Atlanta Vehicle Occupancy Monitoring

R. FRED FISHER, G. JACK WILLIAMS, AND J. PHILLIP BOYD

This report describes the implementation of a statistically defined survey technique for collecting vehicle classification and occupancy data in the Atlanta region. The paper describes the results of a stratified, areawide survey for collecting passenger occupancy rates. The potential movement of people provided by the capacity of our roadway system is virtually an untapped resource, according to the data collected for this study. Sponsored by the U.S. Department of Transportation, efforts are being made to improve the usefulness of passenger vehicles through current programs that include vanpooling, ride-sharing programs, and park-and-ride lots. The success of these ventures, which are likely to become more significant in the future, can be measured by a dependable vehicle-occupancy monitoring program. This research has proved that the Guide for Estimating Urban Vehicle Classification and Occupancy provides a statistically acceptable method to measure vehicle occupancy rates. The minimum sample requirement for determining occupancy rates by area and facility type is desirable for an annual program of this nature.

In recent years the need for monitoring the movement of people by vehicle has touched all aspects of society. Not only is the information valuable to transportation engineers and planners, but it has also become vital to those evaluating energy consumption, environmental quality, and trends in the nation's economy.

In 1979, the Georgia Department of Transportation (GDOT) and the Atlanta Regional Commission (ARC) participated in testing the Guide for Estimating Urban Vehicle Classification and Occupancy (1), which was prepared for the Federal Highway Administration. The overall objective in testing the procedures set forth in the guide was to establish a methodology for estimating vehicle classification and occupancy in the Atlanta region on an annual basis, during peak periods, and for daylight hours. The Atlanta region is the more than 2000-mile^2 area composed of Clayton, Cobb, DeKalb, Douglas, Fulton, Gwinnett, and Rockdale Counties in north central Georgia.

A joint effort was made by local planning agencies in order to define purposes for monitoring vehicle occupancy and classification so that potential use of this data could be maximized. These purposes are as follows:

1. Provide monitoring data for the Atlanta Regional Transportation Plan;
2. Provide a basis for assessing the impact of implementing the Metropolitan Atlanta Rapid Transit Authority's rapid rail lines;
3. Evaluate the effectiveness of such programs as carpooling, park-and-ride facilities, and high-occupancy vehicle lanes;
4. Validate transportation planning models;
5. Assess energy efficiency of travel;
6. Assess air quality related to transportation;
7. Provide data for project planning and design; and
8. Provide data for person miles traveled (PMT) when estimating procedures for vehicle miles of travel (VMT) are implemented in the Atlanta region.

SURVEY DESIGN

The survey was structured to test procedures in the guide and provide results compatible to existing data-collection and modeling techniques. The primary criterion was to develop a survey that would gather statistically reliable data at the lowest possible cost. To meet these objectives, the effort was divided into two surveys.

Stratified Areawide Study

A stratified areawide survey was designed to provide areawide data with eight stratifications. There were two highway functional classifications (freeway and nonfreeway) and four geographic areas, as illustrated in Figure 1. The methodology in the guide (1) was used in determining sample size for the areawide survey. This was considered a critical element in testing the validity of procedures set forth in the guide. The sample size of link days was estimated separately for vehicle classification and automobile occupancy to determine which travel measure controlled the design. A link day was defined as the sampling unit--i.e., a particular highway segment on a particular day.

Vehicle Classification

The first approach considered for determining sample size was by regional vehicle classification. This measure is defined as the ratio of VMT of a particular vehicle type to the total regional VMT. Before a minimum sample size can be computed, the composite standard deviation of the proportion of vehicles must be estimated. After assuming the standard deviation of the proportion of truck traffic, S_t, was the same as that of passenger vehicles, the composite standard deviation was computed for each stratification as follows:

$$S_{th} = (S_{tlh}^2 + S_{tsh} + S_{tdh}^2)^{1/2} \quad (1)$$

where

S_{th} = composite standard deviation of the proportion of trucks in stratum h,
S_{tlh} = standard deviation of the proportion of trucks across link days within a season in stratum h,
S_{tsh} = standard deviation of the proportion of trucks across seasons in stratum h, and
S_{tdh} = standard deviation of the proportion of trucks across time periods during the day (as a result of hourly counts) in stratum h.

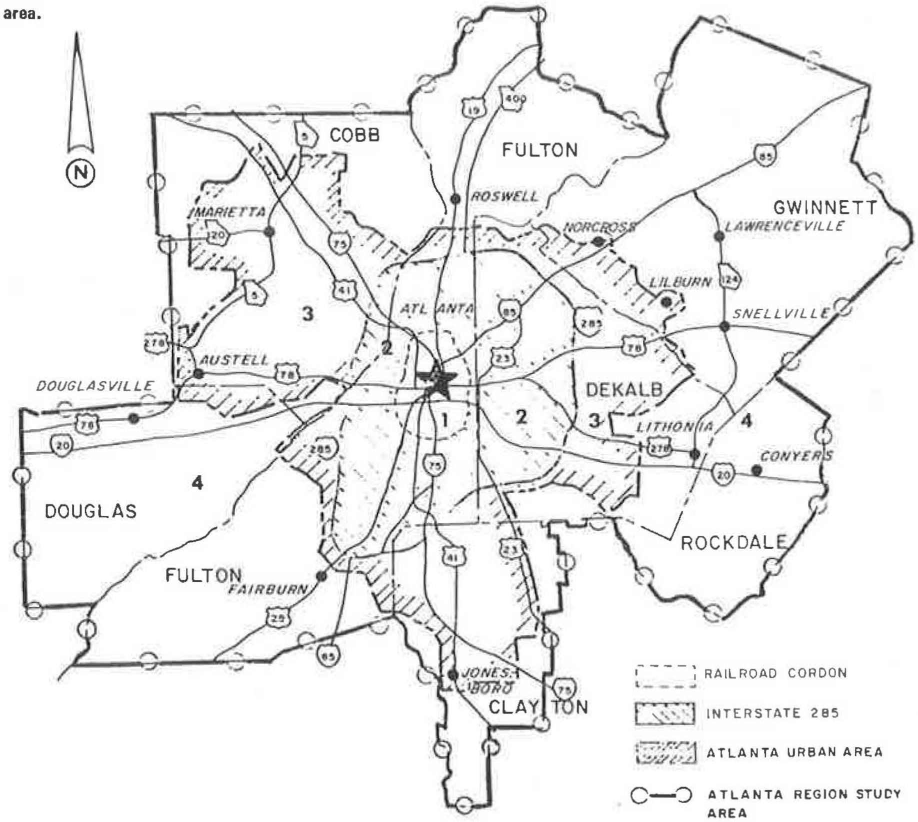
The following assumptions were made in calculating the composite standard deviations for each stratum:

1. S_{tlh} = (maximum % trucks - minimum % trucks)/3.5 (2)

2. S_{tsh} = 0 (the survey would not be conducted across seasons), and
3. S_{tdh} = 0.009 (recommended for all strata, based on default value).

Factor	Computation
Composite standard deviation	$S_{oh} = (S_{olh}^2 + S_{osh}^2 + S_{odh}^2)^{1/2}$
	$= (0.63^2 + .000^2 + .017^2)^{1/2} = 0.065$
Sample size per stratum	$N_h = (Z^2 \cdot S_{oh}^2)/(D_{ot}^2)$
90 percent confidence level	$N_h = (1.645^2 \cdot 0.065^2)/(0.03^2) = 12.7 \approx 13$
80 percent confidence level	$N_h = (1.28^2 \cdot 0.065^2)/(0.03^2) = 7.7 \approx 8$
Areawide tolerance	$D_o = Z(\sum_h W_h^2 \cdot S_{oh}^2/N_h)^{1/2}$
90 percent confidence level	$D_o = 1.645(\sum_h W_h^2 \cdot 0.065^2/13)^{1/2} = 0.011\ 88$
80 percent confidence level	$D_o = 1.280(\sum_h W_h^2 \cdot 0.065^2/8)^{1/2} = 0.011\ 78$

Figure 1. Atlanta vehicle-occupancy study area.



Thus, the lower confidence level could be selected with little loss of precision and would require a sample of only eight link days per stratum as opposed to 13 link days per strataum at the higher confidence level.

In analyzing the two methods for computing sample size, it was apparent that the average passenger-vehicle-occupancy approach should be used for determining the sample size. Eight link days were selected from each stratum for a total of 64 samples required for the areawide survey.

A sample selection technique on random highway links and random days was used to allocate the 64 required samples. A computer listing and map of all the links for the current Atlanta Regional Traffic Assignment Network was used to define the survey population. The network was delineated by direction and by the eight stratifications.

A computer program was developed to generate a series of random numbers within the range of the link-number assignment network. The series was of sufficient quantity that only one draw would be necessary for the selection of the samples in the various stratifications.

The samples were selected by using listings of all the links and random numbers. Because the link-node listing was in numerical order areawide, each link number selected had to be verified as to its stratification. If the link did not meet the required stratification, the selection was rejected, and the next random number was selected. This process was repeated until all 64 required samples were selected for the areawide survey.

Once the standard deviations were computed, the total sample size that would allow for a reliable estimate of the proportion of trucks within desired levels of precision was computed as follows:

$$N = \left[Z^2\left(\sum_h W_h S_{th}\right)^2\right] / D_t^2 \quad (3)$$

where

- N = number of link days required for monitoring vehicle classification and occupancy for the entire region,
- W_h = proportion of VMT occurring in stratum h to total regional VMT,
- Z = normal variate at specific confidence level,
- S_{th} = composite standard deviation for stratum h computed in Equation 1, and
- DT_t = desired tolerance (acceptable difference between estimated and true proportion of trucks).

The following assumptions were made in making this computation:

1. Z = 1.96 (value for a two-tailed statistical test at a 90 percent level of confidence) and
2. DT_t = ±0.02 (desired objective for the tolerance level).

The required minimum sample of link days within each stratum was then computed by the following equation:

$$N_h = N\left[(W_h S_{th}) \Big/ \left(\sum_h W_h S_{th}\right)\right] \quad (4)$$

The results of these computations for computing composite standard deviation and sample sizes are shown in Table 1. As illustrated, these computations resulted in an areawide sample size of 22 link days. Once the sample sizes were determined by strata, the precision of the areawide estimate was computed as follows:

$$DT_t = Z\left[\left(\sum_h W_h^2\right)\left(S_{th}^2/N_h\right)^{1/2}\right] \quad (5)$$

Table 1. Results of computation of sample size by vehicle classification when S_{tsh} = 0, S_{tdh} = 0.009, Z = 1.96, and DT_t = 0.02.

Facility	Factor	Inside Railroad Cordon	Railroad Cordon to I-285	I-285 to Urban Arterial Boundary	Urban Arterial Boundary to Region Boundary
Freeway	W_h (%VMT)	0.05	0.25	0.14	0.13
	S_{tlh}	0.014	0.042	0.057	0.071
	S_{tsh}	0.000	0.000	0.000	0.000
	S_{tdh}	0.009	0.009	0.009	0.009
	S_{th}	0.017	0.043	0.058	0.072
	N_h	1	5	4	4
Nonfreeway	W_h (%VMT)	0.03	0.19	0.10	0.11
	S_{tlh}	0.029	0.042	0.057	0.029
	S_{tsh}	0.000	0.000	0.000	0.000
	S_{tdh}	0.009	0.009	0.009	0.009
	S_{th}	0.030	0.043	0.058	0.030
	N_h	1	4	2	1

The areawide tolerance equaled the desired value of 0.02.

Average Passenger-Vehicle Occupancy

The second method used for determining sample size considered the measure of average passenger-vehicle occupancy. The composite standard deviation was computed as follows:

$$S_{oh} = (S_{olh}^2 + S_{osh}^2 + S_{odh}^2)^{1/2} \quad (6)$$

where

S_{oh} = standard deviation of average occupancy across link days within a season in stratum h,
S_{olh} = standard deviation of average occupancy across link days within a season for stratum h,
S_{osh} = standard deviation of average occupancy across seasons in stratum h, and
S_{odh} = standard deviation of average occupancy across hours during the day (as a result of hourly counting) in stratum h.

The following assumptions were made for this computation:

1. S_{ol} = 0.063 (recommended value based on previous studies),
2. S_{os} = 0 (the survey would not be conducted across seasons), and
3. S_{od} = 0.017 (recommended value based on previous studies).

Once the composite standard deviation was computed, the desired sample size of link days was then determined by using the following equation:

$$N_h = (Z^2 S_{oh}^2)/(DT_{oh})^2 \quad (7)$$

where

N_h = sample size in link days in stratum h,
Z = normal variate,
S_{oh} = composite standard deviation of average occupancy in stratum h, and
DT_{oh} = desired tolerance for stratum h (acceptable difference between the estimated average occupancy and the true value).

Once the sample sizes were determined by strata, the tolerance achieved at the regional level was computed as follows:

$$DT_o = Z\left[\left(\sum_h W_h^2\right)\left(S_{oh}^2/N_h\right)\right]^{1/2} \quad (8)$$

Tests were made of the sample-size computation by using confidence levels of 90 percent and of 80 percent, with a tolerance of ±0.03 for both tests. Results indicated, as shown below, a very small difference in the areawide tolerance (see Figure 2 for W_h) between the two confidence levels:

A random number, ranging in value equivalent to the days of the survey, was selected for each sample. This number was assigned to the sample to designate the date of observation.

Focused Multiple-Location Survey

A focused multiple-location survey was designed to fulfill needs for travel information for specific screenlines and corridors within the region. Some 22 locations were selected to monitor freeways and arterials crossing the railroad cordon, which included the central business district (CBD), and the Interstate perimeter highway (I-285). It was determined that the data from these locations coupled with the ongoing GDOT traffic-counting program would provide a good measure of person travel. In addition, five locations were selected to measure "before" occupancy rates on highways adjacent to proposed park-and-ride lots within the region.

For calculating sample-size requirements, the composite standard deviation was assumed equal to 0.018 for both vehicle classification and occupancy. To achieve a 90 percent confidence and ±0.03 tolerance, one day of data collection was required at each location. Thus, a higher accuracy for total traffic entering each cordon (railroad and perimeter) would be achieved.

Due to the characteristics of this survey, observations were required in both directions at each location (areawide survey observations were one direction only). The focused observations were randomly selected on days by using the same process for random days in the areawide survey.

DATA COLLECTION

Once the monitoring stations were selected, a work plan was prepared for collecting the vehicle classification and occupancy data. A listing of the monitoring stations that described the assigned station number, station type, geographic area, facility type, number of lanes, and location description was prepared.

Scheduling

The technique used to monitor the traffic was to manually observe each lane for a 15-min duration and a total of 45 min/h. The remaining 15-min period was used to record the observed information and provide a break for the monitor. Observations were made by one person parked on the shoulder of the road in such a manner as to maximize observation of passing vehicles and minimize disruption to normal traffic flow due to the presence of the surveyor's vehicle.

A work plan for the monitoring program was then established, and the required work crew was scheduled. Field data collection began on March 31, 1979, and was completed on May 30, 1979. Observations were made from 7:00 a.m. to 7:00 p.m. The individuals selected to make the manual observations were assembled and given their work schedule for the entire study period with the appropriate classifying instructions and counting forms. By scheduling the observations for the duration of the study, the field personnel knew

Figure 2. Form used during Atlanta-region vehicle-occupancy survey.

VEHICLE OCCUPANCY SURVEY

Sheet ____ of ____

Location ______________________________ Station No ________

Date __________ Time Begin ______ AM PM Time End ______ AM PM Direction of Count __________

Recorder __________________ Weather __________ No of Lanes in Count Direction __________

Enter Beginning Hour	Minutes	Lane No (Median) 3 2 1 1 2 1	Passenger Cars; Pickups; Light Trucks — Occupants 1	2	3	4	5	6+	Heavy Trucks — Occupants 1	2	3+	Buses — Occupants 1-25	26-46	47+	TOTAL
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
___ Am Pm	00-15														
	15-30														
	30-45														
	45-00														
Sub-Total															
TOTAL															

their days off in advance. Their knowledge of the complete work schedule served as a performance incentive.

Field Data-Entry Form

The primary objectives for designing the data-entry form were to properly identify the monitoring stations and provide necessary inputs for computing accurate vehicle classification and occupancy rates. The necessary information included the station location description, station identification number, beginning and ending hour of observation, time interval of observation by lane, vehicle-count data, and direction of observation.

The vehicle-count data were recorded in 15-min intervals by lane number in the following classification groups:

1. Passenger cars, pickups, and light trucks;
2. Heavy trucks; and
3. Buses.

Each of these three categories was further classified by the number of occupants per vehicle.

The data form was also designed to include information that could influence the resulting classification and occupancy rates. Such items included the weather conditions, lane configuration, date of observation, and field recorder's name. The field data-entry form is illustrated in Figure 2.

Field Data-Collection Supervision

The field supervisor visited each site during the period of observation to ensure adherence to the schedule and proper conformance to the various procedures previously outlined. The survey forms from the previous day's observation were collected and verified. The field supervisor was also responsible for tabulating the field data into hourly summaries as the data sheets were collected. This enabled the supervisor to assure the accuracy and legibility of the data as they were recorded. The count summaries were later used in compiling classification and occupancy rates. If counts appeared to be incorrectly recorded, based on the supervisor's observation of the traffic flow, questionable entries were resolved. The field supervisor maintained a file of the completed forms. At the conclusion of the field data-collection phase, the data were transmitted to the clerical unit for processing.

DATA ANALYSIS

Data Preparation

The data from each station were input into a computer file. The input items included facility type, geographic area, station identification number, direction, number of lanes, date, beginning hour, and hourly classification subtotals.

The field data were expanded to a common 12-h base according to the number of lanes. The factors were used to expand each vehicle classification for every hour of count data. The following ratios were used to compute the expansion factors: 1-lane factor = total possible count period/actual total period counts = 48/36 = 1.33; 2-lane factor = 48/18 = 2.67; 3-lane factor = 48/12 = 4.00; and 4-lane factor = 48/9 = 5.33.

Station Tabulations

A computer program was written to summarize the

Figure 3. Sample printout of Atlanta automobile-occupancy summary data.

ATLANTA AUTO OCCUPANCY SUMMARY

STATION TYPE 2
AREA 3
FACILITY TYPE 2
STATION NO 780 Memorial Drive between Collinwood Drive and Rockbridge Road
DIRECTION 4

TIME	PASSENGER CAR	HEAVY TRUCK	BUS	ALL VEHICLES
700- 800	1.17	1.29	29.63	1.40
800- 900	1.33	1.38	26.78	1.57
900-1000	1.37	1.30	22.32	1.56
1000-1100	1.37	1.25	22.17	1.50
1100-1200	1.42	1.33	18.30	1.58
1200-1300	1.44	1.36	18.02	1.57
1300-1400	1.44	1.32	21.42	1.53
1400-1500	1.37	1.24	36.67	1.55
1500-1600	1.33	1.14	28.04	1.49
1600-1700	1.39	1.17	30.03	1.60
1700-1800	1.35	1.10	21.42	1.42
1800-1900	1.41	1.21	30.03	1.56
MEAN	1.37	1.21	24.90	1.53
STD DEV	.123	.147	5.465	.140
PEAK HOURS (700-900)	1.25	1.35	28.08	1.49

Table 2. Resulting standard deviation for each stratification after computation of average passenger-vehicle occupancy rates.

	Freeway			Nonfreeway		
Functional Classification Area	Samples	Occupancy Rate	Standard Deviation	Samples	Occupancy Rate	Standard Deviation
Inside railroad cordon	8	1.35	0.033	8	1.40	0.108
Railroad cordon to outside I-285	9	1.35	0.046	6	1.41	0.074
Outside I-285 to urban arterial boundary	7	1.36	0.058	9	1.35	0.105
Urban arterial boundary to region boundary	9	1.57	0.122	8	1.44	0.029

vehicle occupancy data by direction, station number, facility type, geographic area, and station type. These tabulations provided hourly occupancy rates in four categories: passenger cars, heavy trucks, buses, and all vehicles. The mean and standard deviation of the occupancy rates were computed for each category of each station. An occupancy rate for the period between 7:00 and 9:00 a.m. was also computed by category to represent the morning peak-hour rates. Figure 3 shows a typical station summary.

Sample Precision

Due to publication constraints, only the data analysis of the stratified, areawide passenger-vehicle occupancy is presented. Before the precision of the sample data could be assessed, the composite standard deviation had to be computed.

The composite standard deviation of the average passenger-vehicle occupancy rates were computed for each stratification by the following formula:

$$S_{oh} = N_h \left[\sum_i (P_{ih} - OCC_h VPV_{ih})^2 \right] / \left(\sum_i VPV_{ih} \right)^2 \tag{9}$$

where

S_{oh} = composite standard deviation of average occupancy rates in stratum h,
N_h = total number of sample locations in stratum h,
P_{ih} = factored number of persons counted at sample location i in stratum h,
OCC_h = estimated average passenger-vehicle occupancy in stratum h, and
VPV_{ih} = factored number of passenger vehicles counted at sample location i in stratum h.

Table 2 illustrates the resulting standard deviation for each stratification. The objective standard deviation (0.065) was exceeded in four of eight strata. The areawide occupancy rate equaled 1.39 persons/vehicle. The actual precision of the estimates was then determined by using Equation 8 and substituting the S_{oh} values calculated above and the W_h values from Table 1. The resulting areawide precision level was 0.013, which was well within the desired areawide precision level of 0.020.

SUMMARY

The major objective of this study was to test the procedures set forth in the guide (1) and to investigate the possible implementation of an annual monitoring program in the Atlanta region. The statistical tests applied to the results indicated a satisfactory degree of confidence. Only a minimal effort ($23 000) was required to collect this data when compared with similar activities in other urban areas of a comparable size. The 1979 survey results will be used to improve precision of future surveys while (it is hoped) reducing costs. Realignment and aggregation of geographic stratifications will allow a reduction in sample-monitoring sites without compromising precision.

Due to the success of this research, an on-going monitoring program is being established to generate vehicle occupancy and classification rates for the metropolitan Atlanta area. Data will be collected on a quarterly basis at selected sites in order to measure seasonal variations in occupancy rates.

These data will be used extensively by local planning units to evaluate the effectiveness of transportation management programs, validate transportation planning models, and monitor general trends in travel characteristics. Local environmental units will use the data to assess transportation-related air quality measures. The measure of vehicle classification and occupancy rates is an important statistic in today's environment, and all indications are that it will become more important in the future.

REFERENCE

1. Peat, Marwick, Mitchell, and Co. Guide for Estimating Urban Vehicle Classification and Occupancy. Federal Highway Administration, Sept. 1978.