

Data from TRB-Proposed National Monitoring System and Procedures for Analysis of Truck Accident Rates

T. CHIRA-CHAVALA

To follow trends of truck accident involvement rates requires reliable information on truck accidents and travel. Procedures for estimating truck accident involvement rates and their confidence limits on the basis of variabilities inherent in the sample design of the TRB-proposed National Monitoring System (NMS) are presented. Formulas for computing confidence limits of national and state truck accident involvement rates per mile of travel are given for any level of disaggregation. The quality of truck accident and travel data that may be expected from implementing the NMS, together with consistent estimation of confidence limits of accident involvement rates, would represent significant improvement over truck safety statistics available from existing data programs.

Fatalities, personal injuries, and property damage resulting from truck crashes constitute major public health and economic problems. Each year, about 5,000 deaths, 50,000 personal injuries, and up to \$1 billion in property damage are caused by accidents involving heavy or medium trucks (those with gross vehicle weight more than 10,000 lb) (1). Public concern is growing that truck safety is quickly deteriorating. However, existing truck accident and travel data are inadequate to address this concern, support essential government planning functions, guide public and private policy decisions on truck operations, or guide actions to reduce accidents and losses.

The annual numbers of truck accidents reported by NHTSA's National Accident Sampling System (2), the National Safety Council accident statistics (3), and the Office of Motor Carriers (OMC) accident data (4) are not in agreement. Estimates of annual truck miles of travel reported by existing sources also vary greatly [e.g., FHWA's *Highway Statistics* (5), the Census Bureau's *Truck Inventory and Use Survey* (6), and the private National Truck Trip Information System (7)]. As a result, information is not available to reliably assess the magnitude and the trends of the safety performance of various truck types or to determine the extent to which truck safety may be improving or worsening (1). This has helped to fuel the controversy about truck safety.

Recently, two major studies aimed at closing the truck safety information gap have been completed. They are the National Governors' Association (NGA) study on uniform truck accident reporting among states (8) and TRB's *Special Report 228: Data Requirements for Monitoring Truck Safety* (1). The NGA study recommends uniform data elements and criteria for states to use in reporting accidents involving trucks or

buses. *Special Report 228* recommends a plan for developing the National Monitoring System (NMS) that will assemble nationwide accident and travel data of medium and heavy trucks (those with gross vehicle weight ratings more than 10,000 and 26,000 lb, respectively) on an ongoing basis. Recommendations from the two studies are related and complementary in that implementation of the NGA recommendations would bring about uniform truck accident data among states. Uniform state accident data could then be combined into a single national truck accident data base, a major component of the proposed NMS.

Among the recommendations of *Special Report 228* are a detailed data plan for developing the NMS, an implementation timetable, organizational arrangements, and cooperative efforts among federal and state governments and industry. To illustrate intended applications of the NMS, the TRB report identifies benchmark questions that the NMS should, as a minimum, provide the data to address (Figure 1). The TRB report states that accurate information on trends of truck accident frequency, truck miles of travel, and rates (Benchmark Question 1) is particularly important because it is needed for guiding and evaluating policy decisions concerning truck operations.

The TRB report, however, does not describe procedures for analyzing the accident and travel data to be available from the NMS. This paper provides this missing link. Its objectives are to describe truck accident and travel data that can be expected from the implementation of the NMS and to present procedures for estimating truck accident involvement rates and their statistical confidence limits on the basis of variabilities inherent in the sampling of accidents and truck miles.

SUMMARY OF NGA RECOMMENDATIONS ON UNIFORM TRUCK ACCIDENT REPORTING AMONG STATES

The NGA study (8) recommends that states report all truck accidents that result in fatalities, serious injuries (in which the injured is taken from the scene), and property damage only (PDO) in which at least one of the vehicles involved is towed away because it is inoperable as a result of the accident. States may choose lower accident reporting thresholds as long as "towaway" accidents can be distinguished from other PDO accidents. The study also recommends that states include uniform data elements (9) in their accident report forms or supplemental forms.

- Question 1.* What are the trends in numbers of truck accidents, amount of truck travel, and truck accident involvement rates per mile traveled, disaggregated by truck type, road class, region, and carrier type?
- Question 2.* What are the relationships between truck accident involvement rates and driver age?
- Question 3.* What are the kinds of crashes (e.g. overturn, jackknife, separation of units, collision with another vehicle) in which different types of trucks are likely to be involved, and what are the corresponding severity levels?
- Question 4.* What are the trends in the numbers and severity of crashes involving trucks placarded for carrying hazardous cargoes, disaggregated by truck type, road class, region, carrier type, and driver age?
- Question 5.* What are the possible underreporting and other biases of other truck accident and travel data programs, on the basis of comparisons of statistics between these programs and the proposed National Monitoring System?

FIGURE 1 Benchmark questions defining the NMS (I).

The NGA conducted a pilot test in five Midwestern states for 3 months in 1988 to determine whether police officers who report accidents might encounter major reporting problems. The pilot test does not indicate major reporting problems by police officers who participated in the test (9).

OVERVIEW OF DATA EXPECTED FROM TRB-PROPOSED NMS

The proposed NMS calls for the development of two data modules: national truck accident data and national truck travel data. Data elements and their levels of details for these two modules are shown in Figures 2 through 4. Data from the NMS would eventually permit derivation of statistically reliable estimates of truck accident rates both nationwide and for individual states. However, it will take a number of years for all states to adopt uniform truck accident data reporting within their boundaries. *Special Report 228*, therefore, recommends that the NMS be implemented in two phases so that improved

- Vehicle
 - o Vehicle configuration
 - o Cargo body style
 - o Vehicle identification number
 - o Hazardous cargoes involvement:
 - Was hazardous cargo present in the truck?
 - Hazardous cargo class
 - Was hazardous cargo released?
- Carrier
 - o Carrier identification
- Driver
 - o Driver identification
 - o Driver age
- Accident
 - o Accident events
 - o Accident severity
 - Number of fatalities
 - Number of injured people, transported away
 - Number of vehicles towed away
- Roadway
 - o Roadway functional class
 - o Degree of urbanization (Rural or urban)
 - o Trafficway (undivided or divided)
 - o Access control
 - o Road surface condition
- Environment
 - o State
 - o Weather
 - o Time of day
 - o Light condition (day or night)
 - o Day of week
 - o Month
 - o Year

FIGURE 2 Accident data elements required by the NMS (I).

- Vehicle
 - o Vehicle configuration
 - o Cargo body style
- Carrier
 - o Carrier identification number
- Driver
 - o Driver age
- Roadway
 - o Roadway functional class
 - o Degree of urbanization (rural or urban)
- Environment
 - o State
 - o Light condition (day or night)
 - o Season

FIGURE 3 Truck travel data elements required by the NMS (I).

national truck accident and travel data can become available immediately.

Phase 1 Data

The TRB report recommends that a national truck travel component and a short-term national truck accident component be developed in Phase 1.

Short-Term National Truck Accident Data

While states are moving toward adopting the NGA uniform truck accident reporting recommendations, short-term national truck accident data could be developed immediately by augmenting data from the truck subset of NHTSA's Fatal Accident Reporting System (FARS) and General Estimates System (GES). The short-term national truck accident data, to be available on a quarterly basis, would include the following:

- A census of the nation's fatal truck accidents, with a level of detail similar to that shown in Figure 2, would be available. Data from the truck subset of NHTSA's FARS would be augmented by adding carrier ID, cargo body style, and hazardous cargo class. These additional variables will be obtained by matching the FARS cases to OMC accident (50-T) reports and by examining available state sources for the FARS cases with no matching 50-T reports. The short-term fatal truck accident data will permit derivation of statistically reliable estimates of fatal truck accidents at both the national and the state levels.

- A sample of the nation's total truck accidents (fatal and nonfatal) with a level of detail similar to that shown in Figure 2 would be available, possibly with up to 10,000 involvements per year. Data from the truck subset of the NHTSA's GES would be augmented by adding carrier ID, cargo body style, hazardous cargo presence and release, hazardous cargo class, and "towaway." The additional variables will be acquired by matching the GES cases with the OMC 50-T reports and examining available state sources for the GES cases that have no matching 50-T reports. Because the GES accident sample is designed for making national accident estimates, these short-term total truck accident data will permit derivation of sta-

Vehicle configuration	Roadway functional class (source: states)
Single unit truck (two-axle, six-tire)	Interstate
Single unit truck (three-or more axles)	Other principal arterial
Truck/trailer	Major arterial
Truck-tractor (bobtail)	Major collector
Tractor-semitrailer	Minor collector
Tractor (double)	Local road or street
Tractor (triple)	Degree of urbanization (source: states)
Cannot classify	Rural
Cargo body style	Urban
Van	Trafficway
Tank	Undivided two way
Flatbed	Divided, without traffic barriers
Dump	Divided, with traffic barriers
Concrete mixer	One way
Auto Transporter	Access control
Garbage/refuse	Unlimited access
Other	Full control
Was hazardous cargo present in the truck?	Other
Yes	Road surface condition
No	Dry
Hazardous cargo class	Wet
four-digit placard number or name	Snow
one-digit placard number	Ice
Was hazardous cargo released?	Sand, mud, dirt, oil
Yes	Other
No	Unknown
Carrier identification	State
U.S. Department of Transportation number	Weather
Interstate Commerce Commission motor carrier number	No adverse conditions
State number	Rain
Other number	Sleet or hail
None	Snow
Driver age	Fog
Accident severity	Blowing sand, soil, dirt, or snow
Number of fatalities	Severe crosswind
Number of injured people, transported away	Other
Number of vehicles towed away	Unknown
Accident events (in order of occurrence)	Time of day
Ran off road	Light conditions
Jackknife	Daylight
Overturn	Dark (not lighted)
Downhill runaway	Dark (lighted)
Cargo loss or shift	Dawn
Fire	Dusk
Separation of units	Unknown
Collision with pedestrian	Day of week
Collision with motor vehicle	Month
Collision with parked vehicle	Year
Collision with train	
Collision with pedalcycle	
Collision with animal	
Collision with fixed object	
Collision with other object	
Other	

FIGURE 4 Levels of detail of the NMS data elements (1).

tistically reliable national estimates of total truck accidents, but not estimates for individual states.

National Truck Travel Data

National truck travel data with a level of detail similar to that shown in Figure 3 can be assembled immediately in Phase 1 for use in deriving national and state estimates. National truck travel data would be assembled from truck classification-travel data that individual states are collecting for FHWA's Highway Performance Monitoring System (HPMS) on the basis of probability samples of HPMS road sections. The TRB report recommends that states follow guidelines published in FHWA's *Traffic Monitoring Guide* (10) in collecting truck classification-travel data. The state data, however, would not have information on cargo body style, carrier type, or driver age. These three variables will be obtained from the fraction of the state Motor Carrier Safety Assistance Program's (MCSAP's) truck safety inspections that is based on random selections of trucks.

Phase 2 Data

After states adopt the NGA uniform truck accident reporting recommendations, a long-term national truck accident data base will be developed to replace the short-term accident data. Uniform truck accident data from individual states will be combined into a national truck accident data base. The long-term accident data will have a level of detail as shown in Figure 2 and will permit derivation of statistically reliable accident estimates nationwide and for individual states.

National truck travel data developed in Phase 1 will also be applicable in Phase 2.

Quality of the NMS Data

Special Report 228 emphasizes that the NMS accident and travel data must have known quality and limitations that can be accounted for in the analysis of truck accident involvement rates.

GENERALIZED FORMULAS FOR ESTIMATING ACCIDENT RATES AND CONFIDENCE LIMITS

Accident involvement rates per mile of travel are the most commonly accepted measure for comparing safety performance of various truck types or operating conditions. Truck travel data of the NMS will be derived from truck counts on samples of road sections conducted independently by states. Estimates of truck miles of travel within individual states will have random variations due to the sampling. Random variations due to sampling will also exist in national estimates of truck accidents that are obtained from the short-term total truck accident data of Phase 1, because these data will be derived from the truck subset of GES (a probability sample of the nation's police-reported accidents). Truck accident involvement rates derived from the sampled accident and travel data will inherit these sampling variabilities, which must be accounted for in order to correctly interpret trends of accident involvement rates. This can be accomplished by estimating statistical confidence intervals (CIs) of accident involvement rates. Formulas to do this are presented below for the following two cases:

- Both accident and travel data come from probability samples; the source of random variations to be accounted for in estimating confidence limits is that due to the sampling of accidents and truck miles.

- Truck travel data come from probability samples, whereas the accident data come from the census of reported truck accidents. The source of random variations to be accounted for in estimating confidence limits is that due to the sampling of truck miles; there is no random variations due to sampling for the census of accidents.

The following formulas are derived for accident involvement rates per mile of travel, disaggregated by two variables. Disaggregation involving any other number of variables follows the same procedure.

Case A: Both Accident and Travel Data Come from Probability Samples

Let i and j denote the i th truck type and the j th road class; Y_{ij} and X_{ij} denote the number of accidents and truck miles, respectively, for the i th truck type on the j th road class; and R_{ij} denote the accident involvement rate per mile of travel for the i th truck type on the j th road class.

Kish (11) derived a theoretical value of the sampling variance of R_{ij} , $Var(R_{ij})$, as shown in Equation 1. All parameters in Equation 1 are expected values (i.e., theoretical values).

$$\begin{aligned} Var(R_{ij}) &= E[R_{ij} - E(R_{ij})]^2 \\ &= \frac{1}{X_{ij}^2} [Var(Y_{ij}) + R_{ij}^2 Var(X_{ij}) \\ &\quad - 2R_{ij} Cov(Y_{ij}, X_{ij})] \end{aligned} \quad (1)$$

An unbiased estimate of the sampling variance of R_{ij} can be obtained by substituting sample values throughout Equation 1 to yield Equation 2.

For Equation 2 and all the formulas that follow, the parameters shown are sample values (i.e., unbiased estimates of the expected values shown in Equation 1). The same symbols are used in Equation 1 and in all the other equations to eliminate the need for two different sets of symbols—one denoting the theoretical values and the other unbiased sample values. In this way, the volume of notations may be reduced significantly.

$$\begin{aligned} var(R_{ij}) &= \frac{1}{X_{ij}^2} [var(Y_{ij}) + R_{ij}^2 var(X_{ij}) \\ &\quad - 2R_{ij} cov(Y_{ij}, X_{ij})] \end{aligned} \quad (2)$$

where

- $var(X_{ij})$ = an unbiased estimate of the sampling variance of truck-miles, $Var(X_{ij})$;
- $var(Y_{ij})$ = an unbiased estimate of the sampling variance of the number of accidents, $Var(Y_{ij})$; and
- $cov(Y_{ij}, X_{ij})$ = an unbiased estimate of sampling covariance of accidents and truck miles, $Var(Y_{ij}, X_{ij})$.

For truck accident and travel data that are independently collected from different sample designs and sample units, $cov(Y_{ij}, X_{ij})$ does not exist. Therefore, an unbiased estimate of the sampling variance of R_{ij} becomes

$$var(R_{ij}) \approx \frac{1}{X_{ij}^2} [var(Y_{ij}) + R_{ij}^2 var(X_{ij})] \quad (3)$$

The $(1 - \alpha)$ percent CI of R_{ij} is expressed as

$$(1 - \alpha) \text{ percent CI}(R_{ij}) = R_{ij} \pm c\sqrt{var(R_{ij})} \quad (4)$$

where α is the Type I error and c is a two-sided normal variate corresponding to α .

Substituting Equation 3 into Equation 4 yields a generalized formula for CI of accident involvement rate for the i th truck type on the j th road class, as follows:

$$CI(R_{ij}) = R_{ij} \pm c \frac{\sqrt{var(Y_{ij}) + R_{ij}^2 var(X_{ij})}}{X_{ij}} \quad (5)$$

Case B: Accident Data Come from a Census and Travel Data Come from a Sample

For the census of accident involvements, variability due to sampling does not exist. Therefore $var(Y_{ij}) = 0$, and the generalized CI formula of Equation 5 becomes

$$CI(R_{ij}) = R_{ij} \pm c \frac{\sqrt{R_{ij}^2 var(X_{ij})}}{X_{ij}}$$

or

$$CI(R_{ij}) = \frac{R_{ij}}{X_{ij}} [X_{ij} \pm c\sqrt{var(X_{ij})}] \quad (6)$$

PROCEDURES FOR CALCULATING CIs OF INVOLVEMENT RATES USING NMS DATA

Accident involvement rates per mile of travel by truck type, road class, carrier type, and so on can be calculated from future NMS data. Random variations in accident involvement rates due to sampling variabilities of truck miles and accidents can be quantified by estimating statistical confidence intervals of these rates. Because truck accident data in Phases 1 and 2 of the NMS are expected to be different, different procedures for estimating CIs of accident involvement rates are suggested for the two phases, as follows. The notation is first introduced. For ease of illustration, the CI procedures are based on truck accident involvement rates disaggregated by two variables, truck type and road class. Other levels of disaggregation involving any number of variables follow the same procedure.

Notation Meaning

i	subscript denoting the i th truck type
j	subscript denoting the j th road class
k	subscript denoting the k th state
y	unbiased estimate of accident involvement within state
x	unbiased estimate of truck miles within state
r	unbiased estimate of accident involvement rate for the state
Y	unbiased national estimate of accident involvement
X	unbiased national estimate of truck miles
R	unbiased estimate of national accident involvement rate

Confidence Interval Procedure for Phase 1 Data

The NMS accident data in Phase 1 would consist of two short-term components: the census of nationwide fatal truck accident involvements and a sample of up to 10,000 nationwide total (fatal plus nonfatal) truck accident involvements.

Truck-mile estimates by state would be available from the classification counts that individual states carry out on samples of HPMS road sections. Procedures for calculating CIs for fatal and nonfatal involvement rates are as follows.

CI of Fatal Involvement Rates Nationwide and by State

The annual number of fatal truck involvements nationwide, Y_{ij} , would be known without sampling variance, because it is the census of fatal accident involvements. Estimates of truck miles of travel by state (x_{ijk}) and their sampling variances [$\text{var}(x_{ijk})$] would also be known from individual states' samples.

An unbiased estimate of the national fatal involvement rate for the i th truck type on the j th road class, R_{ij} , is obtained as

$$R_{ij} = \frac{Y_{ij}}{X_{ij}} = \frac{Y_{ij}}{\sum_k x_{ijk}}$$

From Equation 6,

$$CI(R_{ij}) = \frac{R_{ij}}{X_{ij}} \left[X_{ij} \pm c \sqrt{\text{var}(X_{ij})} \right]$$

States independently derive estimates of truck miles from

samples of road segments, and then these estimates are combined to yield a national estimate. Therefore

$$CI(R_{ij}) = \frac{R_{ij}}{\sum_k x_{ijk}} \left[\sum_k x_{ijk} \pm c \sqrt{\sum_k \text{var}(x_{ijk})} \right] \quad (7)$$

An unbiased estimate of the fatal involvement rate for the i th truck type on the j th road class in the k th state, r_{ijk} , is

$$r_{ijk} = \frac{y_{ijk}}{x_{ijk}}$$

From Equation 6,

$$CI(r_{ijk}) = \frac{r_{ijk}}{x_{ijk}} \left[x_{ijk} \pm c \sqrt{\text{var}(x_{ijk})} \right] \quad (8)$$

CI of Nonfatal Involvement Rates Nationwide

In Phase 1, only estimates of the national number of nonfatal accident involvements would be available from the NMS. Estimates of nonfatal accident involvements by state would not be statistically reliable. Therefore, only nationwide nonfatal truck involvement rates can be reliably estimated, not rates by state.

An estimate of the annual nationwide number of nonfatal involvements for the i th truck type on the j th road class (Y_{ij}) as well as its sampling variance, $\text{var}(Y_{ij})$, would be known. Estimates of truck miles of travel for individual states (x_{ijk}) and their sampling variances [$\text{var}(x_{ijk})$] would also be known from individual states' samples, which are independent of one another. These states' estimates of truck miles can be combined to yield a national estimate of truck miles.

An unbiased estimate of the national nonfatal involvement rate for the i th truck type on the j th road class is as follows:

$$R_{ij} = \frac{Y_{ij}}{X_{ij}} = \frac{Y_{ij}}{\sum_k x_{ijk}}$$

From Equation 5,

$$CI(R_{ij}) = R_{ij} \pm c \frac{\sqrt{\text{var}(Y_{ij}) + R_{ij}^2 \text{var}(X_{ij})}}{X_{ij}}$$

Substituting $X_{ij} = \sum_k x_{ijk}$ yields

$$CI(R_{ij}) = R_{ij} \pm c \frac{\sqrt{\text{var}(Y_{ij}) + R_{ij}^2 \sum_k \text{var}(x_{ijk})}}{\sum_k x_{ijk}} \quad (9)$$

Confidence Interval Procedure for Phase 2 Data

Recall that in Phase 2, the census of fatal and nonfatal accident involvements would be available nationwide and by state. Estimates of truck miles by state identical to those of Phase

1 would continue to be available in Phase 2. Because fatal and nonfatal accident data of Phase 2 would be derived from the same source, the same CI procedures apply for fatal and nonfatal accident involvement rates.

The numbers of accidents for individual states, y_{ijk} , would be known without any sampling variances. Estimates of truck miles of travel for individual states (x_{ijk}) and their sampling variances [$\text{var}(x_{ijk})$] would be known from individual states' samples, which are independent of one another.

A national estimate of involvement rate for the i th truck type on the j th road class is expressed as

$$R_{ij} = \frac{Y_{ij}}{X_{ij}} = \frac{\sum_k y_{ijk}}{\sum_k x_{ijk}}$$

From Equation 6,

$$\text{CI}(R_{ij}) = \frac{R_{ij}}{X_{ij}} \left[X_{ij} \pm c \sqrt{\text{var}(X_{ij})} \right]$$

Substituting $X_{ij} = \sum_k x_{ijk}$ yields

$$\text{CI}(R_{ij}) = \frac{R_{ij}}{\sum_k x_{ijk}} \left[\sum_k x_{ijk} \pm c \sqrt{\sum_k \text{var}(x_{ijk})} \right] \quad (10)$$

The estimate of involvement rates for the i th truck type on the j th road class in the k th state is expressed as

$$r_{ijk} = \frac{y_{ijk}}{x_{ijk}}$$

From Equation 6,

$$\text{CI}(r_{ijk}) = \frac{r_{ijk}}{x_{ijk}} \left[x_{ijk} \pm c \sqrt{\text{var}(x_{ijk})} \right] \quad (11)$$

CONCLUSION

Once the NMS is implemented, improved national truck accident and travel data will be available for the analysis of truck accident involvement rates by various levels of disaggregation. Reliable truck safety information will be available to support public and private policy decisions concerning truck safety and operations. In particular, various important trends of truck accident involvement rates could be more accurately measured, and random variations of these rates due to sampling variabilities inherent in the accident and travel data could be assessed. This would represent significant improvement over truck safety statistics available from existing data programs.

REFERENCES

1. *Special Report 228: Data Requirements for Monitoring Truck Safety*. TRB, National Research Council, Washington, D.C., 1990.
2. *National Accident Sampling System*. National Highway Traffic Safety Administration, U.S. Department of Transportation, 1985.
3. *Accident Facts*. National Safety Council, 1985.
4. *Accidents of Motor Carriers of Properties 1985*. Office of Motor Carriers, Federal Highway Administration, U.S. Department of Transportation, 1987.
5. *Highway Statistics*. Federal Highway Administration, U.S. Department of Transportation, 1982, 1985.
6. *1982 Census of Transportation: Truck Inventory and Use Survey*. Bureau of the Census, 1985.
7. K. L. Campbell, D. F. Blower, R. G. Gattis, and A. C. Wolfe. *Analysis of Accident Rates of Heavy-Duty Vehicles: Final Report*. DTNH 22-83-07188. University of Michigan Transportation Research Institute, Ann Arbor, 1988.
8. *NGA Motor Carrier Accident Reporting Project: Final Data Elements*. National Governors' Association, Washington, D.C., 1989.
9. R. Craft. NGA Motor Carrier Accident Reporting Project. Presented at the 16th International Forum on Traffic Records Systems, Bal Harbour, Fl., July 16-19, 1990.
10. *Traffic Monitoring Guide*. Federal Highway Administration, U.S. Department of Transportation, 1985.
11. L. Kish. *Survey Sampling*. University of Michigan, John Wiley & Sons, 1965.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.