WATER RESOURCES RESEARCH, VOL. 29, NO. 4, PAGES 1247-1256, APRIL 1993

A Comparison of Selection at List Time and Time-Stratified Sampling for Estimating Suspended Sediment Loads

ROBERT B. THOMAS[1] AND JACK LEWIS

Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Arcata, California

Time-stratified sampling of sediment for estimating suspended load is introduced and compared to selection at list time (SALT) sampling. Both methods provide unbiased estimates of load and variance. The magnitude of the variance of the two methods is compared using five storm populations of suspended sediment flux derived from turbidity data. Under like conditions, the SALT coefficient of variation was 1.4-7.7 times that of time-stratified sampling. Time-stratified sampling performs well if the range of sediment flux in each stratum is small. This requirement can be met by using small sample sizes in many short strata. Theoretically, SALT sampling has the potential for smaller sampling variance; however, it is difficult to select an auxiliary variable that predicts flux well under diverse flow conditions. An "optimum" auxiliary variable formed from the largest storm performed about as well as time-stratified sampling for the larger storms. Time-stratified sampling ensures that specimens are collected in all storms, facilitating load estimation for individual storms. In contrast, SALT can better allocate sampling resources over different size storms, enabling efficient estimation of the total load for longer periods. Because time-stratified sampling is less sensitive to the way measurements are allocated to different parts of the population, it is preferred for estimating storm loads of multiple constituents from the same sample.

INTRODUCTION

The process used to sample data, the methodology employed to make estimates from the data, and the interaction between these factors govern the quality of estimates of any physical quantity. Although practice suggests that this fact is not generally realized, it holds true in particular when estimating suspended sediment loads in rivers.

Although only one sample is usually selected from a given suspended sediment population, an estimating process (i.e., the sampling method and the procedure used to calculate estimates) can be evaluated only by the properties of the distribution of estimates over all possible samples. Estimates of any quantity should have at least the following three basic properties:

1. There should be no (or negligible) systematic difference between the expected value of the distribution of estimates and the true underlying parameter being estimated (i.e., the load in the case of suspended sediment). Estimators with this property are said to be unbiased.
2. A valid estimate of sample variance of the distribution of load estimates should be available from each sample so the quality of the estimate can be assessed from the only sample that will normally be collected. ("Sample" here refers to a set of concentration measurements used to make an estimate of load; a single measurement will be called a "specimen.")
3. The variance of the distribution of sample estimates (i.e., the sample variance) should be "small" to reduce error in estimating load. Variance can often be reduced by using knowledge of population structure, especially when sampling episodic populations such as suspended sediment loads.

Statistical estimation involving some method of random sampling is required in order to satisfy these three requirements. Estimators not based on random samples (i.e., nonstatistical estimators) have indeterminable bias and lack valid variance estimators. Whether due to professional inertia or cost constraints, most methods for sampling and estimating load in rivers have been nonstatistical. Problems with nonstatistical estimators have been amply demonstrated [*Cochran et al*., 1954; *Thomas*, 1988, 1991]. It is a misconception that statistical schemes cannot adapt to variations in constituent load; there are methods of restricting randomization to enhance sampling during higher discharges that still retain the properties listed above.

The selection at list time (SALT) sampling method [*Thomas*, 1985] uses an "auxiliary variable" related to suspended sediment flux to increase sampling density at higher flows. SALT applications use programmable data loggers that combine random numbers with an auxiliary variable based on river stage to make real-time "decisions" about when to operate pumping samplers. Auxiliary variables are typically based on sediment rating curves or "average sampling rates" specified by the operator [*Thomas*, 1989).

Another statistical scheme divides the hydrograph into different length time periods, each of which is sampled randomly and independently. If the periods are selected so that no period has an excessive range in suspended sediment discharge, estimates often have lower variance than those from SALT. This method is called "time-stratified" sampling and can also be applied automaticallyusing a programmable data logger and a pumping sampler [*Eads*, 1991]. Suspended sediment load for a monitored period is obtained by adding estimated loads for the component strata. Similarly, variance estimates for the entire period are calculated by adding the (independent) estimates of stratum variance.

Both time-stratified and SALT sampling are unbiased, have their own valid estimator of sampling variance, and use stage information to reduce variance by raising sampling densityduring periods of high and variable flows.The magnitude of their variance can differ,however, so five populations

[1]Retired

Paper number 92WR02711.

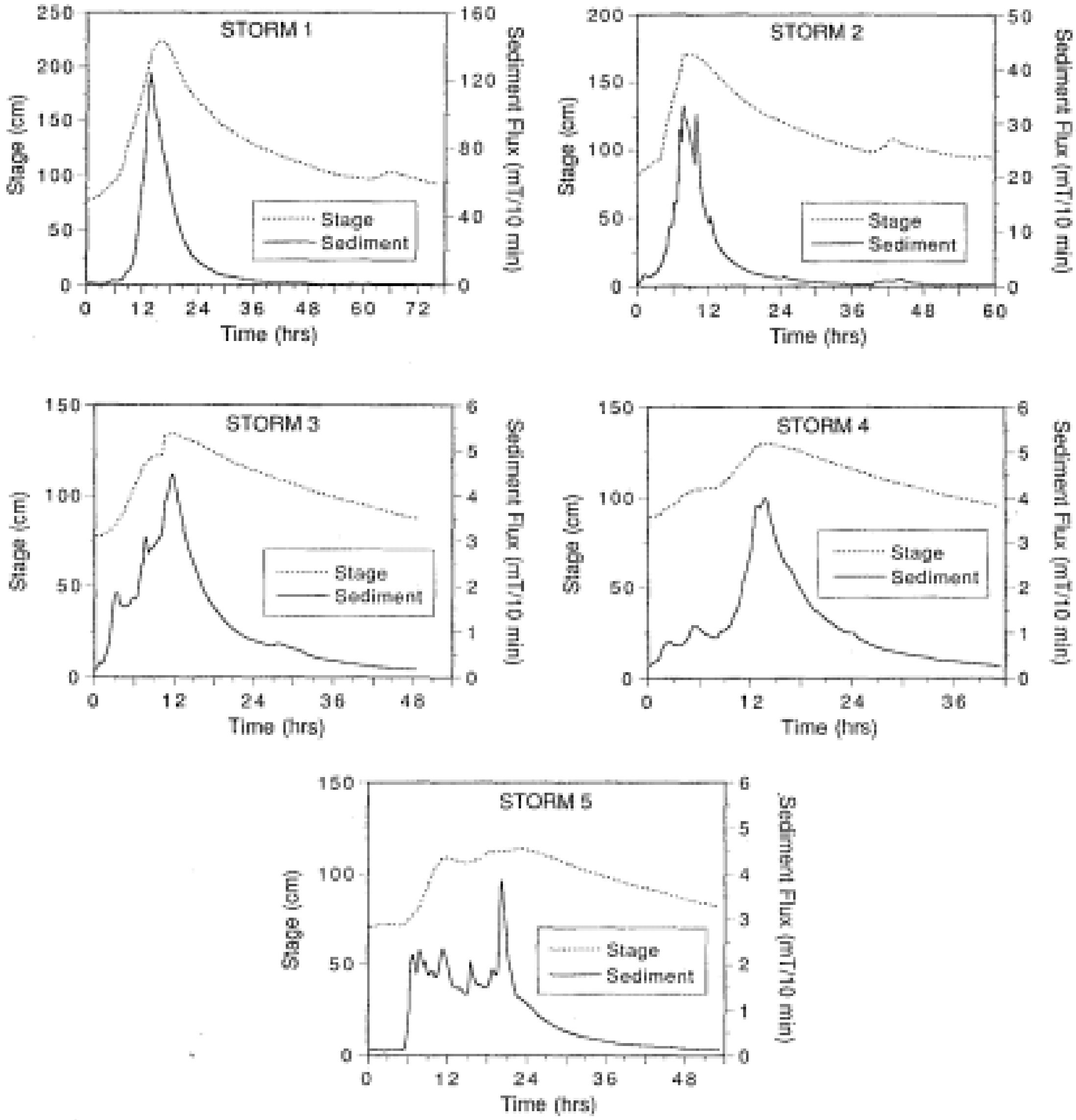

Fig. 1. Stage and suspended sediment flux over time for five storms during February 1983 in the North Fork of the Mad River near Korbel in northern California. The storms are in order of decreasing suspended sediment flux with 5505, 1403, 347, 284, and 265 metric tons, respectively.

of suspended sediment were derived from turbidity data and used to compare the performance of the two schemes.

METHODS

Study Data

Data were collected for five storms over a range of sizes in February and March 1983 from the North Fork of the Mad River in northern California. Storms were defined using the method of *Hewlett and Hibbert* [1967]. Major pre-storm troughs were identified before the primary peaks on the discharge hydrograph and straight lines with slopes of 5.5 x 10^{-6} m^3 s^{-1} ha^{-1} h^{-1} were drawn to the right of each trough. Intersections of these lines with the recession limbs of the hydrographs defined the ends of the storms.

Water discharge and turbidity was recorded continuously on charts and pumped specimens of river water were collected to analyze for suspended sediment concentration. The recorded turbidity varied rapidly over time producing a solid "band" of ink on the chart that widened during periods of higher sediment flux. Because there was no well-defined trace that demarcated the turbidity, we assumed this variation was due to artifacts of the turbidity instrumentation and pumping mechanisms, and we derived turbidities as midpoints between smoothed lines drawn at the upper and lower borders of this band. If this assumption were incorrect our procedure would have produced populations having too low a variance and the sampling variances that we present for both the time-stratified and SALT procedure would thereby be underestimated by an unknown amount. However, the procedure is supported by 72 measurements of sediment concentration recently made at 10-min intervals during a storm at North Fork Caspar Creek in northwestern California. The sedigraph of these data was similar in appearance to those shown in Figure 1, showing very little high-frequency fluctuation.

A regression relationship of In concentration on ln turbidity (R^2 = 0.988 with 52 observations) was used to estimate concentration at any time. Each storm consists of a finite set of 10-min periods with midpoint estimates of concentration based on turbidity. The period length was chosen arbitrarily, but 10 min is probably a short enough period to be represented by a single concentration value. Sediment load for each 10-min period is the product of the midperiod discharge, the concentration, and a constant to account for units and period length. Although these estimated period loads may not exactly match the "true" population units, they represent realistic populations useful for sampling comparisons. Accordingly, we shall refer to these as known populations.

Values of stage and sediment flux for each storm are shown on the same plots to indicate interactions of the two quantities (Figure 1). Although sediment flux is not perfectly related to stage (or discharge), most flux occurs during and shortly after the rise of the hydrograph to its peak. The 10-min loads were summed over the storm periods to determine population totals for each of the five storms. The suspended sediment loads for storms 1 to 5 were 5505, 1403, 347, 284, and 265 metric tons, respectively.

Time-Stratified Sampling

Time-stratified sampling was introduced as a statistical method of estimating suspended sediment load without requiring an intelligent controller. The method enables a hydrologist to collect valid statistical samples and reduce sampling variance by anticipating patterns of water and sediment discharge. The duration of a time period, or "stratum," is chosen over which the supply of pumping sampler bottles is to be filled: shorter strata for expected high or rapidly increasing discharges, and longer ones for lower flows, especially during hydrograph recession. A simple random sample of times is chosen within the stratum and the pumping sampler set to operate at those times. The station is visited at the end of each stratum to service the sampler, choose the next stratum duration, and enter a new set of random times.

The time-stratified procedure gives unbiased estimates of total load and variance using standard formulas (see the appendix). The quality of an estimate (i.e, the magnitude of its variance) depends on how well the hydrologist forecasts changes in the pattern of sediment flux, thereby selecting appropriate stratum lengths. By regulating sampling density over time, therefore, the hydrologist gains some control over the quality of the estimates without compromising their validity.

Because short strata tend to have lower variance, sampling efficiency can be dramatically raised by selecting several strata within the complement of sampler bottles. Suppose 24 specimens are to be taken in 12 hours. Rather than taking a random sample of 24 in the 12 hours, variance can be reduced by dividing the period into four 3-hour strata and choosing six random times within each one. Carrying this process to its logical conclusion, two bottles can be taken in each stratum, the minimum required to calculate variance. Compromising at three bottles per stratum gives some protection against lost sediments. Using shorter strata also enforces 'a certain level of coverage of the population, ensuring that all parts of the population have some representation in the sample.

Time-stratified sampling can also be done with the same instrumentation used for SALT sampling; only the computer program need be different [*Eads*, 1991]. At the beginning of a stratum the computer determines the stage direction and level and sets a stratum duration and sample size based on stored information. (To detect a reversal in stage condition, we required the stage to drop 14 mm below a peak or rise 14 mm above a trough). It then selects the proper number of random times, sorts them, and stores them in memory. The computer operates the pumping sampler at the selected times, stores required information, and repeats the process at the end of each stratum. The major benefit of automatic application of time-stratified sampling is to reduce variance by improving response to changing conditions.

A sampling "schedule" is required to enable a field algorithm to "decide" on appropriate stratum lengths as stream behavior changes. The schedule embodies knowledge about the relationship of stage level and condition to rate of change in sediment flux so that the data can be efficiently distributed over the period to be monitored. The schedule and the use of short strata combine to make time-stratified sampling an efficient procedure for estimating suspended sediment loads.

Eighteen sampling schedules were chosen for comparisons in this paper (Table 1). Each schedule consists of a set of stratum lengths providing a range of sampling intensities for four stage classes and rising and falling stage conditions. The number of schedules was doubled to 36 by using the same set of stratum lengths with stratum sample sizes of two and three. All the schedules tend to sample more heavily on rising limbs and at higher stages. These schedules are not a random set of those possible, but seemed to us a fair representation of those likely to be chosen by hydrologists familiar with patterns of suspended sediment flux. Variation for comparing the SALT and time-stratified schemes derives from these 36 sampling schedules and the five storms.

One application of the time-stratified sampling process is illustrated for storm 2 and schedule 13 with three specimens per stratum (Figure 2, Table 2). This illustrates the way in which time-stratified sampling confines measurements to specific periods of differing lengths throughout the storm. Except for stratum 3, the sample stratum totals matched the true totals fairly well for this sample. The difference between the estimated and true storm totals was due almost entirely to stratum 3. The total for stratum 3 was underestimated because two specimens very close to the minimum were selected. Similarly, the variance was overestimated in stratum 4 and underestimated in stratum 5, because the selected specimens happened to include the extreme minimum and maximum values in stratum 4 and three nearly identical values in stratum 5. These quantities depend on the way the strata overlay a particular storm period and emphasize the need to reduce sampling variance by selecting short strata. Other schedules will give different patterns of stratum lengths for this storm, resulting in different variances.

Comparisons

Simulating the SALT or time-stratified sampling distributions by Monte Carlo methods is unnecessary for these comparisons because all five populations are completely known. Since the auxiliary variables for SALT, the stratum boundaries for each time-stratified sampling schedule, and sample sizes for both schemes are also known, the true

TABLE 1. Some Sampling Schedules for Real-Time Control of Stratum Lengths During Storms

	Stratum Length, min							
	Rising Stage, cm				Falling Stage, cm			
Schedule	0–122	122–152	152–183	>183	>183	183–152	152–122	122–0
1	180	90	50	50	90	180	180	360
2	180	90	50	50	90	180	360	360
3	180	90	90	50	90	180	180	360
4	180	90	90	50	90	180	360	360
5	180	180	90	50	90	180	180	360
6	180	180	90	50	90	180	360	360
7	180	180	90	90	180	180	360	360
8	180	180	90	90	180	360	360	360
9	180	180	90	90	180	360	360	720
10	180	180	90	90	180	360	720	720
11	180	180	180	90	180	180	360	360
12	180	180	180	90	180	360	360	360
13	180	180	180	90	180	360	360	720
14	180	180	180	90	180	360	720	720
15	180	180	180	180	180	180	360	360
16	180	180	180	180	180	360	360	360
17	180	180	180	180	180	360	360	720
18	180	180	180	180	180	360	720	720

The stratum length is governed by stage height and whether the stage is rising or falling. The number of sampling units selected in each stratum is fixed at either two or three, resulting in 36 different schedules.

sample variances for both methods can be calculated directly from the known 10-min period loads. The appendix contains formulas for the load estimates and true variances for the SALT and time-stratified methods as well as the variance estimate for time-stratified sampling.

To compare sampling schemes and schedules we used a kind of "coefficient of variation" calculated by dividing the standard deviation of the estimate of the load by the true load (both known population values in our case). Variation is thereby expressed as a proportion of load, enabling direct comparison of the performance of sampling schemes over storms of differing sizes and facilitating plotting. In using this statistic, the variation in small storms may seem as “important” as that from large ones. The purpose of a study must be known to decide whether obtaining precise estimates is as important for small storms as for large storms.

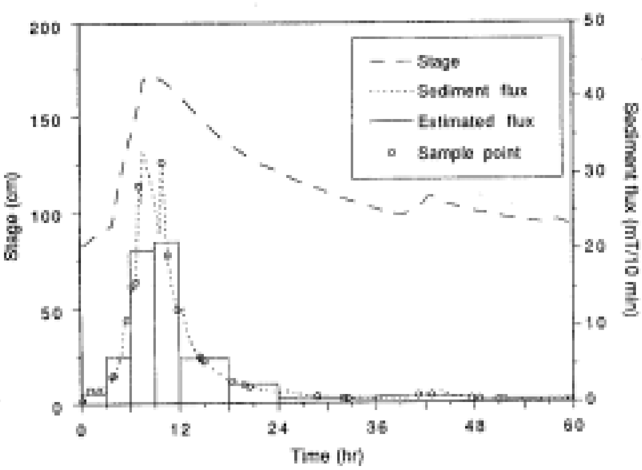

Fig. 2. Details of a single application of time-stratified sampling to storm 2 using schedule 13 (Table 2) with three specimens per stratum. Widths of bars indicate stratum durations. Heights of bars indicate sample estimates of mean stratum flux (metric tons) per 10-min period.

RESULTS

Effects of Time Stratification

Stratification is widely used to reduce variance in sample estimates by taking advantage of information about population structure. The information is used to partition the population into two or more "strata" within which the variable of interest is relatively homogeneous. Stratum sample variances are thereby kept low and population variance, the sum of the stratum variances, is usually reduced. It is harder to define strata in the time-stratified case since it must be done in real time before the population structure is completely known. The key to reducing stratum variation when estimating suspended sediment flux is to limit the range of flux in each stratum, which is best done by restricting stratum length.

The effect of stratum length on variation is seen by plotting the coefficient of variation against stratum length for each of the five storms (Figure 3). All strata within a sample had the same specified lengths for this plot, and sample sizes were proportional to storm lengths (ranging from 42 to 76). The coefficient of variation drops markedly as stratum lengths decrease (i.e., the number of strata increase) even though sample size remains constant within each storm. The coefficients of variation do not always drop smoothly due to interactions between stratum lengths and the idiosyncrasies of each storm hydrograph. Setting fixed short stratum sample sizes of two (or three) still allows variance to be calculated, decreases overall sample variance, and reduces sampling schedule development to choosing stratum lengths. Evidently, the reduction in the sum of squared deviations in a short stratum more than compensates for the reduced stratum sample size.

TABLE 2. Computation of Time-Stratified Sample Estimates of Total and Variance, Applying Sampling Schedule 13 to Storm 2

		Estimated		True	
Stratum	Stratum Length, min	Total per Stratum, t	Variance, t^2	Total per Stratum, t	Variance, t^2
1	180	23.81	51	29.71	35
2	180	109.53	1,612	114.91	1,011
3	180	358.29	5,042	455.73	4,247
4	180	379.57	8,528	365.12	3,452
5	360	212.58	37	224.67	2,901
6	360	90.18	50	75.32	95
7	720	49.06	31	60.92	204
8	720	58.20	132	53.02	148
9	700	27.28	4	23.98	6
Total		1,308.50	15,487	1,403.38	12,099

The unit t denotes metric ton.

A problem in applying time-stratified sampling can occur on recession limbs. Strata must be long under such conditions to reduce sampling intensity during low and falling flows, thereby keeping overall sample sizes low. It is possible for these long-duration strata to extend into the rising limb of a subsequent storm, thus undersampling the next period of high sediment flux. Stations being operated manually must be visited before the hydrograph begins its rise to shift to shorter strata. For automatic operation the program must recognize sudden increases in stage and initiate shorter strata. In either case the last long stratum must be ended and provision made to deal with the possibility of not having collected enough specimens before cutoff. This may introduce some small bias, but loss of information in recession strata is not as critical as inefficient sampling during rising limbs.

Comparisons of SALT and Time-Stratified Sampling

For comparison purposes time-stratified variances were calculated by applying the 36 schedules directly to the five populations, while SALT variances were computed using auxiliary variables derived from the schedules or from an "optimal" scheme based on data from storm 1. The SALT auxiliary variables derived from the schedules were intended to compare the two schemes using the same "information" about the population. They were stepwise functions with the same average sampling rates for all stage classes and conditions as a corresponding schedule. Sampling rates were calculated by dividing the stratum sample size (two or three) by the stratum length in hours. Time-stratified sample sizes from corresponding schedules (i.e., the expected SALT sample sizes) were used to calculate the coefficients of variation for SALT.

The "optimal" SALT auxiliary variable was developed using rising and falling stage rating curves from storm 1 (Figure 4). Two sets of rising and falling steps were visually selected as being reasonably linear (omitting several points with falling flux values at the end of the rising limb) and were fitted by straight-line regressions. This method yielded R^2 values of 0.978 for the rising limb and 0.993 for the falling limb, but the data are more complex than is implied by linear models because the residuals are serially correlated. Nevertheless, this compromise between quality of the auxiliary variable and ease of its development works well for storm 1 except for units toward the end of the rising limb. The performance of this auxiliary variable is not as good for the

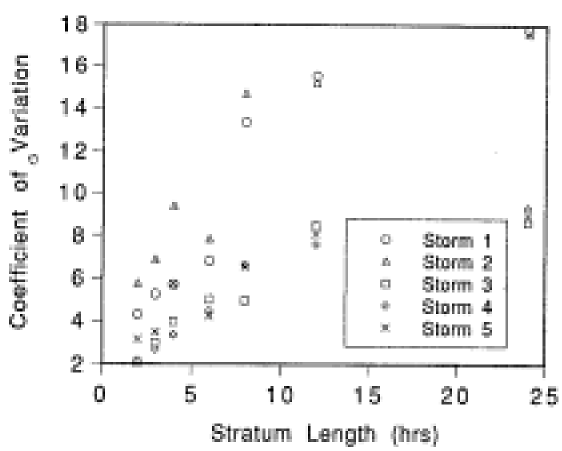

Fig. 3. Coefficients of variation to compare different degrees of stratification with average sampling rates fixed at one bottle per hour. Sample sizes vary from 42 to 76 and are proportional to storm lengths.

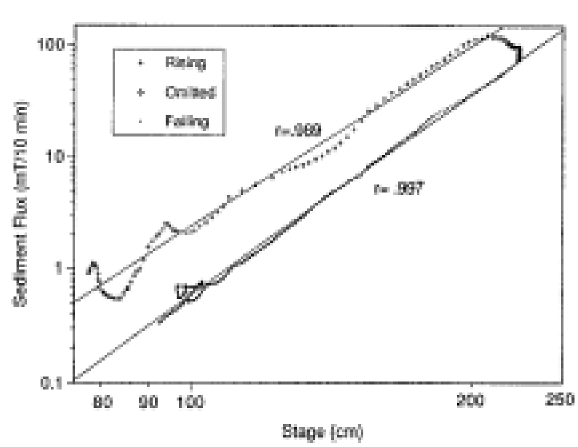

Fig. 4. Rising and falling stage regressions used for optimal SALT auxiliary variable for storm 1. (The abbreviation mT denotes metric tons.)

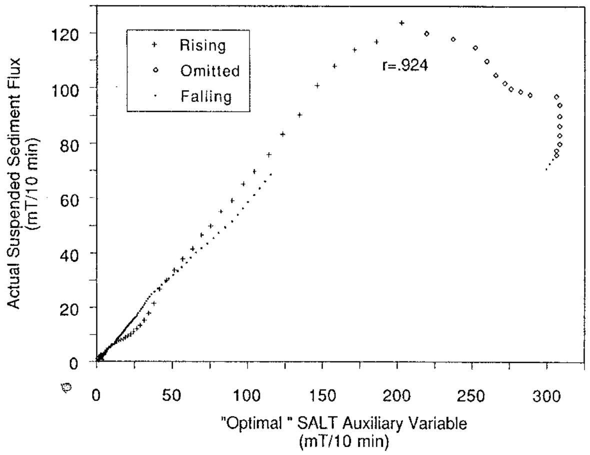

Fig. 5. Plot of suspended sediment flux (metric tons per 10-min period) versus the stage-based "optimal" auxiliary variable. Symbols denote how points had been classified when developing auxiliary variable (Figure 4). The threshold for switching from the rising to falling regression was a decline of 4.5 cm in stage height.

other storms, which illustrates a problem inherent with rating curves (and SALT schemes based on them).

The benefit of using SALT over simple random sampling depends on the quality of the auxiliary variable. A SALT auxiliary variable is optimal when it is proportional to the primary variable. This explains why our "optimal" auxiliary variable does not perform perfectly. Neither regression adequately fitted the points near the peak of the sedigraph when sediment yield started to fall and stage was still rising. Applying both regressions to all points in storm 1, it is apparent (Figure 5) that these points cause a severe departure from proportionality near the peak of the hydrograph.

A major contrast between SALT and time-stratified sampling is that SALT can theoretically have zero variance with a "perfect" auxiliary variable. Therefore, another auxiliary variable might be found that reduces the SALT variance below that for time-stratified sampling. However, at least for populations of suspended sediment, it appears difficult to develop efficient auxiliary variables that are simple functions of stage. Alternatively, time-stratified sampling does not appear to be overly sensitive to which sampling schedule is used as long as the strata are kept short relative to changes in sediment flux so that the range of flux is limited in each stratum.

Because of its robustness, a single time-stratified schedule is likely to be an acceptable compromise for sampling several different water quality constituents. Satisfactory estimates of a number of constituents should be obtainable from the same time-stratified sample. In contrast, a different SALT sample using a different auxiliary variable would probably be needed to satisfactorily estimate each constituent load with SALT.

The results of variance calculations for the 36 schedules and five storms are presented as plots of coefficient of variation against sample size (Figure 6). Each time-stratified and associated SALT average sampling rate schedule is shown by a symbol indicating the (associated) stratum sample size of two or three. The coefficient of variation for "optimal" SALT sampling is a continuous function of sample size; thus it is shown by a solid line. There are fewer than 36 symbols for two- and three-stratum sample sizes for all except storm 1. This is because several of the schedules differ only for higher stages and the smaller storms peaked before these stages were reached. Hence, some of the symbols for plots other than for storm 1 represent more than one schedule.

Time-stratified sampling schedules had markedly lower coefficients of variation than did the associated SALT average sampling rate schemes over the entire range of sample sizes for all storms. The SALT coefficients of variation were from 1.4 to 7.7 times as large as those for the time-stratified scheme. The effect was most pronounced for the larger storms where the short strata evidently ensured that there was little within-stratum range in sediment flux. Since corresponding SALT and time-stratified samples used the same sampling rates, the lower coefficients of variation for time-stratified sampling indicate that greater benefit accrues from stratification than from adjusting sampling rates to changing stage conditions.

The "optimum" SALT scheme did better than the average sampling rate method. It performed similarly to the time-stratified scheme for the larger storms but not as well for storms 3, 4, and 5. In the larger storms, there was some tendency for optimum SALT to perform relatively better than time-stratified sampling for smaller sample sizes and poorer for larger sample sizes. Small time-stratified samples require schedules having longer strata, which tends to raise variance, while larger samples allow the benefits of shorter strata to be realized. For storm 5, optimal SALT did no better than SALT based on average sampling rates.

Of course, in an operational setting it is not possible to use one of the sampled storms to derive the auxiliary variable, so the efficacy of SALT would depend on the data used to derive it and on the relationship between those developmenttal data and the data to which the auxiliary variable is applied. It is not clear what storm characteristics are best for collecting data to develop a good SALT auxiliary variable. Sediment flux appears to respond differently to rising limbs of different storms, so that an auxiliary variable that performs well for one storm may not do so for another. These are manifestations of the same problem that plague the development of rating curves for direct estimation of sediment loads. Sediment flux is clearly not a simple function of simultaneously measured river stage (or discharge).

To assess how the two sampling methods estimate loads overall for the different storms, we compared a single time-stratified schedule to the optimum SALT scheme for a nominal target of 48 specimens for storm 1 (Table 3). Schedule 13 was used for the time-stratified method with three specimens per stratum, which resulted in a sample of size 47 for storm 1. (Sample sizes were not always multiples of 3 since the last stratum in some storms extended beyond the end of a storm and had to be truncated. One, two, or three specimens were used in truncated strata according to whether the truncation was done in the first, second, or last third of the stratum, respectively.) The optimum SALT auxiliary variable was used with a sample size of 48 for storm 1. SALT sample sizes for the other storms were proportions of the storm 1 sample size based on the respective auxiliary variable totals as shown in the fourth column from the left of Table 3. The actual real-time SALT sampling variance is somewhat higher than that reported here due to random sample sizes.

The coefficients of variation for the two sampling methods

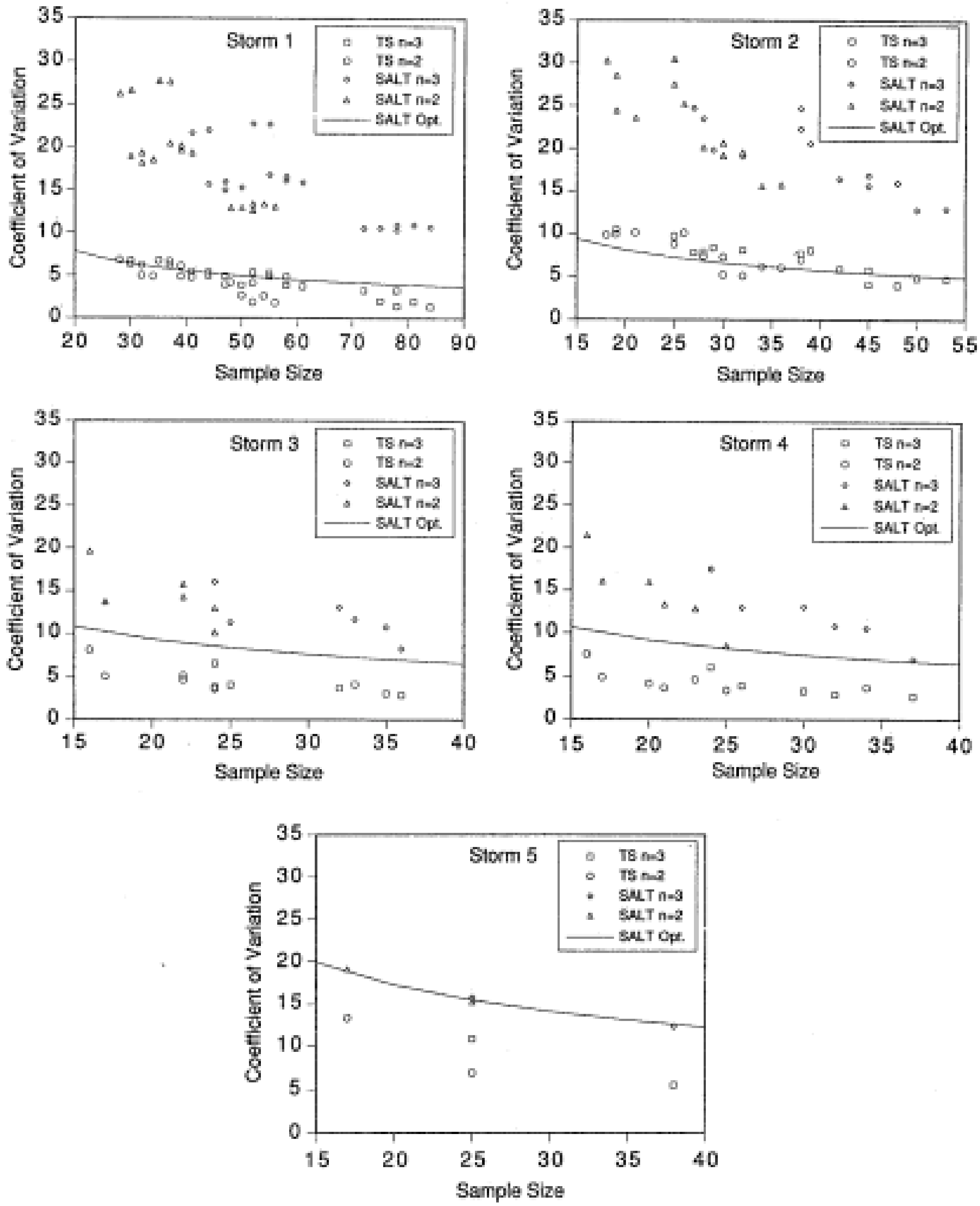

Fig. 6. Coefficients of variation for all 18 time-stratified (TS) sampling schedules with two and three samples per stratum (Table 1), the corresponding SALT stepwise schemes, and the SALT procedure optimized for storm 1.

are nearly identical for storm 1. However, the SALT coefficients get progressively larger than those for the corresponding time-stratified samples as storm size falls. Part of this difference is because SALT allocates fewer specimens to smaller storms relative to time-stratified sampling. This auxiliary variable resulted in a total SALT sample size of 70 for all five storms, which is less than half of the 150 specimens collected in time-stratified sampling.

The coefficients of variation for an estimate of the total over all storms were 3.77 for time-stratified sampling and 4.51 for SALT. Therefore, the SALT coefficient of variation was about 20% higher than that for time-stratified sampling, but was based on a sample less than half the size. The SALT scheme which keeps the same proportion of sample sizes in the several storms, but with an overall sample size of 150, has a coefficient of variation of 3.08, which is about 18% lower than the value for time-stratified sampling.

Choosing between these sampling plans should be governed by the purpose of the study. In some situations, sediment load estimates are wanted for every storm. In this example the time-stratified procedure collected data more evenly over the diferent sized storms. This depends partly on the schedule and auxiliary variable being used, but generally time-stratified sampling "covers" a hydrograph

TABLE 3. Sample Sizes and Coefficients of Variation (CV) Resulting From the Application of Time-Stratified Sampling (Schedule 13) and SALT Sampling

	Time-Stratified Sampling		SALT Sampling*		
Storm	Sample Size	CV	X	Sample Size	CV
1	47	4.92	11,664	48	5.05
2	27	7.84	2,755	11	10.98
3	25	4.09	1,038	4	20.97
4	26	3.99	1,050	4	20.68
5	25	10.81	739	3	44.62

*The auxiliary variable was optimized by a power function of stage on both rising and falling limbs to fit storm 1. SALT sample size was set to 48 for storm 1 and is proportional to X, the sum of the auxiliary variable values.

more reliably than does SALT. If one is interested primarily in total load over a long period (such as a year) a SALT scheme might be preferred to better allocate sampling resources to the more important events.

SUMMARY AND DISCUSSION

Traditionally, suspended sediment in streams has been sampled and loads estimated using "nonstatistical" methods, that is, methods that do not take probabilities of sample selection into account. This practice has produced biased estimates of load with little or no dependable information on sampling and measurement errors. Statistical estimates of loads can eliminate bias and, if rigorously applied, yield good estimates of how well the scheme has performed for each sample: These benefits suggest that hydrologists should seriously consider using statistical sampling and estimating procedures to determine suspended sediment loads, especially when the results are to be used in sensitive projects.

SALT and time-stratified sampling are two statistical methods for sampling populations of suspended sediment flux that are essentially unbiased and provide valid estimates of variance. Moreover, they give the hydrologist a way to take advantage of knowledge of the population to reduce variation. For SALT this knowledge takes the form of a general understanding that higher suspended sediment flux occurs during higher and, especially, rising flows. With time-stratified sampling, the key information is that variation in suspended sediment flux occurs during such flows. Both methods optimally tend to concentrate data collection during the same periods, but with different implementation and with different results.

The variances of the SALT and time-stratified methods were compared for known populations of five different sized storms. The populations were determined using continuous measurements of turbidity which were used to estimate suspended sediment concentration at 10-min intervals. Comparisons were made for 36 “schedules” that specified the sampling intensity by stage class and condition (rising or falling) for both sampling schemes. These were also compared to an "optimum" SALT scheme which used an auxiliary variable developed by maximizing the correlations between sediment flux and stage on both rising and falling stages of the largest storm.

SALT has the higher potential for reducing sampling variance; a "perfect" auxiliary variable would always predict the variable of interest exactly and would thus have a sampling variance of zero. However, at least for estimating suspended sediment loads, it is difficult to find simple functions of stage or discharge that give good predictions of suspended sediment flux over a wide variety of stream conditions. SALT may have wider application when sampling over long periods of time in which sampling resources must be apportioned over storms of different sizes.

Time-stratified sampling had lower variance than SALT for virtually all comparisons for all five storms. The coefficients of variation were from 13 to 70% of those for the corresponding SALT samples of the same size based on average sampling rates taken from the time-stratified sampling schedules.

The critical factor for good performance of time-stratified sampling is that the strata be as short as possible to reduce the variation of within-stratum sediment flux. This can be done by choosing small stratum sample sizes. Stratum sample sizes of two should be used if there is little chance of losing specimens, or three to guard against loss of specimens.

In designing a time-stratified sampling schedule, one must first consider the type of river, the number of measurements that can be feasibly made and the expected range in stage or discharge. Applying hypothetical schedules to existing hydrographs yields probable sample sizes. On small or flashy rivers like the North Fork Mad River, strata should not exceed 1-2 hours when conditions are changing rapidly. In each of our schedules, the longest stratum was 4-8 times the length of the shortest stratum. Stage boundaries might be selected any number of ways. Dividing the range from base flow stage to peak stage into equal intervals worked well here.

Because time-stratified samples are time based they are better at "covering" storms of different sizes, which can be of great benefit in studies that require an estimate of load from each storm. These samples seem not to be as sensitive to how population knowledge is used as is the SALT technique. Most of the time-stratified sampling schedules worked well. For a given sample size, schedules with sample sizes of two gave lower variance than those with sample sizes of three in longer strata. Time-stratified samples are apparently more "robust" in that they can be expected to produce precise load estimates in most situations without having to choose the best schedule for specific conditions. For this reason, the method works well in situations where samples are analyzed for multiple constituents. Because time-stratified sampling is an adaptation of commonly used stratified random sampling, it may also have the advantage of being somewhat easier for those applying the method in the field to understand.

A condition to be aware of with time-stratified sampling is that long recession limb strata designed to reduce total sample size may extend into the early stages of a following storm. It is critical to periodically test for a rising stage condition and switch to short strata as soon as it is detected.

By using the "same" population information, we attempted to make fair comparisons of time-stratified sampling with SALT using the schedule-based auxiliary variables. However, with the "optimal" SALT auxiliary variable based on the relationship between sediment flux and stage, SALT performed much more like (but no better than) time-stratified sampling. This form of optimization may give SALT an undue advantage because in a real-time situation the auxiliary variable would have to be determined before sampling began. Also, it would be unusual to have as much high-flow data available as we used in developing this auxiliary variable. It appears that selecting a SALT auxiliary variable that performs well across a wide range of stream conditions is a difficult problem.

The populations of sediment flux values used for these comparisons were derived from turbidity data that were "smoothed" because of broad chart traces caused by a rapidly moving instrument trace. It is thought that these variations are artifacts of the instrumentation. If not, our populations may not reflect the levels of actual variation and serial correlation that exist in true populations. Other methods of collecting populations of sediment flux are being pursued that will not have these problems and can be used to make further comparisons.

Applying statistical sampling to estimate suspended sediment is a major undertaking. However, such programs can be viewed as transferring much of the effort to earlier parts of the study. Planning is critical, and the logistics of the field operation are demanding and must be carried out reliably and competently. This implies the need for well-trained field crews of adequate size. Rigorous application of these methods requires sampling to be done when flux is highest, which often turns out to be times that are generally least amenable to crew availability and station visits. It is difficult with any flow-based sampling plan such as time-stratified or SALT to know when the stations must be serviced since the complement of pumping sampler bottles is filled as a function of flow, which is generally not known away from the field. There are also significant problems keeping track of the data so that all information is available for making the estimates.

The benefits of statistical sampling plans come at the end of a study when estimates are made. The estimators are known at the start and can be applied "automatically" when the data are available after collection. Load estimates and their variances can be relied on to have the selected properties designed into them at the beginning of the study. The results are defensible according to accepted methods of inference regardless of the population to which they are applied.

Appendix

Let q_i be the water discharge and c_i the measured suspended sediment concentration for the ith sampling period. If Δt is the period length and k is a constant to adjust units, the sediment load for period i is $y_i = kq_ic_i\Delta t$. The corresponding auxiliary variable is $x_i = f(s)\Delta t$, in which $f(s)$ is a function of river stage. If N is the population size, $X = \sum_{i=1}^{N} x_i$, and r_i is a variable indicating the number of random numbers in the ith interval ($r_i = 0$ for all nonsampled periods), then the SALT estimator for the total (suspended sediment load) from a sample of size n can be written as [*Thomas*, 1985]

$$\hat{Y}_{\mathrm{SALT}} = \frac{X}{n} \sum_{i=1}^{N} r_i \frac{y_i}{x_i} \tag{1}$$

and, given that the sample size is n, the true variance of $\hat{Y}_{\mathrm{SALT}}$ is [*Norick*, 1969]

$$\mathrm{Var}\,[\hat{Y}_{\mathrm{SALT}}] = \frac{1}{n}\left(X \sum_{i=1}^{N} \frac{y_i^2}{x_i} - Y^2\right) \tag{2}$$

Divide the population into L strata, each with N_h sample units, y_{hi}. Then the true stratum means, $\bar{Y}_h$, and variances, S_h^2, are given, respectively, by

$$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi} \tag{3}$$

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2 \tag{4}$$

Similarly, the stratum sample means, $\bar{y}_h$, and variances, s_h^2, for a sample of n_h in the hth stratum are

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi} \tag{5}$$

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 \tag{6}$$

The time-stratified estimate, $\hat{Y}_{st}$, of the total (i.e., the load) is

$$\hat{Y}_{st} = \sum_{h=1}^{L} N_y \bar{y}_h \tag{7}$$

and an estimate of the variance based on the sample is

$$\widehat{\mathrm{Var}}\,[\hat{Y}_{st}] = \sum_{h=1}^{L} N_h(N_h - n_h)\frac{s_h^2}{n_h} \tag{8}$$

If all population units are known (as they are for the populations used in this paper) the true variance for the estimator $\hat{Y}_{st}$ is calculated by [*Cochran*, 1963]

$$\mathrm{Var}\,[\hat{Y}_{st}] = \sum_{h=1}^{L} N_h(N_h - n_h)\frac{S_h^2}{n_h} \tag{9}$$

Acknowledgment. We thank the Simpson Timber Company for its cooperation with the Forest Service, U.S. Department of Agriculture, in collecting data at the test site.

References

Cochran, W. G., *Sampling Techniques*, 2nd ed., John Wiley, New York, 1963.

Cochran, W. G., F. Mosteller, and J. W. Tukey, Principles of sampling, *J. Am. Stat. Assoc., 49*(265), 13-35, 1954.

Eads, R. E., Controlling sediment collection with data loggers, in Proceedings of the Fifth Federal Interagency Sedimentation Conference, vol. I, edited by S.-S. Fan and Y.-H. Kuo, pp. 2-41 to 2-48, Fed. Energy Regul. Comm., Washington, D. C., 1991.

Hewlett, J. D., and A. R. Hibbert, Factors affecting the response of small watersheds to precipitation in humid areas, in *Forest Hydrology*, edited by W. E. Sopper and H. W. Lull, pp. 275-290, Pergammon, New York, 1967.

Norick, N. X., SALT (selection at list time) sampling: A new method and its application in forestry, M.A. thesis, 50 pp., Univ. of Calif., Berkeley, 1969.

Thomas, R. B., Estimating total suspended sediment yield with probability sampling, *Water Resour. Res.*, *21*(9), 1381-1388, 1985.

Thomas, R. B., Monitoring baseline suspended sediment in forested basins: The effects of sampling on suspended sediment rating curves, J. *Hydrol. Sci.*, *33*(5), 499-514, 1988.

Thomas, R. B., Piecewise SALT sampling for estimating suspended sediment yields, *Gen. Tech. Rep. PSW-114,* 11 pp., For. Serv., U.S. Dep. of Agric., Berkeley, Calif., 1989.

Thomas, R. B., Systematic sampling for suspended sediment, in Proceedings of the Fifth Federal Interagency Sedimentation Conference, vol. 1, edited by S.-S. Fan and Y.-H. Kuo, pp. 2-17 to 2-24, Fed. Energy Regul. Comm., Washington, D. C., 1991.

J. Lewis and R. B. Thomas, Pacific Southwest Research Station, USDA Forest Service, 1700 Bayview Street, Arcata, CA 95521.

(Received May 28, 1992;
revised October 27, 1992;
accepted November 12, 1992.)