SYSTEMATIC SAMPLING FOR SUSPENDED SEDIMENT

By Robert B. Thomas, Mathematical Statistician,
USDA Forest Service, Arcata, California

ABSTRACT

Because of high costs or complex logistics, scientific populations cannot be measured entirely and must be sampled. Accepted scientific practice holds that sample selection be based on statistical principles to assure objectivity when estimating totals and variances. Probability sampling--obtaining samples with known probabilities--is the only method that assures these results. However, probability sampling is seldom combined with appropriate estimators to determine suspended sediment loads. Many current load-estimating methods, therefore, have unknown bias and variation making the estimates questionable.

Suspended sediment loads are often estimated by sampling concentration at fixed intervals. This type of sampling is-promoted by the widespread use of pumping samplers which can be set to sample at regular intervals. Sampling intensity is sometimes increased during periods of high water discharge.

Randomly started systematic samples are probability samples, and estimates of totals from such samples are unbiased. These estimates tend to have low variance, but the variance cannot be *estimated*, and is not always reduced by increasing sample size. Systematic sampling of concentration distributes data evenly over time, so that most measurements are collected during low flows, and few during the brief high-flow periods when most sediment is transported.

Systematic sampling for estimating suspended sediment loads is investigated for a "complete" sediment record from the Mad River in northern California. The "true" total for the 31-day period is compared to expected estimates from systematic samples without random starts. Systematic sample variance is compared to three other finite sampling schemes and to estimates using the simple random sampling variance formula. The effects of changing systematic sample size are also studied.

If systematic sampling is used to estimate suspended sediment loads, the limitations of the method should be realized and correct estimating formulas used. The best use for systematic "sampling" is to define the sampled population for further sampling by more efficient finite sampling schemes.

INTRODUCTION

Fixed-Interval Sampling

Widespread use of pumping samplers promotes collection of suspended sediment data at regular time intervals. Although ease of data collection must be considered, the dictation of sampling method by technology (or logistical convenience) may result in distorted and misleading estimates and comparisons.

Sampling Suspended Sediment Data

Two factors should guide sampling of suspended sediment populations. One is the difficulty and cost of collecting and processing specimens, which dictates that samples be small. ("Specimen" refers to a bottle of water/sediment, and "sample" indicates a collection of specimens.) A second factor is that most sediment flux occurs during rare and brief periods of high flow. Such "sporadic" populations of suspended sediment are best sampled by emphasizing periods of high sediment flux. Fixed-interval samples do the opposite by spreading specimens evenly over the entire population.

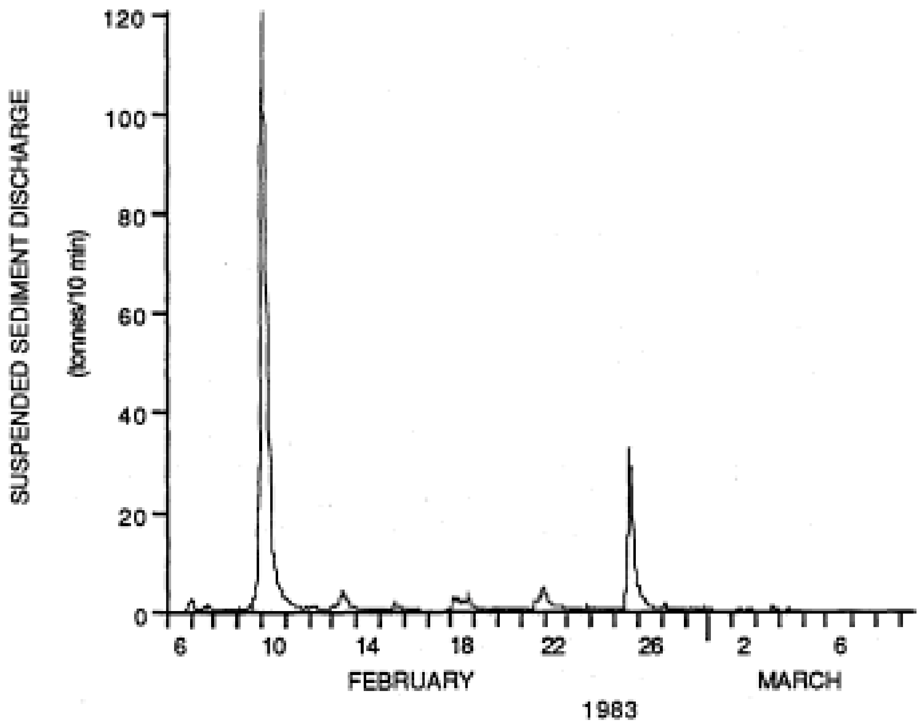

Figure 1- ***Sedigraph of a 31-day period from water year 1983 at the North Fork of the Mad River near Korbel, California. Suspended sediment yields were calculated for 10-minute periods.***

Study Data

Data used in this study were taken from the North Fork of the Mad River near Korbel in northern California (Figure 1). Turbidity was continuously monitored for a 31-day period in 1983, and a good relationship found between turbidity and suspended sediment concentration. This relationship was used to predict the suspended sediment loads for 4450 10-minute intervals. This large finite data set is assumed to be the "true" population for sampling purposes; it was used for computing totals and variances for several sampling schemes and calculating the "true" suspended sediment load for the period.

ESTIMATING LOADS WITH FIXED-INTERVAL DATA

Sediment rating curves are commonly used to estimate suspended sediment loads, regardless of the method of data collection. Rating curves model the logarithm of sediment response as a linear function of the logarithm of the simultaneous water discharge. This model can give highly biased estimates, especially for small streams (Walling, 1977a; 1977b; Walling and Webb, 1981). Ferguson (1986, 1987) suggested applying a correction for bias caused by the logarithmic transformation, but this method is not always satisfactory (Thomas, 1985; Walling and Webb, 1988). Thomas (1988) found that uncorrected and corrected rating-curve estimates using fixed-interval samples of 50 units from the Mad River data were biased. Also, a rating curve of the entire population gave an estimate of population total which was biased about the same as that for the samples. Therefore, rating curves are problematic, and should only be used where the hydrologist is certain that the model fits well.

Time series analysis requires fixed interval data to estimate total suspended load. Transfer function models have been used with fixed-interval data with good success (Gurnell and Fenn, 1984). Time series analysis accounts for the *serial correlation* usually present in closely spaced sequential data. Aside From the complexity of time series analyses, the generally sporadic nature of sediment flux makes it difficult to collect adequate information without frequent samples and consequent high cost. When such expenditures are justified, time series analyses provide information on patterns of sediment flux above that required for estimating suspended sediment loads.

Other techniques for estimating totals of *finite* populations are based on survey sampling theory. These methods are appropriate when the population is finite or when a reasonable finite *sampling population* (the population actually sampled) can be formed from the *target population* (the population of interest). A finite sampling population can be formed from a period of continuous sediment discharge by dividing the period into short equal-length time intervals. The interval length is chosen so that water discharge and sediment concentration measurements made at the midpoint reasonably represent the continuous sediment flux for each interval. This method uses fixed-interval "sampling" to form a finite sampled population from a continuous one.

Sample inferences apply to sampled and not target populations. Therefore, these populations must be similar in essential details, a condition usually based on judgement. The sampled and target populations of sediment flux are logically similar if the sample intervals are "short."

Survey sampling methods can often take advantage of any knowledge of the population to reduce sample size or variance. The fact that most sediment flux is delivered during periods of high water discharge can be used in several ways to obtain higher quality estimates with greater sampling efficiency.

SYSTEMATIC SAMPLING

In survey sampling theory, fixed-interval samples are called *systematic* samples. Consider a sequential population with $kn+c$ $(0 \leq c < k)$ units. A systematic sample includes one of the first k units chosen at random, and every unit at intervals of k after the first. Each of the k possible samples is, in effect, a *cluster* of n or $n+1$ units that covers the population in a regular way. Systematic sampling is often used because it is easy to apply.

Clearly, composition of a systematic sample depends on which of the first k population units is selected at random and on population order. If the population is randomly ordered, each sample will be random. Such data can be treated as independent samples for estimation purposes. Suspended sediment discharges tend not to be random, however, at least when measured intervals are short enough to adequately define the process. Therefore, simple random sampling estimators should not be used with systematic samples of suspended sediment data unless serial correlation is negligible.

Randomly started systematic samples are probability samples (regardless of population order) and the estimates of totals are unbiased. If the ith suspended sediment load for the jth interval is given by Yij, an unbiased estimator for total load during the sampled period is

$$\hat{Y} = k \sum_{i} y_{ij} \qquad (1)$$

(Raj, 1968). Since each cluster covers the population evenly, these estimates tend to have low variance, but the variance cannot be estimated for the usual single-cluster systematic sample. Also, controlling the variance by changing sample size is dependent on the order of units in the population. In some cases increasing n actually increases the variance. Population order does not affect properties of the estimators for other forms of probability sampling in which larger sample sizes produce reduced variance.

SYSTEMATIC SAMPLES WITH AND WITHOUT RANDOM STARTS

Starting times for pumping sampler collection of concentration data usually depend on administrative and weather-related factors. In non-storm periods data collection intervals are long and starting times based on administrative convenience. Such data sets may be best from one statistical standpoint: long intervals can result in independence among the measurements so they can reasonably be treated as simple random samples for estimating totals and variances. However, during low flows, suspended sediment discharge is also low, and the contribution to the overall total and variance is small.

Sampling intervals can be shortened for higher flow periods, generally in response to existing or expected weather or stream conditions. It is reasonable to assume that storms from large frontal disturbances arrive at random times, but the interaction of work schedules and the logistics of getting to stations to start or change sampling programs may still produce nonrandom starts. Logistical and administrative restrictions are real, but they can be surmounted to ensure random starts for systematic samples. Pumping samplers not only have clocks to sample at preset intervals, but they also have time delays that can be used to initiate sampling at random times.

The sampled population is first specified by defining its units. Units are short time-intervals in the monitored period that can be characterized by one water discharge and suspended sediment measurement at mid-interval. Interval length depends on response time of the river; large rivers that react slowly to storm inputs might have units of several hours, while small flashy streams may have units of five to ten minutes. A criterion for suitable interval lengths is confidence that if all intervals could be measured, the resulting total would be the same as that for the continuous target population. It is best to err by selecting the intervals too short, especially since that does not greatly increase the sample size required.

In "ideal" systematic sampling the sampled population size, N, is determined by dividing the monitoring time by the sampling-unit duration. The sample size, n, is then chosen and k found by dividing N by n. If N is not an exact multiple of n, some samples are of size n and others are of size n+l. For sampling suspended sediment, however, n is usually set by the number of bottles in a pumping sampler and N by some vague limit imposed by costs of laboratory processing. Usually, no provision is made for samples of different size, so, implicitly, kn equals N exactly (i.e., c=0).

If an 18-hour period with 10-minute sampling units is to be sampled with a 24-bottle pumping sampler, there are 108 sampling units which cannot be divided evenly by 24. The sampling interval, k, can be chosen as 4 or 5 implying that N is 16 or 20 hours instead of the nominal 18. In either case, a value from 1 to k is chosen at random, and the first pumped sample taken at that time.

The effects of not randomizing over all possible starting times are shown for samples of size 62 (an average of 2 per day) for the Mad River data. Consider restricting starting times to any four-hour period in the k=71 (4450=71*62+48) starting times. There are 48 such four-hour periods from the beginning to eight hours into the record. Random samples in the four-hour periods give expected totals which are plotted against the first interval numbers of the periods (Figure 2). The expected values are generally biased, ranging from about 6750 to over 10,500 tonnes compared to the "true" load of 8307 tonnes.

In this case, if the samples were randomized within the four-hour period beginning at the 20th interval into the record, the expected total would be nearly unbiased. However, the shape of this curve and the point of crossing the true total are dependent on the ordering of units in the population, which is not known, even after the sample is collected. Generally, not enough is known about specific populations to ensure unbiased estimates of total loads using systematic sampling unless starting times are randomly selected over all possible samples.

VARIANCES FOR SEVERAL FINITE SAMPLING PLANS

Sampling plans should be chosen for performance characteristics as well as for ease of application. The performance of finite sampling schemes is measured primarily by bias and variance. Unbiased estimators produce distributions of estimates having expected values equal to the parameters being estimated. Most finite sampling schemes have little or no bias, so the main interest is in methods with minimum variance. Lower variance gives parameter estimates with smaller errors, better sensitivity for detecting differences, or reduced sample size. We now compare the variance of systematic sampling to several other finite sampling schemes suitable for measuring suspended sediment loads.

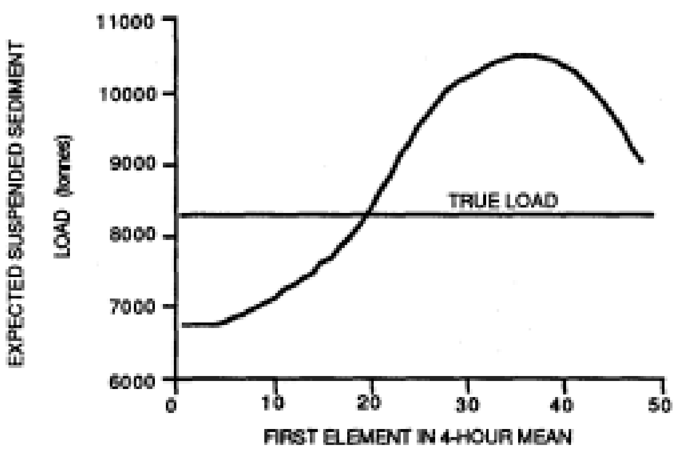

Figure 2 - ***Expected suspended sediment loads for systematic samples of size 62 from the Mad River data. Values were computed by restricting random starts to 4-hour periods starting from the beginning to 8 hours into the record.***

There are formulas to calculate "true" variances for all finite sampling plans, but they depend on knowing the complete population. Since the entire Mad River sample population is known for the 31-day record, the "true" variance of systematic sampling can be calculated even though the variance cannot be *estimated* from a sample.

Systematic sampling was compared to simple random sampling (SRS), stratified random sampling (STRS), and selection at list time sampling (SALT), each sample having 62 observations (Figure 3). SRS was used as a benchmark method because it is the most fundamental finite sampling plan. SRS is not recommended for suspended sediment populations because, like systematic sampling, it does not emphasize periods that produce most sediment flux. Even though SRS estimates of the total and its variance are unbiased, its precision is inferior to that for systematic sampling for the Mad River population as measured by the standard error of the total. This result is expected since systematic samples always have at least some measurements during high-flow periods, while SRS measurements during those periods are due to chance.

Calculating "variance" from *systematic samples* using the *SRS variance formula* illustrates the effects of not using the correct estimators for given sampling schemes. The estimated standard error of systematic samples using SRS formulas averages about 3 times the true standard error, and is nearly equal to that for real SRS samples. Even though systematic sampling in this case is far superior to SRS, its variance is greatly overestimated by the SRS variance formula. This emphasizes the need to match sampling plans and estimators.

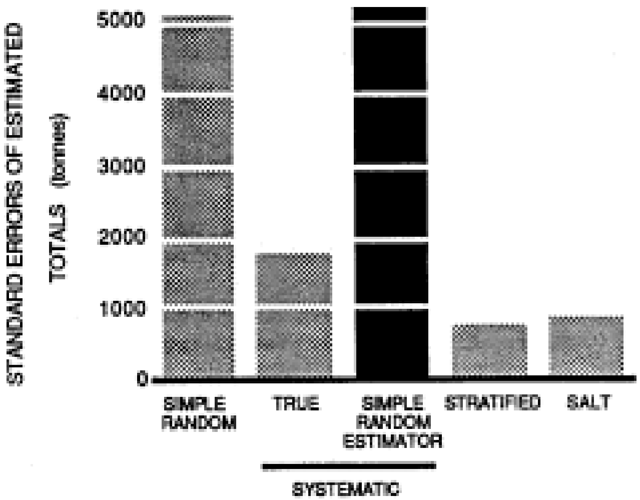

Figure 3 - ***Comparison of sampling standard errors for five estimators of suspended sediment load for the Mad River data. All values are "true" standard errors calculated from the population for samples of size 62. The dark bar shows the mean standard error obtained from using the simple random sampling estimator with systematic samples.***

STRS is widely used to reduce variance. If populations can be divided into homogeneous strata and separate SRS samples are taken in each stratum, unbiased estimates having lower variance are usually obtained. The Mad River record was divided into 11 strata based on 3 stage classes separated at 1.2 and 1.8 meters. The 62 observations were optimally allocated to the strata by Neyman's method (Cochran, 1963). The results reduce the standard error of the total to 41% of that for systematic sampling.

The SALT sampling scheme also uses knowledge of population structure to reduce variance (Thomas, 1985). It is a variable probability scheme relying on a function of stage to make real-time decisions to enhance sampling probabilities during periods of high discharge. SALT samples of 62 specimens using a sediment rating curve of 100 water discharge/concentration pairs collected before the 31-day period give a standard error of estimate of about 845 tonnes. Again, this is an improvement over systematic sampling and is superior to SRS.

These comparisons are illustrative; the magnitudes, both relative and absolute, may not be uniform for all situations. This is especially true for the STRS and SALT schemes. The performance of those methods is heavily dependent on how the population is partitioned and how well sediment concentration can be predicted by stage. There are other ways to optimize both the STRS and SALT schemes, however, so these methods are likely to have lower variances than either systematic or SRS methods for most situations.

THE EFFECTS OF SAMPLE SIZE ON SYSTEMATIC SAMPLING VARIANCE

We now focus on the effects of changing sample size on the variance of systematic samples. As noted, the true variance of systematic sampling does not always respond as expected to changes in sample size. To see this for the Mad River data, the "true" sampling variances were plotted for 400 values of k (Figure 4). Values of k were used instead of n because k does not always divide N evenly so that n may not be unique, especially for large values of k. Lower values of k are essentially inversely proportional to those of n.

For these data, smaller values of k mean that the variance *generally* drops. Locally, however, reducing k can result in increased variance. For example,

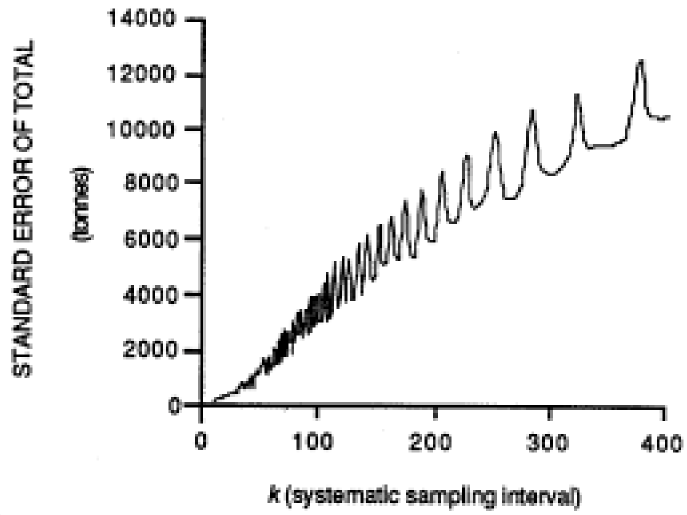

Figure 4 - ***Standard errors for estimates of suspended sediment load using systematic sampling for sampling intervals*****, k,** ***from 1 to 400 (*****n** ***equals 4450 to 11). The true" standard errors were calculated from the Mad River population***.

increasing the sample size 12 percent from 58 to 65 *increases* the standard error about 47 percent. The global pattern is expected; as k drops (and n increases), sampling variation must fall as n approaches the entire population (when k=1, n=N, and equation 1 is just the population total). Also, changes in the standard error are smaller as n becomes larger. Local behavior is more complex and depends on the interaction of the "grid" spacing of the systematic sample with the patterns in the particular record of suspended sediment flux. This behavior depends on the specific circumstances and cannot be predicted unless the complete population is known. Therefore, it is difficult to select the sample size for specified performance of general systematic schemes.

Some patterns of sequential populations produce more predictable relationships between systematic sample size and variance. If a population correlogram is concave upward, increased sample size always results in lower variance (Cochran, 1946). The correlogram for the Mad River data was concave upward over only part of its range, however, so this result does not hold (Figure 4).

A finite sampled population of sediment flux taken at short equal intervals from a continuous sedigraph is really a systematic "sample" of the target population. It has a logically similar total to the continuous target population, and the generally low variance of closely-spaced systematic samples supports using this method to define sampled populations. This kind of systematic sample can therefore be used to define the sampled population of suspended sediment discharge which can in turn be sampled by more efficient finite population schemes.

SUMMARY

Fixed-interval or systematic samples are widely used to collect data to estimate suspended sediment loads; a fact partly due to the convenience of pumping samplers. Systematic samples are inefficient, however, since evenly spaced sampling conflicts with the sporadic nature of suspended sediment populations. Several factors should guide the use of systematic sampling for estimating suspended sediment loads:

- Systematic samples with random starts give unbiased estimates of total loads.

- Variances of these estimates are generally low, but cannot be estimated from samples.

- Variances of systematic samples do not always drop with increased sample size.

- Stratified and variable probability sampling schemes are more efficient for sampling suspended sediment populations.

- Systematic "sampling" is best used for defining sampled populations for sampling by other finite population sampling schemes.

REFERENCES

Cochran, W.G., 1946, Relative accuracy of systematic and stratified random samples for a certain class of populations. Annals of Mathematical Statistics, 17, 164-177.

Cochran, W.G., 1963, Sampling techniques. Wiley and Sons, New York, pp. 413.

Ferguson, R.I., 1986, River loads underestimated by rating curves. Water Resour. Res., 22, 74-76.

Ferguson, R.I., 1987, Accuracy and precision of methods for estimating river loads. Earth Surf. Processes and Landforms, 12, 95-104.

Gurnell, A.M., and Fenn, C.R., 1984, Box-Jenkins transfer function models applied to suspended sediment concentration-discharge relationships in a proglacial stream. Arctic and Alpine Res., 16, 93-106.

Raj, D., 1968, Sampling Theory. McGraw-Hill, New York, pp. 302.

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

Thomas, R.B., 1988, Monitoring baseline suspended sediment in forested basins: the effects of sampling on suspended sediment rating curves. Hyd. Sciences Jour., 33, 499-514.

Walling, D.E., 1977a, Assessing the accuracy of suspended sediment rating curves for a small basin. Water Resour. Res., 13, 531-538.

Walling, D.E., 1977b, Limitations of the rating-curve technique for estimating suspended sediment loads, with particular reference to British rivers. In: Erosion and Solid Matter Transport in Inland Waters (Proc. Paris Symp., July 1977), 34-38. IAHS Publ. no. 122.

Walling, D.E., and Webb, B.W., 1981, The reliability of suspended sediment load data. In: Erosion and Transport Measurement (Proc. Florence Symp., June 1981), 177-194. IAHS Publ. no. 133.

Walling, D.E., and Webb, B.W., 1988, The reliability of rating-curve estimates of suspended sediment load: some further comments. In: Sediment Budgets (Proc. Porto Alegre Symp., December 1988), 337-350. IAHS Publ. no. 174.

PROCEEDINGS

OF THE

FIFTH FEDERAL INTERAGENCY

SEDIMENTATION CONFERENCE

MARCH 18-21, 1991

LAS VEGAS, NEVADA

Edited by: Drs. Shou-Shan Fan and Yung-Huang Kuo

Interagency Advisory Committee on Water Data
Subcommittee on Sedimentation

Agricultural Research Service (ARS)
Bureau of Land Management (BLM)
Bureau of Mines (BOM)
Bureau of Reclamation (USBR)
Corps of Engineers (COE)
Council on Environmental Quality (CEQ)
Department of Housing and Urban Development (HUD)
Environmental Protection Agency (EPA)
Federal Energy Regulatory Commission (FERC)
Federal Highway Administration (FHWA)
Forest Service (USFS)
Geological Survey (USGS)
National Oceanic and Atmospheric Administration (NOAA)
National Park Service (NPS)
Nuclear Regulatory Commission (NRC)
Office of Surface Mining Reclamation and Enforcement (OSMRE)
Soil Conservation. Service (SCS)
Tennessee Valley Authority (TVA)

Printed by

Federal Energy Regulatory Commission