a, and N_B^a are randomly selected. Based on the-3,000 replications, we calculate the averages and standard deviations of E_1 and E_2 and the average of RE.

We develop the simulation-based comparison for different scenarios resulting from the combination of values shown in Table 1. We use three levels of actual traffic volume in the test (N_R^e); five levels of actual traffic volume in a monitoring study (N_R^a); and two levels of the probability of error of omission (p), error of commission (q) and expected false alarms (A.).

RESULTS OF THE COMPARISON

In table 2 we present the averages E_1 and E_2 for different values of N_R^e and N_R^a 9 two levels of λ (λ=1,000 and λ=100) and fixed p and q (p=0.02 and q=0.01). We note that the size of the averages of the relative errors E_2 (columns 4 and 6) are greater than the averages of the relative error E_1 (columns 3 and 5) for most of the traffic volume pairs considered. When the traffic volume in the test is equal to the traffic volume in the monitoring study (i.e., $N_R^e = N_R^a$), the size of the relative errors is the same, with the exception of $N_R^e = N_R^a$ = 5,000 and λ=1,000. In this case, E_2 is slightly greater than E_1.

Moreover, as indicated in Figure 1, we observe that for both values of λ, the average of E_2 is less than zero when the actual traffic volume levels in the test is less than the actual traffic volume in the monitoring study ($N_R^e < N_R^a$), and greater or equal to zero when $N_R^e \geq N_R^a$. This indicates that the aggregate-based formula is expected to either underestimate or overestimate traffic when the actual traffic volume in the test is different than actual traffic volume in the monitoring study.

In Figure 2 we present the average ratio of the size of the relative errors as a function of traffic in the monitoring study (N_R^a), for three levels of traffic in the test (N_R^e) and two levels of λ. We observe that for both λ values, the averages of RE are increasing functions of the size of the difference between N_R^e and N_R^a. When this difference is zero (i.e., $N_R^e = N_R^a$), the average of RE is less than one for λ=1,000 and equal to one for λ=100. This indicates that when the levels of traffic in the test and the monitoring study are the same, the disaggregrate-based formula cases is expected to yield errors at least as large as the aggregate-based formula. Nevertheless, having exactly the same traffic in the test and in the monitoring study is very unlikely.

We also performed the simulations and calculated summary statistics for different values of the probability of error of omission (p) and probability of error of commission (q) In Table 3 we present the averages of the relative error with the disaggregate estimate for N_R^e =5,000, λ=1 ,000, two combinations of p and q, and five levels of traffic in the monitoring study (N_R^a). We observe similar results than those obtained with p=0.02 and q=0.01. In the new cases, in which p=0.10 and q=0.01, and p=0.02 and q=0.05, the sizes of the averages of the relative error E_2 are also greater or equal than the averages of the relative error E_1 . This indicates that the observed results do not depend on the values of the probabilities of error of omission (p) and error of commission type A (q)/

DISCUSSION

The estimates of actual traffic volume in a monitoring study using a formula based on disaggregate (i.e., vehicle-by-vehicle) comparison of actual and observed traffic in a test are better than those using a formula based on aggregate comparison. In the unlikely case in which the traffic volumes in the test are the same as that traffic volumes in the monitoring study, the aggregated approach provides equal or better estimates than the disaggregated approach.

As a result, it would be recommended to use a correction formula based on disaggregated comparison of the observed and actual traffic in a test. Obtaining the parameters of the dissagregate-based formula would be more complicated, however. Specifically, it would be necessary to discern among existing vehicles that are not counted (errors of omission), existing vehicles that are counted more that once (errors of commission type A) and non-existing vehicles that are counted (errors of commission type B).

It would be interesting in future studies to evaluate the advantages of improving accuracy compared to the additional cost and complexity of the required tests. This comparison might lead, for example, to special test designs in which only a subset of the accuracy test is compared at the disaggregate level. This would lead to worse estimates of the correction parameters, but less costs than performing full disaggregate test. A result of the comparison of advantages and disadvantages of alternative test designs could be the selection of a test design that would maximize accuracy while keeping cost and complexity at tolerable levels.

Our results are based on a simple model of the detection process. This model is based on the assumption that the probabilities of committing different types of errors are the same in the monitoring study and the test. We also assume that when multiple counting errors (errors of commission type A) are committed, only two vehicles are reported. Additionally, we assume that the number of non-existent vehicles reported (errors of commission type B) has a Poisson distribution. Nevertheless, we expect that our main results would hold for a more realistic model of the process. Further simulation tests with more complex model specifications would be required to confirm this belief. For instance, it would be interesting to use a model in which errors of commission could result in three or more vehicle reports, or a model in which the distribution of the non-existent vehicles is associated with the level of traffic in other lanes. Still, it appears that looking at errors from a more dissagregate perspective could lead to greater accuracy.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Ohio State University Civil and Environmental Engineering Department and the OSU Center for Intelligent Transportation Research. The interpretations are those of the authors.

REFERENCES

(1) US Department of Transportation - Federal Highway Administration. "Traffic Monitoring Guide" Office of Highway Information Management, Publication No. FHWA-PL-92-O 17, October 1992.

(2) Davies, Peter and Salter, David R. *Reliability of Classified Traffic Count Data,* Transportation Research Record 905, Transportation Research Board, 1983, p17-27.

(3) Duckworth, G.L.; Bing, J.; Carlson, S.H.; Frey, M.L.; Hamilton, M.A.; Heine, J.C., Milligan, S.D.; Mlawski, R.; Ritter, S.; Warner, L.R.; Watters, B.G.; Welsh, R.H. and Whittemore, D. *A Comparative Study of Traffic Monitoring Sensors.* Proceedings, National Traffic Data Acquisition Conference, Rocky Hill, Connecticut, September, 1994, pp. 113-122.

(4) Merry, Carolyn J. and McCord, Mark. The *Feasibility of Traffic Data Collection Using Satellite Imagery.* Proceedings, National Traffic Data Acquisition Conference, Rocky Hill, Connecticut, September, 1994, pp. 477-483.

LIST OF TABLES

LIST OF FIGURES

Table 1. Values of the Input Variables Used to Construct Simulation Scenarios

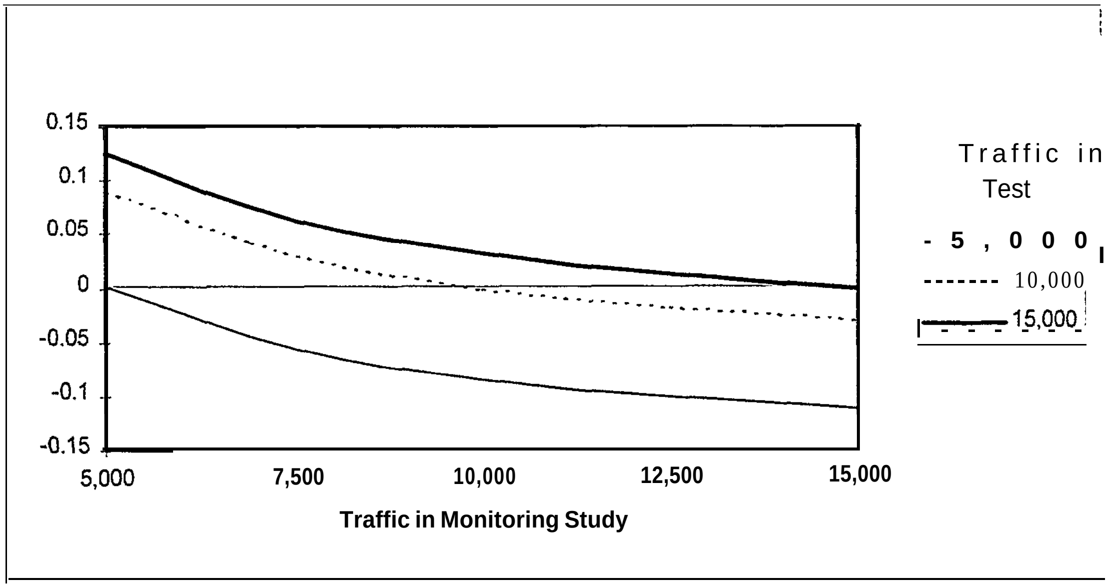
Input Variable	Selected Values
Real traffic volume in the test (N_R^e)	5,000; 10,000; 15,000
Real traffic volume in the monitoring study (N_R^a)	5,000 ; 7,500; 10,000; 12,500; 15,000
Probability Error of Omission (p)	0.02; 0.10
Probability Error of Commission (q)	0.01; 0.05
Expected number of error of omission type B (h)	100; 1,000

Table 2. Results of the Simulation for Different Values of N_R^e and N_R^a (p=0.02, q=O.Ol)

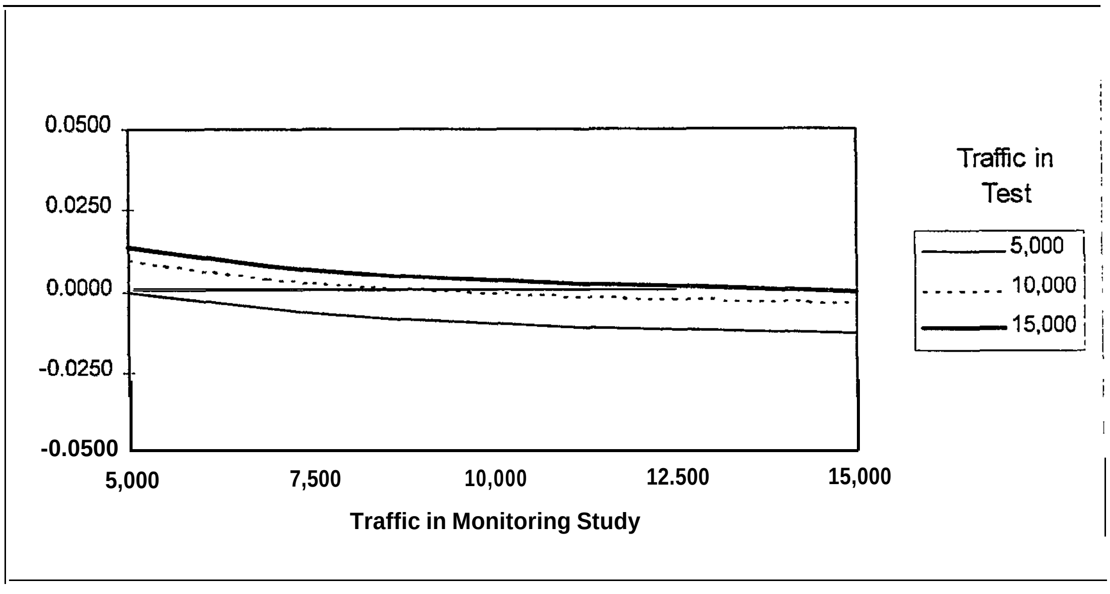
		λ=1,000		λ=100	
N_R^e	N_R^a	Average E_1	Average E_2	Average E_1	Average E_2
5,000	5,000	0.0000	0.0001	0.0000	0.0000
5,000	7,500	0.0000	-0.0559	-0.0001	-0.0067
5,000	10,000	0.0000	-0.0841	0.0000	-0.0098
5,000	12,500	0.0000	-0.1008	0.0001	-0.0118
5,000	15,000	0.0000	-0.1119	-0.0001	-0.0133
10,000	5,000	-0.0001	0.0916	-0.0002	0.0098
10,000	7,500	-0.0001	0.0305	0.0000	0.0030
10,000	10,000	0.0000	0.0000	0.0000	0.0000
10,000	12,500	0.0000	-0.0183	0.0000	-0.0020
10,000	15,000	0.0000	-0.0305	0.0000	-0.0033
15,000	5,000	-0.0001	0.1260	0.0001	0.0134
15,000	7,500	0.0000	0.0630	-0.0001	0.0066
15,000	10,000	0.0000	0.0315	-0.0001	0.0033
15,000	12,500	0.0001	0.0126	0.0001	0.0014
15,000	15,000	0.0000	0.0000	0.0000	0.0000

Table 3. Results of the Simulation of E_1 for Different Values of N_R^a and Two Combinations of p and q (N_R^e =5,000 and λ=1,000)

N_R^a	p=O.l, q=O.Ol		p=0.02, q=0.05	
	Average E_1	Average E_2	Average E_1	Average E_2
5000	**0.0000**	-0.000 1	0.0000	0.0000
7500	**0.0002**	**-0.0602**	**-0.0002**	**-0.0542**
10000	0.0000	-0.0900	0.0001	-0.0812
12500	-0.0001	-0.1081	-0.000 1	-0.0973
15000	0.0001	-0.1201	0.0000	-0.1083

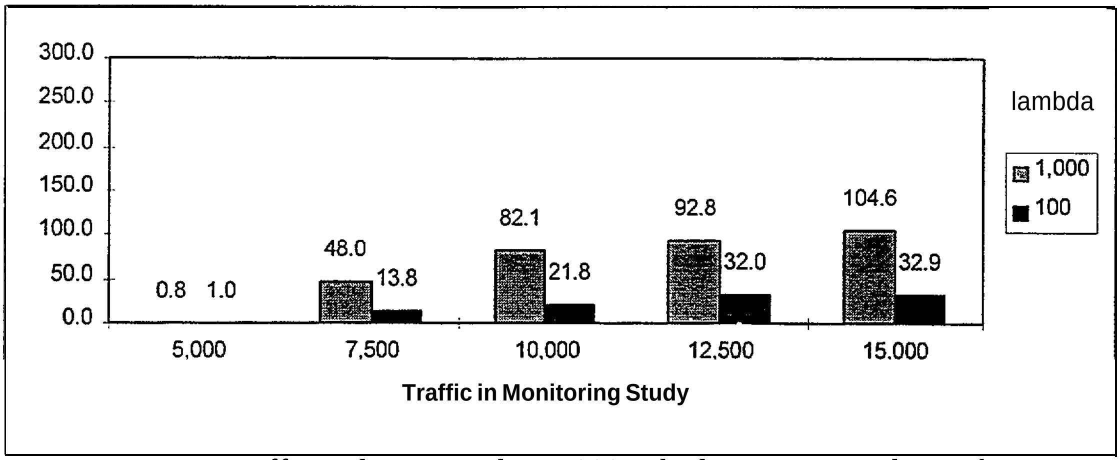


a. p=0.02, q=0.01, and λ=1,000

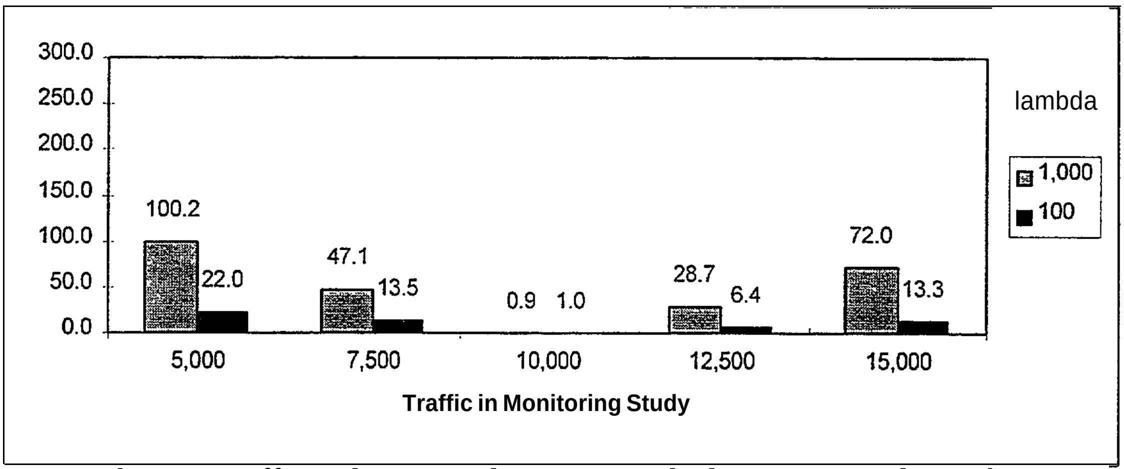


b. p=0.02, q=0.01 , and λ=100

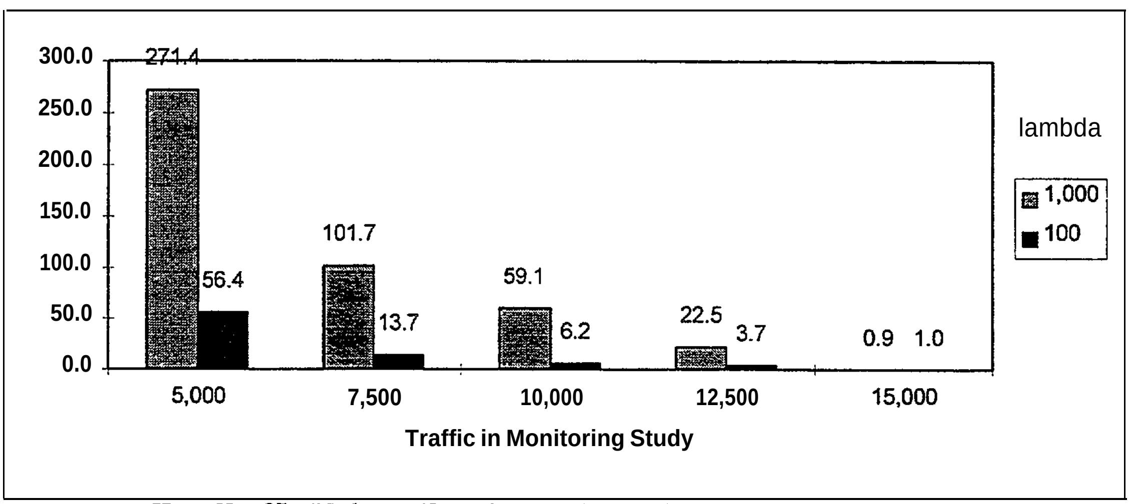
Figure 1. Average Relative Error of Aggregate Approach (E_2) as a Function of Traffic in a Monitoring Study for Different Levels of Traffic in a Test



a. Test Traffic Volume Equal to 5,000 vehicles, p=0.02, and q=0.01



b. Test Traffic Volume Equal to 10,000 vehicles, p=0.02, and q=0.01



c. Test Traffic Volume Equal to 15,000 vehicles, p=0.02, and q=0.01

Figure 2. Average Relative Error Size Ratio as a Function of Traffic in Monitoring Study and λ, for Different Levels of Traffic in Test

ALTERNATIVE CRITERIA FOR SELECTING SHORT TERM CLASSIFICATION COUNTING PERIODS

Barbara K. Ostrom, P.E.
EBA Engineering, Inc.

Presented at
National Traffic Data Acquisition Conference
Albuquerque, New Mexico

May 5-9, 1996

For further information contact:
Barbara K. Ostrom, P.E.
EBA Engineering, Inc.
6166 Leesbury Pike, B302
Falls Church, VA 22044
Ph: (703) 2852514

ALTERNATIVE CRITERIA FOR SELECTING SHORT TERM CLASSIFICATION COUNTING PERIODS

BARBARA K. OSTROM, P.E.

EBA ENGINEERING, INC.

Classification counts are used for a variety of purposes: HPMS estimates of truck percentages, cumulative loads for pavement management systems and site specific pavement design. Resources can constrain data collection to short counts which must then be factored to annual estimates. This research is two-fold: estimating general expansion factors and evaluating several short term count periods in terms of when they should occur and how to define when using site characteristics.

This research did not use the AASHTO expansion methodology to obtain annual estimates. Instead a set of factors was developed that would allow direct factoring of a given count period to an annual estimate. Factors were developed for four categories of facilities: urban freeways, rural freeways, urban other routes and rural other routes. Three counting periods were investigated: 48 hours, 12 hours and 6 hours.

Reflecting the range of uses for heavy vehicle data annual estimates of total trucks, Class 9s, Class 5s and Class 4s were developed. Since the data came for the Long Term Pavement Performance database the estimates pertained to the outside lane only. Only limited data would have been available had estimates of the various annual values for one or both directions been specified.

The second phase of the research was to determine when these short counts should be taken to develop the "best" estimates and what site characteristics could be used to describe when. A "best" estimate was defined to be one that was within plus or minus 10 percent of the actual value after factoring. Site characteristics considered included functional class, number of lanes, freezing index and precipitation. Definitions of when included season, month and day of week