National Nonpoint Source Monitoring Program

Tech NOTES 8

Through the National Nonpoint Source Monitoring Program **(NNPSMP)**, states monitor and evaluate a subset of watershed projects funded by the Clean Water Act Section 319 Nonpoint Source Control Program.

The program has two major objectives:

1. To scientifically evaluate the effectiveness of watershed technologies designed to control nonpoint source pollution
2. To improve our understanding of nonpoint source pollution

NNPSMP Tech Notes is a series of publications that shares this unique research and monitoring effort. It offers guidance on data collection, implementation of pollution control technologies, and monitoring design, as well as case studies that illustrate principles in action.

April 2013

Donald W. Meals, R. Peter Richards[1], and Steven A. Dressing. 2013. Pollutant load estimation for water quality monitoring projects. Tech Notes 8, April 2013. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 21 p. Available online at *https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes*.

[1]Heidelberg University, Tiffin, OH.

Pollutant Load Estimation for Water Quality Monitoring Projects

Introduction

Determination of pollutant load is a key objective for many nonpoint source (NPS) monitoring projects. The mass of nutrients delivered to a lake or estuary drives the productivity of the waterbody. The annual suspended sediment load transported by a river is usually a more meaningful indicator of soil loss in the watershed than is a suspended sediment concentration. The foundation of water resource management embodied in the TMDL (total maximum daily load) concept lies in assessment of the maximum pollutant load a waterbody can accept before becoming impaired and in the measurement of changes in pollutant loads in response to implementation of management measures.

Estimation of pollutant load through monitoring is a complex task that requires accurate measurement of both pollutant concentration and water flow and careful calculation, often based on a statistical approach. It is imperative that a NPS monitoring program be designed for good load estimation at the start. This Tech Note addresses important considerations and procedures for developing good pollutant load estimates in NPS monitoring projects. Much of the material is taken from an extensive monograph written by Dr. R. Peter Richards, of Heidelberg University, ***Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs***. The reader is encouraged to consult that document and its associated tools for additional information on load estimation.

General Considerations

Definitions

Load may be defined as the mass of a substance that passes a particular point of a river (such as a monitoring station on a watershed outlet) in a specified amount of time (e.g., daily, annually). Mathematically, load is essentially the product of water discharge and the concentration of a substance in the water. Flux is a term that describes the loading rate, i.e., the instantaneous rate at which the load passes a point in the river. Water discharge is defined as the volume of water that passes a cross-section of a river in a specified amount of time, while flow refers to the discharge rate, the instantaneous rate at which water passes a point. Refer to ***Tech Notes #3*** (Meals and Dressing 2008) for guidance on appropriate ways to estimate or measure surface water flow for purposes associated with NPS watershed projects.

Basic Terms

Flux – instantaneous loading rate (e.g., kg/sec)

Flow rate – instantaneous rate of water passage (e.g., L/sec)

Discharge – quantity of water passing a specified point (e.g., m^3)

Load – mass of substance passing a specified point (e.g., metric tons)

If we could directly and continuously measure the flux of a pollutant, the results might look like the plot in Figure 1. The load transported over the entire period of time in the graph would simply be equal to the shaded area under the curve.

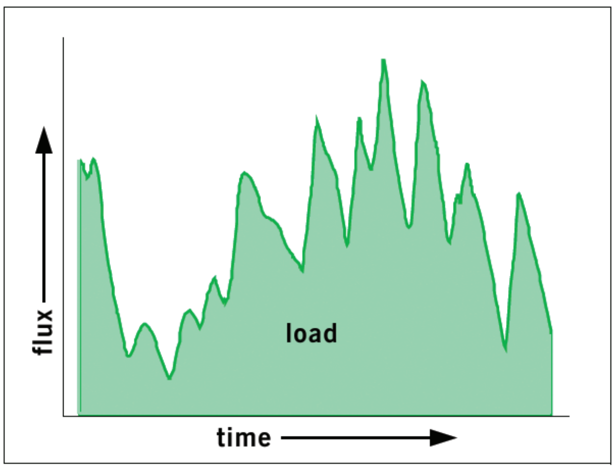


Figure 1. Imaginary plot of pollutant flux over time at a monitoring station (Richards 1998).

However, we cannot measure flux directly, so we calculate it as the product of instantaneous concentration and instantaneous flow:

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

where c is concentration and q is flow, both a function of time (t), and k is a unit conversion factor. Because we must take a series of discrete samples to measure concentration, the load estimate becomes the sum of a set of *n* products of concentration (c), flow (q), and the time interval (Δt) over which the concentration and flow measurements apply:

$$\text{Load} = k \sum_{i=1}^{n} c_i q_i \Delta t$$

The main monitoring challenge becomes how best to take the discrete samples to give the most accurate estimate of load. Note that the total load is the load over the timeframe of interest (e.g., one year) determined by summing a series of unit loads (individual calculations of load as the product of concentration, flow, and time over smaller, more homogeneous time spans). The central problem is to obtain good measures of concentration and flow during each time interval. As a general rule, in cases where sampling frequency is high relative to the timeframe of interest (e.g., daily sampling for annual loads) instantaneous concentrations from single grab samples can be used with discharge at the time of sampling or mean discharge values for the time interval. If less frequent sampling is performed (e.g., weekly for annual loads), however, it is recommended that concentrations from flow-weighted composite samples are used with total discharge estimates for the time period over which each sample is composited. Once these choices—which are described in greater detail on the following pages—have been made, calculation of total load by summing unit loads is simple arithmetic.

Issues of Variability

Both flow and concentration vary considerably over time, especially in NPS situations. Accurate load estimation becomes an exercise in both how many samples to take and when to take them to account for this variability.

Sampling frequency has a major influence on the accuracy of load estimation, as shown in Figure 2. The top panel shows daily suspended solids load (calculated as the products of daily total suspended solids (TSS) concentration and mean daily discharge measured at a continuously recording U.S. Geological Survey (USGS) station) for the Sandusky River in Ohio. The middle panel represents load calculated using weekly TSS samples and mean weekly discharge; the lower panel shows load calculated from monthly TSS samples and mean monthly discharge data. Clearly, very different pictures of suspended solids load emerge from different sampling frequencies, as decreasing sampling frequencies tend to miss more and more short-term but important events with high flow or high TSS concentrations.

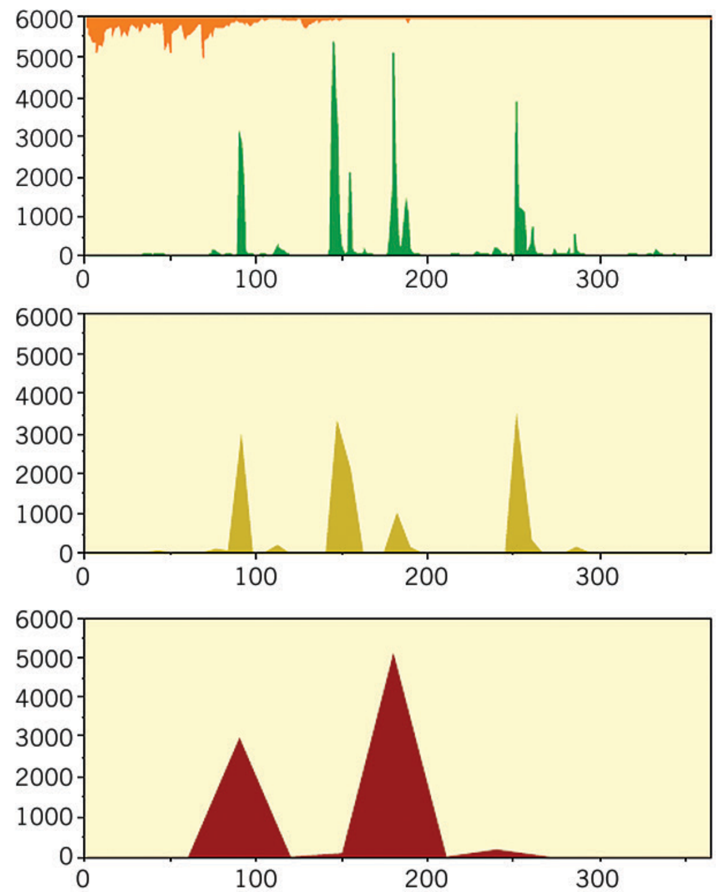


Figure 2. Plot of suspended solids loads for the Sandusky River, water year 1985. Top: daily TSS samples; middle: weekly samples; bottom: monthly samples. Weekly and monthly sample values were drawn from actual daily sample data series (Richards 1998).

Because in NPS situations most flux occurs during periods of high discharge (e.g., ~80–90% of annual load may be delivered in ~10–20% of time), choosing *when* to sample can be as important as how often to sample. The top panel in Figure 3 shows a plot of daily suspended solids load derived from weekly sampling superimposed on daily flux data; the bottom panel shows daily loads derived from monthly and quarterly sampling on top of the same daily flux data. Weekly samples give a reasonably good visual fit over the daily flux pattern. The monthly series gives only a very crude representation of the daily flux, but it is somewhat better than expected, 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 suspended solids load. Quarterly samples result in a poor fit on the actual daily flux pattern.

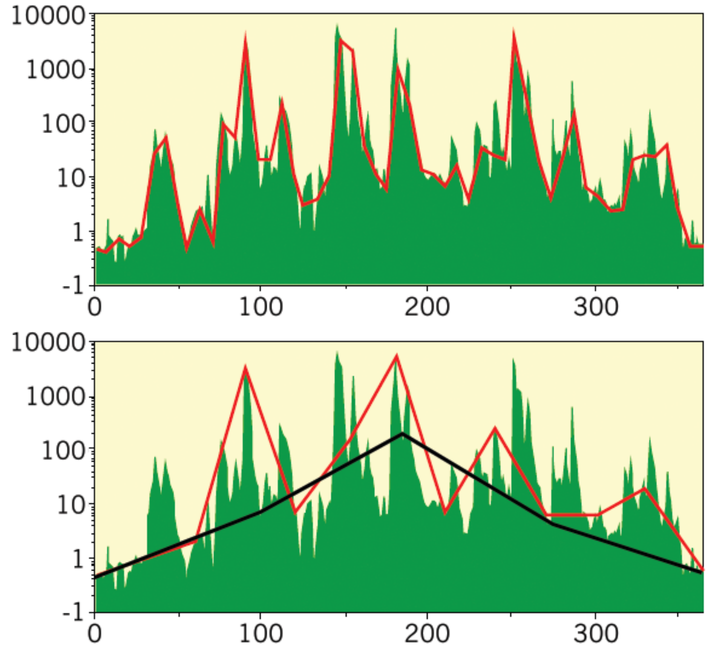


Figure 3. Weekly (red line in top panel) and monthly (red line) and quarterly (black line) (bottom panel) suspended solids load time series superimposed on a daily load time series (Richards 1998).

The key point here is that many samples are typically needed to accurately and reliably capture the true load pattern. Quarterly observations are generally inadequate, monthly observations will probably not yield reliable load estimates, and even weekly observations may not be satisfactory, especially if very accurate load estimates are required to achieve project objectives.

Practical Load Estimation

Ideally, the most accurate approach to estimating pollutant load would be to sample very frequently and capture all the variability. Flow is relatively straightforward to measure continuously (see ***Meals and Dressing 2008***), but concentration is expensive to measure and in most cases impossible to measure continuously. It is therefore critically important to choose a sampling interval that will yield a suitable characterization of concentration.

There are three important considerations involved in sampling for good load estimation: sample type, sampling frequency, and sample distribution in time. Grab samples represent a concentration only at a single point in time and the selection of grab sampling interval must be made in consideration of the issues of variability discussed above. Integrated samples (composite samples made up of many individual grab samples) are frequently used in NPS monitoring. Time-integrated or time-proportional samples are either taken at a constant rate over the time period or are composed of subsamples taken at a fixed frequency. Time-integrated samples are poorly suited for load estimation because they are taken without regard to changes in flow (and concentration) that may occur during the integration period and are usually biased toward the low flows that occur most often. Flow-proportional samples (where a sample is collected for every n units of flow that pass the station), on the other hand, 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. However, collecting flow-proportional samples is technically challenging and may not be suitable for all purposes. Also, even though a flow-proportional sample over a time span (e.g., a week) is a good summation of the variability of that week, ability to see what happened within that week (e.g., a transient spike in concentration) is lost. Flow-proportional sampling is also not compatible with some monitoring demands, such as monitoring for ambient concentrations that are highest at low flow or for documenting exceedance of critical values (e.g., a water quality standard).

Sampling frequency determines the number of unit load estimates that can be computed and summed for an estimate of total load. Using more unit loads increases the probability of capturing variability across the year and not missing an important event (see Figure 3); in general, the accuracy and precision of a load estimate increases as sampling frequency increases. Over a sufficiently short interval between samples, a sampling program will probably not miss a sudden peak in flux. If, for example, unit loads are calculated by multiplying the average concentration for the time unit by the discharge over the same time unit, the annual load that is the sum of four quarterly unit loads will be considerably less accurate than the annual load that is the sum of twelve monthly loads. Note that this example does not mean that an annual load calculated from 12 monthly loads is sufficiently accurate for all purposes.

There is a practical limit to the benefits of increasing sampling frequency, however, due to the fact that water quality data tend to be autocorrelated. The concentration or flux at a certain point today is related to the concentration or flux at the same point yesterday and, perhaps to a lesser extent, to the concentration or flux at that spot last week. Because of this autocorrelation, beyond some point, increasing sampling frequency will accomplish little in the way of generating new information. This is usually not a problem for monitoring programs, but can be a concern, however, when electronic sensors are used to collect data nearly continuously.

Consideration of the basic sampling frequency—n samples per year—does not address the more complex issue of timing. The choice of *when* to collect concentration samples is critical. Most NPS water quality data have a strong seasonal component as well as a strong association with other variable factors such as precipitation, streamflow, or watershed management activities such as tillage or fertilizer application. Selecting when to collect samples for concentration determination is essentially equivalent to selecting when the unit loads that go into an annual load estimate are determined. That choice must consider the fundamental characteristics of the system being monitored. In northern climates, spring snowmelt is often the dominant export event of the year; sampling during that period may need to be more intensive than during midsummer in order to capture the most important peak flows and concentrations. In southern regions, intensive summer storms often generate the majority of annual pollutant load; intensive summer monitoring may be required to obtain good load estimates. For many agricultural pesticides, sampling may need to be focused on the brief period immediately after application when most losses tend to occur. Issues of random sampling, stratified random sampling, and other sampling regimes should be considered. Simple random sampling may be inappropriate for accurate load estimation if, as is likely, the resulting schedule is biased toward low flow conditions. Stratified random sampling—division of the sampling effort or the sample set into two or more parts which are different from each other but relatively homogeneous within—could be a better strategy. In cases where there is a conflict between the number of observations a program can afford and the number needed to obtain an accurate and reliable load estimate, it may be possible to use flow as the basis for selecting the interval between concentration observations. For example, planning to collect samples every x thousand ft^3 of discharge would automatically emphasize high flux conditions while economizing on sampling during baseflow conditions.

How accurate does your load estimate need to be?

The required accuracy of load estimates is driven by project objectives, and should be specified in the project Quality Assurance Project Plan (QAPP)—see the ***Quality Management Tools*** provided by the U.S. Environmental Protection Agency (EPA) for additional information on QAPPs. Fundamentally, the accuracy of a load estimate depends on the accuracy of the component concentration and discharge measurements. Generally, if a monitoring program is conducted to document a difference in loads from one period or one site to the next or to identify a long-term trend (see ***Meals et al. 2011***), the confidence in the load estimate must be at least as great as the anticipated difference or change. In this context, the confidence in the estimate derives not only from inherent uncertainties in measurement of concentration and/or discharge but also from the influence of natural variability and the ability of the monitoring program to address it.

The Minimum Detectable Change (MDC) is the minimum change in a pollutant concentration or load over a given period of time required to be considered statistically significant (see ***Spooner et al. 2011***). The calculation of MDC has several practical uses. Data collected in the first several years of a project or from a similar project can be used to determine how much change must be measured in the water resource to be considered statistically significant and not an artifact of system variability. These calculations facilitate realistic expectations when projecting water quality results. Calculation of the magnitude of the water quality change required can serve as a useful tool to evaluate water quality monitoring designs for their effectiveness in detecting changes in water quality. Closely related, these calculations can also be used to design effective water quality monitoring networks (Spooner et al. 1987, 1988).

The MDC is a function of pollutant variability, sampling frequency, length of monitoring time, explanatory variables (e.g., season, meteorologic, and hydrologic variables) used in the analyses that help explain some of the variability in the measured data, magnitude and structure of the autocorrelation, and statistical techniques and significance level used to analyze the data.

It is recommended that the reader consult ***Tech Notes #7*** (Spooner et al. 2011) for detailed information on how to calculate and interpret issues of MDC.

Planning Monitoring Programs for Effective Load Estimation

Both discharge and concentration data are needed to calculate pollutant loads, but monitoring programs designed for load estimation will usually generate more flow than concentration data. This leaves three basic choices for practical load estimation:

1. Find a way to estimate un-measured concentrations to go with the flows observed at times when chemical samples were not taken;
2. Throw out most of the flow data and calculate the load using the concentration data and just those flows observed at the same time the samples were taken; and
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 unsatisfactory because the frequency of chemical observations is likely to be inadequate to give a reliable load estimate when simple summation is used. Thus almost all effective load estimation approaches are variants of approaches 1 or 3.

Unfortunately, the decision to calculate loads is sometimes made after the data are collected, often using data collected for other purposes. At that point, little can be done to compensate for a data set that contains too few observations of concentration, discharge, or both, collected using an inappropriate sampling design. Many programs choose monthly or quarterly sampling with no better rationale than convenience and tradition. 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 half of the time (Richards and Holloway 1987).

To avoid such problems, the sampling regime needed for load estimation must be established in the initial monitoring design, 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 following steps are recommended to plan a monitoring effort for load estimation:

- Determine whether the project goals require knowledge of load, or if goals can be met using concentration data alone. In many cases, especially when trend detection is the goal, concentration data may be easier to work with and be more accurate than crudely estimated load data.

- If load estimates are required, determine the accuracy and precision needed based on the uses to which they will be put. This is especially critical when the purpose of monitoring is to look for a change in load. It is foolish to attempt to document a 25% load reduction from a watershed program with a monitoring design that gives load estimates ±50% of the true load (see ***Tech Notes #7*** (Spooner et al. 2011)).
- Decide which 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 flow or season.
- Calculate the loads based on the samples obtained after the first full year of monitoring, and compare the precision estimates (of both flow measurement and the sampling program) with the initial goals of the program. Adjust the sampling program if the estimated precision deviates substantially from the goals.

It is possible that funding or other limitations may prevent a monitoring program from collecting the data required for acceptable load estimation. In such a case, the question must be asked: is a biased, highly uncertain load estimate preferable to no load estimate at all? Sometimes the correct answer will be no.

Approaches to Load Estimation

Several distinct technical approaches to load estimation are discussed below. The reader is encouraged to consult ***Richards (1998)*** for details and examples of these calculations.

Numeric Integration

The simplest approach is numeric integration, where the total load is given by

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

where c_i is the concentration in the i^{th} sample, q_i is the corresponding flow, and t_i is the time interval represented by the i^{th} sample, calculated as:

$$\frac{1}{2}(t_{i+1} - t_{i-1})$$

It is not required that t_i be the same for each sample.

The question becomes how fine to slice the pie — few slices will miss much variability, many slices will capture variability but at a higher cost and monitoring effort. Numeric integration is only satisfactory if the sampling frequency is high — often on the order of 100 samples per year or more, and samples must be distributed so that all major runoff events are captured. Selection of sampling frequency and distribution over the year is critical — sampling must focus on times when highest fluxes occur, i.e., periods of high discharge.

As noted above, flow-proportional sampling is a special case of mechanical rather than mathematical integration that assumes that one or more samples can be obtained that cover the entire period of interest, each representing a known discharge and each with a concentration that is in proportion to the load that 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.

Regression

When, as is often the case with NPS-dominated systems, a strong relationship exists between flow and concentration, using regression to estimate load from continuous flow and intermittent concentration data can be highly effective. In this approach, a regression relationship is developed between concentration and flow based on the days for which concentration data exist. Usually, these data are based on grab samples for concentration and mean daily flow for the sampling day. This relationship may involve simple or multiple regression analysis using covariates like precipitation. In most applications, both concentration and flow are typically log-transformed to create a dataset suited for regression analysis (see ***Tech Notes #1*** (Meals and Dressing 2005) for basic information on data transformations). 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. This method requires that there be a strong linear association between flow and concentration that does not change appreciably over the period of interest. If BMP implementation is expected to affect the relationship between flow and concentration, such relationships must be tracked carefully — if BMPs change the relationship, the concentration estimation procedure must be corrected.

Once the regression relationship is established, the regression equation is used to estimate concentrations for each day on which a sample was not taken, based on the mean daily flow for the day. The total load is calculated as the sum of the daily loads that are obtained by multiplying the measured or estimated daily concentration by the total daily discharge.

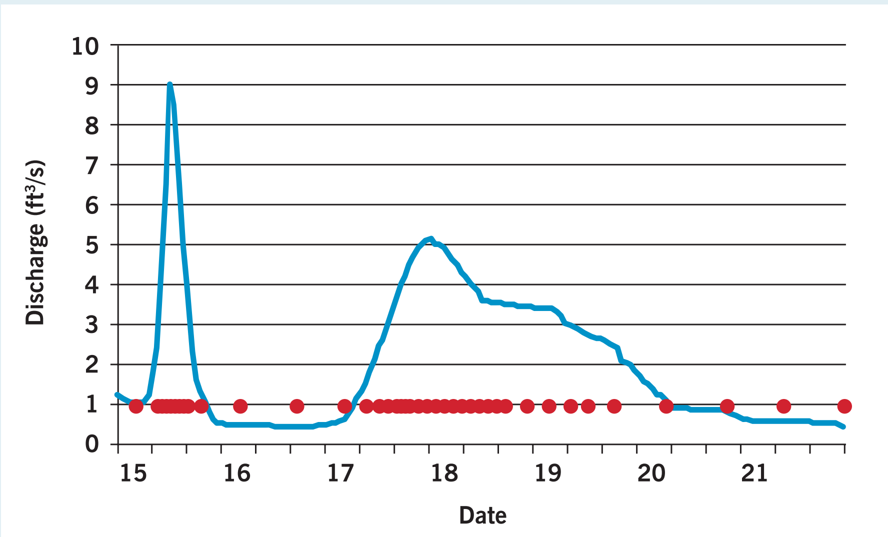
The goal of chemical sampling under this approach is to accurately characterize the relationship between flow and concentration. The monitoring program should be designed

Estimating TP Load from Continuous Discharge and Weekly Flow-Proportional Composite Sampling

Lake Champlain Basin Agricultural Watersheds NNPSMP

Monitoring regime:

- Continuous discharge measurement
- Flow-proportional autosampling (subsample taken every 20,000 ft^3) composited into single container for TP analysis of one sample per week
- Single analysis of flow-proportional composite (equivalent to Event Mean Concentration (EMC))
- Unit load = one week



Station hydrograph for week of July 15 – 21, 2000.
Red dots superimposed on hydrograph represent times of individual subsamples comprising weekly composite sample.

Total weekly discharge = 1,200,000 ft^3

#samples = 60

EMC = 0.58 mg/L

Weekly load = Total weekly discharge * EMC * unit conversion = 19.7 kg

to obtain samples over the entire range of expected flow rates. If seasonal differences in the flow/concentration relationship are likely, the entire range of flows should be sampled in each season. In some cases, separate seasonal flow-concentration regressions may need to be developed and used to estimate seasonal loads. Examples of such flow-concentration regressions are shown in Figure 4. It is clear in Figure 4 that a single regression line would seriously underestimate high concentrations and therefore overall loads. The seasonal regressions shown in the lower panel would yield a better load estimate.

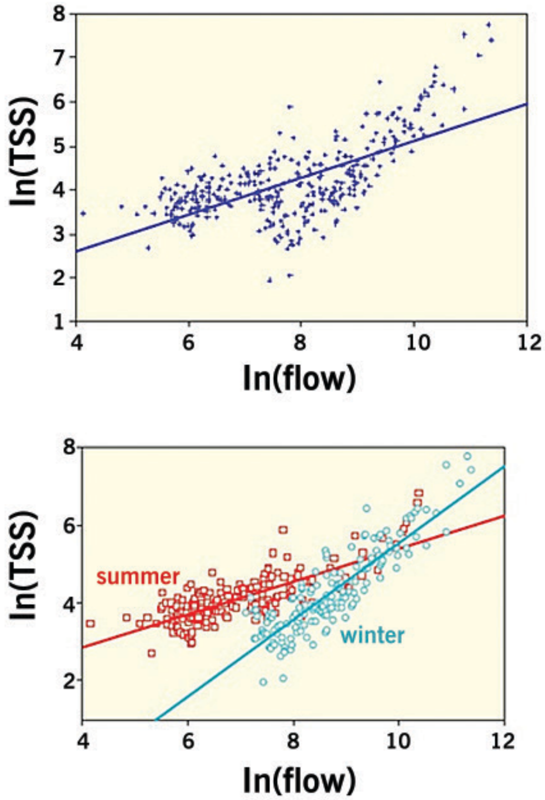


Figure 4. Top panel: plot of regression relationship between log of total suspended solids concentration and log of flow for the 1991 water year dataset from the Maumee River (Ohio). Bottom panel: plot of same data divided into two groups based on time of year. Within each season, the regression model is stronger, has lower error, and provides a more accurate load estimate (Richards 1998).

This approach is especially applicable to situations where continuous flow data already exist, e.g., from an ongoing USGS hydrologic station. Grab samples can be collected as needed and then associated with the appropriate flow observations. Economy is another significant advantage of this approach. After an initial intensive sampling period to develop the regression, it may be possible to maintain the regression model with ~20 samples a year for concentration, focusing on high-flow or critical season events. Software exists to calculate and manage this approach, e.g. ***FLUX*** (Walker 1999). FLUX is an interactive program designed for use in estimating the loadings of nutrients or other water quality components passing a tributary sampling station over a given period of time. Data requirements include (a) grab-sample nutrient concentrations, typically measured at a weekly to monthly frequency for a period of at least 1 year, (b) corresponding flow measurements (instantaneous or daily mean values), and (c) a complete flow record (mean daily flows) for the period of interest. Using six calculation techniques, FLUX maps the flow/concentration relationship developed from the sample record onto the entire flow record to calculate total mass discharge and associated error statistics. An option to stratify the data into groups based upon flow, date, and/or season is also included. The USGS program ***LOADEST*** is also available and is widely used to estimate loads using the regression approach. Another USGS load estimation calculation tool—***FLUXMASTER***—has been used in the SPARROW (SPAtially Referenced Regressions On Watershed attributes) watershed modeling technique to compute unbiased detrended estimates of long-term mean flux, and to provide an estimate of the associated standard error (Schwarz et al. 20006). The two programs include the same statistical model and in most cases will give the same result.

There are a few potential disadvantages to this approach. First, as mentioned earlier, potential changes or trends in the concentration-flow relationship—sometimes a goal of watershed projects—must be tracked and if the relationship changes, a new regression model must be constructed. Second, the monitoring program must be managed to effectively capture the entire range of flows/conditions that occur; the use of data from fixed-interval time-based sampling is not appropriate for this purpose because of bias toward low flow conditions.

Ratio Estimators

The concept of ratio estimators is a powerful statistical tool for estimating pollutant load from continuous flow data and intermittent concentration data. Ratio estimators assume that there is a positive linear relationship between concentration and flow that passes through the origin. On days when chemistry samples are taken, the daily load is calculated as the product of grab-sampled concentration and mean daily flow, and the mean of these loads over the year 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 (e.g., for distinct seasons), 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.

The basic approach to determining sampling frequency assumes a normal distribution for concentration and random sampling. Several formulas are available to calculate the number of samples (random or within strata) required to obtain a load estimate of acceptable accuracy based on known variance of the system (see ***Chapter 2 (Developing a Monitoring Plan) of Monitoring Guidance for Determining the Effectiveness of Nonpoint Source Controls***). Stratification may improve the precision and accuracy of the load estimate by allocating more of the sampling effort to the aspects which are of greatest interest or which are most difficult to characterize because of great variability such as high flow seasons

The Beale Ratio Estimator is one common technique; computer programs are available to implement this calculation method. An example of such a tool is available in *Richards 1998*.

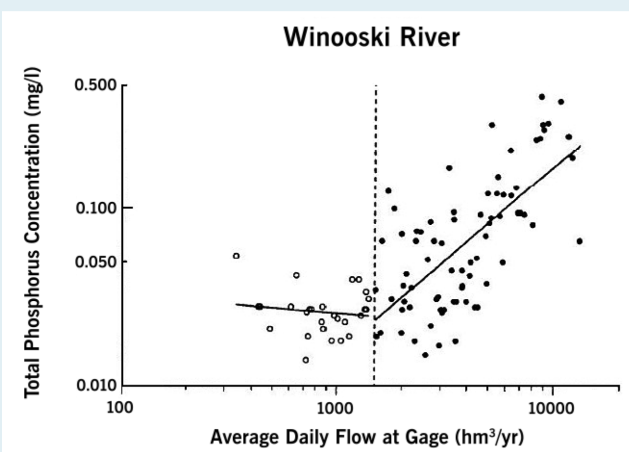
Estimating Phosphorus Load for Lake Champlain Using Flow vs. Concentration Regression

Monitoring regime:

- Tributary flows to Lake Champlain measured continuously at 31 sites, 1990–1992.
- Sampling conducted on 31 tributaries, concentrated on obtaining as high a proportion possible of samples during high flow conditions; samples were also obtained under low and moderate flow conditions. Samples were analyzed for total P, dissolved P, and chloride.

Analysis:

- The total number of samples per stream ranged from 84–107 for the Vermont/Quebec tributaries, and from 36–115 for the New York sites. The percentage of samples taken during high flow conditions ranged from 23–36% at the Vermont/Quebec stations, and from 13–29% in New York. The percentage of all high flow days that were sampled ranged from 25–44% in Vermont and Quebec, and from 12–32% in New York.
- The average daily flow records obtained for each gage station and the tributary sampling data were used in the FLUX program (Walker 1987, 1990) to develop mean load estimates for chloride, total P, and dissolved P for each tributary.
- Examination of concentration vs. flow plots indicated that most of the tributaries had significant relationships between chloride or total phosphorus concentration and daily flow. The concentration vs. flow relationships were stratified with respect to flow and/or by season, to improve the precision of the loading estimates
- A log-log regression relationship was developed between concentration and flow within each flow stratum, and applied with a correction for bias to each daily flow value to produce an estimate of the mean loading rate for the period. Error estimates for the mean loading values were obtained by the FLUX program.



Example of Total Phosphorus concentration vs. Discharge relationship for the Winooski River (VT), March 199–April 1992. Regression lines are shown for the two flow intervals used in the FLUX program load estimation procedure.

Total P load estimates for selected Lake Champlain tributaries, March 1990 – February 1992. C.V. is the coefficient of variation for load and concentration estimates.

Tributary	Mean Discharge (106 m^3/yr)	Total Phosphorus Mean Conc. (mg/L)	Total Phosphorus Mean Load (mt/yr)	Total Phosphorus C.V.
Winooski	2,003	12.3	24,615	0.029
Missisquoi	1,534	6.5	9,899	0.024
Otter	1,427	9.7	13,786	0.016
Lamoille	1,423	8.8	12,515	0.032
Poultney	371	10.5	3,890	0.019
Pike	347	11.9	4,128	0.022

Source: VT DEC and NYS DEC. 1997.

Comparison of Load Estimation Approaches

Although strongly driven by available resources, the monitoring program design (that should have included consideration of load estimation issues from the beginning), and the natural system itself, the choice of load estimation approach can make an enormous difference in the resulting load estimate.

Consider an example drawn from a dataset for the Maumee River (Ohio), water year 1991, consisting of daily TSS concentrations and continuous discharge data (Richards 1998). Different load estimation methods were applied to this dataset and to data taken from seven subsamples representing different weekly sampling schedules (e.g., Monday samples only), the results of which are summarized in Table 1 and Figure 5. The numeric integration estimate based on all daily data was assumed to be the true load for this exercise. The percent error is calculated as 100*(estimate-trueload)/trueload.

Table 1. Suspended solids load estimations from the Maumee River, water year 1991, computed on different subsample datasets and by different load estimation methods (Richards 1998).

Sample	Method	n	Suspended Solids Load (1000s mt/yr)	95% confidence Interval (1000s mt/yr)	% error
All data	**Numeric Integration**	**347**	**2386**		**0**
	Beale Ratio Est.	347	2365	± 415	-1
	Regression	347	1043		-56
	Seasonal Regression	347	1902		-20
Weekly samples on Sundays	Beale Ratio Est.	48	1910	± 1036	-20
	Regression	48	1232		-48
Weekly samples on Mondays	Beale Ratio Est.	53	2544	± 2357	7
	Regression	53	1035		-57
Weekly samples on Tuesdays	Beale Ratio Est.	50	2366	± 1815	-1
	Regression	50	1024		-57
Weekly samples on Wednesdays	Beale Ratio Est.	50	1912	± 1112	-20
	Regression	50	1186		-50
Weekly samples on Thursdays	Beale Ratio Est.	49	1079	± 235	-55
	Regression	49	991		-58
Weekly samples on Fridays	Beale Ratio Est.	48	864	± 173	-64
	Regression	48	867		-64
Weekly samples on Saturdays	Beale Ratio Est.	49	1111	± 301	-53
	Regression	49	1026		-57

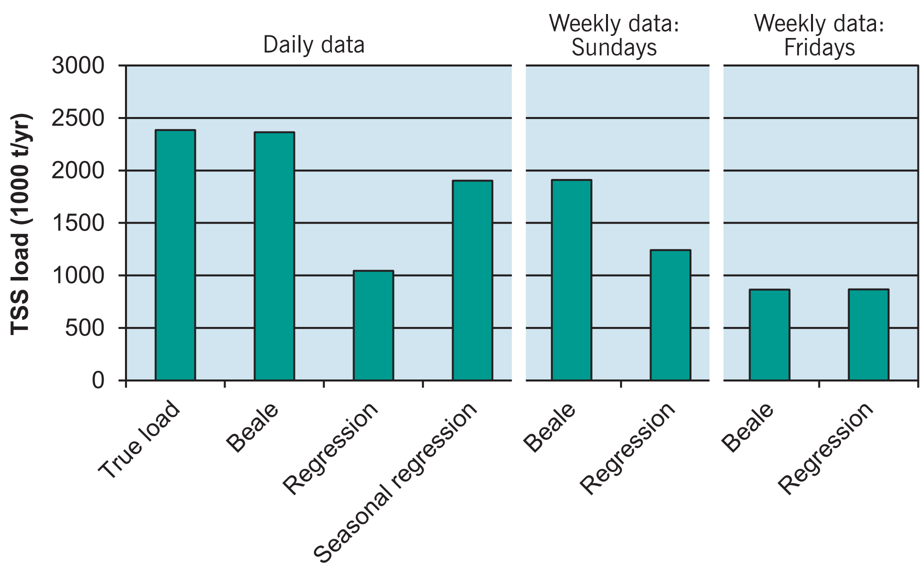


Figure 5. Plots comparing load estimates from Maumee River water year 1991 computed on different subsample datasets and by different load estimation methods. "True load" is represented by numerical integration of daily data.

This example clearly shows that different methods of load estimation applied to different datasets can result in substantially different estimates of pollutant load. Most of the scenarios tested here underestimated suspended solids load, probably because the samples included missed high flow/TSS events and/or the estimation methods were biased toward low flow conditions. In most cases, the Beale Ratio Estimator gave a load estimate closer to the true load than did the regression method. For the full daily dataset, the single flow-concentration regression over the entire year appeared to seriously underestimate suspended solids load; separating the data into summer and winter seasons improved the fit and the accuracy of the load estimate (see Figure 4).

Some scenarios using weekly suspended solids samples for the load estimation gave annual load estimates reasonably close to the "true load," but estimates using weekly concentrations usually underestimated the annual load. This resulted partly from the lower resolution of weekly vs. daily samples, but most of the difference was because in this particular year, suspended solids load was dominated by a single 4-day event that occurred from a Sunday through Wednesday. Samples from Sunday–Wednesday captured at least part of this extreme event, whereas samples from Thursday–Saturday did not. These results show that it was possible to obtain a fairly reasonable load estimate (e.g., within ±20% of the true load) from weekly concentration data, but at the same time, point out the risk of using fixed-interval concentration data in a NPS load estimation.

Load Duration Curves

Once good pollutant load estimates have been derived through monitoring, annual or seasonal loads can be compared over time, between watersheds or tributaries, or against some load objective, as in a TMDL allocation. One particularly useful application of load estimate data is that of the *load duration curve*.

Simply stated, a duration curve is a graph representing the percentage of time during which the value of a given parameter (e.g., flow, concentration, or load) is equaled or exceeded. A load duration curve is therefore a cumulative frequency plot of mean daily flows, concentrations, or daily loads over a period of record, with values plotted from their highest value to lowest without regard to chronological order. For each flow, concentration, or load value the curve displays the corresponding percent of time (0 to 100) that the value was met or exceeded over the specified time—the flow, concentration, or load duration interval. Extremely high values are rarely exceeded and have low duration interval values; very low values are often exceeded and have high duration interval values.

The process of using load duration curves generally begins with the development of a flow duration curve, using existing historical flow data (e.g., from a USGS gage), typically using mean daily discharge values. A basic flow duration curve runs from high to low along the x-axis, as illustrated in Figure 6. The x-axis represents the duration or percent of time, as in a cumulative frequency distribution. The y-axis represents the flow value (e.g., ft^3/sec (cfs)) associated with that percent of time. Figure 6 illustrates that the highest observed flow for the period of record was about 5,400 cfs, while the lowest flow was 6 cfs. The median flow—the flow exceeded 50 percent of the time—was about 200 cfs.

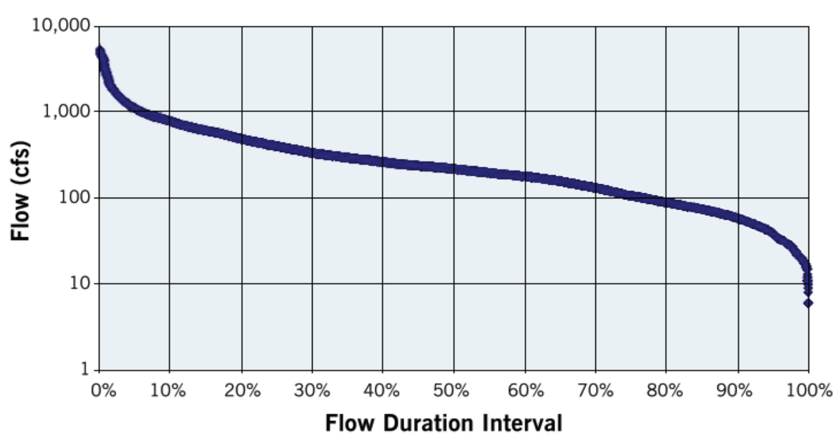


Figure 6. Flow duration curve for the Sevier River near Gunnison, UT, covering the period January 1977 through September 2002.

In the next step, a load duration curve is created from the flow duration curve by multiplying each of the flow values by the applicable numeric water quality target (usually a water quality criterion) and a unit conversion factor, then plotting the results as for the flow duration curve. The x-axis remains as the flow duration interval, and the y-axis depicts the load rather than the flow. This curve represents the allowable load (e.g., the TMDL) at each flow condition over the full range of observed flow. An example is shown in Figure 7 for the same site as shown in the flow duration curve, using a target of 0.05 mg/L total P. Then, observed P load observations associated with the flow intervals are plotted along the same axes. Points located above the curve represent exceedances of the target load, while those plotting below the curve represent compliance with the target and allowable daily loads.

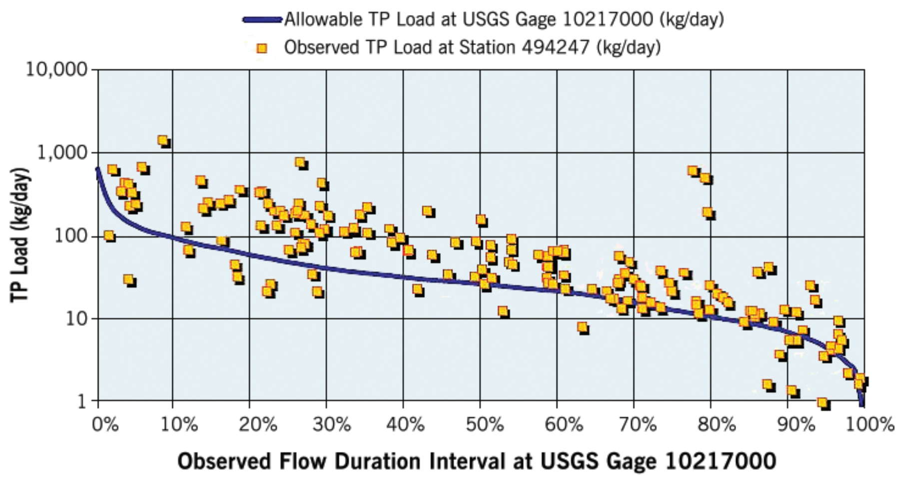


Figure 7. Load duration curve for the Sevier River near Gunnison, UT, January 1977 through September 2002. Blue line represents allowable total P load calculated as the product of each observed flow duration interval and the target total P concentration of 0.05 mg/L. Yellow points represent observed total P loads at the same flow duration intervals.

A key feature of load duration curve analysis is that the pattern of loads—and impairment—can be easily visualized over the full range of flow conditions. Because flow variations usually correspond to seasonal patterns, this feature can address the requirement that TMDLs account for seasonal variations. The pattern of observed loads exceeding target loads can be examined to see if impairments occur only at high flows, only during low flows, or across the entire range of flow conditions. A common way to look at a load duration curve is by dividing it into zones representing, for example: high flows (0–10% flow duration interval), moist conditions (10–40%), mid-range flows (40–60%), dry conditions (60–90%), and low flows (90–100%). Data may also be grouped by season (e.g., spring runoff versus summer base flow). Sometimes, analysis of the load duration curve can provide insight on the source of pollutant loads. Measured loads that plot above the curve during flow duration intervals above 80% (low flow conditions), for example, may suggest point sources that discharge continuously during dry weather. Conversely, measured loads that plot above the curve during flow duration intervals of about 10 to 70% tend to reflect wet weather contributions by NPS such as erosion, washoff, and streambank erosion. Figure 7 illustrates that allowable total P loads in the

Sevier River were exceeded during all flow intervals, suggesting that multiple sources may contribute to the impairment.

It should be noted that a significant weakness of load duration curves is that they usually apply a single target concentration (i.e., the water quality standard) for all flows. For systems with significant NPS inputs this is usually an unreasonable assumption. As a result, the curve will almost always show that most violations occur at higher flows, where load is high due to both high flows and elevated concentrations. We know that concentration increases with flow in these systems, and this is basically a natural phenomenon. Application of different target concentrations to different flow strata may help with this issue.

Finally, note that an individual load duration curve applies only to the point in the stream where the data were collected. A load duration curve developed at a watershed outlet station (e.g., for a TMDL) applies only to loads observed at that point. If significant pollution sources exist upstream, a single load duration analysis at the watershed outlet can underestimate the extent of impairment in upstream segments. For this reason, it is usually wise to develop multiple load duration curves throughout the watershed to address the spatial distribution of impairments. Such an exercise can also be useful in targeting land treatment to critical watershed source areas.

For more detailed discussion of load duration curves, particularly their application to the TMDL process, refer to these sources:

USEPA. 2007. ***An Approach for Using Load Duration Curves in the Development of TMDLs***.

Tetra Tech. (undated) ***Advantages and Disadvantages of Using Load Duration Curves to Estimate Existing and Allowable Loads for the Development of Nutrient TMDLs***.

Summary and Recommendations

Estimation of pollutant load through monitoring is a complex task that requires accurate measurement of both pollutant concentration and water flow, as well as careful calculation, often based on a statistical approach. A NPS monitoring program must be designed for good load estimation at the start. In planning a watershed project, determine whether the project goals require knowledge of load, or if goals can be met using concentration data alone. In many cases, especially when trend detection is the goal, concentration data may be easier to work with and be more accurate than crudely estimated load data.

Good load estimates are usually derived from continuous flow data and intermittent data on pollutant concentration.

Both flow and pollutant concentrations are highly variable. Generally, continuous flow measurement and frequent water quality samples are needed to accurately and reliably capture the true load pattern. Although sampling frequency requirements will vary by the system monitored and the accuracy desired, quarterly concentration observations are generally inadequate, monthly observations will probably not yield reliable load estimates, and even weekly observations may not be satisfactory, especially if very accurate load estimates are required to achieve project objectives.

Water quality sampling for load estimation must capture periods of high flows and pollutant concentrations. Flow-proportional sampling will often provide the most accurate and cost-effective data for load estimation; frequency, timing, and stratification are important considerations for fixed-interval sampling programs.

Computational and statistical techniques appropriate to nonpoint source load estimation include:

- Numeric integration to compute load as the product of flow and concentration over a sequence of observations;
- Regression to estimate un-sampled concentrations based on flow; and
- Ratio estimators to adjust individual unit loads based on flow conditions at the time of sampling.

Different methods of load estimation can result in substantially different estimates of pollutant load. Select the preferred method based on project objectives and monitoring resources.

A load duration curve is a useful approach to compare pollutant load estimates over time, between sites, or against a load-reduction objective.

References

Meals, D.W. and S.A. Dressing. 2005. Monitoring data – exploring your data, the first step, Tech Notes 1, July 2005. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 14 p. Available online at ***https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes***. (Accessed 1-3-2013).

Meals, D.W. and S.A. Dressing. 2008. Surface water flow measurement for water quality monitoring projects, Tech Notes 3, March 2008. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 16 p. Available online at ***https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes***. (Accessed 12-31-2012).

Meals, D.W., J. Spooner, S.A. Dressing, and J.B. Harcum. 2011. Statistical analysis for monotonic trends, Tech Notes 6, November 2011. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 23 p. Available online at ***https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes***. (Accessed 1-3-2013).

Richards, R.P. 1998. Estimation of pollutant loads in rivers and streams: a guidance document for NPS programs. U.S. EPA Region VIII Grant X998397-01-0, Water Quality Laboratory, Heidelberg University, Tiffin, OH. ***http://141.139.110.110/sites/default/files/jfuller/images/Load_Est1.pdf*** (Accessed 12-31-2012).

Richards, R.P. and J. Holloway. 1987. Monte Carlo studies of sampling strategies for estimating tributary loads. *Water Resources Research* 23:1939-1948.

Schwarz, G.E., A.B. Hoos, R.B. Alexander, and R.A. Smith. 2006. The SPARROW surface water-quality model: theory, application and user documentation. U.S. Geological Survey Techniques and Methods Book 6, Section B, Chapter 3. ***http://pubs.usgs.gov/tm/2006/tm6b3/*** (Accessed 1-16-2013).

Spooner, J., S.L. Brichford, D.A. Dickey, R.P. Maas, M.D. Smolen, G.J. Ritter and E.G. Flaig. 1988. Determining the sensitivity of the water quality monitoring program in the Taylor Creek-Nubbin Slough, Florida project. Lake and Reservoir Management, 4(2):113–124.

Spooner, J., S.A. Dressing, and D.W. Meals. 2011. Minimum detectable change analysis. Tech Notes 7, December 2011. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 21 p. Available online at ***https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes***. (Accessed 12-31-2012).

Spooner, J., R.P. Maas, M.D. Smolen and C.A. Jamieson. 1987. Increasing the sensitivity of nonpoint source control monitoring programs. p. 243–257. In: Symposium On Monitoring, Modeling, and Mediating Water Quality. American Water Resources Association, Bethesda, Maryland. 710p.

Tetra Tech. White Paper (l): Advantages and disadvantages of using load duration curves to estimate existing and allowable loads for the development of nutrient TMDLs. ***http://rd.tetratech.com/epa/Documents/White%20Paper%20_l_%20Load%20Duration%20Curves.pdf*** (Accessed 1-7-2013).

USEPA. 2007. An approach for using load duration curves in the development of TMDLs. Watershed Branch, Office of Wetlands, Oceans and Watersheds, Washington, DC. ***http://www.epa.gov/owow/tmdl/techsupp.html*** (Accessed 1-7-2013).

VT DEC and NYS DEC. 1997. A phosphorus budget, model, and load reduction strategy for Lake Champlain. Lake Champlain diagnostic-feasibility study, final report, Vermont Department of Environmental Conservation, Waterbury, VT. *www.vtwaterquality.org/lakes/docs/lp_lcdfs-finalreport.pdf* (Accessed 1-7-2013).

Walker, W.W. 1987. Empirical methods for predicting eutrophication in impoundments. Report 4: Applications Manual. Tech. Rep. E-81-9. Prep. for U.S. Army Corps Eng. Waterways Exp. Sta. Vicksburg, MS.

Walker, W.W. 1990. FLUX stream load computations. Version 4.4. Prep. for U.S. Army Corps Eng. Waterways Exp. Sta. Vicksburg, MS.

Walker, W.W. 1999. Simplified Procedures for Eutrophication Assessment and Prediction: User Manual. Prepared for U.S. Army Corps of Engineers, Water Operations Technical Support Program, Instruction Report W-96-2, September 1996 (Updated April 1999). *http://www.wwwalker.net/bathtub/Flux_Profile_Bathtub_DOS_1999.pdf* (Accessed 1-7-2013).