Journal of Lake and Reservoir Management 14(2-3):197-212, 1998

Analysis of Material Loading to Cannonsville Reservoir: Advantages of Event-Based Sampling

Patricia Longabucco and Michael R. Rafferty
New York State Department of Environmental Conservation
Bureau of Watershed Management
Room 398, 50 Wolf Road
Albany, NY 12233-3508

ABSTRACT

Longabucco, P. and M. R. Rafferty. 1998. Analysis of material loading to Cannonsville Reservoir: advantages of event-based sampling. Lake and Reserv. Manage. 14(2-3):197-212.

Eight years of monitoring data are reviewed and analyzed for the West Branch of the Delaware River, the major tributary supplying Cannonsville Reservoir: 3 years are from the early 1980s, the other 5 years include 1991–1996. Phosphorus, nitrogen, and sediment loads were measured at the mouth of the West Branch using an approach of intensified sampling during runoff events and routine sampling during baseflow intervals. Total river loads of nutrients and sediment are reported on a water-year basis; selected analytes are reported on a monthly basis. Additionally, phosphorus inputs are partitioned into nonpoint and point source loads. The point source phosphorus load decreased about 75% during the 1990's study period primarily due to upgrades at the largest municipal wastewater treatment plant in the watershed. Annual nonpoint source phosphorus loads were variable over the study period: dissolved phosphorus ranged from 6,600 $kg \cdot y^{1}$ to 20,800 $kg \cdot y^{1}$, while particulate phosphorus ranged from 7,400 $kg \cdot y^{-1}$ to 115,000 $kg \cdot y^{-1}$. Most of the annual loads of nonpoint source phosphorus, sediment and ammonia were delivered during runoff events. The loads determined through event-based sampling are compared to loads calculated using two other load estimation approaches: an export coefficient technique and fixed frequency sampling combined with the loading software FLUX. The natural variation in nonpoint source loads is examined in the context of setting management goals for improving water quality.

Key Words: load, runoff, event, sampling, nonpoint, phosphorus, nutrients, tributary.

Material loading largely drives the water quality of lakes and reservoirs, and along with hydrologic loading, is a primary forcing function in mass balance models. Loading can be external and internal, although internal sources are usually reflective of earlier external inputs. External loads can derive from either point sources or nonpoint sources. Point sources are discharges, often continual, which have a well-defined point of entry into a water body, such as a pipe carrying effluent from a wastewater treatment facility. Nonpoint sources are diffuse; inputs from them are often episodic and associated with runoff-producing events. Anthropogenic nonpoint sources are a leading concern as these have often been identified as negatively impacting water quality (e.g., Novotny and Chesters 1981) and, thus, the logical target of control efforts when attempting to remediate or prevent problems.

Loading estimates are invaluable in establishing cause and effect relationships for various lake water quality problems, and essential to drive related mass-balance water quality models (Effler and Whitehead 1996). For example, as phosphorus (P) is generally considered to be the primary nutrient controlling eutrophication (e.g., Vollenweider 1968, Hutchinson 1973, Dillon 1975, Carlson 1977), it is a common management approach to quantify the relationship between P loading and trophic state, and establish a target P load consistent with desired water-quality or trophic-state objectives.

The temporal nature of external loading and related lake impacts will dictate the sampling needs of a tributary monitoring program intended to establish cause and effect relationships between inputs and lake water quality. In the case of nonpoint source loads, monitoring over a range of runoff conditions, with particular emphasis on high flow periods, is generally necessary to obtain reliable estimates of fluvial material loading (Dolan et al. 1981, Richards and Holloway 1987, Young et al. 1988). For a given number of concentration samples, loading estimates will usually be of greater precision if sampling efforts are weighted toward high flow periods and runoff events, as these intervals, while usually short in duration, often account for a large fraction of the annual tributary load (Walker 1987). Additionally, as nonpoint source inputs associated with runoff are highly dependent on meteorological and hydrologic conditions, it is desirable to assess the attendant natural variability in material loadings that can occur, particularly from the standpoint of evaluating the extent to which systematic reductions from control efforts may be masked by these variations.

The purpose of this paper is to present the results of a long-term, event-oriented monitoring program that measured fluvial loading of nutrients and sediments to a eutrophic drinking water supply, Cannonsville Reservoir, from its largest tributary, the West Branch of the Delaware River (WBDR). The loading estimates were used to support testing and calibration of a nutrient-phytoplankton model developed for Cannonsville Reservoir (Doerr et al. 1998), and application of the model in a preliminary assessment of management measures for Cannonsville (Owens et al. 1998a). Specifically, we: 1) reviewed nutrient and sediment concentrations; 2) identified trends in these concentrations; 3) presented estimates of river loads; 4) partitioned nonpoint and point source loads of P; 5) identified long-term trends in point source loading of P, and; 6) documented inter-annual variations in nonpoint source loads. The results of the monitoring program are also used to highlight the importance of event-based sampling when quantifying nonpoint source inputs by comparing our load estimates to estimates made using other less data-intensive techniques. The implications of accurate load estimates and inter-annual variability in nonpoint source contributions are then examined in the context of setting target loads reflective of desired water quality.

Site Description

The WBDR is the primary source of water for Cannonsville Reservoir, an upstate drinking water supply for the City of New York. The river is approximately 80 km in length, drains in a southwesterly direction and is located almost entirely within Delaware County, New York (Fig. 1). Long-term average river discharge rate, as measured at the U.S. Geological Survey automated gauge station in Walton, for the 44-year period of record ending September 1996, is $16.4\ m^3 \cdot s^{-1}$ (watershed runoff of $60.1\ cm \cdot yr^{-1}$) (USGS 1997).

The WBDR drains an area of 910 km^2 and comprises about 80% of Cannonsville Reservoir's total drainage basin. Present land use is little changed from that in 1980 which was 70% wooded, 25% agricultural, and 2% urban/industrial, with dairy farming being the predominant agricultural use (Brown et al. 1986). The watershed's current human population is essentially the same as it was in the early 1980s, about 18,000. In 1990, 8,016 people resided in the four incorporated villages located in the watershed and were served by sewers and wastewater treatment plants that discharge to the WBDR. The watershed and its characteristics are described more fully in Brown et al. (1983).

During the study period, there were seven permitted point sources that discharged directly to the WBDR: five are publicly-owned wastewater treatment plants (WWTP); two are industrial discharges. The Walton WWTP, which receives both domestic sewage and food waste from a dairy process facility, is the largest (~1 MGD) and most significant point source in the watershed. All of the WWTPs provide secondary treatment; Walton upgraded to a new system with P removal in 1993.

History

Cannonsville Reservoir began filling in 1963 after damming of the WBDR at Stilesville was completed. Surveys of the reservoir (Schumacher and Wager 1973,

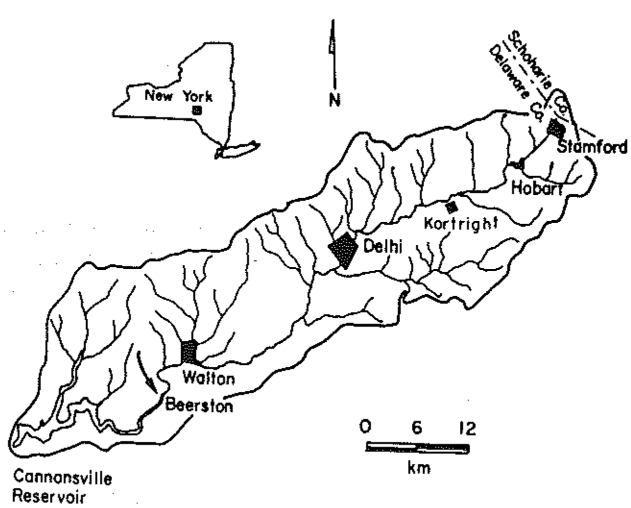

Figure 1.–West Branch of the Delaware River watershed and Cannonsville Reservoir.

USEPA 1974, Wood 1979, Brown et al. 1986, Bader 1993, Effler and Bader 1998) have reported conditions typical of eutrophication since the early 1970s through the present day. Excess loading of nutrients, primarily P, from both point and nonpoint sources in the watershed has been identified as the principal cause of the eutrophic conditions (Wood 1979, Brown et al. 1980, 1983, 1986, Effler and Bader 1998). Internal loading of P appears to be an insignificant contributor to the reservoir's eutrophication (Brown et. al 1986, Effler and Bader 1998). Recent studies (Doerr et al. 1998, Owens et al. 1998a) have identified dissolved P as the most important nutrient in regulating algae growth in the reservoir for much of the year.

Numerous estimates of external P loading to Cannonsville Reservoir have been made using various estimation techniques (Table 1). The estimates from the 1970s and 1980 were based on very limited water quality sampling data from the WBDR and other minor tributaries. Another, more recent, estimate (Bader 1993) used an export coefficient method to determine P loads. All of these analyses considered the WBDR to be the major external source of nutrients to Cannonsville and, while the estimates vary to some degree, all exceed the critical P mass load, based on Vollenweider's (1975) model, of approximately 20,000 kg $\cdot$ y^{1} (Bader 1993) that would be expected to lead to eutrophic conditions in the reservoir.

None of the sampling efforts in the 1970s attempted to resolve the dynamics of P (or other material) concentrations during runoff events, making it difficult to ascertain the reliability of the resulting load estimates. At the time, the importance of characterizing material flux during high flow regimes was perhaps not fully realized. Research has shown that much of the annual loads of nutrients and sediment can be delivered to a receiving water body during extreme hydrologic conditions, i.e., periods of intense rainfall or high stream flows resulting from runoff events (e.g., Alberts et al. 1978, Bloomfield et al. 1984, Brown et al. 1989, Longabucco and Rafferty 1989, Heidtke and Auer 1992). In the Northeast, those runoff events occurring in late winter-early spring particularly influence annual loading as they typically involve combined rainfall/ snowmelt which often produces the highest flows of the year.

Table 1.–Summary of total phosphorus load estimates for Cannonsville Reservoir based on limited tributary water quality data or use of export coefficients.

Reference	Load (kg $\cdot$ 10^{3} $\cdot$ y^{1})
Hydroscience 1975	76 - 117
USEPA 1974	83.7
Bricke 1975	49.2
Goodale 1975	39.7 - 69.7
Wood 1979	26.7 - 32.0
Brown and Rafferty 1980	17.2 - 24.1
Bader 1993	23.9 - 67.7

It was not until 1980 that event sampling became the key element of water quality monitoring in the WBDR. Unlike the earlier estimates (Table 1), loads computed for the March 1980 - September 1982 interval (initially presented in Brown et al. 1983), and the more recent October 1991 - September 1996 interval, were based on a monitoring program that emphasized intensive sampling during runoff periods. The information collected by both of these efforts forms the basis of the analysis presented here.

Methods

Sampling and Data Collection

Precipitation was calculated as the average of monthly totals (NOAA 1980-1982, 1991-1996) recorded at four meteorological stations within the watershed (Walton, Delhi, Kortright and Stamford - Fig. 1). Mean daily and hourly river discharge data (USGS 1981-1983, 1993-1997) were obtained from the U.S. Geological Survey for the Walton gauge station.

The monitoring strategy used during the 1980s (Brown et al. 1983) and the current investigation (1990s) was similar. It was based on a stratified sampling scheme where sub-populations (strata) are separately sampled according to the degree of variability that they exhibit. In the context of hydrological data collection, a minimum of two temporal strata are defined by baseflow, produced by groundwater contributions, and high flow, produced by rainfall and/or snowmelt runoff events (Reckhow et al. 1980). Since there is generally greater variability in concentration and flow rate in the high flow stratum, more frequent measurements will produce a more precise and accurate estimate of the population average (load) (Reckhow et al. 1980). This general approach, the intensification of sample and data collection during the relatively short periods of runoff events, combined with less frequent baseflow sampling, was employed to determine nutrient and sediment losses from the WBDR watershed. A runoff event was defined as any time a rise in river stage of at least 0.5 ft (0.152 m) was measured at the Walton gauge station. Nearly every rainfall or snowmelt event as defined above was sampled. The event was considered to start with the beginning of river rise and was deemed over when concentrations of P and sediment returned to approximately baseflow levels, or when another event began.

The sampling protocol used during the 1980s is described in detail in Brown et al. (1983). Briefly, whenever possible, a baseline sample was collected before significant river rise began. Samples were then collected at 2- to 4-hr intervals on the rise and 6- to 8-hr intervals on the fall, or more frequently during major rainfall/snowmelt events. The stage of the river was monitored by a Stevens Telemark unit installed at the Walton gauge which reported instantaneous river stage via telephone query. Samples were collected at least once a week during dry weather periods.

The sample collection frequency used during the 1990s was refined through examination of data from the 1980s and was dictated by changes in river stage over time. Samples generally were collected systematically at every 0.5-ft change in river stage or at evenly-spaced minimum frequencies within each stratum as defined by stage (Table 2), whichever occurred first. A baseline sample was taken before significant river rise, marking the beginning of the event. River stage was again monitored by telephone through use of the Telemark device installed at the Walton gauge.

Sampling was conducted at Beerston (Fig. 1), located approximately 8 km downstream of the Walton gauge and 0.8 km upstream of Cannonsville Reservoir where State Highway 10 crosses the river. The watershed area increases about 6% between Walton and Beerston. Grab samples were collected during low-flow conditions by wading to approximately mid-channel and immersing a 4-L plastic jug. Under higher flow conditions, a weighted jug was lowered from the road bridge to collect the sample.

Table 2.–Sampling protocol for WBDR at Beerston, 1991-1996.

Stage at USGS Gauge in Walton (ft)	(m)	Frequency of Samples
≤3.00	(≤0.91)	$1 \cdot wk^{-1}$
3.01 - 3.50	(0.92 - 1.07)	$2 \cdot wk^{-1}$
3.51 - 4.00	(1.07 - 1.22)	$3 \cdot wk^{-1}$
4.01 - 4.50	(1.22 - 1.38)	$4 \cdot wk^{-1}$
4.51 - 6.00	(1.38 - 1.83)	$1 \cdot d^{-1}$
6.01 - 7.50	(1.83 - 2.29)	$2 \cdot d^{-1}$
7.51 - 8.50	(2.29 - 2.59)	$3 \cdot d^{-1}$
8.51 - 9.00	(2.59 - 2.74)	$4 \cdot d^{-1}$
9.01 - 9.50	(2.75 - 2.90)	$5 \cdot d^{-1}$
9.51 - 10.00	(2.90 - 3.05)	$6 \cdot d^{-1}$
10.01 - 11.00	(3.05 - 3.35)	$8 \cdot d^{-1}$
11.01 - 13.00	(3.36 - 3.97)	$12 \cdot d^{-1}$
>13.01	(>3.97)	$24 \cdot d^{-1}$

The majority of samples was collected during events. Yearly totals averaged 214 and ranged from 94 to 381, with the more hydrologically active years generally having more samples. In the 30-month study of the 1980s, 31 events were monitored, while in the 60-month study of the 1990s, 70 events were monitored. Large events were relatively infrequent; in fact, during each study period, nearly 70% of the monitored events produced only ≤2.0 cm of runoff, while approximately 10% produced ≥3.5 cm.

Sample analyses for both the 1980's and 1990's investigations were performed at the New York State Department of Health laboratory using identical or updated methodologies. Samples were routinely analyzed for total phosphorus (TP), total dissolved phosphorus (TDP), soluble reactive phosphorus (SRP), nitrate plus nitrite (NO_X), total ammonia ($T\text{-}NH_3$), and total suspended sediment (TSS) (USEPA 1983, APHA 1992). Filtration of water through a 0.45-µm membrane filter prior to analysis was used to define the dissolved P fraction. Particulate P (PP) was computed as the difference of TP and TDP. In 1995 and 1996, analyses for soluble and particulate organic carbon, chloride, and chlorophyll *a* were added to support testing of the nutrient-phytoplankton model for Cannonsville (Doerr et al. 1998) and an investigation of allochthonous sources of trihalomethane (THM) precursors (Stepczuk et al. 1998); loads of these analytes are not considered here.

Computation of Loads from Point and Nonpoint Sources

River flow at Beerston was estimated by increasing mean daily and hourly flow values for the WBDR at Walton by a factor of 1.06 to account for the additional watershed contributing area between Walton and Beerston. Total river loads were computed as the product of flow rate and concentration of samples collected at Beerston. For non-event periods, loads were calculated on a daily basis using the mean daily flow and a sample concentration. For events, hourly loads were calculated using hourly flows and event sample concentrations. These hourly loads were then summed to produce both daily loads and total event loads. Annual loads were calculated as the sum of the daily estimates and, along with river flow, are expressed by water-year (October through September), rather than calendar year, throughout this paper.

Concentrations used to compute daily and hourly loads were either from actual samples or estimated. If a sample was collected on a given day or hour, that sample's concentration was used to compute the daily or hourly load. For intervening times without samples, simple linear interpolation between the preceding and next sample concentrations was used to estimate concentrations for those times. When a baseline sample was missing, the concentration from the last collected non-event sample was assumed to be representative until the river began to rise. Also, when a sample was missed at the end of an event, the concentration from the next non-event sample was used for the event end point. The basic load calculation method is illustrated here for an event in January 1996 (Fig. 2).

Nonpoint source P loads were estimated by subtracting computed point source loads from total river loads calculated at Beerston. Phosphorus load estimates for the Walton WWTP were based on effluent sampling (seven consecutive 24-hr composite samples collected quarterly) by the New York State Department of Environmental Conservation, and sample (usually weekly or twice-weekly collection of a 24-hr composite) and daily flow data from the treatment plant's self-monitoring reports. Loads for the other point sources were provided by the New York City Department of Environmental Protection (Cutietta-Olson 1997).

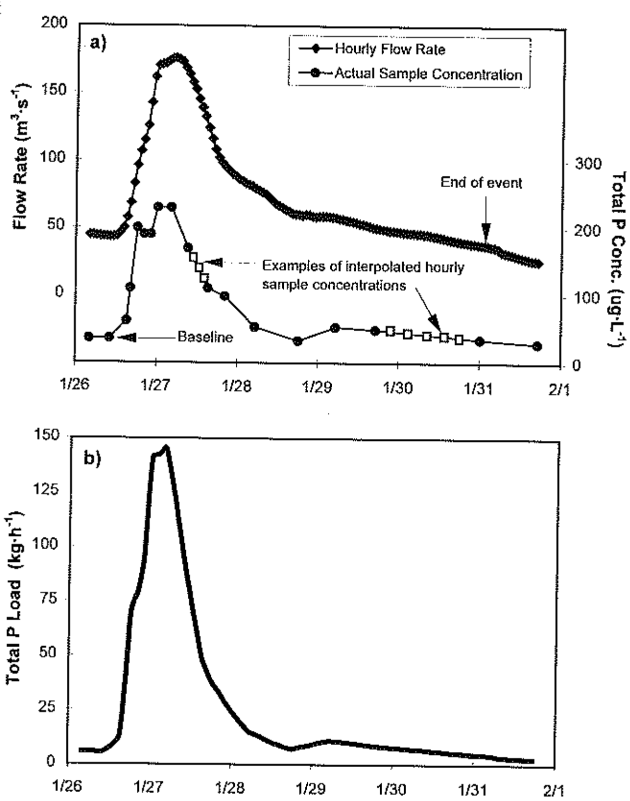

Figure 2.–Load estimation technique illustrated for January 26-31, 1996, event: a) hourly flow rates, actual and interpolated TP concentrations and, b) product of each hourly flow rate and an actual or interpolated concentration.

Analysis of a limited number of Walton WWTP effluent samples for TDP showed that the flow-weighted average dissolved P fraction was 62% (n = 29, range = 27% - 100%) during the 1980's study, and 59% (n = 49, range = 11% - 100%) in the 1990s after the plant was upgraded. Eight samples collected from the Stamford WWTP in the 1980s averaged 92% dissolved (range = 80% - 98%) which is more typical of domestic sewage (Black and Veatch 1971). Based on the above sample results, 60% of Walton's P load was assumed to be dissolved, while P loads from the other WWTPs in the watershed were assumed to be 92% dissolved.

Load Estimation Comparisons

The event-based monitoring efforts of our program were resource-intensive, requiring considerable effort and funding to achieve the level of sampling that was deemed necessary to adequately measure loads from the WBDR. Alternate methods which can reduce time in the field and costs of analyses, without substantially compromising accuracy, are desirable. However, it is difficult to determine how well a method has estimated loads because, generally, the "true" load is not known. The actual load can never be known exactly, but the data collected through our program are so extensive compared to typical (e.g., fixed-frequency) sampling programs that we have used our loads as a basis to evaluate the estimates made prior to 1980 (Table 1) as well as estimates made using two other methods, less labor-intensive than our own.

The first of these two is a pollutant export coefficient approach, which involves no direct water quality sampling. Export coefficients for different types of land use are selected from literature compilations of values (e.g., Reckhow et al. 1980) or are watershed-specific values. The coefficients are typically derived from published results of monitoring studies, which measured loads of one or more materials from a particular land use or combination of land uses over short time periods. Applying the coefficients (e.g., units of $kg \cdot ha^{-1}$) to the area of each type of land use in a watershed gives an estimate of the annual amount of material delivered to the outlet of the watershed. Bader (1993) used this technique to estimate an annual nonpoint source TP load for the entire Cannonsville Reservoir drainage basin. We compare his estimate to the annual loads measured by our event-based monitoring program and analyze the differences.

Techniques that integrate the material concentrations of grab samples collected intermittently or on a fixed interval with continuous (i.e., gauged) flow data to produce unit loads may be used in situations where intensive collection of event-based samples is not feasible. These techniques are more attractive than the export coefficient method because they utilize some actual water quality data from the watershed in question. Typically, regression or ratio estimation (Beale 1962, Tin 1965) is used to estimate concentrations for periods when samples are not available (e.g., Dolan et al. 1981, Young et al. 1988). For these techniques to produce accurate results, though, a strong relationship between flow and concentration is necessary to predict concentrations for periods of no samples from flow data. The load estimation program FLUX, developed by Walker (1987), utilizes such flow-concentration relationships. FLUX is an interactive, user-friendly software that supports estimates of loads from intermittent, routine or event sampling over a period for which a continuous flow record is available (usually as average daily values). We use FLUX in our second comparison to contrast the results of our event-based program with estimates developed from a fixed-frequency sampling protocol typical of many less resource-intensive monitoring programs. Samples collected under a fixed-frequency protocol may include both high- and low-flow samples, but efforts to weight sampling towards high-flow periods are usually not made.

For this comparison, a subset of samples was selected from our database that mimicked a weekly sampling protocol. Only samples taken on a specific day of the week, or the one closest to that day, were included. If more than one sample had been collected on the designated day, only the first sample was used. Two data sets were created that simulated a weekly (i.e., 52 samples $\cdot$ yr^{-1}) sampling program: one for October 1, 1980 - September 30, 1982, and one for October 1, 1991 - September 30, 1996. These sample concentrations were utilized with adjusted (+ 6%) mean daily river flows from the Walton gauge to estimate loads at Beerston with FLUX. The software has several options to improve accuracy and reduce bias based on the number and type of samples (Walker 1987, Effler and Whitehead 1996). For the loading estimates presented here, an option was used that stratifies the samples into groups based on flow. This often has the effect of increasing accuracy and reducing potential bias (Walker 1987). Daily loads of PP, TDP, SRP, TSS, T-NH_3, and NO_X were generated by FLUX and compared to our "true" daily loads determined through event-based sampling. Also, daily loads were summed by water-year and compared on an annual basis.

Results and Discussion

Precipitation and Runoff

Substantial inter-annual differences are evident in precipitation and runoff volume via the WBDR over the 8-year study period (Fig. 3). Annual precipitation (Fig. 3a) and runoff volume (Fig. 3b) were below average for 5 of the 8 years; in fact, runoff volume in 1995 was the third lowest for the 44-year period of record. In contrast, runoff volume in 1996 was the second highest on record, and a new record peak discharge at Walton was set during the January 19-20, 1996, rainfall/snowmelt event, which has been classified as at least a 70-year storm (Lumia 1998).

Seasonal variations in runoff, driven by variations in precipitation and the timing of winter thaws, were observed among the study years as illustrated for 1994, 1995 and 1996 (Fig. 4). For example, while the timing of runoff was quite similar in 1994 (Fig. 4a) and 1996 (Fig. 4c), the magnitude during event periods (identifiable by spikes in runoff volume in Fig. 4) was generally much greater in 1996. Events in 1995 (Fig. 4b) were modest, showing only small increases in river flow. Years 1994 and 1996 had relatively wet summers with 15% and 19%, respectively, of the annual runoff volume occurring from June through September,

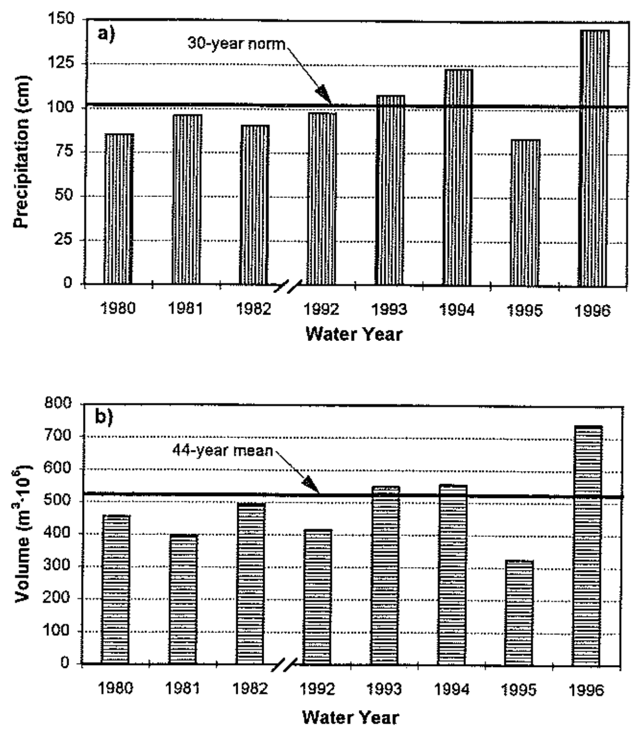

Figure 3.–a) Annual precipitation in the WBDR watershed, and b) annual runoff volume of the WBDR during study period.

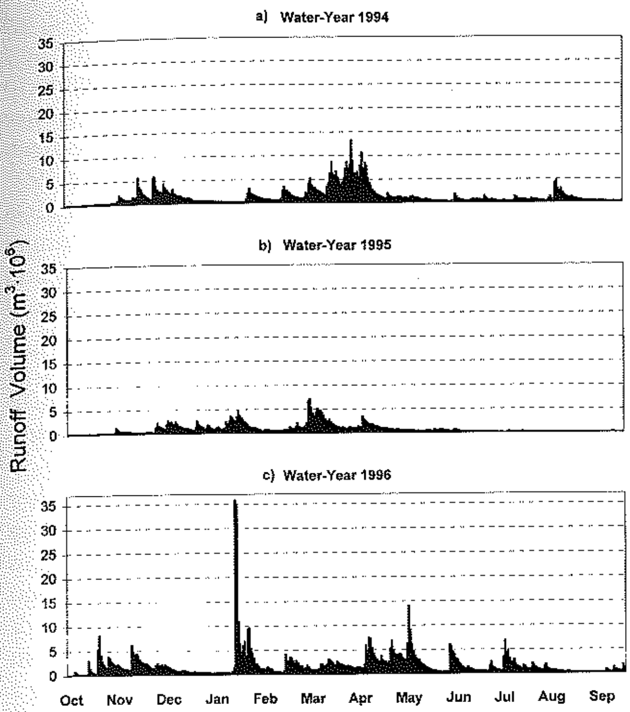

Figure 4.–Daily runoff volumes of the WBDR at Walton for 3 years of the study period: a) water year 1994; b) water year 1995; c) water year 1996.

whereas in 1995 less than 3% occurred during that time. All 3 years were similar in that about 20% of the annual runoff occurred from October through December. As will be shown later, the amount and timing of runoff can have a considerable effect on the loading of nonpoint source pollutants from the watershed.

Nutrient and Sediment Concentrations

Substantial variability in nutrient and sediment concentrations was observed over the course of the two study periods (Table 3). Volume-weighted event means were higher than baseflow means for all analytes except NO_X. Wider ranges in concentration were observed during events than in baseflow for PP, TSS, T-NH_3 in both study periods, and for TDP and SRP in the 1980s. Event and baseflow means and ranges of the 1980s were similar to those of the 1990s for most analytes, with the following exceptions. Wider ranges in PP and TSS event concentrations were observed in the 1990s, but these are directly attributable to the January 19-20, 1996, event, which produced the highest observed concentrations for both of these analytes. Also, concentrations of TDP, SRP and T-NH_3 were often much higher during major runoff events in the 1980s as compared to the 1990s. The reason for these higher concentrations is unknown at this time; however, it is unlikely that analytical differences were the cause, as the same laboratory and methods were used in both study periods.

The proportions of PP, TDP and SRP in WBDR samples collected at Beerston varied considerably. The amount of P that was dissolved ranged from as little as 3% to nearly 100%. The fraction of TDP that was SRP varied from 0 to 100%, but averaged 72%. Particulate P concentration had a strong, direct correlation with TSS concentration ($r=0.96$, $n=1708$), as did SRP with TDP ($r=0.98$, $n=1428$). The lowest concentrations of P were typically observed during periods of high baseflow (~14 $m^3 \cdot s^{-1}$ or more), while the highest concentrations usually occurred during periods of surface runoff. The

Table 3.–Volume-weighted mean concentrations and concentration ranges of nutrients and sediment in the WBDR at Beerston for the two study periods during event and baseflow periods.

Analyte	1980s				1990s			
	Event		Baseflow		Event		Baseflow	
	$\bar{X}$	Range	$\bar{X}$	Range	$\bar{X}$	Range	$\bar{X}$	Range
PP[a]	150	2 - 1,060	6	1 - 50	140	2 - 1,920	8	1 - 76
TDP[a]	82	16 - 420	38	14 - 260	36	5 - 130	31	4 - 440
SRP[a]	81	11 - 390	34	9 - 240	26	1 - 98	15	0 - 400
NO_X[a]	590	60 - 1,300	680	160 - 1,500	730	200 - 1,450	760	10 - 1,430
T-NH_3[a]	210	5 - 650	59	5 - 370	48	3 - 230	19	2 - 170
TSS[b]	97.1	1.3 - 750	5.4	0.2 - 25.9	93.4	1.4 - 1,350	5.3	0.2 - 29.5

[a]Concentrations in $\mu g \cdot L^{-1}$.
[b]Concentrations in $mg \cdot L^{-1}$.

largest and most rapid changes in concentration occurred when river flow rose during runoff events as illustrated by the increase of TP from 36 to 200 $\mu g \cdot L^{-1}$ in 7 hr of the January 27, 1996 event (Fig. 2a). During dry-weather periods of low baseflow (~10 $m^3 \cdot s^{-1}$ or less), P concentrations were observed to increase over time as river discharge decreased. Most of the P was in the dissolved form as TSS, and correspondingly PP, concentrations were usually quite low. This inverse relationship of concentration to flow is characteristic of a constant loading source (e.g., Manczak and Florczyk 1971), presumably in this case the Walton WWTP. The effect of this facility on instream concentrations becomes more discernible as ambient flow decreases.

Nutrient and Sediment Loads

Annual loads delivered to Cannonsville Reservoir by the WBDR varied over the study period (Table 4) with the greatest difference for most analytes observed between the years of 1995 and 1996. Years with greater runoff volumes generally produced higher loads. However, there were inconsistencies. For example, 1993 and 1994 had nearly the same annual runoff, but loads in 1994 were about one-half to two-thirds those of 1993 for most analytes, particularly P forms and TSS. Runoff volume in 1981 was 20% less than in 1982, yet 1981's P and TSS loads were as much as 60% greater.

The widest fluctuations in annual P inputs to Cannonsville Reservoir during the study period have resulted from variations in nonpoint source loading (Table 4). Annual nonpoint PP loading in 1996 (115,000 kg) was 15 times greater than that in 1995 (7,400 kg). Less dramatic, but still major, inter-annual variations were observed in nonpoint loads of TDP and SRP. Three-fold differences were observed in loads of these analytes: TDP ranged from 6,600 kg in 1995 to 20,800 kg in 1996, and SRP ranged from 5,000 kg in 1995 to 16,300 kg in 1981. Most of the annual WBDR loads of PP, TDP and SRP, for all years, were nonpoint-derived.

Nonpoint PP loads were more variable than either nonpoint TDP or SRP loads (Table 4). However, when the PP load from the unusually wet year of 1996 is eliminated from the analysis, the coefficients of variation (CV) for all three forms of P were more similar. Loads of TSS, which is almost completely derived from nonpoint sources, were the most variable, exhibiting a 20-fold difference between 1995 (3,800 metric tons) and 1996 (76,200 metric tons). The vast majority of nitrogen delivered to the reservoir, and available to support phytoplankton growth, was supplied as NO_X. Annual loads of NO_X typically exceeded T-NH_3 loads by

Table 4.–Nutrient and sediment loads ($kg \cdot 10^3$) and runoff volumes ($m^3 \cdot 10^6$) of the WBDR at Beerston by water-year. Phosphorus nonpoint source loads in parentheses.

WY	Runoff	Total Load								
		PP		TDP		SRP		TSS	T-NH_3	NO_X
1980[a]	272	20.4	(18.2)	16.8	(12.5)	13.5	(10.5)	14,300	44.2	145
1981	419	34.5	(32.0)	25.2	(19.8)	20.6	(16.3)	17,700	31.6	305
1982	522	21.2	(19.3)	23.0	(18.4)	NS		13,700	NS	NS
1992	440	18.1	(13.2)	19.0	(9.8)	NS		8,970	10.1	373
1993	581	34.9	(33.3)	22.7	(18.2)	16.9	(13.7)	23,200	23.6	399
1994	587	17.6	(16.6)	15.7	(12.0)	11.6	(9.0)	11,100	15.2	420
1995	342	8.2	(7.4)	9.6	(6.6)	7.1	(5.0)	3,800	9.0	245
1996	785	116.	(115.)	23.6	(20.8)	16.0	(13.9)	76,200	30.6	616
CV[b]			1.10 [.511][c]		.368		.391	1.12 [.521]	.501	.322 [.206]

[a] Runoff volume and loads are for sampling period of March-September 1980.

[b] Coefficients of variation determined using only loads based on entire year of sampling (i.e., 1980 excluded); those for phosphorus are for nonpoint source loads.

[c] Coefficients of variation with 1996 loads excluded for comparison.

NS = Not sampled.

at least a factor of 10. Loads of T-NH_3 were more variable than NO_X, however, ranging from 9,000 kg in 1995 to more than 44,200 kg in 1980 (Table 4). An explanation for the apparently larger T-NH_3 loads of the 1980s (and higher concentrations during events; see Table 3) is lacking since resolution of point versus nonpoint derivation of nitrogen was not supported by the monitoring program.

Annual point source loads of TP [total load of PP+TDP minus nonpoint load of PP+TDP in Table 4] during the 1980's study period ranged from 6,500 kg (1982) to 10,700 kg (1980); loads in the 1990s ranged from 3,400 kg (1996) to 14,100 kg (1992). Records from the Walton WWTP, which in the 1980s and early 1990s accounted for 70 to 80% of the entire point source load to the WBDR, show that discharge from the facility has increased somewhat over the period 1980 to 1992. This fact, along with unusually high effluent P concentrations during part of 1992, explain the larger point source load seen in 1992.

After 1992, point source loads decreased significantly. The 1993 TDP and PP loads were 52% and 68% less, respectively, than 1992 levels. This reduction is largely reflective of the upgraded treatment system constructed at Walton in late 1992, and further improved in 1993, which resulted in much lower effluent P concentrations. Additional operational refinements at Walton have produced even greater reductions since then; the entire point source TP load to the WBDR reached a low of 3,400 kg in 1996, only 32% of which was contributed by the Walton WWTP. Loading of TDP from point sources in the watershed decreased by 70% during the 1992 - 1996 interval.

Phosphorus from the WBDR tended to be loaded to Cannonsville Reservoir in pulses (Fig. 5) coinciding with runoff events. The largest inputs of both PP and TDP were delivered in late winter-early spring, months typically dominated by rainfall/snowmelt events, followed by lesser amounts delivered during rainy periods in the fall and early winter. Cannonsville Reservoir is usually at its lowest level by September or October due to drawdown and reduced summer inflows (see Owens et al. 1998b). Thus, fall and spring runoff recharges the reservoir in terms of both water volume and nutrient supply. Phosphorus loading was lowest during the summer period (June through September) (Fig. 5).

Amounts delivered during runoff events constituted the bulk of the annual load of most analytes (Table 5). However, the number of days on which events occurred was small, often 50 per year or fewer. Similarly, the portion of river runoff volume conveyed during these events was between about a quarter and half of the annual total. Event loads of NO_X, an analyte that exhibited little variation in instream concentration during the course of an event, were approximately proportional to event runoff volume, as evidenced by roughly equivalent percent event volumes and loads (Table 5). However, for PP, TDP, T-NH_3, and TSS, the percent annual load delivered during events was substantially higher than the percent annual runoff volume associated with those events because concentrations of these analytes increased considerably during event periods (see Table 3).

The timing and duration of events can have a significant effect on loads as seen when comparing 1993 to 1994 (Table 5). A greater portion of runoff volume and loading occurred during events in 1993 than did in 1994, which may explain, in part, why the annual loads were higher in 1993 than in 1994 even

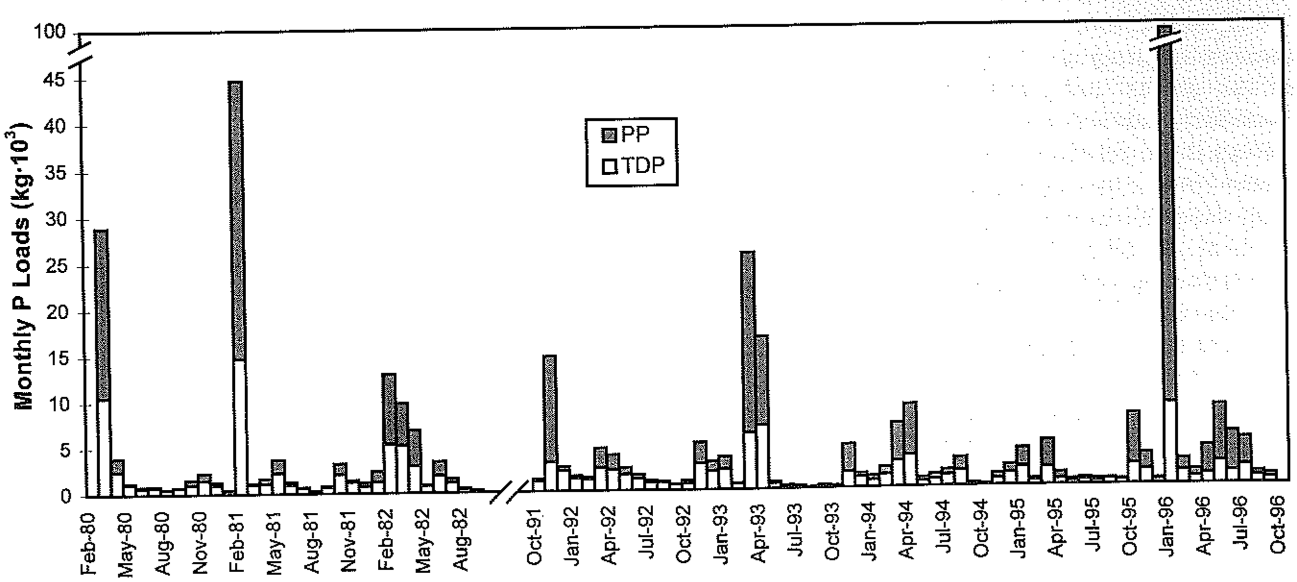

Figure 5.–Monthly P loads from the WBDR during study period.

Table 5.–Number of days on which events occurred, and percentages of annual runoff volume and annual load occurring during those events.

WY	Days (#)	Volume	PP[a]	TDP[a]	TSS	$T\text{-}NH_3$	NO_x
		(%)					
1980[b]	17	33	93	68	92	50	26
1981	27	34	98	72	93	83	33
1982	38	37	85	60	87	NS	NS
1992	53	37	97	52	84	57	36
1993	47	54	95	74	95	81	50
1994	51	40	78	53	82	51	38
1995	22	23	56	34	61	36	22
1996	93	56	97	80	98	77	58

[a] Nonpoint source loads.

[b] Number of days and percentages for partial monitoring year of 224 days.

NS = Not sampled.

though these 2 years had virtually the same annual runoff volume (see Table 4). Furthermore, most of the events of 1993, particularly those producing high river flows, occurred during late winter and early spring when frozen ground and snowmelt would produce more surface runoff, while no events occurred in late spring or summer that year. In contrast, the winter-early spring events of 1994 were generally of lesser magnitude, and several events occurred in late spring and summer when precipitation tends to run off less, producing smaller material loads.

Most of the annual loading of PP and TSS to Cannonsville Reservoir occurred during runoff events in all years (Table 5). This was also the case for TDP and $T\text{-}NH_3$ except in 1995, the lowest runoff year in the monitoring program. Thus, a disproportionate share of the annual nonpoint source P, sediment and ammonia loads is delivered to Cannonsville Reservoir during those relatively few days when events occur. These results, in addition to the considerable variability observed in concentrations during events (see Table 3), establish the need to sample during runoff events to support an accurate determination of loads for these materials.

Based on the extremes of weather that occurred in 1995 and 1996, nonpoint source loads measured in those same years may be considered as approaching a likely range for present-day nutrient and sediment inputs from the WBDR watershed. The 1996 loads were greatly influenced by the January 19-20 event. This single event accounted for 74%, 34%, and 75% of that year's PP, TDP, and TSS loads, respectively. Conversely, 1995 was the third lowest flow year on record with only a few modest events and very little snowmelt. So, it is not surprising that, for most analytes, the nonpoint source loads of other years fall between those measured in 1995 and 1996 (Table 4).

While the January 19-20, 1996, event produced unusually prodigious loads of some analytes, single large events have accounted for a substantial portion of the annual loading of nonpoint source sediment and nutrients in other years as well. In 1981, 1992, 1993 and 1996, one large event produced from one-fifth to nearly half of the year's nonpoint TDP load (the remainder being delivered during other, smaller events and baseflow periods). These single large events also dominated the annual nonpoint source loading of PP and TSS, producing between 60% and 85% of each year's total loads. Since these large events do not appear to be uncommon occurrences, adequate monitoring of them is crucial to accurate determination of nonpoint source loadings. Without it, loads could be seriously underestimated.

Reservoir Response to Changes in Phosphorus Loads

Considerable reductions have been achieved since 1992 in P loads emanating from point sources (Fig. 6), primarily due to treatment upgrades at the Walton WWTP. Yet, Cannonsville Reservoir remained eutrophic in 1993 and beyond (Effler and Bader 1998). Reductions in point source P loading and any resultant amelioration

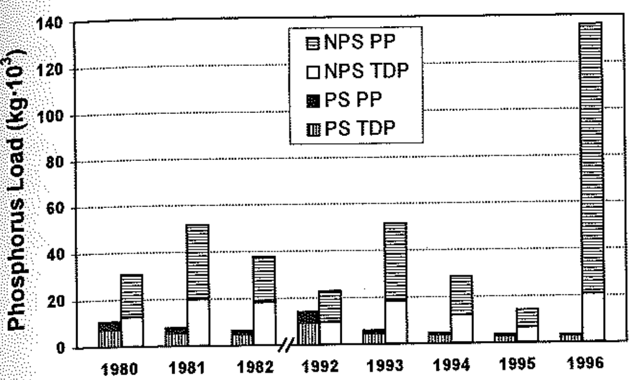

Figure 6.–Annual point and nonpoint source phosphorus loads in the WBDR for the study years. Point source load in 1980 is for entire year; nonpoint source load is for March - September.

of the reservoir's eutrophication problem must be weighed against the major inter-annual variations in nonpoint source P loading (Fig. 6). Phytoplankton growth in the reservoir appears to be largely driven by TDP loading (Doerr et al. 1998). Inter-annual variations in nonpoint source loading of TDP, which are large compared to the reduction in point source loads (Fig. 6), may tend to mask improvements in reservoir water quality expected from point source reductions over the short term. Additionally, even greater inter-annual variations occur in nonpoint source PP loading (Fig. 6). The dynamics of PP and productivity in this system are complex, involving a balance between depositional losses in the reservoir and the reaction rate that defines the conversion of PP to a dissolved form of P that can support phytoplankton growth (Auer et al. 1998). Though this fraction is decidedly less potent than TDP in supporting algae in the reservoir (Auer et al. 1998), model (Doerr et al. 1998) simulations indicated that the very large PP load of 1996 would elicit a noticeable response in phytoplankton growth. This response was manifested in predicted chlorophyll concentrations in Cannonsville Reservoir that averaged 3 $\mu g \cdot L^{-1}$ lower over the growing season when the 1996 PP load was set to zero in model simulation. Simulations conducted for 1994 and 1995 indicated the effect of PP on reservoir chlorophyll concentration was less than in 1996 because the PP loads those years were not nearly as large as that of 1996 (Fig. 6). Thus, natural variations in PP loading should also be expected to mask point source reductions. Based on the results of our loading studies and the modeling analyses of Doerr et al. (1998) and Owens et al. (1998a), it appears that beyond further point source reductions, management efforts in the Cannonsville watershed must be directed primarily at nonpoint sources of dissolved P to be successful in ameliorating the reservoir's eutrophication problem. Nonpoint source reductions of PP should also be sought, though it should be recognized that the initial benefit would likely be less than for the dissolved fraction.

Loads Based on Event Sampling versus Other Estimation Techniques

Utilization of export coefficients for different types of land use to estimate material loads is attractive because it avoids the time and costs of event sampling. Using this method and employing a low and high coefficient value for each of the various land uses in the watershed, Bader (1993) estimated a range for nonpoint source TP loads from the entire Cannonsville Reservoir drainage basin of 11,600 to 54,700 kg with a "most likely" value of 23,100 kg. Our "true" nonpoint TP loads from just the WBDR alone ranged from 14,000 to 136,000 kg in 1995 and 1996, respectively. The loads for the remaining 5 whole water-years that we monitored are probably more typical: the average for these is 38,500 kg. Bader's most likely value for the entire drainage basin, reduced by 20% to represent just the WBDR input, differs from our average load by 20,000 kg, an underestimation of nearly half. Interestingly, several of the export coefficients used in Bader's (1993) analysis are reported as derived from studies conducted in the Cannonsville watershed, which would have been expected to improve the estimate. Use of other, less site-specific, export coefficients taken from the literature may have produced loads that differed even more from the true loads than reported here.

The export coefficient approach as used by Bader (1993) is limited to estimates of annual loads; generating loads for smaller time steps, such as weeks or days, is not possible. Simple methods of load estimation, such as this, provide only rough estimates of material loadings and have very limited predictive capability (USEPA 1992). [A full discussion of the limitations and applicability of the export coefficient method is presented in Reckhow et al. (1980) and Reckhow and Chapra (1983)]. Common uses of this method are as a screening tool in situations where a first approximation of annual loading is sought and when making preliminary relative comparisons of loads from different watersheds. The export coefficient method does not address the seasonal and inter-annual variability in nonpoint loads (e.g., Fig. 5, Table 4) associated with meteorological and hydrological variability and, thus, is of limited value when applying the results to rigorous decision-making for management purposes. There is a danger in estimating loads using export coefficients derived from 1 or 2 years of detailed monitoring as illustrated by comparing calculated coefficients for nonpoint TSS and TDP for our years of monitoring on the WBDR.

Sediment export ranged from 44 kg · ha^{-1} in 1995 to 890 kg · ha^{-1} in 1996 and nonpoint TDP export, in the same years, from 0.078 to 0.243 kg · ha^{-1}. There are obvious differences among years, and if only one of these years were available to represent export coefficients for the WBDR watershed then, clearly, 1995 would seriously underestimate an "average" year, while 1996 would considerably overestimate it.

The load estimation technique of applying the FLUX program to a weekly grab sampling protocol yielded mixed results when comparisons were made between annual FLUX loads and the true loads based on event sampling (Table 6). Annual loads of PP and TSS, materials strongly associated with events, were considerably underestimated by the weekly sampling program and application of FLUX for all years except 1993 for which annual loads were overestimated. Years in which loads were dominated by a large event, such as 1981 and 1996, showed the greatest deviation from the true loads of most analytes. For example in 1981, the annual TDP load was underestimated by 5,100 kg or 20%, while in 1996, the annual PP load was underestimated by 89,000 kg or 77%. Drier years with fewer events, such as 1992 and 1995, show better agreement for most analytes.

Total annual loads of NO_X estimated by the weekly sampling program and application of FLUX were close to true loads (Table 6). FLUX loads of T-NH_3 agreed fairly well with true loads except for a substantial underestimation in 1981. This anomaly is likely associated with the limited T-NH_3 data set of the 1980s (analysis of T-NH_3 was performed for only 1 year) which, when used with FLUX, resulted in a poor flow-concentration relationship for this analyte and, thus, poor concentration estimates during periods when samples were not available.

While some of the annual load estimations produced by FLUX and the weekly sampling program matched well with the corresponding true load, the temporal structure of loading was often not accurately reproduced. These temporal inaccuracies can be seen when the difference between FLUX and true loads of several analytes are plotted on a daily time step for the 1990s (Fig. 7a-c). A good example is NO_X in 1996: the *annual* FLUX load is quite close to the true load (Table 6), but considerable differences emerge in the daily loads (Fig. 7c). Overestimation of NO_X loads on some days cancels the large underestimation in January 1996. The most substantial underestimations by FLUX for all analytes (Fig. 7a-c) tend to coincide with runoff event periods (Fig. 7d). Overestimation of daily loads of all analytes by FLUX resulted during the late March - early April 1993 interval because concentrations predicted by FLUX using a flow-concentration relationship for the recession of this event were higher than what actually occurred. Much of the flow in the recession of that event was contributed by snowmelt which tended to dilute concentrations to less than what was predicted for such high flows using the flow-concentration relationship produced by the weekly sample data set. This again emphasizes the need to adequately sample runoff events since events with similar flows may not have similar concentrations due to differences in ambient conditions preceding and during the event.

The performance of FLUX with a weekly sampling program was best in years that had fewer and smaller events (e.g., 1995). For years with large events where

Table 6.–Difference between loads estimated by weekly grab sampling and application of FLUX (F) and "true" annual loads (T), and difference as percent of true annual load in parentheses. Negative or positive values indicate underestimation or overestimation, respectively, by FLUX.

WY	Total Annual Difference (F-T)					
	PP[a]	TDP[a]	SRP[a]	T-NH_3[a]	NO_X[a]	TSS[b]
1981	-22,400 (-65)	-5,120 (-20)	-4,410 (-21)	-14,500 (-46)	-14,700 (-5)	-12,300 (-70)
1982	-10,100 (-48)	-1,170 (-5)	NS	NS	NS	-7,620 (-56)
1992	-3,060 (-17)	-740 (-4)	NS	-6 (<-1)	-9,830 (-3)	-230 (-3)
1993	8,900 (+26)	900 (+4)	1,470 (+9)	23 (<+1)	7,070 (+2)	2,990 (+13)
1994	-8,020 (-46)	-1,140 (-7)	-560 (-5)	-2,530 (-17)	-5,370 (-1)	-5,590 (-50)
1995	-3,030 (-37)	-1,260 (-13)	-590 (-8)	-1,780 (-20)	260 (<+1)	-1,490 (-39)
1996	-89,500 (-77)	-5,490 (-23)	-3,220 (-20)	-2,590 (-8)	-1,900 (<-1)	-21,700 (-29)

[a]Differences in kg.
[b]Differences in kg · 10^3.
NS = Not sampled.

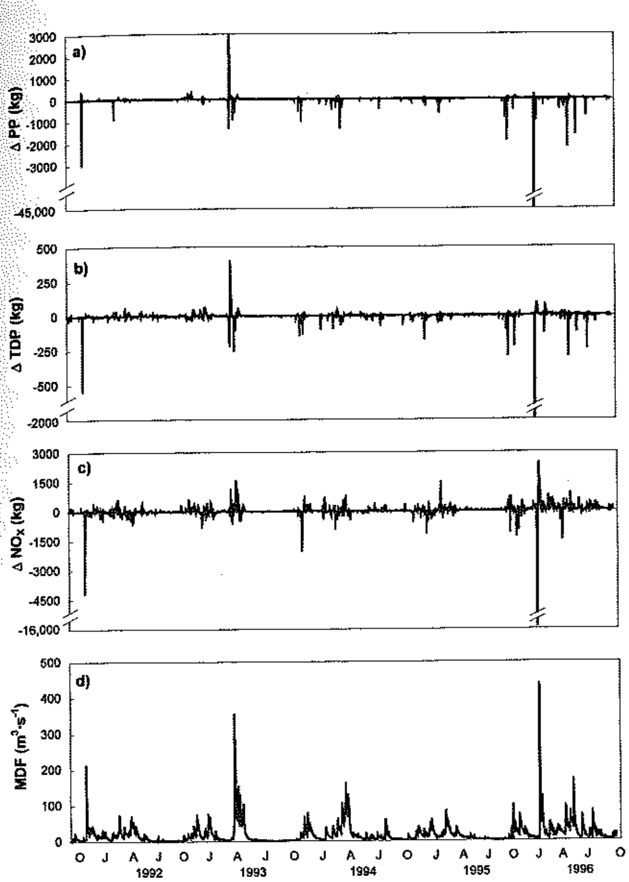

Figure 7.–a-c) Difference between "true" daily loads of PP, TDP, and NO_x and those estimated by weekly sampling and application of FLUX for the 1990s (Δ=FLUX-TRUE; e.g., a negative value indicates underestimation by FLUX); d) Mean daily flow of WBDR at Walton for comparison; spikes indicate events.

material concentrations varied widely over different flow regimes, the loads calculated by FLUX often deviated considerably from true loads. Analytes exhibiting the greatest differences in mean concentration between event and baseflow periods, i.e., PP and TSS (see Table 3), showed the most deviation from true loads when estimated using FLUX and weekly grab sample concentrations. The natural variability in concentrations, and therefore loads, during events does not appear to be adequately represented with the relatively few samples collected each year by a weekly, fixed-frequency sampling program, particularly when only 1 or 2 years of sampling data are available as with T-NH_3 in the 1980s. The comparison presented here underscores the limitations of a fixed-frequency sampling program to support estimates of loads for the WBDR, that doubtless apply to many other systems. Inaccuracies in time-series of daily and seasonal loads of nutrients and sediment (e.g., Fig. 7) that can extend to longer time frames (e.g., years; Table 6) are to be expected.

The effects that temporal inaccuracies in loads based on fixed-frequency sampling can have on model output were evaluated with the nutrient-phytoplankton model developed for Cannonsville Reservoir (Doerr et al. 1998), which runs on a daily time step. Simulations of chlorophyll in the epilimnion of the reservoir during the growing seasons of 3 study years (1994-1996), based on the true loads and the loads estimated by FLUX with fixed-frequency (weekly) samples, are compared (Fig. 8a-c). The chlorophyll concentrations predicted by the model were quite similar in 1994 (Fig. 8a) for the two data sets, averaging only 0.5 $\mu g \cdot L^{-1}$ lower when the FLUX loads were used as inputs. In 1995, however, chlorophyll concentrations predicted using FLUX loads averaged 1 $\mu g \cdot L^{-1}$ lower overall, and as much as 3 $\mu g \cdot L^{-1}$ lower for the mid-July to mid-August period (Fig. 8b). The greatest deviations were observed in 1996 (Fig. 8c), a year in which FLUX loads differed considerably from true loads (Fig. 7, Table 6). Chlorophyll concentrations predicted using FLUX loads averaged 3 $\mu g \cdot L^{-1}$ lower for the entire 1996 growing season, and as much as 8 $\mu g \cdot L^{-1}$ lower for the mid-summer period (Fig. 8c). As these model simulations were run discretely for each year starting at the onset of the growing season (i.e., initial reservoir conditions, including nutrient concentrations, specified from direct

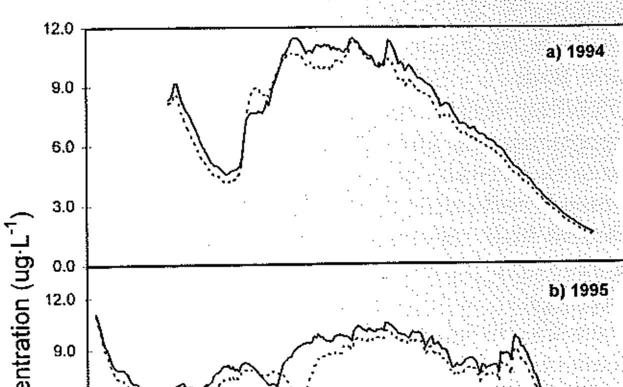

Figure 8.–Comparison of model simulations of daily epilimnetic chlorophyll concentration in Cannonsville Reservoir for the growing seasons of 3 study years using both true loads (solid line) and FLUX loads (dashed line) as daily inputs.

measurements made in April of each year; see Doerr et al. 1998), the effects of load inaccuracies on reservoir water quality predictions might be expected to be even greater if the model were run on a continuous basis for successive years. The results of this analysis demonstrate that output from water quality models can be sensitive to the inaccuracies in loads introduced by a fixed-frequency sampling program. Such loads may be more acceptable in analyses with long time-steps, (e.g., annual) or to support screening-type evaluations.

Implications of Load Inaccuracies When Setting Load Reduction Goals

The reduced ability of a non-event oriented monitoring program to detect inter-annual variability in nonpoint source loads and quantify the disproportionate amount of nutrients and sediments delivered during relatively few event days may hinder efforts to appropriately establish load reduction goals for the purpose of reversing eutrophication or other water quality problems. In the case of Cannonsville Reservoir, a primary forcing function of water quality is tributary P inputs. Reducing tributary inputs will be a principal means of effecting an improvement in future water quality of this reservoir. Yet, without accurate information on tributary loading, reduction goals could be improperly set. As a simple example, we will assume there is no point source TP load to Cannonsville and Bader's (1993) "mostly likely" TP load value of 23,100 kg, determined by the export coefficient method, represents the annual nonpoint input. One might then expect water quality to begin to improve were this amount reduced by 3,100 kg to reach the Vollenweider critical load reported by Bader for Cannonsville of 20,000 $kg \cdot y^{-1}$. However, WBDR nonpoint loads of TP in our "typical" monitored years were considerably greater than 23,100 kg, averaging 38,500 kg. Reducing TP loads in these years by only 3,100 kg would undoubtedly have had little or no noticeable effect and the implemented management program would likely be branded a failure.

Inter-annual variability in nonpoint source loads focuses attention on another factor to consider when setting target loads for water bodies. Traditional allocation methods for wastewater discharges to streams typically consider impacts for a given design condition: usually, the 10-year, 7-day low flow. Based on probability, this assumes that in 9 out of 10 years, stream water quality will be satisfactory due to factors such as sufficient dilution or assimilation. No corresponding design condition is in common use for determining acceptable inputs to lakes or reservoirs, however. If *average* annual or seasonal loads are used as the basis for setting load limits reflective of a desired water quality or trophic state objective in reservoirs or lakes, this implies that, on average, and all other factors being equal, every other year water quality would not be acceptable. Loading *maxima* rather than *averages* may be a more appropriate consideration for these water bodies. If the stream protection analogue were applied to lakes, then the 10-year maximum loading might be reasonable to use when determining the load reductions necessary for achieving the desired water quality condition. This would extend a degree of protection to lakes similar to that given streams and rivers. For lakes that serve as water supplies, 9 out of 10 years with satisfactory quality may be an unacceptable design parameter. In these cases, a 25- or 50-year condition may be more suitable. Certainly the use of an average loading is not appropriate for water supplies unless additional treatment is provided to compensate for the unsatisfactory water quality that would likely occur, based on probability, every other year.

Clearly, inter-annual variability in nonpoint source loads and subsequent effects on lakes and reservoirs must be considered when attempting to establish load limits and implement a successful management program. Thus, it is desirable when setting loading goals to obtain the best possible estimates of current tributary inputs and their inter-annual variability. Based on the results of our monitoring, several years of event-oriented sampling, particularly of the fall and spring runoff periods, combined with routine baseflow sampling, would appear to improve the accuracy of annual loads considerably over export coefficient methods as well as provide estimates of inter-annual variability. Techniques that combine a fixed-frequency or intermittent sampling scheme with actual flow data, such as the weekly grab sampling and FLUX method tested here, may improve estimates over those from export coefficients, but are not as reliable as the event-based program. The FLUX method may be acceptable for some analytes during more "typical years" (e.g., TDP in 1982, 1992 - 1994; Table 6), but for other materials (PP, TSS, $T\text{-}NH_3$) it would appear that inaccuracies in load estimates, both annually and for smaller time steps, may often be too great for informed management decisions to be made. At the very least, a generous margin of safety should be incorporated into the goal-setting process when managers are forced to use load estimates based on limited data sets.

Summary

Our monitoring program measured total annual TP loads from the WBDR that ranged from 17,800 to

140,000 kg. Loading from unmonitored tributaries and direct drainage, added to the WBDR inputs, would result in a present-day TP load range for Cannonsville Reservoir of about 20,000 to 166,000 kg · y^1. Annual load estimates for Cannonsville made by previous investigators (Table 1) using limited water quality data or an export coefficient method, all fell within this range of values, though none predicted a range as large.

The results of our study indicate that event sampling is necessary to reliably assess contributions from nonpoint sources. It is also essential to monitor years with different runoff conditions to observe variability in concentrations and determine possible ranges of expected loadings. This is particularly important when assessing receiving lake water quality, through use of models or other means, over various hydrologic/meteorologic scenarios that cover a realistically wide range of conditions.

The natural variability in nonpoint source contributions, on both seasonal and annual scales, has implications when setting target loads related to improved water quality or trophic state. This variability may mask the effects of point source reductions on lake water quality. Wide ranges in loads should be identified where they exist and appropriately accommodated when establishing material loading goals for Cannonsville Reservoir, or other water bodies, and implementing a management program that will have a chance at success.

ACKNOWLEDGMENTS: Funding for chemical analysis of samples and field staff salaries was provided by the New York City Department of Environmental Protection and the U.S. Environmental Protection Agency. Most of the sample collection in the 1990s was performed by field staff Jeff Lojpersberger and Larry Underwood. Mary Gail Perkins and Susan Doerr of Upstate Freshwater Institute, respectively, performed the program runs of FLUX and reservoir water quality model simulations used in the paper. Special thanks goes to Dr. Steven Effler of Upstate Freshwater Institute for his valuable manuscript review and comments.

References

Alberts, E. E., G. E. Schuman and R. E. Burwell. 1978. Seasonal runoff losses of nitrogen and phosphorus from Missouri Valley loess watersheds. J. Environ. Qual. 7(2):203-208.

APHA. 1992. Standard Methods for the Examination of Water and Wastewater. 18th ed. American Public Health Association, Washington, DC.

Auer, M. T., K. A. Tomasoski, M. J. Babiera, S. M. Doerr, S. W. Effler and J. M. Hansen. 1998. Phosphorus bioavailability and P-cycling in Cannonsville Reservoir. Lake and Reserv. Manage. 14(0):00-00.

Bader, A. P. 1993. A phosphorus loading estimate for New York City's Cannonsville Reservoir. New York City Dept. of Environ. Prot., New York, NY.

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

Black and Veatch Consulting Engineers. 1971. Process design manual for phosphorus removal: USEPA Technology Transfer. Kansas City, MO.

Bloomfield, J. A., J. W. Sutherland, J. Swart and C. Siegfried. 1984. Surface runoff water quality from developed areas surrounding a recreational lake. P. 40-47. *In:* Lake and Reservoir Management, N. Am. Lake Manage. Soc. Proc., EPA 440/5-84-001. Off. Water Reg. & Stand., Wash., DC.

Bricke, K. 1975. An updated computation of the phosphorus loading to the Cannonsville Reservoir. Unpubl. report, EPA Region II, New York, NY.

Brown, M. P. and M. R. Rafferty. 1980. A historical perspective of phosphorus loading to the Cannonsville Reservoir as it relates to the West Branch Model Implementation Program. Technical Report No. 62, New York State Dept. of Environ. Conserv., Albany, NY.

Brown, M. P., M. R. Rafferty and P. Longabucco. 1980. NY-MIP Technical Note: Trophic status of the Cannonsville Reservoir, 1980. Bureau of Water Research Pub. New York State Dept. of Environ. Conserv., Albany, NY.

Brown, M. P., M. R. Rafferty and P. Longabucco. 1983. Phosphorus transport in the West Branch of the Delaware River watershed: Vol. 3 of Nonpoint Source Control of Phosphorus - A Watershed Evaluation. Final report to the U.S. Environ. Prot. Ag., Region II. New York State Dept. of Environ. Conserv., Albany, NY.

Brown, M. P., P. Longabucco and M. R. Rafferty. 1986. The eutrophication of the Cannonsville Reservoir: Vol. 5 of Nonpoint Source Control of Phosphorus - A Watershed Evaluation. Final report to the U.S. Environ. Prot. Ag., Region II. New York State Dept. of Environ. Conserv., Albany, NY.

Brown, M. P., P. Longabucco, M. R. Rafferty, P. D. Robillard, M. F. Walter and D. A. Haith. 1989. Effects of animal waste control practices on nonpoint-source phosphorus loading in the West Branch of the Delaware River watershed. J. Soil and Water Cons. 44(1):67-70.

Carlson, R. E. 1977. A trophic state index for lakes. Limnol. Oceanogr. 23:361-369.

Cutietta-Olson, C. R. 1997. Unpubl. data. New York City Dept. of Environmental Protection, New York, NY.

Dillon, P. J. 1975. The phosphorus budget of Cameron Lake, Ontario: the importance of flushing rate to the degree of eutrophy of lakes. Limnol. Oceanogr. 20:28-39.

Doerr, S. M., E. M. Owens, R. K. Gelda, M. T. Auer and S. W. Effler. 1998. Development and testing of a nutrient-phytoplankton model for Cannonsville Reservoir. Lake and Reserv. Manage. 14(2-3):301-321.

Dolan, D. M., A. K. Yui and R. D. Geist. 1981. Evaluation of river load estimation methods for total phosphorus. J. Great Lakes Res. 7(3):207-214.

Effler, S. W. and A. P. Bader. 1998. A limnological analysis of Cannonsville Reservoir, NY. Lake and Reserv. Manage. 14(2-3): 125-139.

Effler, S. W. and K. A. Whitehead. 1996. Tributaries and discharge. P. 97-199. *In:* S. W. Effler (Ed.), *Limnological and Engineering Analysis of a Polluted Urban Lake*, Springer-Verlag, New York, NY.

Goodale, B. 1975. An analysis of phosphorus input to Cannonsville Reservoir. Technical report, New York State Dept. of Environ. Conserv., Albany, NY.

Heidtke, T. M. and M. T. Auer. 1992. Partitioning phosphorus loads: Implications for lake restoration. J. Water Resour. Plan. Manag. ASCE 118(5):562-579.

Hutchinson, G. E. 1973. Eutrophication: The scientific background of a contemporary practical problem. Amer. Sci. 61:269-279.

Hydroscience. 1975. Water quality analysis of the West Branch of the Delaware River. Westwood, NJ.

Longabucco, P. and M. R. Rafferty. 1989. Delivery of nonpoint source phosphorus from cultivated mucklands to Lake Ontario. J. Env. Qual. 18:157-163.

Lumia, R. 1998. Flood of January 19-20, 1996 in New York State. Water Resources Investigations Report 97-4252. U.S. Geological Survey, Troy, NY.

Manczak, H. and L. Florczyk. 1971. Interpretation of results from the studies of pollution of surface flowing waters. Water Res. 5:575-584.

National Oceanic and Atmospheric Administration (NOAA). 1980-1982, 1991-1996. Climatological data for New York: Vols. 92-94, 103-108. National Climatic Data Center, Asheville, NC.

Novotny, V. and G. Chesters. 1981. Handbook of nonpoint pollution: sources and management. Van Nostrand Reinhold Co., New York, NY.

Owens, E.M., S.W. Effler, S.M. Doerr, R.K. Gelda, E. Schneiderman, D.G. Lounsbury and C. L. Stepczuk. 1998a. A strategy for reservoir model forecasting based on historic meteorological conditions. Lake and Reserv. Manage. 14(2-3):322-331.

Owens, E. M., R. K. Gelda, S. W. Effler and J. M. Hassett. 1998b. Hydrologic analysis and model development for Cannonsville Reservoir. Lake and Reserv. Manage. 14(2-3):140-151.

Reckhow, K. H., M. N. Beaulac and J. T. Simpson. 1980. Modeling phosphorus loading and lake response under uncertainty: A manual and compilation of export coefficients. U.S. Environ. Prot. Ag. (EPA 440/5-80-011).

Reckhow, K. H. and S. C. Chapra. 1983. Engineering approaches for lake management, Volume 1: Data analysis and empirical modeling. Butterworth Publishers, Boston.

Richards, R. P. and J. Holloway. 1987. Monte Carlo studies of sampling strategies for estimating tributary loads. Water Resour. Res. 23(10):1939-1948.

Schumacher, G. J. and D. B. Wager. 1973. A study of the phytoplankton in the Delaware River Basin streams in New York State. Delaware River Basin Commission, Trenton, NJ.

Stepczuk, C., A. B. Martin, P. Longabucco, J. A. Bloomfield and S. W. Effler. 1998. Allochthonous contributions of THM precursors to a eutrophic reservoir. Lake and Reserv. Manage. 14(2-3):344-355.

Tin, M. 1965. Comparison of some ratio estimators. J. Am. Statist. Assoc. 60:294-307.

U.S. Environmental Protection Agency (USEPA). 1974. Report on Cannonsville Reservoir, Delaware Co., NY. EPA Working Paper No. 150, National Eutrophication Survey.

USEPA. 1983. Methods for Chemical Analysis of Water and Wastes. EPA 600/4-79.020. Off. Res. Dev., Cincinnati, OH.

USEPA. 1992. Compendium of watershed-scale models for TMDL development. EPA 841-R-92-002. Off. Water, Washington, DC.

U.S. Geological Survey (USGS). 1981, 1982, 1983. Water Resources Data for New York, Water Years 1980, 1981, 1982. Water Data Reports, U. S. Geological Survey, Albany, NY.

USGS. 1993, 1994, 1995, 1996, 1997. Water Resources Data for New York, Water Years 1992, 1993, 1994, 1995, 1996. Water Data Reports, U. S. Geological Survey, Troy, NY.

Vollenweider, R. A. 1968. The scientific basis of lake and stream eutrophication, with particular reference to phosphorus and nitrogen as eutrophication factors. OECD Tech. Rep., Paris, DAS/CSI 68:21.

Vollenweider, R. A.. 1975. Input-output models with special reference to the phosphorus loading concept in limnology. Schwerz. Z. Hydrol. 37:53-84.

Walker, W. W. 1987. Empirical Methods for Predicting Eutrophication in Impoundments. Report 4: Phase III: Applications Manual. Technical Rep. E-81-9. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Wood, L. W. 1979. The limnology of the Cannonsville Reservoir, Delaware Co., NY. Environ. Health Rep. No. 6, New York State Dept. of Health, Albany, NY.

Young, T. C., J. V. DePinto and T. M. Heidtke. 1988. Factors affecting the efficiency of some estimators of fluvial total phosphorus load. Water Resour. Res. 24(9):1535-1540.