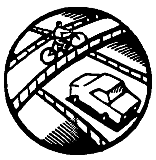
Sources of Exposure Data for Safety Analysis

PUBLICATION NO. FHWA-RD-97-025

NOVEMBER 1997

US. Department of Transportation
Federal Highway Administration

Research and Development
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296



FOREWORD

This report was prepared for the Federal Highway Administration (FHWA) as part of a contract to evaluate sources of exposure data for highway safety research. Several existing and emerging exposure data sources were subsequently selected by FHWA for review.

This report provides highway safety researchers with information to assess the feasibility of using exposure data sources in designing highway safety evaluation studies. One-page summaries are provided for each exposure data source. A longer description covers the purpose of the collection, contents, period covered, sample design, data collection methods, sample size, data quality, data format, possible cautions in using the exposure data, and availability of the data.

Copies of this report are available for a nominal charge from the National Technical Information Service (NTIS), Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

A. George Ostensen, Director
Office of Safety and Traffic Operations
Research and Development

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-RD-97-025		

4. Title and Subtitle	5. Report Date
SOURCES OF EXPOSURE DATA FOR SAFETY ANALYSIS	November 1997
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Campbell, K.L, Joksch, H.C., Blower, D., Kostyniuk, L.P., Pendleton, O.J., and Griffin III, L.I.	UMTRI-96-17

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
The University of Michigan Transportation Research Institute 2901 Baxter Road Ann Arbor, MI 48109-2150 Texas Transportation Institute The Texas A&M University System College Station, TX 77843	3A3B
	11. Contract or Grant No.
	DTFH61-93-C-00123

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Office of Safety and Traffic Operations R&D Federal Highway Administration 6300 Georgetown Pike McLean, Virginia 22101-2296	Task Report October 1993 - March 1996
	14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representatives (COTRs): Michael S. Griffith and Joe Bared, HSR-20

16. Abstract

This report describes existing and emerging exposure data sources for highway safety analysis. Existing exposure data sources reviewed include: Highway Performance Monitoring System (HPMS), Highway Safety Information System (HSIS), Long-Term Pavement Performance (LTPP) Monitoring System, Nationwide Personal Transportation Survey (NPTS), National Truck Trip Information Survey (NTTIS), Operational Exposure Data Sources, Residential Transportation Energy Consumption Survey, Truck Inventory and Use Survey (TIUS), and Weigh-in-Motion (WIM) devices. Emerging data sources are new sources or existing sources that have not been traditionally used to derive exposure estimates. Three areas were reviewed for possible emerging exposure data: Intelligent Transportation Systems (ITS), transportation planning surveys, and traffic volume data collected by the States. One-page summaries are provided for each exposure data source. A longer description covers the purpose of the collection, contents, period covered, sample design, data collection methods, sample size, data quality, data format, possible cautions in using the exposure data, and availability of the data.

17. Key Words	18. Distribution Statement
Exposure data, highway safety.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

Form DOT F 1700.7 Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yards	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3

NOTE: Volumes greater than 1000 l shall be shown in m^3.

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	9
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celcius temperature	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m*	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.71	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
9	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

• SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

(Revised September 1993)

CONTENTS

1. SUMMARY . . . 1
Highway Performance Monitoring System . . . 2
Highway Safety Information System . . . 3
Long-Term Pavement Performance Monitoring System . . . 4
Nationwide Personal Transportation Survey . . . 5
National Truck Trip Information Survey . . . 6
Operational Exposure Data Sources . . . 7
Residential Transportation Energy Consumption Survey . . . 8
Truck Inventory and Use Survey . . .
Weigh in Motion . . . 10
Intelligent Transportation Systems Commercial Vehicle Operations . . . 11
ITS Advanced Traveler Information Systems . . . 12
ITS Advanced Traffic Management Systems . . . 13
Transportation Planning Surveys . . . 14
Census Transportation Planning Package . . . 15

2. EXISTING EXPOSURE DATA SOURCES . . . 17
Highway Performance Monitoring System . . . 18
Highway Safety Information System . . . 23
Long-Term Pavement Performance . . . 45
Nationwide Personal Transportation Survey . . . 49
National Truck Trip Information Survey . . . 51
Operational Exposure Data Sources . . . 53
Residential Transportation Energy Consumption Survey . . . 54
Truck Inventory and Use Survey . . . 59
State Weigh-in-Motion and Automatic Vehicle Counting Devices . . . 61

3. EMERGING EXPOSURE DATA SOURCES . . . 67
Intelligent Transportation Systems . . . 68
Commercial Vehicle Operations of ITS . . . 68
Advanced Traveler Information Systems . . . 70
Advanced Traffic Management System . . . 70
Transportation Planning Surveys . . . 76
Traffic Volume Data — Errors of VMT Estimates Based on Traffic Counts and Section Length . . . 79

APPENDIX: HPMS FORMS AND DATA FORMAT . . . 83

LIST OF TABLES

Table 1. Region 1 WIM 62
Table 2. Region 2 WIM 63
Table 3. Region 3 WIM 64

1. SUMMARY

One-page summaries of both the existing and emerging exposure data sources reviewed for this report are presented in this section. A more complete discussion of each of the existing exposure data sources is presented in Section 2 and emerging data sources are in Section 3. The following exposure data sources are summarized in this section:

Existing Exposure Data Sources

Highway Performance Monitoring System (HPMS)
Highway Safety Information System (HSIS)
Long-Term Pavement Performance (LTPP) Monitoring System
Nationwide Personal Transportation Survey (NPTS)
National Truck Trip Information Survey (NTTIS)
Operational Exposure Data Sources
Residential Transportation Energy Consumption Survey
Truck Inventory and Use Survey (TIUS)
Weigh in Motion (WIM)

Emerging: Data Sources

Intelligent Transportation Systems
- Commercial Vehicle Operations
- Advanced Traveler Information Systems
- Advanced Traffic Management Systems

Transportation Planning Surveys
- Census Transportation Planning Package

Highway Performance Monitoring System (HPMS)
Federal Highway Administration and State Highway Agencies

Purpose:	Assess the length, use, condition, performance, and operating characteristics of the National Highway System
Source:	State highway agencies Vehicle-Miles Traveled (VMT) based on Annual Average Daily Traffic (AADT) Fatal and injury accident data
Coverage:	Annual reporting, initiated in 1978 All public roads in the United States (except local streets and roads) Areawide Universe Standard sample "Donut" sample (for air quality) Geographical Information System (GIS) coding
Sample:	Simple Random Sample (SRS) prescribed by the Federal Highway Administration (FHWA) of ~115,000 road segments
Response:	Data are required by law and, therefore, are complete
Strengths:	National aggregate data in broad categories of highway function, area type, and use Standard format
Limitations:	Accident data not associated with the standard sample (Vehicle classification of VMT not compatible with accident data)
Accuracy:	AADT is improved for the standard sample, but is still the critical element for VMT

Highway Safety Information System (HSIS)
Highway Safety Research Center

Purpose:	Provide linked accident, highway inventory, and traffic count data in SAS® format for selected States to provide an enhanced analysis capability
States Reviewed:	Illinois, Maine, Michigan, Minnesota, and Utah
Source:	VMT from segment lengths and AADT AADT updated from 1 to 5 years, Some estimated or interpolated, Some sites permanent, year-round, Most are temporary sites, 48-h counts Some with vehicle classification, or "commercial"
Coverage:	In most States, a major portion (but not all of the highway system) is covered, usually State-maintained roads
Sample:	Usually a purposefully selected subset Cross-section files in some States contain a sample of segments, usually limited
Strengths:	Large samples Diversity of data in different States SAS® format, documentation Suited for aggregate comparisons
Limitations:	AADT data very coarse, generally not suited for identifying individual, high-risk locations Entering volumes for both roads of an intersection often not available National estimates not possible Diversity of data in different States
Accuracy:	AADT not all observed, not independent, so variance cannot be estimated

Long-Term Pavement Performance (LTPP) Monitoring System
Transportation Research Board/Federal Highway Administration

Purpose:	Satisfy the total range of pavement information needs Collect information to develop models of how various design features, traffic, and environment impact pavement performance Central Traffic Database contains annual estimates of traffic and load data
Source:	Central Traffic Database contains historical and monitored traffic data Yearly estimates of volumes, axle loads, and equivalent single-axle loads available for each site Truck weights and distributions collected at sites quarterly for 7 days 35 percent of sites have weigh-in-motion collectors, the remainder have Automatic Vehicle Classification counters
Coverage:	Data collected in four geographic regions 20-year research program begun in 1987
Sample:	789 sites on key highway routes provide truck weights and distributions Historic traffic data requested where available
Strengths:	With further development, should provide reliable vehicle count and classification data Good data source for location-based safety studies, if sites can be linked with accident histories
Limitations:	Weigh-in-motion data location not always exactly at the site Researcher must verify exact location of traffic data Quality control issues with the data currently a problem Some sites have only a minimal amount of data Currently, only limited amount of data available to the public
Accuracy:	Currently a problem, expected to improve Data quality procedures and standards have been implemented

Nationwide Personal Transportation Survey (NPTS)
Federal Highway Administration

Purpose:	U.S. estimates of personal travel All modes: car, truck, bus, train, subway, airplane, taxi, motorcycle, bicycle, and walking Includes household demographics, person-level information, household vehicles, and trip information
Source:	Conducted by Research Triangle Institute (1990) Random-dialing household telephone survey 12-month survey period 24-h travel-day period 14-day travel period for trips >12 1 km
Coverage:	National coverage, all trips, all modes, all purposes, in all 50 States plus Washington, D.C. Oversample in Connecticut; N.Y. metropolitan planning organization; and Indianapolis, Indiana Approximately 7-year intervals
Sample:	22,000 households 48,000 persons 35,000 licensed drivers 41,000 vehicles
Response:	~85 percent at the household level
Strengths:	Only source for national personal travel Large sample size Stable since 1969 (Home interviews prior to 1990) Good detail at all levels
Limitations:	Households without telephones not included Limited sample for commercial vehicles (trucks) Self-reported information Cannot disaggregate by State 7-year interval
Accuracy:	Sampling errors can be calculated with appropriate software

National Truck Trip Information Survey (NTTIS)
University of Michigan Transportation Research Institute

Purpose:	National estimates of medium and heavy truck population and travel with detailed vehicle and trip-level data that allow cross-classification by configuration, loading, road type, rural/urban, and day/night
Source:	Sample of registered trucks from R.L. Polk Telephone surveys on four randomly assigned dates Conducted by University of Michigan Transportation Research Institute (UMTRI)
Coverage:	48 States plus Washington, D.C. Government-owned vehicles excluded 12-month survey period in 1985-1986 One time only
Sample:	Probability-based sample of 8,144 registered trucks (GVWR>4536 kg) from 1983 R.L. Polk files Trip-level data on a sub-sample of 5,000 vehicles 13,097 trips on 17,660 survey days
Response:	83 percent at the vehicle level 86 percent at the survey-day level
Strengths:	Most accurate identification of trucks > 4536 kg Duplicate registrations deleted from frame Detailed cross-classification of vehicle characteristics, loading, and operating environment unmatched in any other source Extensive edit and consistency checks Some questions overlap Truck Inventory and Use Survey for comparison
Limitations:	Limited sample size Cannot disaggregate by State Self-reported information Now out of date
Accuracy:	Underrepresents newest vehicles due to lag between sample and trip survey Complex sample design can be calculated with appropriate software Large variances for small subsets (doubles)

Operational Exposure Data Sources
State and Local Traffic Agencies

Purpose: State and local traffic agencies collect a variety of traffic data for both long-term and short-term objectives that often go beyond the requirements of the Highway Performance Monitoring System (HPMS) described previously. Typical data include traffic counts from both permanent and temporary stations, Automatic Traffic Recorders, and State highway inventory files. However, data collection beyond the scope of HPMS is often on an ad hoc basis to address specific short-term purposes.

Source: There is no single source. State traffic agencies are often aware of many of the local programs, as well as the State data; but the city, county, or metropolitan planning organization will have to be contacted to obtain detailed information or data.

Coverage: Most States have extensive traffic monitoring programs with a combination of permanent and temporary programs. Major cities often collect Average Daily Traffic (ADT) volumes on many arterial streets as well.

Sample: Some stations may be permanent and coverage of individual routes may be quite complete, but outside of HPMS, there is generally no sample design that would support any extrapolation of the data.

Strengths: Specific projects may be possible, taking advantage of additional details with regard to peak versus off-peak, day-of-week, and site-specific data that might be located.

Limitations: A major limitation is that none of the data is typically automated. Another important limitation is that the often ad hoc nature of the data collection may bias the data.

Residential Transportation Energy Consumption Survey
Energy Information Administration

Purpose:	Obtain information on the vehicles used for personal transportation in the United States Companion survey to the Residential Energy Consumption Survey (RECS) RECS includes household demographics Residential Transportation Energy Consumption Survey (RTECS) includes VMT (from odometer readings), motor vehicle stock, and vehicle fuel consumption and expenditure data.
Source:	RECS is a random household telephone survey (mail questionnaire used when telephone interview is not possible) Multistage probability sample incorporating a rotating panel RTECS is a subsample of RECS households, telephone/mail survey First phase of RTECS done in conjunction with RECS Subsequent three phases conducted at the beginning, middle, and end of the year
Coverage:	All 50 States and Washington, D.C. Families or individuals living in group quarters or with no fixed address excluded Motorcycles, bicycles, and all-terrain vehicles excluded Conducted every 3 years since 1985
Sample:	5,095 households responded to the most recent RECS survey 3,045 households selected for most recent RTECS survey
Response:	75 percent household response rate to RECS Unknown response rate to RECS
Strengths:	Household VMT and vehicle stock data Estimates of VMT by age and gender of primary driver Stable since 1978
Limitations:	Small sample size No trip data Two odometer readings not obtained for large fraction of sample vehicles, annual VMT imputed for these Data do not relate VMT to person-miles of travel, so vehicle occupancy is unknown, and driver age and gender have to be assumed from primary driver data 3-year interval
Accuracy:	Questionable 26 percent of households not followed for the entire year Various imputation techniques used to handle item nonresponse

Truck Inventory and Use Survey (TIUS)
Bureau of the Census

Purpose:	Estimate U.S. population of registered trucks (light, medium, and heavy) and provide descriptive information on the trucks and their use over the past year
Source:	R.L. Polk Stratified probability sample of truck registrations from each State Survey form mailed to each owner
Coverage:	Registered trucks in the 50 States plus Washington, D.C. "Typical" use during the past year Excludes government-owned and passenger vehicles Conducted every 5 years
Sample:	~100,000 vehicles
Response:	Required by law ~80 percent (1987)
Strengths:	Well-defined population Rigorous sample design (SRS) Large sample Good response Stable format back to 1967 Population estimates can be disaggregated by State
Limitations:	Self-reported "Typical" use over the past year underrepresents minority use such as bobtail or infrequent trailers/cargoes Mileage estimated cannot be disaggregated by State Possible duplications in registration data across States Conducted only every 5 years
Accuracy:	Sufficient data to calculate sampling errors not released Approximate error formulas provided Minimal bias, random errors generally small

Weigh in Motion

Purpose:	Provide information about vehicle weights and axle loads or decisions related to planning, funding, operating, and managing highway facilities for enforcement of weight limits
Source:	***Traffic Monitoring Guide (TMG)*** — required by FHWA and collected by State Departments of Transportation (DOTs) Long-Term Pavement Performance (LTPP) data — part of the Strategic Highway Research Program (SHRP) — collected by State DOTs and forwarded to regional SHRP contractors Truck weight enforcement stations — data collected by State police organizations, data usually not retained
Coverage:	National coverage
Sample:	TMG — 1,400 Weigh-in-Motion (WIM) sites throughout the United States LTPP — 777 WIM sites throughout the United States
Data Availability:	National database containing station description, traffic volume, vehicle classification, and truck weight available directly from FHWA in ASCII flat-file format Individual State data must be requested from State DOTs, data formats vary widely
Strengths:	Only national source for exposure by truck weight
Weakness:	Compatibility of TMG data across States — each State determines own experimental design, and number and location of WIM sites Hardware and software problems associated with collecting data
Accuracy:	Varies by State — need to contact State for design and sampling information

Intelligent Transportation Systems (ITS) Commercial Vehicle Operations

Broad Categories of Commercial Vehicle Operations (CVO) User Services:

- Commercial vehicle electronic clearance.
- Automated roadside inspections.
- Commercial vehicle administrative services.
- On-board safety monitoring.
- Hazardous material incident response.
- Commercial fleet management.

Commercial Fleet Management:

Global Positioning System (GPS) recording of vehicle trips by fleet linking with cargo, configuration, and vehicle data

Prospect? Produce the electronic equivalent of a trip diary

Commercial Vehicle Administrative Services:

Vehicle-based GPS technology to get travel by State for International Registration Plan (IRP) purposes (Iowa)

Prospect? Added GPS detail could produce a vehicle-based sample of mileage by road type

Commercial Vehicle Electronic Clearance:

Electronic roadside sampling to transmit compliance data

Prospect? Roadside sampling of vehicle, cargo, and driver characteristics
Identification could allow tracking to subsequent locations to get VMT and travel time

ITS Advanced Traveler Information Systems

Route Navigation:

Vehicle-based navigation system could retain a history of the route followed, plus speed and time, providing an electronic trip diary

Other Uses of ITS Technology:

WIM technology installed on a banked curve could measure vehicle center-of-gravity (cg) height

ITS Advanced Traffic Management Systems

Purpose:	Detailed traffic volume data are collected in many large metropolitan areas to provide real-time information for sophisticated traffic management systems. Details vary from one installation to the next. Each city must be contacted for specific information. Seattle and Minneapolis/St. Paul are reviewed in Section 3 of this report.
Source:	Inductive loops are the primary source for both volume and speed data. Some automatic vehicle classification equipment is used.
Coverage:	High-volume freeways in large metropolitan areas.
Sample:	Coverage of road network under the control of the traffic management system is essentially complete.
Strengths:	Data are automated and all historical data are archived. Level of detail typically is on the order of 1-min counts per lane at 0.8-km intervals in both directions with speed data for a subset of the stations, plus some ramp measurements. A typical installation has several hundred stations.
Limitations:	Limited to the highway network covered.
Accuracy:	Accuracy of the data from inductive loops is not 100 percent, but is comparable with other traffic volume measurement methods. Observations outside the expected range are automatically flagged in the better systems.

Transportation Planning Surveys
(Travel)

Purpose:	Designed primarily as origin-destination surveys for planning purposes like the Census Transportation Planning Package (CTPP), with coverage of more trip purposes, but for a limited geographic region.
Source:	Metropolitan planning organizations, or sometime States, conduct additional surveys, often to support travel demand models and other requirements of the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA).
Coverage:	Limited geographic region Broader coverage of trip purposes
Sample:	Usually a census-based household sample, plus surveys of registered trucks or taxis, and roadside surveys.
Strengths:	More complete coverage of trip purposes and time of day Objective is to get future origin-destination flows by travel mode
Limitations:	Difficult to get VMT estimates Geographic limitation

Census Transportation Planning Package (CTPP)
Bureau of the Census

Purpose:	Provide national data for transportation planners on the journey to work. Focus is on the origin-destination flows between traffic analysis zones
Source:	Questions on a supplement to the U.S. Census that is sent to a sample of households, covering residential location, employment location, mode of journey, starting time, and journey time.
Coverage:	National, but only for the journey to work.
Sample:	Statewide package Urban Package SRS of about one out of six households
Strengths:	Designed for transportation planning purposes.
Limitations:	Journey to work only VMT not available Difficult to imagine application to safety analysis
Accuracy:	Sampling errors can be calculated.

2. EXISTING EXPOSURE DATA SOURCES

Existing exposure data sources for use in highway safety analysis are described in this section. The following exposure data sources are included:

Highway Performance Monitoring System (HPMS)
Highway Safety Information System (HSIS)
Long-Term Pavement Performance (LTPP)
Nationwide Personal Transportation Survey (NPTS)
National Truck Trip Information Survey (NTTIS)
Operational Exposure Data Sources
Residential Transportation Energy Consumption Survey (RTECS)
Truck Inventory and Use Survey (TIUS)
Weigh in Motion (WIM)

A description of each data source has been prepared for a data catalog. The objective of the catalog is to provide the highway safety researcher with sufficient information to assess the feasibility (considering time, level of effort, and cost constraints) of using the exposure data source in designing a highway safety evaluation study. The descriptions contain the following information, as applicable:

- Original purpose of the data collection.
- Brief description of the contents of the data source that would be of interest in highway safety research.
- Discussion about the quality of the data, how the data were archived, and for what time periods.
- Discussion of data collection methods or the performance characteristics of the equipment used in terms of reliability and data quality.
- Discussion of the number of sites and locations of the data collection effort and the statistical reliability of these sample sizes as applied to highway safety research.
- Sample of the data format and details as to how to obtain the data, what software or hardware is necessary to access the data, how often the data are updated, and the frequency of data releases, etc.
- Cautions and potential problems with exposure estimates.

Highway Performance Monitoring System (HPMS)

Contents

The Highway Performance Monitoring System (HPMS) is a nationwide inventory system that includes all of the Nation's public road mileage. The primary purpose of the HPMS is to serve the data and information needs of the FHWA and Congress. The HPMS assesses the system length, use, condition, performance, and operating characteristics of the highway infrastructure.

The HPMS was initiated in 1978 to consolidate and streamline the States' data collection efforts and reporting requirements. In keeping with FHWA's mandate to provide information, the HPMS is reassessed and modified to collect data relevant to emerging issues. In such a way, collection of pavement information was added to the HPMS in 1987. It was modified again in 1993 to respond to the need to monitor travel for the clean air issues. The HPMS also changes with advances in technology. In 1993, States were required to submit a linear referencing system for their road systems. Thus, the structure of HPMS is undergoing changes over time as data items are added and dropped in response to current information needs.

The HPMS organization, guidance, and analyses are the responsibility of the FHWA. Data reporting for the HPMS is accomplished by the State highway agencies in cooperation with local governmental units and metropolitan planning agencies.

The HPMS report submitted annually by each State consists of:

- Areawide data.
- Universe data.
- Data for a standard sample.
- Data for the "donut" sample (new in 1993).
- Linear referencing system (new in 1993).

Areawide Data. The areawide data consist of statewide summaries. These data consist of the totals for mileage, travel, accidents, local system data, land area, population, and travel activity by vehicle type. This information is reported for rural, total small urban, and individual urbanized areas.

Universe Data. Universe data refers to a limited set of data items reported for the entire public roads system as individual sections or grouped length records. The public roads system includes those roads owned by the State highway agency, local governments, and Federal agencies. These data contain a complete inventory of mileage classified by system, jurisdiction, and selected operational characteristics.

Standard Sample Data. The standard sample data include specific inventory, condition, and operational data obtained for the sample panels of highway sections. These data can be expanded to represent the universe of highway mileage.

The data cover:

- Identification relative to functional system, route, jurisdiction, and area type.
- Operational information about volume, lanes, access control, medians, and pavement.
- Geometric information about lane widths, shoulders, right-of-way (ROW), horizontal and vertical alignment, and passing sight distance.
- Traffic volume and capacity information such as AADT, speed limits, design factors, service volumes, and signalization.
- Environmental information such as climate and drainage.
- Intersection and interchange information.
- Information about capital improvements.

"Donut" Sample Data. "Donut" data requirements were added to the HPMS in 1993 in response to a need of the Environmental Protection Agency (EPA). The "donut" sample is a supplementary sample of highway panels from the nonurbanized portion (donut area) of National Ambient Air Quality Standards (NAAQS) nonattainment areas. This additional sampling is required to serve EPA's Section 187 Travel Tracking and Forecasting Procedures for the NAAQS non-attainment areas.

The data items are a subset of the data items provided for the standard sample and include identifiers, AADT, and expansion factors.

Linear Referencing System. A linear referencing system (LRS) was added to the HPMS for the 1993 report. These data will enhance the HPMS with Geographic Information System (GIS) capabilities. The data consist of node data file, inventory route and link data files, and inventory route and node maps for the principal arterial system/national highway system (PAS/NHS), and the rural minor arterial system.

Samples

Standard Sample. The HPMS universe consists of all public highways or roads within a State with the exception of roads functionally classified as local. The reporting strata for the HPMS include type of area (rural, small urban, and individual or collective urbanized areas) and functional class (in rural areas, these are Interstate; other principal arterial, minor arterial, major collector, and minor collector; in urban areas, these are Interstate, other freeway or expressway, other principal arterial, minor arterial, and collector). A third level of stratification based on volume was added as a statistical device to reduce sample size and to ensure inclusion of the higher volume sections of the sample in 1987.

The HPMS sampling element is defined on the basis of road segment, which includes both. directions of travel and all travel lanes within the section. The HPMS standard sample design is a stratified simple random sample.

Donut Area Sample. The donut area sampling universe consists of all highway sections functionally classified as rural minor arterial and major collector, and small urban minor arterial and collector that are located within the defined nonattainment boundary and outside of all urbanized area boundaries. This typically forms an annular spatial area and is, therefore, called a "donut."

The donut universe is stratified into two functional systems (the minor arterial and collector) and a limited number of volume-group strata. The sample is a stratified simple random sample.

Data Quality

Generally, the quality of data is good. There is some variation in quality of the HPMS reports across the States. Since these data are required by the Federal Government and used for developing national policy and determining the funding of highways, the States comply.

The frequency of missing data is very low. However, whenever there is a change in the HPMS, such as the addition of the donut area information in 1993, there are some problems with the new data from some of the States. Typically, such problems are resolved by the second year of the requirement.

Coverage

FHWA has all the HPMS data from 1978 to the present. Individual States generally will have only their most recent few years.

The national universe data for 1 year contain about 3.25 million records. It is stored on tapes. Records go back to 1980.

The total national standard sample contains approximately 115,000 records per year. Again, these data are stored on tape. Records go back to 1978.

The areawide data for each State are submitted on a series of templates. At first, there were five templates that were submitted on paper. Later, spreadsheet templates were allowed. In 1993, the number of templates was increased to seven and spreadsheet templates (Lotus 1-2-3) were mandated.

Annually, FHWA transfers these records to a mainframe file and stores them on tape. One format was used until 1992. A new format (basically an ASCII file) was instituted in 1993.

The first submissions of the donut sample and line referencing systems 'were required in 1994. There are no archives of them at this time.

Measurement

The key variable in the sampling design of the HPMS is AADT. AADT is not directly measured (except for a very small number of continuous permanent counting stations in each State), but is either derived from short counts, factored from previous counts, or estimated in some other manner.

States are asked to maintain at least one automatic traffic recorder (ATR) on each route of the PAS/NHS and a minimum of three on both the rural and urban portions of the non-PAS/NHS highways. These are used to develop day of week and seasonal factors used for expansion of short counts to AADT.

Typically, volumes at the ATRs are measured with pavement loops. Pavement loops are prone to failure, especially in northern climates and from construction vehicles. However, failures at ATR stations are supposed to be repaired as soon as possible. Recently, other more reliable technologies have been introduced.

The HPMS methodology requires that traffic counts of at least 24 h be conducted on one-third of the road sections in the standard sample each year. These counts typically are taken with pneumatic tube-type portable counters. These are reliable and, if a problem is suspected, the count can be easily repeated. The vehicle volume is derived from these counts by adjusting for the number of multi-axle vehicles in the traffic flow.

The AADT for these sections is then calculated from the short period volumes, with the application of adjustment factors developed from volumes at the ATRs.

The AADT at the sites where traffic counts were not made in the current year is factored from previous counts at the site or by other methods (estimation, engineering judgment, tracing volume maps, etc.). The method of AADT estimation for each site is one of the data items for the sample.

Statistical Reliability

The HPMS standard sample design is a stratified simple random sample. The HPMS sample size estimation process was tied to the AADT. Of the approximately 80 data items collected, AADT is perhaps the most variable data item in HPMS. Therefore, the reliability of most other characteristics would be expected to exceed that of AADT.

The sample size for each stratum of the samples is prescribed in the HPMS Manual. The sample sizes per functional system vary by State according to the total number of road sections (universe), the number of predetermined volume groups, the validity of the State's AADT data, and the design precision levels.

For rural, small urban, and collective urbanized areas, sample sizes are based on 90-5 precision levels for volume groups of the Principal Arterial System (PAS), 90-10 for minor arterial system, and 80-10 for the collectors (excluding minor collectors).

For individual urbanized areas with populations > 200,000 that are in NAAQS non-attainment areas, the design precision is 90-10 for the arterial system and 80-10 for collectors.

For individually sampled urban areas with populations < 200,000, the precision levels are 80-10 or 70-15 depending on several other factors.

The only objective of the donut portion of the HPMS is to estimate the daily vehicle-miles traveled (DVMT) within the donut areas with a precision of ±10 percent with 90 percent confidence. DVMT is determined from AADT. Thus, the sample size for a particular donut area is based on the-variability of AADT in that donut area.

Data Format and Access

The templates for the areawide data and the data format for the universe, standard sample, and donut sample data are shown in the appendix. Note that the fields are marked with an A, S, or D indicating that this field is required for all records, standard sample records, or donut records, respectively.

To obtain these data files or some portion of these data, contact the Highway Systems Performance Division of the FHWA.

All data are available on IBM readable mainframe computer tapes. The types of tapes that the data are stored on correspond to tape technology at the time the data were collected.

The universe data file is extremely large, approximately 3.25 million records per year. It does not appear particularly useful for highway safety research. However, should a researcher have a need for this information, he/she would have to contact the Highway Systems Performance Division and work out the details of copying the desired tapes. The researcher would have to provide the tapes.

The standard sample data consist of about 115,000 records per year. All the available data sets (from 1978) can be obtained on mainframe cartridge tape.

The areawide data are available on mainframe computer tape in Extended Binary Coded Decimal Information Code (EBCDIC) format. These files can be obtained from the Highway Systems Performance Division on PC diskettes in ASCII format.

The HPMS is updated annually and a new HPMS is generated at that time. It is important to note that some of the data fields and even some of the overall structure of HPMS may change from year to year.

The HPMS data from the States for the previous year is due at FHWA on June 15. It becomes available outside the FHWA sometime at the end of the year. Thus, a researcher can get data from the 1993 HPMS in December 1994 or January 1995.

The FHWA contact for HPMS is: David R. McElhaney, Director
Office of Highway Information Management
Federal Highway Administration
400 7th Street S.W.
Washington, DC 20590
(202) 366-0180

Reference

Highway Performance Monitoring System Field Manual. Federal Highway Administration. OMB No. 2125-0028. 1993.

Highway Safety Information System (HSIS), FHWA

General: The Highway Safety Information System is produced by the Highway Safety Research Center (HSRC) at the University of North Carolina.

Purpose: The FHWA has selected States for HSIS that provide linked accident, highway inventory, and traffic count data, and has converted the files to SAS format to provide an enhanced analysis capability. This introductory section provides only a general overview of the data. Descriptions specific to each of the States follow.

Source: VMT is estimated from segment lengths and AADT. The AADT volumes are updated from 1 to 5 years. Some values are estimated or interpolated; some sites are permanent and some are year-round. Most are temporary sites, taking 48-h counts. Some have vehicle classification, or "commercial," vehicle counts. 'Commercial" is usually any vehicle with two axles and six tires or more.

Coverage: States covered in this write-up include: California, Illinois, Maine, Michigan, Minnesota, North Carolina, Utah, and Washington. Additional States are being added to HSIS. In most States, a major portion (but not all) of the highway system is covered. Usually, these are the State-maintained roads.

Sample: The highway segments covered are usually a purposefully selected subset. Cross-section files in some States contain a sample of segments, usually limited in number.

Strengths: Sample size is large, and there is a diversity of data in different States. The files are in SAS format for convenience, and the documentation is better than usually available from the States. The data are suited for aggregate comparisons.

Limitations: The AADT data are sometimes coarse, and may not be suited for identifying individual, high-risk locations. Entering volumes for both roads of an intersection often are not available. National estimates are not possible. The diversity of data in different States can also be a disadvantage.

Accuracy: AADT volumes are not all observed and are not independent, so the variance cannot be estimated.

Also included at the end of this section is a brief discussion of the statistical implications of the nature of the traffic volume data in most State files. Issues discussed include the use of a purposeful sample rather than a random selection of sites for counts, and the use of estimated or interpolated counts rather than actual counts. A general conclusion is that the traffic volume data will not support a statistically defensible analysis (except when the HPMS procedures have been followed). However, a purposeful sample can be representative, although the variance is likely to be underestimated. Similarly, estimated or interpolated counts may also be reasonable in value, but again, the variance will be underestimated. When highway sections have been stratified prior to selecting sites, the most rigorous use of the data is to calculate estimates at the strata level. Use of the volume data to simply stratify the data into volume groups is also relatively sound.

Thus, the traffic volume data must be used with caution. The actual extent of any of these problems cannot be estimated without additional data. Estimated or interpolated counts mean that the observations are no longer independent, and most statistical techniques are no longer appropriate. In particular, the variance is underestimated and bias may be introduced. The analyst should be aware of the source of the traffic counts in each State and should use good judgment in the selection of an analytic approach. Though statistically sound analyses of accident rates may not be possible with the currently available exposure information, it may be possible to use this information in a productive way, e.g., for stratifying sites, and to perform within the strata only analyses relying on counts.

HSIS Contacts: Jeffrey Paniati at (703) 285-2057 or Yusuf Mohamedshah at (703) 285-2090

California, HSIS

Coverage: The current accident files cover the years 199 1 to 1995, and there is roadway information for 1993 and 1994. Accident reporting is not uniform in California, with some municipalities using their own report form and reporting threshold, instead of the California Highway Patrol (CHP) form. Accidents occurring on State routes (including those in urban areas that do not use the CHP form) are location coded. There are about 150,000 accidents annually on State routes (all with location codes) out of an estimated statewide total of 500,000 accidents per year. Reporting is also not complete for uninjured occupants. Information on uninjured occupants is only collected if there is al least one injured occupant. Thus, the occupant injury data are biased to overrepresent injured occupants. However, uninjured drivers have been identified in the driver file by Highway Safety Research Center (HSRC) by linking the injury information from the occupant file with the vehicle file. Overall, HSRC estimates that information on uninjured occupants is missing for about 50 percent of non-towaway accidents.

The roadway information is contained in three files: the Roadlog file, the Intersection file, and the Interchange Ramp file. The Roadlog file contains information on approximately 24,461 km of roadway, including about 3943 km of Interstate, 17,702 km of other primary highways, and about 2736 km of secondary/county/township roads. The 24,461 km are divided into about 50,000 records in the Roadlog file, for an average section length of 0.5 km.

The Roadlog file contains information describing the functional class of the road, cross section information such as width and number of lanes, as well as information on design speed, median barriers, and other special features. The intersection file has information of 20,000 intersections, and the Interchange Ramp file has information on 14,000 ramps. Accidents can be linked with all three roadway files and the intersection file can be linked with the associated segments in the Roadlog file, but the Interchange Ramp data cannot be linked with its associated interchange.

Exposure Information: The Roadlog file includes an AADT and a DVMT for each segment (record). Section length is also included. No information on truck travel is available. In the Intersection file, there is an AADT for the mainline road and for the crossing road, as well as descriptive information for both the mainline and cross road. AADT is also included in the Interchange Ramp file.

Traffic Data: As indicated in the preceding three sections, all three inventory files contain AADT information. In addition, the Roadlog file contains information on DVMT, which is computed as the product of the section length and section AADT estimate.

In California, the 12 district offices have the responsibility of collecting traffic data and developing the AADT estimates for each road section within their district. The Division of Traffic Operations of the Caltrans central office oversees the operation and attempts to maintain consistency in the methods and data across all districts as much as possible. If requested, Traffic Operations personnel will assist a district in calculating the AADT estimates. The division also maintains all count data on an on-line computer file for the districts' use.

There are approximately 2,100 permanent count stations on mainline highways operated by Caltrans in California. Of these, approximately 400 are permanent, continuous counting control stations that operate each day in a given year. Every major State-administered route is counted each year. The 400 permanent continuous count stations form a network that covers all major routes. The remaining control stations are permanent, quarterly counting control stations, i.e., in-pavement loops to which a counter/recorder device is attached for 7 to 14 days during each quarter. Caltrans also collects count data at approximately 700 of these quarterly counting control stations once every 3 years. In a given year, there are approximately 1,000 permanent quarterly counting stations where count data are not collected. California has determined that the AADT estimates, which are derived from the simple average of the four (unadjusted) quarterly counts, does indeed account for seasonal fluctuations without further adjustment based on nearby permanent counters. Consequently, there are no additional adjustments or corrections applied to the AADT's estimated from the quarterly counts.

In addition to the permanent control stations, approximately 1,000 coverage counts are collected annually. The intent is to collect coverage counts on a 3-year cycle (for a total of approximately 3,000 coverage counts), although conditions may force longer intervals in certain districts at times. A coverage count is basically a 24-h to l-week count.

Coverage counts are expanded to AADT estimates using factors derived from the combined continuous counts and quarterly count data. For road sections that are not counted in a given year, it is the responsibility of the districts to develop these AADT estimates. In some cases, the districts reply on overall traffic growth trends within the district. However, in most cases, the AADT assigned to the section is developed by studying the traffic growth in counts falling on each side of the section.

It is also noted that 24-h to l-week coverage counts are collected on approximately 3,200 on- and off-ramps per year. These ramp counts are manipulated through ramp balancing to reflect continuity of flow on mainline freeways.

Finally, vehicle classification data are collected at approximately 70 permanent stations across the State. Additional classification counts are collected on an as-requested basis, typically at locations where traffic count data are being collected. Since this is district-based, there is no reliable estimate on how many additional classification counts are collected across all 12 districts per year. Finally, there are approximately 45 weigh-in-motion stations statewide that provide speed, volume, and the "13-bin" vehicle classification information. (Taken from ***HSIS Guidebook for the California State Data Files.)***

Linking Accident and Exposure Information: Accidents can be linked with all three roadway files. Accidents are located manually using the scene diagram on the accident report and maps. Accuracy of the location is believed to be within 0.16 km, and missing data is only a few percentage points.

Illinois, HSIS

Coverage: During 1985 to 1994, this included 26,232 km of roadway of which 2,736 km were Interstate highways; 15,449 km of other primary roadways, and 8,047 km of secondary, county, and township roads.

Exposure Information: All exposure information is contained in the Roadlog file, which contains records for 197,000 sections; each section, on average, is slightly less than 0.16 km.

Exposure, in terms of VMT, can be calculated from AADT and the section length. In addition to the total, AADT for "heavy commercial vehicles" (defined as having two or more axles and six or more tires) is given.

Intersection information is in the Roadlog file and also in an Intersection Location file. They contain the same information, but the Intersection Location file contains one record for each intersection. If there is more than one intersection in a section, the information from the Roadlog file is repeated for each intersection record. Intersections are characterized as "across," "left," and "right. " The crossing road is apparently not identifiable. Thus, it appears that for intersection exposure only, the AADT on the through road is available.

Traffic Data: As indicated earlier, the Roadlog file contains information on AADT, percentage of trucks for 1990 and earlier, and commercial vehicle AADT for 1991 and later. These data are developed in Illinois' traffic volume counting program and are based on a combination of permanent counters that count traffic 24 h each day for 365 days each year and a series of short-term "coverage" counts conducted each year. Illinois has 49 automatic traffic recorders (ATRs), of which 21 are capable of collecting counts by vehicle class in accordance with FHWA's Scheme F. The ATR locations on the 5 different classes of roadway include 7 locations on rural Interstate roadway, 6 locations on urban Interstate locations, 12 locations on other rural roadways, 19 locations on other urban routes, and 5 locations on "recreational" routes.

In addition to the ATR data, short-term traffic counts on Interstate and primary highway systems are done on a 2-year cycle. During even-numbered years, portable counter devices are deployed in combination with pre-established in-pavement loop detectors. Typically, the counter devices are deployed during 1 week of the year at any given site. Short counts (e.g., 24- or 48-h counts) are collected on Monday through Thursday only. It should be noted that a sample of Interstate sections are counted 1 week out of every 4 months. During odd-numbered years, the Illinois DOT conducts a comprehensive interchange ramp counting program on State highways. These ramp counts are used to supplement ADT data for sections where the State did not have monitors (i.e., counter devices). In total, it is estimated that approximately 96 percent of the primary system is covered during each 2-year cycle.

For other non-primary roads (i.e., the "off" marked route system), Illinois collects 48-h coverage counts in approximately 20 percent of the counties once every 5 years. However, the northeast counties are done every 4 years. With the exception of Cook County, which is also on a 4-year cycle, urban areas within counties are counted on a 5-year statewide cycle.

Additional vehicle classification counts are conducted on HPMS sections. These are made at 300 locations over a 3-year cycle (i.e., approximately 100 each year) to form a representative distribution for-the State.

Finally, the districts often have a need for additional traffic data. Consequently, when requested, the State collects 12-hour turning movement counts at intersections and other "special" traffic data to satisfy these needs.

To convert the short-term coverage counts to AADT, Illinois applies adjustments to reflect corrections for number of axles and for seasonal differences in the daily traffic. Axle corrections are developed from both permanent classification counters and from manual (HPMS) counts. For seasonal corrections, each coverage count location is assigned to one of the five categories of roadway where permanent counters are located, as defined above. The seasonal factors are based on averages from all ATRs in that group.

When a road section is not counted during a given year, growth factors are developed and applied to the most recent prior year's count. Average growth factors are created each year for each functional class of roadway using ATR data and data from adjusted short counts for the current year. The growth factor applied to a particular uncounted section is based on its functional class. For sections where no prior AADT exist, AADT/mile averages by functional class are developed and then used in order to "fill in" the AADTs.

Finally, it should be noted that the percentages of truck-related "Heavy Commercial Volumes" include "two-axle trucks with six or more tires plus multi-axle vehicles." Thus, while pick-ups and vans are excluded, this combination would include single trucks, tractor-semi combinations, and buses. Thus, it cannot be considered a count of just the multiple unit (tractor-trailer) trucks that are found on the roadway system. (Taken from the ***HSIS Guidebook for the Illinois State Data Files.)***

Linking Accident and Exposure Information: Data on different files can be linked by a linkage key, which combines county, route prefix, and route number with the station number.

For intersection accidents, the intersecting route number and route prefix are given. However, it does not appear possible to identify which vehicle approached the intersection from the main road and which one approached from the crossing road. The direction of travel for each vehicle is given, but the direction of the road is not given in the Roadlog file.

Maine, HSIS

Coverage: The Link Record file covers all highways in Maine, including local roads and urban streets. The 35,405 km are divided into 67,000 links. Files are currently available for the years 1985 to 1994.

Exposure Information: The Link Record file contains AADT for each link; the year of AADT; and whether it is an actual count, an interpolation, or an estimate. Together with the length of the link, VMT can be estimated.

Information on intersections is available from the Node Records file, which also includes nodes other than intersections. The configuration of each intersection is given, and up to six legs are identified by the corresponding link numbers. As an exposure measure, only the total number of vehicles entering the intersection is given. However, it is possible to obtain the AADT for each leg from the Road Link file.

Traffic Data: With respect to the traffic information on both the Link and the Node files, the traffic counts that are in the system are extracted from a traffic file again prepared within the Bureau of Planning. The counts are extracted from a series of 54 permanent count stations across the State, 6 of which do detailed vehicle classification counts. There are a total of 9 stations on Interstate routes (which collect counts in both directions), approximately 13 stations on U.S. routes, 24 stations on State routes, and 8 stations on other routes.

In addition to the continuous count stations, each summer, 48-h counts are done at between 1,600 to 2,200 locations on all US and State highways. Beginning in 1994, the number of coverage counts increased to between 3,600 and 4,200. Approximately 10 percent of these counts include vehicle classification counts. Classification estimates exist for other locations that are not high-priority locations.

Each year, these counts are done in either the northern, central, or southern areas of the State. The counters move to a different area the following summer, covering the entire State every 5 years. The southern and central areas are counted in alternate years for the first 4 years of a cycle. Then, the northern area, where counts change less per year, is counted during the fifth year of the cycle.

Seasonal adjustment factors for the coverage counts are based on continuous count stations that fall into the same "highway type" category as the coverage count. Based on extensive analysis in the late 1980's, the three categories used are Urban (including suburban locations), Arterial (including all Interstate locations plus other locations in rural areas), and Recreational locations (whether urban or rural). The actual adjustment factor for a given coverage count location is based on the weekly average ADT for all continuous count stations falling into that category.

For years in which no count data were collected within a given area of the State, historical daily traffic flows are factored up on a county-by-county basis. The growth factor used is based primarily on traffic changes at the continuous count stations falling into the same highway-type category described above. Other information used in developing a specific growth factor includes counts from nearby urbanized areas and special counts that may have been conducted at the location for other reasons. The final growth factor used is based on interpolation between points of known growth (such as 2 or more years at the similar continuous count stations), and is done by personnel with a working knowledge of the system's traffic patterns.

In summary, while some of the counts may be off due to roadside development and/or roadway construction within a specific area of the State that occurred within the 2-year period, in general, the count data are felt to be quite adequate for analysis purposes. (Taken from the ***HSIS Guidebook for the Maine State Data Files.)***

Linking Accident and Exposure Information: Accident and exposure data can be linked by the low and high node numbers that identify each segment and by the distance from the low node given in the accident record.

Intersection accidents are identified as such, distinguishing three-, four-, and five-leg intersections. However, the leg from which a vehicle entered an intersection cannot be determined.

Michigan, HSIS

Coverage: Of 189,897 km of roadway in Michigan, the Roadway Segment file covers only 15,449 km of trunkline divided into 43,000 segments. Data for the years 1985 to 1994 are currently available.

Exposure Information: The Roadway Segment file shows AADT categorized into 10 classes. Commercial AADT is also given. No definition of "commercial" is shown. AADT for the segment is given.

A Cross Section file covers 8,047 km of two-lane rural roads with segments selected by a stratified random sample. Very detailed roadside feature information is given. However, there is no information on sample stratum. ADT values are given based on counts in the early 1980s. Counts of accidents by severity are given.

There is an Intersection file that has recently been released for analysis. However, information on AADT or vehicles entering the intersection is not provided.

Traffic Data: As noted above, information on AADT and Commercial Vehicle AADT is found on the Roadlog file. These data are developed in Michigan's traffic counting program, which, like other States, includes both full-time permanent counter locations that operate 365 days each year and short-term coverage counts at a much larger number of locations. Michigan DOT currently operates and maintains 121 permanent traffic recording (PTR) stations. These PTRs include 34 on Interstates, 31 on U.S. routes, 23 on Michigan State highways, and 12 on other routes.

In addition, there are a varying number of short-term "coverage counts" conducted each year. Michigan DOT indicated that approximately 3,300 such 48-h "short" counts were requested in 1995. These coverage counts included the following:

- 950 short counts (volume only).
- 1,300 classification counts (volume by vehicle class).
- 1,000 interchange ramp counts.

Michigan attempts to count every State-maintained road section in a 3-year period. Unless required under the HPMS, Michigan also attempts to collect classification counts over a 6-year cycle. It should be noted that in addition to the State's traffic counting program, other agencies (notably those in urban areas) are also collecting traffic data for HPMS purposes. Furthermore, the Metropolitan Planning Organizations (MPOs) in Michigan have developed and supported urban transportation planning models in accordance with ISTEA requirements. These MPOs subsequently have their own counting programs to support their model development and application.

To factor up the short counts to reflect AADT, seasonal factors are developed. Unlike some States where these seasonal factors are based on PTR counts within the same functional class as the short-count location, Michigan has defined six or seven “cluster-analysis groups.” Each of these groups contains a number of PTRs, and the adjustment factors are based on averaging the PTR counts within that group. Each roadway section (and thus each short count) is assigned to one of these cluster-analysis groups.

When a specific roadway section is not counted in a given year, its count from the previous year must be adjusted to represent traffic growth. Here, Michigan attempts to “look up and down the road” and identify the closest, comparable section for which an ADT was estimated (counted) for the given year. They determine the percentage change (e.g., increase or decrease) in the ADT associated with that “comparable” section, and apply that percentage change to the historical count for the specific section in question. (Taken from the ***HSIS Guidebook for the Michigan State Data Files.)***

Linking Accident and Exposure Information: Though the Roadway Segment file covers less than 10 percent of the total highway mileage, about one-third of all accidents can be matched with locations on the Roadway Segment file. Linking can be done via information on the control section, and the milepost.

Accidents that occur within 30.5 m of an intersection with a trunkline road are coded for that road with the milepost of the intersection.

Minnesota, HSIS

Coverage: Coverage includes the years 1985 to 1994; however, some files are available only for certain years, and there were changes between the years. Files detail 19,311 km of primary roadways, an additional 37,014 km of State-maintained systems, and 157,711 km of county and local roads.

Exposure Information: Two files provide exposure information: (1) the Roadlog file and (2) the Intersection/Interchange file.

The Roadlog file contains information on about 200,000 road sections on which highway characteristics remain constant. Exposure in terms of VMT can be obtained from the values of AADT given for the segment, and the given length of the segment. Also given is "commercial" ADT. Commercial vehicles are defined as having at least two axles and at least six tires. Exposure estimates can be stratified according to the highway characteristics contained in the file (also according to AADT or AADT per lane).

The Roadlog file identifies the type of intersection at the beginning of a segment. However, it does not identify the intersecting road. Thus, intersection exposure cannot be obtained from this file.

The Intersection/Interchange file contains data on 3,500 intersections, 256 interchanges, and 2,800 grade crossings, currently for the years 1987, 1989, and 1991. Intersections were originally selected for the purpose of identifying high accident locations, but are retained in the file.

Intersection type and a code describing it in some detail are given. The route on which each approaching segment is located is identified, and there are up to two legs for each segment. The direction (N, NE, E, etc.) of each leg is also shown. This allows reconstruction of the configuration of the intersection. For each leg of each segment, the AADT for several years is given. For the second leg of a crossing minor roadway, in 10 percent to 30 percent of the cases, AADT is missing. In these cases, it is recommended that the value for the first leg be used. Thus, the available exposure for intersections consists of AADT on the intersection approaches.

Commercial AADT is not given for intersections. However, it appears possible, though cumbersome, to obtain this information from the Roadlog file.

Traffic Data: The Traffic file contains information related to AADT data for all roadway sections across the State. This information is manually derived from sample and continuous counts taken at temporary and permanent count stations throughout the State. It contains total AADTs and AADTs for heavy commercial vehicles (which are defined as vehicles with two axles and six or more tires).

Like other States, Minnesota develops traffic volume estimates based on automatic traffic recorder stations (ATRs) and short-term (48-h) "coverage" counts. There are approximately 120 ATRs that count traffic 24 hours per day, 365 days per year, across the various roadway types. These are located on all classes of both rural and urban highway, with approximately 55 percent of the locations being on urban roadways and 45 percent on rural roadways.

In addition, there are approximately 34,000 coverage (temporary) count locations across the State where 48-h counts are made. Approximately 12,000 of these locations are covered each year. For the trunk highway system (including Interstate roads), these counts are made on a 't-year cycle, as are counts on roads within the Twin Cities metropolitan area. For the lower order County State-Aid Highways and the Municipal State-Aid System outside the Twin Cities metropolitan area, the counts are made on a 4-year cycle.

The seasonal adjustment factor for a given coverage count is based on counts made at ATRs which are similar to the coverage count location. Here, ATRs are grouped into the following classifications:

Outside (i.e., non-metropolitan area)

- Rural, farm-to-market roads.
- Rural, weekend recreational road.
- Rural, summer-peak recreational road.
- Municipal, non-recreational road, less than 5,000 population.
- Municipal, non-recreational road, more than 5,000 population.
- Municipal, recreational road, less than 5,000 population.
- Municipal, recreational road, more than 5,000 population.

Metropolitan Area

- High commuter route.
- Commuter shopper route.
- Low recreational route.

Seasonal adjustment factors, based on the data for the previous 3 years, are developed for each classification and are applied to all coverage counts collected at locations within that classification.

For the "non-count" years, a growth factor is applied to the previous year's data based on changes in counts at the ATR stations located on the same functional class of roadway. When new data are available at the end of the next count cycle, these data for the interim non-count years are readjusted to represent the average of prior and subsequent count years (e.g., a 1987 "non-count" year estimate based on the growth factor would be readjusted to represent the average of 1986 and 1988 counts at that location as soon as the 1988 count year was completed).

In developing AADT estimates for each section of roadway, there are sometimes road sections with no historical count data (e.g., lower order local facilities, including township roadways and local streets). In these cases, an original "baseline" estimate is based-on ATR counts on lowest order roadways with the lowest counted volumes. Growth factors for these uncounted sections are also based on this same ATR group.

MinnDOT also collects vehicle classification counts at about 300 sites per year. These are 16-h (e.g., 6 a.m. to 10 p.m.) manual classification counts usually over 2 different days. In addition, portable vehicle classifiers are deployed to collect 48-h data. Currently, there is no program to seasonally adjust the classification counts. There are an additional 25 weigh-in-motion stations statewide that collect classification data. However, these data are used less than the manual classification counts.

The new count data are placed in the Traffic file within the first 6 months of the subsequent calendar year. While the Traffic file can also be thought of as a "Section" file (with a specified AADT at the beginning count station being assumed constant over the entire section), it differs from the Roadlog file to which it will often be merged in that the beginning and end points (termini) are often located at different points on the roadway. The linking variables are again the route system/route number/reference point (milepost).

There are approximately 208,000 records on the file, but these do not represent a one-to-one match with the 200,000 "true" records on the Roadlog file. Often, there are Roadlog sections with multiple Traffic file records (i.e., multiple count stations), and often there are Roadlog sections with no Traffic file records (i.e., corresponding count stations) located within the section.

Each raw file record contains up to 30 years of AADT information (with the related year "attached"). Thus, to determine the average AADT for a given year for a series of sections on a given route: (1) the traffic section reference points must be matched with the appropriate Roadlog sections by comparing the reference point with the beginning and ending milepoint on Roadlog sections (with the ending milepoint being "assigned" as being equal to the beginning milepoint on the succeeding section), (2) the appropriate yearly AADT for each contained Traffic file record must be extracted, and (3) the counts must be averaged for sections where multiple Traffic file records exist. If no Traffic file record exists for a given Roadlog section, then the section AADT is assumed to be equal to the AADT at the previous (upstream) traffic section on the same route. (This is the assumption made by Minnesota and by HSRC programs. However, other procedures could be followed in calculating AADT if they are felt to be more appropriate for a given research question.) Any AADT assignment program developed must not carry over counts from one route to another; this is a mistake that can easily be made since the Roadlog file is a continuous file in route order. Obviously, averaging traffic over more than 1 year will require additional programming.

Currently, there are two HSIS SAS-formatted Traffic files — one developed for 1987 and earlier data, and one containing data for only 1988 and 1989. Again, please note that traffic data were merged with the Roadlog file for years 1987 through ***1994***. The Traffic file still remains a separate file on the HSIS system for the years 1987 through 1989. It is no longer available as a separate file on the HSIS system after 1989.

The first Traffic file (1987) is similar to the raw file in that it contains up to 10 years of data, with 1987 counts being the most recent data. The second file (1988-1989) contains only counts for 1988 and 1989. Each record on the file contains information on traffic counts for one year for a given location. To combine across years for a given counter location, records with the same location information can be merged.

To make the AADT information even more easily usable in subsequent analyses, HSRC developed a linking program that links the basic AADT information from the SAS Traffic file with the Roadlog file to produce a separate single "Average AADT" variable for each Roadlog section on each of the two Roadlog files (i.e., 1985-1987, 1988-1989). Where necessary, averaging across traffic sections in a given Roadlog section for a given year and "carrying down" AADT information from the prior record have been done in this linkage program. Since the 1987 Roadlog file is used with accident data from 1985-1987, and the 1989 file is used for 1988-1989 accidents, the AADT variable on each Roadlog file represents an average AADT over the respective time periods. That is, the 1987 file contains average AADTs for 1985-1987, and the 1989 Roadlog file contains average AADTs for 1988-1989. Different AADTs (say for individual accident years) could be developed by modifying the existing computer program.

Since it is not possible to perform an independent "check" of the accuracy of the AADT information, it is assumed that the procedure in place in Minnesota to monitor count stations and update the file provides adequate information. As indicated above, these are felt to be excellent data for the trunkline system where they are updated on a 2-year cycle. There are also fairly good data for the county State-aid systems, which are generally updated on a 4-year cycle. (Taken from the ***HSIS Guidebook for the Minnesota State Data Files.)***

Relating Accident and Exposure Data: Accidents are located by information on the route system, route number, and a "reference point." This information allows an accident to be attached to the appropriate section of the Roadlog file.

Accidents in an intersection can also be attached to the Intersection file by using route system and number, and the reference point.

Apparently, the approach from which a vehicle entered an intersection cannot be identified, except possibly by matching the direction of travel with the direction of the approach from the Intersection file.

North Carolina, HSIS

Coverage: The current HSIS files for North Carolina cover the years 1990-1995. Accidents are linked to the Roadway Inventory file with a computerized referencing system that currently covers about 38 percent of the estimated 148,056 total road kilometers in North Carolina. The reference systems covers all 22,530 km of primary routes, and an additional 33,473 km of secondary roads (rural secondary roads and city streets). There are no "county" roads in North Carolina, since all are under State control. This system links about 60 percent of the accidents (118,000 out of 192,000) to a road segment in the Roadway Inventory file.

Exposure Information: The Roadway Inventory file describes homogeneous road segments defined by a beginning and ending milepost. An AADT is provided with the year in which the count was taken and the section length in miles. The percent trucks in peak traffic is available for about 40 percent of the sections and an off-peak percent trucks is available for about 10 percent of the sections. The roadway variables include roadway width, number of lanes, lane width, shoulder type and width, median type and width, surface type, whether the section is in the HPMS sample, a traffic growth factor, and other variables.

Currently, intersection and interchange information cannot be linked with accident as the descriptive information is not available in a suitable format. The available information on roadway segments does not include information on horizontal curvature, vertical grade, or passing sight distance.

Traffic Data; As indicated above, the basic AADT and percent truck information is included on the Roadway Inventory file. The traffic count information used in the development of these variables is developed from a series of permanent control count locations and spot counts across the system. Currently, there are approximately 100 ATRs across the State. These are permanent full-time counters that are used both for counts at their location and to establish seasonal and growth factors used with spot counts from surrounding locations.

In addition to these permanent stations, there are approximately 60,000 points in the State where 24- to 48-h counts are made. The entire primary and Interstate system is covered each year. Fifty percent of the secondary roadway system is covered each year with the remaining 50 percent being done in the alternate year. The spot counts are linked with a group of nearby ATRs in order to establish distributional factors. The data are reviewed internally by the inter-office traffic staff, edited, quality control is checked, and then factors are developed. The traffic counts are closed out for the count year in October of each year and then sent to the roadway inventory staff for inclusion in the Inventory file.

Ramp counts are made each year on all interchange ramps on the Interstate system. These ramp counts are used to generate turning volumes and to balance counts on the mainline for the Interstate and crossing roadways. This represents approximately a 2-week count on each ramp. Past ramp counts are found on paper file, but have been computerized since early 1993.

Truck counts are made on a 3-year cycle at 300 vehicle classification sites across the State. The 300 count locations are not necessarily at all of the ATR sites. There are approximately 90 truck weigh stations in the State related to the SHRP program. In addition, it was noted that truck counts are made every 3 years on all HPMS sections in the State.

Finally, for intersections that are in the State's Traffic Improvement Program, turning counts are done on an as-needed basis. These turning counts include both a.m. and p.m. peak traffic, with each count being conducted for approximately 7 h. It is estimated that approximately 500 of these are done each year. These are found in a paper file, which may be computerized in the next 1 to 2 years.

Examination of the traffic-related variables in the HSIS Inventory file indicates that ADT is present for 99.9 percent of the sections. However, what is missing is data on percent trucks. Here, the variable concerning "Percent Trucks at Peak" is uncoded for approximately 60 percent of the mileage. The variable related to "Off-Peak Percent Trucks" is uncoded for almost 90 percent of the mileage. Conversations with department of highways staff indicated that this is the result of the fact that these variables are only coded if there is fairly high confidence in the percentages. This would occur if a classification count had been done on the section (as in an HPMS sample section) or on an adjacent or nearby section. Thus, while the data present should be fairly accurate, data are missing for a large number of miles.

Linking Accident and Exposure Information: The linking system for the accident data is unusual in that it is based on a "paper" reference system. The linkage information is the county, route, and milepost. However, there are no physical mileposts on the roads. The investigating officer records the distance and direction to a reference point that may be an intersection, bridge, or city boundary. Mileposts are determined in a computerized referencing system, based on the location of the reference given. The accident is linked by using the milepost generated by the computerized reference system to locate the section in the Roadway Inventory file, which includes this milepost within the beginning and ending milepost defining the section. Nearly all accidents on the primary road system are linked with this system, plus a large number of accidents on the secondary roads. About 90 percent of the mileage in the reference system is in rural areas. About 80 percent of the rural accident locations are believed to be accurate within 0.16 km, and 80 to 90 percent of the urban accident locations are thought to be accurate within 30.5 m.

Intersection characteristics are not currently available for linkage with the accident data.

Utah, HSIS

Coverage: Accident data for 1985-1994 are included, but highway data for 1990 are not available.

Of the 80,465 highway kilometers in Utah, 69,200 km are on the Roads file. However, only 20,599 km of these have inventory information and can be used for analytical purposes.

Exposure Information: The Roads file contains AADT for each section. Also given are the percentage of trucks in off-peak periods and the percentage of commercial vehicles in peak periods. No definition of "trucks" and "commercial vehicles" are given. Together with the segment length, VMT can be estimated.

No separate information for intersection exposure is available. The only information given for intersections is the number of intersections by segment, also separated by type of control. The intersecting roads are not identifiable.

For the State-controlled system, a Horizontal Curve file and a Vertical Grade file are also available. They allow disaggregation of exposure by grade and curvature.

For a random sample of sections of two-lane roads, a Cross Section file is available. It contains extensive information on cross-section and roadside features, including trees, posts, hydrants, recovery area, etc. This would allow the inclusion of highly specialized exposure measures, such as the number of trees passed, etc. Counts of accidents by severity are also given.

Traffic Data: As noted earlier, traffic data related to AADT and truck percentages are found on the Roadlog file. These data are based on Utah's traffic count program. In this program, there are 85 permanent ATRs on Interstate and Utah State roads that are in operation 365 days/year. Of these, 53 ATRs capture volume and vehicle classification counts and 32 ATRs count volume only. These ATRs conform with FHWA's HPMS guidelines. In addition, there are approximately 10 ATRs on roads inside National Parks in Utah that are operated by the National Park Service.

In addition to these permanent counts, Utah collects 48-h coverage counts at approximately 1,000 locations per year. Counts on the State-system roadway are done on a 3- to 5-year cycle. Approximately 100 traffic counting machines are used to collect traffic data for 11,426 km of State-system roads in Utah. In terms of coverage, Utah tends to have a better sample coverage of high-volume roads compared to lower functional categories. From a purely statistical perspective, a larger sample might be more appropriate for the lower functional classes of roads. However, Utah believes that limited resources for counting should be devoted to the roads that carry the bulk of the traffic. In addition to these coverage counts, approximately 100 short-term vehicle classification counts are conducted each year.

Short-term counts are expanded to AADT estimates using ATR data for roads with similar characteristics, functional class, and volume group. For a year in which no count is made, the previous year's count for a section is modified by a "growth factor" that is based on data from an "assigned" (similar) ATR station, area count data, and/or estimated statewide averages. In this manner, volume assignments are made to each section of State-system roadway each year. Finally, Utah staff also develop estimates of truck percentages and equivalent single-axle loadings (ESALs) for "on-system" roadways. Traffic information is entered into the Traffic file as it is being collected, but is transferred to the computerized system and, thus, to the Roadlog file at the end of the year.

With respect to the accuracy of the traffic information, Utah staff indicated that the data are currently being corrected so that errors would probably not be greater than ±10 percent for almost all of the sections. (Taken from the ***HSIS Guidebookfor the Utah State Data files***.)

Linking Accident and Exposure Information: Accident and highway files contain the route number and milepost, which allow linking of the data. Intersection accidents can be identified by a code based on the officer's intersection sketch. However, they cannot be linked to a specific intersection in a segment, except if there is only one in a segment.

Washington State, HSIS

Coverage: The current HSIS files for Washington State cover the years 1993-1995. Data for 1991 and 1992 will be added later when it is available. There are approximately 120,000 accidents per year in Washington State. Approximately 42,000 of these occur on State routes, and are location coded manually, based on the scene diagram and location information on the accident report. About 20 percent of these are "citizen" reports. Omission of these citizen reports reduces the located accidents on State routes to about 34,000.

A total of 13,840 km are described in the Roadlog file. This mileage includes 11,748 km of mainline roads, and another 2092 km of ramp front and other non-mainline roads. For example, information on each ramp for 876 interchanges is included. Interstate, U.S., and State routes are included. About 85 percent of the mileage is rural and there are about 1408 freeway kilometers. Each record describes a homogeneous section of road, as created by HSRC from point-by-point files supplied by the State. There are a total of 41,000 sections at an average section length of 0.3 km. Although the points at which intersecting roads cross are identified, there is not sufficient information (milepost) to link in the section data for the crossing road. Thus, the Washington State data do not appear well suited to an analysis of intersection accidents.

Exposure Information: The Roadlog file includes the beginning and ending mileposts and section length, the latter two calculated by HSRC. AADT is also given. By linking with the Traffic file, additional weekday and weekend counts are available, as well as single- and double-trailer truck volume. The available roadway characteristics include surface width, lane width and type, shoulder width and type, median information, functional class, posted speed, and other information.

The Traffic file created by HSRC describes road sections with approximately constant volume. The beginning milepost is identified, and the endpoint is found as the beginning milepost for the next record. However, one must check that the route has not changed. Additional section files describe 33,000 vertical grade sections and 14,600 horizontal curve sections. These can also be linked with the Roadlog file based on beginning and ending mileposts.

Traffic Data: As noted above, traffic count data captured on the Trips file, and thus in the HSIS system, contain a number of variables. These include AADT, average weekday volume, average weekend volume, single-trailer truck percentage, double-trailer truck percentage, and various peak-hour descriptive percentages. While AADT information has been merged into the HSIS Roadlog file to facilitate rate-based analyses, the other variables can be linked with the Roadlog file through linkage variables contained in both files.

In the base traffic file from which this information is derived, a new record is begun when there is a change in the AADT. The traffic census staff go through each of the inventory groups and identify what they feel are "discontinuities" along the routes in terms of volumes. These discontinuities would represent locations where the staff expect there to be significant changes in the AADT, such as an intersection with a significant turning volume or the location of a major traffic generator such as a shopping mall exit. In short, the Traffic file is a set of "homogeneous traffic sections." Thus, even though the file is organized as "point data" with only a "beginning" milepost, the data should not change until the next milepost. (In using and merging the file, some caution should be taken to ensure that the next milepost on the file is within the same route.)

The basis for the traffic information is a series of permanent and non-permanent count stations across the State. There are 117 permanent ATRs in the State as of December 1993; all 117 produced volume counts. Of these permanent count stations, 70 produced vehicle classification counts, 32 produced truck weight plus classification counts, 22 produced vehicle length counts, and 47 produced speed counts.

In addition to the permanent count stations, the traffic census staff conducts approximately 3,500 weekday counts each year. Each of these is a 72-h, Tuesday through Thursday count. Approximately 400 of these include additional vehicle classification counts each year. The counts are not always taken at the exact same sights, but do cover all HPMS locations as well as certain project counts that are conducted each year. In Washington State, there are 3,200 HPMS sections. The traffic staff feel that there are approximately 5,000 unique “homogeneous traffic” sections in the State each year. Counts are made at each of these locations every other year or every third year. In addition to these counts, there are ramp counts done at 120 to 150 interchanges each year.

With respect to accuracy and completeness, the DOT staff feel that they have very good data on approximately 90 to 95 percent of the roadway in the trips system. They feel that the least accurate information on the file is the vehicle classification counts. This is due to the limited number of count stations that are, by necessity, available for these type counts. However, traffic census staff are working toward increasing the accuracy of these truck counts. Their current feeling is that the variable related to daily truck percentage in the peak hour now contains good data. The overall truck count system was redone in 1987. One of the current points of interest is to try to expand the seasonal factors for trucks to make these even more accurate.

As noted under specific variable descriptions in the later format section, certain other variables (such as “Peak Hour Percentage” and “Peak Hour Split”) have significant numbers of uncoded (“zero”) locations. These represent locations where counts were not made or where old, erroneous counts have been deleted from the file. Washington State staff recommend carrying forward values from the preceding valid count location in these cases.

Linking Accident and Exposure Information: County, route, and milepost in the accident files can be used to create an 1 l-character variable that can be linked based on the route identifier and the beginning and ending mileposts in the Roadlog file. In the Traffic file, the beginning milepost is given, and the endpoint is assumed to be the beginning of the next record after checking that the route is the same.

Intersection volume and characteristics are only available for the mainline roads. Information for the crossing road sections cannot be linked.

Exposure Information in Highway Files

Highway files typically contain AADT for each segment in the file. Sometimes additional information is given, e.g., AADT for commercial vehicles or peak ADT. Together with the section length, AADT allows calculation of VMT on that section. If a segment ends at an intersection, AADT provides the number of vehicles entering and leaving the intersection from each approach. For an intersection within a segment, the same values must be assumed for the two approaches on this road.

In a formal sense, this provides enough information to calculate and analyze accident rates. However, if accident rates or accident counts in relation to AADT are used in statistical analyses, then the statistical characteristics of the AADT information in the files need to be known.

There are basically three types of accident studies:

(1) Making and comparing aggregate estimates.

(2) Studying relationships between accidents and highways and other factors using segments or intersections as observations.

(3) Identification of hazardous locations-"black spots."

The statistical characteristics of the AADT information affect these analyses in different ways.

The AADT values for the many sections of a highway file are derived from relatively few actual counts. At continuous counting stations, counts are made 24 hours a day, 365 days a year. At temporary counting stations, counts are made for usually 24 or 48 h, at intervals of 1 or several years.

There are two statistical questions: (1) what are the sampling characteristics of the actual counts, and (2) how are the AADT values for the sections without counts obtained from those for the sections with counts?

The answers to these questions determine the statistical analyses that can be validly performed with accident rates as dependent, or AADT as independent, variables.

To allow generalization beyond the sites with actual counts, sites should be randomly sampled from a well defined "frame," e.g., all sections on Interstate highways. This is often not done. Historically, "judgment" samples have often been made. Sites were selected that experts thought to be "typical" or representing the entire range of highway characteristics. While a judgment sample can give unbiased estimates, one cannot be certain of this. In particular, one cannot validly predict the errors of estimates based on judgment samples.

At the temporary counting stations, there is also sampling over time. If the counting is not done during certain parts of the year only, but year-round, sampling over time may be adequately close to random sampling.

Statistical analyses of a sample obtain estimates for the total sampling frame: totals or averages. In this application, it would be the number of all vehicles entering intersections on the highway network constituting the frame or the AADT representing an average over all sites on this highway network.

If the sample is stratified, then the estimates apply to each stratum separately, and estimates for all strata combined can also be obtained.

Such estimates can be used for studies of broad questions, e.g., comparing accident risks among highway systems, among highways with different numbers of lanes, classes, and intersections, etc. The level of detail such studies can consider is limited, because each stratum provides a single observation. However, if a detailed sampling plan is developed that stratifies according to many factors and their interactions, then even if the minimum of two sampling sites per stratum is used, detailed analyses may be possible.

One limitation of this type of analysis is that it does not allow identification of high-risk sites or "black spots." Highway data files contain information that, in principle, allows identification of such black spots, e.,g., AADT for short highway sections. With this information, an analyst can calculate accident risks for sections and intersections, and identify high-risk locations. However, without fully understanding how the AADT values for the individual sections are obtained from the relatively few sites with actual counts, the analyst cannot assess the statistical characteristic of the AADT values, and analyses based on them may be invalid. One approach is to assign to each section the value of the preceding section, until a section with an actual count is encountered, then carry over this count, etc. An alternative is to linearly interpolate AADT on the sections between connecting stations. While such approaches may give realistic order-of-magnitude estimates, and may even be quite realistic under certain conditions, this is not guaranteed. Thus, estimates of accident rates based on them can be biased and unrealistic. A more subtle, but not less important, aspect is that the estimates are not independent. Usually, the estimates on adjacent sections are positively correlated. A consequence is that analyses, which are using individual sections with their accident counts and AADT values as observations, tend to underestimate the uncertainties and errors of the results. They may also lead to the identification of "black spots," which appear to have unusually high accident risks only because the variability of the calculated rates is underestimated. Therefore, the statistical value of AADT figures by segment, without indication from which stations and by which method they are derived, is very limited.

Long-Term Pavement Performance (LTPP)

Historical Summary and Purpose: The Long-Term Pavement Performance (LTPP) program is a 20-year research project begun in 1987 as part of the Strategic Highway Research Program (SHRP). During the early 1980s, the Transportation Research Board (TRB) of the National Research Council, under the sponsorship of the Federal Highway Administration (FHWA) and with the cooperation of the American Association of State Highway and Transportation Officials (AASHTO), undertook a study of the deterioration of the Nation's highway system!) The SHRP was established on the recommendation of this study to focus research and development activities aimed at improving highway transportation. The Long-Term Pavement Performance program was one of six key research areas identified by this study.[2] The LTPP program is a comprehensive program to "satisfy the total range of pavement information needs" drawing on "technical knowledge of the pavements presently available and seeking to develop models that will better explain how pavements perform . . . this includes specific effects on pavement performance of various design features, **traffic and environment,** etc." The traffic and environmental data contained in the LTPP data collection plan are of potentially extreme interest as measures of exposure for highway safety issues as well. The concept of a traffic database, later named the Central Traffic Database (CTDB), originated in 1989 when the Expert Task Group concluded that the volume of traffic and load data that would be collected over the 20 years of the LTPP program required a separate database.

Data Contents and Structure: The LTPP data are housed in seven modules. A brief description of those modules that could be of interest in highway safety studies is described below:

(1) Climatic module.

Data derived from the National Oceanic and Atmospheric Administration (NOAA) weather data. Climatic data include site-specific estimates (based on the five closest weather stations) of various temperature, precipitation, humidity, and solar data statistics on a monthly basis for each test section, as well as actual values for the weather stations.

(2) Inventory module.

Data that identify the site and describe the pavement at the time the section was chosen. Data include location, material properties, composition, construction improvements, etc.

(3) Maintenance module.

Data describing all maintenance activities associated with the site.

(4) Monitoring module.

Friction, deflection, and distress data that could be of interest in wet pavement accident studies, etc.

(5) Traffic module (Central Traffic Database [CTDB]).

Historical and monitored traffic data. Yearly estimates of volumes, axle loads, and equivalent single-axle loads are available for each site. Also, data on truck weights and distributions are available at 789 sites quarterly for 7 days. Approximately 35 percent of these sites have weigh-in-motion collectors and the rest are Automatic Vehicle Classification (AVC) counters.

Experimental Design, Sample Plan, and Location Distribution: Data are collected in four geographic regions by regional staff members. With regard to traffic data, staff members are responsible for reviewing and processing the traffic counts, classification, and weight data, as well as ensuring acceptable collection procedures. The regional offices transmit their data to the national LTPP Traffic Database. Here, the data are further scrutinized and edited and it is the responsibility of this office to decide what data are of sufficient quality to release to the general public.

Traffic data are collected on more than 789 sites on key highway routes. In addition to new traffic data collection, historic traffic data were also requested where available. There are generally two types of traffic data available — vehicle count and classification data (Automatic Vehicle Classification [AVC] devices) and vehicle count and weight data (Weigh-in-Motion [WIM], either permanent or portable). The location of the WIM data collection may not always be exactly at the site, especially near interchanges. For the purpose of safety analyses, it is important that the researcher verify the exact location of the traffic data. These data have been of varying quality and one of the future objectives will be to back-validate some of the historic data with the new data, incorporating trends established based on the new data. Figure 1 show the geographic regions and Table 1 lists the number of locations by State for these locations. (Note: A revised table will be submitted that identifies locations that have WIM equipment and that have AVC equipment only when it is available).

Data Acquisition and Documentation: Information from the LTPP studies is available from the LTPP Information Management System (IMS), a database developed under SHRP. The pavement performance data are stored in the National Information Management System (NIMS) located at the TRB in Washington, D.C., and are updated on a regular basis. Similarly, the more detailed traffic data are housed in the CTDB and updated on a regular basis. Summary traffic data from the CTDB are periodically sent to NIMS for inclusion with the pavement performance data. These updates include corrections of previous erroneous data. Procedures and standards were established to ensure data quality, and extensive data quality checks are preformed throughout the collection and recording process. Information is also available indicating the level of data reliability. Although data are collected at the regional level and stored in Regional Information Management Systems (RIMS) and regional CTDBs, data are only released to the public after they have passed these checks and are stored in the national databases.

A guide that contains more details on the background and objectives of LTPP — what data are collected, how to request data, data formats, and examples of reports generated — can be found in reference 2. Complete information on how the data are collected, what quality checks are imposed, etc., can be found in other documents.

Data are released on two levels: (1) a sectional release and (2) an experimental analysis release. Data in Level 1 generally should be considered for analysis of a given test section, not comparisons across sections. These data have passed a minimum number of quality checks and, if used in analyses, should be used cautiously. Level 2 data have completed all assurance checks and are considered acceptable for analysis. Many quality control issues are still under development and consideration in an ongoing FHWA contract. Among these is the prospect of grouping sites into classifications according to the completeness of the traffic data at those sites. A classification being considered for the amount of data available is "preferred," meaning that at least 9 months of continuous data are available; "desirable" would mean that at least 6 months of continuous data are available; and "minimum" would mean that anywhere from 1 day to 6 months of data are available. Missing data can be due to lack of continuous WIM devices, equipment failure, etc. These classifications have not been set and could have changed by the time of this report. The researcher is referred to the periodic progress reports produced from this contract. The FHWA contact for this information is Kris Gupta. At this time, there is a limited amount of data available to the public, i.e., data that have passed Quality Assurance/Quality Control (QA/QC) checks. Although the plan is to have at least 50 percent of the data available by the end of 1995, the FHWA contact can best update the researcher on this.

Potential uses of the LTPP traffic data would have to focus on safety studies that are location based. For example, the question of "are double-tractor configurations overly represented in on-/off-ramp accidents as compared to singles?" might be addressed using the LTPP traffic data. First, it would be necessary to ascertain whether or not there are a sufficient number of LTPP sites with complete enough traffic data to supply enough accidents to do an adequate evaluation. Secondly, are accident histories available at these sites and over a sufficient time period? This would be the general process for examining the feasibility of using the LTPP traffic data (or any location-specific traffic database):

1. Formulate the hypothesis.

2. Determine what traffic data best represent the exposure for the data required to address the hypothesis.

3. Determine if there are sufficient sites of the type required by the hypothesis in the CTDB. How complete are the traffic data at these sites?

4. Determine whether accident histories are available and in sufficient numbers to justify the analysis.

These steps should be attainable using only a minimum amount of resources.

The only way to receive LTPP data from the national databases is to submit a complete LTPP Data Request Form to the TRB NIMS Administrator:

Penny Passikoff
National Academy of Sciences
Transportation Research Board
2101 Constitution Avenue, NW
Washington, D.C. 20418
TEL: (202) 334-3259
FAX: (202) 334-3495

Costs for obtaining the data include a $75 handling fee, media costs that depend on the type of media selected on the form, shipping costs, and any costs due to custom requests. State and Federal agencies and international participants do not have to pay the $75 handling fee.

References

(1) Rowshan, Shahed. ***Long-Term Pavement Performance Information Management System Data Users Guide.*** Federal Highway Administration, Report No. FHWA-RD-93-094, July 1993.

(2) Herman, John L.; Charlie R. Copeland; and W.O. Hadley. ***SHRP-LTPP Traffic Data Collection and Analysis: Five-Year Report.*** Texas Research and Development Foundation, Austin, TX. SHRP-P-386, 1994.

Nationwide Personal Transportation Survey (NPTS), FHWA

Purpose: The Nationwide Personal Transportation Survey (NPTS) provides nationally representative estimates of personal travel in the United States. All modes of transport are covered, including passenger cars, trucks, motorcycles, buses, trains, subways, airplanes, taxis, bicycles, and walking. The dataset includes information on demographic characteristics of the household, person-level information on the individuals participating in the survey, descriptive information on each vehicle in the household, and two levels of travel information. The first level of travel information is a detailed account of all trips taken on the survey day. The second level is information on trips longer than 12 1 km that occurred during the 2-week period immediately prior to the survey day. Travel information includes mode, vehicle type, road type, date of travel, time of day, trip purpose, origin and destination, elapsed time, and area type.

Source: The most recent NPTS (1990) was conducted by the Research Triangle Institute of Research Triangle Park, NC, under the sponsorship of the U.S. Department of Transportation?) A random sample of 26,172 households with telephones was selected by means of a random-digit dialing procedure, and almost 22,000 households responded. Responses were collected by means of a telephone interview. (Earlier surveys were done using in-home interviews.) Each household was assigned a 24-h travel day (defined as 4:00 a.m. on the travel day to 3:59 a.m. on the following day) and a 14-day travel period. The survey period was from March 1990 to March 199 1. Person-level interviews were conducted with all household members age 5 years and older. Trip-level interviews were conducted with all household members age 13 and older. The latter respondents supplied travel information on residents 5 to 13 years of age.

Coverage: The current file (1990) is the fourth in the series; earlier NPTS files are for 1969, 1977, and 1983. All personal trips, all modes of transportation, all purposes, and all 50 States and the District of Columbia are covered. Connecticut, the New York Metropolitan Planning Organization (MPO), and the Indianapolis MPO funded oversampling in their respective areas. The file includes weight variables, so that estimates of national totals can be computed.

Strengths: The NPTS file is the only source for national data on personal travel. Sample sizes are large, with 22,3 17 households, 48,385 persons, 35,152 licensed drivers, and 41,178 vehicles in the most recent sample. The survey design includes both driver and passenger travel, so vehicle occupancy rates can be analyzed. NPTS files are now available for 1969, 1977, 1983, and 1990, allowing trends over a period of 2 1 years to be analyzed. Efforts were made to maintain comparability of the major elements of the survey over that period. Travel can be broken down by region and for households in certain metropolitan statistical areas. Detailed information is available on the socioeconomic status of the household; age, gender, and other characteristics of the travelers; purpose of trip; type, make, and model of vehicle; and time, distance, and duration of travel. Interviews are conducted using computer-assisted telephone interviewing techniques, so many inconsistencies could be identified during the interview and addressed by the respondent.

Limitations: Road type is available only for a small subset of day trips. Sample sizes for commercial vehicles are small-the focus of the survey was on personal travel-so the NPTS is not useful for truck travel. The focus of NPTS is on national travel. It is possible to estimate the travel for regions of the country and for certain States and Metropolitan Sampling Areas (MSAs), but estimates for individual local areas, MSAs, or States may not be based on large enough sample sizes and may be imprecise. Households without telephones could not be included in the sample because the sampling procedure was based on a random-digit dialing procedure. In addition, the data are all self-reported.

Sampling Errors: Sampling errors can be calculated using appropriate software. See the User's Guide.

Access: The data are contained in six hierarchical files and can be obtained either as an EBCDIC file (similar to plain ASCII) or formatted for the SAS statistical analysis package. The files can be obtained on magnetic tape through the Volpe National Transportation Systems Center, Cambridge, MA, (617) 494-2450.

References

(1) ***User's Guide for the Public Use Tapes: 1990 Nationwide Personal Transportation Survey,*** December 1991, Report No. FHWA-PL-92-007.

National Truck Trip Information Survey (NTTIS), UMTRI

Purpose: The National Truck Trip Information Survey (NTTIS) provides national estimates of truck travel that can be cross-classified by truck configuration and loading, road type, area type, and time of day. Details on truck configuration and loading include cabstyle, number of trailers (if any), number of axles for each unit, empty weight and length for each unit, cargo body style, cargo type for each unit, and cargo weight for each unit. Road type is divided into three categories: limited access, U.S. and State numbered routes, and other roads. Area is classified using Federal Highway Administration definitions of urban or rural. The time of operation is classified as either day or night.

Source: The NTTIS was conducted by the Center for National Truck Statistics, part of the University of Michigan Transportation Research Institute (UMTRI).(1) The work was supported primarily by the Motor Vehicle Manufacturers Association, the Western Highway Institute, the Engine Manufacturers Association, and the American Trucking Associations. An initial sample of 8,144 trucks was drawn from registration files maintained by the R.L. Polk Company. The sampling frame was stratified by State and within each State, and by whether the truck appeared to be a tractor, straight truck, or unknown type. An interval selection procedure with a random start was used to draw the sample. Interviewers contacted current owners and operators of the vehicles by telephone to obtain a general description of the vehicle and company that operated it. Questions included estimates of annual travel that were checked against estimates from the TIUS.

A subsample of approximately 5,000 trucks was drawn for the travel survey. On four randomly selected days over a year, each truck was surveyed as to its use for the previous 24-h period. The survey method was to essentially follow the truck for 24 h. Survey staff collected information on the actual route the vehicle followed, cargo carried (if any) and where it was loaded or unloaded, and a complete description of the truck's configuration. The route was then followed on a map and the mileage was classified by road type, time of day, and urban/rural. All data were subject to extensive editing to ensure accuracy. To the extent possible and where necessary, inconsistencies and inaccuracies were cleared up by more phone calls to survey respondents.

Coverage: The NTTIS was a one-time survey. The sampling frame was trucks registered in the United States in 1983. The phone survey to collect the initial vehicle description and then the follow-up calls for trip information took place between November 1985 and February 1987. The file covers all medium and heavy trucks (GVWR > 4536 kg) registered in the United States, except for trucks owned by any level of government.

Strengths: Travel estimates can be cross-classified by truck configuration, loading, and operating environment — a level of detail unmatched in any other file of travel data.(2) It is possible, for example, to compare the travel of loaded and unloaded two-axle tank trailers on limited-access roads in urban areas at night. All data were carefully reviewed by editors experienced with the trucking industry. Ambiguous or unusual responses were clarified, where possible, with respondents. It is expected that the data are as accurate-as is feasible.

Limitations: Data are all self-reported, although subject to careful evaluation and consistency checking. Given the frequent contact between interview staff and respondents, and the ability to check responses, it is felt that the data are not systematically biased. Estimates from the file are all national. It is not possible to retrieve travel information for particular routes or even particular States. Moreover, by 1995, the file is clearly dated. There have been several important changes in the trucking industry since 1987 — for example, an increasing reliance on multiple-trailer trucks — that the file cannot reflect.

Sampling Errors: All sampling strata variables are included in the analysis file. Sampling errors can be calculated with appropriate software.

Access: The NTTIS file is a hierarchical dataset consisting of three parts: (1) a truck file with data describing the power unit, (2) a tractor trip file with data on trips by tractors, and (3) a straight truck file with comparable information about straight truck trips. The trip files contain one record for each trip taken by a survey vehicle on a survey day. Access to the data is provided through the Center for National Truck Statistics at UMTRI. Contact Kenneth L. Campbell or Daniel Blower at (313) 764-0248.

References

(1) Blower, Daniel and Leslie C. Pettis. ***National Truck Trip Information Survey.*** University of Michigan Transportation Research Institute, Ann Arbor, MI, Report No. UMTRI-88-11, March 1988.

(2) Massie, Dawn L.; Kenneth L. Campbell; and Daniel F. Blower. "Large-Truck Travel Estimates From the National Truck Trip Information Survey." ***Transportation Research Record No. 1407, Large-Vehicle Safety Research.*** Transportation Research Board, Washington, D.C., 1993, pp. 42-49.

Operational Exposure Data Sources

Historical Summary and Purpose: Researchers in the field of highway operations are often in need of exposure data in the form of both quantity of traffic and traffic congestion. Several researchers at Texas Transportation Institute were queried as to their knowledge of these data sources and the following reports resulted:

Kevin Balke's understanding is that the State of Texas (and probably others) has an extensive traffic monitoring program. His personal experience included collecting ADT volumes on many arterials and highways in major cities every 4 years. These studies were managed by local MPOs and these counts were published in a report. The Texas Department of Transportation maintains permanent count stations. A map is published annually with the AADT volumes displayed by location. However, none of this has been automated — this seems to be the major drawback in most operations study data sources. And, of course, there is the State roadway inventory file to which operations researchers often turn. Gerald Ullman relies on these State roadway inventory databases, as well as the State's ATR stations. With regard to urban area operations, some cities have systematic count programs and some do not, according to Ray Krammes. Dallas, for example, has a machine count program. Specific personnel in each city would serve as the contact for obtaining this information (in Dallas, it would be Ken Melston). State highway departments would probably be the best source for this information. In Dallas, the initial goal was to have manual counts on every 1.6-km segment of arterial road every 3 years. However, lack of funding seriously reduced this effort. Dallas still collects much of the data and stores 24-h and peak counts in a computer program and publishes two reports every January — one that lists the most recent count on each link and one that lists historical data, i.e., all counts on all links. Fifteen-minute counts could also be attained on paper copy. The only other city in the North Texas region that has some count data is Fort Worth. Most cities in the Metroplex do counts only on an ad hoc basis and generally hire consultants to do this work. In a review of Texas cities, this was generally the case (Austin, Houston, etc.). The counts are done on an ad hoc, nonsystematic basis for specific purposes.

It may be possible to design a highway safety research project using some of these site-specific count data. For example, Dallas would appear to have sufficient count data to address a particular urban problem. Consider the comparison of accident severities as a function of congestion — peak vs. off-peak times, weekend vs. weekdays, etc., or issues such as alcohol-related crashes in urban areas by time of day. However, due to the erratic nature of the data collection, one must be concerned about what biases such non-systematic data collection might be introducing into the safety analysis. Also, the fact that most data sources appear to be unautomated, at least in Texas, is a serious drawback.

For the most part, it appeared that operations researchers are interested primarily in very site-specific data and rely on ad hoc, often manual, procedures for obtaining exposure information. However, when they are interested in more global issues, they rely heavily on the Highway Performance Monitoring System (HPMS), described separately.

Residential Transportation Energy Consumption Survey (RTECS)

Historical Summary and Purpose: The Residential Transportation Energy Consumption Survey (RTECS) is a survey designed and administered by the Energy Information Administration (EIA). The objective of the survey is to obtain information on vehicles used for personal transportation in the United States. It is a companion survey to the Residential Energy Consumption Survey (RECS).

The first RTECS was done in 1978 and has been repeated triennially since 1985. The most recent survey for which published data are available is 1991. The following discussion relates to the 1991 survey. A survey was done in 1994, but the data are not available as of the date of this publication. The survey has been done five times. The RTECS is a follow-up survey and companion to the RECS. The RECS collects data on the households and includes preliminary information on the vehicles available to the household, while the RTECS consists of three stages in which additional data are collected on the vehicles available and the use of the vehicles by members of the household.

The data collected in the RTECS and RECS may have applicability in different areas of highway safety research. Primary data elements of interest in highway safety are the estimates of vehicle-miles of travel and the motor vehicle stock available to households for personal travel. These data elements may be linked to characteristics of households to allow computations concerning the amount of exposure (both vehicle-miles of travel and vehicle type) for similar households. Since the primary driver of each vehicle in sampled households was identified, as well as the age of the driver, the vehicle-miles of travel and vehicle used by age of primary driver may be estimated by implication. Since the data were not collected for trips by individuals within the household, the use of these estimates of exposure for different age groups may be questionable. It does appear the data are disaggregate enough for computing vehicle-miles of travel for households stratified by different household characteristics. This would provide a means for the estimation of exposure for those households and the applicability of those estimates to specific regions where similar stratifications of households could be obtained.

Data Contents and Structure: Household data collected in the RECS through personal interview that may be of interest in highway safety research include the following:

- Census region and division where household was located.
- Urban status of the household location (whether urban or rural area).
- Number of persons in the household.
- Data on the household composition (e.g., number with/without children, age of householder, etc.).
- Race of householder.
- 1990 family income (these were reported in nine different ranges).
- Number of drivers in household.

- Age and sex of primary driver for each vehicle in household.
- Average number of vehicles available to household during the year.
- Model year and vehicle type for vehicles available.
- Whether vehicle was used for commuting to and from work.

For the household data collected, data on the number of vehicles available and the vehicle-miles of travel for those vehicles were obtained. Vehicular data were not collected in the RTECS for motorcycles, bicycles, all-terrain vehicles (ATVs), and other related vehicles.

Experimental Design, Sample Plan, and Location Distribution: The focus of the RTECS is to obtain data on the vehicle-miles of travel, motor vehicle stock, and vehicle fuel consumption and expenditure data. Its companion survey (RECS) collects data on household energy consumption and expenditure. The sampling units in both the RECS and RTECS are households, with the universe being all housing units occupied as the primary residence in the 50 States and the District of Columbia. The sample of households selected in the 1991 RTECS was based on the 1990 RECS. The 1990 RECS was a multistage probability sample that incorporated a rotating panel to allow the observation of changes in energy use over time for households that fall in successive panels.

The 1990 RECS initial sample consisted of 6,757 units. Of these units, 848 were found to be ineligible for reasons such as the dwelling being uninhabitable, currently vacant, or used for seasonal occupancy. Energy-related data were collected from 4,828 households via telephone interviews, and an additional 267 units were surveyed through a mail follow-up, for a total of 5,095 responding households. The RTECS sample of households was selected from the 5,095 housing units that responded to the 1990 RECS survey. The number of RECS housing units selected for the RTECS survey was 3,045. Of those units, 2,842 were contacted by telephone and 200 were identified as households that had to be contacted by mail. The number for contact by mail was subsequently increased to 485 due to an increased number of households with unlisted or disconnected telephones.

The RTECS data collection effort consists of four phases, with the first phase being done in conjunction with the RECS. The first phase (during the RECS personal interviews) collected data on the household's vehicle stock, the vehicle identification numbers (VIN) of the vehicles, and initial odometer reading for each vehicle. The subsequent three phases were conducted at the beginning of the year (B-O-Y), mid-year (M-Y), and the end of the year (E-O-Y). These data collection efforts were done by telephone interview and, where this was not possible, the data were collected via a mail questionnaire. The B-O-Y and E-O-Y phases updated the data on the vehicle stock and collected data on the vehicle characteristics (including the vehicle make, model and model year, the vehicle odometer readings, and VIN). The M-Y phase was an inventory update where respondents were asked to complete a vehicle update worksheet and keep it for use during the telephone interview or mail it back if the household was classified as a no-telephone household.

The data collected during the RTECS allow for the computation of actual vehicle-miles of travel from the recorded odometer readings. These data represent total travel between the two points in time (i.e., B-O-Y and E-O-Y). Data were also collected on the disposition of vehicles and acquisition of new vehicles during the survey period.

Quality of Data: The data collected in the RECS and RTECS appear to be of relative high quality. Since the surveys produce estimates based on randomly chosen subsets of the entire population of occupied housing units, the estimates will always differ from the true population values and will include sources of nonsampling and sampling errors. The following sections discuss various sources of potential error in estimates produced from these surveys:

Noncovered Residential Vehicles. Since the sample of households surveyed in the RTECS were selected from the RECS, any household excluded from the RECS would not be represented in the RTECS, and the subsequent survey data would not include vehicles available to those households. Specifically, those families or individuals not included in the RECS were those living in group quarters such as college dormitories, military barracks, or large boarding houses; those living in recreational or other types of vehicles; and those with no fixed address. The effect of these exclusions is an underestimation of the total number of vehicles and related data.

Date of Reference for Survey. Since the survey design requires households to be followed for an entire year, changes in household structure and composition may not be accurately reflected. For example, the survey sample may have an overrepresentation of older established households and an underrepresentation of new households or families. Resulting estimates of vehicles and related data may have a negative bias induced by established households separating and only one portion being followed in the RTECS, vehicles acquired by household members that leave the household are not captured in the survey, and the total estimated households (used for expansion) is based on the July 1991 Current Population Survey (Bureau of the Census).

Item Nonresponse. Item nonresponse refers to the inability to collect full information when respondents either do not know the answer or refuse to answer selected questions. It can also occur when an interviewer fails to ask a question or record an answer. In the RTECS, item nonresponses were imputed to provide an estimate of the most probable response. Three techniques were used: hot-decking, predictive mean matching, and regression.

Hot-decking is a technique by which a household is randomly selected and its response to the missing data item is used as the response for the household with the missing item. The items imputed in the RTECS by this method were pre-1975 vehicle characteristics and fuel grade. Household demographic items, such as family income and ethnic background, were also imputed by this method for the RECS.

Predictive mean matching was used for imputing changes in vehicle stock for households not followed for the complete duration of the RTECS. In the 1991 RTECS, 26 percent (i.e., 795 households) were not followed for the entire year and imputations were computed to estimate the number that acquired and/or disposed of vehicles during the year. For households with no vehicles that were lost, a hot-deck procedure was used to impute the changes in vehicle stock.

Multiple regressions were used to impute annual vehicle-miles of travel for those vehicles that were imputed as being acquired. Linear and multiple regressions were also used for estimated annual mileage for vehicles where two odometer readings were not obtained in the survey. For 26 percent (i.e., 1,576) of the sample vehicles, no odometer span was available. An estimate of the annual vehicle-miles of travel had been obtained from the respondent during the RECS interview. Vehicle-miles of travel were imputed from a regression on the estimate of vehicle-miles of travel obtained in the RECS interview. For an additional 19 percent (i.e., 1,150) of the sample vehicles, no odometer span was available and an estimate of annual vehicle-miles of travel was not obtained in the RECS interview. Estimates of vehicle-miles of travel for these sample vehicles were imputed using a multiple regression using number of drivers, household income, age of household head, type of vehicle, and use of vehicle on the job as independent variables. This same method was used for imputing the vehicle-miles of travel for vehicles that were imputed as being acquired and/or disposed. Various other adjustments to the vehicle-miles of travel data were necessary to put each in terms of the same time period. Data from the Federal Highway Administration on monthly vehicle-miles of travel were used for this purpose.

Potential Problems: The RTECS data provide reasonable estimates of vehicle-miles of travel for households and vehicle types. These data will produce reasonable estimates of exposure relative to household estimates and estimates by vehicle type. However, the data do not include travel by motorcycles, bicycles, all-terrain vehicles, or similar types of vehicles, which may be critical in safety analyses. In addition, the data do not relate vehicle-miles of travel to person-miles of travel. The data are collected for vehicles and related to the households that own or have those vehicles available. While the exposure may be computed for vehicles in terms of type and vehicle-miles of travel, the data do not indicate the number of persons that may be in the vehicle on an average basis. Other data sources on average vehicle occupancy would have to be used to impute that estimate. The use of the data to compute exposure estimates by age of individuals would have to be based on the implication of primary driver for each vehicle in the survey. This is a relatively weak implication and is not considered an accurate estimate. Thus, it is not considered appropriate to use data from this source for estimating exposure for persons by age.

Data Acquisition and Documentation: Data from the RTECS and RECS are available in a variety of media. The following published reports may be purchased from the Government Printing Office (GPO):

- ***Household Vehicles Energy Consumption** 1991*; December 1993, DOE/EIA-0464(91) ☎☠□ GPO Stock No.).

- ***Household Vehicles Energy Consumption** 1988*; February 1990, DOE/EIA-0464(88), GPO Stock No. 061-003-00652-3.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns** of **Household Vehicles,** 1985*; April 1987, DOE/EIA-0464(85), GPO Stock No. 061-003-00521-7.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, 1983***; January 1985, DOE/EIA-0464(83), GPO Stock No. 061-003-00420-2.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, Supplement: January 1981 to September 1981***; February 1973, DOE/EIA-0328, GPO Stock No. 061-003-00297-8.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, June 1979 to December 1980***; April 1982, DOE/EIA-0319 (No GPO Stock No.).

The above documents are not the only ones available, but are considered to represent those report data that are of interest to highway safety engineers. In addition to the published reports, data tapes and diskettes may be ordered directly from the National Technical Information Service (NTIS). Information on how to order these may be obtained by telephoning NTIS at (703) 487-4807, FAX number (703) 321-8547. Detailed technical questions on topics of interest to highway safety engineers may be addressed to the following:

RTECS Manager	Ronald Lambrecht	(202) 586-4962
Vehicle-Miles of Travel	John Pearson	(202) 586-6160
Trends in Household Vehicle Stock	Ronald Lambrecht	(202) 586-4962

References

(1) ***Household Vehicles Energy Consumption 1991***; December 1993, DOE/EIA-0464(91) (No GPO Stock No.).

Truck Inventory and Use Survey (TIUS), Bureau of the Census

Purpose: The Truck Inventory and Use Survey (TIUS) is one of a number of economic censuses performed by the U.S. Bureau of the Census. It is designed to provide information on the population and use of trucks for government, business, industry, and the general public. The TIUS is conducted every 5 years. The most recent data year currently available is 1992.

The TIUS provides annualized estimates of the primary uses of trucks. Data include a physical description of the truck (axle count, cabstyle, cargo body style, overall length, empty weight, typical loaded weight, maximum loaded weight); a general description of the industry in which the vehicle is used; and a breakdown of the vehicle's use over the course of a year. For example, respondents report any placarded hazardous materials carried in the vehicle and then estimate the percentage of the total annual travel in which hazardous materials were carried. Similarly, respondents estimate the proportion of annual travel accumulated off-road, less than 80.5 km from the truck's home base, 80.5 to 321.9 km from base, and more than 321.9 km from base.

The TIUS is useful for estimating broad categories of annual truck use. Given the way the data are reported, however, it is not possible to break down or cross-classify travel estimates by road type, area type, or any other feature of the operating environment. It is also not possible to estimate travel by State, month, or season.

Source: The TIUS is a stratified probability sample of trucks registered in the 50 States and the District of Columbia. Within each State, trucks are stratified by body style. Within each stratum, a fixed number of trucks are sampled randomly. Roughly 3,000 trucks are sampled per State. Survey forms are then mailed to the registered owners of the sampled trucks. By law, the surveys must be completed and returned. The data are all self-reported and are all estimates of use for a particular year. Reports are subject to computer editing. Apparently erroneous responses are reviewed and corrected, if possible.

Coverage: The sampling frame for the TIUS covers all vehicles registered as trucks in the 50 States and the District of Columbia. This includes pickups, small vans, and other utility vehicles registered as trucks. The file excludes vehicles owned by any unit of government, passenger -- vehicles, ambulances, buses, and motor homes. Vehicles used exclusively off-road do not have to be registered, and thus are also excluded.

Strengths: The TIUS has a very large sample size. Roughly 154,000 vehicles were selected for the survey in 1992. Nearly 132,000 trucks are represented in the file. Estimates of population totals and annual travel from the TIUS have been compared with estimates generated by other techniques (e.g., NTTIS; for a description of NTTIS, see the discussion in an earlier section) and are in general agreement. Data collection procedures and survey questions have been fairly stable for a number of surveys, so comparisons among survey years are valid.

Limitations: The main limitation in the use of the TIUS file for safety-related exposure data is that the data represent typical or primary use only. Consequently, configurations that represent secondary use, such as bobtails or doubles, are not represented at all or are under-estimated. There is very little ability to cross-classify the travel estimates by operational characteristics that are known to be associated with differences in accident-involvement risk. For example, straight trucks do a large share of their travel in urban areas and on non-limited-access roads. Tractor-semitrailer combinations accumulate a much larger fraction of their travel on limited-access roads, which are typically the safest in the highway system. The TIUS data do not provide any means of controlling for such environmental confounding factors.

Sampling Errors: Variables representing the sampling strata are not released with the file, so it is not possible to calculate sampling errors for particular estimates. However, the published ***Census of Transportation*** includes an appendix with equations for approximating relative standard errors.

Access: Available on CD-ROM from the Bureau of Transportation Statistics and from Customer Services, Bureau of the Census, Washington, D.C. 20233. The data are the raw records from the survey, modified to limit the possibility of identifying particular individuals or businesses.

State Weigh-in-Motion (WIM) and Automatic Vehicle Counting (AVC) Devices

Historical Summary and Purpose: Truck weighing equipment is required for meeting a wide variety of public, private, and institutional needs. In the public sector, there are two major functional areas of application of these devices: data collection and enforcement. Statistically representative truck weight data are collected and used as the primary basis for engineering analyses and decisions related to planning, funding, design, operation, maintenance, and management of highway facilities. Measurements of the weights of individual trucks are needed to provide enforcement agencies with the capability to protect the highway infrastructure from damage due to unexpectedly high loads. In both data collection and enforcement, it is necessary to weigh large numbers of individual trucks.

A weigh-in-motion (WIM) system is used to attempt to approximate the gross weight of a vehicle or the portion of the vehicle weight carried by a wheel, an axle, or a group of axles by measuring, during a short time interval, the vertical component of dynamic (continually changing) force that is applied to a smooth, level road surface by the tires of the moving vehicle. Although the weight of a vehicle does not change as it moves over the surface of the road, the dynamic force applied to the roadway surface by a rolling tire on a vehicle varies dramatically when the tire/wheel mass accelerates vertically. This acceleration can be induced by roughness in the road surface and/or by an out-of-round or out-of-balance wheel/tire assembly.

Data Contents and Structure: WIM data are collected in the United States by the States under three programs. One is specified and required by the FHWA under the provisions of its ***Traffic Monitoring Guide*** (TMG). The States have designated and collected data at approximately 1,400 WIM sites in the United States. The data are stored as individual truck records by the individual States and are transmitted to FHWA.

Additional WIM data are obtained under the Long-Term Pavement Performance monitoring aspect of the Strategic Highway Research Program. Data are acquired quarterly for 7 continuous days at 777 sites throughout the United States and are transmitted to regional SHRP contractors.

The last type of WIM data is collected at truck weight enforcement stations during the weighing and sorting of trucks to determine whether they exceed legal limits. These data are not normally retained.

Each State is required to submit vehicle classification and truck weight data to the FHWA either annually or quarterly. Where continuous weigh-in-motion data are available, 1 week of data per quarter is required. These data provide input to national databases that are maintained by the FHWA. These databases include the Traffic Volume Trends System and the Vehicle Travel Information System. The Traffic Volume Trends System is a database management system that is based on state-supplied ATR data. The Vehicle Travel Information System is a microcomputer database management system that validates, summarizes, and maintains vehicle classification and truck weight study data. Tables 1 through 3 contain State-by-State information on the number of WIM sites, type of equipment, level of monitoring, the existence of historical data, and monitoring frequency. Level of monitoring refers to the amount of data collected. The preferred, minimum, etc. categories are the ones described in the LTPP traffic data, although these may not be the levels adopted by the CTDB.

Table 1. Region 1 WIM.

STATE	NO. SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA
Illinois	18	GK Instruments 6000 AWACS	Preferred	Y
Indiana	18	IRD Bending Plate	Preferred	Y
Iowa	12	GK Instruments 6701	Preferred	Y
Kansas	17	GK Instruments 6701	Preferred 1, Desirable 16	Y
Kentucky	7	Unknown (Portable)	Preferred 1, Minimum 6	Y
Michigan	13	GK Instruments 6012 (Piezo)	Preferred	Y
Minnesota	24	IRD Bending Plate	Preferred 21, Unknown 3	Y
Missouri	20	IRD 100 and GK 6701	Minimum	Y
Nebraska	15	Golden River Portable	Minimum	Y
North Dakota	4	GK Instruments 6701	Preferred	Y
Ohio	11	Pat Equipment	Preferred	Y
South Dakota	9	In-House Bridge WIM	Preferred	Y
Wisconsin	16	Pat Equipment	Preferred 5, Minimum 11	Y

Table 2. Region 2 WIM.

STATE	NO. SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA	MONITORING FREQ
Alabama	18	Bending Plate and Piezo Cable	Preferred 1, Desirable 17	Y	1 continuous, rest 7 days per season
Arkansas	14	Cap Pads and Piezo	Preferred 1, Desirable 13	Y	1 continuous, rest 7 days per season
Florida	29	Portable	Desirable	Y	7 days per season
Georgia	23	Cap Pads and Bridge	Preferred 2, Desirable 20, Minimum 1	Y	2 continuous, rest 7 days per season
Louisiana	2	Cap Pads	Desirable	Y	7 days per season
Mississippi	25	Piezo	Preferred	Y	Continuous
New Mexico	12	Cap Pads	Desirable	Y	7 days per season
Oklahoma	21	IRD Piezo	Preferred	Y	Continuous
Puerto Rico	4	Cap Pads	Desirable	Y	7 days per season
South Carolina	9	Portable	Minimum 8, Below Minimum 1	Y	Seasonal
Tennessee	15	Piezo	Preferred 2, Desirable 13	Y	2 Continuous, rest 7 days per season
Texas	90	Cap Pads	Below Minimum	Y	27 days annually

Table 3. Region 3 WIM.

STATE	NO.SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA	MONITORING FREQ.
Alaska	6	IRD	Preferred 5, Continuous 1	Y	Preferred
Arizona	25	Portable	Minimum	Y	Seasonal 7 day
California	37	Pat	Preferred 3, Continuous 15, Minimum 11, Below Minimum 8	Y	Continuous or seasonal
Colorado	16	IRD	Preferred	Y	Preferred
Hawaii	4	IRD	. Minimum	Y	Seasonal 7 day
Idaho	13	Portable	Preferred 1, Continuous 12	Y	Seasonal 7 day
Montana	7	Portable	Below Minimum	Y	Seasonal 7 day
Nevada	8	Portable	Preferred 1, Minimum 7	Y	Seasonal 7 day
Oregon	11	Pat	Minimum	Y	Seasonal 7 day
Utah	14	Portable	Minimum 2, Below Minimum 12	Y	Seasonal 7 day
Washington	19	IRD	Preferred	Y	Preferred
Wyoming	14	Pat	Minimum	Y	Seasonal 7 day

Experimental Design, Sample Plan, and Location Distribution: Each State determined their own experimental design and determined the number and location of the sites based on differing economic and policy-making factors. When using WIM data from any State for highway safety evaluation purposes, the researcher should contact the respective State's DOT and request specific information regarding site-selection criteria.

Potential uses of the WIM databases must be location-oriented, similar to the ones described for the LTPP WIM.

Data Acquisition and Documentation: Data from the national database must be requested from the FHWA directly. These data include: station description data, traffic volume data, vehicle classification data, and truck weight data. Each type of data has its own individualized record format. All data files are in ASCII flat files.

Individual State data can be requested of the individual State DOTs. The formats will vary. For example, Illinois currently has 18 active WIM sites dispersed throughout the State. The WIM system has not consistently provided the necessary data to the national database due to hardware and/or software problems. Illinois DOT collects data biweekly and stores all data that are required by the FHWA. The data are processed and kept on the mainframe computer in a hexadecimal format. Their data on the continuous count ATR network are located at 21 sites. These data provide vehicle count and classification data and are kept on personal computers in ASCII format.

Washington State DOT has 41 active WIM sites — 5 use bending plates and the rest use piezoelectric sensors. The sites are continuous monitoring sites and the data are downloaded weekly. The data provide the standard vehicle classification and truck weight data required by the FHWA. The data are converted by the State from 13-bin to 4-bin format for storage on a mainframe computer. Data from 1990 to the present are available.

Reference

(1) Parsons, Brinckerhoff, Quade & Douglas, Inc. And URS Consultants, ***Inc. Pavement Damage Factors Derived From Weigh-In-Motion Data Measured by Portable vs. Permanent Systems.*** Florida Department of Transportation Statistics Office, Traffic and Roadway Data General Consultant Task Work Order Number 4, Sub-Task 3.2, December 1993.

3. EMERGING EXPOSURE DATA SOURCES

Emerging exposure data sources are new sources or existing sources that have not been traditionally used to derive exposure estimates. Three areas were reviewed for possible emerging exposure data: Intelligent Transportation Systems, transportation planning surveys, and traffic volume data collected by the States. The scope of each area is described briefly in the following paragraphs.

Intelligent Transportation Systems (ITS)

Within the broad Intelligent Transportation Systems (ITS) area, three subareas were examined: Commercial Vehicle Operations (CVO), Advanced Traveler Information Systems (ATIS), and Advanced Traffic Management Systems (ATMS). Specific projects in the CVO area are the Crescent project in the western States and Advantage I-75 in the east. Each includes some automatic provisions for trucks to communicate various required information about the vehicle and driver, such as license status, vehicle permits, and inspection data. These are all multistate projects intended to minimize the stops a truck needs to make to demonstrate compliance with all the applicable regulations. Since the information is recorded electronically, there may be some way to get descriptive information and counts that could be used as exposure measures. Similar potential to gather exposure data may be present in the other two ITS areas reviewed.

Transportation Planning Surveys

The second area covers a range of transportation planning surveys. These are usually household surveys conducted by mail or telephone. Examples are the Transportation Planning Package of the U.S. Census (CTPP). This survey provides nationwide data that form the basis for many State and local transportation planning efforts. However, only trips to and from work are included. The other general source in this area is regional planning surveys. These are also household surveys patterned after the CTPP. The geographic coverage is limited, of course, but more detailed information is frequently collected, often for a broader range of trip purposes than just travel to and from work.

Traffic Volume Data - Errors of VMT Estimates Based on Traffic Counts and Section Length

The third area reviews the traffic volume data that are available from many States, and that form the basis of the traffic volume data in HSIS. Most traffic volume data are collected by State and local highway departments. Consequently, we need a good understanding of the accuracy and timeliness of the available data. How often are the counts actually taken at the site and, if taken some distance away, how accurate will they be for the site in question?

The remaining material is organized under these three headings.

Intelligent Transportation Systems (ITS)

The development of Intelligent Transportation Systems (ITS) technologies and services offers new opportunities to obtain exposure information. Since the primary objectives of ITS are not related to exposure data collection, it is important to recognize such opportunities and identify processes by which exposure data could be obtained. This section explores possible exposure data sources within the commercial vehicle operations portion of ITS.

Commercial Vehicle Operations (CVO) of ITS

Commercial Vehicle Operations (CVO) has been divided into six user services:

- Commercial vehicle electronic clearance.
- Automated roadside inspections.
- Commercial vehicle administrative services.
- On-board safety monitoring.
- Hazardous material incident response.
- Commercial fleet management.

Of these services, commercial fleet management, commercial vehicle administrative services, and commercial vehicle electronic clearance have potential as sources of data on commercial vehicle exposure in terms of vehicle-miles traveled over specific types of roads by various categories of commercial vehicles. There is also a possibility of applying some of the technology being developed for ITS research purposes to collect specialized exposure data.

Vehicle tracking systems for commercial fleet management that keep dispatchers appraised of the current locations of all their fleet vehicles could provide a source of exposure data. Such a system would need to include an automatic vehicle location (AVL) system, probably a global positioning system (GPS) and map matching software that would locate the vehicle on a map. If the system could preserve the history of travel of an individual vehicle over the course of the trip, the equivalent of a trip diary could be generated for every vehicle in a fleet with such a system The record of the configuration and cargo of the commercial vehicle for the trip could also be included in the trip record. The data file from the individual records could yield the miles traveled by each vehicle by road class and by vehicle configuration for the fleet.

A problem with commercial fleet management systems as sources for exposure data is that the data would be collected by the motor carriers. They might prefer to treat this information as proprietary and would not be willing to share this information with others. Even if some fleets decide to share this information with researchers, there may still be a problem with obtaining cooperation from enough fleets of appropriate sizes and diversity for a desired sample.

Another application of CVO systems that might overcome the problem with proprietary information is the commercial vehicle administrative process. States need to know the mileage of commercial vehicles on their roads for the purpose of fuel tax allocation. A specific system currently being tested in Iowa for this purpose is the on-board automated mileage system. The system uses GPS vehicle location technology and map-matching algorithms and software to determine the mileage a given commercial vehicle equipped with the system has traveled within a State. The map-matching algorithm identifies the route traveled. This information is transmitted electronically to the State authorities.

This will give the State a database from which mileage by commercial vehicles of various types on various types of roads can be obtained.

This seems like a promising source of exposure data. It is reasonable to assume that all States will eventually go to automatic systems of collecting commercial vehicle mileage information for fuel tax allocation. The system will also streamline reporting and paperwork for the carriers and they may be willing to install the units in their fleets.

The electronic vehicle clearance services identify a vehicle at a point, but do not track it over a route. These services will enable transponder-equipped trucks to have their safety status, credentials, and weight checked at mainline speeds. Vehicles that are safe and legal and have no outstanding out-of-service citations will be allowed to pass the inspection/weigh facility without delay. To use this system for collecting exposure information, a researcher would have to follow the vehicle from one inspection station to the next. There is currently much work being done on transponders that have "read-write" capabilities. Thus, a commercial vehicle passing through the inspection station could have the unique identification of the station recorded or the station could keep the record of the identification of the vehicle that passes through. If the vehicle kept a record of stations visited, the information would have to go into map-matching software to get the routes and then be entered into a database. If the stations kept the records, then the station data would have to be processed to find the paths of the vehicles and develop the vehicle mileage. The system, as conceptualized here, would be computationally challenging and does not appear to be a promising source of exposure data.

One of the technological developments brought about by ITS is better motion detectors, which were needed to study the actual paths, speeds, and accelerations of vehicles performing maneuvers in traffic. This information is needed to understand the micro-behavior of vehicles in traffic, which, in turn, is needed to design ITS systems.

There is a potential for using this advanced motion-detection technology together with WIM systems to collect information about the distribution of centers of gravity of commercial vehicles. Center of gravity is a surrogate for roll stability of vehicles and its distribution and exposure are often desired in analyses of rollover accidents.

The measurement of the center of gravity of a truck could be obtained by having the vehicle travel over a superelevated curve (of known superelevation) with a WIM system. The motion-detection system would precisely follow the vehicle's path and determine the radius of curvature of the vehicle's tires and also the record of the velocity over the path.' The forces acting on the vehicle would be measured at certain locations by the WIM. The information is sufficient to determine the vehicle's center of gravity, which would be calculated by microprocessor.

The center-of-gravity information would be recorded for each vehicle that passes over the instrumented curve. Information on the vehicle type could also be read from the vehicle's bar code or by an automatic vehicle identification system and could be added to the record. It is conceivable that a series of such stations could be built at sites selected by a sampling design to get the distribution of roll stability of commercial vehicles.

Advanced Traveler Information Systems (ATIS)

Advanced Traveler Information Systems (ATIS) provide the motorist with highway maps and other traffic and geographic information. For example, if a car is equipped with a map-based route planning system, this system might retain information on the route followed and provide more accurate data of the type that is traditionally sought through a trip diary. Speeds and travel times might also be incorporated.

Route guidance is a feature that holds the best potential for exposure data. At the basic level, route guidance is a static map. The map can be used to plan routes and provide directions to a destination. More sophisticated features would combine certain real-time (or dynamic) information on congestion, construction, and alternate routes with the map display. Route guidance (or navigation) systems may be either mobile- or infrastructure-based. "Mobile-based" means it is self-contained in the vehicle, while "infrastructure-based" implies that the capability resides in a central location and the information is communicated to the vehicle. The navigation capability requires position determination. The system must be able to track the position of the vehicle on a real-time basis using GPS or other methods. This is true for both the mobile- and infrastructure-based systems. A current program supported by FHWA is the In-Vehicle Routing and Navigation System (INRANS).

The attraction for exposure measurement would be the capability of the system to store the actual route followed by the vehicle. Traditional survey methods have drivers keep a diary to record where they went and when. This would provide much more accurate information. In principle, the travel could be linked with roadway characteristics, vehicle characteristics (including perhaps cargo weight and type for trucks), and driver characteristics. A sampling scheme to select vehicles and days could provide representative data for any geographic region, or vehicle or driver population.

ATIS may have a very different implementation in the trucking industry. Although some independent operators may be interested in a route planning system like that being developed for passenger cars, fleets are more likely to be interested in tracking systems that keep dispatchers appraised of the current location of all vehicles. A communication capability may also be part of such a system. Such a tracking system might also be able to preserve a history of the travel of individual vehicles. Information on the vehicle status and condition might be communicated back to the system over the course of the trip. Again, the equivalent of trip diaries may be generated for every vehicle in a fleet with such a system.

Advanced Traffic Management System (ATMS)

Historical Summary and Purpose: ITS technologies offer considerable improvements in data collection and dissemination in all areas of transportation. They are promising sources of exposure data for highway safety analyses. To date, however, little attention has been given to this application of data from ITS sources. The principle guiding documents for ITS developments in the United States — IVHS America's ***Strategic Plan for Intelligent Vehicle-Highway Systems in the United States,*** and the U.S. Department of Transportation's ***IVHS Strategic Plan: Report to Congress*** — make scant mention of the potential for integrating data from Intelligent Vehicle-Highway Systems (IVHS) sources into highway safety databases. FHWA is currently evaluating proposals for the national ITS system architecture study. Highway safety applications are addressed in the system architecture study to ensure that the architecture accommodates these applications. Therefore, the results of this proposed study are urgently needed.

Several opportunities for extracting exposure data from IVHS technologies are readily identifiable:

- Roadway-based exposure data from improved traffic surveillance systems.
- Vehicle-based exposure data from improved commercial vehicle monitoring systems.
- Individual-based exposure data from proposed route guidance systems.

Advanced Traffic Management Systems (ATMS) are the foundation for ITS, and more accurate and widespread surveillance of traffic conditions is a keystone of advanced traffic management. ITS America has proposed a long-term (20-year) goal of 30,577 km of freeway and 64,372 km of urban arterial roadways covered by surveillance systems. These systems will provide more accurate traffic volume data on the most important roadways in the major metropolitan areas of the United States.

The Commercial Vehicle Operations (CVO) component of ITS is a promising source of exposure data for large trucks. Since commercial vehicle applications will be one of the earliest areas of ITS implementation, this area deserves special attention in the proposed research. Automatic vehicle identification, classification, and location systems will become more widespread in commercial vehicle fleets. One application of data from these systems that will be the subject of an operational test during the next several years is the use of these data for determining vehicle-miles traveled in a State for taxation purposes. The same data would be a valuable measure of exposure for highway safety analyses.

One feature of the Advanced Traveler Information Systems (ATIS) component of ITS is in-vehicle route guidance, which requires a communications link between individual travelers and the centralized traffic management center. The concept, simply stated, is that travelers starting a trip enter their current location and intended destination into an on-board computer that has a two-way communications link to the traffic management center, and the computer — through some combination of the in-vehicle database of historical traffic conditions and updates on current traffic conditions from the traffic management center — identifies a recommended travel route. Information on the traveler and his/her trip origins and destinations would be a valuable source of individual-based exposure data.

Traffic management systems are an important source of the traffic information upon which Intelligent Transportation Systems are based. Traffic management systems are also a potential source of exposure data for highway safety studies. Most of the traffic management systems currently in operation or being designed are limited in scope to freeways. System functions include surveillance, control, and information. Surveillance involves real-time monitoring of traffic conditions (traffic volume and occupancy and, in some cases, speed) on a link-by-link basis in the freeway system. The control function may include ramp metering, for example. The information function refers to advising travelers about accidents or poor traffic conditions ahead via changeable message signs, highway advisory radio, traffic reports on commercial radio stations, etc.

Data Contents and Structure: The traffic volume data available from traffic management systems are generally aggregated over shorter time periods and are measured at more closely spaced intervals than the exposure data typically used for highway safety studies. In fact, the level of detail of the volume data is likely to exceed the needs of many, if not most, highway safety study objectives.

Typical current practice employed by traffic management systems for measuring traffic conditions includes detector stations at 0.8-km intervals along the freeway. The detector stations commonly consist of one inductive loop detector in each freeway lane to measure traffic volume and occupancy. At a subset of those stations, pairs of loop detectors may be used so that speed can also be measured. Twenty- to sixty-second traffic volumes are counted and then transmitted from a local control unit at the detector station to a traffic management center at which volume data from all stations are gathered, processed, monitored in real time, disseminated (in some centers), and stored.

Transportation Research Circular 378 lists freeway traffic management systems currently in operation or in the planning, design, or construction phase. As of 1991, the following areas had operational freeway traffic management systems with a significant number of traffic volume measurement locations: Chicago, Detroit, Long Island, Los Angeles, Minneapolis/St. Paul, Northern Virginia, Phoenix, San Diego, and Seattle. Dozens of urban areas are planning, designing, or constructing systems.

Experimental Design, Sample Plan, and Location Distribution: Each system operates independently and is unique with respect to the scope of surveillance coverage; location of detector stations; detector and communications technologies; and data collection, processing, and storage procedures. To illustrate the similarities and differences among systems, more detailed descriptions will be provided for two urban areas: Seattle and Minneapolis/St. Paul.

Seattle Traffic Management System: The Seattle traffic management system is operated by the Washington State Department of Transportation. The system has grown and evolved since the early 1970s. Traffic volume data are collected at approximately 200 stations. The stations are spaced at approximately 0.8-km intervals. This system provides traffic condition monitoring for approximately 113 km of freeway. Currently, four freeways are monitored: I-5, I-90, SR-405, and SR-520. The system will be expanded within the next several years to add a fifth freeway (SR-167).

Detector stations typically consist of inductive loop detectors in each freeway lane to measure traffic volume and occupancy. At a limited number of stations, pairs of loop detectors in each lane are used to measure speed. Traffic measurements at a detector station are recorded at a local control unit and transmitted to the traffic management center every 20 s. At the center, the volume data are aggregated to 5-min, 15-min, and l-h volumes. Both per lane and total directional volumes are transmitted to the center. Volume data from the detector stations are not disaggregated by vehicle type. There are, however, separate vehicle classification data collection sites in the Seattle area.

The occupancy data are displayed on a dynamic map that is updated every 20 s. Real-time monitoring of the map display is one of several methods used to identify potential incident locations.

The volume data from the detector stations have several uses. The traffic management center uses the volume data to evaluate changes in the ramp metering system, including adjusting metering rates at ramps or analyzing additions to the ramp metering system. Other groups within the Washington State Department of Transportation also make frequent use of the volume data, including design, traffic operations, and traffic data offices.

All volume data from all detector stations are stored. Data are stored as 5-min, 15-min, and l-h volumes. The data are stored on the center's computer system within the system's memory capacity; currently, approximately 10 months of data are available on-line. Older data are archived on magnetic tape or diskette. With some exceptions, data for a given detector station are available for as long as that station has been in operation, some for as long as 25 years. Exceptions include gaps in available data due to detectors being temporarily out of service for maintenance, system expansion, or during freeway reconstruction activities. No assurances can be given that data requested for specific detectors and for specific time periods are available. The availability of data can be determined only through the processing required to access and download the data.

Loop detector data cannot be considered 100 percent accurate. The accuracy of data from loop detectors, however, is generally comparable to other standard methods of measuring traffic volumes. The volume data transmitted to the center from the local control units at each detector station are checked to ensure its quality. Volume counts for an individual lane that fall beyond specified minimums or maximums or that differ more than a specified amount from the volume counts for other lanes at the detector station are flagged as either bad or suspect. These flags are recorded in the files containing the volume data. The flagging process is considered conservative-i.e., some data flagged as suspect because of differences between lanes may, in fact, be correct. Flagged data are excluded from station-wide measures.

Minneapolis/St. Paul Traffic Management System: The Minneapolis Department of Transportation operates a Traffic Management Center to manage traffic on the freeways in the Minneapolis/St. Paul Twin Cities metropolitan area. The center was constructed in 1972. Traffic volume data are collected at approximately 650 stations spaced at approximately 0.8-km intervals. This system monitors traffic on approximately 402 km (805 directional kilometers) of freeway. The freeways monitored include six Interstate highways (I-35E, I-35W, I-94, I-394, I-494, and I-694), as well as seven State highways (Routes 5, 36, 62, 77, 100, 169, and 212).

Detector stations typically consist of inductive loop detectors in each freeway lane to measure traffic volume and occupancy. Traffic speed is calculated based upon these measures. Detectors are also located on entrance and exit ramps. The detectors operate and transmit data to the center 24 hours per day. For control purposes, the center uses 1-min running averages that are updated every 30 s.

All data are archived. The basic time interval for archived data is a 5-min period. The archived data are stored in compressed binary format. Access programs transform the data, extract subsets that are requested, and aggregate data to the desired form. Traffic volume and occupancy data and calculated speeds can be aggregated in 5-, 15-, and 30-min; hourly; and daily time periods. Data can be provided by lane or aggregated for all lanes at a detector station. Data are available for approximately the past 2 or 3 years.

The data are provided "as is." There is no filtering to extract erroneous data, such as due to detector malfunctioning. Volume and occupancy data that deviate from certain thresholds are flagged, and those flags are included in the database. Appropriate use of the data requires familiarity with the area and with this type of data.

Data Acquisition and Documentation: Requests for Minneapolis/St. Paul volume data are handled by the Traffic Management Center on a case-by-case basis. The center has limited staff resources to process requests. The staff can handle requests for small amounts of data and provide the data for specified stations and time periods on diskette to the requester. If the amount of data requested is large, then it may be necessary for the requester to come to the center; the center provides access and the necessary software for the requester to decompress and download the data. The center is considering providing access to data through Internet at some future date. There are no confidentiality requirements or other restrictions on the use of volume data obtained from the center.

Minneapolis/St. Paul data are routinely used in-house and are provided to researchers and government agencies. Several periodic reports are routinely developed using the data, including a congestion report identifying congestion hot spots, a lane closure report that identifies allowable lane closures, a traffic report for traffic forecasting personnel, and a quarterly report on peak-hour volumes and AADT. There is no cost for obtaining the data and there are neither limitations nor confidentiality requirements on the use of the data.

Requests for data should be directed to:

Jim Aswegan
Freeway Operations
Metropolitan Division
Waters Edge
1500 West County Road, B2
Roseville, MN 55113

Reference

(1) Transportation Research Board. ***Transportation Research Circular** 378, **Freeway Operations Project Summary.*** September 1991.

Transportation Planning Surveys

This area covers a range of transportation planning surveys. These are usually household surveys conducted by mail or telephone. Examples are the Transportation Planning Package of the U.S. Census (CTPP). This survey provides nationwide data that form the basis for many State and local transportation planning efforts. However, only trips to and from work are included. The other general source in this area is regional planning surveys. These are also household surveys patterned after the CTPP. The geographic coverage is limited of course, but more detailed information is frequently collected, often for a broader range of trip purposes than just travel to and from work.

Census Transportation Planning Package (CTPP)

Purpose: The Census Transportation Planning Package (CTPP) is a set of special tabulations of the 1990 census data tailored to meet the data needs of transportation planners. The 1990 CTPP was produced by the Bureau of the Census and was sponsored by State Departments of Transportation under a pooled funding arrangement with the American Association of State Highway and Transportation Officials. The CTPP program was coordinated and is technically supported by the Federal Highway Administration of the U.S. Department of Transportation.

The CTPP consists of tables of sociodemographic and journey-to-work information. These tables provide information on commuter travel flows and characteristics; baseline origin-destination data on local work trips; household characteristics; and worker characteristics for use in travel forecasting models and for monitoring car-pooling and transit use. The CTPP data on commuter flows are also used to evaluate and select projects, develop traffic congestion management systems, and identify transportation corridors that need capacity expansion.

In addition, the CTPP also provides travel-to-work and vehicle availability information used in the preparation of vehicular travel and pollutant emissions profiles, computation of regional average rates of vehicle occupancy in the commute to work, and the evaluation of the impact of long-range transportation plans on air quality in compliance with the Clean Air Act Amendments of 1990.

Source: The source of information for the CTPP is the U.S. decennial census, particularly questions 23a and b, and 24a and b, that were asked of a sample of households. These questions asked for mode to work last week, vehicle occupancy, and time the work trip was started and how many minutes it took. This information, together with information on employment location, residential location, and sociodemographics, is the basis of the CTPP.

Organization: Two sets of data packages were produced: (1) statewide packages for each State and the District of Columbia and (2) urban packages for each "CTPP region” as defined by Metropolitan Planning Organizations (MPO).

The statewide CTPP consists of six parts (A through F). Part A contains characteristics of persons, workers, and housing units by county and by place of residence of 2,500 or more population (city, town, village, etc.). Part B contains characteristics of workers by county and place of work of 2,500 or more population. Part C contains characteristics of workers in journey-to-work flows between counties and places of residence of 2,500 or more population and counties and places of work of 2,500 or more population. Parts D, E, and F are similar to parts A, B, and C except for more detailed cross-tabulations of counties of 750,000 or more population and places of 75,000 or more population.

The urban CTPP has eight parts. Part 1 contains the characteristics of persons, workers, and housing units by traffic analysis zone or census tract (MPO option) of residence. Part 2 contains the characteristics of workers by traffic analysis zone or census tract. Part 3 contains characteristics of workers in journey-to-work flows from traffic analysis zone to traffic analysis zone, or from census tract to census tract. Part 4 contains detailed cross-tabulations of trip generation characteristics for the urbanized area, transportation study area, and metropolitan area. Part 5 does not exist, but is a "place-holder" to retain comparability with the 1988 Urban Transportation Planning Package (UTPP). Part 6 contains detailed cross-tabulations of workers in journey-to-work flows between "super districts" (aggregations of traffic analysis zones or census tracts) in CTPP regions of 1,000,000 or more population. Part 7 contains characteristics of workers by census tract of work with an emphasis on economic characteristics. Part 8 contains detailed cross-tabulations of characteristics of workers in journey-to-work flows between traffic analysis zones or census tracts for CTPP regions of 1,000,000 or more population.

Coverage: The 1990 CTPP is the fourth in a series of special transportation-oriented tabulations from the decennial census. In 1960, information on the place of work, mode of travel to work, and automobiles available at home was collected. Tabulations of worker streams were available in a special report for Standard Metropolitan Statistical Areas of more than 250,000 population. Information on automobile availability could be obtained in the series of census reports on housing.

The key transportation-related data collected in the 1970 census were again: place of work, mode of travel to work, and automobiles available in the home. The main difference between the 1960 and 1970 data was the level of geographic coding of the work place. In 1970, specific work addresses were required, while in 1960, only the city or county was identified. A special census product of sociodemographic and journey-to-work information could be ordered by the States and MPOs for transportation planning purposes.

In the 1980 decennial census, additional information on vehicle occupancy, travel time to work, and car and van availability was collected. The place-of-work data were coded to census tracts or blocks. As in 1970, States and MPOs could order special tabulations of demographic and journey-to-work information (now called the Urban Transportation Planning Package).

Strengths and Limitations: The CTPP provides detailed information on the journey-to-work trip for the entire country. Information includes mode, time of journey start, journey time, vehicle occupancy, and sociodemographics of the workers. Since the journey to work is the dominant trip purpose in the morning peak-traffic period, the data in the CTPP could be used to determine exposures for that particular time period. Obviously, any study using this approach would have to consider the portion of traffic in that time period not associated with the work journey.

The availability of similar journey-to-work information from previous censuses allows for the analysis of trends and changes in exposure for the morning peak-traffic period.

Since the information in the CTPP is limited to the journey to work, the CTPP is not a good source of exposure information for any times other than morning traffic-peak periods.

Sampling Errors: Variable sampling rates were used in the sample portion of the census. In general, in less densely populated areas, one in two households was sampled; while in densely populated areas, the rate was one in eight households. When all sampling rates are taken into account across the country, one in every six households was sampled.

The standard error of sample estimates can be calculated using tables and procedures given in ***Appendix C - Accuracy of the Data*** of the CTPP documentation.

Access: CTPP data are available from the Bureau of Transportation Statistics of the U.S. Department of Transportation on CD-ROM, together with the software (TransVU — CTPP Edition) to display and retrieve the data. TransVU — CTPP Edition is a Microsoft Windows application that provides both map and tabular view of CTPP data and simplifies extraction of CTPP tables in dBASE, Lotus, and comma-delimited or fixed-format test files. The CTPP CD-ROM and a copy of TransVU — CTPP Edition software are available from the Bureau of Transportation Statistics without charge.

Traffic Volume Data — Errors of VMT Estimates Based on Traffic Counts and Section Length

Typically, vehicle-miles traveled (VMT) are estimated from traffic counts and highway mileage. While the basic idea is simple, it can be implemented in several ways, which lead to different estimates with different errors.

This is the summary of a brief analysis of these techniques, including the method recommended in the HPMS for estimating VMT. Only the results are shown, not the sometimes tedious algebra. Two of the three procedures involve nonlinear expressions; therefore, linear approximations were used as usual. Therefore, the formulas are good approximations only if the coefficients of variation of the data are "small." A value of 0.1 is, for nearly all practical purposes, "small," 0.2 is small for most, and even 0.3 might be adequate for some approximate estimates.

Basic Definitions

The highway (system) studied has the length L and is divided into N sections of lengths l_i; their average is l_o. A sample of n sections is used; each section has the same probability of being selected. On section I, the average daily traffic is x_i. Its mean overall section is x_o. Variables s(x) and s(l) are the standard deviations of x_i and l_i. Their coefficients of variations are $c_x = s(x)/x_o$, and $c_l = s(l)/l_o$. One also needs the correlation coefficient ρ between the x_i and l_i. For instance, if in more densely settled areas traffic is heavier and sections are shorter, there is a negative correlation. On the other hand, if highways of a different character are combined, those with heavier traffic might have longer sections than those with lighter traffic. Then, there would be a positive correlation. Such correlations can appreciably influence the errors of VMT estimates. Therefore, they must be empirically determined and incorporated into the calculations. Formula ***3*** on page 3-3-9 of the ***traffic monitoring guide*** appears to do this implicitly.[1] However, this is a formula for the standard error of a biased estimate that is less relevant than the mean square error (see below).

The total vehicle-miles traveled on the L miles of highway are:

$$V=\sum l_i x_i = L x_o (1 + c_x c_l \rho) \qquad (1)$$

where the second term in the parentheses reflects the effects of correlations between section length and volume.

[1] This formula is, aside from a misprint, equivalent to formula (6.10) in section 6.4 of W.G. Cochran, Sampling Techniques, Third Edition, Wiley, 1977.

The Unbiased Estimator

If n highway sections are randomly selected out of N with equal probabilities, the unbiased estimator of total **VMT** is:

$$\hat{V}_1 = \frac{N}{n} \sum x_i l_i \tag{2}$$

where the sum is over the n elements of the sample. It has a standard deviation (equal to the mean square error, because the estimator is unbiased) given by:

$$SD(\hat{V}_1)^2 = (Lx_o c_x)^2 \left(\frac{1}{n}\left(1 + 2\frac{c_l}{c_x}\rho + \left(\frac{c_l}{c_x}\right)^2\right) - c_l^2 \rho\right) \tag{3}$$

if the finite population correction is ignored. The effect of a correlation between section length and volume is complex. If n is large, the expression in the right parentheses can become negative. This means simply that the linear approximation used for the product $x_I l_I$ is no longer valid.

A "Quick and Dirty" Estimator

This estimator averages the observed x_I and multiplies the average by the length of the highway system:

$$\hat{V}_2 = L \sum x_i / n \tag{4}$$

It is a biased estimator. Its expected value is:

$$E(\hat{V})_2 = Lx_o \tag{5}$$

It differs from the unbiased estimator by a factor of $1/(1+c_x c_l \rho)$. The bias disappears if the x_I and l_I are uncorrelated (ρ = o); it does not decrease when the sample size is increased. For a negative correlation and large coefficients of variation, $1+c_x c_l \rho$ can be small, and $\hat{V}_2$ can be a gross overestimate of **V**, no matter how large the sample. The standard error is given by:

$$SD(\hat{V}_2)^2 = (Lx_o c_x)^2 / n \tag{6}$$

However, because it is a biased estimator, the mean square error given by:

$$MSQE(\hat{V}_2)^2 = E(\hat{V}_2 - V)^2 \tag{7}$$

is more meaningful, because it includes the bias into the error calculation:

$$MSQE(\hat{V}_2)^2 = (Lx_o c_x)^2 (\frac{1}{n} + c_l^2 \rho) \tag{8}$$

The second term in the parentheses reflects the effect of the bias. The first term decreases with increasing sample site n; the second remains constant. Thus, if ρ and c_l are not negligible, this is not a good estimator.

The Ratio Estimator Recommended by HPMS

The unbiased estimator calculates VMT on the sample sections and then divides it by the sample fraction-the ratio of sampled sections to total sections. The ratio estimator also calculates VMT on the sample sections, but then divides it by the ratio of the combined length of the sample sections and the total length L:

$$\hat{V}_3 = \frac{L}{\sum l_i} \sum x_i l_i \tag{9}$$

The advantage of this is that it reduces the effect of the varying length of the sample sections on the variance of the estimate; its disadvantage is that the estimate is biased. The expected value is

$$E(\hat{V}_3) = x_o L(1 + \rho c_x c_l - \frac{1}{n} \rho c_x c_l) = V(1 - \frac{1}{n}) \frac{\rho c_x c_l}{1 + \rho c_x c_l} \tag{10}$$

For this estimator, the bias decreases with increasing sample size; it also decreases with decreasing correlation ρ and with decreasing coefficients of variation c_x and c_l. Its mean square error is given by:

$$MSQE(\hat{V}_3)^2 = (Lx_o c_x)^2 ((1 + 2c_l^2 \rho^2)/n - c_l^2 \rho^2) \tag{11}$$

Again, the right parentheses can become zero or negative if the linear approximations are no longer valid.

Comparing the Unbiased Estimator and FHWA's Estimator

The difference between equation (9) and V is the bias of FHWA's estimator. Thus,

$$\frac{BIAS}{v} = -\frac{I}{n}\frac{c_x c_l \rho}{1 + c_x c_l \rho} \tag{12}$$

is the bias as a proportion of the actual value. This bias is the price to pay for the reduction of the variance achieved by the ratio estimator. Whether it is worthwhile depends on the difference between the mean square error of the two estimators. The difference of their squares is

$$MSQE(\hat{V}_1)^2 - MSQE(\hat{V}_3)^2 = \frac{(Lx_o c_x)^2}{n}\left(2\frac{c_l}{c_x}\rho + \left(\frac{c_l}{c_x}\right)^2 - 2c_l^2\rho\right) \tag{13}$$

This difference can be positive as well as negative. It can become large with either sign, but the relevance of this is limited because before very large values are reached, the linear approximations become invalid.

However, it appears worthwhile to check in real applications how large an improvement of the variance is provided by using a biased estimator, and whether despite the bias, the mean square error will be improved.

APPENDIX: HPMS FORMS AND DATA FORMAT

The appendix contains selected forms reproduced from the 1993 edition of the FHWA Highway Performance Monitoring System Field Manual, OMB No. 2125-0028.

Template – 1

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

TOTALS OF URBANIZED AREAS, SMALL URBAN AREAS, RURAL AREAS, AND STATEWIDE

STATE:__________ STATE FIPS CODE:______ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: ______ DATE:______

URBANIZED AREAS TOTAL

POPULA-TION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM						
			PRINCIPAL ARTERIALS			MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
			INTERSTATE	OTHER FREEWAYS & EXPRESSWAYS	OTHER				
		LENGTH							
		TRAVEL (1,000)							

SMALL URBAN AREAS TOTAL

POPULA-TION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM						
			PRINCIPAL ARTERIALS			MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
			INTERSTATE	OTHER FREEWAYS & EXPRESSWAYS	OTHER				
		LENGTH							
		TRAVEL (1,000)							

RURAL AREAS TOTAL

POPULA-TION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM						
			PRINCIPAL ARTERIALS		MINOR ARTERIAL	COLLECTOR		LOCAL	TOTAL
			INTERSTATE	OTHER		MAJOR	MINOR		
		LENGTH							
		TRAVEL (1,000)							

STATEWIDE TOTALS

POPULA-TION (1,000)	NET LAND AREA	TOTAL LENGTH	TOTAL TRAVEL (1,000)

1/ English units for length and travel are miles and vehicle–miles (in thousands), respectively.

2/ Metric units for length and travel are kilometers and vehicle–kilometers (in thousands), respectively.

Template – 2

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

INDIVIDUAL URBANIZED AREAS

STATE:________ STATE FIPS CODE: UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR______ DATE:________

Shaded cells are reserved for titles and computer software generated values. Enter data in the unshaded cells only.

URBANIZED AREA CODE	NONATTAINMENT AREA CODE 3/	POPULATION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM						
					PRINCIPAL ARTERIAL			MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
					INTERSTATE	OTHER FREEWAYS & EXPRESSWAYS	OTHER				
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							

1/ English units for length and travel are miles and daily vehicle–miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and daily vehicle–kilometers (in thousands), respectively.
3/ The National Ambient Air Quality Standards Nonattainment Area Code is the same as the Urbanized Area Code of the primary urbanized area contained in the nonattainment area. When the Urbanized Area is not in a nonattainment area, code zero.
4/ Average vehicle occupancy is reported to the nearest tenth of a person.

Template – 3

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

DONUT AREA DATA FOR INDIVIDUAL NAAQS NONATTAINMENT AREAS

STATE: ____________ STATE FIPS CODE: ______ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: ________ DATE: ____________

Shaded cells are resewed for titles and computer sofware generated values. Enter data In the unshaded cells only.

NONATTAIN-MENT AREA CODE 3/	POPULATION (1,000)	NET LAND AREA	DATA TYPE	RURAL AND SMALL URBAN FUNCTIONAL SYSTEMS COMBINED			
				PRINCIPAL ARTERIAL	MINOR ARTERIAL AND (MAJOR) COLLECTOR	MINOR COLLECTOR AND LOCAL	TOTAL
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				

1/ English units for length and travel are miles and daily vehicle-miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and daily vehicle-kilometers (in thousands), respectively.
3/ The National Ambient Air Quality Standards (NAAQS) Nonattainment Area Code is the same as the Urbanized Area Code of the primary urbanized area contained in the nonattainment area.

AIRQUAL1 6/3/93

Template – 4

MINOR COLLECTOR AND LOCAL FUNCTIONAL SYSTEM LENGTH

BY SURFACE TYPE AND VOLUME GROUP

STATE: ____________________ STATE FIPS CODE: __________ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: ______ DATE: __________

Shaded cells are reserved for titles and computer software generated values. Enter data in the unshaded cells only.

FUNCTIONAL SYSTEM	SURFACE TYPE					TOTAL LENGTH
	PAVED				UNPAVED	
	HIGH	INTERMEDIATE	LOW	SUBTOTAL		
RURAL MINOR COLLECTOR						

FUNCTIONAL SYSTEM / SURFACE TYPE	TRAFFIC VOLUME GROUP					TOTAL LENGTH
RURAL MINOR COLLECTOR	UNDER 100	100– 199	200– 499	500– 1,999	2,000 & OVER	
PAVED						
UNPAVED						
TOTAL						
RURAL LOCAL	UNDER 50	50– 199	200– 499	500– 1,999	2,000 & OVER	
PAVED						
UNPAVED						
TOTAL						
URBAN LOCAL	UNDER 200	200– 499	500– 1,999	2,000– 3,999	4,000 & OVER	
PAVED						
UNPAVED						
TOTAL						

1/ English units consist of miles.
2/ Metric units consist of kilometers.

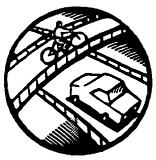
Sources of Exposure Data for Safety Analysis

PUBLICATION NO. FHWA-RD-97-025

NOVEMBER 1997

US. Department of Transportation
Federal Highway Administration

Research and Development
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296



FOREWORD

This report was prepared for the Federal Highway Administration (FHWA) as part of a contract to evaluate sources of exposure data for highway safety research. Several existing and emerging exposure data sources were subsequently selected by FHWA for review.

This report provides highway safety researchers with information to assess the feasibility of using exposure data sources in designing highway safety evaluation studies. One-page summaries are provided for each exposure data source. A longer description covers the purpose of the collection, contents, period covered, sample design, data collection methods, sample size, data quality, data format, possible cautions in using the exposure data, and availability of the data.

Copies of this report are available for a nominal charge from the National Technical Information Service (NTIS), Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

A. George Ostensen, Director
Office of Safety and Traffic Operations
Research and Development

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-RD-97-025		

4. Title and Subtitle	5. Report Date
SOURCES OF EXPOSURE DATA FOR SAFETY ANALYSIS	November 1997
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Campbell, K.L, Joksch, H.C., Blower, D., Kostyniuk, L.P., Pendleton, O.J., and Griffin III, L.I.	UMTRI-96-17

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
The University of Michigan Transportation Research Institute 2901 Baxter Road Ann Arbor, MI 48109-2150 Texas Transportation Institute The Texas A&M University System College Station, TX 77843	3A3B 11. Contract or Grant No. DTFH61-93-C-00123

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Office of Safety and Traffic Operations R&D Federal Highway Administration 6300 Georgetown Pike McLean, Virginia 22101-2296	Task Report October 1993 - March 1996 14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representatives (COTRs): Michael S. Griffith and Joe Bared, HSR-20

16. Abstract

This report describes existing and emerging exposure data sources for highway safety analysis. Existing exposure data sources reviewed include: Highway Performance Monitoring System (HPMS), Highway Safety Information System (HSIS), Long-Term Pavement Performance (LTPP) Monitoring System, Nationwide Personal Transportation Survey (NPTS), National Truck Trip Information Survey (NTTIS), Operational Exposure Data Sources, Residential Transportation Energy Consumption Survey, Truck Inventory and Use Survey (TIUS), and Weigh-in-Motion (WIM) devices. Emerging data sources are new sources or existing sources that have not been traditionally used to derive exposure estimates. Three areas were reviewed for possible emerging exposure data: Intelligent Transportation Systems (ITS), transportation planning surveys, and traffic volume data collected by the States. One-page summaries are provided for each exposure data source. A longer description covers the purpose of the collection, contents, period covered, sample design, data collection methods, sample size, data quality, data format, possible cautions in using the exposure data, and availability of the data.

17. Key Words	18. Distribution Statement
Exposure data, highway safety.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

Form DOT F 1700.7 **Reproduction of completed page authorized**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yards	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: Volumes greater than 1000 I shall be shown in m^3.				
		MASS		
oz	ounces	28.35	grams	9
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celcius temperature	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m*	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By .	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.71	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
9	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

- SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

(Revised September 1993)

CONTENTS

1. SUMMARY . . . 1
Highway Performance Monitoring System . . . 2
Highway Safety Information System . . . 3
Long-Term Pavement Performance Monitoring System . . . 4
Nationwide Personal Transportation Survey . . . 5
National Truck Trip Information Survey . . . 6
Operational Exposure Data Sources . . . 7
Residential Transportation Energy Consumption Survey . . . 8
Truck Inventory and Use Survey . . .
Weigh in Motion . . . 10
Intelligent Transportation Systems Commercial Vehicle Operations . . . 11
ITS Advanced Traveler Information Systems . . . 12
ITS Advanced Traffic Management Systems . . . 13
Transportation Planning Surveys . . . 14
Census Transportation Planning Package . . . 15

2. EXISTING EXPOSURE DATA SOURCES . . . 17
Highway Performance Monitoring System . . . 18
Highway Safety Information System . . . 23
Long-Term Pavement Performance . . . 45
Nationwide Personal Transportation Survey . . . 49
National Truck Trip Information Survey . . . 51
Operational Exposure Data Sources . . . 53
Residential Transportation Energy Consumption Survey . . . 54
Truck Inventory and Use Survey . . . 59
State Weigh-in-Motion and Automatic Vehicle Counting Devices . . . 61

3. EMERGING EXPOSURE **DATA SOURCES** . . . 67
Intelligent Transportation Systems . . . 68
Commercial Vehicle Operations of ITS . . . 68
Advanced Traveler Information Systems . . . 70
Advanced Traffic Management System . . . 70
Transportation Planning Surveys . . . 76
Traffic Volume Data — Errors of VMT Estimates Based on Traffic Counts and Section Length . . . 79

APPENDIX: **HPMS** FORMS **AND DATA FORMAT** . . . 83

LIST OF TABLES

Table 1. Region 1 WIM 62
Table 2. Region 2 WIM 63
Table 3. Region 3 WIM 64

1. SUMMARY

One-page summaries of both the existing and emerging exposure data sources reviewed for this report are presented in this section. A more complete discussion of each of the existing exposure data sources is presented in Section 2 and emerging data sources are in Section 3. The following exposure data sources are summarized in this section:

<u>Existing Exposure Data Sources</u>

Highway Performance Monitoring System (HPMS)
Highway Safety Information System (HSIS)
Long-Term Pavement Performance (LTPP) Monitoring System
Nationwide Personal Transportation Survey (NPTS)
National Truck Trip Information Survey (NTTIS)
Operational Exposure Data Sources
Residential Transportation Energy Consumption Survey
Truck Inventory and Use Survey (TIUS)
Weigh in Motion (WIM)

<u>Emerging: Data Sources</u>

Intelligent Transportation Systems
- Commercial Vehicle Operations
- Advanced Traveler Information Systems
- Advanced Traffic Management Systems

Transportation Planning Surveys
- Census Transportation Planning Package

Highway Performance Monitoring System (HPMS)
Federal Highway Administration and State Highway Agencies

Purpose:	Assess the length, use, condition, performance, and operating characteristics of the National Highway System
Source:	State highway agencies Vehicle-Miles Traveled (VMT) based on Annual Average Daily Traffic (AADT) Fatal and injury accident data
Coverage:	Annual reporting, initiated in 1978 All public roads in the United States (except local streets and roads) Areawide Universe Standard sample "Donut" sample (for air quality) Geographical Information System (GIS) coding
Sample:	Simple Random Sample (SRS) prescribed by the Federal Highway Administration (FHWA) of ~115,000 road segments
Response:	Data are required by law and, therefore, are complete
Strengths:	National aggregate data in broad categories of highway function, area type, and use Standard format
Limitations:	Accident data not associated with the standard sample (Vehicle classification of VMT not compatible with accident data)
Accuracy:	AADT is improved for the standard sample, but is still the critical element for VMT

Highway Safety Information System (HSIS)
Highway Safety Research Center

Purpose:	Provide linked accident, highway inventory, and traffic count data in SAS® format for selected States to provide an enhanced analysis capability
States Reviewed:	Illinois, Maine, Michigan, Minnesota, and Utah
Source:	VMT from segment lengths and AADT AADT updated from 1 to 5 years, Some estimated or interpolated, Some sites permanent, year-round, Most are temporary sites, 48-h counts Some with vehicle classification, or "commercial"
Coverage:	In most States, a major portion (but not all of the highway system) is covered, usually State-maintained roads
Sample:	Usually a purposefully selected subset Cross-section files in some States contain a sample of segments, usually limited
Strengths:	Large samples Diversity of data in different States SAS® format, documentation Suited for aggregate comparisons
Limitations:	AADT data very coarse, generally not suited for identifying individual, high-risk locations Entering volumes for both roads of an intersection often not available National estimates not possible Diversity of data in different States
Accuracy:	AADT not all observed, not independent, so variance cannot be estimated

Long-Term Pavement Performance (LTPP) Monitoring System
Transportation Research Board/Federal Highway Administration

Purpose:	Satisfy the total range of pavement information needs Collect information to develop models of how various design features, traffic, and environment impact pavement performance Central Traffic Database contains annual estimates of traffic and load data
Source:	Central Traffic Database contains historical and monitored traffic data Yearly estimates of volumes, axle loads, and equivalent single-axle loads available for each site Truck weights and distributions collected at sites quarterly for 7 days 35 percent of sites have weigh-in-motion collectors, the remainder have Automatic Vehicle Classification counters
Coverage:	Data collected in four geographic regions 20-year research program begun in 1987
Sample:	789 sites on key highway routes provide truck weights and distributions Historic traffic data requested where available
Strengths:	With further development, should provide reliable vehicle count and classification data Good data source for location-based safety studies, if sites can be linked with accident histories
Limitations:	Weigh-in-motion data location not always exactly at the site Researcher must verify exact location of traffic data Quality control issues with the data currently a problem Some sites have only a minimal amount of data Currently, only limited amount of data available to the public
Accuracy:	Currently a problem, expected to improve Data quality procedures and standards have been implemented

Nationwide Personal Transportation Survey (NPTS)
Federal Highway Administration

Purpose:	U.S. estimates of personal travel All modes: car, truck, bus, train, subway, airplane, taxi, motorcycle, bicycle, and walking Includes household demographics, person-level information, household vehicles, and trip information
Source:	Conducted by Research Triangle Institute (1990) Random-dialing household telephone survey 12-month survey period 24-h travel-day period 14-day travel period for trips >12 1 km
Coverage:	National coverage, all trips, all modes, all purposes, in all 50 States plus Washington, D.C. Oversample in Connecticut; N.Y. metropolitan planning organization; and Indianapolis, Indiana Approximately 7-year intervals
Sample:	22,000 households 48,000 persons 35,000 licensed drivers 41,000 vehicles
Response:	~85 percent at the household level
Strengths:	Only source for national personal travel Large sample size Stable since 1969 (Home interviews prior to 1990) Good detail at all levels
Limitations:	Households without telephones not included Limited sample for commercial vehicles (trucks) Self-reported information Cannot disaggregate by State 7-year interval
Accuracy:	Sampling errors can be calculated with appropriate software

National Truck Trip Information Survey (NTTIS)
University of Michigan Transportation Research Institute

Purpose:	National estimates of medium and heavy truck population and travel with detailed vehicle and trip-level data that allow cross-classification by configuration, loading, road type, rural/urban, and day/night
Source:	Sample of registered trucks from R.L. Polk Telephone surveys on four randomly assigned dates Conducted by University of Michigan Transportation Research Institute (UMTRI)
Coverage:	48 States plus Washington, D.C. Government-owned vehicles excluded 12-month survey period in 1985-1986 One time only
Sample:	Probability-based sample of 8,144 registered trucks (GVWR>4536 kg) from 1983 R.L. Polk files Trip-level data on a sub-sample of 5,000 vehicles 13,097 trips on 17,660 survey days
Response:	83 percent at the vehicle level 86 percent at the survey-day level
Strengths:	Most accurate identification of trucks > 4536 kg Duplicate registrations deleted from frame Detailed cross-classification of vehicle characteristics, loading, and operating environment unmatched in any other source Extensive edit and consistency checks Some questions overlap Truck Inventory and Use Survey for comparison
Limitations:	Limited sample size Cannot disaggregate by State Self-reported information Now out of date
Accuracy:	Underrepresents newest vehicles due to lag between sample and trip survey Complex sample design can be calculated with appropriate software Large variances for small subsets (doubles)

Operational Exposure Data Sources
State and Local Traffic Agencies

Purpose: State and local traffic agencies collect a variety of traffic data for both long-term and short-term objectives that often go beyond the requirements of the Highway Performance Monitoring System (HPMS) described previously. Typical data include traffic counts from both permanent and temporary stations, Automatic Traffic Recorders, and State highway inventory files. However, data collection beyond the scope of HPMS is often on an ad hoc basis to address specific short-term purposes.

Source: There is no single source. State traffic agencies are often aware of many of the local programs, as well as the State data; but the city, county, or metropolitan planning organization will have to be contacted to obtain detailed information or data.

Coverage: Most States have extensive traffic monitoring programs with a combination of permanent and temporary programs. Major cities often collect Average Daily Traffic (ADT) volumes on many arterial streets as well.

Sample: Some stations may be permanent and coverage of individual routes may be quite complete, but outside of HPMS, there is generally no sample design that would support any extrapolation of the data.

Strengths: Specific projects may be possible, taking advantage of additional details with regard to peak versus off-peak, day-of-week, and site-specific data that might be located.

Limitations: A major limitation is that none of the data is typically automated. Another important limitation is that the often ad hoc nature of the data collection may bias the data.

Residential Transportation Energy Consumption Survey
Energy Information Administration

Purpose: Obtain information on the vehicles used for personal transportation in the United States
Companion survey to the Residential Energy Consumption Survey (RECS)
RECS includes household demographics
Residential Transportation Energy Consumption Survey (RTECS) includes VMT (from odometer readings), motor vehicle stock, and vehicle fuel consumption and expenditure data.

Source: RECS is a random household telephone survey (mail questionnaire used when telephone interview is not possible)
Multistage probability sample incorporating a rotating panel
RTECS is a subsample of RECS households, telephone/mail survey
First phase of RTECS done in conjunction with RECS
Subsequent three phases conducted at the beginning, middle, and end of the year

Coverage: All 50 States and Washington, D.C.
Families or individuals living in group quarters or with no fixed address excluded
Motorcycles, bicycles, and all-terrain vehicles excluded
Conducted every 3 years since 1985

Sample: 5,095 households responded to the most recent RECS survey
3,045 households selected for most recent RTECS survey

Response: 75 percent household response rate to RECS
Unknown response rate to RECS

Strengths: Household VMT and vehicle stock data
Estimates of VMT by age and gender of primary driver
Stable since 1978

Limitations: Small sample size
No trip data
Two odometer readings not obtained for large fraction of sample vehicles, annual VMT imputed for these
Data do not relate VMT to person-miles of travel, so vehicle occupancy is unknown, and driver age and gender have to be assumed from primary driver data
3-year interval

Accuracy: Questionable
26 percent of households not followed for the entire year
Various imputation techniques used to handle item nonresponse

Truck Inventory and Use Survey (TIUS)
Bureau of the Census

Purpose:	Estimate U.S. population of registered trucks (light, medium, and heavy) and provide descriptive information on the trucks and their use over the past year
Source:	R.L. Polk Stratified probability sample of truck registrations from each State Survey form mailed to each owner
Coverage:	Registered trucks in the 50 States plus Washington, D.C. "Typical" use during the past year Excludes government-owned and passenger vehicles Conducted every 5 years
Sample:	~100,000 vehicles
Response:	Required by law ~80 percent (1987)
Strengths:	Well-defined population Rigorous sample design (SRS) Large sample Good response Stable format back to 1967 Population estimates can be disaggregated by State
Limitations:	Self-reported "Typical" use over the past year underrepresents minority use such as bobtail or infrequent trailers/cargoes Mileage estimated cannot be disaggregated by State Possible duplications in registration data across States Conducted only every 5 years
Accuracy:	Sufficient data to calculate sampling errors not released Approximate error formulas provided Minimal bias, random errors generally small

Weigh in Motion

Purpose: Provide information about vehicle weights and axle loads or decisions related to planning, funding, operating, and managing highway facilities for enforcement of weight limits

Source: ***Traffic Monitoring Guide (TMG)*** — required by FHWA and collected by State Departments of Transportation (DOTs)
Long-Term Pavement Performance (LTPP) data — part of the Strategic Highway Research Program (SHRP) — collected by State DOTs and forwarded to regional SHRP contractors
Truck weight enforcement stations — data collected by State police organizations, data usually not retained

Coverage: National coverage

Sample: TMG — 1,400 Weigh-in-Motion (WIM) sites throughout the United States
LTPP — 777 WIM sites throughout the United States

Data Availability: National database containing station description, traffic volume, vehicle classification, and truck weight available directly from FHWA in ASCII flat-file format
Individual State data must be requested from State DOTs, data formats vary widely

Strengths: Only national source for exposure by truck weight

Weakness: Compatibility of TMG data across States — each State determines own experimental design, and number and location of WIM sites
Hardware and software problems associated with collecting data

Accuracy: Varies by State — need to contact State for design and sampling information

Intelligent Transportation Systems (ITS) Commercial Vehicle Operations

Broad Categories of Commercial Vehicle Operations (CVO) User Services:

- Commercial vehicle electronic clearance.
- Automated roadside inspections.
- Commercial vehicle administrative services.
- On-board safety monitoring.
- Hazardous material incident response.
- Commercial fleet management.

Commercial Fleet Management:

Global Positioning System (GPS) recording of vehicle trips by fleet linking with cargo, configuration, and vehicle data

Prospect? Produce the electronic equivalent of a trip diary

Commercial Vehicle Administrative Services:

Vehicle-based GPS technology to get travel by State for International Registration Plan (IRP) purposes (Iowa)

Prospect? Added GPS detail could produce a vehicle-based sample of mileage by road type

Commercial Vehicle Electronic Clearance:

Electronic roadside sampling to transmit compliance data

Prospect? Roadside sampling of vehicle, cargo, and driver characteristics
Identification could allow tracking to subsequent locations to get VMT and travel time

ITS Advanced Traveler Information Systems

Route Navigation:

Vehicle-based navigation system could retain a history of the route followed, plus speed and time, providing an electronic trip diary

Other Uses of ITS Technology:

WIM technology installed on a banked curve could measure vehicle center-of-gravity (cg) height

ITS Advanced Traffic Management Systems

Purpose: Detailed traffic volume data are collected in many large metropolitan areas to provide real-time information for sophisticated traffic management systems. Details vary from one installation to the next. Each city must be contacted for specific information. Seattle and Minneapolis/St. Paul are reviewed in Section 3 of this report.

Source: Inductive loops are the primary source for both volume and speed data. Some automatic vehicle classification equipment is used.

Coverage: High-volume freeways in large metropolitan areas.

Sample: Coverage of road network under the control of the traffic management system is essentially complete.

Strengths: Data are automated and all historical data are archived. Level of detail typically is on the order of 1-min counts per lane at 0.8-km intervals in both directions with speed data for a subset of the stations, plus some ramp measurements. A typical installation has several hundred stations.

Limitations: Limited to the highway network covered.

Accuracy: Accuracy of the data from inductive loops is not 100 percent, but is comparable with other traffic volume measurement methods. Observations outside the expected range are automatically flagged in the better systems.

Transportation Planning Surveys
(Travel)

Purpose:	Designed primarily as origin-destination surveys for planning purposes like the Census Transportation Planning Package (CTPP), with coverage of more trip purposes, but for a limited geographic region.
Source:	Metropolitan planning organizations, or sometime States, conduct additional surveys, often to support travel demand models and other requirements of the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA).
Coverage:	Limited geographic region Broader coverage of trip purposes
Sample:	Usually a census-based household sample, plus surveys of registered trucks or taxis, and roadside surveys.
Strengths:	More complete coverage of trip purposes and time of day Objective is to get future origin-destination flows by travel mode
Limitations:	Difficult to get VMT estimates Geographic limitation

Census Transportation Planning Package (CTPP)
Bureau of the Census

Purpose:	Provide national data for transportation planners on the journey to work. Focus is on the origin-destination flows between traffic analysis zones
Source:	Questions on a supplement to the U.S. Census that is sent to a sample of households, covering residential location, employment location, mode of journey, starting time, and journey time.
Coverage:	National, but only for the journey to work.
Sample:	Statewide package Urban Package SRS of about one out of six households
Strengths:	Designed for transportation planning purposes.
Limitations:	Journey to work only VMT not available Difficult to imagine application to safety analysis
Accuracy:	Sampling errors can be calculated.

2. EXISTING EXPOSURE DATA SOURCES

Existing exposure data sources for use in highway safety analysis are described in this section. The following exposure data sources are included:

Highway Performance Monitoring System (HPMS)
Highway Safety Information System (HSIS)
Long-Term Pavement Performance (LTPP)
Nationwide Personal Transportation Survey (NPTS)
National Truck Trip Information Survey (NTTIS)
Operational Exposure Data Sources
Residential Transportation Energy Consumption Survey (RTECS)
Truck Inventory and Use Survey (TIUS)
Weigh in Motion (WIM)

A description of each data source has been prepared for a data catalog. The objective of the catalog is to provide the highway safety researcher with sufficient information to assess the feasibility (considering time, level of effort, and cost constraints) of using the exposure data source in designing a highway safety evaluation study. The descriptions contain the following information, as applicable:

- Original purpose of the data collection.
- Brief description of the contents of the data source that would be of interest in highway safety research.
- Discussion about the quality of the data, how the data were archived, and for what time periods.
- Discussion of data collection methods or the performance characteristics of the equipment used in terms of reliability and data quality.
- Discussion of the number of sites and locations of the data collection effort and the statistical reliability of these sample sizes as applied to highway safety research.
- Sample of the data format and details as to how to obtain the data, what software or hardware is necessary to access the data, how often the data are updated, and the frequency of data releases, etc.
- Cautions and potential problems with exposure estimates.

Highway Performance Monitoring System (HPMS)

Contents

The Highway Performance Monitoring System (HPMS) is a nationwide inventory system that includes all of the Nation's public road mileage. The primary purpose of the HPMS is to serve the data and information needs of the FHWA and Congress. The HPMS assesses the system length, use, condition, performance, and operating characteristics of the highway infrastructure.

The HPMS was initiated in 1978 to consolidate and streamline the States' data collection efforts and reporting requirements. In keeping with FHWA's mandate to provide information, the HPMS is reassessed and modified to collect data relevant to emerging issues. In such a way, collection of pavement information was added to the HPMS in 1987. It was modified again in 1993 to respond to the need to monitor travel for the clean air issues. The HPMS also changes with advances in technology. In 1993, States were required to submit a linear referencing system for their road systems. Thus, the structure of HPMS is undergoing changes over time as data items are added and dropped in response to current information needs.

The HPMS organization, guidance, and analyses are the responsibility of the FHWA. Data reporting for the HPMS is accomplished by the State highway agencies in cooperation with local governmental units and metropolitan planning agencies.

The HPMS report submitted annually by each State consists of:

- Areawide data.
- Universe data.
- Data for a standard sample.
- Data for the "donut" sample (new in 1993).
- Linear referencing system (new in 1993).

Areawide Data. The areawide data consist of statewide summaries. These data consist of the totals for mileage, travel, accidents, local system data, land area, population, and travel activity by vehicle type. This information is reported for rural, total small urban, and individual urbanized areas.

Universe Data. Universe data refers to a limited set of data items reported for the entire public roads system as individual sections or grouped length records. The public roads system includes those roads owned by the State highway agency, local governments, and Federal agencies. These data contain a complete inventory of mileage classified by system, jurisdiction, and selected operational characteristics.

Standard Sample Data. The standard sample data include specific inventory, condition, and operational data obtained for the sample panels of highway sections. These data can be expanded to represent the universe of highway mileage.

The data cover:

- Identification relative to functional system, route, jurisdiction, and area type.
- Operational information about volume, lanes, access control, medians, and pavement.
- Geometric information about lane widths, shoulders, right-of-way (ROW), horizontal and vertical alignment, and passing sight distance.
- Traffic volume and capacity information such as AADT, speed limits, design factors, service volumes, and signalization.
- Environmental information such as climate and drainage.
- Intersection and interchange information.
- Information about capital improvements.

"Donut" Sample Data. "Donut" data requirements were added to the HPMS in 1993 in response to a need of the Environmental Protection Agency (EPA). The "donut" sample is a supplementary sample of highway panels from the nonurbanized portion (donut area) of National Ambient Air Quality Standards (NAAQS) nonattainment areas. This additional sampling is required to serve EPA's Section 187 Travel Tracking and Forecasting Procedures for the NAAQS non-attainment areas.

The data items are a subset of the data items provided for the standard sample and include identifiers, AADT, and expansion factors.

Linear Referencing System. A linear referencing system (LRS) was added to the HPMS for the 1993 report. These data will enhance the HPMS with Geographic Information System (GIS) capabilities. The data consist of node data file, inventory route and link data files, and inventory route and node maps for the principal arterial system/national highway system (PAS/NHS), and the rural minor arterial system.

Samples

Standard Sample. The HPMS universe consists of all public highways or roads within a State with the exception of roads functionally classified as local. The reporting strata for the HPMS include type of area (rural, small urban, and individual or collective urbanized areas) and functional class (in rural areas, these are Interstate; other principal arterial, minor arterial, major collector, and minor collector; in urban areas, these are Interstate, other freeway or expressway, other principal arterial, minor arterial, and collector). A third level of stratification based on volume was added as a statistical device to reduce sample size and to ensure inclusion of the higher volume sections of the sample in 1987.

The HPMS sampling element is defined on the basis of road segment, which includes both. directions of travel and all travel lanes within the section. The HPMS standard sample design is a stratified simple random sample.

Donut Area Sample. The donut area sampling universe consists of all highway sections functionally classified as rural minor arterial and major collector, and small urban minor arterial and collector that are located within the defined nonattainment boundary and outside of all urbanized area boundaries. This typically forms an annular spatial area and is, therefore, called a "donut."

The donut universe is stratified into two functional systems (the minor arterial and collector) and a limited number of volume-group strata. The sample is a stratified simple random sample.

Data Quality

Generally, the quality of data is good. There is some variation in quality of the HPMS reports across the States. Since these data are required by the Federal Government and used for developing national policy and determining the funding of highways, the States comply.

The frequency of missing data is very low. However, whenever there is a change in the HPMS, such as the addition of the donut area information in 1993, there are some problems with the new data from some of the States. Typically, such problems are resolved by the second year of the requirement.

Coverage

FHWA has all the HPMS data from 1978 to the present. Individual States generally will have only their most recent few years.

The national universe data for 1 year contain about 3.25 million records. It is stored on tapes. Records go back to 1980.

The total national standard sample contains approximately 115,000 records per year. Again, these data are stored on tape. Records go back to 1978.

The areawide data for each State are submitted on a series of templates. At first, there were five templates that were submitted on paper. Later, spreadsheet templates were allowed. In 1993, the number of templates was increased to seven and spreadsheet templates (Lotus 1-2-3) were mandated.

Annually, FHWA transfers these records to a mainframe file and stores them on tape. One format was used until 1992. A new format (basically an ASCII file) was instituted in 1993.

The first submissions of the donut sample and line referencing systems 'were required in 1994. There are no archives of them at this time.

Measurement

The key variable in the sampling design of the HPMS is AADT. AADT is not directly measured (except for a very small number of continuous permanent counting stations in each State), but is either derived from short counts, factored from previous counts, or estimated in some other manner.

States are asked to maintain at least one automatic traffic recorder (ATR) on each route of the PAS/NHS and a minimum of three on both the rural and urban portions of the non-PAS/NHS highways. These are used to develop day of week and seasonal factors used for expansion of short counts to AADT.

Typically, volumes at the ATRs are measured with pavement loops. Pavement loops are prone to failure, especially in northern climates and from construction vehicles. However, failures at ATR stations are supposed to be repaired as soon as possible. Recently, other more reliable technologies have been introduced.

The HPMS methodology requires that traffic counts of at least 24 h be conducted on one-third of the road sections in the standard sample each year. These counts typically are taken with pneumatic tube-type portable counters. These are reliable and, if a problem is suspected, the count can be easily repeated. The vehicle volume is derived from these counts by adjusting for the number of multi-axle vehicles in the traffic flow.

The AADT for these sections is then calculated from the short period volumes, with the application of adjustment factors developed from volumes at the ATRs.

The AADT at the sites where traffic counts were not made in the current year is factored from previous counts at the site or by other methods (estimation, engineering judgment, tracing volume maps, etc.). The method of AADT estimation for each site is one of the data items for the sample.

Statistical Reliability

The HPMS standard sample design is a stratified simple random sample. The HPMS sample size estimation process was tied to the AADT. Of the approximately 80 data items collected, AADT is perhaps the most variable data item in HPMS. Therefore, the reliability of most other characteristics would be expected to exceed that of AADT.

The sample size for each stratum of the samples is prescribed in the HPMS Manual. The sample sizes per functional system vary by State according to the total number of road sections (universe), the number of predetermined volume groups, the validity of the State's AADT data, and the design precision levels.

For rural, small urban, and collective urbanized areas, sample sizes are based on 90-5 precision levels for volume groups of the Principal Arterial System (PAS), 90-10 for minor arterial system, and 80-10 for the collectors (excluding minor collectors).

For individual urbanized areas with populations > 200,000 that are in NAAQS non-attainment areas, the design precision is 90-10 for the arterial system and 80-10 for collectors.

For individually sampled urban areas with populations < 200,000, the precision levels are 80-10 or 70-15 depending on several other factors.

The only objective of the donut portion of the HPMS is to estimate the daily vehicle-miles traveled (DVMT) within the donut areas with a precision of ±10 percent with 90 percent confidence. DVMT is determined from AADT. Thus, the sample size for a particular donut area is based on the-variability of AADT in that donut area.

Data Format and Access

The templates for the areawide data and the data format for the universe, standard sample, and donut sample data are shown in the appendix. Note that the fields are marked with an A, S, or D indicating that this field is required for all records, standard sample records, or donut records, respectively.

To obtain these data files or some portion of these data, contact the Highway Systems Performance Division of the FHWA.

All data are available on IBM readable mainframe computer tapes. The types of tapes that the data are stored on correspond to tape technology at the time the data were collected.

The universe data file is extremely large, approximately 3.25 million records per year. It does not appear particularly useful for highway safety research. However, should a researcher have a need for this information, he/she would have to contact the Highway Systems Performance Division and work out the details of copying the desired tapes. The researcher would have to provide the tapes.

The standard sample data consist of about 115,000 records per year. All the available data sets (from 1978) can be obtained on mainframe cartridge tape.

The areawide data are available on mainframe computer tape in Extended Binary Coded Decimal Information Code (EBCDIC) format. These files can be obtained from the Highway Systems Performance Division on PC diskettes in ASCII format.

The HPMS is updated annually and a new HPMS is generated at that time. It is important to note that some of the data fields and even some of the overall structure of HPMS may change from year to year.

The HPMS data from the States for the previous year is due at FHWA on June 15. It becomes available outside the FHWA sometime at the end of the year. Thus, a researcher can get data from the 1993 HPMS in December 1994 or January 1995.

The FHWA contact for HPMS is:

David R. McElhaney, Director
Office of Highway Information Management
Federal Highway Administration
400 7th Street S.W.
Washington, DC 20590
(202) 366-0180

Reference

Highway Performance Monitoring System Field Manual. Federal Highway Administration. OMB No. 2125-0028. 1993.

Highway Safety Information System (HSIS), FHWA

General: The Highway Safety Information System is produced by the Highway Safety Research Center (HSRC) at the University of North Carolina.

Purpose: The FHWA has selected States for HSIS that provide linked accident, highway inventory, and traffic count data, and has converted the files to SAS format to provide an enhanced analysis capability. This introductory section provides only a general overview of the data. Descriptions specific to each of the States follow.

Source: VMT is estimated from segment lengths and AADT. The AADT volumes are updated from 1 to 5 years. Some values are estimated or interpolated; some sites are permanent and some are year-round. Most are temporary sites, taking 48-h counts. Some have vehicle classification, or "commercial," vehicle counts. 'Commercial" is usually any vehicle with two axles and six tires or more.

Coverage: States covered in this write-up include: California, Illinois, Maine, Michigan, Minnesota, North Carolina, Utah, and Washington. Additional States are being added to HSIS. In most States, a major portion (but not all) of the highway system is covered. Usually, these are the State-maintained roads.

Sample: The highway segments covered are usually a purposefully selected subset. Cross-section files in some States contain a sample of segments, usually limited in number.

Strengths: Sample size is large, and there is a diversity of data in different States. The files are in SAS format for convenience, and the documentation is better than usually available from the States. The data are suited for aggregate comparisons.

Limitations: The AADT data are sometimes coarse, and may not be suited for identifying individual, high-risk locations. Entering volumes for both roads of an intersection often are not available. National estimates are not possible. The diversity of data in different States can also be a disadvantage.

Accuracy: AADT volumes are not all observed and are not independent, so the variance cannot be estimated.

Also included at the end of this section is a brief discussion of the statistical implications of the nature of the traffic volume data in most State files. Issues discussed include the use of a purposeful sample rather than a random selection of sites for counts, and the use of estimated or interpolated counts rather than actual counts. A general conclusion is that the traffic volume data will not support a statistically defensible analysis (except when the HPMS procedures have been followed). However, a purposeful sample can be representative, although the variance is likely to be underestimated. Similarly, estimated or interpolated counts may also be reasonable in value, but again, the variance will be underestimated. When highway sections have been stratified prior to selecting sites, the most rigorous use of the data is to calculate estimates at the strata level. Use of the volume data to simply stratify the data into volume groups is also relatively sound.

Thus, the traffic volume data must be used with caution. The actual extent of any of these problems cannot be estimated without additional data. Estimated or interpolated counts mean that the observations are no longer independent, and most statistical techniques are no longer appropriate. In particular, the variance is underestimated and bias may be introduced. The analyst should be aware of the source of the traffic counts in each State and should use good judgment in the selection of an analytic approach. Though statistically sound analyses of accident rates may not be possible with the currently available exposure information, it may be possible to use this information in a productive way, e.g., for stratifying sites, and to perform within the strata only analyses relying on counts.

HSIS Contacts: Jeffrey Paniati at (703) 285-2057 or Yusuf Mohamedshah at (703) 285-2090

California, HSIS

Coverage: The current accident files cover the years 199 1 to 1995, and there is roadway information for 1993 and 1994. Accident reporting is not uniform in California, with some municipalities using their own report form and reporting threshold, instead of the California Highway Patrol (CHP) form. Accidents occurring on State routes (including those in urban areas that do not use the CHP form) are location coded. There are about 150,000 accidents annually on State routes (all with location codes) out of an estimated statewide total of 500,000 accidents per year. Reporting is also not complete for uninjured occupants. Information on uninjured occupants is only collected if there is al least one injured occupant. Thus, the occupant injury data are biased to overrepresent injured occupants. However, uninjured drivers have been identified in the driver file by Highway Safety Research Center (HSRC) by linking the injury information from the occupant file with the vehicle file. Overall, HSRC estimates that information on uninjured occupants is missing for about 50 percent of non-towaway accidents.

The roadway information is contained in three files: the Roadlog file, the Intersection file, and the Interchange Ramp file. The Roadlog file contains information on approximately 24,461 km of roadway, including about 3943 km of Interstate, 17,702 km of other primary highways, and about 2736 km of secondary/county/township roads. The 24,461 km are divided into about 50,000 records in the Roadlog file, for an average section length of 0.5 km.

The Roadlog file contains information describing the functional class of the road, cross section information such as width and number of lanes, as well as information on design speed, median barriers, and other special features. The intersection file has information of 20,000 intersections, and the Interchange Ramp file has information on 14,000 ramps. Accidents can be linked with all three roadway files and the intersection file can be linked with the associated segments in the Roadlog file, but the Interchange Ramp data cannot be linked with its associated interchange.

Exposure Information: The Roadlog file includes an AADT and a DVMT for each segment (record). Section length is also included. No information on truck travel is available. In the Intersection file, there is an AADT for the mainline road and for the crossing road, as well as descriptive information for both the mainline and cross road. AADT is also included in the Interchange Ramp file.

Traffic Data: As indicated in the preceding three sections, all three inventory files contain AADT information. In addition, the Roadlog file contains information on DVMT, which is computed as the product of the section length and section AADT estimate.

In California, the 12 district offices have the responsibility of collecting traffic data and developing the AADT estimates for each road section within their district. The Division of Traffic Operations of the Caltrans central office oversees the operation and attempts to maintain consistency in the methods and data across all districts as much as possible. If requested, Traffic Operations personnel will assist a district in calculating the AADT estimates. The division also maintains all count data on an on-line computer file for the districts' use.

There are approximately 2,100 permanent count stations on mainline highways operated by Caltrans in California. Of these, approximately 400 are permanent, continuous counting control stations that operate each day in a given year. Every major State-administered route is counted each year. The 400 permanent continuous count stations form a network that covers all major routes. The remaining control stations are permanent, quarterly counting control stations, i.e., in-pavement loops to which a counter/recorder device is attached for 7 to 14 days during each quarter. Caltrans also collects count data at approximately 700 of these quarterly counting control stations once every 3 years. In a given year, there are approximately 1,000 permanent quarterly counting stations where count data are not collected. California has determined that the AADT estimates, which are derived from the simple average of the four (unadjusted) quarterly counts, does indeed account for seasonal fluctuations without further adjustment based on nearby permanent counters. Consequently, there are no additional adjustments or corrections applied to the AADT's estimated from the quarterly counts.

In addition to the permanent control stations, approximately 1,000 coverage counts are collected annually. The intent is to collect coverage counts on a 3-year cycle (for a total of approximately 3,000 coverage counts), although conditions may force longer intervals in certain districts at times. A coverage count is basically a 24-h to l-week count.

Coverage counts are expanded to AADT estimates using factors derived from the combined continuous counts and quarterly count data. For road sections that are not counted in a given year, it is the responsibility of the districts to develop these AADT estimates. In some cases, the districts reply on overall traffic growth trends within the district. However, in most cases, the AADT assigned to the section is developed by studying the traffic growth in counts falling on each side of the section.

It is also noted that 24-h to l-week coverage counts are collected on approximately 3,200 on- and off-ramps per year. These ramp counts are manipulated through ramp balancing to reflect continuity of flow on mainline freeways.

Finally, vehicle classification data are collected at approximately 70 permanent stations across the State. Additional classification counts are collected on an as-requested basis, typically at locations where traffic count data are being collected. Since this is district-based, there is no reliable estimate on how many additional classification counts are collected across all 12 districts per year. Finally, there are approximately 45 weigh-in-motion stations statewide that provide speed, volume, and the "13-bin" vehicle classification information. (Taken from ***HSIS Guidebook for the California State Data Files.)***

Linking Accident and Exposure Information: Accidents can be linked with all three roadway files. Accidents are located manually using the scene diagram on the accident report and maps. Accuracy of the location is believed to be within 0.16 km, and missing data is only a few percentage points.

Illinois, HSIS

Coverage: During 1985 to 1994, this included 26,232 km of roadway of which 2,736 km were Interstate highways; 15,449 km of other primary roadways, and 8,047 km of secondary, county, and township roads.

Exposure Information: All exposure information is contained in the Roadlog file, which contains records for 197,000 sections; each section, on average, is slightly less than 0.16 km.

Exposure, in terms of VMT, can be calculated from AADT and the section length. In addition to the total, AADT for "heavy commercial vehicles" (defined as having two or more axles and six or more tires) is given.

Intersection information is in the Roadlog file and also in an Intersection Location file. They contain the same information, but the Intersection Location file contains one record for each intersection. If there is more than one intersection in a section, the information from the Roadlog file is repeated for each intersection record. Intersections are characterized as "across," "left," and "right. " The crossing road is apparently not identifiable. Thus, it appears that for intersection exposure only, the AADT on the through road is available.

Traffic Data: As indicated earlier, the Roadlog file contains information on AADT, percentage of trucks for 1990 and earlier, and commercial vehicle AADT for 1991 and later. These data are developed in Illinois' traffic volume counting program and are based on a combination of permanent counters that count traffic 24 h each day for 365 days each year and a series of short-term "coverage" counts conducted each year. Illinois has 49 automatic traffic recorders (ATRs), of which 21 are capable of collecting counts by vehicle class in accordance with FHWA's Scheme F. The ATR locations on the 5 different classes of roadway include 7 locations on rural Interstate roadway, 6 locations on urban Interstate locations, 12 locations on other rural roadways, 19 locations on other urban routes, and 5 locations on "recreational" routes.

In addition to the ATR data, short-term traffic counts on Interstate and primary highway systems are done on a 2-year cycle. During even-numbered years, portable counter devices are deployed in combination with pre-established in-pavement loop detectors. Typically, the counter devices are deployed during 1 week of the year at any given site. Short counts (e.g., 24- or 48-h counts) are collected on Monday through Thursday only. It should be noted that a sample of Interstate sections are counted 1 week out of every 4 months. During odd-numbered years, the Illinois DOT conducts a comprehensive interchange ramp counting program on State highways. These ramp counts are used to supplement ADT data for sections where the State did not have monitors (i.e., counter devices). In total, it is estimated that approximately 96 percent of the primary system is covered during each 2-year cycle.

For other non-primary roads (i.e., the "off" marked route system), Illinois collects 48-h coverage counts in approximately 20 percent of the counties once every 5 years. However, the northeast counties are done every 4 years. With the exception of Cook County, which is also on a 4-year cycle, urban areas within counties are counted on a 5-year statewide cycle.

Additional vehicle classification counts are conducted on HPMS sections. These are made at 300 locations over a 3-year cycle (i.e., approximately 100 each year) to form a representative distribution for-the State.

Finally, the districts often have a need for additional traffic data. Consequently, when requested, the State collects 12-hour turning movement counts at intersections and other "special" traffic data to satisfy these needs.

To convert the short-term coverage counts to AADT, Illinois applies adjustments to reflect corrections for number of axles and for seasonal differences in the daily traffic. Axle corrections are developed from both permanent classification counters and from manual (HPMS) counts. For seasonal corrections, each coverage count location is assigned to one of the five categories of roadway where permanent counters are located, as defined above. The seasonal factors are based on averages from all ATRs in that group.

When a road section is not counted during a given year, growth factors are developed and applied to the most recent prior year's count. Average growth factors are created each year for each functional class of roadway using ATR data and data from adjusted short counts for the current year. The growth factor applied to a particular uncounted section is based on its functional class. For sections where no prior AADT exist, AADT/mile averages by functional class are developed and then used in order to "fill in" the AADTs.

Finally, it should be noted that the percentages of truck-related "Heavy Commercial Volumes" include "two-axle trucks with six or more tires plus multi-axle vehicles." Thus, while pick-ups and vans are excluded, this combination would include single trucks, tractor-semi combinations, and buses. Thus, it cannot be considered a count of just the multiple unit (tractor-trailer) trucks that are found on the roadway system. (Taken from the ***HSIS Guidebook for the Illinois State Data Files.)***

Linking Accident and Exposure Information: Data on different files can be linked by a linkage key, which combines county, route prefix, and route number with the station number.

For intersection accidents, the intersecting route number and route prefix are given. However, it does not appear possible to identify which vehicle approached the intersection from the main road and which one approached from the crossing road. The direction of travel for each vehicle is given, but the direction of the road is not given in the Roadlog file.

Maine, HSIS

Coverage: The Link Record file covers all highways in Maine, including local roads and urban streets. The 35,405 km are divided into 67,000 links. Files are currently available for the years 1985 to 1994.

Exposure Information: The Link Record file contains AADT for each link; the year of AADT; and whether it is an actual count, an interpolation, or an estimate. Together with the length of the link, VMT can be estimated.

Information on intersections is available from the Node Records file, which also includes nodes other than intersections. The configuration of each intersection is given, and up to six legs are identified by the corresponding link numbers. As an exposure measure, only the total number of vehicles entering the intersection is given. However, it is possible to obtain the AADT for each leg from the Road Link file.

Traffic Data: With respect to the traffic information on both the Link and the Node files, the traffic counts that are in the system are extracted from a traffic file again prepared within the Bureau of Planning. The counts are extracted from a series of 54 permanent count stations across the State, 6 of which do detailed vehicle classification counts. There are a total of 9 stations on Interstate routes (which collect counts in both directions), approximately 13 stations on U.S. routes, 24 stations on State routes, and 8 stations on other routes.

In addition to the continuous count stations, each summer, 48-h counts are done at between 1,600 to 2,200 locations on all US and State highways. Beginning in 1994, the number of coverage counts increased to between 3,600 and 4,200. Approximately 10 percent of these counts include vehicle classification counts. Classification estimates exist for other locations that are not high-priority locations.

Each year, these counts are done in either the northern, central, or southern areas of the State. The counters move to a different area the following summer, covering the entire State every 5 years. The southern and central areas are counted in alternate years for the first 4 years of a cycle. Then, the northern area, where counts change less per year, is counted during the fifth year of the cycle.

Seasonal adjustment factors for the coverage counts are based on continuous count stations that fall into the same "highway type" category as the coverage count. Based on extensive analysis in the late 1980's, the three categories used are Urban (including suburban locations), Arterial (including all Interstate locations plus other locations in rural areas), and Recreational locations (whether urban or rural). The actual adjustment factor for a given coverage count location is based on the weekly average ADT for all continuous count stations falling into that category.

For years in which no count data were collected within a given area of the State, historical daily traffic flows are factored up on a county-by-county basis. The growth factor used is based primarily on traffic changes at the continuous count stations falling into the same highway-type category described above. Other information used in developing a specific growth factor includes counts from nearby urbanized areas and special counts that may have been conducted at the location for other reasons. The final growth factor used is based on interpolation between points of known growth (such as 2 or more years at the similar continuous count stations), and is done by personnel with a working knowledge of the system's traffic patterns.

In summary, while some of the counts may be off due to roadside development and/or roadway construction within a specific area of the State that occurred within the 2-year period, in general, the count data are felt to be quite adequate for analysis purposes. (Taken from the ***HSIS Guidebook for the Maine State Data Files.)***

Linking Accident and Exposure Information: Accident and exposure data can be linked by the low and high node numbers that identify each segment and by the distance from the low node given in the accident record.

Intersection accidents are identified as such, distinguishing three-, four-, and five-leg intersections. However, the leg from which a vehicle entered an intersection cannot be determined.

Michigan, HSIS

Coverage: Of 189,897 km of roadway in Michigan, the Roadway Segment file covers only 15,449 km of trunkline divided into 43,000 segments. Data for the years 1985 to 1994 are currently available.

Exposure Information: The Roadway Segment file shows AADT categorized into 10 classes. Commercial AADT is also given. No definition of "commercial" is shown. AADT for the segment is given.

A Cross Section file covers 8,047 km of two-lane rural roads with segments selected by a stratified random sample. Very detailed roadside feature information is given. However, there is no information on sample stratum. ADT values are given based on counts in the early 1980s. Counts of accidents by severity are given.

There is an Intersection file that has recently been released for analysis. However, information on AADT or vehicles entering the intersection is not provided.

Traffic Data: As noted above, information on AADT and Commercial Vehicle AADT is found on the Roadlog file. These data are developed in Michigan's traffic counting program, which, like other States, includes both full-time permanent counter locations that operate 365 days each year and short-term coverage counts at a much larger number of locations. Michigan DOT currently operates and maintains 121 permanent traffic recording (PTR) stations. These PTRs include 34 on Interstates, 31 on U.S. routes, 23 on Michigan State highways, and 12 on other routes.

In addition, there are a varying number of short-term "coverage counts" conducted each year. Michigan DOT indicated that approximately 3,300 such 48-h "short" counts were requested in 1995. These coverage counts included the following:

- 950 short counts (volume only).
- 1,300 classification counts (volume by vehicle class).
- 1,000 interchange ramp counts.

Michigan attempts to count every State-maintained road section in a 3-year period. Unless required under the HPMS, Michigan also attempts to collect classification counts over a 6-year cycle. It should be noted that in addition to the State's traffic counting program, other agencies (notably those in urban areas) are also collecting traffic data for HPMS purposes. Furthermore, the Metropolitan Planning Organizations (MPOs) in Michigan have developed and supported urban transportation planning models in accordance with ISTEA requirements. These MPOs subsequently have their own counting programs to support their model development and application.

To factor up the short counts to reflect AADT, seasonal factors are developed. Unlike some States where these seasonal factors are based on PTR counts within the same functional class as the short-count location, Michigan has defined six or seven "cluster-analysis groups." Each of these groups contains a number of PTRs, and the adjustment factors are based on averaging the PTR counts within that group. Each roadway section (and thus each short count) is assigned to one of these cluster-analysis groups.

When a specific roadway section is not counted in a given year, its count from the previous year must be adjusted to represent traffic growth. Here, Michigan attempts to "look up and down the road" and identify the closest, comparable section for which an ADT was estimated (counted) for the given year. They determine the percentage change (e.g., increase or decrease) in the ADT associated with that "comparable" section, and apply that percentage change to the historical count for the specific section in question. (Taken from the ***HSIS Guidebook* for *the Michigan State Data Files.)***

Linking Accident and Exposure Information: Though the Roadway Segment file covers less than 10 percent of the total highway mileage, about one-third of all accidents can be matched with locations on the Roadway Segment file. Linking can be done via information on the control section, and the milepost.

Accidents that occur within 30.5 m of an intersection with a trunkline road are coded for that road with the milepost of the intersection.

Minnesota, HSIS

Coverage: Coverage includes the years 1985 to 1994; however, some files are available only for certain years, and there were changes between the years. Files detail 19,311 km of primary roadways, an additional 37,014 km of State-maintained systems, and 157,711 km of county and local roads.

Exposure Information: Two files provide exposure information: (1) the Roadlog file and (2) the Intersection/Interchange file.

The Roadlog file contains information on about 200,000 road sections on which highway characteristics remain constant. Exposure in terms of VMT can be obtained from the values of AADT given for the segment, and the given length of the segment. Also given is "commercial" ADT. Commercial vehicles are defined as having at least two axles and at least six tires. Exposure estimates can be stratified according to the highway characteristics contained in the file (also according to AADT or AADT per lane).

The Roadlog file identifies the type of intersection at the beginning of a segment. However, it does not identify the intersecting road. Thus, intersection exposure cannot be obtained from this file.

The Intersection/Interchange file contains data on 3,500 intersections, 256 interchanges, and 2,800 grade crossings, currently for the years 1987, 1989, and 1991. Intersections were originally selected for the purpose of identifying high accident locations, but are retained in the file.

Intersection type and a code describing it in some detail are given. The route on which each approaching segment is located is identified, and there are up to two legs for each segment. The direction (N, NE, E, etc.) of each leg is also shown. This allows reconstruction of the configuration of the intersection. For each leg of each segment, the AADT for several years is given. For the second leg of a crossing minor roadway, in 10 percent to 30 percent of the cases, AADT is missing. In these cases, it is recommended that the value for the first leg be used. Thus, the available exposure for intersections consists of AADT on the intersection approaches.

Commercial AADT is not given for intersections. However, it appears possible, though cumbersome, to obtain this information from the Roadlog file.

Traffic Data: The Traffic file contains information related to AADT data for all roadway sections across the State. This information is manually derived from sample and continuous counts taken at temporary and permanent count stations throughout the State. It contains total AADTs and AADTs for heavy commercial vehicles (which are defined as vehicles with two axles and six or more tires).

Like other States, Minnesota develops traffic volume estimates based on automatic traffic recorder stations (ATRs) and short-term (48-h) "coverage" counts. There are approximately 120 ATRs that count traffic 24 hours per day, 365 days per year, across the various roadway types. These are located on all classes of both rural and urban highway, with approximately 55 percent of the locations being on urban roadways and 45 percent on rural roadways.

In addition, there are approximately 34,000 coverage (temporary) count locations across the State where 48-h counts are made. Approximately 12,000 of these locations are covered each year. For the trunk highway system (including Interstate roads), these counts are made on a 't-year cycle, as are counts on roads within the Twin Cities metropolitan area. For the lower order County State-Aid Highways and the Municipal State-Aid System outside the Twin Cities metropolitan area, the counts are made on a 4-year cycle.

The seasonal adjustment factor for a given coverage count is based on counts made at ATRs which are similar to the coverage count location. Here, ATRs are grouped into the following classifications:

Outside (i.e., non-metropolitan area)

- Rural, farm-to-market roads.
- Rural, weekend recreational road.
- Rural, summer-peak recreational road.
- Municipal, non-recreational road, less than 5,000 population.
- Municipal, non-recreational road, more than 5,000 population.
- Municipal, recreational road, less than 5,000 population.
- Municipal, recreational road, more than 5,000 population.

Metropolitan Area

- High commuter route.
- Commuter shopper route.
- Low recreational route.

Seasonal adjustment factors, based on the data for the previous 3 years, are developed for each classification and are applied to all coverage counts collected at locations within that classification.

For the "non-count" years, a growth factor is applied to the previous year's data based on changes in counts at the ATR stations located on the same functional class of roadway. When new data are available at the end of the next count cycle, these data for the interim non-count years are readjusted to represent the average of prior and subsequent count years (e.g., a 1987 "non-count" year estimate based on the growth factor would be readjusted to represent the average of 1986 and 1988 counts at that location as soon as the 1988 count year was completed).

In developing AADT estimates for each section of roadway, there are sometimes road sections with no historical count data (e.g., lower order local facilities, including township roadways and local streets). In these cases, an original "baseline" estimate is based-on ATR counts on lowest order roadways with the lowest counted volumes. Growth factors for these uncounted sections are also based on this same ATR group.

MinnDOT also collects vehicle classification counts at about 300 sites per year. These are 16-h (e.g., 6 a.m. to 10 p.m.) manual classification counts usually over 2 different days. In addition, portable vehicle classifiers are deployed to collect 48-h data. Currently, there is no program to seasonally adjust the classification counts. There are an additional 25 weigh-in-motion stations statewide that collect classification data. However, these data are used less than the manual classification counts.

The new count data are placed in the Traffic file within the first 6 months of the subsequent calendar year. While the Traffic file can also be thought of as a "Section" file (with a specified AADT at the beginning count station being assumed constant over the entire section), it differs from the Roadlog file to which it will often be merged in that the beginning and end points (termini) are often located at different points on the roadway. The linking variables are again the route system/route number/reference point (milepost).

There are approximately 208,000 records on the file, but these do not represent a one-to-one match with the 200,000 "true" records on the Roadlog file. Often, there are Roadlog sections with multiple Traffic file records (i.e., multiple count stations), and often there are Roadlog sections with no Traffic file records (i.e., corresponding count stations) located within the section.

Each raw file record contains up to 30 years of AADT information (with the related year "attached"). Thus, to determine the average AADT for a given year for a series of sections on a given route: (1) the traffic section reference points must be matched with the appropriate Roadlog sections by comparing the reference point with the beginning and ending milepoint on Roadlog sections (with the ending milepoint being "assigned" as being equal to the beginning milepoint on the succeeding section), (2) the appropriate yearly AADT for each contained Traffic file record must be extracted, and (3) the counts must be averaged for sections where multiple Traffic file records exist. If no Traffic file record exists for a given Roadlog section, then the section AADT is assumed to be equal to the AADT at the previous (upstream) traffic section on the same route. (This is the assumption made by Minnesota and by HSRC programs. However, other procedures could be followed in calculating AADT if they are felt to be more appropriate for a given research question.) Any AADT assignment program developed must not carry over counts from one route to another; this is a mistake that can easily be made since the Roadlog file is a continuous file in route order. Obviously, averaging traffic over more than 1 year will require additional programming.

Currently, there are two HSIS SAS-formatted Traffic files — one developed for 1987 and earlier data, and one containing data for only 1988 and 1989. Again, please note that traffic data were merged with the Roadlog file for years 1987 through *1994*. The Traffic file still remains a separate file on the HSIS system for the years 1987 through 1989. It is no longer available as a separate file on the HSIS system after 1989.

The first Traffic file (1987) is similar to the raw file in that it contains up to 10 years of data, with 1987 counts being the most recent data. The second file (1988-1989) contains only counts for 1988 and 1989. Each record on the file contains information on traffic counts for one year for a given location. To combine across years for a given counter location, records with the same location information can be merged.

To make the AADT information even more easily usable in subsequent analyses, HSRC developed a linking program that links the basic AADT information from the SAS Traffic file with the Roadlog file to produce a separate single "Average AADT" variable for each Roadlog section on each of the two Roadlog files (i.e., 1985-1987, 1988-1989). Where necessary, averaging across traffic sections in a given Roadlog section for a given year and "carrying down" AADT information from the prior record have been done in this linkage program. Since the 1987 Roadlog file is used with accident data from 1985-1987, and the 1989 file is used for 1988-1989 accidents, the AADT variable on each Roadlog file represents an average AADT over the respective time periods. That is, the 1987 file contains average AADTs for 1985-1987, and the 1989 Roadlog file contains average AADTs for 1988-1989. Different AADTs (say for individual accident years) could be developed by modifying the existing computer program.

Since it is not possible to perform an independent "check" of the accuracy of the AADT information, it is assumed that the procedure in place in Minnesota to monitor count stations and update the file provides adequate information. As indicated above, these are felt to be excellent data for the trunkline system where they are updated on a 2-year cycle. There are also fairly good data for the county State-aid systems, which are generally updated on a 4-year cycle. (Taken from the ***HSIS Guidebookfor the Minnesota State Data Files.)***

Relating Accident and Exposure Data: Accidents are located by information on the route system, route number, and a "reference point." This information allows an accident to be attached to the appropriate section of the Roadlog file.

Accidents in an intersection can also be attached to the Intersection file by using route system and number, and the reference point.

Apparently, the approach from which a vehicle entered an intersection cannot be identified, except possibly by matching the direction of travel with the direction of the approach from the Intersection file.

North Carolina, HSIS

Coverage: The current HSIS files for North Carolina cover the years 1990-1995. Accidents are linked to the Roadway Inventory file with a computerized referencing system that currently covers about 38 percent of the estimated 148,056 total road kilometers in North Carolina. The reference systems covers all 22,530 km of primary routes, and an additional 33,473 km of secondary roads (rural secondary roads and city streets). There are no "county" roads in North Carolina, since all are under State control. This system links about 60 percent of the accidents (118,000 out of 192,000) to a road segment in the Roadway Inventory file.

Exposure Information: The Roadway Inventory file describes homogeneous road segments defined by a beginning and ending milepost. An AADT is provided with the year in which the count was taken and the section length in miles. The percent trucks in peak traffic is available for about 40 percent of the sections and an off-peak percent trucks is available for about 10 percent of the sections. The roadway variables include roadway width, number of lanes, lane width, shoulder type and width, median type and width, surface type, whether the section is in the HPMS sample, a traffic growth factor, and other variables.

Currently, intersection and interchange information cannot be linked with accident as the descriptive information is not available in a suitable format. The available information on roadway segments does not include information on horizontal curvature, vertical grade, or passing sight distance.

Traffic Data; As indicated above, the basic AADT and percent truck information is included on the Roadway Inventory file. The traffic count information used in the development of these variables is developed from a series of permanent control count locations and spot counts across the system. Currently, there are approximately 100 ATRs across the State. These are permanent full-time counters that are used both for counts at their location and to establish seasonal and growth factors used with spot counts from surrounding locations.

In addition to these permanent stations, there are approximately 60,000 points in the State where 24- to 48-h counts are made. The entire primary and Interstate system is covered each year. Fifty percent of the secondary roadway system is covered each year with the remaining 50 percent being done in the alternate year. The spot counts are linked with a group of nearby ATRs in order to establish distributional factors. The data are reviewed internally by the inter-office traffic staff, edited, quality control is checked, and then factors are developed. The traffic counts are closed out for the count year in October of each year and then sent to the roadway inventory staff for inclusion in the Inventory file.

Ramp counts are made each year on all interchange ramps on the Interstate system. These ramp counts are used to generate turning volumes and to balance counts on the mainline for the Interstate and crossing roadways. This represents approximately a 2-week count on each ramp. Past ramp counts are found on paper file, but have been computerized since early 1993.

Truck counts are made on a 3-year cycle at 300 vehicle classification sites across the State. The 300 count locations are not necessarily at all of the ATR sites. There are approximately 90 truck weigh stations in the State related to the SHRP program. In addition, it was noted that truck counts are made every 3 years on all HPMS sections in the State.

Finally, for intersections that are in the State's Traffic Improvement Program, turning counts are done on an as-needed basis. These turning counts include both a.m. and p.m. peak traffic, with each count being conducted for approximately 7 h. It is estimated that approximately 500 of these are done each year. These are found in a paper file, which may be computerized in the next 1 to 2 years.

Examination of the traffic-related variables in the HSIS Inventory file indicates that ADT is present for 99.9 percent of the sections. However, what is missing is data on percent trucks. Here, the variable concerning "Percent Trucks at Peak" is uncoded for approximately 60 percent of the mileage. The variable related to "Off-Peak Percent Trucks" is uncoded for almost 90 percent of the mileage. Conversations with department of highways staff indicated that this is the result of the fact that these variables are only coded if there is fairly high confidence in the percentages. This would occur if a classification count had been done on the section (as in an HPMS sample section) or on an adjacent or nearby section. Thus, while the data present should be fairly accurate, data are missing for a large number of miles.

Linking Accident and Exposure Information: The linking system for the accident data is unusual in that it is based on a "paper" reference system. The linkage information is the county, route, and milepost. However, there are no physical mileposts on the roads. The investigating officer records the distance and direction to a reference point that may be an intersection, bridge, or city boundary. Mileposts are determined in a computerized referencing system, based on the location of the reference given. The accident is linked by using the milepost generated by the computerized reference system to locate the section in the Roadway Inventory file, which includes this milepost within the beginning and ending milepost defining the section. Nearly all accidents on the primary road system are linked with this system, plus a large number of accidents on the secondary roads. About 90 percent of the mileage in the reference system is in rural areas. About 80 percent of the rural accident locations are believed to be accurate within 0.16 km, and 80 to 90 percent of the urban accident locations are thought to be accurate within 30.5 m.

Intersection characteristics are not currently available for linkage with the accident data.

Utah, HSIS

Coverage: Accident data for 1985-1994 are included, but highway data for 1990 are not available.

Of the 80,465 highway kilometers in Utah, 69,200 km are on the Roads file. However, only 20,599 km of these have inventory information and can be used for analytical purposes.

Exposure Information: The Roads file contains AADT for each section. Also given are the percentage of trucks in off-peak periods and the percentage of commercial vehicles in peak periods. No definition of "trucks" and "commercial vehicles" are given. Together with the segment length, VMT can be estimated.

No separate information for intersection exposure is available. The only information given for intersections is the number of intersections by segment, also separated by type of control. The intersecting roads are not identifiable.

For the State-controlled system, a Horizontal Curve file and a Vertical Grade file are also available. They allow disaggregation of exposure by grade and curvature.

For a random sample of sections of two-lane roads, a Cross Section file is available. It contains extensive information on cross-section and roadside features, including trees, posts, hydrants, recovery area, etc. This would allow the inclusion of highly specialized exposure measures, such as the number of trees passed, etc. Counts of accidents by severity are also given.

Traffic Data: As noted earlier, traffic data related to AADT and truck percentages are found on the Roadlog file. These data are based on Utah's traffic count program. In this program, there are 85 permanent ATRs on Interstate and Utah State roads that are in operation 365 days/year. Of these, 53 ATRs capture volume and vehicle classification counts and 32 ATRs count volume only. These ATRs conform with FHWA's HPMS guidelines. In addition, there are approximately 10 ATRs on roads inside National Parks in Utah that are operated by the National Park Service.

In addition to these permanent counts, Utah collects 48-h coverage counts at approximately 1,000 locations per year. Counts on the State-system roadway are done on a 3- to 5-year cycle. Approximately 100 traffic counting machines are used to collect traffic data for 11,426 km of State-system roads in Utah. In terms of coverage, Utah tends to have a better sample coverage of high-volume roads compared to lower functional categories. From a purely statistical perspective, a larger sample might be more appropriate for the lower functional classes of roads. However, Utah believes that limited resources for counting should be devoted to the roads that carry the bulk of the traffic. In addition to these coverage counts, approximately 100 short-term vehicle classification counts are conducted each year.

Short-term counts are expanded to AADT estimates using ATR data for roads with similar characteristics, functional class, and volume group. For a year in which no count is made, the previous year's count for a section is modified by a "growth factor" that is based on data from an "assigned" (similar) ATR station, area count data, and/or estimated statewide averages. In this manner, volume assignments are made to each section of State-system roadway each year. Finally, Utah staff also develop estimates of truck percentages and equivalent single-axle loadings (ESALs) for "on-system" roadways. Traffic information is entered into the Traffic file as it is being collected, but is transferred to the computerized system and, thus, to the Roadlog file at the end of the year.

With respect to the accuracy of the traffic information, Utah staff indicated that the data are currently being corrected so that errors would probably not be greater than ±10 percent for almost all of the sections. (Taken from the ***HSIS Guidebookfor the Utah State Data files***.)

Linking Accident and Exposure Information: Accident and highway files contain the route number and milepost, which allow linking of the data. Intersection accidents can be identified by a code based on the officer's intersection sketch. However, they cannot be linked to a specific intersection in a segment, except if there is only one in a segment.

Washington State, HSIS

Coverage: The current HSIS files for Washington State cover the years 1993-1995. Data for 1991 and 1992 will be added later when it is available. There are approximately 120,000 accidents per year in Washington State. Approximately 42,000 of these occur on State routes, and are location coded manually, based on the scene diagram and location information on the accident report. About 20 percent of these are "citizen" reports. Omission of these citizen reports reduces the located accidents on State routes to about 34,000.

A total of 13,840 km are described in the Roadlog file. This mileage includes 11,748 km of mainline roads, and another 2092 km of ramp front and other non-mainline roads. For example, information on each ramp for 876 interchanges is included. Interstate, U.S., and State routes are included. About 85 percent of the mileage is rural and there are about 1408 freeway kilometers. Each record describes a homogeneous section of road, as created by HSRC from point-by-point files supplied by the State. There are a total of 41,000 sections at an average section length of 0.3 km. Although the points at which intersecting roads cross are identified, there is not sufficient information (milepost) to link in the section data for the crossing road. Thus, the Washington State data do not appear well suited to an analysis of intersection accidents.

Exposure Information: The Roadlog file includes the beginning and ending mileposts and section length, the latter two calculated by HSRC. AADT is also given. By linking with the Traffic file, additional weekday and weekend counts are available, as well as single- and double-trailer truck volume. The available roadway characteristics include surface width, lane width and type, shoulder width and type, median information, functional class, posted speed, and other information.

The Traffic file created by HSRC describes road sections with approximately constant volume. The beginning milepost is identified, and the endpoint is found as the beginning milepost for the next record. However, one must check that the route has not changed. Additional section files describe 33,000 vertical grade sections and 14,600 horizontal curve sections. These can also be linked with the Roadlog file based on beginning and ending mileposts.

Traffic Data: As noted above, traffic count data captured on the Trips file, and thus in the HSIS system, contain a number of variables. These include AADT, average weekday volume, average weekend volume, single-trailer truck percentage, double-trailer truck percentage, and various peak-hour descriptive percentages. While AADT information has been merged into the HSIS Roadlog file to facilitate rate-based analyses, the other variables can be linked with the Roadlog file through linkage variables contained in both files.

In the base traffic file from which this information is derived, a new record is begun when there is a change in the AADT. The traffic census staff go through each of the inventory groups and identify what they feel are "discontinuities" along the routes in terms of volumes. These discontinuities would represent locations where the staff expect there to be significant changes in the AADT, such as an intersection with a significant turning volume or the location of a major traffic generator such as a shopping mall exit. In short, the Traffic file is a set of "homogeneous traffic sections." Thus, even though the file is organized as "point data" with only a "beginning" milepost, the data should not change until the next milepost. (In using and merging the file, some caution should be taken to ensure that the next milepost on the file is within the same route.)

The basis for the traffic information is a series of permanent and non-permanent count stations across the State. There are 117 permanent ATRs in the State as of December 1993; all 117 produced volume counts. Of these permanent count stations, 70 produced vehicle classification counts, 32 produced truck weight plus classification counts, 22 produced vehicle length counts, and 47 produced speed counts.

In addition to the permanent count stations, the traffic census staff conducts approximately 3,500 weekday counts each year. Each of these is a 72-h, Tuesday through Thursday count. Approximately 400 of these include additional vehicle classification counts each year. The counts are not always taken at the exact same sights, but do cover all HPMS locations as well as certain project counts that are conducted each year. In Washington State, there are 3,200 HPMS sections. The traffic staff feel that there are approximately 5,000 unique "homogeneous traffic" sections in the State each year. Counts are made at each of these locations every other year or every third year. In addition to these counts, there are ramp counts done at 120 to 150 interchanges each year.

With respect to accuracy and completeness, the DOT staff feel that they have very good data on approximately 90 to 95 percent of the roadway in the trips system. They feel that the least accurate information on the file is the vehicle classification counts. This is due to the limited number of count stations that are, by necessity, available for these type counts. However, traffic census staff are working toward increasing the accuracy of these truck counts. Their current feeling is that the variable related to daily truck percentage in the peak hour now contains good data. The overall truck count system was redone in 1987. One of the current points of interest is to try to expand the seasonal factors for trucks to make these even more accurate.

As noted under specific variable descriptions in the later format section, certain other variables (such as "Peak Hour Percentage" and "Peak Hour Split") have significant numbers of uncoded ("zero") locations. These represent locations where counts were not made or where old, erroneous counts have been deleted from the file. Washington State staff recommend carrying forward values from the preceding valid count location in these cases.

Linking Accident and Exposure Information: County, route, and milepost in the accident files can be used to create an 1 l-character variable that can be linked based on the route identifier and the beginning and ending mileposts in the Roadlog file. In the Traffic file, the beginning milepost is given, and the endpoint is assumed to be the beginning of the next record after checking that the route is the same.

Intersection volume and characteristics are only available for the mainline roads. Information for the crossing road sections cannot be linked.

Exposure Information in Highway Files

Highway files typically contain AADT for each segment in the file. Sometimes additional information is given, e.g., AADT for commercial vehicles or peak ADT. Together with the section length, AADT allows calculation of VMT on that section. If a segment ends at an intersection, AADT provides the number of vehicles entering and leaving the intersection from each approach. For an intersection within a segment, the same values must be assumed for the two approaches on this road.

In a formal sense, this provides enough information to calculate and analyze accident rates. However, if accident rates or accident counts in relation to AADT are used in statistical analyses, then the statistical characteristics of the AADT information in the files need to be known.

There are basically three types of accident studies:

(1) Making and comparing aggregate estimates.

(2) Studying relationships between accidents and highways and other factors using segments or intersections as observations.

(3) Identification of hazardous locations-"black spots."

The statistical characteristics of the AADT information affect these analyses in different ways.

The AADT values for the many sections of a highway file are derived from relatively few actual counts. At continuous counting stations, counts are made 24 hours a day, 365 days a year. At temporary counting stations, counts are made for usually 24 or 48 h, at intervals of 1 or several years.

There are two statistical questions: (1) what are the sampling characteristics of the actual counts, and (2) how are the AADT values for the sections without counts obtained from those for the sections with counts?

The answers to these questions determine the statistical analyses that can be validly performed with accident rates as dependent, or AADT as independent, variables.

To allow generalization beyond the sites with actual counts, sites should be randomly sampled from a well defined "frame," e.g., all sections on Interstate highways. This is often not done. Historically, "judgment" samples have often been made. Sites were selected that experts thought to be "typical" or representing the entire range of highway characteristics. While a judgment sample can give unbiased estimates, one cannot be certain of this. In particular, one cannot validly predict the errors of estimates based on judgment samples.

At the temporary counting stations, there is also sampling over time. If the counting is not done during certain parts of the year only, but year-round, sampling over time may be adequately close to random sampling.

Statistical analyses of a sample obtain estimates for the total sampling frame: totals or averages. In this application, it would be the number of all vehicles entering intersections on the highway network constituting the frame or the AADT representing an average over all sites on this highway network.

If the sample is stratified, then the estimates apply to each stratum separately, and estimates for all strata combined can also be obtained.

Such estimates can be used for studies of broad questions, e.g., comparing accident risks among highway systems, among highways with different numbers of lanes, classes, and intersections, etc. The level of detail such studies can consider is limited, because each stratum provides a single observation. However, if a detailed sampling plan is developed that stratifies according to many factors and their interactions, then even if the minimum of two sampling sites per stratum is used, detailed analyses may be possible.

One limitation of this type of analysis is that it does not allow identification of high-risk sites or "black spots." Highway data files contain information that, in principle, allows identification of such black spots, e.,g., AADT for short highway sections. With this information, an analyst can calculate accident risks for sections and intersections, and identify high-risk locations. However, without fully understanding how the AADT values for the individual sections are obtained from the relatively few sites with actual counts, the analyst cannot assess the statistical characteristic of the AADT values, and analyses based on them may be invalid. One approach is to assign to each section the value of the preceding section, until a section with an actual count is encountered, then carry over this count, etc. An alternative is to linearly interpolate AADT on the sections between connecting stations. While such approaches may give realistic order-of-magnitude estimates, and may even be quite realistic under certain conditions, this is not guaranteed. Thus, estimates of accident rates based on them can be biased and unrealistic. A more subtle, but not less important, aspect is that the estimates are not independent. Usually, the estimates on adjacent sections are positively correlated. A consequence is that analyses, which are using individual sections with their accident counts and AADT values as observations, tend to underestimate the uncertainties and errors of the results. They may also lead to the identification of "black spots," which appear to have unusually high accident risks only because the variability of the calculated rates is underestimated. Therefore, the statistical value of AADT figures by segment, without indication from which stations and by which method they are derived, is very limited.

Long-Term Pavement Performance (LTPP)

Historical Summary and Purpose: The Long-Term Pavement Performance (LTPP) program is a 20-year research project begun in 1987 as part of the Strategic Highway Research Program (SHRP). During the early 1980s, the Transportation Research Board (TRB) of the National Research Council, under the sponsorship of the Federal Highway Administration (FHWA) and with the cooperation of the American Association of State Highway and Transportation Officials (AASHTO), undertook a study of the deterioration of the Nation's highway system!) The SHRP was established on the recommendation of this study to focus research and development activities aimed at improving highway transportation. The Long-Term Pavement Performance program was one of six key research areas identified by this study.(2) The LTPP program is a comprehensive program to "satisfy the total range of pavement information needs" drawing on "technical knowledge of the pavements presently available and seeking to develop models that will better explain how pavements perform . . . this includes specific effects on pavement performance of various design features, **traffic and environment,** etc." The traffic and environmental data contained in the LTPP data collection plan are of potentially extreme interest as measures of exposure for highway safety issues as well. The concept of a traffic database, later named the Central Traffic Database (CTDB), originated in 1989 when the Expert Task Group concluded that the volume of traffic and load data that would be collected over the 20 years of the LTPP program required a separate database.

Data Contents and Structure: The LTPP data are housed in seven modules. A brief description of those modules that could be of interest in highway safety studies is described below:

(1) Climatic module.

Data derived from the National Oceanic and Atmospheric Administration (NOAA) weather data. Climatic data include site-specific estimates (based on the five closest weather stations) of various temperature, precipitation, humidity, and solar data statistics on a monthly basis for each test section, as well as actual values for the weather stations.

(2) Inventory module.

Data that identify the site and describe the pavement at the time the section was chosen. Data include location, material properties, composition, construction improvements, etc.

(3) Maintenance module.

Data describing all maintenance activities associated with the site.

(4) Monitoring module.

Friction, deflection, and distress data that could be of interest in wet pavement accident studies, etc.

(5) Traffic module (Central Traffic Database [CTDB]).

Historical and monitored traffic data. Yearly estimates of volumes, axle loads, and equivalent single-axle loads are available for each site. Also, data on truck weights and distributions are available at 789 sites quarterly for 7 days. Approximately 35 percent of these sites have weigh-in-motion collectors and the rest are Automatic Vehicle Classification (AVC) counters.

Experimental Design, Sample Plan, and Location Distribution: Data are collected in four geographic regions by regional staff members. With regard to traffic data, staff members are responsible for reviewing and processing the traffic counts, classification, and weight data, as well as ensuring acceptable collection procedures. The regional offices transmit their data to the national LTPP Traffic Database. Here, the data are further scrutinized and edited and it is the responsibility of this office to decide what data are of sufficient quality to release to the general public.

Traffic data are collected on more than 789 sites on key highway routes. In addition to new traffic data collection, historic traffic data were also requested where available. There are generally two types of traffic data available — vehicle count and classification data (Automatic Vehicle Classification [AVC] devices) and vehicle count and weight data (Weigh-in-Motion [WIM], either permanent or portable). The location of the WIM data collection may not always be exactly at the site, especially near interchanges. For the purpose of safety analyses, it is important that the researcher verify the exact location of the traffic data. These data have been of varying quality and one of the future objectives will be to back-validate some of the historic data with the new data, incorporating trends established based on the new data. Figure 1 show the geographic regions and Table 1 lists the number of locations by State for these locations. (Note: A revised table will be submitted that identifies locations that have WIM equipment and that have AVC equipment only when it is available).

Data Acquisition and Documentation: Information from the LTPP studies is available from the LTPP Information Management System (IMS), a database developed under SHRP. The pavement performance data are stored in the National Information Management System (NIMS) located at the TRB in Washington, D.C., and are updated on a regular basis. Similarly, the more detailed traffic data are housed in the CTDB and updated on a regular basis. Summary traffic data from the CTDB are periodically sent to NIMS for inclusion with the pavement performance data. These updates include corrections of previous erroneous data. Procedures and standards were established to ensure data quality, and extensive data quality checks are preformed throughout the collection and recording process. Information is also available indicating the level of data reliability. Although data are collected at the regional level and stored in Regional Information Management Systems (RIMS) and regional CTDBs, data are only released to the public after they have passed these checks and are stored in the national databases.

A guide that contains more details on the background and objectives of LTPP — what data are collected, how to request data, data formats, and examples of reports generated — can be found in reference 2. Complete information on how the data are collected, what quality checks are imposed, etc., can be found in other documents.

Data are released on two levels: (1) a sectional release and (2) an experimental analysis release. Data in Level 1 generally should be considered for analysis of a given test section, not comparisons across sections. These data have passed a minimum number of quality checks and, if used in analyses, should be used cautiously. Level 2 data have completed all assurance checks and are considered acceptable for analysis. Many quality control issues are still under development and consideration in an ongoing FHWA contract. Among these is the prospect of grouping sites into classifications according to the completeness of the traffic data at those sites. A classification being considered for the amount of data available is "preferred," meaning that at least 9 months of continuous data are available; "desirable" would mean that at least 6 months of continuous data are available; and "minimum" would mean that anywhere from 1 day to 6 months of data are available. Missing data can be due to lack of continuous WIM devices, equipment failure, etc. These classifications have not been set and could have changed by the time of this report. The researcher is referred to the periodic progress reports produced from this contract. The FHWA contact for this information is Kris Gupta. At this time, there is a limited amount of data available to the public, i.e., data that have passed Quality Assurance/Quality Control (QA/QC) checks. Although the plan is to have at least 50 percent of the data available by the end of 1995, the FHWA contact can best update the researcher on this.

Potential uses of the LTPP traffic data would have to focus on safety studies that are location based. For example, the question of "are double-tractor configurations overly represented in on-/off-ramp accidents as compared to singles?" might be addressed using the LTPP traffic data. First, it would be necessary to ascertain whether or not there are a sufficient number of LTPP sites with complete enough traffic data to supply enough accidents to do an adequate evaluation. Secondly, are accident histories available at these sites and over a sufficient time period? This would be the general process for examining the feasibility of using the LTPP traffic data (or any location-specific traffic database):

1. Formulate the hypothesis.

2. Determine what traffic data best represent the exposure for the data required to address the hypothesis.

3. Determine if there are sufficient sites of the type required by the hypothesis in the CTDB. How complete are the traffic data at these sites?

4. Determine whether accident histories are available and in sufficient numbers to justify the analysis.

These steps should be attainable using only a minimum amount of resources.

The only way to receive LTPP data from the national databases is to submit a complete LTPP Data Request Form to the TRB NIMS Administrator:

Penny Passikoff
National Academy of Sciences
Transportation Research Board
2101 Constitution Avenue, NW
Washington, D.C. 20418
TEL: (202) 334-3259
FAX: (202) 334-3495

Costs for obtaining the data include a $75 handling fee, media costs that depend on the type of media selected on the form, shipping costs, and any costs due to custom requests. State and Federal agencies and international participants do not have to pay the $75 handling fee.

References

(1) Rowshan, Shahed. ***Long-Term Pavement Performance Information Management System Data Users Guide.*** Federal Highway Administration, Report No. FHWA-RD-93-094, July 1993.

(2) Herman, John L.; Charlie R. Copeland; and W.O. Hadley. ***SHRP-LTPP Traffic Data Collection and Analysis: Five-Year Report.*** Texas Research and Development Foundation, Austin, TX. SHRP-P-386, 1994.

Nationwide Personal Transportation Survey (NPTS), FHWA

Purpose: The Nationwide Personal Transportation Survey (NPTS) provides nationally representative estimates of personal travel in the United States. All modes of transport are covered, including passenger cars, trucks, motorcycles, buses, trains, subways, airplanes, taxis, bicycles, and walking. The dataset includes information on demographic characteristics of the household, person-level information on the individuals participating in the survey, descriptive information on each vehicle in the household, and two levels of travel information. The first level of travel information is a detailed account of all trips taken on the survey day. The second level is information on trips longer than 12 1 km that occurred during the 2-week period immediately prior to the survey day. Travel information includes mode, vehicle type, road type, date of travel, time of day, trip purpose, origin and destination, elapsed time, and area type.

Source: The most recent NPTS (1990) was conducted by the Research Triangle Institute of Research Triangle Park, NC, under the sponsorship of the U.S. Department of Transportation?) A random sample of 26,172 households with telephones was selected by means of a random-digit dialing procedure, and almost 22,000 households responded. Responses were collected by means of a telephone interview. (Earlier surveys were done using in-home interviews.) Each household was assigned a 24-h travel day (defined as 4:00 a.m. on the travel day to 3:59 a.m. on the following day) and a 14-day travel period. The survey period was from March 1990 to March 199 1. Person-level interviews were conducted with all household members age 5 years and older. Trip-level interviews were conducted with all household members age 13 and older. The latter respondents supplied travel information on residents 5 to 13 years of age.

Coverage: The current file (1990) is the fourth in the series; earlier NPTS files are for 1969, 1977, and 1983. All personal trips, all modes of transportation, all purposes, and all 50 States and the District of Columbia are covered. Connecticut, the New York Metropolitan Planning Organization (MPO), and the Indianapolis MPO funded oversampling in their respective areas. The file includes weight variables, so that estimates of national totals can be computed.

Strengths: The NPTS file is the only source for national data on personal travel. Sample sizes are large, with 22,3 17 households, 48,385 persons, 35,152 licensed drivers, and 41,178 vehicles in the most recent sample. The survey design includes both driver and passenger travel, so vehicle occupancy rates can be analyzed. NPTS files are now available for 1969, 1977, 1983, and 1990, allowing trends over a period of 2 1 years to be analyzed. Efforts were made to maintain comparability of the major elements of the survey over that period. Travel can be broken down by region and for households in certain metropolitan statistical areas. Detailed information is available on the socioeconomic status of the household; age, gender, and other characteristics of the travelers; purpose of trip; type, make, and model of vehicle; and time, distance, and duration of travel. Interviews are conducted using computer-assisted telephone interviewing techniques, so many inconsistencies could be identified during the interview and addressed by the respondent.

Limitations: Road type is available only for a small subset of day trips. Sample sizes for commercial vehicles are small-the focus of the survey was on personal travel-so the NPTS is not useful for truck travel. The focus of NPTS is on national travel. It is possible to estimate the travel for regions of the country and for certain States and Metropolitan Sampling Areas (MSAs), but estimates for individual local areas, MSAs, or States may not be based on large enough sample sizes and may be imprecise. Households without telephones could not be included in the sample because the sampling procedure was based on a random-digit dialing procedure. In addition, the data are all self-reported.

Sampling Errors: Sampling errors can be calculated using appropriate software. See the User's Guide.

Access: The data are contained in six hierarchical files and can be obtained either as an EBCDIC file (similar to plain ASCII) or formatted for the SAS statistical analysis package. The files can be obtained on magnetic tape through the Volpe National Transportation Systems Center, Cambridge, MA, (617) 494-2450.

References

(1) ***User's Guide for the Public Use Tapes: 1990 Nationwide Personal Transportation Survey,*** December 1991, Report No. FHWA-PL-92-007.

National Truck Trip Information Survey (NTTIS), UMTRI

Purpose: The National Truck Trip Information Survey (NTTIS) provides national estimates of truck travel that can be cross-classified by truck configuration and loading, road type, area type, and time of day. Details on truck configuration and loading include cabstyle, number of trailers (if any), number of axles for each unit, empty weight and length for each unit, cargo body style, cargo type for each unit, and cargo weight for each unit. Road type is divided into three categories: limited access, U.S. and State numbered routes, and other roads. Area is classified using Federal Highway Administration definitions of urban or rural. The time of operation is classified as either day or night.

Source: The NTTIS was conducted by the Center for National Truck Statistics, part of the University of Michigan Transportation Research Institute (UMTRI).(1) The work was supported primarily by the Motor Vehicle Manufacturers Association, the Western Highway Institute, the Engine Manufacturers Association, and the American Trucking Associations. An initial sample of 8,144 trucks was drawn from registration files maintained by the R.L. Polk Company. The sampling frame was stratified by State and within each State, and by whether the truck appeared to be a tractor, straight truck, or unknown type. An interval selection procedure with a random start was used to draw the sample. Interviewers contacted current owners and operators of the vehicles by telephone to obtain a general description of the vehicle and company that operated it. Questions included estimates of annual travel that were checked against estimates from the TIUS.

A subsample of approximately 5,000 trucks was drawn for the travel survey. On four randomly selected days over a year, each truck was surveyed as to its use for the previous 24-h period. The survey method was to essentially follow the truck for 24 h. Survey staff collected information on the actual route the vehicle followed, cargo carried (if any) and where it was loaded or unloaded, and a complete description of the truck's configuration. The route was then followed on a map and the mileage was classified by road type, time of day, and urban/rural. All data were subject to extensive editing to ensure accuracy. To the extent possible and where necessary, inconsistencies and inaccuracies were cleared up by more phone calls to survey respondents.

Coverage: The NTTIS was a one-time survey. The sampling frame was trucks registered in the United States in 1983. The phone survey to collect the initial vehicle description and then the follow-up calls for trip information took place between November 1985 and February 1987. The file covers all medium and heavy trucks (GVWR > 4536 kg) registered in the United States, except for trucks owned by any level of government.

Strengths: Travel estimates can be cross-classified by truck configuration, loading, and operating environment — a level of detail unmatched in any other file of travel data.(2) It is possible, for example, to compare the travel of loaded and unloaded two-axle tank trailers on limited-access roads in urban areas at night. All data were carefully reviewed by editors experienced with the trucking industry. Ambiguous or unusual responses were clarified, where possible, with respondents. It is expected that the data are as accurate-as is feasible.

Limitations: Data are all self-reported, although subject to careful evaluation and consistency checking. Given the frequent contact between interview staff and respondents, and the ability to check responses, it is felt that the data are not systematically biased. Estimates from the file are all national. It is not possible to retrieve travel information for particular routes or even particular States. Moreover, by 1995, the file is clearly dated. There have been several important changes in the trucking industry since 1987 — for example, an increasing reliance on multiple-trailer trucks — that the file cannot reflect.

Sampling Errors: All sampling strata variables are included in the analysis file. Sampling errors can be calculated with appropriate software.

Access: The NTTIS file is a hierarchical dataset consisting of three parts: (1) a truck file with data describing the power unit, (2) a tractor trip file with data on trips by tractors, and (3) a straight truck file with comparable information about straight truck trips. The trip files contain one record for each trip taken by a survey vehicle on a survey day. Access to the data is provided through the Center for National Truck Statistics at UMTRI. Contact Kenneth L. Campbell or Daniel Blower at (313) 764-0248.

References

(1) Blower, Daniel and Leslie C. Pettis. ***National Truck Trip Information Survey.*** University of Michigan Transportation Research Institute, Ann Arbor, MI, Report No. UMTRI-88-11, March 1988.

(2) Massie, Dawn L.; Kenneth L. Campbell; and Daniel F. Blower. "Large-Truck Travel Estimates From the National Truck Trip Information Survey." ***Transportation Research Record No. 1407, Large-Vehicle Safety Research.*** Transportation Research Board, Washington, D.C., 1993, pp. 42-49.

Operational Exposure Data Sources

Historical Summary and Purpose: Researchers in the field of highway operations are often in need of exposure data in the form of both quantity of traffic and traffic congestion. Several researchers at Texas Transportation Institute were queried as to their knowledge of these data sources and the following reports resulted:

Kevin Balke's understanding is that the State of Texas (and probably others) has an extensive traffic monitoring program. His personal experience included collecting ADT volumes on many arterials and highways in major cities every 4 years. These studies were managed by local MPOs and these counts were published in a report. The Texas Department of Transportation maintains permanent count stations. A map is published annually with the AADT volumes displayed by location. However, none of this has been automated — this seems to be the major drawback in most operations study data sources. And, of course, there is the State roadway inventory file to which operations researchers often turn. Gerald Ullman relies on these State roadway inventory databases, as well as the State's ATR stations. With regard to urban area operations, some cities have systematic count programs and some do not, according to Ray Krammes. Dallas, for example, has a machine count program. Specific personnel in each city would serve as the contact for obtaining this information (in Dallas, it would be Ken Melston). State highway departments would probably be the best source for this information. In Dallas, the initial goal was to have manual counts on every 1.6-km segment of arterial road every 3 years. However, lack of funding seriously reduced this effort. Dallas still collects much of the data and stores 24-h and peak counts in a computer program and publishes two reports every January — one that lists the most recent count on each link and one that lists historical data, i.e., all counts on all links. Fifteen-minute counts could also be attained on paper copy. The only other city in the North Texas region that has some count data is Fort Worth. Most cities in the Metroplex do counts only on an ad hoc basis and generally hire consultants to do this work. In a review of Texas cities, this was generally the case (Austin, Houston, etc.). The counts are done on an ad hoc, nonsystematic basis for specific purposes.

It may be possible to design a highway safety research project using some of these site-specific count data. For example, Dallas would appear to have sufficient count data to address a particular urban problem. Consider the comparison of accident severities as a function of congestion — peak vs. off-peak times, weekend vs. weekdays, etc., or issues such as alcohol-related crashes in urban areas by time of day. However, due to the erratic nature of the data collection, one must be concerned about what biases such non-systematic data collection might be introducing into the safety analysis. Also, the fact that most data sources appear to be unautomated, at least in Texas, is a serious drawback.

For the most part, it appeared that operations researchers are interested primarily in very site-specific data and rely on ad hoc, often manual, procedures for obtaining exposure information. However, when they are interested in more global issues, they rely heavily on the Highway Performance Monitoring System (HPMS), described separately.

Residential Transportation Energy Consumption Survey (RTECS)

Historical Summary and Purpose: The Residential Transportation Energy Consumption Survey (RTECS) is a survey designed and administered by the Energy Information Administration (EIA). The objective of the survey is to obtain information on vehicles used for personal transportation in the United States. It is a companion survey to the Residential Energy Consumption Survey (RECS).

The first RTECS was done in 1978 and has been repeated triennially since 1985. The most recent survey for which published data are available is 1991. The following discussion relates to the 1991 survey. A survey was done in 1994, but the data are not available as of the date of this publication. The survey has been done five times. The RTECS is a follow-up survey and companion to the RECS. The RECS collects data on the households and includes preliminary information on the vehicles available to the household, while the RTECS consists of three stages in which additional data are collected on the vehicles available and the use of the vehicles by members of the household.

The data collected in the RTECS and RECS may have applicability in different areas of highway safety research. Primary data elements of interest in highway safety are the estimates of vehicle-miles of travel and the motor vehicle stock available to households for personal travel. These data elements may be linked to characteristics of households to allow computations concerning the amount of exposure (both vehicle-miles of travel and vehicle type) for similar households. Since the primary driver of each vehicle in sampled households was identified, as well as the age of the driver, the vehicle-miles of travel and vehicle used by age of primary driver may be estimated by implication. Since the data were not collected for trips by individuals within the household, the use of these estimates of exposure for different age groups may be questionable. It does appear the data are disaggregate enough for computing vehicle-miles of travel for households stratified by different household characteristics. This would provide a means for the estimation of exposure for those households and the applicability of those estimates to specific regions where similar stratifications of households could be obtained.

Data Contents and Structure: Household data collected in the RECS through personal interview that may be of interest in highway safety research include the following:

- Census region and division where household was located.
- Urban status of the household location (whether urban or rural area).
- Number of persons in the household.
- Data on the household composition (e.g., number with/without children, age of householder, etc.).
- Race of householder.
- 1990 family income (these were reported in nine different ranges).
- Number of drivers in household.

- Age and sex of primary driver for each vehicle in household.
- Average number of vehicles available to household during the year.
- Model year and vehicle type for vehicles available.
- Whether vehicle was used for commuting to and from work.

For the household data collected, data on the number of vehicles available and the vehicle-miles of travel for those vehicles were obtained. Vehicular data were not collected in the RTECS for motorcycles, bicycles, all-terrain vehicles (ATVs), and other related vehicles.

Experimental Design, Sample Plan, and Location Distribution: The focus of the RTECS is to obtain data on the vehicle-miles of travel, motor vehicle stock, and vehicle fuel consumption and expenditure data. Its companion survey (RECS) collects data on household energy consumption and expenditure. The sampling units in both the RECS and RTECS are households, with the universe being all housing units occupied as the primary residence in the 50 States and the District of Columbia. The sample of households selected in the 1991 RTECS was based on the 1990 RECS. The 1990 RECS was a multistage probability sample that incorporated a rotating panel to allow the observation of changes in energy use over time for households that fall in successive panels.

The 1990 RECS initial sample consisted of 6,757 units. Of these units, 848 were found to be ineligible for reasons such as the dwelling being uninhabitable, currently vacant, or used for seasonal occupancy. Energy-related data were collected from 4,828 households via telephone interviews, and an additional 267 units were surveyed through a mail follow-up, for a total of 5,095 responding households. The RTECS sample of households was selected from the 5,095 housing units that responded to the 1990 RECS survey. The number of RECS housing units selected for the RTECS survey was 3,045. Of those units, 2,842 were contacted by telephone and 200 were identified as households that had to be contacted by mail. The number for contact by mail was subsequently increased to 485 due to an increased number of households with unlisted or disconnected telephones.

The RTECS data collection effort consists of four phases, with the first phase being done in conjunction with the RECS. The first phase (during the RECS personal interviews) collected data on the household's vehicle stock, the vehicle identification numbers (VIN) of the vehicles, and initial odometer reading for each vehicle. The subsequent three phases were conducted at the beginning of the year (B-O-Y), mid-year (M-Y), and the end of the year (E-O-Y). These data collection efforts were done by telephone interview and, where this was not possible, the data were collected via a mail questionnaire. The B-O-Y and E-O-Y phases updated the data on the vehicle stock and collected data on the vehicle characteristics (including the vehicle make, model and model year, the vehicle odometer readings, and VIN). The M-Y phase was an inventory update where respondents were asked to complete a vehicle update worksheet and keep it for use during the telephone interview or mail it back if the household was classified as a no-telephone household.

The data collected during the RTECS allow for the computation of actual vehicle-miles of travel from the recorded odometer readings. These data represent total travel between the two points in time (i.e., B-O-Y and E-O-Y). Data were also collected on the disposition of vehicles and acquisition of new vehicles during the survey period.

Quality of Data: The data collected in the RECS and RTECS appear to be of relative high quality. Since the surveys produce estimates based on randomly chosen subsets of the entire population of occupied housing units, the estimates will always differ from the true population values and will include sources of nonsampling and sampling errors. The following sections discuss various sources of potential error in estimates produced from these surveys:

Noncovered Residential Vehicles. Since the sample of households surveyed in the RTECS were selected from the RECS, any household excluded from the RECS would not be represented in the RTECS, and the subsequent survey data would not include vehicles available to those households. Specifically, those families or individuals not included in the RECS were those living in group quarters such as college dormitories, military barracks, or large boarding houses; those living in recreational or other types of vehicles; and those with no fixed address. The effect of these exclusions is an underestimation of the total number of vehicles and related data.

Date of Reference for Survey. Since the survey design requires households to be followed for an entire year, changes in household structure and composition may not be accurately reflected. For example, the survey sample may have an overrepresentation of older established households and an underrepresentation of new households or families. Resulting estimates of vehicles and related data may have a negative bias induced by established households separating and only one portion being followed in the RTECS, vehicles acquired by household members that leave the household are not captured in the survey, and the total estimated households (used for expansion) is based on the July 1991 Current Population Survey (Bureau of the Census).

Item Nonresponse. Item nonresponse refers to the inability to collect full information when respondents either do not know the answer or refuse to answer selected questions. It can also occur when an interviewer fails to ask a question or record an answer. In the RTECS, item nonresponses were imputed to provide an estimate of the most probable response. Three techniques were used: hot-decking, predictive mean matching, and regression.

Hot-decking is a technique by which a household is randomly selected and its response to the missing data item is used as the response for the household with the missing item. The items imputed in the RTECS by this method were pre-1975 vehicle characteristics and fuel grade. Household demographic items, such as family income and ethnic background, were also imputed by this method for the RECS.

Predictive mean matching was used for imputing changes in vehicle stock for households not followed for the complete duration of the RTECS. In the 1991 RTECS, 26 percent (i.e., 795 households) were not followed for the entire year and imputations were computed to estimate the number that acquired and/or disposed of vehicles during the year. For households with no vehicles that were lost, a hot-deck procedure was used to impute the changes in vehicle stock.

Multiple regressions were used to impute annual vehicle-miles of travel for those vehicles that were imputed as being acquired. Linear and multiple regressions were also used for estimated annual mileage for vehicles where two odometer readings were not obtained in the survey. For 26 percent (i.e., 1,576) of the sample vehicles, no odometer span was available. An estimate of the annual vehicle-miles of travel had been obtained from the respondent during the RECS interview. Vehicle-miles of travel were imputed from a regression on the estimate of vehicle-miles of travel obtained in the RECS interview. For an additional 19 percent (i.e., 1,150) of the sample vehicles, no odometer span was available and an estimate of annual vehicle-miles of travel was not obtained in the RECS interview. Estimates of vehicle-miles of travel for these sample vehicles were imputed using a multiple regression using number of drivers, household income, age of household head, type of vehicle, and use of vehicle on the job as independent variables. This same method was used for imputing the vehicle-miles of travel for vehicles that were imputed as being acquired and/or disposed. Various other adjustments to the vehicle-miles of travel data were necessary to put each in terms of the same time period. Data from the Federal Highway Administration on monthly vehicle-miles of travel were used for this purpose.

Potential Problems: The RTECS data provide reasonable estimates of vehicle-miles of travel for households and vehicle types. These data will produce reasonable estimates of exposure relative to household estimates and estimates by vehicle type. However, the data do not include travel by motorcycles, bicycles, all-terrain vehicles, or similar types of vehicles, which may be critical in safety analyses. In addition, the data do not relate vehicle-miles of travel to person-miles of travel. The data are collected for vehicles and related to the households that own or have those vehicles available. While the exposure may be computed for vehicles in terms of type and vehicle-miles of travel, the data do not indicate the number of persons that may be in the vehicle on an average basis. Other data sources on average vehicle occupancy would have to be used to impute that estimate. The use of the data to compute exposure estimates by age of individuals would have to be based on the implication of primary driver for each vehicle in the survey. This is a relatively weak implication and is not considered an accurate estimate. Thus, it is not considered appropriate to use data from this source for estimating exposure for persons by age.

Data Acquisition and Documentation: Data from the RTECS and RECS are available in a variety of media. The following published reports may be purchased from the Government Printing Office (GPO):

- ***Household Vehicles Energy Consumption** 1991*; December 1993, DOE/EIA-0464(91) ☎☠□ GPO Stock No.).

- ***Household Vehicles Energy Consumption** 1988*; February 1990, DOE/EIA-0464(88), GPO Stock No. 061-003-00652-3.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns** of **Household Vehicles,** 1985*; April 1987, DOE/EIA-0464(85), GPO Stock No. 061-003-00521-7.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, 1983***; January 1985, DOE/EIA-0464(83), GPO Stock No. 061-003-00420-2.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, Supplement: January 1981 to September 1981***; February 1973, DOE/EIA-0328, GPO Stock No. 061-003-00297-8.

- ***Residential Transportation Energy Consumption Survey: Consumption Patterns of Household Vehicles, June 1979 to December 1980***; April 1982, DOE/EIA-0319 (No GPO Stock No.).

The above documents are not the only ones available, but are considered to represent those report data that are of interest to highway safety engineers. In addition to the published reports, data tapes and diskettes may be ordered directly from the National Technical Information Service (NTIS). Information on how to order these may be obtained by telephoning NTIS at (703) 487-4807, FAX number (703) 321-8547. Detailed technical questions on topics of interest to highway safety engineers may be addressed to the following:

RTECS Manager	Ronald Lambrecht	(202) 586-4962
Vehicle-Miles of Travel	John Pearson	(202) 586-6160
Trends in Household Vehicle Stock	Ronald Lambrecht	(202) 586-4962

References

(1) ***Household Vehicles Energy Consumption 1991***; December 1993, DOE/EIA-0464(91) (No GPO Stock No.).

Truck Inventory and Use Survey (TIUS), Bureau of the Census

Purpose: The Truck Inventory and Use Survey (TIUS) is one of a number of economic censuses performed by the U.S. Bureau of the Census. It is designed to provide information on the population and use of trucks for government, business, industry, and the general public. The TIUS is conducted every 5 years. The most recent data year currently available is 1992.

The TIUS provides annualized estimates of the primary uses of trucks. Data include a physical description of the truck (axle count, cabstyle, cargo body style, overall length, empty weight, typical loaded weight, maximum loaded weight); a general description of the industry in which the vehicle is used; and a breakdown of the vehicle's use over the course of a year. For example, respondents report any placarded hazardous materials carried in the vehicle and then estimate the percentage of the total annual travel in which hazardous materials were carried. Similarly, respondents estimate the proportion of annual travel accumulated off-road, less than 80.5 km from the truck's home base, 80.5 to 321.9 km from base, and more than 321.9 km from base.

The TIUS is useful for estimating broad categories of annual truck use. Given the way the data are reported, however, it is not possible to break down or cross-classify travel estimates by road type, area type, or any other feature of the operating environment. It is also not possible to estimate travel by State, month, or season.

Source: The TIUS is a stratified probability sample of trucks registered in the 50 States and the District of Columbia. Within each State, trucks are stratified by body style. Within each stratum, a fixed number of trucks are sampled randomly. Roughly 3,000 trucks are sampled per State. Survey forms are then mailed to the registered owners of the sampled trucks. By law, the surveys must be completed and returned. The data are all self-reported and are all estimates of use for a particular year. Reports are subject to computer editing. Apparently erroneous responses are reviewed and corrected, if possible.

Coverage: The sampling frame for the TIUS covers all vehicles registered as trucks in the 50 States and the District of Columbia. This includes pickups, small vans, and other utility vehicles registered as trucks. The file excludes vehicles owned by any unit of government, passenger -- vehicles, ambulances, buses, and motor homes. Vehicles used exclusively off-road do not have to be registered, and thus are also excluded.

Strengths: The TIUS has a very large sample size. Roughly 154,000 vehicles were selected for the survey in 1992. Nearly 132,000 trucks are represented in the file. Estimates of population totals and annual travel from the TIUS have been compared with estimates generated by other techniques (e.g., NTTIS; for a description of NTTIS, see the discussion in an earlier section) and are in general agreement. Data collection procedures and survey questions have been fairly stable for a number of surveys, so comparisons among survey years are valid.

Limitations: The main limitation in the use of the TIUS file for safety-related exposure data is that the data represent typical or primary use only. Consequently, configurations that represent secondary use, such as bobtails or doubles, are not represented at all or are under-estimated. There is very little ability to cross-classify the travel estimates by operational characteristics that are known to be associated with differences in accident-involvement risk. For example, straight trucks do a large share of their travel in urban areas and on non-limited-access roads. Tractor-semitrailer combinations accumulate a much larger fraction of their travel on limited-access roads, which are typically the safest in the highway system. The TIUS data do not provide any means of controlling for such environmental confounding factors.

Sampling Errors: Variables representing the sampling strata are not released with the file, so it is not possible to calculate sampling errors for particular estimates. However, the published ***Census of Transportation*** includes an appendix with equations for approximating relative standard errors.

Access: Available on CD-ROM from the Bureau of Transportation Statistics and from Customer Services, Bureau of the Census, Washington, D.C. 20233. The data are the raw records from the survey, modified to limit the possibility of identifying particular individuals or businesses.

State Weigh-in-Motion (WIM) and Automatic Vehicle Counting (AVC) Devices

Historical Summary and Purpose: Truck weighing equipment is required for meeting a wide variety of public, private, and institutional needs. In the public sector, there are two major functional areas of application of these devices: data collection and enforcement. Statistically representative truck weight data are collected and used as the primary basis for engineering analyses and decisions related to planning, funding, design, operation, maintenance, and management of highway facilities. Measurements of the weights of individual trucks are needed to provide enforcement agencies with the capability to protect the highway infrastructure from damage due to unexpectedly high loads. In both data collection and enforcement, it is necessary to weigh large numbers of individual trucks.

A weigh-in-motion (WIM) system is used to attempt to approximate the gross weight of a vehicle or the portion of the vehicle weight carried by a wheel, an axle, or a group of axles by measuring, during a short time interval, the vertical component of dynamic (continually changing) force that is applied to a smooth, level road surface by the tires of the moving vehicle. Although the weight of a vehicle does not change as it moves over the surface of the road, the dynamic force applied to the roadway surface by a rolling tire on a vehicle varies dramatically when the tire/wheel mass accelerates vertically. This acceleration can be induced by roughness in the road surface and/or by an out-of-round or out-of-balance wheel/tire assembly.

Data Contents and Structure: WIM data are collected in the United States by the States under three programs. One is specified and required by the FHWA under the provisions of its ***Traffic Monitoring Guide*** (TMG). The States have designated and collected data at approximately 1,400 WIM sites in the United States. The data are stored as individual truck records by the individual States and are transmitted to FHWA.

Additional WIM data are obtained under the Long-Term Pavement Performance monitoring aspect of the Strategic Highway Research Program. Data are acquired quarterly for 7 continuous days at 777 sites throughout the United States and are transmitted to regional SHRP contractors.

The last type of WIM data is collected at truck weight enforcement stations during the weighing and sorting of trucks to determine whether they exceed legal limits. These data are not normally retained.

Each State is required to submit vehicle classification and truck weight data to the FHWA either annually or quarterly. Where continuous weigh-in-motion data are available, 1 week of data per quarter is required. These data provide input to national databases that are maintained by the FHWA. These databases include the Traffic Volume Trends System and the Vehicle Travel Information System. The Traffic Volume Trends System is a database management system that is based on state-supplied ATR data. The Vehicle Travel Information System is a microcomputer database management system that validates, summarizes, and maintains vehicle classification and truck weight study data. Tables 1 through 3 contain State-by-State information on the number of WIM sites, type of equipment, level of monitoring, the existence of historical data, and monitoring frequency. Level of monitoring refers to the amount of data collected. The preferred, minimum, etc. categories are the ones described in the LTPP traffic data, although these may not be the levels adopted by the CTDB.

Table 1. Region 1 WIM.

STATE	NO. SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA
Illinois	18	GK Instruments 6000 AWACS	Preferred	Y
Indiana	18	IRD Bending Plate	Preferred	Y
Iowa	12	GK Instruments 6701	Preferred	Y
Kansas	17	GK Instruments 6701	Preferred 1, Desirable 16	Y
Kentucky	7	Unknown (Portable)	Preferred 1, Minimum 6	Y
Michigan	13	GK Instruments 6012 (Piezo)	Preferred	Y
Minnesota	24	IRD Bending Plate	Preferred 2 1, Unknown 3	Y
Missouri	20	IRD 100 and GK 6701	Minimum	Y
Nebraska	15	Golden River Portable	Minimum	Y
North Dakota	4	GK Instruments 670 1	Preferred	Y
Ohio	11	Pat Equipment	Preferred	Y
South Dakota	9	In-House Bridge WIM	Preferred	Y
Wisconsin	16	Pat Equipment	Preferred 5, Minimum 11	Y

Table 2. Region 2 WIM.

STATE	NO. SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA	MONITORING FREQ
Alabama	18	Bending Plate and Piezo Cable	Preferred 1, Desirable 17	Y	1 continuous, rest 7 days per season
Arkansas	14	Cap Pads and Piezo	Preferred 1, Desirable 13	Y	1 continuous, rest 7 days per season
Florida	29	Portable	Desirable	Y	7 days per season
Georgia	23	Cap Pads and Bridge	Preferred 2, Desirable 20, Minimum 1	Y	2 continuous, rest 7 days per season
Louisiana	2	Cap Pads	Desirable	Y	7 days per season
Mississippi	25	Piezo	Preferred	Y	Continuous
New Mexico	12	Cap Pads	Desirable	Y	7 days per season
Oklahoma	21	IRD Piezo	Preferred	Y	Continuous
Puerto Rico	4	Cap Pads	Desirable	Y	7 days per season
South Carolina	9	Portable	Minimum 8, Below Minimum 1	Y	Seasonal
Tennessee	15	Piezo	Preferred 2, Desirable 13	Y	2 Continuous, rest 7 days per season
Texas	90	Cap Pads	Below Minimum	Y	27 days annually

Table 3. Region 3 WIM.

STATE	NO.SITES	TYPE OF EQUIPMENT	LEVEL OF MONITORING	HIST. DATA	MONITORING FREQ.
Alaska	6	IRD	Preferred 5, Continuous 1	Y	Preferred
Arizona	25	Portable	Minimum	Y	Seasonal 7 day
California	37	Pat	Preferred 3, Continuous 15, Minimum 11, Below Minimum 8	Y	Continuous or seasonal
Colorado	16	IRD	Preferred	Y	Preferred
Hawaii	4	IRD	. Minimum	Y	Seasonal 7 day
Idaho	13	Portable	Preferred 1, Continuous 12	Y	Seasonal 7 day
Montana	7	Portable	Below Minimum	Y	Seasonal 7 day
Nevada	8	Portable	Preferred 1, Minimum 7	Y	Seasonal 7 day
Oregon	11	Pat	Minimum	Y	Seasonal 7 day
Utah	14	Portable	Minimum 2, Below Minimum 12	Y	Seasonal 7 day
Washington	19	IRD	Preferred	Y	Preferred
Wyoming	14	Pat	Minimum	Y	Seasonal 7 day

Experimental Design, Sample Plan, and Location Distribution: Each State determined their own experimental design and determined the number and location of the sites based on differing economic and policy-making factors. When using WIM data from any State for highway safety evaluation purposes, the researcher should contact the respective State's DOT and request specific information regarding site-selection criteria.

Potential uses of the WIM databases must be location-oriented, similar to the ones described for the LTPP WIM.

Data Acquisition and Documentation: Data from the national database must be requested from the FHWA directly. These data include: station description data, traffic volume data, vehicle classification data, and truck weight data. Each type of data has its own individualized record format. All data files are in ASCII flat files.

Individual State data can be requested of the individual State DOTs. The formats will vary. For example, Illinois currently has 18 active WIM sites dispersed throughout the State. The WIM system has not consistently provided the necessary data to the national database due to hardware and/or software problems. Illinois DOT collects data biweekly and stores all data that are required by the FHWA. The data are processed and kept on the mainframe computer in a hexadecimal format. Their data on the continuous count ATR network are located at 21 sites. These data provide vehicle count and classification data and are kept on personal computers in ASCII format.

Washington State DOT has 41 active WIM sites — 5 use bending plates and the rest use piezoelectric sensors. The sites are continuous monitoring sites and the data are downloaded weekly. The data provide the standard vehicle classification and truck weight data required by the FHWA. The data are converted by the State from 13-bin to 4-bin format for storage on a mainframe computer. Data from 1990 to the present are available.

Reference

(1) Parsons, Brinckerhoff, Quade & Douglas, Inc. And URS Consultants, ***Inc. Pavement Damage Factors Derived From Weigh-In-Motion Data Measured by Portable vs. Permanent Systems.*** Florida Department of Transportation Statistics Office, Traffic and Roadway Data General Consultant Task Work Order Number 4, Sub-Task 3.2, December 1993.

3. EMERGING EXPOSURE DATA SOURCES

Emerging exposure data sources are new sources or existing sources that have not been traditionally used to derive exposure estimates. Three areas were reviewed for possible emerging exposure data: Intelligent Transportation Systems, transportation planning surveys, and traffic volume data collected by the States. The scope of each area is described briefly in the following paragraphs.

Intelligent Transportation Systems (ITS)

Within the broad Intelligent Transportation Systems (ITS) area, three subareas were examined: Commercial Vehicle Operations (CVO), Advanced Traveler Information Systems (ATIS), and Advanced Traffic Management Systems (ATMS). Specific projects in the CVO area are the Crescent project in the western States and Advantage I-75 in the east. Each includes some automatic provisions for trucks to communicate various required information about the vehicle and driver, such as license status, vehicle permits, and inspection data. These are all multistate projects intended to minimize the stops a truck needs to make to demonstrate compliance with all the applicable regulations. Since the information is recorded electronically, there may be some way to get descriptive information and counts that could be used as exposure measures. Similar potential to gather exposure data may be present in the other two ITS areas reviewed.

Transportation Planning Surveys

The second area covers a range of transportation planning surveys. These are usually household surveys conducted by mail or telephone. Examples are the Transportation Planning Package of the U.S. Census (CTPP). This survey provides nationwide data that form the basis for many State and local transportation planning efforts. However, only trips to and from work are included. The other general source in this area is regional planning surveys. These are also household surveys patterned after the CTPP. The geographic coverage is limited, of course, but more detailed information is frequently collected, often for a broader range of trip purposes than just travel to and from work.

Traffic Volume Data - Errors of VMT Estimates Based on Traffic Counts and Section Length

The third area reviews the traffic volume data that are available from many States, and that form the basis of the traffic volume data in HSIS. Most traffic volume data are collected by State and local highway departments. Consequently, we need a good understanding of the accuracy and timeliness of the available data. How often are the counts actually taken at the site and, if taken some distance away, how accurate will they be for the site in question?

The remaining material is organized under these three headings.

Intelligent Transportation Systems (ITS)

The development of Intelligent Transportation Systems (ITS) technologies and services offers new opportunities to obtain exposure information. Since the primary objectives of ITS are not related to exposure data collection, it is important to recognize such opportunities and identify processes by which exposure data could be obtained. This section explores possible exposure data sources within the commercial vehicle operations portion of ITS.

Commercial Vehicle Operations (CVO) of ITS

Commercial Vehicle Operations (CVO) has been divided into six user services:

- Commercial vehicle electronic clearance.
- Automated roadside inspections.
- Commercial vehicle administrative services.
- On-board safety monitoring.
- Hazardous material incident response.
- Commercial fleet management.

Of these services, commercial fleet management, commercial vehicle administrative services, and commercial vehicle electronic clearance have potential as sources of data on commercial vehicle exposure in terms of vehicle-miles traveled over specific types of roads by various categories of commercial vehicles. There is also a possibility of applying some of the technology being developed for ITS research purposes to collect specialized exposure data.

Vehicle tracking systems for commercial fleet management that keep dispatchers appraised of the current locations of all their fleet vehicles could provide a source of exposure data. Such a system would need to include an automatic vehicle location (AVL) system, probably a global positioning system (GPS) and map matching software that would locate the vehicle on a map. If the system could preserve the history of travel of an individual vehicle over the course of the trip, the equivalent of a trip diary could be generated for every vehicle in a fleet with such a system The record of the configuration and cargo of the commercial vehicle for the trip could also be included in the trip record. The data file from the individual records could yield the miles traveled by each vehicle by road class and by vehicle configuration for the fleet.

A problem with commercial fleet management systems as sources for exposure data is that the data would be collected by the motor carriers. They might prefer to treat this information as proprietary and would not be willing to share this information with others. Even if some fleets decide to share this information with researchers, there may still be a problem with obtaining cooperation from enough fleets of appropriate sizes and diversity for a desired sample.

Another application of CVO systems that might overcome the problem with proprietary information is the commercial vehicle administrative process. States need to know the mileage of commercial vehicles on their roads for the purpose of fuel tax allocation. A specific system currently being tested in Iowa for this purpose is the on-board automated mileage system. The system uses GPS vehicle location technology and map-matching algorithms and software to determine the mileage a given commercial vehicle equipped with the system has traveled within a State. The map-matching algorithm identifies the route traveled. This information is transmitted electronically to the State authorities.

This will give the State a database from which mileage by commercial vehicles of various types on various types of roads can be obtained.

This seems like a promising source of exposure data. It is reasonable to assume that all States will eventually go to automatic systems of collecting commercial vehicle mileage information for fuel tax allocation. The system will also streamline reporting and paperwork for the carriers and they may be willing to install the units in their fleets.

The electronic vehicle clearance services identify a vehicle at a point, but do not track it over a route. These services will enable transponder-equipped trucks to have their safety status, credentials, and weight checked at mainline speeds. Vehicles that are safe and legal and have no outstanding out-of-service citations will be allowed to pass the inspection/weigh facility without delay. To use this system for collecting exposure information, a researcher would have to follow the vehicle from one inspection station to the next. There is currently much work being done on transponders that have "read-write" capabilities. Thus, a commercial vehicle passing through the inspection station could have the unique identification of the station recorded or the station could keep the record of the identification of the vehicle that passes through. If the vehicle kept a record of stations visited, the information would have to go into map-matching software to get the routes and then be entered into a database. If the stations kept the records, then the station data would have to be processed to find the paths of the vehicles and develop the vehicle mileage. The system, as conceptualized here, would be computationally challenging and does not appear to be a promising source of exposure data.

One of the technological developments brought about by ITS is better motion detectors, which were needed to study the actual paths, speeds, and accelerations of vehicles performing maneuvers in traffic. This information is needed to understand the micro-behavior of vehicles in traffic, which, in turn, is needed to design ITS systems.

There is a potential for using this advanced motion-detection technology together with WIM systems to collect information about the distribution of centers of gravity of commercial vehicles. Center of gravity is a surrogate for roll stability of vehicles and its distribution and exposure are often desired in analyses of rollover accidents.

The measurement of the center of gravity of a truck could be obtained by having the vehicle travel over a superelevated curve (of known superelevation) with a WIM system. The motion-detection system would precisely follow the vehicle's path and determine the radius of curvature of the vehicle's tires and also the record of the velocity over the path.' The forces acting on the vehicle would be measured at certain locations by the WIM. The information is sufficient to determine the vehicle's center of gravity, which would be calculated by microprocessor.

The center-of-gravity information would be recorded for each vehicle that passes over the instrumented curve. Information on the vehicle type could also be read from the vehicle's bar code or by an automatic vehicle identification system and could be added to the record. It is conceivable that a series of such stations could be built at sites selected by a sampling design to get the distribution of roll stability of commercial vehicles.

Advanced Traveler Information Systems (ATIS)

Advanced Traveler Information Systems (ATIS) provide the motorist with highway maps and other traffic and geographic information. For example, if a car is equipped with a map-based route planning system, this system might retain information on the route followed and provide more accurate data of the type that is traditionally sought through a trip diary. Speeds and travel times might also be incorporated.

Route guidance is a feature that holds the best potential for exposure data. At the basic level, route guidance is a static map. The map can be used to plan routes and provide directions to a destination. More sophisticated features would combine certain real-time (or dynamic) information on congestion, construction, and alternate routes with the map display. Route guidance (or navigation) systems may be either mobile- or infrastructure-based. "Mobile-based" means it is self-contained in the vehicle, while "infrastructure-based" implies that the capability resides in a central location and the information is communicated to the vehicle. The navigation capability requires position determination. The system must be able to track the position of the vehicle on a real-time basis using GPS or other methods. This is true for both the mobile- and infrastructure-based systems. A current program supported by FHWA is the In-Vehicle Routing and Navigation System (INRANS).

The attraction for exposure measurement would be the capability of the system to store the actual route followed by the vehicle. Traditional survey methods have drivers keep a diary to record where they went and when. This would provide much more accurate information. In principle, the travel could be linked with roadway characteristics, vehicle characteristics (including perhaps cargo weight and type for trucks), and driver characteristics. A sampling scheme to select vehicles and days could provide representative data for any geographic region, or vehicle or driver population.

ATIS may have a very different implementation in the trucking industry. Although some independent operators may be interested in a route planning system like that being developed for passenger cars, fleets are more likely to be interested in tracking systems that keep dispatchers appraised of the current location of all vehicles. A communication capability may also be part of such a system. Such a tracking system might also be able to preserve a history of the travel of individual vehicles. Information on the vehicle status and condition might be communicated back to the system over the course of the trip. Again, the equivalent of trip diaries may be generated for every vehicle in a fleet with such a system.

Advanced Traffic Management System (ATMS)

Historical Summary and Purpose: ITS technologies offer considerable improvements in data collection and dissemination in all areas of transportation. They are promising sources of exposure data for highway safety analyses. To date, however, little attention has been given to this application of data from ITS sources. The principle guiding documents for ITS developments in the United States — IVHS America's *Strategic Plan for Intelligent Vehicle-Highway Systems in the United States,* and the U.S. Department of Transportation's *IVHS Strategic Plan: Report to Congress* — make scant mention of the potential for integrating data from Intelligent Vehicle-Highway Systems (IVHS) sources into highway safety databases. FHWA is currently evaluating proposals for the national ITS system architecture study. Highway safety applications are addressed in the system architecture study to ensure that the architecture accommodates these applications. Therefore, the results of this proposed study are urgently needed.

Several opportunities for extracting exposure data from IVHS technologies are readily identifiable:

- Roadway-based exposure data from improved traffic surveillance systems.
- Vehicle-based exposure data from improved commercial vehicle monitoring systems.
- Individual-based exposure data from proposed route guidance systems.

Advanced Traffic Management Systems (ATMS) are the foundation for ITS, and more accurate and widespread surveillance of traffic conditions is a keystone of advanced traffic management. ITS America has proposed a long-term (20-year) goal of 30,577 km of freeway and 64,372 km of urban arterial roadways covered by surveillance systems. These systems will provide more accurate traffic volume data on the most important roadways in the major metropolitan areas of the United States.

The Commercial Vehicle Operations (CVO) component of ITS is a promising source of exposure data for large trucks. Since commercial vehicle applications will be one of the earliest areas of ITS implementation, this area deserves special attention in the proposed research. Automatic vehicle identification, classification, and location systems will become more widespread in commercial vehicle fleets. One application of data from these systems that will be the subject of an operational test during the next several years is the use of these data for determining vehicle-miles traveled in a State for taxation purposes. The same data would be a valuable measure of exposure for highway safety analyses.

One feature of the Advanced Traveler Information Systems (ATIS) component of ITS is in-vehicle route guidance, which requires a communications link between individual travelers and the centralized traffic management center. The concept, simply stated, is that travelers starting a trip enter their current location and intended destination into an on-board computer that has a two-way communications link to the traffic management center, and the computer — through some combination of the in-vehicle database of historical traffic conditions and updates on current traffic conditions from the traffic management center — identifies a recommended travel route. Information on the traveler and his/her trip origins and destinations would be a valuable source of individual-based exposure data.

Traffic management systems are an important source of the traffic information upon which Intelligent Transportation Systems are based. Traffic management systems are also a potential source of exposure data for highway safety studies. Most of the traffic management systems currently in operation or being designed are limited in scope to freeways. System functions include surveillance, control, and information. Surveillance involves real-time monitoring of traffic conditions (traffic volume and occupancy and, in some cases, speed) on a link-by-link basis in the freeway system. The control function may include ramp metering, for example. The information function refers to advising travelers about accidents or poor traffic conditions ahead via changeable message signs, highway advisory radio, traffic reports on commercial radio stations, etc.

Data Contents and Structure: The traffic volume data available from traffic management systems are generally aggregated over shorter time periods and are measured at more closely spaced intervals than the exposure data typically used for highway safety studies. In fact, the level of detail of the volume data is likely to exceed the needs of many, if not most, highway safety study objectives.

Typical current practice employed by traffic management systems for measuring traffic conditions includes detector stations at 0.8-km intervals along the freeway. The detector stations commonly consist of one inductive loop detector in each freeway lane to measure traffic volume and occupancy. At a subset of those stations, pairs of loop detectors may be used so that speed can also be measured. Twenty- to sixty-second traffic volumes are counted and then transmitted from a local control unit at the detector station to a traffic management center at which volume data from all stations are gathered, processed, monitored in real time, disseminated (in some centers), and stored.

Transportation Research Circular 378 lists freeway traffic management systems currently in operation or in the planning, design, or construction phase. As of 1991, the following areas had operational freeway traffic management systems with a significant number of traffic volume measurement locations: Chicago, Detroit, Long Island, Los Angeles, Minneapolis/St. Paul, Northern Virginia, Phoenix, San Diego, and Seattle. Dozens of urban areas are planning, designing, or constructing systems.

Experimental Design, Sample Plan, and Location Distribution: Each system operates independently and is unique with respect to the scope of surveillance coverage; location of detector stations; detector and communications technologies; and data collection, processing, and storage procedures. To illustrate the similarities and differences among systems, more detailed descriptions will be provided for two urban areas: Seattle and Minneapolis/St. Paul.

Seattle Traffic Management System: The Seattle traffic management system is operated by the Washington State Department of Transportation. The system has grown and evolved since the early 1970s. Traffic volume data are collected at approximately 200 stations. The stations are spaced at approximately 0.8-km intervals. This system provides traffic condition monitoring for approximately 113 km of freeway. Currently, four freeways are monitored: I-5, I-90, SR-405, and SR-520. The system will be expanded within the next several years to add a fifth freeway (SR-167).

Detector stations typically consist of inductive loop detectors in each freeway lane to measure traffic volume and occupancy. At a limited number of stations, pairs of loop detectors in each lane are used to measure speed. Traffic measurements at a detector station are recorded at a local control unit and transmitted to the traffic management center every 20 s. At the center, the volume data are aggregated to 5-min, 15-min, and l-h volumes. Both per lane and total directional volumes are transmitted to the center. Volume data from the detector stations are not disaggregated by vehicle type. There are, however, separate vehicle classification data collection sites in the Seattle area.

The occupancy data are displayed on a dynamic map that is updated every 20 s. Real-time monitoring of the map display is one of several methods used to identify potential incident locations.

The volume data from the detector stations have several uses. The traffic management center uses the volume data to evaluate changes in the ramp metering system, including adjusting metering rates at ramps or analyzing additions to the ramp metering system. Other groups within the Washington State Department of Transportation also make frequent use of the volume data, including design, traffic operations, and traffic data offices.

All volume data from all detector stations are stored. Data are stored as 5-min, 15-min, and l-h volumes. The data are stored on the center's computer system within the system's memory capacity; currently, approximately 10 months of data are available on-line. Older data are archived on magnetic tape or diskette. With some exceptions, data for a given detector station are available for as long as that station has been in operation, some for as long as 25 years. Exceptions include gaps in available data due to detectors being temporarily out of service for maintenance, system expansion, or during freeway reconstruction activities. No assurances can be given that data requested for specific detectors and for specific time periods are available. The availability of data can be determined only through the processing required to access and download the data.

Loop detector data cannot be considered 100 percent accurate. The accuracy of data from loop detectors, however, is generally comparable to other standard methods of measuring traffic volumes. The volume data transmitted to the center from the local control units at each detector station are checked to ensure its quality. Volume counts for an individual lane that fall beyond specified minimums or maximums or that differ more than a specified amount from the volume counts for other lanes at the detector station are flagged as either bad or suspect. These flags are recorded in the files containing the volume data. The flagging process is considered conservative-i.e., some data flagged as suspect because of differences between lanes may, in fact, be correct. Flagged data are excluded from station-wide measures.

Minneapolis/St. Paul Traffic Management System: The Minneapolis Department of Transportation operates a Traffic Management Center to manage traffic on the freeways in the Minneapolis/St. Paul Twin Cities metropolitan area. The center was constructed in 1972. Traffic volume data are collected at approximately 650 stations spaced at approximately 0.8-km intervals. This system monitors traffic on approximately 402 km (805 directional kilometers) of freeway. The freeways monitored include six Interstate highways (I-35E, I-35W, I-94, I-394, I-494, and I-694), as well as seven State highways (Routes 5, 36, 62, 77, 100, 169, and 212).

Detector stations typically consist of inductive loop detectors in each freeway lane to measure traffic volume and occupancy. Traffic speed is calculated based upon these measures. Detectors are also located on entrance and exit ramps. The detectors operate and transmit data to the center 24 hours per day. For control purposes, the center uses 1-min running averages that are updated every 30 s.

All data are archived. The basic time interval for archived data is a 5-min period. The archived data are stored in compressed binary format. Access programs transform the data, extract subsets that are requested, and aggregate data to the desired form. Traffic volume and occupancy data and calculated speeds can be aggregated in 5-, 15-, and 30-min; hourly; and daily time periods. Data can be provided by lane or aggregated for all lanes at a detector station. Data are available for approximately the past 2 or 3 years.

The data are provided "as is." There is no filtering to extract erroneous data, such as due to detector malfunctioning. Volume and occupancy data that deviate from certain thresholds are flagged, and those flags are included in the database. Appropriate use of the data requires familiarity with the area and with this type of data.

Data Acquisition and Documentation: Requests for Minneapolis/St. Paul volume data are handled by the Traffic Management Center on a case-by-case basis. The center has limited staff resources to process requests. The staff can handle requests for small amounts of data and provide the data for specified stations and time periods on diskette to the requester. If the amount of data requested is large, then it may be necessary for the requester to come to the center; the center provides access and the necessary software for the requester to decompress and download the data. The center is considering providing access to data through Internet at some future date. There are no confidentiality requirements or other restrictions on the use of volume data obtained from the center.

Minneapolis/St. Paul data are routinely used in-house and are provided to researchers and government agencies. Several periodic reports are routinely developed using the data, including a congestion report identifying congestion hot spots, a lane closure report that identifies allowable lane closures, a traffic report for traffic forecasting personnel, and a quarterly report on peak-hour volumes and AADT. There is no cost for obtaining the data and there are neither limitations nor confidentiality requirements on the use of the data.

Requests for data should be directed to:

Jim Aswegan
Freeway Operations
Metropolitan Division
Waters Edge
1500 West County Road, B2
Roseville, MN 55113

Reference

(1) Transportation Research Board. ***Transportation Research Circular 378, Freeway Operations Project Summary.*** September 1991.

Transportation Planning Surveys

This area covers a range of transportation planning surveys. These are usually household surveys conducted by mail or telephone. Examples are the Transportation Planning Package of the U.S. Census (CTPP). This survey provides nationwide data that form the basis for many State and local transportation planning efforts. However, only trips to and from work are included. The other general source in this area is regional planning surveys. These are also household surveys patterned after the CTPP. The geographic coverage is limited of course, but more detailed information is frequently collected, often for a broader range of trip purposes than just travel to and from work.

Census Transportation Planning Package (CTPP)

Purpose: The Census Transportation Planning Package (CTPP) is a set of special tabulations of the 1990 census data tailored to meet the data needs of transportation planners. The 1990 CTPP was produced by the Bureau of the Census and was sponsored by State Departments of Transportation under a pooled funding arrangement with the American Association of State Highway and Transportation Officials. The CTPP program was coordinated and is technically supported by the Federal Highway Administration of the U.S. Department of Transportation.

The CTPP consists of tables of sociodemographic and journey-to-work information. These tables provide information on commuter travel flows and characteristics; baseline origin-destination data on local work trips; household characteristics; and worker characteristics for use in travel forecasting models and for monitoring car-pooling and transit use. The CTPP data on commuter flows are also used to evaluate and select projects, develop traffic congestion management systems, and identify transportation corridors that need capacity expansion.

In addition, the CTPP also provides travel-to-work and vehicle availability information used in the preparation of vehicular travel and pollutant emissions profiles, computation of regional average rates of vehicle occupancy in the commute to work, and the evaluation of the impact of long-range transportation plans on air quality in compliance with the Clean Air Act Amendments of 1990.

Source: The source of information for the CTPP is the U.S. decennial census, particularly questions 23a and b, and 24a and b, that were asked of a sample of households. These questions asked for mode to work last week, vehicle occupancy, and time the work trip was started and how many minutes it took. This information, together with information on employment location, residential location, and sociodemographics, is the basis of the CTPP.

Organization: Two sets of data packages were produced: (1) statewide packages for each State and the District of Columbia and (2) urban packages for each "CTPP region” as defined by Metropolitan Planning Organizations (MPO).

The statewide CTPP consists of six parts (A through F). Part A contains characteristics of persons, workers, and housing units by county and by place of residence of 2,500 or more population (city, town, village, etc.). Part B contains characteristics of workers by county and place of work of 2,500 or more population. Part C contains characteristics of workers in journey-to-work flows between counties and places of residence of 2,500 or more population and counties and places of work of 2,500 or more population. Parts D, E, and F are similar to parts A, B, and C except for more detailed cross-tabulations of counties of 750,000 or more population and places of 75,000 or more population.

The urban CTPP has eight parts. Part 1 contains the characteristics of persons, workers, and housing units by traffic analysis zone or census tract (MPO option) of residence. Part 2 contains the characteristics of workers by traffic analysis zone or census tract. Part 3 contains characteristics of workers in journey-to-work flows from traffic analysis zone to traffic analysis zone, or from census tract to census tract. Part 4 contains detailed cross-tabulations of trip generation characteristics for the urbanized area, transportation study area, and metropolitan area. Part 5 does not exist, but is a "place-holder" to retain comparability with the 1988 Urban Transportation Planning Package (UTPP). Part 6 contains detailed cross-tabulations of workers in journey-to-work flows between "super districts" (aggregations of traffic analysis zones or census tracts) in CTPP regions of 1,000,000 or more population. Part 7 contains characteristics of workers by census tract of work with an emphasis on economic characteristics. Part 8 contains detailed cross-tabulations of characteristics of workers in journey-to-work flows between traffic analysis zones or census tracts for CTPP regions of 1,000,000 or more population.

Coverage: The 1990 CTPP is the fourth in a series of special transportation-oriented tabulations from the decennial census. In 1960, information on the place of work, mode of travel to work, and automobiles available at home was collected. Tabulations of worker streams were available in a special report for Standard Metropolitan Statistical Areas of more than 250,000 population. Information on automobile availability could be obtained in the series of census reports on housing.

The key transportation-related data collected in the 1970 census were again: place of work, mode of travel to work, and automobiles available in the home. The main difference between the 1960 and 1970 data was the level of geographic coding of the work place. In 1970, specific work addresses were required, while in 1960, only the city or county was identified. A special census product of sociodemographic and journey-to-work information could be ordered by the States and MPOs for transportation planning purposes.

In the 1980 decennial census, additional information on vehicle occupancy, travel time to work, and car and van availability was collected. The place-of-work data were coded to census tracts or blocks. As in 1970, States and MPOs could order special tabulations of demographic and journey-to-work information (now called the Urban Transportation Planning Package).

Strengths and Limitations: The CTPP provides detailed information on the journey-to-work trip for the entire country. Information includes mode, time of journey start, journey time, vehicle occupancy, and sociodemographics of the workers. Since the journey to work is the dominant trip purpose in the morning peak-traffic period, the data in the CTPP could be used to determine exposures for that particular time period. Obviously, any study using this approach would have to consider the portion of traffic in that time period not associated with the work journey.

The availability of similar journey-to-work information from previous censuses allows for the analysis of trends and changes in exposure for the morning peak-traffic period.

Since the information in the CTPP is limited to the journey to work, the CTPP is not a good source of exposure information for any times other than morning traffic-peak periods.

Sampling Errors: Variable sampling rates were used in the sample portion of the census. In general, in less densely populated areas, one in two households was sampled; while in densely populated areas, the rate was one in eight households. When all sampling rates are taken into account across the country, one in every six households was sampled.

The standard error of sample estimates can be calculated using tables and procedures given in ***Appendix C - Accuracy of the Data*** of the CTPP documentation.

Access: CTPP data are available from the Bureau of Transportation Statistics of the U.S. Department of Transportation on CD-ROM, together with the software (TransVU — CTPP Edition) to display and retrieve the data. TransVU — CTPP Edition is a Microsoft Windows application that provides both map and tabular view of CTPP data and simplifies extraction of CTPP tables in dBASE, Lotus, and comma-delimited or fixed-format test files. The CTPP CD-ROM and a copy of TransVU — CTPP Edition software are available from the Bureau of Transportation Statistics without charge.

Traffic Volume Data — Errors of VMT Estimates Based on Traffic Counts and Section Length

Typically, vehicle-miles traveled (VMT) are estimated from traffic counts and highway mileage. While the basic idea is simple, it can be implemented in several ways, which lead to different estimates with different errors.

This is the summary of a brief analysis of these techniques, including the method recommended in the HPMS for estimating VMT. Only the results are shown, not the sometimes tedious algebra. Two of the three procedures involve nonlinear expressions; therefore, linear approximations were used as usual. Therefore, the formulas are good approximations only if the coefficients of variation of the data are "small." A value of 0.1 is, for nearly all practical purposes, "small," 0.2 is small for most, and even 0.3 might be adequate for some approximate estimates.

Basic Definitions

The highway (system) studied has the length L and is divided into N sections of lengths l_i; their average is l_o. A sample of n sections is used; each section has the same probability of being selected. On section I, the average daily traffic is x_i. Its mean overall section is x_o. Variables s(x) and s(l) are the standard deviations of x_i and l_i. Their coefficients of variations are $c_x = s(x)/x_o$, and $c_l = s(l)/l_o$. One also needs the correlation coefficient ρ between the x_i and l_i. For instance, if in more densely settled areas traffic is heavier and sections are shorter, there is a negative correlation. On the other hand, if highways of a different character are combined, those with heavier traffic might have longer sections than those with lighter traffic. Then, there would be a positive correlation. Such correlations can appreciably influence the errors of VMT estimates. Therefore, they must be empirically determined and incorporated into the calculations. Formula *3* on page 3-3-9 of the ***traffic monitoring guide*** appears to do this implicitly.[1] However, this is a formula for the standard error of a biased estimate that is less relevant than the mean square error (see below).

The total vehicle-miles traveled on the L miles of highway are:

$$V=\sum l_i x_i = L x_o (1 + c_x c_l \rho) \tag{1}$$

where the second term in the parentheses reflects the effects of correlations between section length and volume.

[1] This formula is, aside from a misprint, equivalent to formula (6.10) in section 6.4 of W.G. Cochran, Sampling Techniques, Third Edition, Wiley, 1977.

The Unbiased Estimator

If n highway sections are randomly selected out of N with equal probabilities, the unbiased estimator of total VMT is:

$$\hat{V}_1 = \frac{N}{n} \sum x_i l_i \qquad (2)$$

where the sum is over the n elements of the sample. It has a standard deviation (equal to the mean square error, because the estimator is unbiased) given by:

$$SD(\hat{V}_1)^2 = (Lx_o c_x)^2 (\frac{1}{n}(1 + 2\frac{c_l}{c_x}\rho + (\frac{c_l}{c_x})^2) - c_l^2 \rho) \qquad (3)$$

if the finite population correction is ignored. The effect of a correlation between section length and volume is complex. If n is large, the expression in the right parentheses can become negative. This means simply that the linear approximation used for the product $x_l l_l$ is no longer valid.

A "Quick and Dirty" Estimator

This estimator averages the observed x_l and multiplies the average by the length of the highway system:

$$\hat{V}_2 = L \sum x_i / n \qquad (4)$$

It is a biased estimator. Its expected value is:

$$E(\hat{V})_2 = Lx_o \qquad (5)$$

It differs from the unbiased estimator by a factor of $1/(1+c_x c_l \rho)$. The bias disappears if the x_l and l_l are uncorrelated ($\rho = o$); it does not decrease when the sample size is increased. For a negative correlation and large coefficients of variation, $1+c_x c_l \rho$ can be small, and $\hat{V}_2$ can be a gross overestimate of V, no matter how large the sample. The standard error is given by:

$$SD(\hat{V}_2)^2 = (Lx_o c_x)^2 / n \qquad (6)$$

However, because it is a biased estimator, the mean square error given by:

$$MSQE(\hat{V}_2)^2 = E(\hat{V}_2 - V)^2 \tag{7}$$

is more meaningful, because it includes the bias into the error calculation:

$$MSQE(\hat{V}_2)^2 = (Lx_o c_x)^2 (\frac{1}{n} + c_l^2 \rho) \tag{8}$$

The second term in the parentheses reflects the effect of the bias. The first term decreases with increasing sample site n; the second remains constant. Thus, if ρ and c_l are not negligible, this is not a good estimator.

The Ratio Estimator Recommended by HPMS

The unbiased estimator calculates VMT on the sample sections and then divides it by the sample fraction—the ratio of sampled sections to total sections. The ratio estimator also calculates VMT on the sample sections, but then divides it by the ratio of the combined length of the sample sections and the total length L:

$$\hat{V}_3 = \frac{L}{\sum l_i} \sum x_i l_i \tag{9}$$

The advantage of this is that it reduces the effect of the varying length of the sample sections on the variance of the estimate; its disadvantage is that the estimate is biased. The expected value is

$$E(\hat{V}_3) = x_o L(1 + \rho c_x c_l - \frac{1}{n} \rho c_x c_l) = V(1 - \frac{1}{n}) \frac{\rho c_x c_l}{1 + \rho c_x c_l} \tag{10}$$

For this estimator, the bias decreases with increasing sample size; it also decreases with decreasing correlation ρ and with decreasing coefficients of variation c_x and c_l. Its mean square error is given by:

$$MSQE(\hat{V}_3)^2 = (Lx_o c_x)^2 (\,(1 + 2c_l^2 \rho^2)/n - c_l^2 \rho^2) \tag{11}$$

Again, the right parentheses can become zero or negative if the linear approximations are no longer valid.

Comparing the Unbiased Estimator and FHWA's Estimator

The difference between equation (9) and V is the bias of FHWA's estimator. Thus,

$$\frac{\text{BIAS}}{v} = -\frac{1}{n}\frac{c_x c_l \rho}{1 + c_x c_l \rho} \tag{12}$$

is the bias as a proportion of the actual value. This bias is the price to pay for the reduction of the variance achieved by the ratio estimator. Whether it is worthwhile depends on the difference between the mean square error of the two estimators. The difference of their squares is

$$MSQE(\hat{V}_1)^2 - MSQE(\hat{V}_3)^2 = \frac{(Lx_o c_x)^2}{n}\left(2\frac{c_l}{c_x}\rho + \left(\frac{c_l}{c_x}\right)^2 - 2c_l^2\rho\right) \tag{13}$$

This difference can be positive as well as negative. It can become large with either sign, but the relevance of this is limited because before very large values are reached, the linear approximations become invalid.

However, it appears worthwhile to check in real applications how large an improvement of the variance is provided by using a biased estimator, and whether despite the bias, the mean square error will be improved.

APPENDIX: HPMS FORMS AND DATA FORMAT

The appendix contains selected forms reproduced from the 1993 edition of the FHWA Highway Performance Monitoring System Field Manual, OMB No. 2125-0028.

Template – 1

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

TOTALS OF URBANIZED AREAS, SMALL URBAN AREAS, RURAL AREAS, AND STATEWIDE

STATE:__________ **STATE FIPS** CODE:______ **UNITS: [ ] English 1/ [ ] Metric 2/** **DATA YEAR:** ______ **DATE:**______

URBANIZED AREAS TOTAL

POPULATION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM: PRINCIPAL ARTERIALS: INTERSTATE	OTHER FREEWAYS & EXPRESSWAYS	OTHER	MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
		LENGTH							
		TRAVEL (1,000)							

SMALL URBAN AREAS TOTAL

POPULATION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM: PRINCIPAL ARTERIALS: INTERSTATE	OTHER FREEWAYS & EXPRESSWAYS	OTHER	MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
		LENGTH							
		TRAVEL (1,000)							

RURAL AREAS TOTAL

POPULATION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM: PRINCIPAL ARTERIALS: INTERSTATE	PRINCIPAL ARTERIALS: OTHER	MINOR ARTERIAL	COLLECTOR: MAJOR	COLLECTOR: MINOR	LOCAL	TOTAL
		LENGTH							
		TRAVEL (1,000)							

STATEWIDE TOTALS

POPULATION (1,000)	NET LAND AREA	TOTAL LENGTH	TOTAL TRAVEL (1,000)

1/ English units for length and travel are miles and vehicle–miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and vehicle–kilometers (in thousands), respectively.

Template – 2

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

INDIVIDUAL URBANIZED AREAS

STATE:________ STATE FIPS CODE: UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR______ DATE:________

Shaded cells are reserved for titles and computer software generated values. Enter data in the unshaded cells only.

URBANIZED AREA CODE	NONATTAINMENT AREA CODE 3/	POPULATION (1,000)	NET LAND AREA	DATA TYPE	FUNCTIONAL SYSTEM: PRINCIPAL ARTERIAL: INTERSTATE	PRINCIPAL ARTERIAL: OTHER FREEWAYS & EXPRESSWAYS	PRINCIPAL ARTERIAL: OTHER	MINOR ARTERIAL	COLLECTOR	LOCAL	TOTAL
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							
				LENGTH							
				TRAVEL (1,000)							
				OCCUPANCY 4/							

1/ English units for length and travel are miles and daily vehicle–miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and daily vehicle–kilometers (in thousands), respectively.
3/ The National Ambient Air Quality Standards Nonattainment Area Code is the same as the Urbanized Area Code of the primary urbanized area contained in the nonattainment area. When the Urbanized Area is not in a nonattainment area, code zero.
4/ Average vehicle occupancy is reported to the nearest tenth of a person.

Template - 3

SYSTEM LENGTH AND DAILY VEHICLE TRAVEL

DONUT AREA DATA FOR INDIVIDUAL NAAQS NONATTAINMENT AREAS

STATE: ____________ STATE FIPS CODE: ______ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: ________ DATE: __________

Shaded cells are resewed for titles and computer sofware generated values. Enter data In the unshaded cells only.

NONATTAINMENT AREA CODE 3/	POPULATION (1,000)	NET LAND AREA	DATA TYPE	RURAL AND SMALL URBAN FUNCTIONAL SYSTEMS COMBINED			
				PRINCIPAL ARTERIAL	MINOR ARTERIAL AND (MAJOR) COLLECTOR	MINOR COLLECTOR AND LOCAL	TOTAL
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				
			LENGTH				
			TRAVEL (1,000)				

1/ English units for length and travel are miles and daily vehicle-miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and daily vehicle-kilometers (in thousands), respectively.
3/ The National Ambient Air Quality Standards (NAAQS) Nonattainment Area Code is the same as the Urbanized Area Code of the primary urbanized area contained in the nonattainment area.

AIRQUAL1 6/3/93

Template – 4

MINOR COLLECTOR AND LOCAL FUNCTIONAL SYSTEM LENGTH

BY SURFACE TYPE AND VOLUME GROUP

STATE: ____________ STATE FIPS CODE: ________ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: ______ DATE: ________

Shaded cells are reserved for titles and computer software generated values. Enter data In the unshaded cells only.

FUNCTIONAL SYSTEM	SURFACE TYPE					TOTAL LENGTH
	PAVED				UNPAVED	
	HIGH	INTERMEDIATE	LOW	SUBTOTAL		
RURAL MINOR COLLECTOR						

FUNCTIONAL SYSTEM / SURFACE TYPE	TRAFFIC VOLUME GROUP					TOTAL LENGTH
RURAL MINOR COLLECTOR	UNDER 100	100– 199	200– 499	500– 1,999	2,000 & OVER	
PAVED						
UNPAVED						
TOTAL						
RURAL LOCAL	UNDER 50	50– 199	200– 499	500– 1,999	2,000 & OVER	
PAVED						
UNPAVED						
TOTAL						
URBAN LOCAL	UNDER 200	200– 499	500– 1,999	2,000– 3,999	4,000 & OVER	
PAVED						
UNPAVED						
TOTAL						

1/ English units consist of miles.

2/ Metric units consist of kilometers.

Template – 5

FATAL AND INJURY MOTOR VEHICLE ACCIDENTS

STATE: ______________ STATE FIPS CODE: ________ UNITS: [ ] English 1/ [ ] Metric 2/ DATA YEAR: __________ DATE: ______________

Shaded cells are reserved for titles and computer software generated values. Enter data in the unshaded cells only.

HIGHWAY SYSTEM	SYSTEM LENGTH	DAILY TRAVEL (1,000)	ACCIDENTS FATAL	ACCIDENTS NONFATAL INJURY	PERSONS 3/ FATALITIES	PERSONS 3/ TOTAL NONFATAL INJURY 4/	PERSONS 3/ MOST SERIOUS NONFATAL INJURY	PEDESTRIANS FATALITIES	PEDESTRIANS NONFATAL INJURY
RURAL									
INTERSTATE									
OTHER PRINCIPAL ARTERIAL									
NATIONAL HIGHWAY SYSTEM (NHS)									
OTHER (NON-NHS)									
SUBTOTAL									
MINOR ARTERIAL									
NATIONAL HIGHWAY SYSTEM (NHS)									
OTHER (NON-NHS)									
SUBTOTAL									
MAJOR COLLECTOR									
MINOR COLLECTOR									
LOCAL									
TOTAL – RURAL									
URBAN									
INTERSTATE									
OTHER FREEWAYS & EXPRESSWAYS									
NATIONAL HIGHWAY SYSTEM (NHS)									
OTHER (NON-NHS)									
SUBTOTAL									
OTHER PRINCIPAL ARTERIAL									
NATIONAL HIGHWAY SYSTEM (NHS)									
OTHER (NON-NHS)									
SUBTOTAL									
MINOR ARTERIAL									
NATIONAL HIGHWAY SYSTEM (NHS)									
OTHER (NON-NHS)									
SUBTOTAL									
COLLECTOR									
LOCAL									
TOTAL – URBAN									
TOTAL – STATEWIDE									

1/ English units for length and travel are miles and daily vehicle–miles (in thousands), respectively.
2/ Metric units for length and travel are kilometers and daily vehicle–kilometers (in thousands), respectively.
3/ Includes pedestrians.
4/ Includes most serious nonfatally injured persons.

Template - 6

TRAVEL ACTIVITY BY VEHICLE TYPE

BASIC DATA

STATE:______________ STATE FIPS CODE:____ DATA YEAR:________ DATE:______________

Shaded cells are reserved fo[illegible]itles and computer software generated values. Enter data in the unshaded cells only.

FUNCTIONAL SYSTEM	PERCENT OF TRAVEL													
	MOTORCYCLES (OPTIONAL)	PASSENGER CARS (2 AXLE, 4 TIRE)	LIGHT TRUCKS (OTHER 2 AXLE, 4 TIRE)	BUSES	SINGLE-UNIT TRUCKS			SINGLE-TRAILER TRUCKS			MULTI-TRAILER TRUCKS			TOTAL
					2 AXLE, 6 TIRE	3 AXLE	4 AXLE OR MORE	4 AXLE OR LESS	5 AXLE	6 AXLE OR MORE	5 AXLE OR LESS	6 AXLE	7 AXLE OR MORE	
RURAL														
INTERSTATE														
OTHER PRINCIPAL ARTERIAL														
MINOR ARTERIAL														
MAJOR COLLECTOR														
MINOR COLLECTOR														
LOCAL														
URBAN														
INTERSTATE														
OTHER FREEWAYS & EXPRESSWAYS														
OTHER PRINCIPAL ARTERIAL														
MINOR ARTERIAL														
COLLECTOR														
LOCAL														

Template – 7

TRAVEL ACTIVITY BY VEHICLE TYPE

SUPPLEMENTAL DATA

STATE:__________ STATE FIPS CODE:______ DATA YEAR:__________ DATE:__________

1. VEHICLE CLASSIFICATION DATA ON TEMPLATE 6 ARE REPRESENTATIVE OF DATA NORMALLY COLLECTED DURING THE INDICATED HOURS, DAYS OF THE WEEK, AND MONTHS:

___AM/PM TO ___AM/PM, [ ] ALL HOURS OF DAY

[ ] ALL DAYS
[ ] SUNDAY
[ ] MONDAY
[ ] TUESDAY
[ ] WEDNESDAY
[ ] THURSDAY
[ ] FRIDAY
[ ] SATURDAY

[ ] ALL MONTHS
[ ] JANUARY
[ ] FEBRUARY
[ ] MARCH
[ ] APRIL
[ ] MAY
[ ] JUNE
[ ] JULY
[ ] AUGUST
[ ] SEPTEMBER
[ ] OCTOBER
[ ] NOVEMBER
[ ] DECEMBER

2. VEHICLE CLASSIFICATION DATA ON TEMPLATE 6 ARE REPRESENTATIVE OF DATA NORMALLY COLLECTED ON THE FOLLOWING HIGHWAY SYSTEMS:

[ ] ALL SYSTEMS
[ ] RURAL
[ ] URBAN
[ ] STATE OWNED
[ ] INTERSTATE
[ ] OTHER PRINCIPAL ARTERIAL
[ ] MINOR ARTERIAL
[ ] (MAJOR) COLLECTOR

3. INDICATE BELOW WHERE EACH OF THE SPECIFIC VEHICLE TYPES, LISTED IN THE LEFT COLUMN, ARE INCLUDED ON TEMPLATE 6:

SPECIFIC VEHICLE TYPE	PREFERABLE VEHICLE TYPE	REPORTED VEHICLE TYPE IS CONTAINED IN THE FOLLOWING CATEGORY ON TEMPLATE 6
P-AXLE, 4-TIRE TRUCKS WITHOUT A TRAILER	3	
2-AXLE, 4-TIRE TRUCKS WITH A TRAILER	3	
2-AXLE, 6-TIRE PICKUP TRUCKS WITHOUT A TRAILER	5	
2-AXLE, 6-TIRE PICKUP TRUCKS WITH A TRAILER	8 - 10 ASAPPROPRIATE	
OTHER SINGLE-UNIT TRUCKS WITH SEMI-TRAILERS	8 - 13 ASAPPROPRIATE	
OTHER SINGLE-UNIT TRUCKS WITH FULL-TRAILERS	8 - 13 ASAPPROPRIATE	

4. COMMENTS.

Template – 8

U. S. TERRITORY INFORMATION

TERRITORY:________________ TERRITORY FIPS ____CODE: UNITS [ ] English 1/ [ ] Metric 2/ DATA YEAR:________ DATE:____________

Shaded cells are reserved for titles and computer software generated values. Enter data in the unshaded cells only.

CATEGORY	RURAL	URBAN	TOTAL
POPULATION (1,000)			
NET LAND AREA			
FEDERAL-AID TERRITORIAL HIGHWAY SYSTEM --- ARTERIAL:			
PAVED LENGTH			
UNPAVED LENGTH			
SUBTOTAL			
DAILY TRAVEL (1,000)			
FEDERAL-AID TERRITORIAL HIGHWAY SYSTEM --- COLLECTOR:			
PAVED LENGTH			
UNPAVED LENGTH			
SUBTOTAL			
DAILY TRAVEL (1,000)			
OTHER PUBLIC ROADS:			
PAVED LENGTH			
UNPAVED LENGTH			
SUBTOTAL			
DAILY TRAVEL (1,000)			
ALL PUBLIC ROADS:			
TOTAL LENGTH			
TOTAL DAILY TRAVEL (1,000)			
MOTOR VEHICLE ACCIDENTS ON PUBLIC ROADS:			
NUMBER OF FATAL ACCIDENTS			
NUMBER OF NONFATAL INJURY ACCIDENTS			
NUMBER OF FATALITIES			
NUMBER OF NONFATALLY INJURED PERSONS			

1/ English units for length and travel are miles and daily vehicle-miles (in thousands), respectively.

2/ Metric units for length and travel are kilometers and daily vehicle- kilometers (in thousands), respectively.

DATA ITEM SUMMARY TABLE

Data Item Requirements

Under the columns headed "Required Universe Items" and "Required Sample Items," in the data item summary table, an "A" indicates that the item is required for "All" of that-system's section records, both universe and sample (standard and donut area). An "S" indicates that the item is only required if the section record is part of the "Standard" sample panel. A "D" indicates that the item is only required if the section record is part of the "Donut" area sample panel. The following abbreviations are used in the column headings:

All Records - Universe and Sample Data

Prin Art/ Oth NHS	Report these items for all principal arterial and other National Highway System sections. The principal arterial system includes the rural and urban Interstate, urban other freeways and expressways and rural and urban other principal arterial functional systems. The National Highway System is made up primarily of these same systems, but may include a minor amount of roadways on other functional systems.	
Int	Interstate	Rural and Urban
OFE	Other Freeways and Expressways	Urban
OPA	Other Principal Arterial	Rural and Urban
MA	Minor Arterial	Rural and Urban
MaC	Major Collector	Rural
MiC	Minor Collector	Rural
Col	Collector	Urban
Loc	Local	Rural and Urban
Pos	This column indicates the position of the item in the section record as reported to FHWA.	
Len	This column indicates the length of the field used for the data item.	

Caution Regarding the Data Item Coding Summary

Several data items in both the universe and sample data portions of these records require additional discussion regarding the type of section for which the data item is applicable. For example, Percent Passing Sight Distance (Item 62) is required only for rural paved, two-lane facilities. The summary table only indicates that this item is required for the rural standard sample sections. Do not rely solely on the data item summary table for system coding requirements; each data item description must be consulted for complete details.

Universe Data

			Required Universe Items								
			<---- Rural --->				<--- Urban ---->				
Item No.	Pos	Len	Prin Art/ Oth NHS	MA	MaC	MiC & Loc	Prin Art/ Oth NHS	MA	Col	Loc	Data Item
											Identification
1	1-100	100									State Control Field
2	101	1	A	A	A	A	A	A	A	A	Reporting Units
3	102-103	2	A	A	A	A	A	A	A	A	Year
4	104-105	2	A	A	A	A	A	A	A	A	State code
5	106	1	A	A	A	A	A	A	A	A	Type of Section
6	107-109	3	A	A	A	A	A	A	A	A	County code
7	110-133	24	A	A	A	A	A	A	A	A	Section Identification
8	134-147	14	A	A			A				LRS Milepoint/ Kilometerpoint
9	148	1	A	A	A	A	A	A	A	A	Rural/Urban Designation
10	149-152	4	A	A	A	A	A	A	A	A	Urbanized Area Sampling Technique and Code
11	153-155	3	A	A	A	A	A	A	A	A	Nonattainment Area Code
											System
12	156-157	2	A	A	A	A	A	A	A	A	Functional system
13	158	1	A	A	A	A	A	A	A	A	Generated Functional System Code
14	159	1	A	A	A	A	A	A	A	A	National Highway System
15	160	1	A				A				Unbuilt Facility
16	161-165	5	A				A				Official Interstate Route Number
17	166	1	A	A			A				Route Signing
18	167	1	A	A			A				Route Qualifier
19	168-175	8	A	A			A				Signed Route Number

Key: A - Code for "All" universe, standard and donut area sample sections.
S - Code for all "Standard" sample sections.
D - Code for all "Donut" area sample sections.

Universe Data (Cont.)

Item No.	Pos	Len	Rural: Prin Art/ Oth NHS	Rural: MA	Rural: MaC	Rural: MiC & Loc	Urban: Prin Art/ Oth NHS	Urban: MA	Urban: Col	Urban: Loc	Data Item
											Jurisdiction
20	176-177	2	A	A	A	A	A	A	A	A	Governmental Ownership
21	178-179	2	A	A	A	A	A	A	A	A	Special Systems
											Operation
22	180	1	A	A	A	A	A	A	A	A	Type of Facility
23	181	1	A	A	A	A	A	A	A	A	Designated Truck Route/ Parkway
24	182	1	A	A	A	A	A	A	A	A	Toll
											Other
25	183-188	6	A	A	A	A	A	A	A	A	**Section Length**
26	189	1		A	A			A	A		**Donut Area Sample** Panel **AADT** Volume **Group**[1]
27	190-191	2	A	A	A		A	A	A		Standard Sample Panel **AADT** Volume Group
28	192-197	6	A	S&D	S&D		A	S&D	S&D		**AADT**
29	198	1	A	S&D			A	S&D			**AADT** Derivation
30	199-200	2	A	S	S		A	S	S		Number of Through Lanes

Key:- Code for **"All"** universe, standard and donut area sample sections.
S - Code for all "Standard" sample sections.
D - Code for all **"Donut"** area sample sections.

[1] The **"A"** in the **summary** table cells for the Donut Area Volume Group (Item **26**) is meant to indicate that all data records (universe **and** sample) **for** the noted functional systems in a donut area are to include these **data.**

Universe Data (Cont.)

Item No.	Pos	Len	Rural: Prin Art/ Oth NHS	Rural: MA	Rural: MaC	Rural: MiC & Loc	Urban: Prin Art/ Oth NHS	Urban: MA	Urban: Col	Urban: Loc	Data Item
											Other (Cont.)
31	201	1					A	S			Urban Location
32	202	1	A	S	S		A	S	S		Access Control
33	203	1	A	S	S		A	S	S		Median Type
34	204-206	3	A	S	S		A	S	S		Median Width
35	207-209	3	A	S			A				Roughness (IRI)
36	210-211	2	A	S	S		A	S	S		Pavement Condition (PSR)
37	212-225	14									Reserved for Federal Use
38	226-229	4	A	A	A	A	A	A	A	A	Record Type

Required Universe Items: <---- Rural ---> ; <--- urban --(1)->

(A Universe section record ends here unless the section contains HOV Operations, and/or Surveillance Systems. If one or both of these exist on the applicable PAS section; data Items 81 and/or 82 must be added to the universe record.)

Key:
- \- Code for "All" universe, standard and donut area sample sections.
- S - Code for all "Standard" sample sections.
- D - Code for all "Donut" area sample sections.

Sample Data

Item No.	Pos	Len	Required Sample Items <--- Rural ---> Int	OPA	MA	MaC	<----- Urban ---> Int	OFE	OPA	MA	Col	Data Item
												Identification
39	230-241	12	S	S	S&D	S&D	S	S	S	S&D	S&D	Sample Number
40	242	1	S	S	S&D	S&D	S	S	S	S&D	S&D	Sample Subdivision
												Computational
41	243-248	6			D	D				D	D	Donut Area Expansion Factor

(A Donut area sample section record ends here, unless it is also a standard sample section record.)

Item No.	Pos	Len	Rural Int	OPA	MA	MaC	Urban Int	OFE	OPA	MA	Col	Data Item
42	249-254	6	S	S	S	S	S	S	S	S	S	Standard Expansion Factor
												Pavement
43	255-256	2	S	S	S	S	S	S	S	S	S	Surface/ Pavement Type
44	257	1	S	S	S	S	S	S	S	S	S	Pavement Section
45	258-260	3	S	S			S	S	S			SN or D
46	261	1	S	S			S	S	S			Type of Base
47	262	1	S	S			S	S	S			Type of Subgrade
												Improvements
48	263-265	3	S	S	S	S	S	S	S	S	S	Overlay or Pavement Structure Thickness
49	266-269	4	S	S	S	S	S	S	S	S	S	Year of Surface Improvement
50	270-271	2	S	S	S	S	S	S	S	S	S	Type of Improvement

Key:
- \- Code for "All" universe, standard and donut area sample sections.
- S - Code for all "Standard" sample sections.
- D - Code for all "Donut" area sample sections.

Sample Data (Cont.)

Item No.	Pos	Len	Rural Int	Rural OPA	Rural MA	Rural MaC	Urban Int	Urban OFE	Urban OPA	Urban MA	Urban Col	Data Item
												Geometrics
51	272-273	2	S	S	S	S	S	S	S	S	S	Lane Width
52	274	1	S	S	S	S	S	S	S	S	S	Shoulder Type
53	275-278	4	S	S	S	S	S	S	S	S	S	Shoulder Width
54	279	1					S	S	S	S	S	Peak Parking
55	280-282	3	S	S	S	S	S	S	S	S	S	ROW Width
56	283	1	S	S	S	S	S	S	S	S	S	Widening Feasibility
57	284	1				S						Horizontal Alignment Adequacy
59	285-375	91	S	S	S		S	S	S			Curves by Class
	376	1	S	S	S	S						Type of Terrain
60	377	1				S						Vertical Alignment Adequacy
61	378-419	42	S	S	S		S	S	S			Grades by Class
62	420-422	3	S	S	S	S						Percent Passing Sight Distance
												Traffic/Capacity
63	423-425	3	S	S	S	S	S	S	S	S	S	Speed Limit
64	426-428	3				S						Weighted Design Speed (calculated)
65A	429-432	4	S	S	S	S	S	S	S	S	S	Percent Single Unit Comm. Vehicles
65B	433-436	4	S	S	S	S	S	S	S	S	S	Percent Combination Comm. Vehicles
66	437-438	2	S	S	S	S	S	S	S	S	S	K-Factor
67	439-441	3	S	S	S	S	S	S	S	S	S	Directional Factor
68	442-446	5					S	S	S	S	S	Peak Capacity
69	447-449	3	S	S	S	S	S	S	S	S	S	V/SF Ratio (calculated)
70	450-451	2					S	S	S	S	S	Turning Lanes
72	453-454	2					S	S	S	S	S	Signalization
							S	S	S	S	S	% Green Time

Key: A - Code for "All" universe, standard and donut area sample sections.
S - Code for all "Standard" sample sections.
D - Code for all "Donut" area sample sections.

Sample Data (Cont.)

Item No.	Pos	Len	Rural Int	Rural OPA	Rural MA	Rural MaC	Urban Int	Urban OFE	Urban OPA	Urban MA	Urban Col	Data Item
												Traffic/Capacity
73	455-460	6	S	S	S	S	S	S	S	S	S	Future AADT
74	461-462	2	S	S	S	S	S	S	S	S	S	Future AADT Year
												Environment
75	463-464	2	S	S	S	S	S	S	S	S	S	Climate Zone[2]
76	465	1	S	S	S	S	S	S	S	S	S	Drainage Adequacy
77	466	1	S	S	S	S	I	I	I	I		Type of Development
78	467-468	2	S	S	S	S	S	S	S	S	S	Number Grade Separated Interchanges
79	469-474	6	S	S	S	S	S	S	S	S	S	Number At-Grade Intersections
80	475-476	2	S	S	S	S	S	S	S	S	S	Number At-Grade Railroad Crossings

The following supplemental data are reported only if HOV Operations and/or Highway Surveillance Systems exist on the applicable PAS (universe or standard sample). Do not report these data items if the features do not exist.
3

Item No.	Pos	Len	Rural Int	Rural OPA	Rural MA	Rural MaC	Urban Int	Urban OFE	Urban OPA	Urban MA	Urban Col	Data Item
												Supplemental
81	Varies[3]	58	A	A			A	A	A			HOV Operations
82	Varies[3]	7	A	A			A	A	A			Surv. Systems

Key: A - Code for "All" universe, standard and donut area sample sections.
S - Code for all "Standard" sample sections.
D - Code for all "Donut" area sample sections.

[2] The Climate Zone entry (Item 75) is made by the Submittal Software Package. It may be changed by the State.

[3] The positions for these data items depend on whether they are attached to a universe record or to a standard sample record, and whether one or both exist on the section. For universe records, the positions are 230-287 for Item 81 and 288-294 for Item 82, if they both exist. For a standard sample record the positions are 477-534 for Item 81 and 535-541 for Item 82, if they both exist. If only one of the data items exist, it will begin at position 230 for a universe record and at position 477 for a standard sample record. The ending position depends on the data item length.

HNR-20/11-97(400)QE