ESTIMATION OF POLLUTANT LOADS IN RIVERS AND STREAMS: A GUIDANCE DOCUMENT FOR NPS PROGRAMS

R. Peter Richards
Water Quality Laboratory
Heidelberg College
310 East Market Street
Tiffin, OH 44883

Prepared under Grant X998397-01-0
U.S. Environmental Protection Agency, Region VIII
David Rathke, Project Officer

TABLE OF CONTENTS

INTRODUCTION

The Loading Concept

The basic concept of the pollutant load of a river or stream is deceptively simple. The pollutant **load** is the mass or weight of pollutant which passes a cross-section of the river in a specific amount of time. Loads are expressed in mass units (e.g. tons, kilograms), but the interval of time over which the load occurs is always implicit and should be clear from context. A related concept is that of **discharge**, which is the volume of water which passes a cross-section of the river in a specific amount of time. Discharge has units of volume, usually cubic meters or cubic feet.

The loading rate, or **flux**, is the instantaneous rate at which the load is passing a point of reference on a river, such as a sampling station, and has units of mass/time such as grams/second or tons/day. The discharge rate, or **flow**, is the instantaneous rate at which water is passing the reference point, and has units of volume/time such as cubic feet per second.

If we could directly and continuously measure the flux of a pollutant in a typical river or stream, we would see that the flux changes continuously with time. A hypothetical example is shown in Figure 1. The load for the period of time in the graph would be equal to the area under the curve.

Mathematically, the load is the integral over time of the flux:

$$\text{Load} = \int_t \text{flux}(t)\,dt \qquad (1)$$

where k is a constant which handles the conversion of units from, for example, mg/L and ft^3/sec to tons/year.

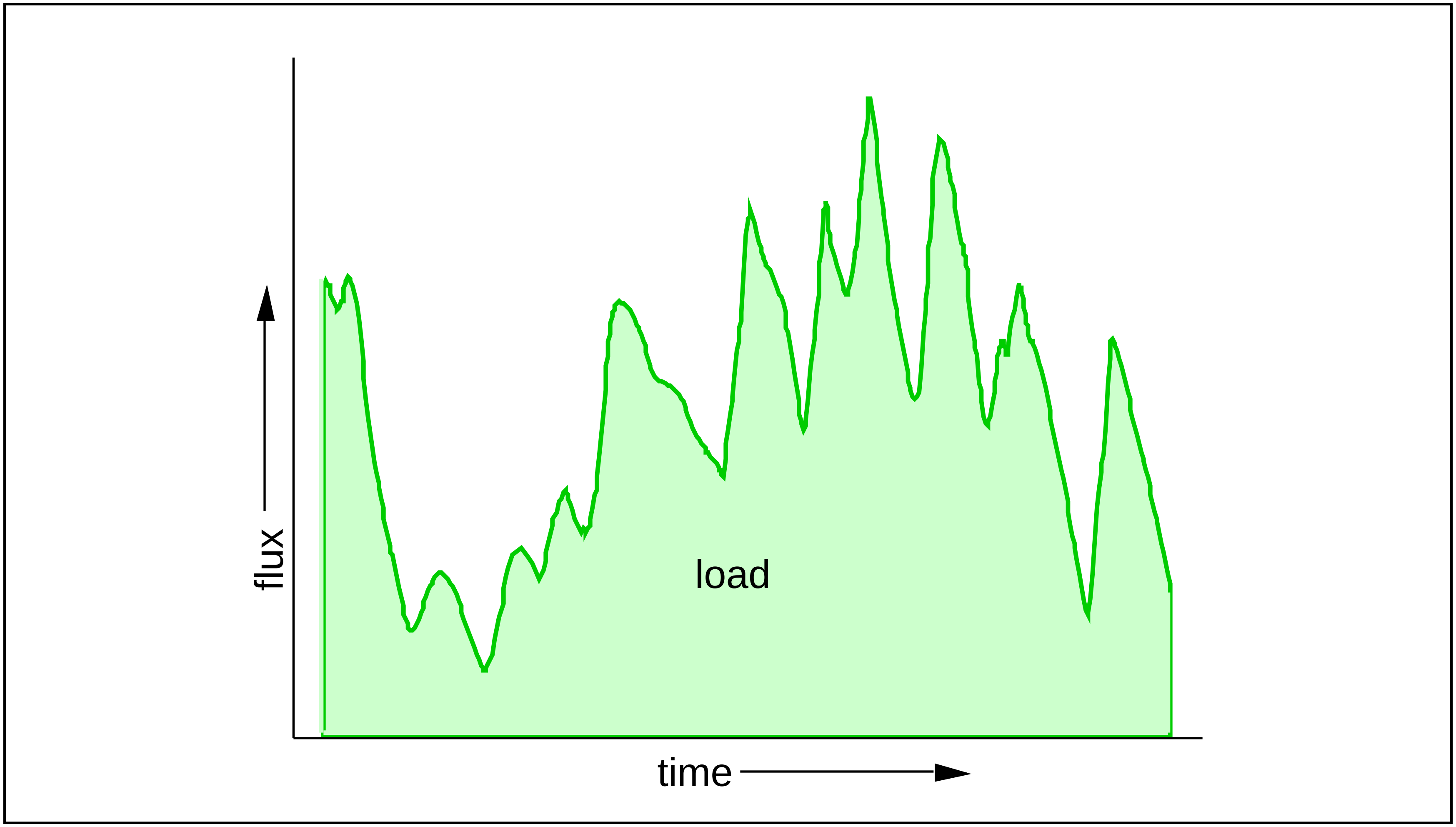

Figure 1. A hypothetical graph of flux over time. The area under the curve is the load for the time interval.

There are several problems with the practical application of this concept of the load. The first is that we cannot measure flux directly. Instead, we measure flux as the product of concentration and flow.

The expression for the load then becomes:

$$\text{Load} = k\int_t c(t)q(t)dt \qquad (2)$$

where c(t) is concentration and q(t) is flow, both continuous functions of time.

The second problem is that, while concentration and flow are both continuous functions of time, we cannot measure them continuously. Thus the integral which is the load must be estimated by summing the products of a sequence of discrete measurements of concentration and flow:

$$\text{Load} = k\sum_{i=1}^{n} c_i q_i Dt \qquad (3)$$

where c_i is the ith observation of concentration, q_i is the corresponding observation of flow, and Δt is the interval between observations. If the observations are not equally spaced, the load is given by:

$$\text{Load} = k\sum_{i=1}^{n} c_i q_i t_i \qquad (4)$$

where t_i is the time interval represented by the ith sample. When t_i is given by $\frac{1}{2}(t_{i+1} - t_{i-1})$, this formula is equivalent to the trapezoidal rule of numeric integration.

Especially for particulate pollutants of non-point origin, the flux varies drastically over time, with fluxes during snowmelt and storm runoff events often several orders of magnitude greater than those during low flow periods. It is not uncommon for 80 to 90% or more of the annual load to be delivered during the 10% of the time with the highest fluxes, as is illustrated in Table 1. Clearly it is critical to sample during these periods, if an accurate load estimate is to be obtained.

Table 1. Delivery of pollutants in Lake Erie tributaries, during selected periods of time with the highest fluxes. Data from Baker (1988).

Parameter	Percent of time	Percent of annual load delivered				
		Raisin	Maumee	Sandusky	Rock	Cuyahoga
Suspended solids	0.5	17.8	17.3	24.3	42.7	28.3
	1.0	26.9	27.1	36.4	59.5	38.1
	10.0	79.6	81.6	87.7	97.6	81.5
	20.0	91.2	93.9	95.4	99.0	90.9
Total Phosphorus	0.5	14.8	9.8	14.8	30.9	13.2
	1.0	23.9	17.2	22.8	47.2	18.0
	10.0	67.9	67.5	77.3	93.9	51.3
	20.0	81.3	85.6	90.2	97.2	64.8
Nitrate + Nitrite	0.5	5.3	5.0	6.9	17.9	3.0
	1.0	9.5	8.7	12.3	28.8	5.2
	10.0	54.2	52.2	56.7	81.0	25.9
	20.0	76.4	75.4	77.2	91.4	40.5

Several facets of the problem of load estimation are represented graphically in Figures 2 to 4. Figure 2 compares flux time series for the Grand River (Michigan) and for the Sandusky River (Ohio). Successive panels show the time series with observations at daily, weekly, and monthly intervals. In each case, the area under the plot is an estimate of the load. Clearly, the quality of the load estimate decreases as the time resolution of the data becomes poorer. Figure 3 shows the daily time series for both rivers with flux plotted on a log axis. With the data represented in this way, the much greater variability of the Sandusky River data is apparent. In Figure 4, the weekly and monthly time series for the Sandusky River are superimposed on the daily series to facilitate comparison. Areas between the two series represent contributions to the error of the load estimate based on the monthly series, as compared to the load estimate from the daily series.

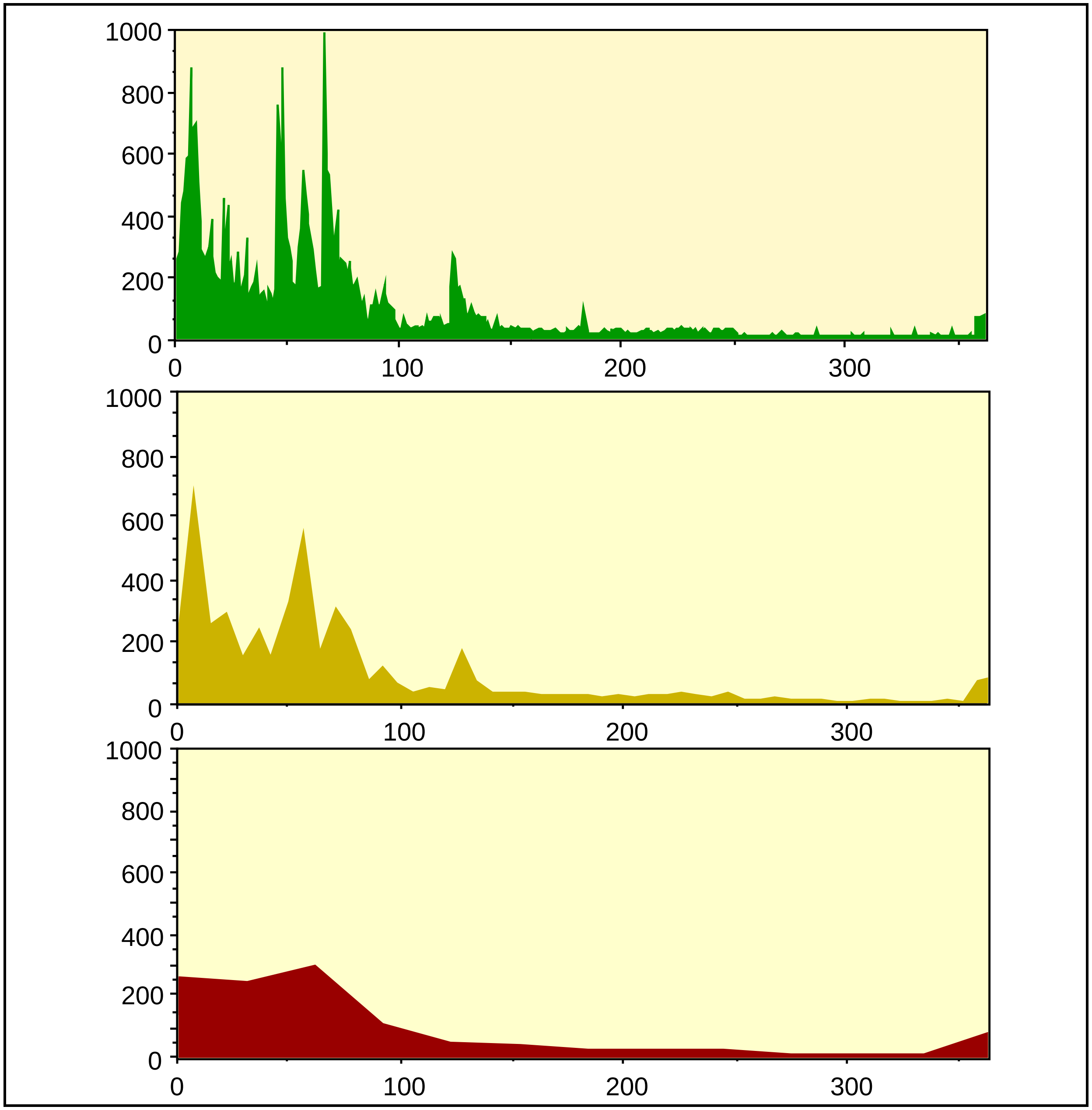

Figure 2a. Loads for the Grand River for the year beginning March 1, 1976 and ending February 28, 1977. Top: daily samples; middle: weekly samples; bottom: monthly samples.

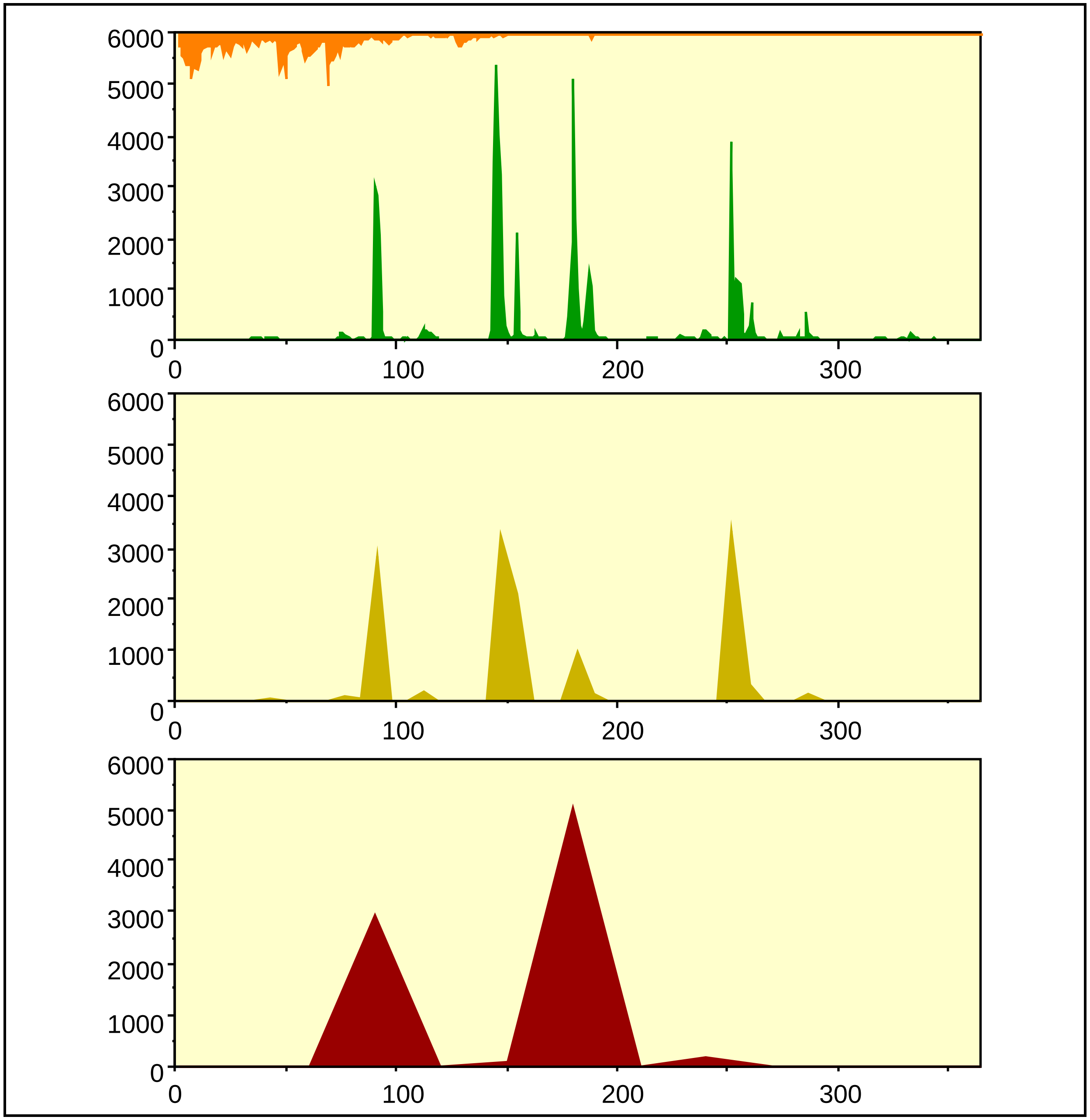

Figure 2b. Loads for the Sandusky River, for the 1985 water year (10/1/84 to 10/1/85). Top: daily samples; middle: weekly samples; bottom: monthly samples. The daily load series for the Grand River is reproduced, inverted and at the same scale as the Sandusky data, along the top of the top panel.

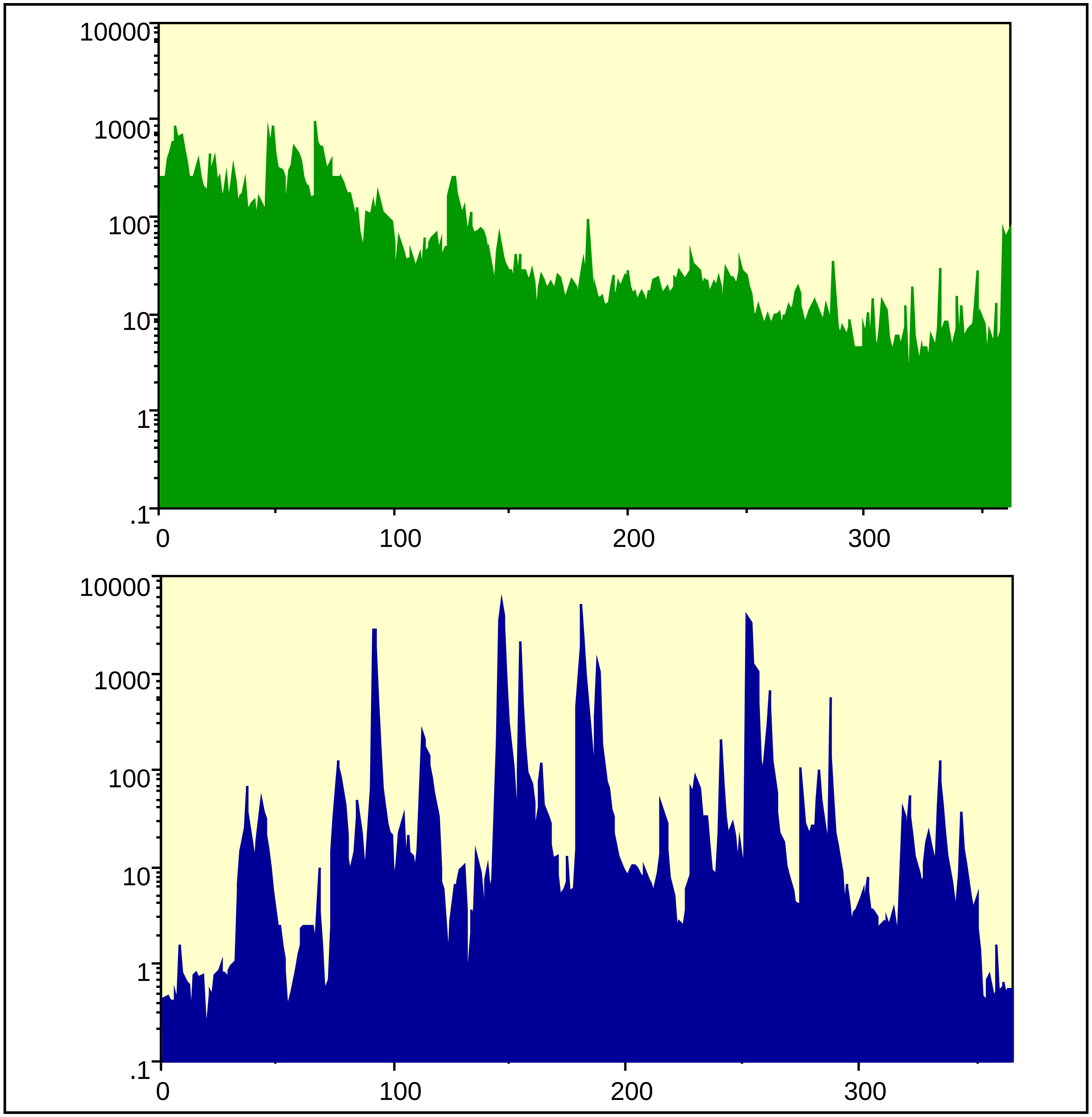

Figure 3. Daily load time series for the Grand River (top) and the Sandusky River (bottom), with the loads plotted on a logarithmic axis, illustrating the difference in variability between the two systems.

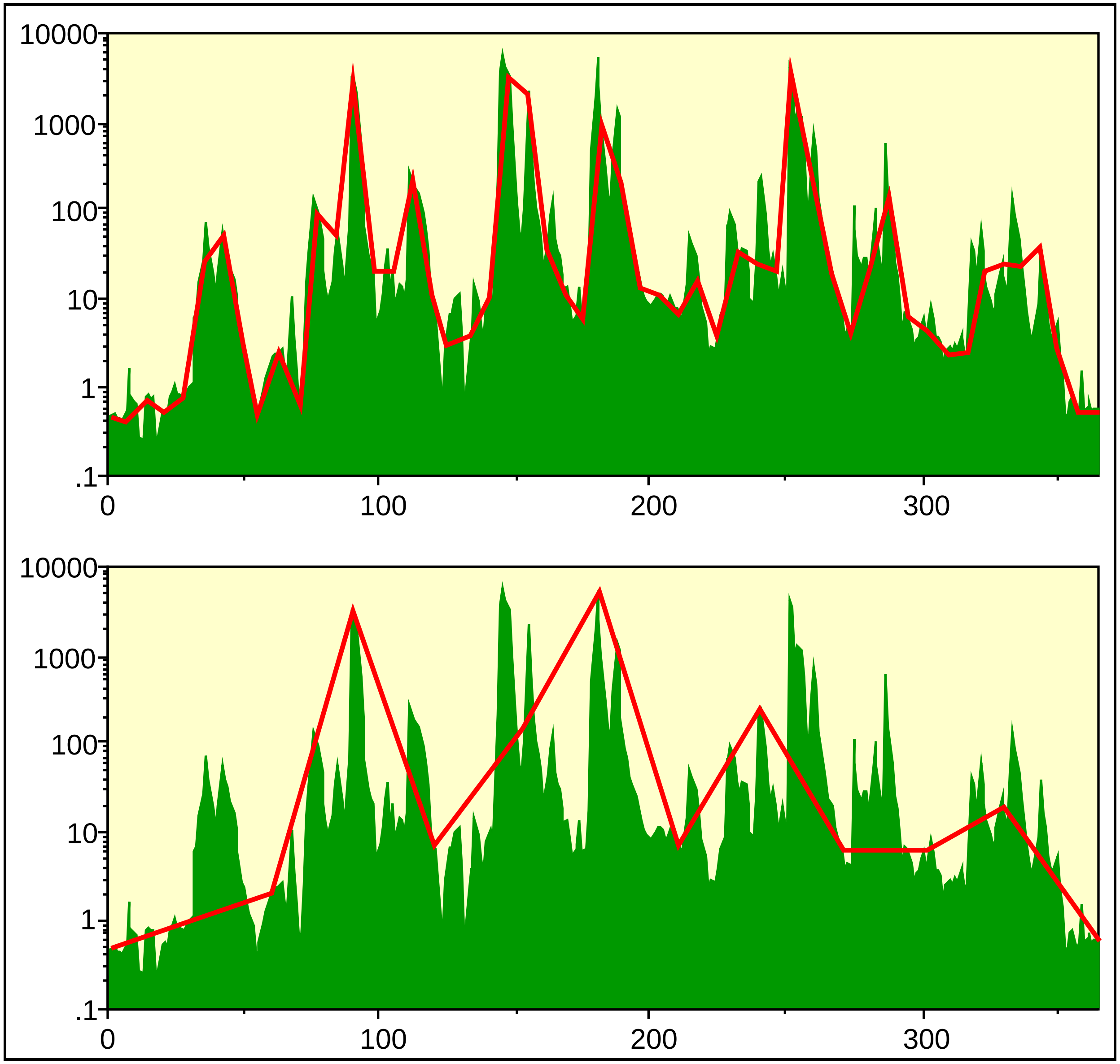

Figure 4. Weekly (top) and monthly (bottom) load time series superimposed on the daily load time series to facilitate comparison. The weekly series captures much of the information contained in the daily series, but the monthly series does much more poorly. In fact, this particular monthly series does much better than most, because it includes the peak flux associated with two major runoff events.

This particular monthly series, while only a crude representation of the daily series, is probably better than most, because it happens to include the peaks of two of the four major storms for the year. A monthly series based on dates about 10 days later than these would have included practically no storm observations, and would have seriously underestimated the load.

It is clear from these figures that many samples will be needed to accurately and reliably capture the true load. Monthly observations will probably not yield reliable load estimates, and even weekly observations may be less reliable than would be wished. There is often a conflict between the number of observations a program can afford and the number needed to obtain an accurate and reliable load estimate.

For this reason, anything which can be done to use other information to make estimates for the intervals between observations will be very important. The information usually used is flow information.

Flow is usually determined on a routine basis by measuring the stage, or water height, and using a previously established **rating curve** to convert stage into flow. Once a gaging station is established, stage measurements are made inexpensively using automated equipment, and converted to flow by computer programs. Hence flow measurements are relatively inexpensive to make and are available on an almost continuous basis, i.e. they are made sufficiently frequently that major changes in flow between measurements will not occur. Many rivers and larger streams have established gaging stations, often maintained by the U.S. Geological Survey, and these stations are often used for load estimation because the stage/flow relationships are already established and a prior record of flow patterns is available. Stage measurements are typically made at hourly intervals at U.S. Geological Survey stations, or at quarter-hourly intervals at stations on smaller rivers and streams, where the flow can change more quickly in response to storm runoff or other causes. The measurement of flow is a well-established science, and many excellent books and reports are available to describe the methods involved. Measuring flow in smaller streams may require different methodologies than are used on larger streams and rivers.

Concentration measurements can range in cost from a few dollars to more than a thousand dollars per sample, depending on what parameter or parameters are being measured. Obtaining concentration measurements usually involves taking water samples to a laboratory for chemical analysis, though some parameters can be measured directly in the stream by ion-specific electrodes. Such in-stream measurements can be made with the same frequency as flow measurements, and are inexpensive to make after the initial investment in equipment. However, the list of parameters that can be measured in this way is short, and most of the parameters of interest to NPS programs not included. Nor is in-stream measurement the worry-free come-back-in-a-month-and-get-your-data approach it might appear to be. Aside from basic malfunctions which lead to lost observations, there are problems with fouling of the sensor(s) by algae, bacteria, sediment and/or detritus, which may lead to biased results which "look O.K." and are therefore misleading but almost impossible to detect unless discrete samples have also been periodically taken for lab analysis. Also, the response of the sensors changes with time, with the result that the measured concentrations may be inaccurate unless frequent calibration checks are performed or less-frequent lab analyses are used to adjust the sensor results and compensate for the drift.

Analytical chemistry is part science and part art. A number of different techniques can be used to measure the same chemical, with differing degrees of precision and accuracy. The quality of the results is strongly affected by the skill and care of the analyst. Detailing specific analytical techniques is far beyond the scope of this document. Analytical techniques which represent the best currently available technology are described, sometimes in excruciating detail and contorted prose, in the EPA methods series and books such as the Standard Methods series (e.g. Greenberg *et al.*, 1992)).

Because most parameters of interest to NPS programs require samples analyzed in the lab, and because these are much more expensive than flow measurements, it is almost always true that chemical observations are available less frequently than flow observations. This creates the basic problem of practical load estimation. To do the summation described above, one has three choices of basic approach:

1. Find a way to estimate "missing" concentrations: i.e. concentrations to go with the flows observed at times when chemical samples were not taken.

2. Abandon most of the flow data and calculate the load using the concentration data and just those flows which were observed at the same time the samples were taken.

3. Do something in between - find some way to use the more detailed knowledge of flow to adjust the load estimated from matched pairs of concentration and flow.

The second approach is usually totally unsatisfactory because the frequency of chemical observations is inadequate to lead to a reliable load estimate when simple summation is used. Thus almost all of the load estimation approaches which have been shown to give good results are variants of approaches 1 or 3.

Total load vs. unit load

If all of the flow and concentration observations are available for all n intervals in formula (3) or (4) above, the summation is an easy task. The summation is the **total load** for the time period of interest; each individual product, $c_i q_i \Delta t$ or $c_i q_i t_i$, could be called the **unit load**. If the total load is an annual load, the unit load might be the daily load. If the total load is a weekly load, the unit load might be the hourly load.

This distinction is important, because we tend to focus on the total load as our goal. The considerable complexities of the various methods for calculating loads, however, are almost always related to trying to accurately characterize the unit loads. Once this is done, the total load is easily obtained. Two examples will briefly illustrate.

One approach to load estimation develops a regression relationship between concentration and flow, based on those intervals during which both flow and concentration were measured. This regression relationship is then used to estimate concentration for the other intervals. Unit loads are then calculated for each interval, using either the observed or the estimated concentrations. The total load is the sum of the (now complete) set of unit loads.

The other approach calculates the average unit load, based on the intervals when flow and concentration were measured, and adjusts the average unit load using the ratio between the average flow for all intervals and that for intervals when concentrations were measured. The total load is the average unit load times the number of intervals, e.g. the average daily load times 365.

How many intervals are there in a continuous process like the flux in a river over the course of a year? How do we decide whether the unit load should be a weekly, daily, hourly load? These are not easy questions to answer, and there is no single right answer. Often, though, an answer is imposed on us from outside. For example, if what is available to us is mean daily flows, then of necessity the interval load is daily (or at least cannot be more frequent than daily) because we do not have the flow data available on a more frequent basis. If the load is to be calculated by a program which uses daily flows, we cannot decide to use weekly intervals.

In a sense, our goal is to be sure that we did not miss any surprises in our sampling. Fortunately, water and pollutant runoff and other watershed processes have a kind of temporal inertia (expressed by the mathematical concept of autocorrelation) - the flux at a given spot today is related to the flux there yesterday, more weakly related to the flux two days ago, and more strongly related to the flux an hour ago. This inertia means that, given a sufficiently short time interval, no radical change in the flux will occur between observations. In particular, we won't completely miss either a sudden peak in flux or a sudden minimum, thereby obtaining a distorted total load estimate. Beyond a certain point, sampling more and more frequently would do us little good. The reality, however, is that sampling programs rarely can afford to take enough samples to reach this point of diminished returns.

We generally cannot observe the flux record directly to examine this temporal inertia or autocorrelation, but we can observe the flow or stage record. The absence of sudden jumps in flow, and the presence of frequent successive flow observations which are identical gives good assurance that the flow interval is frequent enough to be well within the region of strong autocorrelation, and more than adequate to avoid undetected surprises.

Another issue in choosing the unit load interval is that the more intervals there are contributing to the annual load, the smaller the distortion which is imposed by a non-representative unit load. Thus we would feel more comfortable estimating monthly loads using daily loads as units rather than weekly loads. For estimating an annual load, weekly unit loads might be satisfactory, depending on the behavior of the river involved and the confidence we wish to place on the load estimate.

These general perspectives will be put on a firmer footing in the following sections. For the moment it is sufficient to understand the distinction between total loads and unit loads, and to realize that the problem of load estimation is the problem of adequately characterizing the unit loads. Adding them up is easy.

Working Assumptions for This Document

It is assumed that the total load of interest is the annual load. Monthly or other uniform interval loads and storm event loads may be calculated by appropriate application of the same methods, but the reliability of the methods for calculating loads over these shorter periods of time has not been well evaluated.

It is assumed that the pollutants of interest are primarily of non-point source origin. The methods described below can be used in some circumstances for the estimation of loads from point sources as well, but with some differences in the details of their application.

It is assumed that flow is well characterized. Ideally, instantaneous flow should be available at the time of each chemical sample, and mean daily flows should be available for all days. In many monitoring programs, the chemical sampling is carried out at U.S. Geological Survey gaging stations, which assures that both instantaneous and mean daily flow values will be available. If flow must be monitored by the project, provision should be made for a continuous flow monitor, usually via stage or hydrostatic pressure, and automated recording of values at least hourly.

It is assumed that chemical samples will be available less frequently than flow data, for example that daily flows are available but only 25 chemical samples per year. If flow data and chemical data are available at the same frequency, and particularly if the frequency is high (>100 samples per year), the relatively straight-forward numeric integration approach of El-Shaarawi *et al.* (1986) may be as satisfactory as the approaches described later in this document.

Requirements of a Practical Approach to Load Estimation

The literature on load estimation contains a number of papers reporting very elegant studies in which the details of transport of a pollutant in a river were carefully worked out and incorporated into a very precise load estimation model for that river. Other papers describe elaborate computer-driven sampling procedures which compensate for autocorrelation or other problematic aspects of pollutant transport in rivers. While such studies are valuable contributions to our knowledge of load estimation, they are often difficult and expensive to implement, and sometimes require continuous evaluation and modification to provide accurate and precise results.

In the context of non-point source programs, where load estimation is merely one part of a broader agenda, such specialized approaches are economically unfeasible, and a more general approach is needed which can be relatively easily adapted to the needs of each project. To be broadly useful, such an approach should include the following attributes.

High Precision and Accuracy, i.e. Efficiency

Precision and **accuracy** measure two related but different aspects of the behavior of a measurement system. If repeated measurements are made of an object, the measurement process is called precise if the difference among measurements is small, and it is called accurate if the average measurement is close to the true value. **Bias** is the lack of accuracy; a measurement system which is unbiased is highly accurate. Load estimation approaches which produce low bias and high precision using relatively few samples are described as **efficient**.

Generally we want an approach to load estimation which is as precise and accurate as possible for the number of samples taken, i.e. as efficient as possible. It is especially important to have *a priori* confidence in the lack of bias, since bias cannot be directly evaluated. However, when detecting a change in loads is more important than the actual level of the loads, we may choose a method which produces precise estimates, even if they are biased. Walling and Webb (1981) showed in a simulation study that the product of annual discharge and average concentration was strongly biased but quite precise, and pointed out that it might be useful for trend studies, in spite of the bias.

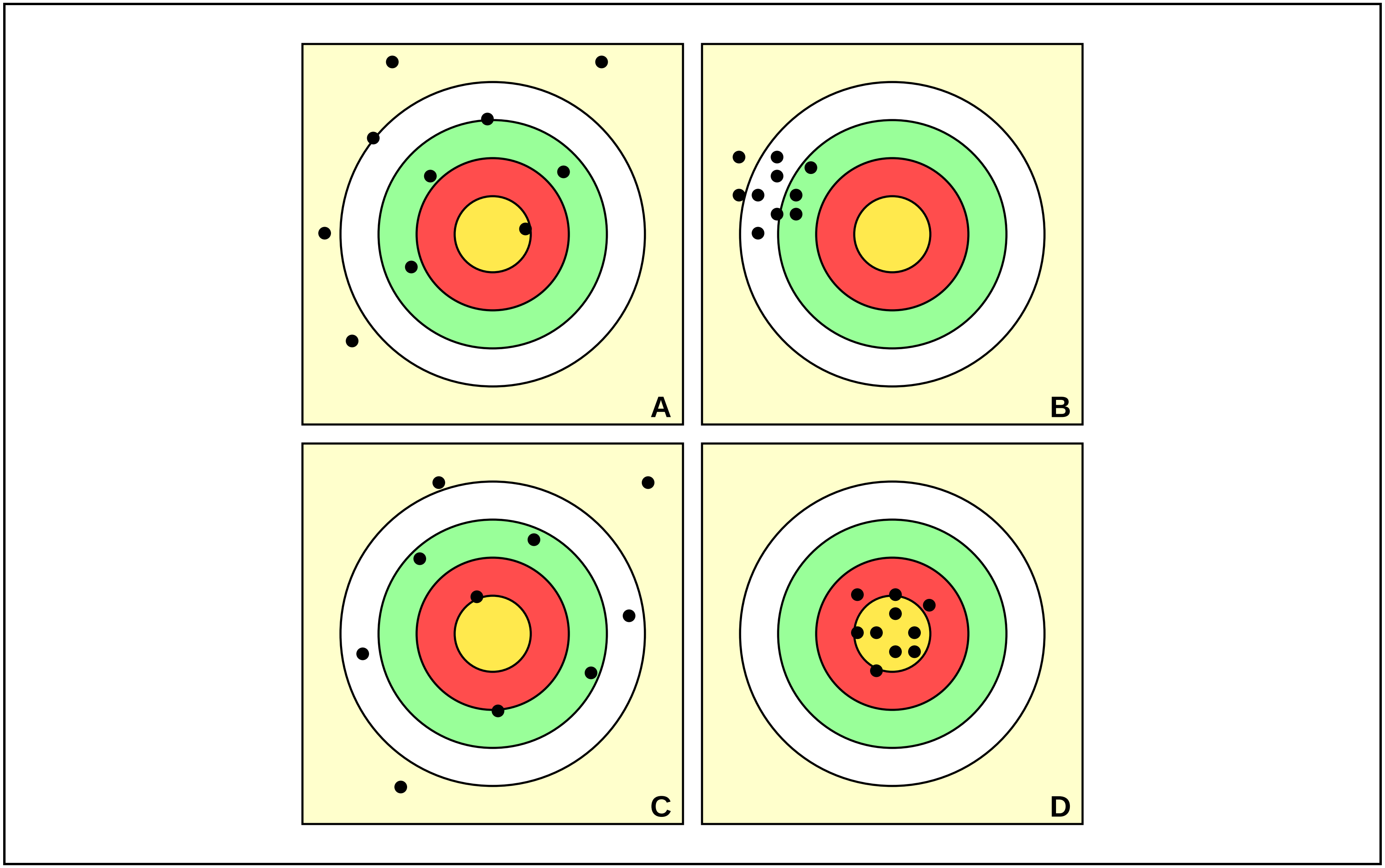

Figure 5. Precision and accuracy. A: Imprecise and inaccurate. B: Precise but inaccurate. C: Accurate but imprecise. D: Accurate and precise.

Robustness

To the extent possible, the approach should be insensitive to the attributes of the data which may be gathered. One approach may be precise, accurate, and powerful when its statistical assumptions are met, but imprecise and biased otherwise. Such an approach is often less useful than an alternative which is always moderately precise, accurate, and powerful, even if it is not always the best approach. The method which is always pretty good is better than the one which is sometimes best and sometimes very bad.

Ease of Use

In projects with main goals other than load estimation, specific expertise in sampling theory and load estimation are not likely to be available. Hence a method is only useful if it can be implemented successfully by researchers without special training.

Broad Applicability

A useful method should be applicable in different climates, on different rivers with dissimilar flow characteristics, and for different chemical parameters. Some specialized approaches can only be applied to one river and one parameter. While such approaches may be justified for certain large rivers and priority pollutants, they cannot be applied to other problems without extensive modification. This lack of transferability is a disadvantage in the context addressed by this document.

Objectivity

Some methods rely on "best professional judgment" to supplement monitored data or to determine patterns of stratification prior to calculation of loads. Ideally, such subjective components should be avoided, and a specific method should arrive at the same result for a given set of data, no matter who processes the data.

Elements of a Load Estimation Program: The Planning Process

All too frequently the decision to calculate loads is made after the data are collected, often using data collected for other purposes. While the quality of the load estimate will still depend on the proper application of the most appropriate method, little can be done to compensate for a data set which contains an insufficient number of observations collected using an inappropriate sampling design.

Many programs choose monthly or quarterly sampling with no better rationale than that it's convenient and it's what other programs did. A simulation study for some Great Lakes tributaries revealed that data from a monthly sampling program, combined with a simple load estimation procedure, gave load estimates which were biased low by 35% or more 50% of the time (Richards and Holloway, 1987). Many of the Rural Clean Water Program projects, and other similar programs (Model Implementation Program, Hydrologic Unit Area, and Management System Evaluation Areas) were unable to document water quality benefits of land management changes because of insufficient sampling (Gale *et al.*, 1992).

To avoid such problems, the sampling which will be needed for load estimation must be established **in the initial planning process**, based on quantitative statements of the precision required for the load estimate. The resources necessary to carry out the sampling program must be known, and budgeted for, **from the beginning**. The planning process involves at least the following steps.

• Determine whether the project goals require load estimation, or whether they can be met using concentration data. In many cases, especially when trend detection is the goal, concentration data may be easier to work with and may provide a more powerful approach.

• If load estimates are to be made, determine the precision needed, based on the uses to which they will be put. Such a determination usually is of the form, "I want the load estimates to be within x% (or within y tons) of the true loads in z% of the years for which calculations will be made.

• Decide what approach will be used to calculate the loads, based on known or expected attributes of the data.

• Use the precision goals to calculate the sampling requirements for the monitoring program. Sampling requirements include both the total number of samples and, possibly, the distribution of the samples with respect to some auxiliary variable such as time or flow.

• Implement the sampling program with appropriate scientific skepticism, perhaps doing additional sampling to check assumptions. For example, if the sampling program is daily at the same time each day, we are implicitly assuming that all times of day are equivalent and thus that there is no diurnal periodicity in concentration or flow. Sampling hourly for a few days, including a weekday and a weekend day, is a wise check of this assumption.

• Calculate the loads based on the samples obtained, and compare the precision estimates with the initial goals of the program. Adjust the sampling program if the results deviate strongly from the goals.

ASPECTS OF LOAD ESTIMATION: AN OVERVIEW

A review of the many papers on load estimation suggests that, even within the context of loads primarily of non-point source origin, major differences exist between the approaches and needs of load estimation exercises in urban and rural settings.

Load estimation in the urban environment usually takes place in the context of planning and is done using models. Often it is not the load for a specific interval of time which is of interest, but the typical annual load or some other planning estimate such as a "unit storm load". Monitoring is used primarily for model calibration, and perhaps secondarily for ongoing model validation. Between-storm accumulation on impervious surfaces is the main source of non-point pollutants, consequently the time between storms is an important factor in many models. The parameters of interest include sediment, nutrients, COD/BOD, bacteria, metals, and organics. Acute toxicity is often an issue in the receiving waters, and concentrations may be more important in this context than loads, especially for flowing waters.

In the rural environment, loads are typically dominated by non-point contributions, especially from agriculture, although construction and logging may be locally or regionally more important. Load estimation may be model-driven, particularly for evaluating management scenarios at the field or plot scale. Models generally do not work reliably at the watershed scale, however, or the data demands for a successful model are prohibitive. The scale of interest is often larger and more complex than in the urban setting. For BMP evaluation in particular, modeling is often inadequate because of mixed land uses which change annually and are characterized by different loading rates, and because of the need to document success in reducing loads. Consequently load estimation tends to be based on monitoring data. The main source of pollutants is often a somewhat uniform mass of soil which erodes during storms; the time since the last storm is less important than it is in the urban environment. The pollutants of concern are sediment, nutrients, pesticides, bacteria; others are less important in most situations. Habitat modification, eutrophication, and chronic effects are generally more important than acute effects.

This document is primarily about rural loadings, and describes techniques traditionally applied in rural or mixed land-use settings, though the approaches could be applied in the urban setting.

Sampling Patterns

Sampling pattern issues include whether observations should be taken at random or systematically from the entire population, or whether the population is better divided into sub-populations, each of which may be sampled with a different pattern. In load estimation, the population of daily (or other interval) loads unfolds over time, which adds further complications. Furthermore, because pollutant transport is a continuous process, a sample is not a well-defined entity, as is the case when drawing marbles from a bag or measuring the weights of a randomly selected group of people.

<u>Point Samples vs. Integrated Samples</u>: Point (or instantaneous) samples are taken at an instant in time, or over a short enough time interval that they can be considered to represent a discrete point in time. Because point samples are analogous to samples from a population of discrete entities, the principles of sampling theory and other branches of statistics can be applied to them, albeit sometimes awkwardly. It is less clear how these principles apply to integrated samples. As a consequence, most approaches to load calculation assume that the samples are instantaneous.

By contrast, integrated or composite samples are taken over a finite interval of time. Integrated samples have the advantage that each physical sample captures more of the changes which may occur during the period of time it represents than is true with point samples at the same frequency. Thus a more representative result is gained for the money spent on the analysis of the sample. However, if integrated sampling is used to reduce analytical costs, for example by compositing daily samples for a week and analyzing the composite sample rather than analyzing the daily samples, a whole week's data is lost if a sample is corrupted by missing an aliquot, or lost due to spillage or failure of the analytical system. Detailed information is also lost: a composite sample provides an average concentration over the compositing interval, but the concentrations of each aliquot are, of course, unknown.

Integrated samples may be collected continuously, e.g. with a low-volume pump running constantly, or they may be composited of a number of instantaneous aliquots placed in the same container, e.g. 10 milliliters every half hour for a day. Two kinds of integrated samples are most common: flow-proportional samples and time-proportional samples.

Time-proportional or time-integrated samples are either taken at a constant rate over the time period or are composed of aliquots taken at a fixed frequency. Time-integrated samples are not very well suited for load estimation because they ignore changes in flow which may occur during the integration period. Shih *et al.* (1992) provide estimates of the bias involved in using time composited samples for load estimation and offer formulas for correcting this bias. Still, they conclude that eight time-integrated samples are required per runoff event to obtain a good load estimate, and it is unclear that compositing holds any advantage over instantaneous samples at such a high frequency.

Flow-proportional samples are ideally suited for load estimation, and in principle should provide a precise and accurate load estimate if the entire time interval is properly sampled. Unfortunately, flow-proportional sampling is relatively difficult to implement, though autosamplers capable of flow-proportional sampling are commercially available and have been used successfully in many projects, and "passive samplers" can be used in some applications. Flow-proportional sampling usually requires a "smart sampler" which can monitor the stage, convert it into flow, and either vary the pumping rate to match changes in flow rate or collect an aliquot of sample every time a fixed amount of water has passed the sampling point. Since the rate of sample accumulation varies with flow, it is not possible to know how soon the sample bottle(s) will be filled, so some form of communication with the sampler is needed, if the sampler is in a remote site. It is also impossible to know how many samples will be collected in a given year, making budgeting difficult, though the analytical costs should be lower than those required to provide a load estimate of the same precision using (many more) point samples.

<u>Random vs. Systematic Sampling:</u> Random sampling is a process which is intended to insure that every member of the population has an equal probability of being drawn into the sample, and that the

observations in the sample are independent. In systematic sampling, the population is usually ordered by some means (index cards in a box, an alphabetical list, or daily loads listed in calendar order), and the observations are drawn systematically (e.g. every 10th observation starting with the 4th). When, as is true of daily loads, the population is arrayed in time, random sampling is difficult (but possible), and systematic sampling is the norm (e.g. one observation per week).

Random sampling is a "safe" approach" designed to avoid inadvertent bias or false estimates of precision due to unrecognized structure in the population. However, it is not as efficient (i.e. as precise for a given sample size) as systematic sampling under certain circumstances (see Cochran, 1963, 208-213), because systematic sampling can spread the observations out over the population, providing better "coverage". Random sampling for load estimation would involve using a random process to determine on which days to sample. Note, however, that since the time series of daily loads typically is autocorrelated, randomly chosen observations closely spaced in time will not be independent, even though they were selected at random, and therefore will not contain as much information as uniformly spaced observations.

Systematic sampling is as precise as, or more precise than random sampling under ideal circumstances. However, there are pitfalls to systematic sampling which can only be avoided by knowledge of the properties of the population being sampled. These pitfalls are due to autocorrelation and periodicity in the time series of daily loads. For example, inputs of some materials to some rivers may show a weekly periodicity because they are produced primarily during the work week (or primarily on the weekend). A program which always sampled on Tuesday might give a mean daily load which was biased high or low, depending on where Tuesdays fell in the weekly cycle of pollutant output. This program would also be unrealistically precise - it would have a small confidence interval because it failed to detect the periodic fluctuations which characterized the population it was sampling. A similar problem exists with rivers which have regulated flow over a diurnal cycle and which are sampled at the same time on each sampling date. Cochran gives criteria for determining when systematic sampling is more precise than simple random sampling; unfortunately their strict application requires a level of knowledge of the population which is usually not available. A practical guideline would be that systematic sampling is as efficient as, or more efficient than simple random sampling if the sampling interval is not equal to a multiple of any strong period of fluctuation in the population. This criterion obviously requires information not revealed by the sampling program, so previous knowledge of the system or reconnaissance sampling is required before it can be applied.

The central question is whether or not sampling which is systematic in time differs from random sampling of the population of interest. If the population exhibits no important time-based structure, then the fact that sampling is systematic in time is irrelevant. For example, if we determine the mean weight of stones in a bucket by weighing examples chosen at random from the bucket, there would be no reason to expect the result to differ whether we weighed one each day for 15 days, or weighed 15 stones in succession on the same day. In such cases it would be appropriate to consider the sampling of the population to be random, and all of the statistics of random sampling would apply. If the population has important time-based properties, such as autocorrelation or periodicity, sampling which is systematic in time would not be random, and assuming so could lead to unrecognized bias and/or inappropriate precision estimates. For this reason it is important to obtain enough information to evaluate the time structure of the population before final design of a sampling program.

Simple sampling vs. stratified sampling: Under simple sampling, which may be either random or systematic, the population is sampled as a whole on the same basis. Under stratified sampling, the population is segmented on some basis, and a different sampling pattern or frequency is applied to each segment. As with simple (unstratified) sampling, stratified sampling can be either systematic or random, within strata. The advantage of stratification is that if one part of the population is much more variable than another, and therefore requires more observations to characterize precisely, a greater percent of the overall effort can be allocated to that part than would occur under simple sampling. The disadvantage of stratification is that it results in a more complex sampling program which may be more costly on a per observation basis. Thus a stratified sampling program is justified if it reduces total sampling effort or increases precision enough to offset the increased cost per observation.

The Relationship between Sampling Frequency and Precision

Intuitively, one would expect the precision of a load estimate to increase as the sampling frequency increases. For a parameter with a normal distribution, precision can be measured by the magnitude of the error which occurs with a stated probability, given the variance of the distribution. Sampling theory tells us that the magnitude of this error is inversely proportional to the square root of the number of samples. Thus, to halve the error (double the precision), one would have to take four times as many samples. Several simulation studies using real or realistic environmental data which is not normally distributed have reached a similar conclusion (Richards and Holloway, 1987; Burn, 1990; Preston *et al.*, 1992), suggesting that the relationship is general enough to be useful for planning purposes.

Methods of Load Estimation

Many different approaches have been used to calculate loads from the observed concentration and flow data. Some are more precise and/or more accurate than others; some are only appropriate under special circumstances. This brief review is intended to indicate the range of approaches which has been used, but not necessarily to recommend them for future use.

The simplest approach is direct **numeric integration**, and the total load is given by

$$\text{Load} = \sum_{i=1}^{n} c_i q_i t_i \tag{5}$$

where c_i is the concentration in the ith sample, q_i is the corresponding flow, and t_i is the time interval represented by the ith sample, given by $\frac{1}{2}(t_{i+1} - t_{i-1})$. It is not required that t_i be the same for each sample.

Numeric integration is only satisfactory if the sampling frequency is high - often on the order of 100 samples per year or more, and sufficiently frequent that all major runoff events are well sampled. Roman-Mas *et al.* (1994) suggested that a sampling frequency sufficient to obtain 20 samples over a typical runoff hydrograph was necessary in order to obtain load estimates with an error less than 5%. Yaksich and Verhoff (1983) suggest 12 samples over the hydrograph.

The **worked record** is a procedure used in some offices of the U.S. Geological Survey. Chemical observations are superimposed on a plot of the more thoroughly sampled hydrograph, and a smooth curve is drawn through these points based on the hydrologist's experience with the relationship between concentration and flow. This interpolated curve is used to estimate a representative concentration for each day (or other unit load interval), and the load is estimated by summing the daily fluxes over the year (or other total load interval). The worked record has the advantage that it allows for the possible inclusion of a peak concentration greater than the largest observed concentration. However, this approach has the serious drawback that it is subjective and relies on the experience and good judgment of the hydrologist.

Averaging approaches use, not surprisingly, some form of average in the calculation of the loads. The simplest approach involves multiplying the average concentration for some period of time by the mean daily flow for each day in the time period to obtain a succession of estimated daily (unit) loads. Another approach involves multiplying the average observed concentration by the average flow based on all days of the year to obtain an "average" daily load, which is then converted to the total load. Other variants are the monthly average concentration times the average flow for the month, the quarterly average concentration times the average flow for the quarter, etc. Several of these averaging approaches are described and evaluated in Dolan *et al.* (1981) and Walling and Webb (1981). Generally, averaging approaches tend to be biased if concentration is correlated with flow: the calculated load is too low if the correlation is positive and too high if the correlation is negative. However, some averaging approaches have shown relatively high precision in some studies, and might be useful in special situations, for example if the goal is to detect a change in the load, and detecting the change is more important that knowing the actual magnitude of the load.

A special case involves the flow-weighted mean concentration (FWMC), which is then multiplied by the discharge to obtain the load. This is a trivial case, since the FWMC is defined as the load divided by the discharge.

The **flow interval** technique (Yaksich and Verhoff, 1983) is a semi-graphical technique which begins with a plot of the year's observed instantaneous fluxes as a function of instantaneous flows at the time the samples were taken. The plot is divided into several intervals of uniform size covering the range of mean daily flows for all days of the year. For each interval, the average flux is calculated and the number of days with mean daily flows in the interval is determined. The interval load is calculated as the product of the average flux, the number of days in the interval, and the appropriate units conversion factor. The annual load is calculated by summing the interval loads. Formulas for estimating a confidence interval are available.

This technique is a form of stratified averaging approach. The use of *uniform* intervals seems unnecessary, and probably is less efficient than some other stratification schemes. While it is more efficient than unstratified averaging approaches, it is presumably less efficient than stratified approaches which make use of more information, such as the ratio estimators described briefly in the next paragraph, and in detail in later sections of this document.

Ratio estimators determine the average daily load for the days with concentration observations, adjust it proportionally by reference to some parameter which is more thoroughly sampled (ideally each day), and then calculate the total load by multiplying the adjusted daily load by 365. The most common parameter used for adjustment is discharge, though an average daily phosphorus load might be adjusted by ratio with a better-sampled average daily sediment load. Multivariate ratio estimators involving more than one adjustment parameter are described in the statistical literature, but the mathematics are very difficult, and such estimators have not been applied to load estimation problems.

Ratio estimators assume that there is a linear relationship between the daily loads and the adjustment parameter, which passes through the origin. Ratio estimators are often biased, but some estimators have been developed which include correction terms which eliminate or greatly reduce the bias. Cochran (1963, p. 150-186) presents a thorough treatment including several nearly unbiased ratio estimators.

In most applications of ratio estimators to pollutant load estimation, the calculations and sometimes the sampling program have been stratified, usually by flow and/or season.

Regression approaches develop a relationship between concentration and flow based on the samples taken, then use the relationship to estimate a representative concentration for days not sampled, usually using the mean daily flow as input to the regression equation. Multivariate regression relationships have been developed in some studies. The relationship is sometimes developed using flux and flow rather than concentration and flow, but the results are identical.

Most regression estimators are based on a linear regression model, though this is often applied after transformation. The log transformation is frequently used, because many environmental parameters are approximately log-normally distributed. Regression relationships between log-transformed concentration or flux and flow are often called **rating curves** in the engineering literature. In a few instances, non-parametric estimators such as the LOWESS Smooth have been used (Helsel and Hirsch, 1992).

A problem which is most commonly encountered with regression estimators is the so-called retransformation bias, which can lead to large errors in estimated loads. A discussion of this problem can be found in Ferguson (1986, 1987) and Koch and Smillie (1987); further information is included later in this document. Researchers at the U.S. Geological Survey have developed retransformation techniques which are largely free of this bias (see Cohn *et al.*, 1992).

Flow-proportional sampling is a totally different approach, mechanical rather than mathematical, which essentially assumes that one or more samples can be obtained which cover the entire period of interest, each representing a known discharge and each with a concentration which is in proportion to the load which passed the sampling point during the sample's accumulation. If this assumption is met, the load for each sample is easily calculated as the discharge times the concentration, and the total load for the year is derived by summation. In principle, this is a very efficient and cost-effective method of obtaining a total load. It has several disadvantages, however, which limit its use:

1. it is not compatible with other goals, such as monitoring for ambient concentrations that are highest at low flow,
2. unit loads are not available for regularly spaced time intervals, such as days,
3. there is no obvious means of recovery in case the mechanical system does not perform as required,
4. there is no way to obtain a precision estimate, and
5. the accuracy and precision of the method cannot be evaluated by the methods described in the next section, which means that direct comparison with the other methods is not possible.

Performance of the methods

Some methods of estimating loads have the additional desirable feature of providing a measure of the uncertainty of the load estimate. Unfortunately, many do not. Furthermore, the uncertainty estimates of different load calculation methods cannot be directly compared, because they reflect different kinds of "error". Further still, the estimated “error” may be different from the error we are interested in. For example, the uncertainty estimate for the Beale Ratio Estimator includes a contribution which is due to differences between individual daily loads and the mean daily load in each stratum. If we are interested in the annual load for that year, we would not consider this to be a source of error, but rather a part of the natural variation of the system we are studying. We would prefer to confine our notion of error to the difference between our estimated mean daily loads and the actual (but unknown) mean daily loads, a difference which is due only to sampling and analytical error.

For these reasons, uncertainty measures do not provide a reliable means to choose between methods. Consequently, evaluation of load estimation methods must rely on comparative studies in which several methods are used to calculated loads from the same data, and the results are compared with the "true" load which is independently known.

There are two basic approaches to simulation: systematic subsampling and Monte Carlo simulation. Systematic subsampling involves taking a dense dataset and splitting it into subsamples of the same size. For example, Walling and Webb took a 7-year record of turbidity with samples every 15 minutes, used a regression relationship to convert it to an equivalent suspended solids record, and then divided it into sample sets which would have been obtained with less frequent sampling. To simulate monthly sampling, one set would be formed from the first sample each month, another from the second sample each month, etc. For this dataset, more than 18,000 such monthly sample sets can be obtained. The "true" load is the load based on the entire dense dataset.

Monte Carlo simulation involves random sampling of an empirical distribution of observations, or sampling of parametric distributions generalized from such observations, to produce any desired number of simulated sample sets. Typically 500, 1000, or 2000 sample sets are generated and evaluated. The "true" load is the load based on the entire dataset or calculated from the parametric distributions used.

There are advantages and disadvantages to each method, and several excellent evaluative studies of loading approaches have been published based on both methods (Burn, 1990; Dickenson, 1981; Dolan *et al.*, 1981; Preston *et al*., 1989, 1992; Richards and Holloway, 1987; Young *et al*., 1988; Walling, 1987; Walling and Webb, 1988). These studies investigated different load calculation methods, sampling frequencies and sampling patterns. Several points of consensus emerge from these studies:

• Not surprisingly, accuracy and precision increase with increased frequency of sampling.

• Averaging methods are generally biased, and the bias increases as the size of the averaging window increases. For example, a monthly load can be calculated by multiplying the average concentration for the month by the discharge for the month and an appropriate conversion factor to account for the change of units (Appendix A), and a quarterly load can be similarly derived using the quarterly discharge and average concentration. In general, the annual load which is the sum of the four quarterly loads will be more biased than the annual load which is the sum of the 12 monthly loads.

• In most studies, ratio approaches performed better than regression approaches, and both performed better than averaging approaches. In particular, ratio approaches which include a bias correction factor and are used in a stratified mode generally showed low to no bias, relatively high precision, and resistance to undue influence of unusual observations.

• Regression approaches can perform well if the relationship between flow and concentration is sufficiently well-defined, linear throughout the range of flows, and constant throughout the year. Stratification may allow these requirements to be met piece-wise. However, the regression approach may lead to large errors in estimated loads if the available data contains unusual observations which fall away from the trend of the rest of the data, especially if these are associated with high flows.

• When it was evaluated, stratified sampling with most samples taken during periods of high flow was of great importance in providing increased precision for a given number of samples.

• Stratification applied at the time of calculation produced more accurate load estimates in some cases, and had little effect or actually produced less accurate load estimates in other cases.

A NUMERIC INTEGRATION APPROACH

General Description

Sampling effort is concentrated in high flow periods, during which enough samples are gathered to characterize the interval loads. A limited number of samples are also gathered during low flow periods to allow loads for these less critical times to be calculated. For most applications, an autosampler will be required which can be triggered by rising stage.

This method, and particularly the sampling strategy, assumes that flows are highly variable and that concentrations increase with flow. If either of these assumptions is not met, this method may not give reliable load estimates. This method is recommended only if concentrations at high flow are higher than those at low flow, and if at least 70% of the annual discharge occurs during the 30% of the time with the highest flows. An alternative is provided in the next section if this is not the case, but it may involve too many samples to be feasible.

Sampling Needs and Sampling Strategy

The sampling strategy is based on the assumption that most of the load occurs in a short period of time during storm runoff events, and that accurate loads can be obtained by sampling primarily during that period of time. The strategy also assumes that if samples are sufficiently close together in time, the concentration and flow between samples will change smoothly with time, i.e. we will miss no peaks or valleys between samples except for the single peak in concentration and flow which must occur sometime during the storm runoff event, and which can never be sampled at exactly the right moment. As such, the method acknowledges and takes advantage of the autocorrelation which is typically present in the runoff process.

To determine the sampling interval during storm runoff events, divide the length of a runoff event by 16. The result may be rounded somewhat for convenience. For example, a sampling interval of 7.3 hours can be rounded to 8 hours.

To determine when sampling should start, do one of the following:

1. By inspection of existing records of stage, determine a stage which separates early storm runoff from base flow, and program the autosampler to begin when this stage is exceeded. Different triggering stages may be appropriate in different seasons.

2. By inspection of existing records of stage, determine a rate of change of stage which characterizes the onset of storm runoff. Program the autosampler to begin when this rate of change is exceeded.

To determine when sampling should stop, do one of the following:

1. Trigger the autosampler to stop sampling, or turn it off manually, when the stage decreases to less than 110% of the stage at which sampling started.

2. Turn off the autosampler manually when the water level and turbidity indicate that storm runoff has ceased, but not before 16 samples have been obtained.

3. Allow the autosampler to complete its cycle of sampling (typically 24 samples), at which time it will stop sampling automatically.

In addition to storm sampling, take one sample during low flow conditions during each month.

Stages must be recorded at hourly intervals for rivers for which a typical storm lasts four days or more, at 15 minute intervals for rivers with storm durations between one and four days, and at 5 minute intervals for rivers with storm durations less than one day. These stages must be converted to flows for use in calculating the loads, using an established and verified rating curve.

The number of samples obtained in a year can be estimated from existing flow records, but will fluctuate from year to year depending on the number of storms.

Load Estimation Approach

Numeric integration is used to calculate the load. This approach assumes that sampling during high flow periods is frequent enough that the sampled fluxes (concentrations and flows) closely match the continuous pattern of the actual fluxes, and in particular that the peak flux for each storm is not too badly underestimated. The likely validity of this assumption should be tested by comparing the flow pattern based on the flows at the times of the samples with that based on all of the flow data.

Express the time of each chemical sample and each flow observation as decimal days of the year. For example, noon on January 2 would be day 1.5, 6:00 p.m. would be day 1.75. For each chemical sample, establish a time window which starts halfway between the sample and the previous sample, and ends halfway between the sample and the next sample. The time window for the first sample during a storm should include only the interval based on the time to the next sample (i.e. should start at the beginning of the storm); the time window for the last sample during a storm should include only the interval based on the time from the preceding sample (i.e. should end at the end of the storm). Multiply the flow at the time of the sample by the concentration at the time of the sample and the time interval for the sample; multiply this by .002447 to convert cfs-mg/L-days to metric tons, or use a related conversion factor (see Appendix B) as appropriate. The result is the load for the time interval. Add all these loads together for the year; this is the annual load.

If chemical samples were lost due to analytical problems or autosampler failure, the load may be in error. It is good practice to compare the total annual discharge for the year based on the flows used in calculating the load, with the total annual discharge based on all observed flows. An adjusted load may be calculated by multiplying the observed load by the ratio of the total annual discharge to the annual discharge for samples.

A somewhat more accurate load estimate can be obtained by the following variant. Define the time window for each sample as above. Average the observations of flow which occurred within this time window, and use the average flow to calculate the interval load rather than the flow at the time of the chemical sample.

Computer Program

A computer program which calculates loads by numeric integration is included on the CD-ROM. It is called Integrator. Specifics of its use are provided in the accompanying user's guide.

Uncertainty Estimate

This method is not a statistically-based approach to load estimation, and therefore a confidence interval in the strict statistical sense cannot be calculated. However, an uncertainty interval can be estimated if it is assumed that the (unsampled) concentrations change linearly with time between each pair of sampled concentrations. This assumption is unlikely to be strictly true during storms, but it is likely to be approximately true if the sampling interval was calculated correctly and adhered to. The assumption may not be valid during low flow periods, but these periods make only small contributions to the loads, and errors are relatively unimportant for the annual load.

For each series of observed concentrations, interpolate (by time) between observed concentrations to obtain concentration estimates corresponding to the boundaries of the time windows for each sample. Compare each original concentration with the two extrapolated concentrations which surround it in time. Under the assumption of linear change, no concentration in this time interval should be larger than the largest of these three values, and no concentration should be smaller than the smallest of the three. Create two new series of estimated concentrations accordingly: the series of maximum values and the series of minimum values. Use the series of minimum values to calculate a lower-bound estimate of the load, and the series of maximum values to calculate an upper-bound estimate of the load. Use the same values of flow and time interval which were used to calculate the initial annual load.

The observed maximum concentration for each storm is likely to be less than the true maximum and cannot be more. As a consequence, the upper-bound estimate of the annual load is likely to be biased slightly low. As far as is known, no studies have evaluated the magnitude of this bias.

Alternative approach

If it is not the case that concentration increases with flow and that loads occur mostly during storm runoff, the load may be fairly well distributed over the course of the year. Note that this is very much the exception in non-point pollution studies, especially in small rivers and streams. However, if these assumptions are not met, a uniform sampling frequency of twice daily is recommended for initial investigations. Flows should be measured at least hourly. Over several years, examine the behavior of the concentrations and flows. If a pattern can be identified which will allow sampling to be allocated more efficiently by concentrating sampling at certain times, the schedule can be adjusted. For example, pesticide concentrations tend to be strongly seasonal, and loads during the late fall to early spring may be inconsequential.

THE REGRESSION APPROACH

General Description

A regression relationship is developed between concentration and flow, based on the days on which samples are obtained. This relationship may involve simple or multiple regression, and concentration or flux may be used as the dependent variable. In most applications, both concentration (or flux) and flow are log-transformed to create a dataset which is better suited for regression analysis. The regression relationship may be based entirely on the current year's samples, or it may be based on samples gathered in previous years, or both. Time may be used as a variable to account for possible linear trends.

Once the regression relationship is established, it is used to estimate concentrations for each day on which a sample was not taken, based on the flow (usually the mean daily flow) for the day. The total load is calculated as the sum of the daily loads, obtained by multiplying the measured or estimated concentration by the flow, and including a conversion factor (Appendix A) to account for the change in units of measurement from, for example, $\frac{\text{mg}}{\text{L}} * \frac{\text{ft}^3}{\text{sec}}$ to $\frac{\text{tons}}{\text{yr}}$.

Assumptions

Regression approaches assume only that there is a linear relationship between a dependent variable, concentration or flux, and one or more independent variables, typically flow but sometimes also higher powers of flow, time, seasonality, and other variables. Concentration, flux, and flow are often log-transformed, using either natural logs or logs to the base 10, to create a more linear relationship and/or to reduce the influence of the highest concentrations.

When log transformations are applied, the inverse transformation (exponentiation) is required to obtain estimated concentrations, since the regression model yields estimates of log-concentration, not of concentration itself. This transformation creates a bias in the loads, and further assumptions must be made (and validated) about the distribution of the residuals of the estimated concentrations in order to correct the bias. This complex subject will be dealt with in a following section.

Sampling Needs and Sampling Strategy

The goal of sampling is to thoroughly characterize the relationship between flow and concentration (or flux). However, it is difficult to translate this into a statistical statement which can be used to calculate the number of samples required. Since regression approaches to load estimation use more of the information contained in a sample of concentration/flow data than ratio approaches, a conservative approach might be to collect the number of samples calculated for the ratio approach, according to the techniques given in the section on the Beale Ratio Estimator.

Cohn *et al.* (1992) used 75 samples to establish their regression models. Cohn (personal communication, March 1997) recommends that the regression relationship be established using 75 samples taken over a two year period, about half of them collected during high flow periods and half at times selected at random or using a fixed interval between samples.

Since the goal of sampling is to thoroughly characterize the relationship between flow and concentration or flux, the program should be designed to obtain samples over the entire range of expected flows. If seasonal differences in the flow/concentration relationship are possible, the entire range of flows should be sampled in each season.

Concentration/flow relationships within a storm runoff event are usually much more homogeneous than those in different runoff events. For this reason it is important to avoid the temptation to sample one or two storms in great detail, rather than sampling many storms with fewer samples per storm.

In addition, Cohn recommends collecting 25 samples per year after the initial two-year calibration period, to verify that the regression relationship is not changing with time. As in the initial sampling period, approximately half of the samples should be collected during high flow conditions. In the unusual cases where long-term datasets exist, the preferred practice of the USGS is to develop the regression relationship using a ten-year period of data in which the year for which the load is to be calculated is the ninth (second most recent).

Load estimation using regression approaches has usually been done without stratified calculations. However, Walling and Webb (1981) demonstrated reduced bias and increased precision by increasing the number of high flow samples (stratified sampling, as recommended above) and by calculating regression relationships separately for winter and summer, high and low flow regimes (stratified calculations). Other workers have suggested other forms of stratification (see Cohn *et al.*, 1992 for a summary and references). Cohn *et al.* (1992) used a multivariate regression approach rather than multiple bivariate regression relationships (stratification).

Load Estimation Approach

The Basic Approach

The desired number of observations of instantaneous concentration and flow are made, distributed over the flow regime. These data are used to establish a regression relationship of the form

$$c = mq + b \tag{6}$$

where c is concentration, q is flow, m is the slope of the linear relationship, and b is its intercept, as determined by the least-squares regression procedure. Such regression calculations are available in almost any modern spreadsheet, statistics, or data analysis program.

Once the regression relationship is calculated, it is used to estimate concentrations for each day of the year, by substituting the mean daily flow into the equation and solving for the estimated concentration:

$$\hat{c} = m\bar{q} + b \qquad (7)$$

Here $\hat{c}$ is used instead of c to remind us that the concentration is an estimate, and $\bar{q}$ is used instead of q to remind us that the flow is the mean daily flow, not the instantaneous flow.

Finally, the annual load is calculated as the sum of the daily fluxes based on the estimated daily loads and mean daily flows, applying a conversion factor for the change of units:

$$\text{Load} = k\sum_{i=1}^{365} \hat{c}_i \bar{q}_i \qquad (8)$$

The USGS Seven-parameter Model

In several studies of nutrient transport in rivers entering Chesapeake Bay, the U. S. Geological Survey (USGS) found that a more complex regression model gave better results. The approach is given by Cohn *et al.*, 1992. This model can be written, using their notation, as

$$\ln[C] = \beta_0 + \beta_1 \ln\left[\frac{Q}{\tilde{Q}}\right] + \beta_2 \ln\left[\frac{Q}{\tilde{Q}}\right]^2 + \beta_3\left[T - \tilde{T}\right] + \beta_4\left[T - \tilde{T}\right]^2 + \beta_5 \sin[2\pi T] + \beta_6 \cos[2\pi T] + \varepsilon \qquad (9)$$

where ln[ ] means the natural logarithm of the parameter, C is concentration, Q is discharge, and T is time measured in years. The errors, denoted by ε, are assumed to be independent and normally distributed with a mean value of 0 and variance σ_ε^2. β_0 through β_6 are the seven parameters which must be estimated by regression, and $\tilde{Q}$ and $\tilde{T}$ are "centering variables", which simplify the mathematics but have no effect on the load estimate. They are calculated by

$$\tilde{W} = \overline{W} + \frac{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^3}{2\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2} \tag{10}$$

where W=ln(Q), and $\tilde{Q} = e^{\tilde{W}}$

$$\tilde{T} = \overline{T} + \frac{\sum_{i=1}^{n}\left(T_i - \overline{T}\right)^3}{2\sum_{i=1}^{n}\left(T_i - \overline{T}\right)^2} \tag{11}$$

Note that if β_2 through β_6 are zero, this complex formula is comparable to the simple regression relationship (6), except that log-transformed variables are used and discharge is "centered". β_0 is the intercept term and β_1 is the slope term.

This model includes the capability of accounting for curvilinear relationships between concentration and flow, through the Q^2 term, for trends over time (T and T^2 terms), and for seasonality (sin and cos terms). The magnitude (or amplitude, A) of the seasonal effect can be calculated as

$$A = \sqrt{\beta_5^2 + \beta_6^2} \tag{12}$$

and date D of the peak (maximum or minimum) value can be calculated as

$$D = \frac{365}{2\pi}\left[\tan^{-1}\left(\frac{\beta_5}{\beta_6}\right)\right] \tag{13}$$

Transformation, Back-transformation, and Bias Avoidance

In order for concentrations estimated from the regression model to be reliable, the residuals (the differences between the predicted and observed concentrations used to calculate the regression model) must be normally distributed. In addition, it is desirable for the data to be well spread out over the range of observations. For these and several other reasons, regression models relating concentration to flow usually use log-transformed values. In order to be of much use, however, the resulting data must be back-transformed before calculating the loads. The obvious way to do this is by taking the anti-logs of the estimated concentrations.

Statistical theory tells us, however, that when these back-transformed values are used to calculate average daily loads or total annual loads, the results will be biased low (Ferguson, 1986a, 1986b; Koch and Smillie, 1986a, 1986b; Cohn *et al.*, 1989, 1992). In order to avoid this bias, a value must be added to each estimated log-concentration before it is back-transformed.

According to Ferguson (1986a), under the assumption that the residuals are normally distributed, the appropriate procedure is to estimate the concentrations using

$$\hat{c} = e^{\hat{y} + \frac{\sigma^2}{2}} \tag{14}$$

where $\hat{y}$ is the log-concentration estimated from the regression model, and σ^2 is the variance of the residuals of the regression model. This is referred to by Cohn *et al.* (1989), as the quasi-maximum likelihood estimate, or QMLE. If the original transformation used common logs (base 10) rather than natural logs, the equivalent would be

$$\hat{c} = 10^{\hat{y} + 2.65\sigma^2} \tag{14a}$$

Koch and Smillie (1986a, b) pointed out that if the assumption of normally-distributed residuals is violated by the actual data, the application of 12 or 12a can actually lead to overestimates which are farther from the true value than those provided by the uncorrected back-transformation. They applied a non-parametric correction factor called the smearing estimate (Duan, 1983), with even worse results, and concluded that there may be no single approach to bias correction which will work reliably in all cases. However, as Ferguson (1986b) pointed out, Koch and Smillie failed to distinguish between systematic error (bias) and random error (precision), and the smearing estimate cannot be generally condemned on the basis of their results. The smearing estimate is a constant by which the estimated concentration is multiplied after exponentiation. The smearing estimate constant is

$$k_{sm} = \frac{1}{n}\sum_{i=1}^{n} e^{\varepsilon_i} \tag{15}$$

where ε_i is the ith residual from the regression model; this is equivalent to the mean of the exponentiated residuals.

Cohn *et al.* (1989) proposed a minimum variance unbiased estimator (MVUE)

$$\hat{L}_{MVUE} = \hat{L}_{RC}\, g_m\left(\frac{m+1}{2m}[1-V]s^2\right) \tag{16}$$

or, alternatively,

$$\hat{C}_{MVUE} = \hat{C}_{RC}\, g_m\left(\frac{m+1}{2m}[1-V]s^2\right) \tag{16a}$$

where $\hat{L}$ is the estimated load, $\hat{C}$ the estimated concentration, RC refers to the rating curve or regression relationship between log-load and log-flow, and everything to the right of $\hat{L}_{RC}$ is the MVUE bias correction factor. g_m is a Bessel function described in their paper, m is the number of observations used to establish the regression relationship minus the number of parameters in it, s^2 is the variance of the residuals from the regression relationship between load and flow, and V is a “leverage” term which is a function of the values of the independent variables from which a specific load (or concentration) estimate is to be calculated. As a consequence, the MVUE bias correction factor is not a constant.

The authors show that this correction factor performs better than the alternatives presented above. Unfortunately, it is also quite cumbersome to calculate, though they offer computer code to evaluate it. Furthermore, Thomas (1988) points out that bias correction methods, including this one, depend on the validity of generally untested hypotheses, particularly that of normally distributed residuals (in log space).

In a later simulation study based on data from several rivers tributary to Chesapeake Bay, Cohn *et al.* (1992) found that the MVUE estimator generally gave fairly accurate results. They concluded that it was fairly insensitive to modest violations of the assumptions of the regression approach.

Cohn (1995) states that the three bias correction methods give nearly identical results if:

1. the assumed linear model is approximately correct,
2. the regression model is based on 30 or more observations, and
3. the model is not being used to extrapolate beyond the range of data used to calibrate it.

He says further that if only the first condition is satisfied, the MVUE estimate is the best. If all conditions are satisfied, the QMLE is a good choice because it is relatively easy to calculate. If the regression residuals are not normally distributed, the smearing estimate may be preferred, but in this case one must verify that the use of the regression model is appropriate.

Computer Programs

A program which calculates loads using the seven-parameter model has been developed by Tim Cohn at the USGS in Reston, VA. This program, named ESTIMATOR, calculates loads using a user specified subset of the independent variables in the seven-parameter model described above. Several diagnostics are provided as well as the load estimates. The regression model is established using a calibration set of data, and can then be used to estimate loads for any year for which mean daily flow data are provided in a separate flow file. The program is rather user-unfriendly, and in particular the structure of the input files is difficult to follow, because the program was written to use input files of a pre-determined format produced by the USGS Automated Data Processing System (ADAPS) and Water Quality Data Base (QWDATA). However, a thorough user's manual is provided and it describes the data structures in detail. Operational versions of the program for IBM and Macintosh computers is provided on the CD-ROM, along with sample files. This author has limited experience using ESTIMATOR, so those who chose to use this program may expect to spend some time on a fairly steep learning curve.

An Excel spread-sheet version of the seven-parameter model is also provided on the CD-ROM. This version uses the QMLE and smearing back-transformation bias corrections. While this version does not employ the most elaborate of the three bias correction techniques (MVUE), it has the advantage of relative simplicity of operation, and operation in the spread-sheet mode, which may be relatively familiar to many potential users. A worked example is presented below, and a set of instructions for the use of the workbook are provided on the CD-ROM.

A Simple Example: Cuyahoga River Sediment Loads

The detailed Water Quality Lab dataset for the Cuyahoga River at Independence, Ohio, contains 271 observations of flow and suspended sediment concentration for the calendar year 1992. This dataset was subsampled to extract data at weekly intervals, such as a more typical sampling program might obtain. The data are shown in Table 2. Because there were some long gaps in the original data set, there are only 33 "weekly" observations.

Concentration and flow were log-transformed, and the regression relationship between log-flow and log-concentration was found to be

$$\hat{C} = 0.88093174 * \ln(Q) - 2.3791111$$

with an adjusted R^2 of 0.375. The variance of the residuals was equal to 0.85710476.

Table 2. Flow and suspended solids concentrations in the Cuyahoga River, 1992.

Date	Time	Flow	SS	ln(flow)	ln(ss)
Jan 02	1.5	257.0	6.6	5.549	1.879
Jan 09	8.5	272.0	9.4	5.606	2.242
Jan 16	15.5	454.0	30.4	6.118	3.414
Jan 30	29.5	424.0	14.7	6.050	2.688
Feb 06	36.5	401.0	7.7	5.994	2.041
Feb 13	43.5	279.0	5.0	5.631	1.609
Apr 16	106.5	631.0	59.3	6.447	4.083
Apr 23	113.5	1390.0	69.0	7.237	4.234
Apr 30	120.5	870.0	39.2	6.768	3.669
May 07	127.5	542.0	12.2	6.295	2.501
May 14	134.5	268.0	17.0	5.591	2.833
May 21	141.5	253.0	15.9	5.533	2.766
May 28	148.5	296.0	19.2	5.690	2.955
Jun 04	155.5	309.7	27.5	5.735	3.314
Jun 18	169.5	669.0	370.0	6.506	5.914
Jun 25	176.5	380.3	46.4	5.941	3.837
Jul 02	183.5	197.0	22.3	5.283	3.105
Jul 09	190.5	253.0	33.9	5.533	3.523
Jul 30	211.5	2400.0	11.5	7.783	2.442
Aug 06	218.5	933.0	14.3	6.838	2.660
Aug 13	225.5	767.2	69.6	6.643	4.243
Oct 15	288.5	799.0	56.0	6.683	4.025
Oct 22	295.5	445.0	11.9	6.098	2.477
Oct 29	302.5	407.0	9.7	6.009	2.269
Nov 05	309.5	1130.0	38.5	7.030	3.651
Nov 12	316.5	2410.0	200.0	7.787	5.298

Table 2. Flow and suspended solids concentrations in the Cuyahoga River, 1992, concluded.

Date	Time	Flow	SS	ln(flow)	ln(ss)
Nov 19	323.5	1440.0	51.8	7.272	3.947
Nov 26	330.5	1750.0	71.6	7.467	4.271
Dec 03	337.5	1170.0	21.8	7.065	3.082
Dec 10	344.5	865.0	15.5	6.763	2.741
Dec 17	351.5	1030.0	12.0	6.937	2.485
Dec 24	358.5	874.0	11.2	6.773	2.416
Dec 31	365.5	9780.0	1420.0	9.188	7.258

Mean daily values for flow were obtained from the U.S. Geological Survey web site (http://waterdata.usgs.gov/nwis-w/OH/) for calendar year 1992. These daily values were used with the regression equation to estimate concentrations for each day of the year (including the days for which samples were available). The estimated concentrations were then back-transformed, using half the variance of the residuals as a bias correction factor. The resulting estimated concentrations were multiplied by the corresponding mean daily flows and the conversion factor 0.002447 to obtain daily loads in metric tons. The daily loads were summed to obtain the annual load estimate of 105,525 metric tons. Exotic software is not required: all of the calculations for this example were done using Excel (however, not using the Excel version of the seven-parameter model). The daily estimated concentrations and loads are tabulated in Appendix C.

The load calculated using a different subset containing 24 observations, chosen by uniform subsampling of the ranked flows, was 136,195. Clearly, there can be considerable variations in load estimates, due merely to when the samples were taken!

Since these are synthetic data sets derived by reduction of the same much more detailed data set, it is interesting to compare the loads above with that obtained using all of the data, which is 131,024. The load calculated without using bias correction is very much lower: 70,824 or 54% of the bias-corrected load. The load calculated for the same data using the Beale Ratio Estimator (see later) is 123,518.

Example of the USGS approach: Choptank River Total Nitrogen Loads

Cohn *et al.* (1992), in a test of the seven-parameter model, worked with datasets for several water quality constituents measured at several stations on tributaries to Chesapeake Bay. After summarizing their results for total nitrogen (TN) in the Choptank River, we will derive these results independently from the raw data and apply the resulting model to calculate a monthly load.

For the Choptank River, 274 observations of total nitrogen were available over a 13 year period. Applying their model to this data, Cohn *et al.* arrived at the relationship

$$\ln(\mathrm{TN}) = .464 + .056\ \ln\left(\frac{Q}{\tilde{Q}}\right) + .002\ \ln\left(\frac{Q}{\tilde{Q}}\right)^2 + .023\left(T - \tilde{T}\right) - .0025\left(T - \tilde{T}\right)^2 + .055\ \sin(2\pi T) + .0037\ \cos(2\pi T)$$

From the sine and cosine terms they calculated the amplitude of the seasonal effect using Formula (12):

$$A = \sqrt{.055^2 + .0037^2} = \sqrt{.003039} = .055$$

This value is small compared to the constant, indicating that seasonality is not particularly pronounced in this data.

Similarly, the day of maximum amplitude was calculated using Formula (13):

$$D = \frac{365}{2\pi}\left[\tan^{-1}\left(\frac{.055}{.0037}\right)\right] = \frac{365}{2\pi}\left[\tan^{-1}(14.86)\right] = \frac{365}{2\pi}[1.50] = 87.3,$$

the Julian day which corresponds to March 28. Note that the value returned by the arctan function should be in radians, not degrees.

The Choptank total nitrogen and flow data were obtained from USGS and the regression exercise was repeated, using the Excel workbook "Regression with Centering", which is on the CD-ROM. Due to slight differences in data and calculation methods, the results differ slightly. Using the data supplied, the workbook calculated the following centering constants:

$$\tilde{T} = \overline{T} + \frac{\sum_{i=1}^{274}\left(T_i - \overline{T}\right)^3}{2\sum_{i=1}^{274}\left(T_i - \overline{T}\right)^2} = 9.263 + \frac{-7713}{2 * 3605} = 8.194$$

$$\tilde{W} = \overline{W} + \frac{\sum_{i=1}^{274}\left(W_i - \overline{W}\right)^3}{2\sum_{i=1}^{274}\left(W_i - \overline{W}\right)^2} = 4.202 + \frac{116.28}{2 * 397.47} = 4.3554$$

$$\tilde{Q} = e^{\tilde{W}} = e^{4.3554} = 77.90$$

Note that the zero of the time scale for this workbook is set at January 1, 1975. Had we used the decimal years themselves, the centering constant would have been larger by 1975.0 and the centered time variable would wind up the same. Following the instructions for the workbook, we chose the regression option from the Tools:Data Analysis... menu. The regression coefficients indicate the following model:

$$\ln(TN) = .4634 + .0580\ln\left(\frac{Q}{\tilde{Q}}\right) - .0009\ln\left(\frac{Q}{\tilde{Q}}\right)^2 + .0237\left(T - \tilde{T}\right) - .0025\left(T - \tilde{T}\right)^2 + .0418\sin(2\pi T) + .0290\cos(2\pi T)$$

From the sine and cosine terms the workbook calculated the amplitude of the seasonal effect:

$$A = \sqrt{.0419^2 + .0285^2} = \sqrt{.002568} = .0508$$

This value is similar to the one reported by Cohn *et al.*

The day of maximum amplitude is:

$$D = \frac{365}{2\pi}\left[\tan^{-1}\left(\frac{.0418}{.0290}\right)\right] = \frac{365}{2\pi}\left[\tan^{-1}(1.449)\right] = \frac{365}{2\pi}[0.9735] = 56.15,$$

the Julian day which corresponds to February 25, more than a month earlier than their result.

We can now apply our regression model to a sample set of flow data for a different time period and calculate a load. We enter our flow data with dates into the workbook, which then estimates concentrations for each flow, using the regression model. Concentrations are estimated without retransformation bias correction, and with bias correction using the QMLE approach and the smearing estimate. The variance of the residuals from the regression model is 0.05558, hence the additive QMLE bias correction factor is $\frac{s^2}{2} = .02778$. The smearing estimate correction multiplier, calculated as the average of the exponentiated residuals, is 1.0277. The MVUE approach is not implemented in this workbook.

On the following pages we show the calculation of a monthly load, done by applying the regression relationship to mean daily flows obtained from the U.S. Geological Survey's database accessible on the World Wide Web (http: //waterdata.usgs.gov/nwis-w/). The flow data for the month are shown in Table 3.

Table 3. Flows for the Choptank River near Greensboro, MD during May 1990.

Date	Flow	Date	Flow
5/1/90	235	5/13/90	349
5/2/90	200	5/14/90	256
5/3/90	165	5/15/90	219
5/4/90	147	5/16/90	184

Table 3. Flows for the Choptank River near Greensboro, MD, concluded.

Date	Flow	Date	Flow
5/5/90	167	5/17/90	166
5/6/90	206	5/18/90	154
5/7/90	186	5/19/90	139
5/8/90	154	5/20/90	126
5/9/90	137	5/21/90	122
5/10/90	162	5/22/90	122
5/11/90	766	5/23/90	112
5/12/90	680	5/24/90	104
5/25/90	97	5/29/90	549
5/26/90	134	5/30/90	1700
5/27/90	195	5/31/90	850
5/28/90	172		

Using these values with the Choptank total nitrogen regression model, the workbook calculates the estimated log-concentration for each day of the month.

Table 4. Calculation of estimated daily average log-concentrations for May 1990.

Date	T	Q	$T-\tilde{T}$	$\ln(Q/\tilde{Q})$	$\ln(\hat{C})$
5/1/90	14.3397	235	7.147	1.104	0.589
5/2/90	14.3425	200	7.150	0.943	0.578
5/3/90	14.3452	165	7.153	0.751	0.566
5/4/90	14.3479	147	7.156	0.635	0.558
5/5/90	14.3507	167	7.158	0.763	0.565
5/6/90	14.3534	206	7.161	0.972	0.577
5/7/90	14.3562	186	7.164	0.870	0.570
5/8/90	14.3589	154	7.167	0.682	0.558
5/9/90	14.3616	137	7.169	0.565	0.550
5/10/90	14.3644	162	7.172	0.732	0.559
5/11/90	14.3671	766	7.175	2.286	0.652
5/12/90	14.3699	680	7.177	2.167	0.644

Table 4. Calculation of estimated daily average log-concentrations, concluded.

Date	T	Q	$T-\tilde{T}$	$\ln(Q/\tilde{Q})$	$\ln(\hat{C})$
5/13/90	14.3726	349	7.180	1.500	0.602
5/14/90	14.3753	256	7.183	1.190	0.583
5/15/90	14.3781	219	7.186	1.034	0.572
5/16/90	14.3808	184	7.188	0.859	0.561
5/17/90	14.3836	166	7.191	0.757	0.554
5/18/90	14.3863	154	7.194	0.682	0.549
5/19/90	14.3890	139	7.197	0.579	0.542
5/20/90	14.3918	126	7.199	0.481	0.535
5/21/90	14.3945	122	7.202	0.449	0.532
5/22/90	14.3973	122	7.205	0.449	0.531
5/23/90	14.4000	112	7.208	0.363	0.525
5/24/90	14.4027	104	7.210	0.289	0.520
5/25/90	14.4055	97	7.213	0.219	0.515
5/26/90	14.4082	134	7.216	0.542	0.533
5/27/90	14.4110	195	7.219	0.918	0.554
5/28/90	14.4137	172	7.221	0.792	0.546
5/29/90	14.4164	549	7.224	1.953	0.615
5/30/90	14.4192	1700	7.227	3.083	0.685
5/31/90	14.4219	850	7.230	2.390	0.641

Note that the centering values used are the same ones listed above: we do not recalculate the centering values when applying the regression model to a new set of data. Similarly, time (T) is expressed as years after January 1, 1975, as it was in the model calibration.

The estimated log-concentrations must next be back-transformed, then multiplied by the mean daily flows and the constant 0.002447 to calculate daily loads expressed as metric tons, which are then summed to obtain the monthly load. Estimated log-concentrations are back-transform in three ways: without a bias correction factor (the “naive” estimate) and with the bias correction factors based on the standard deviation of the residuals (QMLE) and the smearing estimate.

The results are shown in Table 5. In this example, the choice of bias correction factor does not have a major impact on the load estimate.

Table 5. Daily and monthly loads calculated using three different approaches to bias correction.

Date	$\ln(\hat{C})$	$\hat{C}_{Naive}$	$\hat{C}_{QMLE}$	$\hat{C}_{smear}$	Load Naive	Load QMLE	Load smear
5/1/90	0.589	1.802	1.853	1.852	1.035	1.064	1.064
5/2/90	0.578	1.783	1.859	1.832	0.872	0.896	0.896
5/3/90	0.566	1.761	1.837	1.810	0.710	0.731	0.730
5/4/90	0.558	1.748	1.823	1.796	0.628	0.646	0.646
5/5/90	0.565	1.759	1.835	1.808	0.718	0.739	0.738
5/6/90	0.577	1.780	1.856	1.829	0.897	0.922	0.921
5/7/90	0.570	1.768	1.843	1.817	0.804	0.827	0.826
5/8/90	0.558	1.746	1.821	1.795	0.658	0.676	0.676
5/9/90	0.550	1.733	1.807	1.781	0.580	0.597	0.597
5/10/90	0.559	1.749	1.824	1.797	0.693	0.712	0.712
5/11/90	0.652	1.920	2.002	1.973	3.595	3.697	3.695
5/12/90	0.644	1.904	1.986	1.957	3.165	3.255	3.253
5/13/90	0.602	1.826	1.904	1.877	1.558	1.602	1.601
5/14/90	0.583	1.791	1.867	1.840	1.121	1.152	1.152
5/15/90	0.572	1.772	1.848	1.821	0.949	0.976	0.975
5/16/90	0.561	1.752	1.828	1.801	0.788	0.811	0.810
5/17/90	0.554	1.740	1.815	1.788	0.706	0.726	0.726
5/18/90	0.549	1.731	1.805	1.779	0.652	0.670	0.670
5/19/90	0.542	1.719	1.793	1.766	0.584	0.601	0.600
5/20/90	0.535	1.707	1.781	1.755	0.526	0.541	0.541
5/21/90	0.532	1.703	1.776	1.750	0.508	0.522	0.522
5/22/90	0.531	1.701	1.774	1.748	0.507	0.522	0.521
5/23/90	0.525	1.691	1.764	1.738	0.463	0.476	0.476
5/24/90	0.520	1.682	1.754	1.729	0.428	0.440	0.440

Table 5. Daily and monthly loads calculated using three different approaches to bias correction, concluded.

Date	$\ln(\hat{C})$	$\hat{C}_{Naive}$	$\hat{C}_{QMLE}$	$\hat{C}_{smear}$	Load Naive	Load QMLE	Load smear
5/25/90	0.515	1.674	1.746	1.720	0.397	0.408	0.408
5/26/90	0.533	1.704	1.777	1.751	0.558	0.574	0.574
5/27/90	0.554	1.741	1.816	1.789	0.830	0.853	0.853
5/28/90	0.546	1.727	1.801	1.774	0.726	0.747	0.746
5/29/90	0.615	1.850	1.930	1.902	2.484	2.554	2.552
5/30/90	0.685	1.984	2.069	2.039	8.246	8.478	8.474
5/31/90	0.641	1.898	1.979	1.950	3.944	4.055	4.053
Total load for the month, in metric tons:					40.33	41.47	41.45

THE RATIO APPROACH: THE BEALE RATIO ESTIMATOR

General Description

On days on which samples are taken, the daily load is calculated as the product of concentration and flow, and the mean of these loads is also calculated. The mean daily load is then adjusted by multiplying it by a flow ratio, which is derived by dividing the average flow for the year as a whole by the average flow for the days on which chemical samples were taken. A bias correction factor is included in the calculation, to compensate for the effects of correlation between discharge and load. The adjusted mean daily load is multiplied by 365 to obtain the annual load.

When used in a stratified mode, the same process is applied within each stratum, and the stratum load is calculated by multiplying the mean daily load for the stratum by the number of days in the stratum. The stratum loads are then summed to obtain the total annual load.

Assumptions

Ratio estimators assume that there is a positive linear relationship between dependent and independent variables which passes through the origin. In addition, if the variance of the dependent variable is proportional to the magnitude of the independent variable, the ratio estimator is known to be the best linear unbiased estimator, i.e. the most precise among the class of unbiased estimators which assume a linear relationship. Both of these conditions are often satisfied, at least approximately, by relationships between load and discharge (Figure 6).

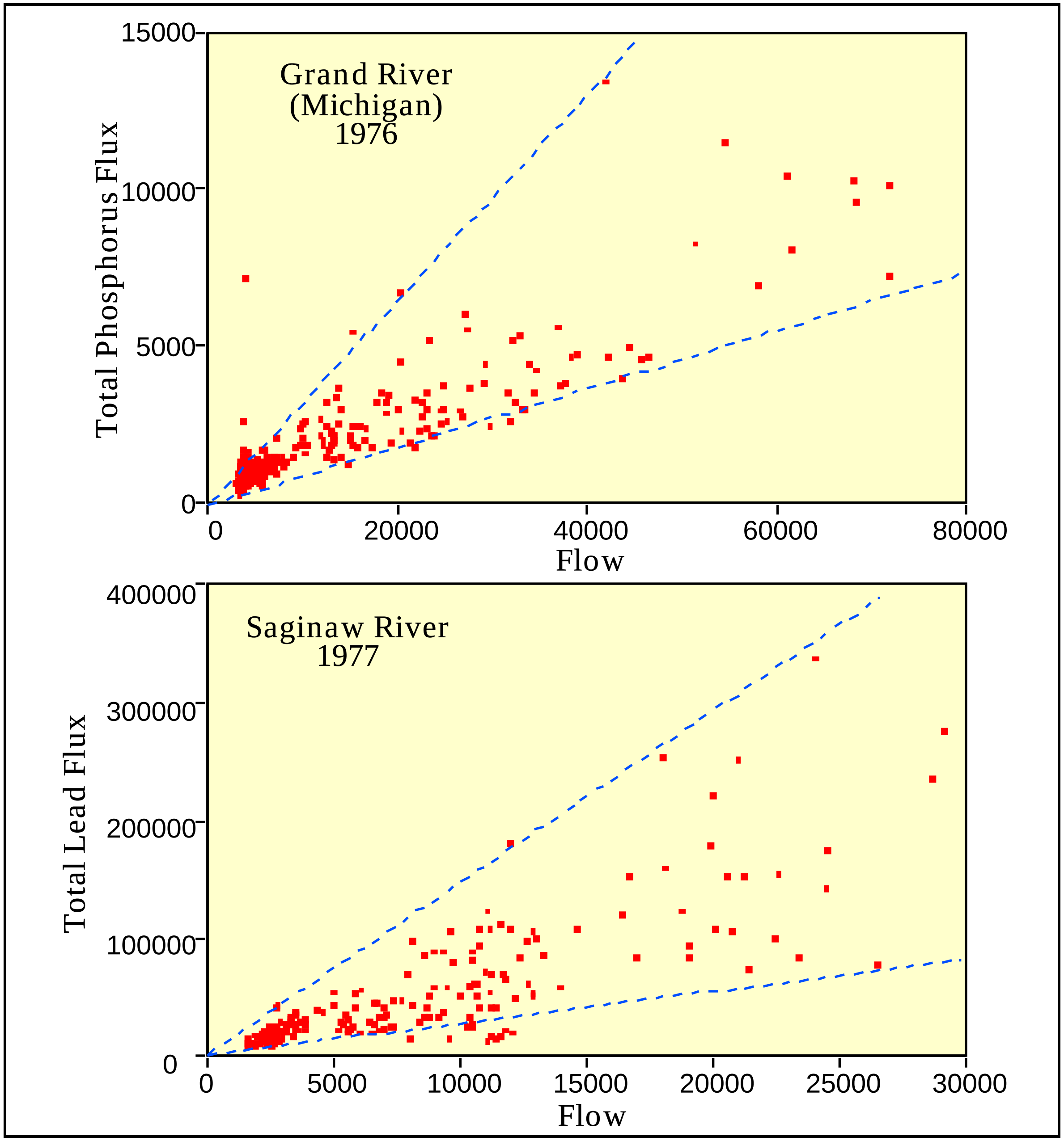

Figure 6. Flux as a function of flow for two different parameters measured in two Michigan rivers. In both instances, the spread of the data increases as the flow increases. Data courtesy of Steve Preston, U.S. Geological Survey, Towsend, Maryland.

Cochran (1977, p. 157) demonstrates the conditions under which the uncertainty associated with a ratio estimate will be smaller than that associated with a mean estimate. In loading terms, the total load estimated using the ratio estimator will be more precise than that derived by multiplying the average daily load by the number of days (without adjusting using the ratio of flows), if

$$G_{lq} \frac{C_l}{C_q} > 0.5 \qquad (17)$$

where G_{lq} is the correlation coefficient between flux and flow, C_l is the coefficient of variation of flux, and C_q is the coefficient of variation of flow. This criterion appears not to have been applied or tested in the context of pollutant load estimation, but it is worth observing that it is a criterion related to precision only, whereas one of the appealing features of the ratio estimator is that it helps to correct individual annual loads for non-representativeness of samples, as judged by flow. Even if the criterion is not satisfied, this adjustment may justify the use of the ratio estimator.

Cochran (1977, p. 156) also states that the variance estimate associated with ratio estimators is only reliable if the sample size exceeds 30 and is also large enough that the coefficients of variation of mean discharge and load are both less than 10%. This limitation also applies to each stratum if ratio estimation is applied within a stratified sampling scheme. The sample size criterion is likely to be met by most monitoring programs, but the coefficient of variation criteria may often not be met.

Moderate violations of these assumptions and criteria will introduce some error into the adjusted mean load and particularly into the estimate of the mean square error or variance, but do not totally invalidate the application of the method. Empirical studies show that the Beale Ratio Estimator is fairly robust against violations of the assumptions.

Sampling Strategy

Determining Sampling Frequency

The basic approach to determining sampling frequency assumes a normal distribution and random sampling. Under these assumptions, the number of observations for an unstratified sample is given by the formula

$$n = \frac{t^2 s^2}{E^2} \tag{18}$$

This formula requires an estimate of the variance of the population, s^2, a decision about the acceptable deviation from the mean daily load (E) and the probability α of the deviation occurring: you must be able to say something like "The variance of the population of daily loads is going to be about 400 tons^2. I want to get the mean daily load within 5 tons and accept only a .05 probability of failure". t is the value of the two tailed t-statistic for probability $\alpha/2$ with n degrees of freedom.

An alternative version of the formula may be used when the acceptable error is expressed as a percentage of the mean: “I want to get the mean daily load within ±15% and accept only a .05 probability of failure”. This version uses the coefficient of variation (cv) and the percent error expressed as a proportion p (for 15%, p=0.15):

$$n = \frac{t^2 (cv)^2}{p^2} \tag{18a}$$

Note that, in these formulas, one must know the number of samples to obtain the value of t, but the value of t depends on the number of samples, an obvious problem. If the number of samples is greater than about 30, the value for t at $n=\infty$ can be used without undue error (Sanders *et al.*, 1983, p. 158).

Ponce (1980) gives an iterative approach to calculate n if n is smaller than 30 or greater security is desired; the procedure converges rather quickly (usually 3 or 4 iterations). Make a reasonable guess as to the number of samples, use it to determine the t value, use the t value to calculate n. The correct value for n will lie somewhere between your guess and the calculated n, but closer to the calculated n. Choose another n accordingly, and repeat the procedure. The successive calculated values of n will show smaller and smaller fluctuations, eventually within one whole number. Use the next whole number as the sample size.

If the resulting number of samples exceeds about 10% of the number of possible observations (365 for daily loads on an annual basis), the estimated sample size should be revised using the finite population correction (Cochran, 1977):

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \qquad (19)$$

where n_0 is the initial estimate, N is the total number of possible observations, and n is the adjust estimate of the sample size needs.

To calculate the number of samples needed under flow-stratified sampling, the effective degrees of freedom for the combined strata must be determined, in order to obtain the proper value for t. The method is given in Darnell (1977), with an example. The effective degrees of freedom, f, is given by:

$$f_{eff} = \frac{\left(\sum_i \frac{N_i^2}{f_i + 1} s_i^2\right)^2}{\sum_i \frac{\left(\frac{N_i^2}{f_i + 1}\right)^2 s_i^4}{f_i}} \tag{20}$$

where f_i is the degrees of freedom associated with the estimate of the variance of stratum i. Once this value for f is determined, the corresponding value for t is found and the total number of samples is calculated according to:

$$n = \frac{t^2 \left(\sum_i \frac{N_i}{N} s_i\right)^2}{E^2} \tag{21}$$

Samples are allocated among strata by formula (25) (see below).

It must be emphasized that these approaches are only directly appropriate for random sampling of normally distributed populations. If the population is strongly skewed, or if systematic sampling is employed, and is more precise than random sampling, these calculations will indicate a larger sample than is required to obtain the specified precision. The amount of "overkill" cannot be known beforehand, and can be considerable in some cases. Similarly, the calculations assume no serial correlation and no periodicity in the data. If these are present, estimates of the variance will be distorted and lead to oversampling or undersampling. See Loftis and Ward (1980) for a study of the effects of autocorrelation and seasonality at different sampling frequencies, and Cochran (1963, p. 218ff) on systematic sampling of autocorrelated and periodic populations. Gilbert (1987, p. 50-53) also provides approaches for estimating sampling needs when observations are made systematically in time and are autocorrelated.

In the event that no prior evidence is available on the variance of the population, Sanders*et al.* (1983) suggest using formula 22 as an estimate of the variance, where RA is an estimate of the range of values in the population.

$$s^2 = \left(\frac{RA}{4}\right)^2 \tag{22}$$

Another alternative is to use the relationship

$$s^2 = \left(\frac{3}{4} IQR\right)^2 \tag{23}$$

where IQR is the interquartile range, i.e. the difference between the 75th percentile value and the 25th percentile value.

For some tributaries, there is a good correlation between flow variance and pollutant flux variance, particularly of particulate pollutants. In the absence of other information, flow data can be used to estimate a flux variance for particulate parameters. An example is provided in Richards (1989a).

Because the requirements of parametric approaches to calculating sample size are rarely met in environmental data, and the other approaches given above are approximate and/or empirical, the calculated sample size must be considered a first estimate. Furthermore, year-to-year variability will guarantee that a sampling program will perform better in some years than in others. For these reasons, the calculated sampling program should be considered a starting point, and adjusted as necessary in light of the results it provides.

Stratification

Stratification is the division of the sampling effort or the sample set into two or more parts which are different from each other but relatively homogeneous within. Stratifying the sampling program permits more of the sampling effort to be allocated to the aspects which are of greatest interest or which are most difficult to characterize because of great variability. Stratifying the data set may allow the calculation of a load which is both more precise and more accurate.

In load estimation for tributaries dominated by non-point source inputs (particularly particulate inputs), storm runoff is characterized by greatly increased flow and concentration. Thus much of the total variance of a sample of daily loads is due to runoff events. Stratification by flow allows runoff events to be sampled separately (more intensely) than low flow periods. For Lake Erie tributaries in northwest Ohio, experience has shown that the 20th percentile of flow (that flow exceeded 20% of the time) is a good cutoff for separating runoff events from low flow periods. Optimal allocation (see below) with this cutoff involves taking 60% to 85% of the samples from this stratum, depending on the river and the parameter of interest.

Seasonal stratification may be useful in some tributaries. For example, it may be important to treat the spring snow-melt period as a separate stratum in rivers in which most of the annual load is thought to be delivered during this time.

In some tributaries, particulate pollutants like suspended solids and total phosphorus rise in concentration faster than flow, and peak sooner. Thus one could consider creating separate strata for the rising and falling sides of hydrographs, though this stratification would be harder to implement than those discussed above.

Other stratification schemes may suggest themselves for specific rivers and parameters of interest. As a general rule, any appropriate stratification, properly executed, will result in more precise load estimates for a given sampling effort (but not necessarily for a given cost). The greatest benefit is generally derived from the first, most obvious stratification, and further stratification based on more subtle criteria may not lead to enough improvement to justify the added complexity.

Allocation of samples among strata

Many approaches are possible to allocating observations among the strata. The simplest is proportional sampling, in which the frequency of observations is in proportion to the expected populations of the strata, based on previous experience (see formula 24). Optimal allocation, also known as Neyman allocation, minimizes the estimate of the variance of the mean daily load, and is given by formula 25.

$$n_i = n\frac{N_i}{N} = n\frac{N_i}{\sum N_i} \qquad (24)$$

$$n_i = n\frac{N_i s_i}{\sum N_i s_i} \qquad (25)$$

Notation:
- n_i = the number of observations in stratum i,
- n = the total number of observations in the sample,
- N_i = the expected population of stratum i,
- N = the total population,
- s_i = the estimated standard deviation of the population in stratum i.

These formulas assume a fixed cost per observation regardless of stratum. If this assumption is not valid, see Cochran (1963, page 96) for a way to optimize the cost-effectiveness of the sampling program. Other allocations may be chosen, e.g. to intentionally oversample a stratum of particular concern.

A priori stratification vs. *a posteriori* stratification

A priori stratification is done before sampling and must be based on prior knowledge of the system. It allows different sampling patterns to be applied in different strata, and is the form of stratification discussed above. It is a stratification of sampling. *A posteriori* stratification is done after the observations are made, based on the observations themselves. It allows load calculations to employ strata not used in the collection of the sample. It of course does not permit concentration of sampling effort in selected strata (except by omission of observations, which would gain nothing). It is a stratification of the load calculation. *A posteriori* stratification can be superimposed on a sample collected with *a priori* stratification. Whereas *a priori* stratification almost always leads to improved precision for a given sampling effort, *a posteriori* stratification is not as successful, and can actually reduce precision. It is recommended that, if *a posteriori* stratification is contemplated, load calculations be made from the sample both with it and without it, and the more precise estimate be used. In the planning of a sampling program, it should be remembered that the main advantage of stratification is in concentrating sampling effort in the variable portions of the population; this can only be done with *a priori* stratification.

A posteriori stratification is used in the automated version of the Beale ratio estimator program described below. This program iteratively seeks out the stratification which minimizes the variance of the load estimate for a given set of data. As such, it automatically compares stratified with unstratified calculations, and only accepts stratification patterns which reduce the variance.

Load Estimation Approach

General Description

As suggested earlier, ratio estimators use the year's data to calculate a mean daily load, then use the mean flow from days lacking concentration data to adjust the mean daily load. The annual load is obtained by multiplying the mean daily load by 365 (or 366 for leap years).

The basis of ratio estimators is the assumption that the ratio of load to flow for the entire year should be the same as the ratio of load to flow on the days concentration was measured. Thus

$$\frac{\bar{l}_a}{\bar{q}_a} = \frac{\bar{l}_o}{\bar{q}_o} \qquad (26)$$

where the subscript a refers to an average for the year, and o refers to an average over the days on which concentration was observed. Assuming that flow is known for each day of the year, then

$$\bar{l}_a = \bar{l}_o \frac{\bar{q}_a}{\bar{q}_o} \qquad (27)$$

and

$$L = 365\bar{l}_a \qquad (28)$$

When the two parameters involved are correlated, as is almost always the case with flow and load, ratio estimators are biased, and a bias correction factor must be used.

The **Beale Ratio Estimator**, which has been widely used in Great Lakes loading calculations, is thoroughly discussed in Baun (1982), and is given by

$$\bar{l}_a = \bar{l}_o \frac{q_a}{q_o} \left[\frac{1 + \left(\frac{1}{n} - \frac{1}{N} \right) \frac{s_{lq}}{\bar{l}_o \bar{q}_o}}{1 + \left(\frac{1}{n} - \frac{1}{N} \right) \frac{s_{qq}}{\bar{q}_o^2}} \right] \qquad (29)$$

The term in square brackets is the bias correction term. In this term, s_{lq} is the covariance of flux and flow, s_{qq} is the variance of flow based on the days on which concentration was measured, and N is the expected population size. This requires some explanation. If the sampling interval is daily, the expected population size is 365 (except in leap years). Under stratified sampling, the expected population size is the probability of being in the given stratum times the sampling frequency in the stratum. For example, for flow stratified sampling using a cutoff at the 20th percentile (by time) of flow, and sampling four times a day, the expected population of the high flow stratum is 365*.20*4=292. If USGS mean daily flow values are used for adjusting the mean stratum loads, as is generally done, the unit loads must be daily, and N is forced to be 365*p, where p is the proportion of time occupied by the stratum. If chemical sampling is carried out more frequently than daily, daily average concentrations may be calculated, or one sample per day may be used, either chosen at random or systematically (e.g. the first sample or the middle sample in the day).

In the version of the formula normally seen in the literature, it is assumed that n is small compared to N, i.e. that most of the possible measurements of concentration are not made. As the sampling program approaches saturation (most of the possible observations are made) the bias correction converges to 1.

The expected bias of the Beale Ratio Estimator varies approximately as $\frac{1}{n^2}$ and thus approaches zero very quickly as n increases (Tin, 1965). Both theory and simulations show that when enough samples are taken to give an acceptable level of precision, the bias is unimportant.

The mean square error of the daily load is given by

$$MSE = \bar{l}^2\left[\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{s_{qq}}{\bar{q}^2}+\frac{s_{ll}}{\bar{l}^2}-2\frac{s_{lq}}{\bar{l}\bar{q}}\right)+\left(\frac{1}{n}-\frac{1}{N}\right)^2\left(2\frac{s_{qq}^2}{\bar{q}^4}-4\frac{s_{qq}}{\bar{q}^2}\frac{s_{lq}}{\bar{l}\bar{q}}+\frac{s_{lq}^2}{\left(\bar{l}\bar{q}\right)^2}+\frac{s_{qq}}{\bar{q}^2}\frac{s_{ll}}{\bar{l}^2}\right)\right] \quad (30)$$

In this formula, all terms involving q are based on the flows corresponding to concentration observations. The RMSE is the square root of this number. The RMSE of the annual load is 365 times the RMSE of the daily load. Since the bias is small, the MSE is equivalent to the variance estimate.

Formula (30) is essentially that given by Baun (1982) except for the factor of 2 in the third term, which was inadvertently omitted by Baun. It differs from that of Tin (1965) in omitting Tin's third term, which is negligible for n greater than about 10. Tin's formula is for the MSE of the ratio of load to flow, rather than the load itself; thus the term to the left of the square brackets is the ratio squared rather than the mean daily load squared. Further discussion and formulas are to be found in Cochran (1977), Chapter 6.

It is not recommended that the Beale Ratio Estimator be used with log-transformed data. It does not assume a normal distribution, and corrections for possible back-transformation biases such as those discussed for regression estimators above have not been worked out, and would be extremely complex.

Computer Programs

Two versions of the Beale Ratio Estimator have been implemented as computer programs available for Macintosh personal computers, and are provided on the CD-ROM. One allows the manual entry and adjustment of stratum boundaries, with stratification by flow and/or by time. This program is interactive, with a graphic user interface which displays the flow and concentration data and the current stratification. It reports the load and its confidence interval given the current stratification, and continuously updates these figures as the stratification is changed. This version of the program can be used to calculate the load from a stratified sampling program, or to explore the possible advantages of post-stratification.

The other version is an automated iterative program which seeks out the post-stratification (by time) which yields the lowest mean square error for a given set of data. This version is not specifically designed to work with data resulting from a stratified sampling program, and its use with a stratified sample might (or might not) produce biased results. For example, if used with a sampling program which obtained a disproportionate number of its samples during high flow periods, it might yield load estimates which were biased high.

Brief Example: Atrazine Loads for the Sandusky River, 1993

The Heidelberg College Water Quality Laboratory collected 76 samples for analysis of atrazine concentrations during the pesticide year 1993, defined as May 1, 1993 to April 30, 1994. The sampling program is stratified by time and flow: during May through August ("summer" for short), samples are scheduled to be taken twice weekly during low flow and three times daily during high flow; during the rest of the year samples are taken about every two weeks. (Since the Beale Ratio Estimator program is not designed to work with more than one sample per day, it averages the multiple samples taken during the summer high flow days, and uses the daily average concentration for each such day in its calculations.)

The annual hydrograph, with superimposed atrazine concentrations, is shown in Figure 7. The sampling stratification is also indicated. Note that most of the flow occurs during the winter when concentrations are low, and during a brief period of time in early June, when concentrations are at their highest for the year.

The first approach to calculating the atrazine load is a bad one: no stratification is applied in the calculation, so Stratum 1 has all the data. This is a bad idea because observations of concentration were made more frequently during high flow periods in summer, when concentrations tend to be high, than at other times of year. Treating all the data together will give undue weight to the high concentration period, leading to a biased load estimate. This approach is applied for illustration purposes only. In general, the stratification applied at sampling time should always be respected.

The second approach is to use the stratification defined by the sampling program. Stratum 1 contains the low-flow summer data, Stratum 2 contains the high flow summer data, and Stratum 3 contains the data for the rest of the year. These strata are demarked by the heavy lines in Figure 7.

The third approach introduces additional time stratification to create an early summer stratum (no high flows occurred during this time, so flow stratification was not necessary), mid-summer low- and high-flow strata, a late summer (low flow) stratum, and two winter strata, one characterized by low concentrations and low flows and the other by high flows but nearly zero concentrations. The additional time boundaries are shown with dashed lines in Figure 7.

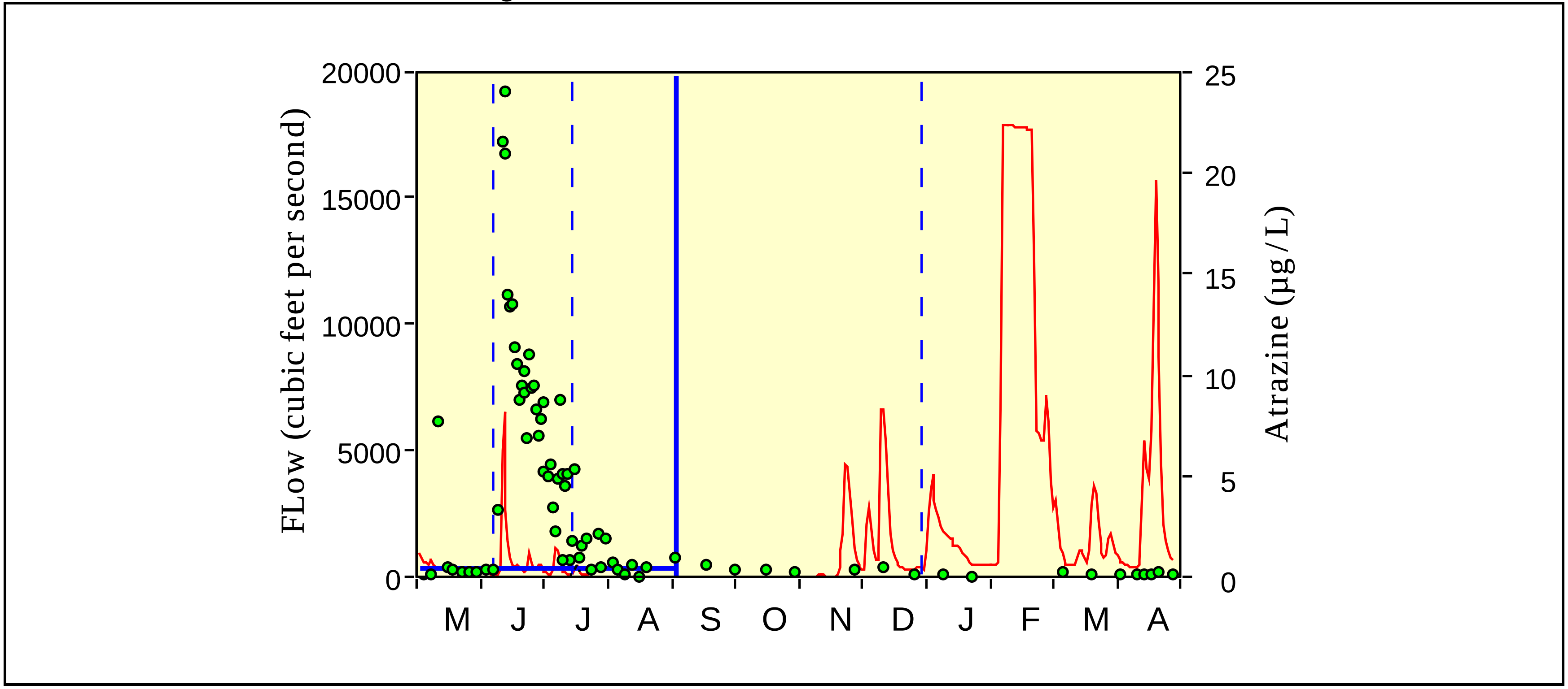

Figure 7. Sandusky River atrazine and flow data, May 1993 through April 1994, showing flow and time stratification applied in calculating the loads.

The fourth approach is to offer the data to a version of the program which automatically seeks out the time stratification which yields the smallest confidence interval. With this data, this program established 7 stratum boundaries, too numerous to list in the table.

The results of these calculations are shown in Table 6. The load calculated using unstratified data is more than five times as high as the other two, and also has a broader confidence interval, both absolutely and as a percent of the load. This occurs because the wintertime high flows cause the

Table 6. Results of the load calculations. Each stratum cell reports the mean daily load for the stratum in kg/day, the number of days in the stratum, and the flow ratio used to adjust the load.

Approach	Individual Stratum Results						Total Load (kg)	95% Confidence Interval
	1	2	3	4	5	6		
1	27.9 365 1.90						10,183	±7664 (±75.3%)
2	1.79 98 0.82	50.0 25 0.82	1.78 233 1.16				1887	±519 (±27.5%)
3	1.0 34 1.01	4.44 19 0.96	68.93 15 1.00	0.48 55 0.71	0.80 117 0.49	2.36 116 1.57	1587	±382 (±24.1%)
4							1519	±129 (±8.5%)

average load, based mostly on the summer high concentration values, to be adjusted upward and then to be applied to all days, including the winter when the concentrations are obviously relatively low. While there is no way to know the true load for this year, the patterns in the data clearly indicate that a load calculated in this way is strongly biased. The other two loads, calculated using different stratification schemes, are much more comparable. Indeed, if the confidence intervals are accurate, these two load estimates are not significantly different from each other.

The addition of further strata in the third approach has led to a smaller confidence interval, but the change in this case is not dramatic. The automated approach found a more detailed stratification which has a substantially smaller confidence interval, but the load is not greatly different from those generated by the other stratified approaches.

Detailed Example: A Sampling Program for Fowl Creek

Fowl Creek is a small hypothetical river in an unspecified midwestern state. It is the site of a BMP implementation program, the goal of which is to reduce phosphorus loading entering Carp Lake. A monitoring station was set up near the mouth of Fowl Creek where it enters the lake. The project plan called for annual loads to be calculated for a five year pre-implementation period, and a five year post-implementation period, separated by a six year implementation period. In order to have a good chance of detecting a change in the loads, the project managers set a goal of measuring loads with an error of not more than 15%.

No previous monitoring data were available for Fowl Creek. Mean daily flow records were available from the U.S. Geological Survey for a nearby creek with similar land use. A small set of observations of phosphorus concentrations and flows were available from another creek in the next county.

In a paper by Richards (1989a), the project scientist found the following empirical formulas relating the coefficient of variation of fluxes of various parameters to a measure of the variability of flows called CVLF5, which is the coefficient of variation of the logs of flows corresponding to the percentiles of the flow distribution {5, 10, 15, 20, ..., 80, 85, 90, 95}.

Table 7. Empirical relationships between CVLF5 and the coefficient of variation of fluxes (corrected from Richards, 1989a)

Parameter	Formula
Suspended Solids	$cv = 8.09\ CVLF5^{.6259}$
Total Phosphorus	$cv = 5.82\ CVLF5^{.5976}$
Chloride	$cv = 2.84\ CVLF5^{.7074}$

Working with the mean daily flow data from the nearby creek, she determined that the relevant percentiles of log-flow were {0.40, 0.52, 0.61, 0.67, 0.73, 0.79, 0.82, 0.86, 0.92, 0.98, 1.04, 1.11, 1.23, 1.30, 1.46, 1.62, 1.79, 1.96, 2.34}. These percentiles have a standard deviation of 0.52 and a mean of 1.11, so the CVLF5 is 0.52/1.11 or 0.47. She entered this value into the relationship for phosphorus, and obtained an estimated coefficient of variation:

$$cv = 5.82 * (0.47)^{.5076} = 3.97$$

Working with the small set of concentration and flow data, she multiplied them together to obtain flux values and calculated the variance and other distribution statistics shown on the next page.

Table 8. TP flux dataset and distribution statistics

Date	Flow	TP concentration	TP flux
890113	280.00	.030	8.300
890205	6.22	.036	.224
890319	24.40	.067	1.635
890409	52.00	.065	3.380
890430	11.00	.031	.341
890521	11.00	.066	.726
890611	7.24	.092	.666
890702	4.34	.080	.347
890730	8.76	.381	3.338
890820	4.02	.104	.418
890910	3.40	.111	.377
891001	2.16	.033	.071
891022	14.60	.214	3.124
891112	5.59	.051	.285
891203	4.65	.027	.126
891224	4.96	.024	.119
		Mean	1.467
		Std. Dev.	2.178
		Variance	4.742
		Coeff. of Var.	1.484
		Interquartile Range	1.737
		Range	8.229

Concerned about the small size of the data set and its strongly skewed distribution, she also used formulas 22 and 23 to estimate variances for this data:

$$s^2 = \left(\frac{\text{Range}}{4}\right)^2 = \left(\frac{8.23}{4}\right)^2 = 4.23$$

$$s^2 = \left(\frac{3}{4}\,\text{IQR}\right)^2 = \left(\frac{3}{4}1.74\right)^2 = 1.70$$

These two variance estimates are quite different because the largest flux is very large compared to the rest, and stretches the range but does not affect the interquartile range. The parametric variance estimate is the largest of the three, presumably due to the extreme nature of the highest value.

The variability estimates were then entered into the proper version of formula 18/18a, along with the acceptable error value for formula 18, calculated as 0.15 times the mean flux:

18a: $$n = \frac{t^2 (cv)^2}{p^2} = \frac{1.96^2 * 3.97^2}{.15^2} = 2691$$

18a: $$n = \frac{t^2 (cv)^2}{p^2} = \frac{1.96^2 * 1.48^2}{.15^2} = 373$$

18: $$n = \frac{t^2 s^2}{E^2} = \frac{1.96^2 * 4.23}{(.15 * 1.467)^2} = 336$$

18: $$n = \frac{t^2 s^2}{E^2} = \frac{1.96^2 * 1.70}{(.15 * 1.467)^2} = 135$$

Since the number of possible observations of the unit (daily) loads is 365, all of these estimates require more than a 10% sampling rate, and the finite population correction (Formula 19) should be used to adjust them:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{2691}{1 + \frac{2691}{365}} = 321$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{373}{1 + \frac{373}{365}} = 184$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{336}{1 + \frac{336}{365}} = 175$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{135}{1 + \frac{135}{365}} = 99$$

While these estimates of sample size offer a little more hope than the initial estimates, they still imply a very intense sampling program. The project scientist reasons that most of the variability is due to storm runoff periods, which occupy a fairly small percentage of the time but account for much of the annual load. It should be adequate to sample intensively during storm runoff periods, less intensively during low flow periods, and calculate the loads using the Stratified Beale Ratio Estimator. Using the general rule of thumb that at least 80 percent of the flow and flux occur in the 20 percent of the days with highest flows, the decision is made to stratify at the 20th (upper) percentile of flow, which is calculated to be 42 cfs. During high flow periods, sampling will occur at the rate indicated by the above calculations. In addition, one sample per month will be taken during low flow periods. Noting that the intermediate two sample size estimates are comparable and correspond approximately to sampling every other day, the expectation is that this sampling program will require 49 samples per year, based on an expected 73 days of high flow per year sampled every other day, plus 12 low flow samples.

This sampling program is implemented on Fowl Creek and operated for the first year of the pre-implementation period; 47 samples were taken. The resulting data are shown below, as are the total phosphorus loads and percent error estimates, as calculated by T&F Beale (see next section) using the same stratification as the sampling program, and as calculated by Autobeale using a more detailed time stratification which seeks to minimize the error estimate.

Table 9. Data gathered during the first year of sampling on Fowl Creek.

Date	Flow (cfs)	Total Phosphorus (mg/L)	Flow stratum
1-Jan-93	178.0	.300	HIGH
4-Jan-93	606.0	.989	HIGH
6-Jan-93	180.0	.441	HIGH
13-Jan-93	282.0	.253	HIGH
15-Jan-93	44.4	.116	HIGH
16-Jan-93	28.0	.081	low
23-Jan-93	87.4	.226	HIGH
25-Jan-93	220.0	.387	HIGH
15-Feb-93	16.3	.067	low
24-Feb-93	63.2	.064	HIGH
3-Mar-93	68.7	.214	HIGH
5-Mar-93	274.0	.311	HIGH
7-Mar-93	197.0	.136	HIGH
9-Mar-93	351.0	.243	HIGH
11-Mar-93	118.0	.112	HIGH
13-Mar-93	52.2	.073	HIGH
15-Mar-93	54.9	.043	HIGH
15-Mar-93	33.6	.036	low
17-Mar-93	338.0	.321	HIGH
19-Mar-93	133.0	.108	HIGH

Table 9. Data gathered during the first year of sampling on Fowl Creek, concluded.

Date	Flow (cfs)	Total Phosphorus (mg/L)	Flow stratum
21-Mar-93	241.0	.378	HIGH
23-Mar-93	116.0	.206	HIGH
25-Mar-93	241.0	.226	HIGH
27-Mar-93	77.0	.093	HIGH
2-Apr-93	384.0	.389	HIGH
4-Apr-93	45.7	.100	HIGH
10-Apr-93	684.0	.611	HIGH
12-Apr-93	48.3	.142	HIGH
15-Apr-93	21.7	.043	low
16-Apr-93	120.0	.381	HIGH
17-Apr-93	81.2	.270	HIGH
15-May-93	8.2	.038	low
9-Jun-93	508.0	.164	HIGH
11-Jun-93	70.1	.270	HIGH
15-Jun-93	11.2	.149	low
21-Jun-93	285.0	.937	HIGH
15-Jul-93	6.2	.176	low
15-Aug-93	2.9	.044	low
15-Sep-93	2.9	.059	low
15-Oct-93	4.3	.026	low
15-Nov-93	29.0	.088	low
17-Nov-93	325.0	.786	HIGH
19-Nov-93	54.9	.474	HIGH
28-Nov-93	143.0	.535	HIGH
4-Dec-93	61.8	.233	HIGH
6-Dec-93	120.0	.492	HIGH
15-Dec-93	9.7	.033	low

Table 10. Loads produced by the samples obtained in the first year

Loads calculated by T&F Beale			
Mean Daily Load (kg)	39.67	Annual Load (tonnes)	14.48
RMSE	8.50	RMSE	3.102
cv (%)	21.4	cv (%)	21.4
Loads calculated by Autobeale			
Mean Daily Load (kg)	39.05	Annual Load (tonnes)	14.25
RMSE	2.55	RMSE	0.932
cv (%)	6.54	cv (%)	6.54

The achieved error rate is larger than the target ±15%, but only slightly, when the load is calculated using the same stratification as was applied in sampling. Furthermore, the achieved error rate is substantially lower than the target when Autobeale is allowed to impose optimal stratification. Therefore, adjustment of the sampling program is not called for at this point. However, if adjustment were necessary, it could be done by noting that, for a given station and t-value, Formula 18a reduces to:

$$n = \frac{t^2 (cv)^2}{p^2} = \frac{k_1 * k_2}{p^2} = \frac{k}{p^2} \qquad (31)$$

Taking ratios of both sides, we get

$$\frac{n_0}{n_1} = \frac{\dfrac{k}{p_0^2}}{\dfrac{k}{p_1^2}} \qquad (32)$$

which can be re-expressed as:

$$n_1 = n_0 \frac{p_0^2}{p_1^2} \qquad (33)$$

where subscript 1 refers to the adjusted sample size and subscript 0 refers to the original sample size. If adjustment were desired, the calculation would give (using the 21.4% cv as an example)

$n_1 = 47 \dfrac{21.4^2}{15^2} = 95$, about twice as many samples.

ACCESS TO TOOLS: HOW TO AVOID DOING IT ALL YOURSELF

This section provides further descriptions of the programs described above, plus information about several other programs which are available for calculating loads. Most of the programs are provided on the CD-ROM, along with instructions for their use and sample data sets.

Beale programs

T&F Beale is a Macintosh implementation of the Beale Ratio Estimator which includes a graphic user interface, by means of which stratification can be defined and adjusted. There is a single input file which lists the mean daily flows for each day of the year and the chemical observations on the days when samples were taken. The graphic user interface displays the mean daily flow data for the year, the chemical observations, and the stratification applied. Stratification can be defined according to flow or time, or combinations of time and flow, and can be entered by typing values into a dialog box or by use of the graphic user interface. As new stratum boundaries are added or existing stratum boundaries are moved, the load estimate based on the current configuration is reported, along with its error estimate and the effective degrees of freedom. When the user is satisfied with the stratification applied, the final load is calculated. Output file options include a short report of the load and its error; a stratification file which can be used to automatically reproduce the final stratum configuration; and a detailed report showing the placement of each observation into each stratum, the load and error estimate for each stratum, and the overall load. Data for more than one load estimate can be combined in a single input file, and the corresponding output file will list loads for each year/parameter for which data is present in the input file. Technical note: T&F Beale does not use the finite population correction (the 1/N term in formulas 29 and 30).

Autobeale is a Macintosh implementation of the Beale Ratio Estimator which automatically and iteratively seeks out the time stratification which produces the lowest error estimate for a given set of data. It requires two input files, one of which lists the mean daily flows for each day of the year, and the other of which lists the chemical observations on the days when samples were taken. A graphic interface displays the mean daily flow data for the year, the chemical observations, and the stratification, as it is being automatically applied and adjusted. As new stratum boundaries are added or existing stratum boundaries are moved, the load estimate based on the current configuration is reported, along with its error estimate and the effective degrees of freedom, though usually these values are changing too rapidly to read! Output file options include a short report of the load and its error; a merged flow and concentration file which can be used as input to **T&F Beale**; and a detailed report showing the placement of each observation into each stratum, the load and error estimate for each stratum, and the overall load. Data for more than one load estimate can be combined in a single input file, however all load estimates must be for the same station and year, since they will all use the same mean daily flow data. The corresponding output file will list loads for each parameter for which data is present in the input file. Technical note: Autobeale does not use the finite population correction (the 1/N term in formulas 29 and 30).

The USGS Estimator program

Estimator implements the USGS seven-parameter regression approach described above. It develops a relationship between the log of concentration and log-flow, log-flow squared, decimal time, decimal time squared, and seasonality, represented by sine and cosine terms. The program requires two input files. The first contains dates with observations of flow and concentration, and is used to calibrate the regression model. The second contains dates and mean daily flows for the period for which loads are to be calculated. The mean daily flows and times are entered into the regression model, and an estimated daily load is calculated for each date, using the MVUE bias correction approach (see description earlier in this document). A total load for the interval is also calculated as the sum of the daily loads. The output file contains the input parameters, the results of the regression analysis, a number of diagnostics for evaluating the appropriateness of the regression model, and a tabular listing of the loads and their uncertainties by month and for the year(s) for which flow data were provided. An optional output file provides a listing of the daily loads.

Regression with Centering is an Excel workbook version of the seven-parameter model. The same types of input information are required, but the format is different and more comfortable for those accustomed to working in spreadsheet programs. The final sheet of the workbook provides a listing of daily loads for the period of interest, calculated using the QMLE and Smearing bias correction approaches (see description earlier in this document) as well as without bias correction for comparison purposes. The necessary input formats are described in the accompanying document Regression Instructions.

Integrator

Integrator is an implementation of the numeric integration method of load estimation, using the trapezoidal cell approach. It is appropriate for use only if the sampling program is detailed enough to provide good resolution of times of rapidly changing flux. In practice, this means that for non-point pollutants there should be at least a dozen samples of concentration and flow during the runoff period from each storm in the year. If such highly detailed datasets are available, it is probably the most straight-forward way to calculate a load. Since it is not a statistical approach, it is not accompanied by an error estimate. However, a novel approach to estimating upper- and lower-bound limits on the load has been added in Integrator, and these can serve as a measure of uncertainty in the load estimate. A detailed discussion of the approach is provided in the User's Manual, as is information on the input and output formats.

Walker's Army Corps set of programs

FLUX is one of three programs developed for the Army Corps of Engineers by Wm. W. Walker, for use in assessing and predicting eutrophication in reservoirs. Flux calculates loads by six different approaches, including two ratio estimates and three regression estimates. Uncertainty estimates are provided with each approach. Stratification may be applied to the data. Flux provides a series of graphical and tabular output formats. One feature of this program is its focus on providing information on how the sampling program might be modified to improve the load estimates. FLUX runs on IBM computers in the DOS environment. It is accompanied by an extensive manual.

FLUX is available from the US Army Engineer Waterways Experiment Station, Vicksburg, MS 39180-6199. The User Manual is Instruction Report W-96-2.

A FINAL, SOBERING EXAMPLE

As a comparative experiment, a year's worth of daily suspended sediment data were extracted from the Water Quality Laboratory's Maumee River database, for the water year 1991 (October 1990 - September 1991). Several days had missing data, and the total number of observations was 347. These data were split into seven groups based on the day of the week, to simulate weekly sampling. This gives seven presumably equivalent datasets, each of which would be considered highly detailed by most agencies. Each dataset was used to calculate annual loads by three methods:

1. regressing log-concentration on log-flow, and using half the variance of the residuals for bias correction during back transformation (this is not the full USGS regression model),

2. applying the Beale Ratio Estimator with stratification at the 80th percentile of flow, and

3. applying the Autobeale program and allowing it to choose the stratification which minimizes the confidence interval around the annual load estimate.

In addition, annual loads were calculated by each method using all of the data. Since samples were taken on nearly every day of the year, it would seem reasonable to consider these estimates the "true" loads for the year and compare the results from weekly sampling with them. We expect these "true" loads to be very similar. A final "true" load was calculated using numeric integration. In this approach, the time interval for any sample was limited to one day, and the final load was adjusted by dividing by the proportion of days in the year with samples. This load estimate was chosen as the basis against which to compare the other loads estimates.

The results are listed in Table 11 and shown in Figure 8. In the figure, diamonds show the annual loads for the Beale Ratio Estimator stratified at the 80th percentile of flow, squares the results of the Autobeale stratification, and circles the results of the regression approach. The four lines show five "true" load values: the top line corresponds to the numeric integration load and the one from Beale Ratio Estimator with flow stratification (which are too similar to be resolved in the graph), the next one down represents the Autobeale. The third from the top will be discussed later. The lowest line represents the “true” load for the regression approach, and it is clearly very much lower than the others.

Table 11. Load estimates produced by three methods applied to Maumee River 1991 water year suspended sediment data, to seven subsamples representing weekly sampling, and to a subsample representing flow-stratified sampling. The loads and confidence intervals are expressed in thousands of metric tons per year. The percent error is calculated as 100*(estimate-trueload)/trueload.

Sample	Method	n	load	95% conf. int.	% error
All data	Numeric Integration (the “true” load)	347	2386		0
	Autobeale	347	2365	± 415	-1
	T&F Beale 80	347	2405	± 1021	1
	Regression	347	1043		-56
	Seasonal Regression	347	1902		-20
Weekly samples on Sundays	Autobeale	48	1910	± 1036	-20
	T&F Beale 80	48	3146	± 2649	32
	Regression	48	1232		-48
Weekly samples on Mondays	Autobeale	53	2544	± 2357	7
	T&F Beale 80	53	4409	± 4976	85
	Regression	53	1035		-57
Weekly samples on Tuesdays	Autobeale	50	2366	± 1815	-1
	T&F Beale 80	50	3566	± 3542	49
	Regression	50	1024		-57

Table 11. Load estimates produced by three methods, concluded.

Sample	Method	n	load	95% conf. int.	% error
Weekly samples on Wednesdays	Autobeale	50	1912	± 1112	-20
	T&F Beale 80	50	2541	± 2013	6
	Regression	50	1186		-50
Weekly samples on Thursdays	Autobeale	49	1079	± 235	-55
	T&F Beale 80	49	1120	± 245	-53
	Regression	49	991		-58
Weekly samples on Fridays	Autobeale	48	864	± 173	-64
	T&F Beale 80	48	1023	± 257	-57
	Regression	48	867		-64
Weekly samples on Saturdays	Autobeale	49	1111	± 301	-53
	T&F Beale 80	49	1579	± 854	-34
	Regression	49	1026		-57
Flow stratified sampling	Autobeale	47	2872	± 418	20
	T&F Beale 90	47	2404	± 964	1
	Regression	47	1917		-20
	Flow-stratified Regression	47	2026		-15

There is a wide range of load estimates using the flow-stratified Beale calculations, a consequence entirely of differences in the data which happen to fall on different days of the week. Several of the load estimates have very large confidence intervals; in fact, the loads for Mondays and Tuesdays are not significantly different from 0! The Autobeale results are less variable, but either have rather large uncertainties or are substantially too low. An examination of the data reveals the reason. Half of the suspended solids load *for the entire year* occurred during four successive days in a major storm runoff event. The one-sample-per-week data subsets which captured one of these days (Sunday -

Wednesday) had high loads and high variability as a consequence of this extreme event; the others had relatively low variability but low annual load estimates because they missed all but the tail end of the one important event of the year.

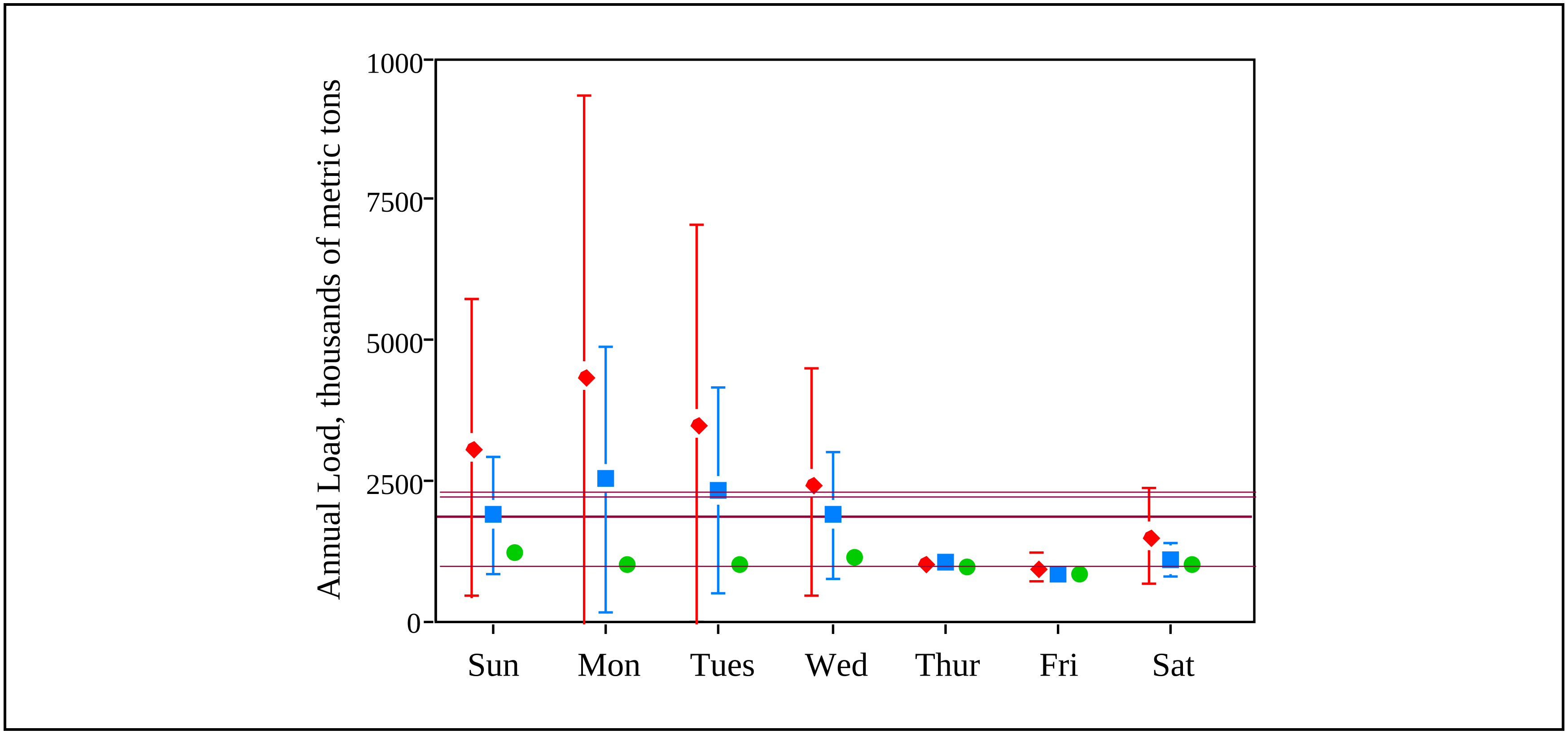

Figure 8. Loads from seven “equivalent” sets of weekly sampling for suspended solids, 1991 Water Year, Maumee River. Error bars are 95% confidence intervals. Diamonds: Beale Ratio Estimator stratified at 80th percentile of flow. Squares: Autobeale. Circles: Regression approach.

Thus these inconsistent results are not due to some weakness of the load calculation methodology, but result from a sampling program with too few samples to reliably capture the action, at least in this particular year. A stratified sampling program which took daily samples during the upper 10% of the flows and one sample per month during the lower 90% of the flows would have captured this and other runoff events and produced a reliable load estimate with no more samples. Simulating this program by subsampling the 1991 dataset gave a dataset of 47 observations and a load estimate of 2404 thousand tons ± 964 using T&F Beale with calculations stratified at the 90th percentile of flow. Autobeale, applied to the same data, produced an estimate of 2872 ± 418.

The loads calculated using the regression approach are much more consistent, but are clearly biased low relative to most of the loads calculated using other approaches. Since the very straight-forward numeric integration approach gives a “true” load similar to the ratio estimator results, we conclude that the regression results are wrong, even though they are more consistent from subset to subset of the data.

In an effort to understand why the regression results are so different, we examine a scatterplot of log-concentration against log-flow, shown in Figure 9. It is immediately clear that the regression relationship badly underestimates concentrations at high flows. Since the days with high flows contribute very heavily to the annual load, this underestimation is the cause of the low bias. It also appears that the data may fall into two clusters, one with a steeper slope than the other. A little exploration soon shows that there are two coherent clusters based on seasons of the year, one a “winter” cluster containing data for the months November through April, and the other a “summer” cluster containing the data for May through October. As shown in Figure 10, these two subsets of data have very different regression relationships.

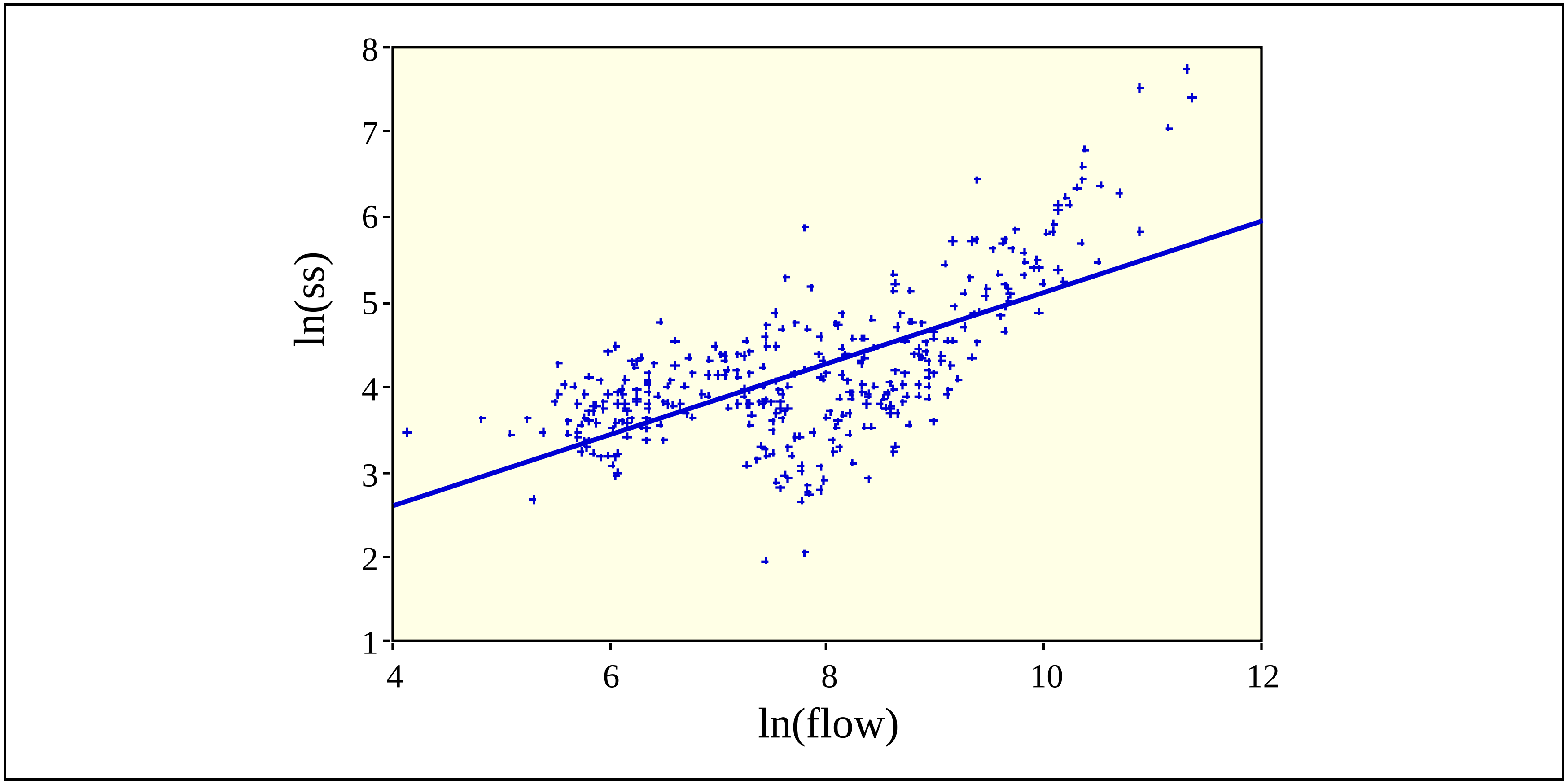

Figure 9. Relationship between log of suspended sediment concentration and log of flow for the entire 1991 water year dataset.

Armed with this information, we returned to the calculation of the “true” load using the regression approach. We split the data into the two seasons defined above, and calculated daily loads for each season using the regression relationship based on that season’s data. These daily loads were summed up to provide estimated winter and summer loads, and these were added together to provide a “true” annual load based on seasonal regression. This load value is the remaining line shown in Figure 8, and it is much closer to the other “true” loads, though it is still on the low side. This confirms that the structure of the data led to a serious low bias when the regression approach was applied to the data for the year as a whole.

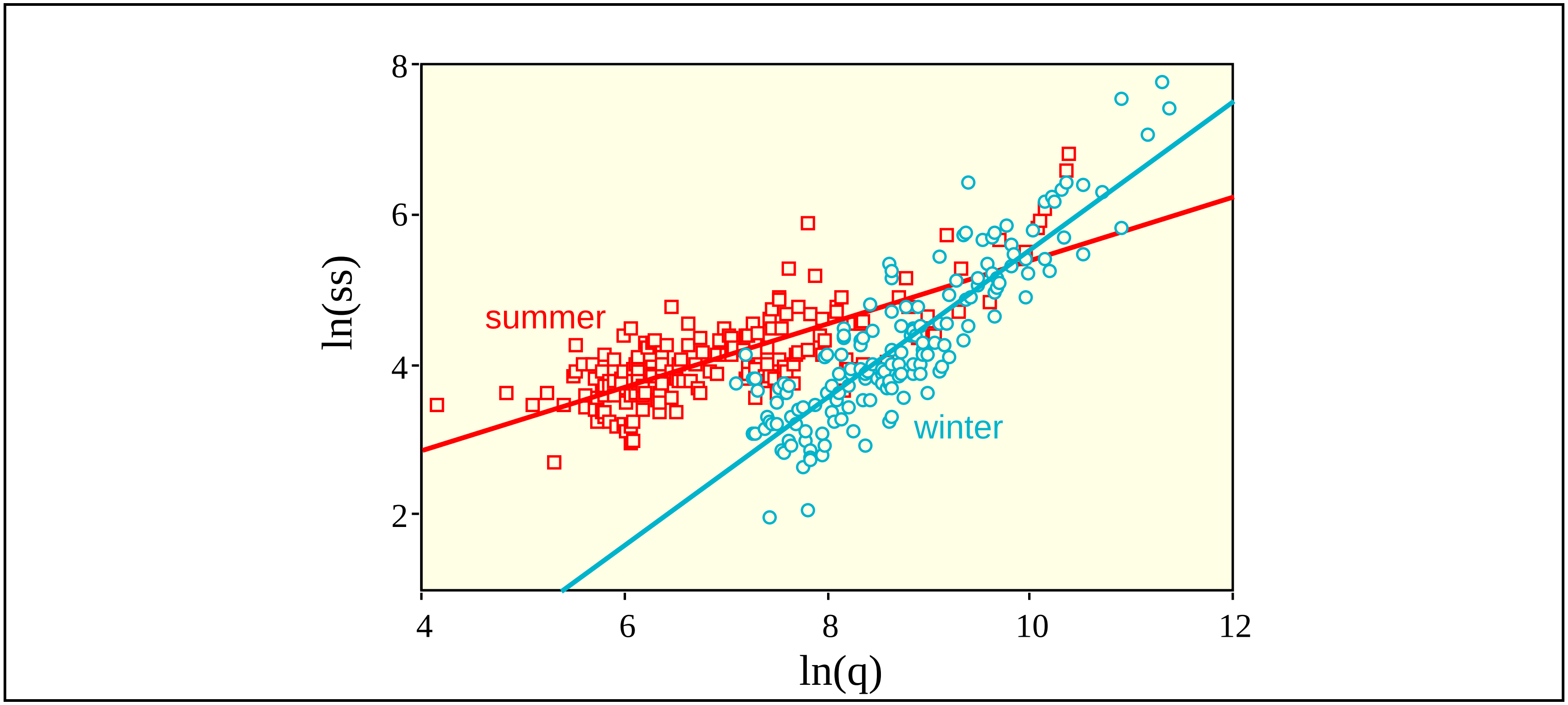

Fig 10. Suspended sediment concentration and flow data divided into two groups based on time of year.

We might revisit the subsets of data which represent weekly sampling, separate them into the same seasonal groups, and recalculate the loads by season. We might also consider applying a seasonal stratification to the Beale Ratio Estimator calculations, in addition to or in place of the flow stratification. However, it is worth pointing out that only the detail present in the entire dataset allowed us to identify the two clusters. Had we sampled only weekly, we might not have suspected that the clusters were present.

The regression approach was also applied to the flow-stratified dataset, both as a whole and treating the two flow groups separately. For the stratified calculations, the mean daily flow values were split into high and low flow groups at their 90th percentile. The stratified calculation gave a better result than the unstratified one, but neither result was as satisfactory as that from the flow-stratified Beale Ratio Estimator.

This example provides several important lessons:

• The performance of a load estimation approach may suffer severely when its underlying assumptions are not met. In this case, the assumption that there was one linear regression relationship between log-concentration and log-flow was not valid. Forcing a single relationship onto the data produced load estimates which were consistently 50% or less of the true load.

• For most non-point source pollutants, the majority of the load is transported in a very short period of time of high flow. Flow-stratified sampling is a very efficient way to obtain detailed data for this critical time without the cost of large numbers of samples. When flow-stratified sampling is employed, the load calculations should be stratified using the same flow cutoff(s), particularly with the Beale Ratio Estimator. Additional stratification at calculation time may not be necessary with flow-stratified sampling.

• It is very important to sample as thoroughly as possible, to **examine the resulting data**, to approach load estimation thoughtfully, and to use all available tools to try to understand the structure in the data. Calculating loads by blindly running data through a computer program is a hazardous exercise!

AFTERWORD

Designing adequate sampling programs for tributary load estimation is difficult. Populations generally fail to satisfy the assumptions of normality which are implicit in sampling program design and often in load calculation procedures. Autocorrelation and seasonality are generally present. Systematic sampling is generally used for convenience, and the possible differences from random sampling ignored. Several suggestions are offered which can help avoid the pitfalls that these problems can offer. They can be summarized by the paradox: You can't design a good sampling program without first having some data from one.

Whenever possible, one or more years of observations should be made with high density compared to the contemplated sampling program. Daily observations would be ideal; more frequent observations may be necessary for small tributaries. Where analytical costs are prohibitive, it may be possible to find a less expensive parameter, the behavior of which correlates strongly with the parameter of interest. This proxy parameter can then be used to provide much of the detailed information (with obvious hazards). Flow data should be scrutinized in as much detail as available, at least for some intervals of time. Many tributaries have nearly continuous traces of flow (hourly or more frequently). These observations should be plotted, and the characteristic patterns identified. What does the spring thaw period look like? How do flow and concentration behave during storm runoff periods? How do storm runoff periods compare in importance with spring thaw? Are there signs of seasonality? What happens to concentrations in winter? To flows in late summer? Is there any evidence of weekly periodicity in concentrations, flows, or fluxes? Do detailed flow records indicate any diurnal fluctuations in flow? How variable are flows and fluxes? Does the variance increase with increasing flux? What is the shape of the histogram of flux observations? Of flows?

Much can be learned just from studying plots of detailed data. Time series analysis can reveal patterns not apparent visually. With this information in hand, a good attempt can be made to design a sampling program which will meet design goals, and do so reasonably efficiently. Proposed sampling programs can be evaluated by simulation from the detailed data sets. This is often an important verification step, because the systems are so badly behaved statistically. Indeed, when this has been done, it has often revealed differences between actual behavior and that expected from theory, and virtually always revealed differences in results as a function of methodology.

As experience with the system develops, fine tuning of the sampling program can be done. A sampling program should never be considered static. It should evolve in response to increasing understanding of the populations being sampled, and in response to the evolving needs of the program of which it is a part. At the same time, it should always be remembered that load estimates are a function not only of what goes down the river, but of how we sample and calculate loads. Sudden drastic changes in sampling strategy are as hazardous as sudden changes in techniques of chemical analysis.

In spite of all difficulties, a well thought out sampling program can perform very well. In contrast, programs based on expediency and ignorance can perform so poorly that when a really significant observation comes along, such as a daily load during a major storm, it is thrown out as a statistical outlier! Inadequate programs have been common in the past; hopefully they will be rare in the future!

AN ANNOTATED BIBLIOGRAPHY OF CITED PAPERS
AND OTHER LITERATURE ON LOAD ESTIMATION

This section includes entries for a number of papers which were not specifically cited in the main body of the document. In carrying out the literature search, emphasis was placed on works which describe, develop, or evaluate specific techniques for load estimation based on monitoring data. Little emphasis was placed on model-based loading approaches, and on works which report calculated loads without comments on the methodology.

Baker, D.B. 1993. The Lake Erie Agroecosystem Program: water quality assessments. *Agriculture, Ecosystems and Environment* 46: 197-215.
A good general summary of the highly detailed watershed-scale monitoring project of the Water Quality Laboratory, which has produced daily and more frequent data for over 20 years at several stations located on Lake Erie tributaries.

Baker, D.B. 1988. Sediment, Nutrient, and Pesticide Transport in Selected Lower Great Lakes Tributaries. EPA-905/4-88-001. U.S. EPA, Great Lakes Program Office, Chicago, IL. 225 pp.
A description of a set of Lake Erie tributaries which are among the most thoroughly characterized anywhere. Unfortunately, the report is not readily available. Other descriptions of many aspects of these data are presented in other papers of D.B. Baker and R.P. Richards listed in this bibliography.

Baun, K. 1982. *Alternative Methods of Estimating Pollutant Loads in Flowing Water*. Tech. Bulletin 133, Dept. Natural Resources, Madison, Wisconsin. 11 pages.
Presents the concept of stratified sampling and the Beale Ratio Estimator, and compares it with "integration analysis" which is a combination of numeric integration, rating curve techniques, and the worked record in the terms of this document. Also compares composite sampling. Good discussion of pros and cons of each method. Interesting that the work was done in the context of urban runoff studies, whereas the BRE has been used mostly in rural or large watershed settings. Comes out in favor of stratified sampling and the BRE.

Beale, E.M.L. 1962. Some uses of computers in operational research. *Industrielle Organisation 31*: 51-52.

A rather strange little paper from a perspective early in the computer era, which is remembered because it apparently is the origin of the Beale Ratio Estimator, though the paper presents only the barest outline of the current version of this estimator.

Bierman, V.J., Jr., S.D. Preston, and S.E. Silliman. 1988. *Development of Estimation Methods for Tributary Loading Rates of Toxic Chemicals*. Technical Report #183, Water Resources Research Center, Purdue University, W. Lafayette, IN. 58 pp.
Part of Steve Preston's thesis work, involving Monte Carlo simulation study of methods for metals and organics. Approaches examined included averaging methods, regression methods, and ratio methods. No one method was always best. Flow the best auxiliary variable. Event sampling coupled with stratified calculations useful. Ratio estimators were not necessarily the most precise, but were always unbiased, generally a desirable feature.

Burn, D.H. 1990. Real-time sampling strategies for estimating nutrient loadings. *J. Water Resources Planning and Management* 116: 727-741.
A simple ratio estimator and the Beale Ratio Estimator performed about the same, and better than averaging estimators. Describes an adjusted stratified sampling program which shifts frequency for wet or dry years to keep total number of samples about equal. This sampling strategy is the paper's primary focus; it might be useful but is probably more complicated than most programs will want to deal with.

Cochran, W.G. 1963. *Sampling Techniques* (2nd edition). Wiley Publications in Statistics. John Wiley and Sons, New York, 413 p.
A classic textbook on sampling, geared primarily to the social sciences, but widely used as a source of approaches to sampling in the environmental sciences as well.

Cohn, T.A. 1995. Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers. Reviews of Geophysics, Supplement, July 1995, pages 1117-1123.
A review of several load estimation procedures, especially the USGS approaches. A good integrated overview of the theory and practical aspects of load estimation. States that the QMLE, MVUE, and smearing corrections for back-transformation bias give about the same results if the data are reasonably appropriately behaved statistically, the model is approximately correct, and the application data set does not exceed the range of the calibration set.

Cohn, T.A., L.L. DeLong, E.J. Gilroy, and R.M. Hirsch, and D.K. Wells. 1989. Estimating Constituent Loads. *Water Resources Research* 25: 937-942.

Presents the log-log regression model which USGS calls the minimum variance unbiased estimator (MVUE), which is a ratio correction to the back-transformed concentration estimate, and assert on theoretical grounds that it is superior to the simple back-transformed value which is known to be biased, and also to the quasi-maximum likelihood estimator (QMLE) which adjusts the back-transformed concentration by multiplying it by half the variance of the regression residuals.

Cohn, T.A., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers. 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. *Water Resources Research* 28: 2353-2363.

Validation study of the MVUE approach, using relationship between log load and log flow and the MVUE bias correction. Other references are needed for complex formulae not given in this paper. While very much more complex than the Beale estimator, the multivariate rating curve approach with MVUE is probably the second method of choice, especially if the existing program does the calculations and can be made more user-friendly.

Crawford, C.G. 1991. Estimation of suspended-sediment rating curves and mean suspended-sediment loads. *Journal of Hydrology* 129: 331-348.

Crawford, C.G. 1996. Estimating mean constituent loads in rivers by the rating-curve and flow-duration, rating-curve methods. Ph.D. dissertation, Indiana University, Bloomington, Indiana, 245 p.

Dickenson, W.T. 1981. Accuracy and precision of suspended sediment loads. Pages 195-202 in *Erosion and Sediment Transport Measurement* (Proceedings of the Florence Symposium, June 1981). IAHS Publication 133.

Effects of five methods and four sampling frequencies on accuracy and precision of suspended solids loads explored. His use of the Beale ratio estimator appears to be confined to post-stratification using a single arbitrary flow cutoff, which is far from ideal. Nine years of data generated by subsampling of three years of daily values (not a very large evaluation set...) at different frequencies. He likes best a "moving rating curve" approach which uses local subsets of the data to establish continuously changing regression relationships between flow and concentration - a local regression approach which, however, is not adequately explained. Precision and accuracy need not go hand-in-hand - annual average conc times annual discharge was precise but not accurate, Beale ratio estimator was accurate but not precise.

Dolan, D. and A.H. El-Shaarawi. 1989. Inferences about point source loadings from upstream/downstream river monitoring data. *Environmental Monitoring and Assessment* 12: 323-357.
Paired sampling upstream and downstream along Niagara River and Detroit River to determine net loading from multiple sources in the interval. Consideration of what to do with non-detects. Of some interest (on a much smaller scale) for 319 projects, but not of direct interest for the loading SOP.

Dolan, D. and A.H. El-Shaarawi. 1991. Applications of mass balance approach with censored data. *Jour. Great Lakes Res.* 17: 220-228.
Mostly about dealing with censored data; not directly relevant for 319 project load estimation, and certainly not relevant for the SOP.

Dolan, D.M., A.K. Yui and R.D. Geist. 1981. Evaluation of river load estimation methods for total phosphorus. *J. Great Lakes Research 7:* 207-214.
Monte Carlo study of methods for load estimation of total phosphorus based on Grand River (MI) data. Sampling at 25 samples/year, in some case with concentration of sampling in high flow periods. Beale ratio estimator was best overall.

Ebise, S. and T. Goda. 1985. Regression models for estimating storm runoff load and its application to Lake Kasumigaura. *Internat. J. Environmental Studies* 25: 73-85.
Total loadings to a lake the focus of the paper. Loadings during dry periods estimated from a single day's measurement. Loadings during storm runoff periods estimated using regression relationships between log storm event load and log storm event discharge, each expressed on a unit area basis, and using data from a number of tributaries to the lake. r values ranged from .62 to .96 with best results for N, worst for P, and intermediate for COD. The regression equations were used to estimate storm loads for all storms, these were summed and added to estimated loads on dry days.

El-Shaarawi, A.H. and D.M. Dolan. 1989. Maximum likelihood estimation of water quality concentrations from censored data. *Canadian Jour. Fish. and Aquatic Sci.* 46: 1033-1039.

Dealing with censored data when the probable concentration matters, e.g. when loads are of interest and the corresponding flow is large. Not very relevant.

El-Shaarawi, A. H., K.W. Kunz and A. Sylvestre. 1986. Estimation of Loading by Numerical Integration. Pages 469-478 in *Statistical Aspects of Water Quality Monitoring*, A.H. El-Shaarawi and R.E. Kwiatkowski, eds., Elsevier, New York, 502 p.

Numeric integration using the trapezoidal rule, including an expression for the mean square error. If observations of concentration are unequally spaced, it is necessary to interpolate or otherwise estimate values for the non-sampled days, or omit some of the flow values, which seems like a bad idea. Of course, that's what we at the WQL do, in effect, by using just the flows at the times of the sample. If values of concentration are estimated to go along with all flow values, it is not clear how this is different from the USGS regression approach or several others. El-Shaarawi's method does not assume a finite population approach. It assumes that the errors in the individual load estimates have a constant variance and an autocorrelation structure which is scaled to the variance. Not clear whether the errors are assumed to be normally distributed. Not clear how one can evaluate the nature of the error term, its variance or its autocorrelation. Interesting, but as usual I do not understand how to apply it completely.

Ferguson, R.I. 1986a. River loads underestimated by rating curves. *Water Resources Research 22:* 74-76.

"Statistical considerations show that the sediment, solute, or pollutant load of a river is likely to be underestimated by methods in which unmeasured concentrations are estimated from discharge using a least-squares regression for the logarithm of concentration. The degree of underestimation increases with the degree of scatter and can reach 50%. A simple correction factor is proposed and tested successfully on simulated and real data sets." -Author's abstract.

Ferguson, R.I. 1986b. Reply (to Koch and Smillie).*Water Resources Research 22:* 2123-2124.

Flemal, R.C. 1978. Formulas and methods for calculating concentration and loads at any point on a stream. Illinois Water Information System Group, Report of Investigations 4, 9 p.

Basically deals with routing load estimates downstream. Does not make any statements about how to obtain the load estimate in the first place, beyond that load equals concentration times flow.

Fraser, A.S. and K.E. Wilson. 1981. Loading estimates to Lake Erie, 1967-1976. CCIW/NWRI Scientific Series 120, 23 p.
Uses Beale estimator to derive loads for several nutrient parameters. Shows substantial increasing trends in NO23 and TKN, decreasing trends in TP and SRP over the period 1967 through 1976

Gale, J.A., S.W. Coffey, D.E. Line, J. Spooner, D.L. Osmond, and J.A. Arnold. 1992. Summary Report: Evaluation of the Expermental Rural Clean Water Program. North Carolina Cooperative Extension Service, Raleigh, NC. 38 p.
Review of the RCWP program which preceeded the NPS demonstration and monitoring projects. Among many other findings, strongly supports the design and implementation of monitoring adequate to document program success, and finds that most previous programs have failed in this respect.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York, 320 pp.
A very useful but sometimes difficult text on the application of statistics to environmental problems, including but not limited to aquatic systems and load estimation.

Greenberg, A.E., L.S. Clesceri, and A.D. Eaton, eds. 1992. *Standard Methods for the Examination of Water and Wastewater, 18th Edition*; American Public Health Association: Washington, DC
One of several standard compendia of widely accepted techniques for chemical analyses of water samples.

Grobler, D.C., C.A. Bruwer, P.J. Kemp, and G.C. Hall. 1982. A comparison of chemical load estimation algorithms using data obtained by sampling four South African rivers at varying frequencies. *Water SA* 8: 121-129.
Compares several approaches at different sampling frequencies, and with wet/dry season stratification. None of the approaches is very sophisticated. Substantial differences exist as a function of sampling frequency, method, and parameter(of course). Methods all involve summing loads for periods of time defined by the chemical sampling to get the annual load. The interval loads are variously defined by the interval discharge times the average of beginning and ending concentration, the interval discharge times the ending concentration, the average concentration times the annual discharge (apparently - the notation is kinda messy), and this last stratified into wet and dry seasons. Several methods, especially the last two, were prone to large positive biases

with conservative substances and large negative biases with particulate substances, presumably due to correlations between flow and concentration. Good list of early literature.

Heidke, T.M., T.C. Young and J.V. DePinto. 1987. Assessment of Alternatives for Calculating Annual Total Phosphorus Tributary Loadings. Pages 367-379 in *Symposium on Monitoring, Modeling, and Mediating Water Quality*. Am. Water Resources Assoc.
Monte Carlo investigation of several versions of regression and ratio estimators for estimating total phosphorus load in the Saginaw River. Examines role of sampling frequency, stratification, allocation of samples among strata. Concludes no one method is always best, but the Beale stratified ratio estimator is the most consistently accurate.

Hellmann, H. 1986. Zum Problem der Frachtberechnung in Fließgewässern. *Zeit. Wasser-Abwasser-Forschung* 19: 133-139.
A general discussion of the relationships of concentration and load to flow, focused on large rivers. Bemoans the inhomogeneity and temporal variability of concentrations, and despairs of ever being to characterize the load exactly. No great insights as far as I can tell; nothing useful for SOP except for reminding us about the importance of proper siting of the station.

Helsel, D.R. and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. Studies in Environmental Science 49. Elsevier, Amsterdam, 522 pp.
An excellent specialized text presenting approaches to analyzing water quality (and water quantity) data. The authors emphasize a pragmatic approach and methods which work well with the ill-tempered data which characterizes environmental research. A number of novel methods are presented which are not frequently encountered elsewhere.

Khare, B.B. and S.R. Srivastava. 1981. A generalized regression ratio estimator for the population mean using two auxiliary variables. *Aligarh Jour. Statistics* 1: 43-51.

Klaine, S.J. and A. Roman-Mas. 1993. *Optimization of Sampling Strategy to Assess Agricultural Nonpoint Source Pollutant Loads*. Progress report to Water Environment Research Foundation, 13 pp.
Focus is on concentration distributions during runoff events, not directly on loads. They conclude that "the temporal characterization of concentrations for selected constituents is very sensitive to sampling intensity,

particularly during storm flow" and that "a sampling interval equal to 0.05 of the duration of storm flow is adequate to characterize concentration for the selected constituents during storm flow". Parameters were NO23, organic N, and suspended sediment. No attempt made to do flow-proportional sampling. See also Roman-Mas *et al*. (1994).

Koch, R.W. and G.M. Smillie. 1986a. Bias in hydrologic prediction using log-transformed regression models. *Water Resources Bulletin* 22: 717-723.

Koch, R.W. and G.M. Smillie. 1986b. Comment on "River loads underestimated by rating curves" by R.I. Ferguson. *Water Resources Research 22:* 2121-2122.

Line, D.E. and others. 1993. Nonpoint sources. *Water Environment Research* 65: 558-570.
Annual review paper on non-point source publications, a source of access to the primary literature.

Line, D.E. and others. 1994. Nonpoint sources. *Water Environment Research* 66: 585-601.
Annual review paper on non-point source publications, a source of access to the primary literature.

Line, D.E. and others. 1995. Nonpoint sources. *Water Environment Research* 67: 685-700.
Annual review paper on non-point source publications, a source of access to the primary literature.

Loftis, J.C. and R.C. Ward. 1980. Water quality monitoring - some practical sampling frequency considerations. *Environmental Management 4:* 521-526.
Network design and sampling frequency determination in the face of seasonality and autocorrelation. At more than 30 samples/year, autocorrelation dominates; from 10-30 samples/year, autocorrelation and seasonality cancel out; less than 10 samples per year, seasonality dominates.

Loftis, J. and R.C. Ward. 1980. Sampling frequency selection for regulatory water quality monitoring. *Water Resources Bulletin 16*: 501-507.
Allocating sampling effort among stations to achieve uniform precision across the network, including considerations of seasonality and autocorrelation.

Loftis, J. and R.C. Ward. 1978. Statistical tradeoffs in monitoring network design. Pages 36-48 in L.G. Everett and K.D. Schmidt, eds., *Establishment of Water Quality Monitoring Programs*, American Water Resources Association, Minneapolis, MN.
Allocating sampling effort among stations in a network. Two methods, a simple one for non-statisticians and a more complex one including autocorrelation considerations for the more statistically astute.

Marsalek, J. 1991. Pollutant loads in urban stormwater: Review of methods for planning-level estimates. *Water Resources Bulletin* 27: 283-291.
Evaluation of urban nonpoint sources is difficult. A review of methods for planning-level estimates of pollutant loads in urban stormwater was conducted, focusing on transfer of characteristic runoff quality data to unmonitored sites, runoff monitoring, and simulation models. Load estimation by transfer of runoff quality data is the least expensive, but the accuracy of estimates is unknown. Runoff monitoring methods provide best estimates of existing loads, but cannot be used to predict load changes resulting from runoff controls, or other changes of the urban system. Simulation models require extensive calibration for reliable application. Models with optional formulations of pollutant buildup, washoff, and transport can be better calibrated and the selection of options should be based on a statistical analysis of calibration data. Calibrated simulation models can be used for evaluation of control alternatives. Not really relevant for this project.

McBride, G.B. and D.G. Smith. 1997. Sampling and analytical tolerance requirements for detecting trends in water quality. J. Am. Wat. Res. Assoc. 33: 367-373.
A useful and interesting paper, but not directly relevant to this guidance, because the goal is trend detection, not load estimation. Provides calculations which take into account Type I and II errors as well as analytical uncertainty. May contain some useful insights and food for thought if the goal of the load estimation program is ultimately to assess trends.

Meyer, D.H. and J. Harris. 1991. Prediction of phosphorus load from non-point sources to South African rivers. *Water SA* 17: 211-216.
Log-linear model between monthly loads and monthly flows not adequate, autocorrelation important and must be incorporated. A quadratic flow term is needed for some rivers to prevent underestimation of loads. A correction factor is needed to prevent bias in back-transformation; they used a modification of the standard s2 correction. This is based on the estimated mean and variance of the error term in the time series transfer model, determined separately for each river.

Olkin, Ingram. 1958. Multivariate ratio estimation for finite populations. *Biometrika* 45: 154-165.

Pacheco-Ceballos, R. 1989. Transport of sediments: Analytical solution. *Jour. Hydraulic Research* 27: 501-518.
Total load of sediment calculated from engineering principles based on conservation of energy and on power formulas. Seems to give pretty good agreement with observed loads. Unclear how practical it is for general application, however, given the kinds of information you have to have.

Pinter, J. and L. Somlyody. 1986. Optimization of regional water-quality monitoring strategies. In *Integrated Design of Hydrological Networks (Proceedings of the Budapest Symposium, July 1986)*. IAHS Publ. no. 158, 259-268.
Linear programming approach to network design, including variable cost scenarios. Interesting, complex, but irrelevant for this project.

Ponce, S.L. 1980. *Water Quality Monitoring Programs.* Watershed Systems Development Group, USDA Forest Service, Fort Collins, Colorado. 66 pp.
A Forest Service manual for planning sampling programs. Includes an iterative procedure for determining the number of samples needed to achieve a specified error with a specified probability, hence its inclusion here.

Preston, S.D., V.J. Bierman Jr., and S.E. Silliman. 1992. Impact of flow variability on error in estimation of tributary mass loads. *J. Environ. Engineering* 118: 402-419.
Monte Carlo evaluation of various estimators: sum of monthly average conc times monthly discharge, Beale Ratio estimator, and Cohn's MVUE regression approach. Stratified Beale Ratio Estimator the only one to consistently perform reliably with good precision and low bias.

Preston, S.D., V.J. Bierman, and S.E. Silliman. 1989. Evaluation of methods for the estimation of tributary mass loads. *Water Resources Research* 25: 1379-1389.
Tributary loading estimation methods were evaluated by conducting retrospective studies with comprehensive sets of field data from the Grand and Saginaw Rivers, Michigan, for flow rates, nutrients, heavy metals, and PCBs. Three broad classes of loading estimation methods were investigated: simple averaging methods, ratio estimation methods, and regression methods. Estimators were evaluated using Monte Carlo sampling studies in which random subsamples of complete loading records were used to estimate annual loadings. These

estimates were then compared to 'true' loadings determined by calculations using the entire record. No group of estimators was superior for all test cases considered. However, individual estimation approaches within each group often provided low error estimates. Results were inconsistent among test cases and these inconsistencies appeared to be related to specific test case characteristics such as the strength and form of the flow-concentration relationship and the nature of the annual hydrograph. Ratio estimators appeared to be more robust to sources of bias than other estimation approaches.

Ramachandran, V. and S.S. Pillai. 1975?. Multivariate unbiased ratio-type estimation in finite population. *Jour. Indian Soc. Agric. Stat.*
Develops an unbiased ratio estimator using two or more auxiliary variables.

Rast, W. and G.F. Lee. 1983. Nutrient Loading Estimates for Lakes. *J. Environ. Engineering* 109: 502-517.
Nutrient loads (N and P) estimated from land use pattern and export coefficients for the major land use types. Vollenweider-type phosphorus models also worked well. A study of 38 U.S. lakes. Not useful for 319 projects because we seek to change the export coefficients!

Reinelt, L.E. and A. Grimvall. 1992. Estimation of nonpoint source loadings with data obtained from limited sampling programs. *Environmental Monitoring and Assessment* 21: 173-192.
They compare numeric integration with "stepwise constant concentrations and linearly interpolated concentrations" with regression relationships between conc and flow (stratified) and between log load and log flow. In all cases, they use the record to estimate concs (or loads) for each day, apparently. They re-discover, in a new form, the bias which results from ignoring correlations between flow and concentration. While they compare the four methods, they have no firm basis on which to evaluate which one performs better because they do not know the true values of the loads. Arguing from details of the hydrograph, they infer that numeric integration methods overestimated the loads, which is probably true since they sampled the snow melt and attributed its concentration to a longer period of time than is reasonable. They argue that regression of conc vs flow is not very reliable for various reasons. They then conclude that the only method they did not criticize must be the best - a process of elimination which is not very satisfactory philosophically, though the conclusion is probably correct.... They also had very high correlations between log load and log flow - 0.95 to 0.98. They did not deal with the inverse transformation bias question at all. Although it sounds promising, this is not a very sophisticated paper.

Reinelt, L.E., R.R. Horner, and R. Castensson. 1992. Non-point source water pollution management: improving decision-making information through water quality monitoring. *J. Environ. Management* 34: 15-30.

Well planned programs provide better information at a lower cost than unplanned or poorly planned ones. (Duh.) Evaluation of a monitoring design for load estimation on a Swedish river. Loads determined using a log load vs. log flow regression relationship. Reliability measured by bootstrapping the regression coefficients to provide a variance estimate on the loads. Now that's interesting! However, the paper is very unclear about how they did this, and I am not convinced that what they did was valid. Most of their improvements in overall efficiency came from dropping superfluous parameters and targeting sampling to catch snowmelt and other important events during their 13 samples per year.

Richards, R.P. 1990. Measures of flow variability and a new flow-based classification of Great Lakes tributaries. *J. Great Lakes Research* 16: 53-70.

Develops and evaluates several different measures to characterize the variability of flow, and applies them to 120 Great Lakes tributaries.

Richards, R.P. 1989a. Measures of flow variability for Great Lakes tributaries. *Environmental Monitoring and Assessment* 12: 361-377.

Flow variability can be used as a substitute for flux variability when the latter is unknown, for the purposes of estimating sampling needs. Formulas are given for suspended solids, total phosphorus, and chloride.

Richards, R.P. 1989b. *Evaluation of some approaches to estimating non-point pollutant loads for unmonitored areas*. Water Resources Bulletin 25: 891-903.

Evaluates various approaches to estimating loads for portions of basins which are not monitored, such as areas downstream of the monitoring station, and by implication for entire basins which are not monitored. Adjustments by land area ratio, by land area ratio and average USLE C-factor ratio, by discharge ratio, and by applying various regression relationships from an adjacent basin or from a previous history in the same basin. Discharge ratio adjustment was most successful overall, but no method produced consistently reliable results.

Richards, R.P., D.B. Baker, J.W. Kramer, and D.E. Ewing. 1996. Annual loads of herbicides in Lake Erie tributaries in Ohio and Michigan. *Journal of Great Lakes Research* 22: 414-428.

Reports annual loads for a number of Lake Erie tributaries for a period as long as 11 years, and compares them with load estimates from other studies, particularly of Goolsby and colleagues at USGS. Finds that loads are highly variable from year to year as function of weather patterns, but are higher in basins with larger percentages of the land in agriculture, and in basins with clay-rich soils and tiles than in basins with sandier soils with better infiltration. Atrazine and metolachlor loads are 2-5 g/ha/yr, alachlor 1-2 g/ha/yr, and cyanazine and metribuzin less than 1.5 g/ha/yr.

Richards, R.P. and D.B. Baker. 1993. Pesticide concentration patterns in agricultural drainage networks in the Lake Erie basin. *Environ. Toxicology and Chemistry* 12: 13-26.

General patterns of pesticide concentrations in Lake Erie tributaries draining predominantly agricultural watersheds at the edge of the corn belt. Probably broadly applicable to other parts of the country where most of the pesticide runoff is expected to be from row crop agriculture, though the periods of peak runoff in the year will vary with the cropping patterns.

Richards, R.P. and J. Holloway. 1987. Monte Carlo Studies of Sampling Strategies for Estimating Tributary Loads. *Water Resources Research 23*:1939-1948.

Monte Carlo simulations were used to evaluate the effects of sampling frequency and pattern and load calculation method on precision and accuracy of loads of suspended sediment, total phosphorus, soluble reactive phosphorus, nitrate plus nitrite, and specific conductance (proxy for total dissolved solids). The Beale ratio estimator combined with flow-stratified sampling gave the best results among the methods evaluated. Precision showed an approximate square root relationship to sampling frequency for a given method. Monthly sampling programs combined with numeric integration performed very poorly. Stratified programs which were forced to stop when a fixed number of samples had been taken (fixed cost) performed poorly. Post-stratification did not improve precision.

Roman-Mas. A., R.W. Stogner, U.H. Doyle and S.J. Klaine. 1994. Assessment of agricultural non-point source pollution and Best Management Practices for the Beaver Creek watershed, west Tennessee. In G.L. Pederson, ed. *Proceedings*, American Water Resources Association National Symposium on Water Quality, Nashville, TN, April 17-20, 1994. American Water Resources Association, Bethesda, MD.

Not seen, but makes same points as Klaine and Roman-Mas (1993), according to S.J. Klaine, and easier to obtain.

Russ, H.-J. and M. Uhl. 1990. Comparison of pollutant load calculation methods based on measured data - state of the investigation. *Water Science Technology* 22: 95-102.

Urban storm water problems and models for estimation. What is being evaluated is the quality of model results when applied by their developers. Nothing of use here, except the comment that models tend to provoke complacency and complex models are more prone to error and the errors are harder to discover.

Sahoo, L.N. 1984. A note on estimation of the population mean using two auxiliary variables. *Aligarth Journal of Statistics* 3&4: 63-66.
Ratio estimator using two aux. variables works better than using one alone. Difficult to apply; too complex for EPA SOP. Proposes a criterion for the suitability of a ratio estimator, apparently from some other source:

$$g01(C0/C1) > 0.5$$

where g01 is the correlation coefficient between the dependent and independent variables, C0 is the coefficient of variation of the dependent variable, and C1 is the coefficient of variation of the independent variable.

Cochran gives the same criterion as the condition under which the variance of a population total based on a ratio estimate will be smaller than the variance of the population total derived by multiplying the sample mean by the population size (i.e. without ratio adjustment).

Sanders, T. G., R.C. Ward, J.C. Loftis, T.D. Steele, D.D. Adrian, and V. Yevjevich. 1983. *Design of Networks for Monitoring of Water Quality.* Water Resources Publications, Littleton, Colorado, 328 pages.
An excellent reference on planning water quality monitoring programs. It has somewhat of a tendency toward a point-source perspective, however, and does not deal directly with load calculation methods.

Shih, G. 1994. Accuracy of nutrient runoff load calculations using time-composite sampling. *Trans. Am. Soc. Agric. Eng.* 37: 417.
Time compositing leads to under-estimation of loads. Estimates as reliable as those from flow-composite sampling can be obtained using 8 time-composite samples per runoff event.

Shih, G., W. Abtew, and J. Obeysekera. 1994. Accuracy of nutrient runoff load calculations using time-composite sampling. *Trans. Am. Soc. Agric. Eng.* 37: 419-429.
Evaluates the bias introduced by using time-composite sampling to calculate loads, rather than flow-composite sampling, using both analytical and simulation approaches. Context is pumping from storage ponds on farms in Florida, but results in general should apply to other runoff processes. Provides a correction approach for the bias introduced by time-composite sampling, and concludes that 8 time-composite samples per hydrograph plus

correction will yield a load with accuracy comparable to a single flow-composite sample. Not clear that this is simpler or cheaper than just doing the flow-composite sample (or more discrete samples) in the first place.

Shukla, G.K. 1966. An alternative multivariate ratio estimate for finite population. *Calcutta Statistical Assoc. Bull.* 15: 127-134.
More multivariate ratio estimator stuff - too complex!

Spooner, J. and others. 1991. Nonpoint sources. *Research Journal WPCF* 63: 527-535.
Annual review paper on non-point source publications, a source of access to the primary literature.

Spooner, J. and others. 1992. Nonpoint sources. *Water Environment Research* 64: 503-513.
Annual review paper on non-point source publications, a source of access to the primary literature.

Tin, M. 1965. Comparison of some ratio estimators. *J. Am. Stat. Assoc. 60:* 294-307.
"...four ratio estimators designated as simple, Quenouille's, Beale's, and modified ratio estimators are compared with respect to bias, efficiency, approach to normality and computational convenience. They are shown to be asymptotically minimum variance bound estimators. Some additional ratio estimators are discussed briefly and compared with these. Quenouille's, Beale's and modified ratio estimators are found to be more attractive than the alternatives compared." -Author's abstract.

Thomas, R.B. 1985. Estimating Total Suspended Sediment Yield with Probability Sampling. *Water Resources Research 21:* 1381-1388.
"The 'Selection at List Time' (SALT) scheme controls sampling of concentration for estimating total suspended sediment yield. The probability of taking a sample is proportional to its estimated contribution to total suspended sediment discharge. This procedure gives unbiased estimates of total suspended sediment yield and the variance of the estimate while automatically emphasizing sampling at higher flows. When applied to real data with know yield, the SALT method underestimated total suspended sediment yield by less than 1%, whereas estimates by the flow duration sediment rating curve method averaged about 51% underestimation. Implementing the SALT scheme requires obtaining samples with a pumping sampler, stage sensing device, and small battery-powered computer." -Author's abstract.

Thompson, M.E. and K. Bischoping. 1986. On the Estimation of Monthly Mean Phosphorus Loadings. Pages 460-468 in *Statistical Aspects of Water Quality Monitoring,* A.H. El-Shaarawi and R.E. Kwiatkowski, eds. Elsevier, Amsterdam, 502 pages.

Uri, N.D. and B. Hyberg. 1990. Stream sediment loading and rainfall – a look at the issue. *Water, Air, and Soil Pollution* 51: 95-104.
"This paper investigates the issue of the nature of the relationship between stream loading and storm intensity and whether stream sediment loading can best be explained by storm intensity (rainfall) or whether a more general average rainfall measure is superior. Based on data covering the years 1947-1985 for the Iowa River watershed north of Iowa City, a nonlinear relationship between stream sediment loading and rainfall is indicated. Moreover, average monthly rainfall better explains sediment loading than do other measures of storm (rainfall) intensity. Finally, when the structural stability of the estimated relationships are [sic] explored, the indications are that the relationships are stable over the sample period." -Author's abstract.

Work is related to USLE and other modeling approaches, and of tangential interest for this document, since the basic data evaluated are annual sediment loads

Walling, D.E. 1978. Reliability considerations in the evaluation and analysis of river loads. *Zeitschrift für Geomorphologie N.F. 29:* 29-42.
"Although measurements of river load have a relatively long history, little attention has been paid to their reliability. Many problems exists both as regards interpretation and accuracy of the resultant data. Accuracy considerations can be subdivided into those of the comparability of various load estimation procedures and their absolute accuracy. Results from several rivers in Devon, England have been used to study these further. In general, solute load data are more reliable than those for suspended sediment load. Absolute errors associated with suspended sediment could be as high as +60% for annual loads and between +400% and -80% for monthly loads and the corresponding values for solution loads could be up to ±25% and ±60%, respectively." -Author's abstract. Very similar to Walling and Webb (1981).

Walling, D.E. and B.W. Webb. 1981. The reliability of suspended sediment load data. In *Erosion and Sediment Transport Measurement (Proceedings of the Florence Symposium, June 1981)* IAHS Publ. No. 133.

Detailed sampling required to define chemograph. Sediment samples less frequent than hydrological measurements. Either numeric integration or regression estimators can involve large errors when sampling is inadequate.

Continuous turbidity records for a 7 year period were converted to hourly sediment records and subsampled to evaluate different estimators. Numeric integration approach produced unbiased but seriously unreproducible results, as did a flow ratio-adjusted version of the same. (Average c * average q) had a serious low bias (80%) but was much more reproducible. Of course, the magnitude of the bias would normally not be known, but if changes (trends) are more important than levels, the bias may be less important than the high precision. Sum of conc*average flow for interval is BIASED LOW by amounts which vary with the sampling interval.

Various rating curve approaches all biased low, mostly pretty precise. Not clear whether a retransformation bias correction was used. Better results occurred when the sampling used to define the rating curve was skewed toward high flows and when separate rating curves were calculated for winter and summer, rising and falling portions of hydrographs.

The load interval method (cf. Yaksich and Verhoff, 1983) is unbiased when daily flows are used, and can have relatively high precision. Perhaps it has not been adequately tested, though it seems to be very similar to the Stratified Beale Ratio Estimator. An excellent paper, though Beale ratio estimator is not considered, nor is stratification used except with rating curves. Also, the study is confined to one set of 7 years of data, so does not reflect effects of annual variability.

Walker, W.W. 1996. Simplified procedures for eutrophication assessment and prediction: User Manual. Instruction Report W-96-2, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
User's manual for FLUX, PROFILE, and BATHTUB, programs for reservoir management. FLUX calculates loads by a number approaches, some rather naive for general use. It runs on a PC.

Whitt, D.M. 1977. Quality Control Handbook for Pilot Watershed Studies. International Joint Commission, Windsor.
Admittedly, an obscure reference.... Sorry! But one source for the Beale Ratio Estimator formulae.

Wu, J. and R.C. Ahlert. 1978. Assessment of methods for computing storm runoff loads. *Water Resources Bulletin* 14: 429-439.

Loadographs (instantaneous flux as function of time) and hydrographs have similar shapes; hydrology dominates the load pattern. Methods proposed are often of planning level of detail/precision, rather urban runoff oriented, and not very useful for this project. More sophisticated (and complex) methods allow the calibration of a storm runoff model to predict flux from flow, but it is unclear if the calibration is valid from one storm to the next.

Yaksich, S.M. and F.H. Verhoff. 1983. Sampling strategy for river pollutant transport. *J. Environ. Eng.* 109: 219-231.
Compares several methods for estimating loads, given frequent flow measurements and infrequent chemical measurements. Describes the flow-interval technique, a variety of ratio estimation.
The other techniques compared are not stratified and some are naive, so not surprisingly they do not perform as well.

Young, G.K., T.R. Bondelid, and D.N. Athayde. 1979. Urban runoff pollution method. *J. Water Resources Planning and Management Division (ASCE)* 105: 353-369.
Nonpoint source loads for urban areas. A quasi-deterministic model approach to estimate runoff for a model storm, including sewers, CSOs etc. Clearly not relevant to the interests of this project!

Young, T.C., J.V. DePinto, and T.M. Heidke. 1988. Factors affecting the efficiency of some estimators of fluvial total phosphorus load. *Water Resources Research* 24: 1535-1540.
Monte Carlo methods used to evaluate five load calculation approaches, including regression, ratio, and robust methods, which use additional flow data to supplement less frequent concentration data. Stratified Beale Ratio Estimator performed best, but was only stratified method. Log-log regression next best. Better results were obtained when high flow periods were sampled more heavily. Post-stratification often improved results.

Young, T.C. and D.M. Dolan. 1995. Assessment and control of loading uncertainty for managing eutrophication and toxic chemical fate in lakes. In: V.P. Singh, ed. *Environmental Hydrology,* Kluwer Academic Publishers, Netherlands.
Interesting discussion of factors which contribute uncertainty to load estimates. Good background but does not directly present any load estimation methodology.

Zeman, L.J. and H.O. Slaymaker. 1985. Estimation of phosphorus flux in a regulated channel. *Water Research* 19: 757-762.

Relationship between phosphorus conc and discharge not significant, because of the regulated nature of the river, hence regression approaches relying on conc/flow relationships could not be used (or so they said). Loads were calculated using a "partial load method" and the flow interval method, and found to be similar. Initial study characterized cross-sectional variability, diurnal variations in concentration, etc. Monthly manual cross-sectional sampling was used to establish error limits around point samples collected using autosamplers over time. Sequential sampling done on two consecutive randomly chosen days each week, four evenly spaced samples per day starting at 3:00.

The partial load method involves calculating a load and c.i. for each of several cross-channel slices, then summing to get total load. Details of the method are presented elsewhere. The "flow interval method" is just to sum the hourly flows symmetrically displaced around the sampling time, and multiply this total discharge times the concentration. The daily load is the sum of the four quarter-daily loads. I.e. just a variety of numeric integration. While the estimation of precision is prominently mentioned, little discussion is given to it other than the results of cross-sectional sampling giving a set of load values at a given moment in time, shown as an error bar.

APPENDIX A: CONVERSION FACTORS

To convert from concentration and flow to daily load in metric tons, multiply the product of concentration and flow by the appropriate constant:

Concentration Units	Flow Units	Constant
g/L	ft^3/sec	2.447
mg/L	ft^3/sec	0.002447
µg/L	ft^3/sec	0.000002447
g/L	m^3/sec	86.4
mg/L	m^3/sec	0.0864
µg/L	m^3/sec	0.000864

For example, if the flow is 375 ft^3/sec and the concentration is 1.32 mg/L, the daily load is

$$L = 375 * 1.32 * 0.002447 = 1.21 \text{ metric tons}$$

Other useful conversion factors are shown on the next page.

Conversion factors for weights, distances, areas, and volumes between metric and English units.

Metric Units		English Units
1 gram	=	0.035 ounces
1 kilogram	=	2.205 pounds
1 metric ton	=	1.1 short tons (U.S.)
1 centimeter	=	0.3937 inches
1 meter	=	39.37 inches
1 meter	=	3.283 feet
1 kilometer	=	0.6214 miles
1 square meter	=	10.76 square feet
1 hectare	=	107,600 square feet
1 hectare	=	2.47 acres
1 hectare	=	0.0039 square miles
1 square kilometer	=	0.3861 square miles
1 cubic centimeter	=	0.061 cubic inches
1 cubic meter	=	35.31 cubic feet

English Units		Metric Units
1 ounce	=	28.57 grams
1 pound	=	0.454 kilograms
1 short ton	=	0.9072 metric tons
1 inch	=	2.540 centimeters
1 inch	=	0.0254 meters
1 foot	=	0.3048 meters
1 mile	=	1.609 kilometers
1 square foot	=	0.0929 square meters
1 square foot	=	9.29×10^{-6} hectares
1 acre	=	0.405 hectares
1 square mile	=	259.0 hectares
1 square mile	=	2.590 sq. kilometers
1 cubic inch	=	16.39 cubic centimeters
1 cubic foot	=	0.0283 cubic meters

APPENDIX B: GLOSSARY

Accuracy: a set of measurements is accurate when its average value is close to the true value of the thing being measured.

Aliquot: a portion of a sample.

Bias: a measure of accuracy, inverse in the sense that measurements that are inaccurate are biased

Discharge: the volume of water which passes a given point in a river or stream in a given period of time, a total quantity.

Efficient: in measurement, having relatively high accuracy and precision for a given amount of effort.

Flow: the rate at which water passes a given point in a river or stream at a given moment. The integral of flow over time is the discharge.

Flow-proportional sample: a composite sample composed of aliquots taken in proportion to the flow rate. This can either be done by taking aliquots at fixed intervals of time, but varying the size of the aliquot in proportion to the discharge since the last sample, or by taking an aliquot of a fixed size every time a specified discharge has passed since the previous aliquot. "Flow-proportional sample" is a misnomer; it really should be called a "discharge-proportional sample".

Flow-weighted mean concentration: the load divided by the discharge; the loading rate. Computational approaches are given in the text.

Flux: the rate at which a pollutant load passes a given point in a river or stream at a given moment. The integral of flux over time is the load. The flux is equal to the concentration times the flow at the time of the sample.

FWMC: flow-weighted mean concentration.

Load: the mass of a chemical substance which passes a given point in a river or stream in a given period of time, a total quantity.

Mean daily flow: the average of the flow measurements made on a particular day. At USGS gaging stations, stages are measured and converted to flows usually at hourly or quarter-hourly intervals, depending on the size of the river.

Mean daily load: the average of the daily loads for a number of days. Many loading programs estimate a daily load for each day of the year or the month. The mean daily load is the average of these. The total load for the month or the year is computed by multiplying the mean daily load by the number of days in the month or year.

Precision: a set of measurements is precise when each measurement is about equally close to the true value of the thing being measured. If the measurements are of the same thing, the set of measurements is precise if the measurements are all similar to each other. A set of measurements can be badly biased but still be considered precise.

Random sampling: Any scheme for sampling a population in such a way that every member of the population has an equal chance of being included in the sample. In sampling over time, a useful technique for avoiding inadvertant bias due to unrecognized periodicities in the system being sampled.

Rating curve: In the broad sense, any curve which allows one parameter to be estimated or calculated from another. In hydrology, a rating curve expresses the empirical relationship between stage, or height of the water, and flow. In load estimation, a rating curve is a relationship used to calculate daily loads or concentrations from flow and other independent variables, usually using some form of regression.

Residuals: In regression analysis, the difference between observed values and those estimated from the regression model for the same value of the independent variable(s).

Retransformation bias: In load estimation by regression using log-transformed flow and concentration, a bias which occurs when log-concentrations or log-loads estimated from the regression model are back-transformed by exponentiation to arrive at the corresponding concentrations or loads.

Stratified sampling: A sampling program which uses a different sampling frequency and/or pattern at different times of year and/or at different flows. Stratified sampling often leads to more efficient estimates of loads, particularly with ratio estimators.

Systematic sampling: sampling which provides a regular subsample of the population - every third house, or one sample per day, for example. In water quality, systematic sampling can be more efficient than random sampling, but can produce biased results, particularly if periodic fluctuations are an important part of the behavior of the parameter being sampled.

Time-proportional sample: A composite sample composed of aliquots taken without regard to flow. Equal volume aliquots are taken at equal intervals.

Total load: The load for an entire period of interest, usually a month or a year. The sum of the daily loads in the period, or the product of the average daily load and the number of days. If the time interval represented by the unit loads is not constant, the sum of the unit loads over the period of interest.

Unit load: The load for a single measurement interval, usually a day. The product of a single concentration and a single flow and the interval of time they represent.

APPENDIX C: CUYAHOGA RIVER DAILY LOADS

Table C1: Estimated concentrations and daily loads for suspended sediment in the Cuyahoga River, calendar year 1992, as calculated by a simple ratio estimator. See main text for explanation.

Date	Q	ln(Q)	ln(c^)	ln(c^)+$s^2/2$	c^	daily load
1/1/92	279	5.631	2.582	3.010	20.291	13.853
1/2/92	261	5.565	2.523	2.951	19.133	12.219
1/3/92	266	5.583	2.540	2.968	19.455	12.664
1/4/92	268	5.591	2.546	2.975	19.584	12.843
1/5/92	265	5.580	2.536	2.965	19.391	12.574
1/6/92	244	5.497	2.464	2.892	18.031	10.766
1/7/92	240	5.481	2.449	2.878	17.770	10.436
1/8/92	237	5.468	2.438	2.866	17.574	10.192
1/9/92	253	5.533	2.495	2.924	18.615	11.525
1/10/92	311	5.740	2.677	3.106	22.327	16.991
1/11/92	310	5.737	2.674	3.103	22.264	16.889
1/12/92	251	5.525	2.488	2.917	18.486	11.354
1/13/92	266	5.583	2.540	2.968	19.455	12.664
1/14/92	1840	7.518	4.243	4.672	106.897	481.301
1/15/92	839	6.732	3.552	3.980	53.520	109.879
1/16/92	503	6.221	3.101	3.529	34.102	41.974
1/17/92	545	6.301	3.171	3.600	36.598	48.808
1/18/92	519	6.252	3.128	3.557	35.056	44.521
1/19/92	453	6.116	3.009	3.437	31.097	34.471
1/20/92	478	6.170	3.056	3.484	32.604	38.136
1/21/92	511	6.236	3.115	3.543	34.579	43.239
1/22/92	450	6.109	3.003	3.431	30.916	34.043
1/23/92	1210	7.098	3.874	4.303	73.894	218.789
1/24/92	1960	7.581	4.299	4.728	113.015	542.033
1/25/92	942	6.848	3.654	4.082	59.268	136.617

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
1/26/92	700	6.551	3.392	3.820	45.627	78.154
1/27/92	596	6.390	3.250	3.679	39.599	57.752
1/28/92	595	6.389	3.249	3.677	39.541	57.570
1/29/92	496	6.207	3.088	3.517	33.684	40.882
1/30/92	483	6.180	3.065	3.494	32.905	38.890
1/31/92	588	6.377	3.238	3.667	39.131	56.302
2/1/92	578	6.360	3.223	3.652	38.544	54.515
2/2/92	509	6.232	3.111	3.540	34.460	42.921
2/3/92	483	6.180	3.065	3.494	32.905	38.890
2/4/92	480	6.174	3.060	3.488	32.725	38.437
2/5/92	488	6.190	3.074	3.503	33.205	39.651
2/6/92	433	6.071	2.969	3.397	29.885	31.664
2/7/92	484	6.182	3.067	3.495	32.965	39.042
2/8/92	448	6.105	2.999	3.427	30.795	33.759
2/9/92	381	5.943	2.856	3.285	26.699	24.892
2/10/92	368	5.908	2.826	3.254	25.895	23.319
2/11/92	357	5.878	2.799	3.227	25.212	22.025
2/12/92	265	5.580	2.536	2.965	19.391	12.574
2/13/92	270	5.598	2.553	2.981	19.713	13.024
2/14/92	312	5.743	2.680	3.109	22.390	17.094
2/15/92	1270	7.147	3.917	4.345	77.112	239.640
2/16/92	3860	8.258	4.896	5.325	205.315	1939.291
2/17/92	2430	7.796	4.488	4.917	136.575	812.103
2/18/92	1350	7.208	3.971	4.399	81.375	268.820
2/19/92	2960	7.993	4.662	5.091	162.500	1177.009
2/20/92	2910	7.976	4.647	5.076	160.080	1139.891
2/21/92	2130	7.664	4.372	4.801	121.607	633.829
2/22/92	1230	7.115	3.889	4.317	74.968	225.641

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
2/23/92	1240	7.123	3.896	4.324	75.505	229.104
2/24/92	1110	7.012	3.798	4.227	68.486	186.021
2/25/92	948	6.854	3.659	4.088	59.600	138.258
2/26/92	813	6.701	3.524	3.952	52.056	103.562
2/27/92	725	6.586	3.423	3.851	47.059	83.487
2/28/92	841	6.735	3.554	3.982	53.633	110.372
2/29/92	1560	7.352	4.098	4.526	92.429	352.831
3/1/92	1040	6.947	3.741	4.169	64.667	164.570
3/2/92	970	6.877	3.679	4.108	60.817	144.355
3/3/92	930	6.835	3.642	4.071	58.602	133.362
3/4/92	840	6.733	3.553	3.981	53.576	110.125
3/5/92	670	6.507	3.353	3.782	43.900	71.973
3/6/92	1060	6.966	3.757	4.186	65.761	170.573
3/7/92	1730	7.456	4.189	4.618	101.247	428.608
3/8/92	1590	7.371	4.115	4.543	93.993	365.701
3/9/92	1140	7.039	3.822	4.250	70.114	195.590
3/10/92	1080	6.985	3.774	4.202	66.853	176.677
3/11/92	1100	7.003	3.790	4.219	67.943	182.881
3/12/92	1010	6.918	3.715	4.143	63.021	155.754
3/13/92	897	6.799	3.610	4.039	56.766	124.600
3/14/92	761	6.635	3.466	3.894	49.112	91.454
3/15/92	691	6.538	3.381	3.809	45.110	76.275
3/16/92	600	6.397	3.256	3.685	39.833	58.483
3/17/92	690	6.537	3.379	3.808	45.052	76.067
3/18/92	816	6.704	3.527	3.956	52.226	104.282
3/19/92	937	6.843	3.649	4.077	58.991	135.256
3/20/92	1080	6.985	3.774	4.202	66.853	176.677
3/21/92	1070	6.975	3.766	4.194	66.308	173.613

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
3/22/92	962	6.869	3.672	4.101	60.375	142.123
3/23/92	1020	6.928	3.724	4.152	63.570	158.668
3/24/92	994	6.902	3.701	4.129	62.141	151.146
3/25/92	1380	7.230	3.990	4.418	82.966	280.166
3/26/92	2220	7.705	4.409	4.837	126.122	685.139
3/27/92	2830	7.948	4.623	5.051	156.197	1081.662
3/28/92	2240	7.714	4.417	4.845	127.123	696.795
3/29/92	1710	7.444	4.179	4.607	100.215	419.336
3/30/92	1670	7.421	4.158	4.586	98.147	401.076
3/31/92	1640	7.402	4.142	4.570	96.592	387.631
4/1/92	1630	7.396	4.137	4.565	96.073	383.198
4/2/92	1540	7.340	4.087	4.515	91.384	344.371
4/3/92	1290	7.162	3.930	4.359	78.181	246.788
4/4/92	1130	7.030	3.814	4.242	69.572	192.375
4/5/92	956	6.863	3.667	4.095	60.043	140.461
4/6/92	829	6.720	3.541	3.969	52.958	107.428
4/7/92	919	6.823	3.632	4.060	57.991	130.410
4/8/92	559	6.326	3.194	3.622	37.425	51.193
4/9/92	467	6.146	3.035	3.464	31.943	36.502
4/10/92	492	6.198	3.081	3.510	33.444	40.264
4/11/92	651	6.479	3.328	3.757	42.801	68.182
4/12/92	909	6.812	3.622	4.051	57.435	127.754
4/13/92	625	6.438	3.292	3.721	41.292	63.150
4/14/92	528	6.269	3.144	3.572	35.591	45.984
4/15/92	468	6.148	3.037	3.466	32.003	36.649
4/16/92	583	6.368	3.231	3.659	38.837	55.405
4/17/92	2480	7.816	4.506	4.935	139.048	843.818
4/18/92	3630	8.197	4.842	5.270	194.499	1727.661

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
4/19/92	2790	7.934	4.610	5.039	154.250	1053.085
4/20/92	1900	7.550	4.272	4.700	109.962	511.245
4/21/92	1960	7.581	4.299	4.728	113.015	542.033
4/22/92	2060	7.630	4.343	4.771	118.079	595.217
4/23/92	1470	7.293	4.046	4.474	87.715	315.518
4/24/92	1410	7.251	4.009	4.437	84.553	291.731
4/25/92	1450	7.279	4.033	4.462	86.663	307.492
4/26/92	1280	7.155	3.924	4.352	77.647	243.202
4/27/92	883	6.783	3.597	4.025	55.985	120.967
4/28/92	797	6.681	3.506	3.935	51.153	99.761
4/29/92	894	6.796	3.607	4.036	56.599	123.817
4/30/92	950	6.856	3.661	4.090	59.711	138.807
5/1/92	876	6.775	3.590	4.018	55.594	119.170
5/2/92	733	6.597	3.433	3.861	47.516	85.228
5/3/92	759	6.632	3.463	3.892	48.998	91.003
5/4/92	600	6.397	3.256	3.685	39.833	58.483
5/5/92	599	6.395	3.255	3.683	39.775	58.300
5/6/92	596	6.390	3.250	3.679	39.599	57.752
5/7/92	542	6.295	3.167	3.595	36.421	48.304
5/8/92	492	6.198	3.081	3.510	33.444	40.264
5/9/92	526	6.265	3.140	3.569	35.472	45.657
5/10/92	492	6.198	3.081	3.510	33.444	40.264
5/11/92	460	6.131	3.022	3.451	31.520	35.480
5/12/92	431	6.066	2.965	3.393	29.763	31.390
5/13/92	343	5.838	2.764	3.192	24.339	20.428
5/14/92	308	5.730	2.669	3.097	22.137	16.684
5/15/92	305	5.720	2.660	3.089	21.947	16.380
5/16/92	306	5.724	2.663	3.092	22.011	16.481

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
5/17/92	296	5.690	2.634	3.062	21.376	15.483
5/18/92	515	6.244	3.122	3.550	34.818	43.877
5/19/92	369	5.911	2.828	3.256	25.957	23.438
5/20/92	319	5.765	2.700	3.128	22.832	17.823
5/21/92	291	5.673	2.619	3.047	21.057	14.994
5/22/92	252	5.529	2.492	2.920	18.550	11.439
5/23/92	274	5.613	2.566	2.994	19.970	13.389
5/24/92	1050	6.957	3.749	4.178	65.215	167.559
5/25/92	500	6.215	3.096	3.524	33.923	41.505
5/26/92	408	6.011	2.916	3.345	28.359	28.313
5/27/92	352	5.864	2.786	3.215	24.901	21.448
5/28/92	314	5.749	2.686	3.114	22.517	17.301
5/29/92	303	5.714	2.654	3.083	21.820	16.179
5/30/92	605	6.405	3.263	3.692	40.125	59.403
5/31/92	956	6.863	3.667	4.095	60.043	140.461
6/1/92	584	6.370	3.232	3.661	38.896	55.584
6/2/92	465	6.142	3.032	3.460	31.822	36.209
6/3/92	384	5.951	2.863	3.292	26.885	25.262
6/4/92	310	5.737	2.674	3.103	22.264	16.889
6/5/92	309	5.733	2.672	3.100	22.201	16.786
6/6/92	380	5.940	2.854	3.282	26.638	24.769
6/7/92	279	5.631	2.582	3.010	20.291	13.853
6/8/92	264	5.576	2.533	2.961	19.326	12.485
6/9/92	239	5.476	2.445	2.874	17.705	10.354
6/10/92	217	5.380	2.360	2.789	16.261	8.635
6/11/92	205	5.323	2.310	2.739	15.466	7.758
6/12/92	201	5.303	2.293	2.721	15.200	7.476
6/13/92	197	5.283	2.275	2.704	14.933	7.199

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
6/14/92	191	5.252	2.248	2.676	14.532	6.792
6/15/92	356	5.875	2.796	3.225	25.150	21.909
6/16/92	232	5.447	2.419	2.848	17.247	9.791
6/17/92	226	5.421	2.396	2.825	16.854	9.320
6/18/92	669	6.506	3.352	3.781	43.842	71.771
6/19/92	570	6.346	3.211	3.640	38.073	53.104
6/20/92	348	5.852	2.776	3.205	24.651	20.992
6/21/92	279	5.631	2.582	3.010	20.291	13.853
6/22/92	267	5.587	2.543	2.971	19.520	12.753
6/23/92	257	5.549	2.509	2.938	18.874	11.870
6/24/92	554	6.317	3.186	3.614	37.130	50.335
6/25/92	399	5.989	2.897	3.325	27.808	27.150
6/26/92	343	5.838	2.764	3.192	24.339	20.428
6/27/92	300	5.704	2.646	3.074	21.630	15.879
6/28/92	248	5.513	2.478	2.906	18.291	11.100
6/29/92	223	5.407	2.384	2.813	16.656	9.089
6/30/92	225	5.416	2.392	2.821	16.788	9.243
7/1/92	217	5.380	2.360	2.789	16.261	8.635
7/2/92	214	5.366	2.348	2.776	16.063	8.411
7/3/92	213	5.361	2.344	2.772	15.997	8.338
7/4/92	252	5.529	2.492	2.920	18.550	11.439
7/5/92	231	5.442	2.415	2.844	17.182	9.712
7/6/92	371	5.916	2.833	3.261	26.081	23.677
7/7/92	245	5.501	2.467	2.896	18.096	10.849
7/8/92	223	5.407	2.384	2.813	16.656	9.089
7/9/92	224	5.412	2.388	2.817	16.722	9.166
7/10/92	254	5.537	2.499	2.927	18.680	11.610
7/11/92	706	6.560	3.399	3.828	45.971	79.419

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
7/12/92	465	6.142	3.032	3.460	31.822	36.209
7/13/92	2660	7.886	4.568	4.997	147.901	962.688
7/14/92	1270	7.147	3.917	4.345	77.112	239.640
7/15/92	3010	8.010	4.677	5.105	164.916	1214.684
7/16/92	2170	7.682	4.389	4.817	123.616	656.402
7/17/92	3370	8.123	4.776	5.205	182.173	1502.270
7/18/92	2870	7.962	4.635	5.063	158.140	1110.598
7/19/92	1560	7.352	4.098	4.526	92.429	352.831
7/20/92	1130	7.030	3.814	4.242	69.572	192.375
7/21/92	1350	7.208	3.971	4.399	81.375	268.820
7/22/92	1070	6.975	3.766	4.194	66.308	173.613
7/23/92	908	6.811	3.621	4.050	57.379	127.490
7/24/92	2120	7.659	4.368	4.797	121.104	628.243
7/25/92	2660	7.886	4.568	4.997	147.901	962.688
7/26/92	2500	7.824	4.513	4.942	140.035	856.664
7/27/92	1720	7.450	4.184	4.612	100.731	423.960
7/28/92	1150	7.048	3.829	4.258	70.656	198.829
7/29/92	839	6.732	3.552	3.980	53.520	109.879
7/30/92	3730	8.224	4.866	5.294	199.212	1818.267
7/31/92	8140	9.005	5.553	5.982	396.165	7891.047
8/1/92	4400	8.389	5.011	5.440	230.418	2480.863
8/2/92	2470	7.812	4.503	4.931	138.554	837.430
8/3/92	1340	7.200	3.964	4.393	80.844	265.087
8/4/92	1580	7.365	4.109	4.538	93.472	361.387
8/5/92	1180	7.073	3.852	4.281	72.277	208.698
8/6/92	822	6.712	3.533	3.962	52.564	105.728
8/7/92	651	6.479	3.328	3.757	42.801	68.182
8/8/92	652	6.480	3.329	3.758	42.859	68.379

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
8/9/92	925	6.830	3.637	4.066	58.325	132.016
8/10/92	477	6.168	3.054	3.483	32.544	37.986
8/11/92	841	6.735	3.554	3.982	53.633	110.372
8/12/92	631	6.447	3.301	3.729	41.641	64.296
8/13/92	845	6.739	3.558	3.986	53.857	111.361
8/14/92	612	6.417	3.274	3.702	40.534	60.702
8/15/92	1190	7.082	3.859	4.288	72.817	212.037
8/16/92	2880	7.966	4.638	5.067	158.625	1117.888
8/17/92	1960	7.581	4.299	4.728	113.015	542.033
8/18/92	1200	7.090	3.867	4.295	73.355	215.401
8/19/92	948	6.854	3.659	4.088	59.600	138.258
8/20/92	802	6.687	3.512	3.940	51.435	100.942
8/21/92	691	6.538	3.381	3.809	45.110	76.275
8/22/92	555	6.319	3.187	3.616	37.189	50.506
8/23/92	457	6.125	3.016	3.445	31.339	35.046
8/24/92	423	6.047	2.948	3.377	29.276	30.303
8/25/92	746	6.615	3.448	3.877	48.258	88.093
8/26/92	415	6.028	2.931	3.360	28.788	29.234
8/27/92	1150	7.048	3.829	4.258	70.656	198.829
8/28/92	4520	8.416	5.035	5.464	235.945	2609.654
8/29/92	3480	8.155	4.805	5.233	187.401	1595.827
8/30/92	2000	7.601	4.317	4.745	115.044	563.027
8/31/92	1410	7.251	4.009	4.437	84.553	291.731
9/1/92	1260	7.139	3.910	4.338	76.577	236.104
9/2/92	1200	7.090	3.867	4.295	73.355	215.401
9/3/92	1790	7.490	4.219	4.648	104.334	456.995
9/4/92	2170	7.682	4.389	4.817	123.616	656.402
9/5/92	1250	7.131	3.903	4.331	76.041	232.591

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
9/6/92	996	6.904	3.703	4.131	62.251	151.718
9/7/92	868	6.766	3.581	4.010	55.147	117.131
9/8/92	874	6.773	3.588	4.016	55.482	118.659
9/9/92	873	6.772	3.587	4.015	55.426	118.403
9/10/92	3260	8.089	4.747	5.176	176.925	1411.366
9/11/92	1920	7.560	4.281	4.709	110.981	521.414
9/12/92	1290	7.162	3.930	4.359	78.181	246.788
9/13/92	1030	6.937	3.732	4.161	64.119	161.606
9/14/92	941	6.847	3.653	4.081	59.212	136.344
9/15/92	872	6.771	3.585	4.014	55.370	118.148
9/16/92	649	6.475	3.325	3.754	42.685	67.789
9/17/92	488	6.190	3.074	3.503	33.205	39.651
9/18/92	735	6.600	3.435	3.863	47.631	85.666
9/19/92	1290	7.162	3.930	4.359	78.181	246.788
9/20/92	622	6.433	3.288	3.716	41.117	62.582
9/21/92	876	6.775	3.590	4.018	55.594	119.170
9/22/92	3310	8.105	4.761	5.189	179.313	1452.357
9/23/92	2200	7.696	4.401	4.829	125.121	673.575
9/24/92	1440	7.272	4.027	4.456	86.136	303.516
9/25/92	1140	7.039	3.822	4.250	70.114	195.590
9/26/92	1010	6.918	3.715	4.143	63.021	155.754
9/27/92	932	6.837	3.644	4.073	58.713	133.902
9/28/92	749	6.619	3.452	3.880	48.429	88.761
9/29/92	641	6.463	3.314	3.743	42.221	66.226
9/30/92	566	6.339	3.205	3.633	37.838	52.406
10/1/92	542	6.295	3.167	3.595	36.421	48.304
10/2/92	537	6.286	3.158	3.587	36.125	47.469
10/3/92	956	6.863	3.667	4.095	60.043	140.461

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
10/4/92	984	6.892	3.692	4.120	61.590	148.298
10/5/92	614	6.420	3.276	3.705	40.651	61.076
10/6/92	407	6.009	2.914	3.343	28.298	28.183
10/7/92	310	5.737	2.674	3.103	22.264	16.889
10/8/92	294	5.684	2.628	3.056	21.249	15.287
10/9/92	316	5.756	2.691	3.120	22.643	17.509
10/10/92	289	5.666	2.613	3.041	20.930	14.801
10/11/92	361	5.889	2.809	3.237	25.461	22.491
10/12/92	341	5.832	2.758	3.187	24.214	20.205
10/13/92	298	5.697	2.640	3.068	21.503	15.680
10/14/92	627	6.441	3.295	3.723	41.408	63.531
10/15/92	871	6.770	3.584	4.013	55.314	117.894
10/16/92	1490	7.307	4.057	4.486	88.765	323.641
10/17/92	1190	7.082	3.859	4.288	72.817	212.037
10/18/92	753	6.624	3.456	3.885	48.657	89.654
10/19/92	693	6.541	3.383	3.812	45.225	76.691
10/20/92	526	6.265	3.140	3.569	35.472	45.657
10/21/92	539	6.290	3.162	3.590	36.243	47.802
10/22/92	470	6.153	3.041	3.470	32.123	36.945
10/23/92	446	6.100	2.995	3.423	30.674	33.476
10/24/92	541	6.293	3.165	3.594	36.362	48.137
10/25/92	523	6.260	3.135	3.564	35.294	45.168
10/26/92	486	6.186	3.071	3.499	33.085	39.346
10/27/92	449	6.107	3.001	3.429	30.855	33.901
10/28/92	420	6.040	2.942	3.370	29.093	29.900
10/29/92	406	6.006	2.912	3.341	28.237	28.053
10/30/92	390	5.966	2.877	3.305	27.254	26.010
10/31/92	359	5.883	2.804	3.232	25.337	22.257

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
11/1/92	390	5.966	2.877	3.305	27.254	26.010
11/2/92	1170	7.065	3.844	4.273	71.737	205.383
11/3/92	1850	7.523	4.248	4.677	107.408	486.233
11/4/92	1180	7.073	3.852	4.281	72.277	208.698
11/5/92	998	6.906	3.704	4.133	62.361	152.292
11/6/92	1140	7.039	3.822	4.250	70.114	195.590
11/7/92	1000	6.908	3.706	4.135	62.471	152.866
11/8/92	887	6.788	3.601	4.029	56.209	122.000
11/9/92	804	6.690	3.514	3.943	51.548	101.416
11/10/92	611	6.415	3.272	3.701	40.476	60.516
11/11/92	884	6.784	3.598	4.026	56.041	121.225
11/12/92	4180	8.338	4.966	5.395	220.238	2252.695
11/13/92	5740	8.655	5.246	5.674	291.224	4090.475
11/14/92	2840	7.952	4.626	5.054	156.683	1088.863
11/15/92	2020	7.611	4.326	4.754	116.057	573.664
11/16/92	1800	7.496	4.224	4.653	104.847	461.809
11/17/92	1850	7.523	4.248	4.677	107.408	486.233
11/18/92	1590	7.371	4.115	4.543	93.993	365.701
11/19/92	1400	7.244	4.003	4.431	84.025	287.852
11/20/92	1000	6.908	3.706	4.135	62.471	152.866
11/21/92	1150	7.048	3.829	4.258	70.656	198.829
11/22/92	2470	7.812	4.503	4.931	138.554	837.430
11/23/92	4130	8.326	4.956	5.384	217.916	2202.279
11/24/92	2450	7.804	4.496	4.924	137.565	824.721
11/25/92	2660	7.886	4.568	4.997	147.901	962.688
11/26/92	1720	7.450	4.184	4.612	100.731	423.960
11/27/92	1690	7.432	4.168	4.597	99.182	410.158
11/28/92	1550	7.346	4.092	4.521	91.907	348.589

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, continued.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
11/29/92	1410	7.251	4.009	4.437	84.553	291.731
11/30/92	1280	7.155	3.924	4.352	77.647	243.202
12/1/92	1130	7.030	3.814	4.242	69.572	192.375
12/2/92	1010	6.918	3.715	4.143	63.021	155.754
12/3/92	1030	6.937	3.732	4.161	64.119	161.606
12/4/92	945	6.851	3.656	4.085	59.434	137.436
12/5/92	902	6.805	3.615	4.044	57.045	125.910
12/6/92	857	6.753	3.570	3.999	54.530	114.355
12/7/92	801	6.686	3.511	3.939	51.379	100.705
12/8/92	767	6.642	3.472	3.901	49.453	92.815
12/9/92	743	6.611	3.444	3.873	48.087	87.428
12/10/92	813	6.701	3.524	3.952	52.056	103.562
12/11/92	1050	6.957	3.749	4.178	65.215	167.559
12/12/92	1230	7.115	3.889	4.317	74.968	225.641
12/13/92	1140	7.039	3.822	4.250	70.114	195.590
12/14/92	1080	6.985	3.774	4.202	66.853	176.677
12/15/92	1020	6.928	3.724	4.152	63.570	158.668
12/16/92	1140	7.039	3.822	4.250	70.114	195.590
12/17/92	1060	6.966	3.757	4.186	65.761	170.573
12/18/92	967	6.874	3.677	4.105	60.651	143.516
12/19/92	928	6.833	3.640	4.069	58.491	132.823
12/20/92	1630	7.396	4.137	4.565	96.073	383.198
12/21/92	1370	7.223	3.983	4.412	82.437	276.359
12/22/92	1120	7.021	3.806	4.235	69.030	189.186
12/23/92	997	6.905	3.704	4.132	62.306	152.005
12/24/92	899	6.801	3.612	4.041	56.878	125.123
12/25/92	785	6.666	3.493	3.921	50.474	96.955
12/26/92	689	6.535	3.378	3.807	44.995	75.860

Table C1: Estimated concentrations and daily loads in the Cuyahoga River, concluded.

Date	Q	ln(Q)	ln(c^)	ln(c^)+s^2/2	c^	daily load
12/27/92	657	6.488	3.336	3.765	43.149	69.369
12/28/92	629	6.444	3.298	3.726	41.524	63.913
12/29/92	691	6.538	3.381	3.809	45.110	76.275
12/30/92	2810	7.941	4.616	5.045	155.224	1067.329
12/31/92	8940	9.098	5.636	6.064	430.271	9412.679