Conditions that appear to warrant concern are (*a*) *X*-error approaching the level of *Y*-error combined with moderate degrees of slope or (*b*) lesser degrees of *X*-error combined with greater degrees of slope.

It was also demonstrated by computer simulation that Mandel's method, which might be termed "oblique least squares" because of the manner in which it minimizes the sum of squared residuals, is extremely effective at removing most of the bias introduced by error in the *X*-variable. Figure 3 and Tables 4 and 5 clearly show that, in general, Mandel's method provides substantially more accurate results than ordinary least squares and Figure 4 illustrates this fact with a specific example based on concrete strength tests. The complete theoretical development, along with a more quantitative guideline to determine when it is advisable to use it, is contained in the original source document (*3*). The FORTRAN coding necessary to apply the procedure is contained in the project report (*2*).

REFERENCES

1. J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Regression Models.* Richard D. Irwin, Homewood, Ill., 1983.
2. R. T. Barros and R. M. Weed. *Applied Regression in the Presence of X Error.* FHWA, U.S. Department of Transportation, 1987.
3. J. Mandel. Fitting Straight Lines When Both Variables Are Subject to Error. *Journal of Quality Technology*, Vol. 16, No. 1, 1984, pp. 1–14.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Validation of a Nonautomated Speed Data Collection Methodology

FRED R. HANSCOM

The objective of this research was to develop a validated spot speed study procedure that does not rely on automated equipment. The field study procedure applied a variety of speed collection techniques and compared results against baseline speeds obtained with reliable pavement instrumentation. A recommended manual-timing technique was based on observed accuracies with various vehicle-selection strategies, site conditions, sample sizes, observation period lengths, and observer characteristics.

The conduct of spot speed studies with nonautomated equipment involves a variety of methodological considerations (*1*). Although such studies have long been used in traffic engineering, a number of factors have hampered their valid application (*2*). Among these factors are observer vehicle-selection bias (e.g., the human ability to select a truly random sample), impact of vantage point (e.g., cosine error associated with radar measurement), technique reliability (e.g., stopwatch timing measurement error), and observer human factors (e.g., experience, fatigue).

Transportation Research Corporation, 2710 Ridge Road, Haymarket, Va. 22069.

The objective of this research was to address the effects of the foregoing factors in order to develop a spot speed data collection procedure that does not rely on automated equipment. The field study procedure involved applying a variety of speed collection techniques and comparing results against baseline speeds obtained with reliable pavement instrumentation. A recommended manual timing technique was based on achieved accuracies with various vehicle-selection strategies, site conditions, sample sizes, observation period lengths, and observer characteristics.

VEHICLE-SELECTION STRATEGIES

Basic Application

Specific techniques were evaluated that controlled observer bias in selecting vehicles for speed measurement. Thus, two speed collection methods (radar and manual timing) were applied by using the following vehicle-selection strategies:

1. *Subjective random (all vehicles)*: The observer designates vehicles that appear to be traveling at a speed representative of overall traffic characteristics. Observer instructions are merely to collect a "random, representative" sample of vehicle speeds.

2. *Subjective random (free-flow vehicles)*: The observer designates vehicles that appear to be traveling at a speed representative of overall traffic characteristics and that appear sufficiently isolated in the traffic stream that drivers can select their desired speeds. Observer instructions are to designate "vehicles in which drivers can select their own speeds, unimpeded by other vehicles."

3. *Systematic (Nth vehicle)*: The observer designates vehicle arrivals at some predetermined interval. Example observer instructions are to collect speeds "on every tenth vehicle."

4. *Randomized (vehicle arrival time)*: The observer uses a scientific random time generator to designate times at which the next vehicle arrival will be selected for speed measurement. The applied technique was to program a hand-held computer to wait a random time (e.g., ranging from 5 to 15 sec) and then to instruct the observer to measure the speed of the next vehicle arrival.

5. *Randomized (designated vehicle)*: The observer uses a scientific random procedure (e.g., modified random number table) to designate vehicle arrival. The applied technique was to program a hand-held computer to randomly select a vehicle (e.g., ranging from the first to the fifth vehicle arrival) for speed measurement.

6. *Subjective platoon weighting*: The observer measures speed for the lead vehicle in a platoon and weights this speed by the total number of platooned vehicles. When radar is applied, this method is known as the radar-platoon technique (*3*).

Speed measurement was conducted with the foregoing sampling techniques on roadway sections instrumented with the Traffic Evaluator System (TES) as a source of baseline data against which to establish the reliability of each technique. [TES is a large-scale data acquisition system developed by FHWA. It consists of electronic roadway sensors and recording apparatus designed to retain information on all passing vehicles. Its accuracy has been established in previous research (*4*).] Sufficient samples were obtained to establish statistical confidence of 1 mph or better: sample sizes were approximately 250 vehicles over a period of 2 hr. Both experienced and inexperienced observers applied the techniques.

Average measurement error (i.e., the difference between TES baseline and sample speeds) for each tested vehicle-selection technique is shown in Table 1. The baseline is taken to be the all-vehicle population at the site during each data collection period. (An exception is made for free-flow sampling; the baseline free-flow population in this case comprises only vehicles with headways of 9 sec or greater.)

Close agreement is shown between the traffic baseline speeds and those gathered by each selection technique. No statistical differences were noted for mean, 15th-, or 85th-percentile speeds gathered either by radar or by manual timing. Little difference in accuracy was noted between radar and manual timing. Average all-vehicle error in miles per hour for these two methods for each speed parameter is as follows:

	Mean Speed	*15th Percentile*	*85th Percentile*
Radar	0.4	0.7	0.1
Manual timing	0.4	0.5	0.4

The vehicle-selection strategy for free-flow vehicles proved to be well suited for that specific application. Both manual timing and radar demonstrated very good ability to match all-vehicle free-flow samples, with the following accuracies:

	Mean Speed	*15th Percentile*	*85th Percentile*
Radar	0.4	0.4	0.1
Manual timing	0.1	0.3	0.2

The remaining five strategies were designated to estimate speeds for the all-vehicle population. In order to distinguish among these strategies in terms of accuracy, accuracies are ranked as follows (1 = most accurate to 5 = least accurate):

	Manual Timing		*Radar*	
Strategy	*Mean Speed*	*85th Percentile*	*Mean Speed*	*85th Percentile*
Subjective random	1	3	4	4
Systematic (Nth vehicle)	5	4	5	5
Randomized (vehicle arrival time)	4	2	2	1
Randomized (designated vehicle)	1	1	1	1
Subjective platoon weighting	3	5	2	1

The rankings indicate consistent superiority of the randomized (designated vehicle) strategy, which ranked first (although twice tying with others) as the most accurate to measure both mean and 85th-percentile speeds by using either radar or manual timing. Furthermore, measurement error (average of mean and 85th-percentile measurement differences for both radar and manual timing) indicated the following relative accuracy associated with each technique:

Strategy	*Avg Error (mph)*
Randomized (designated vehicle)	0.2
Randomized (vehicle arrival time)	0.3
Subjective random	0.4
Subjective platoon weighting	0.6
Systematic (Nth vehicle)	0.8

Again, the randomized (designated vehicle) strategy is seen to be slightly superior.

The data conclusively demonstrate that although all tested vehicle-selection strategies produce acceptable (e.g., not statistically different) agreement with baseline traffic speeds, the randomized (designated vehicle) strategy is preferable. The desirability of its use with either radar or manual timing will be

TABLE 1 SPEED MEASUREMENT ERROR ASSOCIATED WITH VARIOUS SAMPLING STRATEGIES, RURAL FREEWAY

	Manual Timing			Radar[a]		
Strategy	Mean Speed	15th Percentile	85th Percentile	Mean Speed	15th Percentile	85th Percentile
Subjective random (all vehicles)	−0.2	−0.5	−0.5	−0.6	−1.1	−0.4
Subjective random (free-flow vehicles)	+0.1	+0.3	−0.2	−0.4	−0.4	−0.1
Systematic (*N*th vehicle)	−0.7	−0.7	−1.1	−0.7	−0.4	−0.9
Randomized (vehicle arrival time)	−0.6	−1.1	+0.3	−0.2	+0.8	−0.2
Randomized (designated vehicle)	−0.2	+0.2	−0.2	−0.1	−1.0	−0.2
Subjective platoon weighting	−0.5	−0.4	−1.3	+0.2	−0.3	−0.2

NOTE: Measurement error is in miles per hour.
[a]Corrected for cosine error.

addressed in a subsequent section dealing with applied vehicle-selection strategies in varied highway settings.

Lane Specificity

One variation to the applied vehicle-selection strategies just discussed was to designate vehicles by lane for speed measurement. The underlying rationale for this procedure was an attempt to account for the fact that vehicles in the right-hand lane tend to travel more slowly than do vehicles in the left lane. The applied procedure involved making a lane-specific volume count immediately before commencing speed observation. Thus, a hand-held computer was programmed to randomly select vehicles by lane: the proportion of selected vehicles in each lane was based on observed lane occupancy. This lane specificity selection option was applied to both randomized (vehicle arrival time) and randomized (designated vehicle) strategies. A comparison of measurement error (baseline all-vehicle versus selected sample) is as follows:

	Lane Specific		*Not Lane Specific*	
Strategy	*Mean Speed*	*85th Percentile*	*Mean Speed*	*85th Percentile*
Vehicle arrival time	−0.3	0.3	−0.6	+0.3
Designated vehicle	−0.4	0	−0.2	−0.2

A slight overall improvement in accuracy was found with lane-specific vehicle selection for both tested strategies. Average 85th-percentile speed error was 0.15 mph (versus 0.25 mph) and average mean speed error was 0.35 mph (versus 0.40 mph) when the lane specificity option was applied. However, this improvement is so slight (and not statistically significant) that it, in and of itself, cannot constitute a basis for using lane-specific vehicle selection on an operational basis.

A recommendation regarding lane-specific vehicle selection (as opposed to random arrivals regardless of lane presence) must consider the operational application of this procedure and its trade-offs against the potential gain in accuracy. For this reason, application of lane-specific selection cannot be justified in view of the insignificant demonstrated increase in accuracy. The following operational considerations provide the basis for this recommendation.

First, the lane-specific vehicle-selection procedure is time consuming and cumbersome to initiate in the field. A volume count must first be gathered and entered into the hand-held computer. Further, operation of the hand-held computer would be encumbered by the more complex procedure and programming required to accommodate situations of varied lane number. Second, and more important, data collection with the lane-specific selection option is much more time consuming and thus reduces the overall data collection efficiency. This is especially true under low to moderate volume conditions where long intervals exist between vehicle arrivals in the left lane. Greater statistical accuracy can be expected because of the larger sample obtained, within a given time frame, when a straightforward random arrival selection technique is applied.

VARIED SITE CONDITIONS

Limited validation of speed collection techniques was conducted across a variety of site conditions. The purpose of this activity was to examine the possible effect of differing highway conditions (e.g., available vantage point) on speed observation results. Speeds were collected by using radar and manual timing at four site types: urban four lane, urban two lane, rural Interstate, and rural two lane. The vehicle-selection strategy applied at each site was subjective platoon weighting.

Measurement error (difference between platoon-weighted sample and all-vehicle population and difference between lead vehicle sample and free-flow vehicle population) obtained at each site type with both radar and manual timing is shown in Table 2. (Recall that subjective platoon weighting involves measuring lead vehicle speed and weighting this value by the number of vehicles in the platoon.)

Results shown in Table 2 indicate reasonably small measurement error for all-vehicle speed estimation when the platoon-weighting technique is applied with radar as the speed collection method. This average error, across sites, is 0.5 mph for mean speed and 0.7 mph for 85th-percentile speed: a statistical match with baseline speed was achieved under all site conditions. Another tested application of speed collection techniques across sites was to estimate free-flow speed parameters based on the lead vehicle sample used for platoon weighting. Again, when radar was applied as the speed collection method, the technique was seen to work fairly well. (Across sites, average

TABLE 2 SPEED MEASUREMENT ERROR ASSOCIATED WITH SUBJECTIVE PLATOON WEIGHTING AT FOUR SITE TYPES

Sample	Urban, Four Lanes		Urban, Two Lanes		Rural Interstate		Rural, Two Lanes	
	Mean Speed	85th Percentile	Mean Speed	85th Percentile	Mean Speed	85th Percentile	Mean Speed	85th Percentile
Platoon weighted								
Radar	−0.9	−1.0	−0.7	−1.1	+0.2	−0.2	+0.2	−0.6
Manual timing	−2.5[a]	−2.9[a]	−0.8	−1.7[a]	−0.5	−1.3	−3.8[a]	−4.0[a]
Lead vehicle								
Radar	−0.5	−1.4	0.3	−1.7[a]	−0.4	−0.9	0	−1.3
Manual timing	−1.9[a]	−2.9[a]	−0.6	−1.9[a]	−1.1	−1.4	−3.8[a]	−4.9[a]

NOTE: Measurement error is in miles per hour.
[a]Statistically significant ($\alpha = .05$).

errors were 0.3 mph for mean speed and 1.3 mph for 85th-percentile speed.) A statistically different 85th-percentile speed (1.7 mph error) at the urban two-lane site was likely due to small sample size (e.g., 137 lead vehicles).

Manual timing as a speed collection method was not shown to be reliable under all tested conditions. Statistical differences were evident at three of the four sites that did not have elevated observer vantage points. However, somewhat promising results (e.g., mean speed error of 0.7 mph) were found at the urban two-lane site, where the observer was standing at street level. As noted earlier, the significant (1.9 mph) error in 85th-percentile speed may be due to sampling conditions. A detailed evaluation (i.e., a vehicle-by-vehicle error determination) was not possible to fully assess the maintenance of manual timing speed measurement accuracy under this condition.

In summary, radar and manual speed timing methods using the platoon-weighting technique were applied at four site types: urban two- and four-lane highways, rural Interstate, and rural two-lane highways. Radar platoon weighting demonstrated good results across site conditions. All-vehicle population speeds were estimated by using this technique with the following average accuracies: mean speed, 0.5 mph; 85th-percentile speed, 0.7 mph. Radar sampling of lead vehicles was shown to estimate mean free-flow speeds with an accuracy of 0.3 mph. Manual timing was not shown to be reliable at sites without elevated observer vantage points.

RELIABILITY OF MANUAL SPEED TIMING

Stopwatch timing is a frequently applied manual method of speed measurement. In order to examine the accuracy of this technique, the following studies were conducted: (*a*) vehicle-by-vehicle comparison of manually timed speeds with those obtained from a commercial speed-monitoring device, (*b*) intercoder reliability study comparing between-observer results, and (*c*) minimum required observation period.

Vehicle-by-Vehicle Comparisons

Manually timed speeds from two coders were compared on a vehicle-specific basis with results obtained from a commercial automated speed-monitoring device. The applied manual technique used a hand-held computer configured as an electronic stopwatch to time vehicles between two pavement markers spaced 300 ft apart. One coder was relatively inexperienced (but highly motivated), with approximately 2 days of previous speed data collection experience; the second was the project principal investigator, who had considerable speed data collection experience.

The following speed parameters were compared on a vehicle-by-vehicle basis for the automated device (pavement loop) and manual collection (stopwatch timing): mean speed, 15th percentile, and 85th percentile. Average difference (i.e., error between techniques) for each observer is shown (in miles per hour) in Table 3. A minus sign indicates that speed obtained manually was slower than that obtained with the automated device. Sufficient sample sizes were obtained in each trial to establish mean speed confidence (at the 0.01 level) within 1.0 mph. Sample sizes ranged from 215 to 241 vehicles per trial.

Error associated with individual vehicle speed measurements was also examined. Results obtained for each observer are summarized in Table 4. Although individual vehicle measurement errors were shown to be surprisingly large (e.g., approximately 40 percent exceeded 1 mph), the errors were shown to be largely compensating in nature as evident from corresponding mean speed differences between techniques, which ranged from 0.1 to 0.9 mph. The resulting assessment of the manual speed-timing procedure is that the method produced mean data

TABLE 3 AVERAGE DIFFERENCE BETWEEN SPEEDS TIMED MANUALLY AND BY AUTOMATED DEVICE

	Mean Speed		15th Percentile		85th Percentile	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
Inexperienced	−0.1	+0.1	+0.6	+0.1	−1.0	−0.1
Experienced	−0.9	−0.3	−1.2	+0.3	−0.5	−0.2

TABLE 4 ERROR BY EXPERIENCE OF CODER

	Avg Error[a]	Maximum Error	Percent Error by Speed (mph)				
			>5	>4	>3	>2	>1
Inexperienced	1.1 ± 0.14	11.4	1.6	2.6	9.3	19.5	46.9
Experienced	1.4 ± 0.22	12.8	1.1	1.7	6.6	11.8	37.9

[a]0.05 confidence interval.

statistically equivalent to speeds gathered with the automated speed-monitoring device. However, as noted in the previous section, manual speed timing can be considered reliable only in those highway settings that provide an elevated vantage point.

Intercoder Reliability

To determine intercoder reliability, two observers manually timed speeds for specifically selected vehicles. Matched vehicle-by-vehicle speed data were then analyzed to examine agreement between coders. The operational application of the intercoder reliability study is its use as a training aid.

In this experiment, two intercoder reliability studies were undertaken. The first, conducted by the research team, consisted of a traffic engineer with considerable speed collection experience and an assistant with only one day of previous experience. The second study involved two FHWA employees, both of whom were familiar with speed collection procedures. In each study differences (in miles per hour) between observers were found not to be significant (Table 5). Error between coders in each of the studies is shown in Table 6.

Despite the relatively large magnitude of individual measurement differences, results indicate close overall between-coder agreement (approximately 0.5 mph mean speed difference in both tests). The larger sample obtained in the TRC study resulted in additional opportunity for larger individual speed measurement differences (thus explaining the maximum measurement error of 18.3 mph). However, as shown in the previous section, manual timing speed errors were seen to be compensating (i.e., approximately equal in both positive and negative directions). This is further substantiated by the fact that two-thirds of between-coder speed measurements differed by more than 1.0 mph, yet averaged speeds differed by only 0.5 mph.

These tests provide results of a procedure to assess speed measurement ability between observers of varying skill levels. Two precautions must be noted. First, observers are aware that results are being monitored and may therefore perform with more vigilance. Second, manual speed timing must be conducted from an elevated vantage point. Nevertheless, these intercoder reliability studies demonstrate comparable between-observer results for manual speed timing.

TABLE 5 MEASUREMENT DIFFERENCE BY CODING TEAM

	Observer	Sample Size	Mean Speed	Standard Deviation	85th Percentile
TRC	1	200	57.5	4.80	61.9
	2	200	57.0	4.49	61.3
FHWA	1	40	58.6	3.76	62.7
	2	40	59.1	3.53	63.0

NOTE: TRC = Transportation Research Corporation.

TABLE 6 ERROR BY CODING TEAM

	Avg Error[a]	Maximum Error	Percent Error by Speed (mph)	
			>5	>1
TRC	0.49 ± 0.78	18.3	28	66
FHWA	0.54 ± 0.64	6.5	20	68

[a]0.05 confidence interval.

Observation Period

In order to assess the suitability of manual timing to estimate speeds in a one-time study, comparisons were made between sampled speeds and the all-vehicle population for a variety of observation conditions (e.g., length of observation period, time of day, and previous experience). Period durations (10, 20, and 45 min) were randomly ordered throughout each of the two data collection days. In addition, systematic scheduling ensured that both long periods (45 min) and short periods occurred both early and late on different days as a check on coder fatigue. Ten-minute rest breaks were taken between each data collection period, and a 1-hr lunch break was taken at midday. Two observers, one experienced and one inexperienced, participated in this experiment.

Speed measurement accuracy was determined by comparison of manually timed speeds for each observer with an all-vehicle baseline consisting of TES data for each collection period. Summary results contrasting mean speed error (difference in miles per hour between TES data and manual timing speeds) for the experienced and inexperienced observers are as follows:

Observation Period (min)	*Observer*	
	Experienced	*Inexperienced*
10	0.5	2.6
20	0.9	1.3
45	0.4	0.7

Data collected by the inexperienced coder (who regrettably exhibited a lackluster motivation) demonstrated a distinct error effect associated with period duration. The results from the 10-min observation period for the experienced coder indicated surprisingly close agreement between manually coded and TES

TABLE 7 RESULTS FROM 45-MIN OBSERVATION PERIOD

Trial No. and Data Source	Sample Size	Mean Speed	SD	15th Percentile	85th Percentile	95th Percentile	Period of Day	Result
1								
Coder	94	57.4	4.0	54	61	63	1	No statistical differences
TES	171	58.0	4.3	54	62	64		
2								
Coder	81	57.9	5.5	52	63	65	2	Coder mean speed low by 1.5 mph
TES	144	59.4	4.8	54	64	66		
3								
Coder	112	58.8	6.2	54	63	67	4	Coder variance high
TES	208	58.8	4.5	55	63	66		
4								
Coder	103	57.7	4.4	53	62	65	5	Amazing
TES	166	57.5	4.5	53	62	65		
5								
Coder	88	56.3	5.3	53	61	63	5	No statistical differences
TES	168	57.0	4.9	51	62	63		
6								
Coder	91	56.7	6.3	52	62	65	8	No statistical differences
TES	179	56 7	5.7	52	62	64		
7								
Coder	144	58.1	4.7	54	62	64	8	No statistical differences
TES	440	58.1	4.1	54	62	64		
8								
Coder	115	57.7	5.4	53	62	66	10	Coder variance high
TES	323	57.7	4.2	53	61	65		

(all-vehicle) speeds. During each 10-min period, the coder measured speeds on samples ranging in size from 20 to 37 vehicles. This sample represented between 38 and 69 percent of the total vehicle population measured by TES data. Subsequent 20- and 45-min periods resulted in similar sampling percentages. The 20-min period data resulted in a lesser degree of mean speed accuracy. However, the results from the 20-min period showed improved agreement in measured speed variance (no statistical difference).

As expected, closer overall agreement was obtained between TES and coder speeds (both means and all selected percentiles) during the 45-min observation periods. Examination of results from eight individual periods (Table 7) indicates that although statistical differences were found during three trials, a minimal effect was realized in terms of measurement error magnitude. The single incidence of significantly different mean speed was 1.5 mph. The average mean speed error was 0.36 mph. An examination of the raw data indicated that the mean 85th-percentile speed was in error by only 0.45 mph.

The impact of observer fatigue was approached by using observation period duration as a surrogate. The appropriateness of this surrogate lies in the fact that tested conditions represent time requirements to gather statistically suitable samples. With this approach, the effect of fatigue was examined by two procedures. First, within-period fatigue was examined for the data from the 45-min period; yet no degradation in accuracy was found for speed measurements obtained late in any specific period. Second, mean speed error (all-vehicle versus sample difference) demonstrated a trend for less error to occur later in the day. Ranked period-specific mean measurement errors associated with time of day are as follows (Period 1 begins at 9:00 a.m.; Period 10 ends at 5:00 p.m.):

Error (mph)	*Period*
0	8
0.1	9
0.3	6
0.4	10, 1
0.8	5
0.8	3
0.9	7
1.1	4
1.3	2

Results of this experiment indicated that although 45 min is the minimum acceptable period duration, specifying period duration alone does not ensure an adequate sampling requirement. Both observation duration and sample size must be specified. Therefore, sample-size effects were studied next.

DETERMINATION OF MINIMUM SAMPLE SIZE

In this experiment the suitability of small spot speed samples to estimate all-vehicle speed populations was investigated. The objective was to determine minimum sample requirements in order to optimize manpower and financial resources without sacrificing statistical integrity of the study.

The applied procedure involved comparing results obtained with varied sample sizes versus results from the all-vehicle population. Two days of speed observation were applied at a rural Interstate site during hours of uniform traffic flow. Subsamples consisting of 10, 20, 50, 100, and 200 vehicles were randomly selected from the all-vehicle population. Five iterations (random selections) were conducted in each sample size category. Samples were extracted from specific durations (e.g.,

TABLE 8 AVERAGE AND MAXIMUM SPEED MEASUREMENT ERROR FOR VARYING SAMPLE SIZES

Sample Size	Mean Speed		85th Percentile	
	Day 1	Day 2	Day 1	Day 2
Average Error				
10	0.8	0.5	1.0	0.8
20	0.2	0.3	0.4	0.6
100	0.1	0.1	0.1	0.2
200	0.1	0.1	0.1	0.2
Worst-Case Error				
10	3.0	2.8	5.1	5.4
20	1.3	3.1	3.0	2.9
50	0.9	1.1	1.3	2.3
100	0.8	0.7	0.6	0.7
200	0.5	0.6	1.0	0.9

NOTE: $N = 50$ observation periods for each sample size.

a half-hour) in the database so as to represent operational data-gathering periods. A total of 50 observation trials were made for each tested sample size.

No statistical differences ($\alpha = .05$) were found between samples and population mean speeds. In certain instances, standard deviations differed for samples of 10, 20, and 50 vehicles. Speed measurement error (i.e., all-vehicle populations versus sample groups) is summarized in Table 8. Average mean and 85th-percentile speed differences are shown in the top portion of the table. These averages represent magnitude of error without regard to direction (i.e., a +1.0-mph error and a −1.0-mph error would average to 1.0 mph). The results in the upper portion of the table imply that very good results were obtained with relatively small sample sizes. That is, average precision of better than 1.0 mph was achieved with sample sizes as small as 20 vehicles.

However, in order to examine the maximum sampling error likely to be associated with each sample size, the worst-case difference from all 50 trials within each size category is shown in the lower portion of the table. These results indicate that sample sizes of 10 to 50 vehicles can result in mean or 85th-percentile speed sampling errors ranging from 0.9 to 5.4 mph. However, a sharp reduction in error was noted for 85th-percentile speeds as sample size increased from 50 to 100 vehicles. A further increase to 200 vehicles did not yield any real benefit. Thus, maximum expected measurement error associated with a random sample of 100 vehicles was shown to be 0.75 mph for mean speed and 0.65 mph for 85th-percentile speed. Results indicate that under uniform flow conditions (e.g., during non-rush periods), mean and 85th-percentile speeds can be measured with an accuracy of better than 1.0 mph if two sampling minimums are met: a 45-min observation period (as seen from the previous section) and a sample of 100 vehicles.

EFFECT OF OBSERVER EXPERIENCE

Emphasis in this research was placed on the relative accuracy achieved with tested techniques used by experienced versus inexperienced observers. Each of the foregoing vehicle-selection strategies was applied by both an experienced observer (i.e., a traffic engineer with 14 years' experience) and an inexperienced observer (i.e., part-time personnel with short training session) using both radar and manual timing techniques. Four inexperienced observers were used; the same experienced observer conducted data collection for each tested technique as a basis for comparison.

No significant effect in radar application was noted as a function of observer experience. Table 9 gives the manual speed timing accuracies associated with the inexperienced observers. Two speed measurement criteria were applied: (*a*) difference between speed results coded by inexperienced observers and actual speeds of the vehicle population, and (*b*) difference between results of the experienced versus the inexperienced coders. Relative error is shown for each inexperienced coder, averaged across all trials. Three women and one man were used as the inexperienced observers; ages ranged from 25 to 39. Measurement differences for each observer are ranked in the table according to age. Results generally indicate that improved accuracy is associated with greater motivation, as was shown by the results for the two older observers. Mean speed measurement errors recorded by the inexperienced coders ranged from 0.1 to 2.0 mph; differences between the experienced and inexperienced observers ranged from 0 (exact agreement) to 1.4 mph. In one case, however, the inexperienced coder produced more accurate results than did the experienced coder.

TABLE 9 COMPARATIVE MANUAL-TIMING SPEED MEASUREMENT ACCURACIES FOR INEXPERIENCED OBSERVERS

Coder and Age	Inexperienced Coder Versus Actual		Experienced Versus Inexperienced Coder	
	Mean Speed	85th Percentile	Mean Speed	85th Percentile
Eleanor, 25	2.0	1.5	1.4	1.0
Barbara, 27	1.5	1.9	0.9	0.5
Dave, 35	0.1	0.6	−0.5[a]	0.3
Carol, 39	0.3	0.5	0	0.4

[a]Minus sign indicates superior performance by comparison with experienced coder.

The interpretation of these results leads to the following conclusion regarding observer experience and its effect on manual speed timing accuracy. Although generally improved results were associated with age (i.e., observers in their thirties demonstrated improved results in comparison with those in their twenties), no consistent difference was noted between male and female coders. The field experience during this research demonstrated two important factors. First, motivation is more significant than specific observer characteristics in determining suitability for this task. Those personnel who demonstrated a serious attitude and who appeared to genuinely want to do the work proved to be more accurate in their results.

Second, intercoder reliability trials (i.e., vehicle-by-vehicle data comparisons between observers) are essential in order to predetermine the suitability of any employee to conduct manual speed timing. In the case of a motivated observer, one 2-hr training session is likely to be sufficient to provide needed experience. A second training session, conducted on a different day, is recommended to control for within-observer performance variation.

SUMMARY OF RESULTS

A field evaluation of speed data collection techniques was conducted by comparing actual traffic speed characteristics with those measured with the following procedural variations. Six vehicle-selection strategies were tested in order to eliminate observer bias. These strategies included subjective, systematic, and random vehicle-selection procedures. Lane-specific vehicle selection was also tested. The reliability of two methods (radar and manual timing) using the platoon-weighting technique was assessed on four highway types: rural Interstate, rural two lane, urban four lane, and urban two lane. The effect of observer experience (age and practice) was examined. Relative precision for spot speed measurement was determined for a variety of observation period effects (e.g., duration, time of day, and within-period observer fatigue). Spot speed sampling accuracies were determined for minimum cost-effective sample sizes.

Results of the series of field studies are as follows:

1. Six vehicle-selection strategies were tested in order to eliminate observer bias: subjective, systematic, computer-assisted random, and platoon-weighted procedures using both radar and manual timing methods. Lane-specific vehicle selection was also tested but was determined not to be beneficial. All the tested strategies yielded results that were statistically equivalent to real traffic speeds. However, the randomized (designated vehicle, not lane specific) strategy consistently proved best and resulted in mean and 85th-percentile speed error of 0.2 mph or less.
2. The reliability of two methods (radar and manual timing) using the platoon-weighting technique was assessed on four highway types: rural Interstate, rural two lane, urban four lane, and urban two lane. Radar produced the following accuracies: mean speed, 0.5 mph, and 85th-percentile speed, 0.7 mph. However, manual timing was not shown to be reliable in highway settings that do not afford an elevated vantage point.
3. The accuracy of manually timed speed observation was determined from vehicle-by-vehicle comparisons with an automated speed collection device. Despite considerable vehicle-specific error (i.e., approximately 40 percent of the measurements were in error by 1.0 mph or more), these errors were largely compensating in nature. Averaged trials for two observers resulted in sample means and 85th-percentile speeds within 0.5 mph accuracy.
4. Relative precision for spot speed measurement was determined for a variety of observation period effects (e.g., duration, time of day, and within-period observer fatigue). Results showed that a minimum 45-min observation period is required but that no accuracy degradation due to fatigue during this time is expected. Rest periods throughout the day resulted in no manual speed timing accuracy reduction at the end of an 8-hr day.
5. Spot speed sampling accuracies were determined for minimum cost-effective sample sizes. A minimum of 100 vehicles is required for mean speed accuracy of 0.5 mph and 85th-percentile speed accuracy of 1.0 mph.

The product of this series of field studies is a recommended manual technique for speed determination. Manual observation is suggested for applications such as assessment of traffic control device effectiveness and other uses where continuous speed monitoring with automated equipment is not feasible. Application of manual procedures developed in this series of field experiments was determined to yield mean speeds accurate to 0.5 mph at the 0.01 confidence level.

The recommended manual speed collection method consists of the following procedure:

1. Speed-timing personnel should be trained with at least two intercoder reliability trials (on separate days), requiring mean agreement between coders of 0.5 sec or better for individual speed measurements.
2. Speeds should be clocked by using an electronic stopwatch capable of measuring and displaying time to an accuracy of 0.01 sec.
3. Overhead observation points, such as overpasses, should be used.
4. Speed-timing markings should be painted on the pavement at a minimum spacing of 270 ft.
5. A minimum observation period of 45 min and total sample size of 100 vehicles should be used.
6. Observations should be conducted at times of day exhibiting stable speed conditions (e.g., only rush or only nonrush conditions).

ACKNOWLEDGMENT

The work reported in this paper was conducted under a contract with FHWA. The contract monitor was Davey Warren.

REFERENCES

1. F. R. Hanscom. *An Assessment of Speed Measurement Techniques and Practices: A Literature Review.* FHWA, U.S. Department of Transportation, 1983.
2. W. J. Homburger and J. H. Kell. *Fundamentals of Traffic Engineering.* Institute of Transportation Studies, University of California, Berkeley, 1977.
3. M. Van Aerde and S. Yagar. Radar-Platoon Technique for Recording Speeds and Volumes. In *The Effects of Speed Limits on Traffic Accidents and Transport Energy Use,* Organization for Economic Cooperation and Development, Paris, 1981.
4. G. S. Kolsrud. *Diagrammatic Guide Signs for Use on Controlled Access Highways,* Vol. 3. Report FHWA-RD-73-26. FHWA, U.S. Department of Transportation, 1972.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.