SAMPLING CONSIDERATIONS FOR ESTABLISHMENT OF BASELINE LOADINGS FROM FORESTED WATERSHEDS FOR TMDL APPLICATION

PAMELA J. EDWARDS[1]*, KARL W. J. WILLIARD[2] and
JAMES N. KOCHENDERFER[1]
[1] *USDA Forest Service, Northeastern Research Station, Parsons, West Virginia, U.S.A.;*
[2] *Department of Forestry, Southern Illinois University, Carbondale, Illinois, U.S.A.*
(author for correspondence, e-mail: pjedwards@fs.fed.us)*

(Received 4 March 2003; accepted 15 September 2003)

Abstract. Five methods for estimating maximum daily and annual nitrate (NO_3) and suspended sediment loads using periodic sampling of varying intensities were compared to actual loads calculated from intensive stormflow and baseflow sampling from small, forested watersheds in north central West Virginia to determine if the less intensive sampling methods were accurate and could be utilized in TMDL development. There were no significant differences between the annual NO_3 load estimates using non-intensive sampling methods and the actual NO_3 loads. However, maximum daily NO_3 loads were estimated less accurately than annual loads. The ability to estimate baseline NO_3 loads fairly accurately with non-intensive concentration data is attributed to the small fluctuation in NO_3 concentrations over flow and time, particularly during storms. By contrast, suspended sediment exports determined by any of the non-intensive methods varied significantly and widely from and compared poorly to the actual exports for both daily and annual methods. Weekly sampling better approximated actual annual exports, but there were no significant statistical differences among weekly, monthly, and quarterly estimates. Suspended sediment concentrations varied widely within and among storm events, so that accurate estimates of total annual or maximum daily loads could not be obtained from infrequent sampling.

Keywords: nitrate, sampling frequency, stormflow concentrations, suspended sediment, total maximum daily loads

1. Introduction

The concept of total maximum daily loads (TMDLs) was established in 1972 in section 303(d) of the Federal Water Pollution Control Act Amendments to the Clean Water Act (P.L. 92-500) of the United States of America. Section 303(d) requires states to identify water that remained polluted after the adoption of technology standards aimed at point source polluters, prioritize the impaired waters, and develop TMDLs to meet applicable water quality standards. The formula for TMDL development includes waste load allocations from point sources, load allocations from nonpoint sources, a margin of safety for uncertainty in the calculations, and a margin for discharges from future development within the watershed. Action on TMDL development remained dormant until the late 1980s–early 1990s when

Environmental Monitoring and Assessment **98:** 201–223, 2004.

a series of citizen lawsuits spurred the United States Environmental Protection Agency (U.S. EPA) and states into action (Houck, 1999). Most recently, the focus of TMDL development has been on nonpoint sources, such as agriculture and silviculture.

Before effluent loads from different pollution sources or land uses can be allocated to meet water pollution standards, current water pollutant levels contributed from each source within the watershed must be assessed and quantified. The most reliable way to assess these baseline contributions is to sample water for pollutants of interest, at levels in both space and time that adequately quantify these conditions. After current levels are established, pollutant allocation to the various sources within the watershed can be accomplished in a variety of ways so that permissible effluent loads along with a margin of safety (33 U.S.C. Section 1313) are achieved.

Unfortunately, rigorous establishment of baseline pollutant levels is not easy and often may not be accurate because sufficient numbers of samples are not collected or sampling times do not correspond to important discharge events covering the range of baseline conditions. Resulting allocation of pollutant loads, therefore, may not result in improved water quality. Establishment of baseline conditions may be difficult from forested lands on which silvicultural operations and other nonpoint sources of pollution exist because effluent discharges are not necessarily predicted as easily as from point sources. For example, the timing of discharges and associated concentrations of pollutants from industrial point sources often are controlled by the industry, so corresponding loadings are quantified easily. By contrast, the highest concentrations of pollutants from forested watersheds often are attributable to storm events (Edwards *et al.* 2001; Eshleman *et al.*, 1998; Swistock *et al.*, 1997), which occur randomly. All ranges of flows and associated constituent concentrations may need to be included in the data set to establish baseline loadings within an acceptable level of accuracy; therefore, it may not be reasonable to use infrequent sampling during baseline establishment of pollutant loads. If initial estimates of current conditions are inaccurate, allocation of TMDLs will be based on flawed values and assumptions that may not reduce impairment.

State water quality agencies throughout the country measure baseline concentrations and loadings for TMDLs in most streams and rivers from in-place state water quality sampling programs or from federal programs (e.g., United States Geological Survey sampling). Due to costs and available resources, these programs typically employ infrequent sampling strategies; sampling from a given stream or river is rarely more frequent than seasonal (i.e., 4 times per year). Consequently, our objective was to examine how suspended sediment and NO_3 loadings, using various sampling frequencies, compare to actual loadings. We employed sampling frequencies that states are using or are likely to use to determine baseline conditions, as well as more intensive frequencies than are commonly used. These comparisons provide an indication of whether typical sampling frequencies will yield reasonable estimates of baseline conditions, which largely will determine if TMDL success can be measured. We selected suspended sediment and NO_3

because they are common nonpoint source pollutants from forestry operations (Reinhart *et al.*, 1963; Vitousek and Melillo, 1979; Martin *et al.*, 2000).

Our objectives in this paper should not be confused with the many and varied modeling approaches used later in the TMDL process, such as development of models to estimate concentrations, flows, or loads during periods that have not been sampled or measured, or allocations of loads to the various land owners and pollution sources in the watershed. Such modeling or allocation efforts are built upon the assumption that current water quality conditions are known with some acceptable degree of accuracy, and these known conditions are incorporated in various ways in these models. If calculated baseline loads have substantial error, these later modeling efforts will be similarly flawed.

2. Methods

2.1. Field and Laboratory Procedures for NO_3

Stream NO_3 and suspended sediment concentration data used in this analysis were collected on separate watersheds. Stream NO_3 concentrations were obtained from 1988 to 1993 from watershed 4 (WS4) on the Fernow Experimental Forest, Tucker County, West Virginia (WV). This 38.7 ha catchment has been relatively undisturbed since approximately 1905. Around 1905 the native hardwood vegetation growing on WS4 was harvested heavily. Since that time it has regenerated naturally and now supports a mature overstory stand dominated by sugar maple (*Acer saccharum*), red maple (*A. rubrum*), American beech (*Fagus grandifolia*), northern red oak (*Quercus rubra*), and black cherry (*Prunus serotina*).

Streamflow on WS4 has been measured continuously since 1951 with a 120° sharp-crested V-notch weir in combination with a Belfort FW-1 water-level recorder. 'Routine samples' of stream water were taken every Tuesday at approximately the same time each week; thus, both baseflow and various levels of stormflows were represented by these grab samples.

In addition, 'stormflow samples' were collected from the WS4 stream during 63 rainfall and snowmelt events in the study period. Precipitation totals for these events ranged from 0.84 to 11.6 cm, though 76% of these had between 1.78 and 3.81 cm precipitation. Stormflow samples were collected throughout storm hydrographs using ISCO automatic samplers. Discrete samples were collected every 0.5 hr during the growing season and every 1 hr during the dormant season on the rising limb of virtually every storm hydrograph. Falling limb samples were collected on 1 hr or longer intervals regardless of season. The time between falling limb samples was largely dependent upon the storm length; the longer the falling limb, the greater the time between samples. Six hours generally was the longest duration between samples, and this length was for falling limbs that required days to return to baseflow levels. The longest runoff events typically were rain-on-snow events.

Samples were retrieved from the field and analyzed for NO_3 by ion chromatography using EPA-approved methods at the USDA Forest Service's Timber and Watershed Laboratory in Parsons, WV. Specific information on field and laboratory set-ups, equipment, and protocols are provided in Kochenderfer and Edwards (1990) and Edwards and Wood (1993).

2.2. FIELD AND LABORATORY PROCEDURES FOR SUSPENDED SEDIMENT

Suspended sediment samples were collected from 1994–1998 from two adjacent watersheds on the Monongahela National Forest in Randolph County, WV. One watershed (52.3 ha) has been undisturbed since 1938 when it came under Forest Service ownership (unharvested watershed). Prior to 1938, it had been farmed for many years, probably since the mid 1800s. Only a small amount of land in the upper portion of the watershed, well away from the stream, was plowed. A second watershed (104.7 ha) also had been farmed prior to 1938 (harvested watershed). Cultivation appears to have resulted in greater erosion and in-stream siltation on the harvested watershed compared to the unharvested watershed. A timber sale on one side of the harvested watershed was conducted approximately 25 yr ago. Another 27.1 ha were clearcut from June to November 1995. A skid road and haul road network occupying approximately 8–10% of the watershed area was constructed for the 1995 harvesting. All wood removal was done using wheeled skidders. Streamflow was measured continuously using a 0.76 m rectangular weir in the unharvested and a 0.91 m rectangular weir in the harvested watershed in combination with Belfort FW-1 water-level recorders.

The adjacent unharvested and harvested watersheds are located ~10.5 km air distance from WS4. Calvin channery silt loam soils and Hampshire geology predominate in all three watersheds. Topographic characteristics are also similar among the three watersheds. Mean slopes for the unharvested, harvested, and WS4 watersheds are 35, 25, and 25%, respectively. Elevation ranges for the unharvested, harvested, and WS4 watersheds are 700–1000, 700–1000, and 740–870 m, respectively.

'Routine' suspended sediment samples were collected every 6 hr throughout the study using ISCO automatic samplers. Due to the random nature of storm events, routine samples included baseflow and various event samples. 'Stormflow' samples were collected with separate ISCO automatic samplers on 1 hr intervals throughout the hydrograph of most of the rainstorms and snowmelt events ($n = 82$ for unharvested, $n = 80$ for harvested) that occurred during these years, although some samples in the early part of the study were collected on 30 min intervals. Since suspended sediment concentrations vary greatly throughout the storm hydrograph in these headwater streams, sampling the entire storm hydrograph assured that sediment concentrations were characterized well. Suspended sediment concentrations were determined at the Parsons laboratory. Total suspended solids were determined using U.S. EPA method 160.2; the mineral portion of the suspended sediment was further isolated by burning off the organic material at 823 °K in a muffle furnace using U.S. EPA method 160.4 (Keith, 1991).

2.3. Data Analyses

Even though the term total maximum daily loads suggests yields are based upon daily loads, TMDL legislation permits expressing yields over any logical time period for the constituent in question (40 CFR 130.2(i)). Therefore, TMDLs commonly are reported on an annual basis. Consequently, we considered both maximum daily loads and total annual loads.

Total annual loads were calculated on a water year, rather than a calendar year, basis. The water year for NO_3 was defined as May–April, which has been found to be most appropriate for this area. The growing season extends from May–October, and the 1st, 2nd, 3rd, and 4th quarters within a water year extend from May–July, August–October, November–January, and February–April, respectively.

To maximize the number of full years for sediment samples, water years for suspended sediment correspond to the initiation and termination of data collection. The water year for suspended sediment sampling extends from October–September. While this definition of water year is not quite as appropriate for this area as the May–April water year used for NO_3, it is used commonly in the region, so it is adequate for the comparative purposes of this paper. Under the October–September water year, the 1st, 2nd, 3rd, and 4th quarters of the year extend from October–December, January–March, April–June, and July–September, respectively.

2.3.1. *Annual Load Methods*

Estimates of actual total annual loads (kg ha^{-1} yr^{-1}) of NO_3 and suspended sediment were determined by applying flow from the watershed of interest to all of the available NO_3 or sediment concentrations from that watershed for each year. Every routine and storm sample that was collected was used to calculate actual loads. Each concentration (mg L^{-1}) was applied to the accumulated flow since the last sample was taken along with appropriate conversion factors to obtain loading values. These loads were summed annually and are considered 'actual annual exports'. While there are many ways to calculate annual loads, (e.g., Walling and Webb, 1981; Swistock *et al.*, 1997), we used summation procedures, which are generally more accurate than rating curves, at least for suspended sediment (Bunte and MacDonald, 1998). We used instantaneous values rather than mean values prior to summing to obtain better load estimates (Walling and Webb, 1981). Thomas (1991) suggested that random sampling may improve estimates of annual sediment loads, so we also included random sampling via the RANUNI function in SAS in our scheme (SAS, 1988). The RANUNI function in SAS created a series of random numbers between 0 and 1 for each time period of interest (i.e., each month or quarter). Seed numbers were created randomly in each SAS program so they were different for the monthly versus quarterly random samples. Resulting random numbers were assigned to categories based upon equal distance ranges.

To compare actual annual exports to estimates of exports using other less frequent sampling schemes, NO_3 and suspended sediment exports were calculated using the five methods below. Weekly, monthly, or quarterly loads were summed for each method to obtain estimates of total annual export.

1. Weekly exports (W): routine sample concentrations of NO_3 and suspended sediment were multiplied by the total accumulated flow for that week (Sunday–Saturday).
2. Monthly fixed exports (MF): the first routine sample concentration of NO_3 and suspended sediment in each month was multiplied by the total flow for that month.
3. Monthly random exports (MR): a randomly selected routine sample concentration of NO_3 and suspended sediment in each month was multiplied by the total flow for that month.
4. Quarterly fixed exports (QF): the first routine sample concentration of NO_3 and suspended sediment in each quarter was multiplied by the total flow for that quarter.
5. Quarterly random exports (QR): a randomly selected routine sample concentration of NO_3 and suspended sediment in each quarter was multiplied by the total flow for that quarter.

Methods W, MF, and QF correspond most closely to formalized sampling programs that employ repetitive timing of samples, while methods MR and QR correspond to programs that are less formal and depend on the schedules of the individuals collecting the samples rather than predefined fixed sampling times.

2.3.2. *Maximum Daily Load Methods*

Actual maximum daily exports were calculated using only the stormflow data set since the storms included the largest flows and concentration values (compared to the routine data set). The NO_3 or suspended sediment concentration (mg L^{-1}) was applied to the accumulated flow since the previous sample for that day. These loads were summed to obtain the actual daily load for each day. The single largest daily load in each year was then used as 'actual maximum daily load'.

Estimates of actual maximum daily loads for each year using less-frequent and non-intensive sampling (i.e., no discrete sampling throughout individual storms) were calculated using routine concentrations and employing the same sampling frequencies and fixed or random sample strategies used for annual exports. The same random samples selected for the annual exports were used in the daily export calculations.

1. Weekly sample (W): each routine sample concentration of NO_3 and suspended sediment was multiplied by the total flow for the day (00:00–24:00 hr) the sample was collected.

2. Monthly fixed sample (MF): the first routine sample concentration of NO_3 and suspended sediment in each month was multiplied by the largest total daily flow (00:00–24:00 hr) that month.

3. Monthly random sample (MR): a randomly selected routine sample concentration of NO_3 and suspended sediment in each month was multiplied by the largest total daily flow (00:00–24:00 hr) that month.

4. Quarterly fixed sample (QF): the first routine sample concentration of NO_3 and suspended sediment in each quarter was multiplied by the largest total daily flow (00:00–24:00 hr) that quarter.

5. Quarterly random sample (QR): a randomly selected routine sample concentration of NO_3 and suspended sediment in each quarter was multiplied by the largest total daily flow (00:00–24:00 hr) that quarter.

The largest total daily flow during the monthly and quarterly sampling periods were used in methods MF, MR, QF, and QR because the concentration measured during non-intensive sampling is assumed to be representative of that entire period, and loadings are controlled largely by flows.

2.3.3. *Statistical Analysis*

The difference between each actual annual or daily value and each estimated value was determined for each year and method. Negative values indicate that the actual value was less than the estimated value, while positive values indicate the actual value was greater than the estimated value.

Duncan's multiple range tests were used to determine differences among mean actual values and mean estimates in SAS (SAS, 1988). A significance level of 0.05 was used for all comparisons of means. To meet the assumptions of Duncan's test, the comparisons must involve data that are independent. However, because the computation procedure for each method employs data from the same data set, this assumption is not entirely met. The actual values and the weekly method comparisons involve the least independence, particularly for the annual loads. However, the other methods use data that are more independent from the other values due to the relatively small amount of overlapping data. Consequently, we employed Duncan's multiple range comparisons as a relative index of similarity or difference. The calculated values by each method also were similar enough or different enough that the reader likely would come to the same conclusions as the tests without statistical comparisons.

TABLE I

Annual maximum concentrations of all data (storms and routines) for NO_3 on WS4 (undisturbed) and suspended sediment on the unharvested and harvested watersheds

Year	NO_3	Year	Suspended sediment	
	WS4		Unharvested	Harvested
	(mg L^{-1})		(mg L^{-1})	
1988	4.43	1994	624.22	1235.39
1989	7.08	1995	214.30	573.42
1990	6.33	1996	1363.46	1749.30
1991	5.53	1997	77.48	172.63
1992	4.65	1998	219.59	174.57
1993	5.36			
1994	3.94			

3. Results

3.1. CONCENTRATIONS

Annual maximum NO_3 concentrations for WS4 and annual maximum suspended sediment concentrations for the unharvested and harvested watersheds are presented in Table I. Concentrations of NO_3 varied little, even during large discharges, compared to suspended sediment, which varied substantially with flow (Figure 1). Minimum concentrations for both NO_3 and suspended sediment were below the minimum detection limits (0.04 and 0.0001 mg L^{-1}, respectively). The maximum NO_3 concentration for all the monitored years was 7.08 mg L^{-1} in 1989, and all of the annual maximum concentrations occurred during storm runoff events. Maximum sediment concentrations were much greater than NO_3 concentrations. Suspended sediment maximums ranged from approximately 77 mg L^{-1} to more than 1700 mg L^{-1} (Table I). With the exception of 1998, maximum suspended sediment concentrations on the unharvested watershed were all less than on the harvested watershed.

3.2. ANNUAL NO_3 EXPORTS

Actual annual NO_3 exports from WS4 ranged from approximately 16 to 38 kg ha^{-1} yr^{-1} (Table II). There were no significant differences ($\alpha = 0.05$) among the mean loads estimated from each of the five methods and the mean of the actual annual NO_3 exports (Table II). The largest differences between the actual and estimated loads ranged from −1.04 kg ha^{-1} yr^{-1} (overestimated) to 4.9 kg

TABLE II

Difference between actual annual NO_3 loads and estimated annual loads using the five methods described in the text for WS4

Water year	Actual daily export[a]	Method									
		W		MF		MR		QF		QR	
		Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference
		($kg\ ha^{-1}\ yr^{-1}$)									
1988	18.31	19.35	−1.04	18.68	−0.37	18.66	−0.35	17.86	0.45	17.45	0.86
1989	28.73	28.33	0.40	27.41	1.32	25.75	2.98	23.83	4.90	25.30	3.43
1990	37.36	35.01	2.35	36.27	1.09	38.04	−0.68	36.10	1.26	32.91	4.45
1991	16.47	14.60	1.87	17.22	−0.75	17.24	−0.77	16.30	0.17	15.59	0.88
1992	19.62	18.61	1.01	18.59	1.03	17.83	1.79	18.32	1.30	18.27	1.35
1993	26.15	25.05	1.10	25.35	0.80	25.56	0.59	26.31	-0.16	22.82	3.33
Mean[b]	24.44a	23.49a	0.95	23.92a	0.52	23.85a	0.59	23.12a	1.32	22.06a	2.38

[a] The actual annual export was determined using all stormflow and routine sample concentrations from WS4.

[b] Means of actual values and estimates with different letters are significantly different from one another.

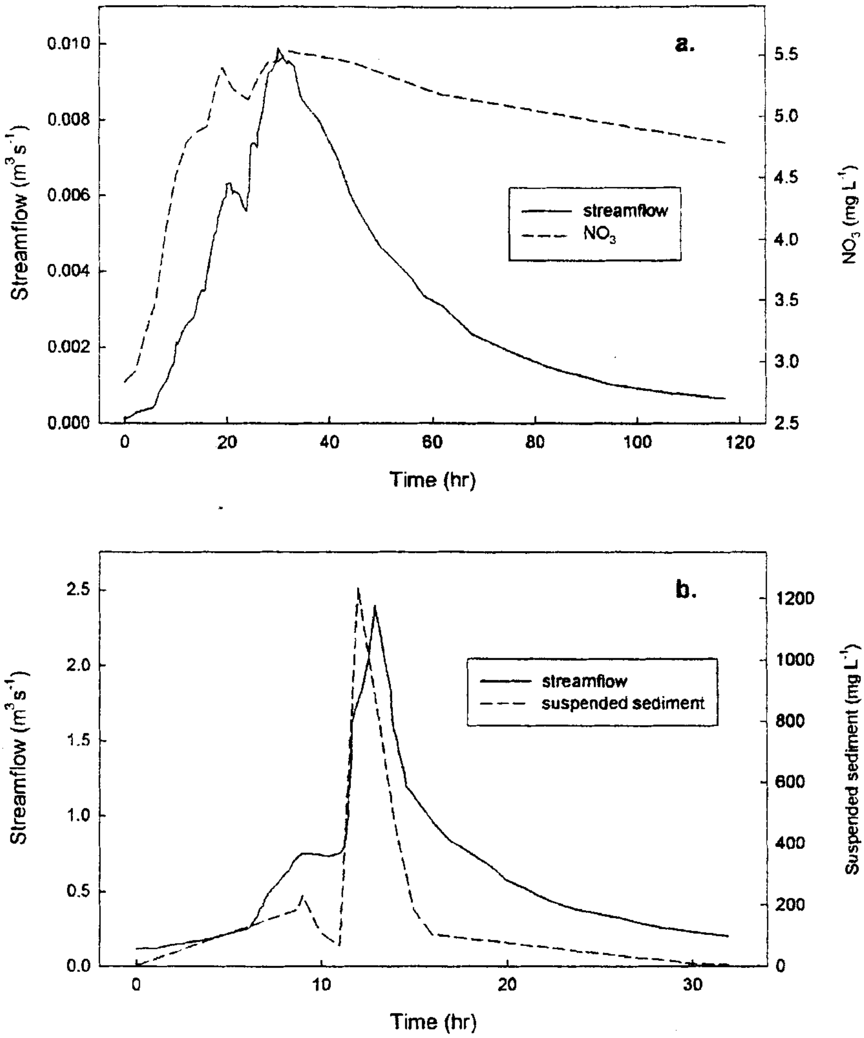

Figure 1. Examples of NO_3 and suspended sediment concentration changes with flow. (a) WS4, 21–26 November 1991, and (b) Harvested watershed, 7–8 May 1994. These storms had the greatest concentrations for their respective years, as in Table I.

ha^{-1} yr^{-1} (underestimated). Almost 80% of the individual differences were positive (Table II), indicating the five estimation methods tended to underestimate annual exports. Five of the seven negative differences occurred for the monthly calculation procedures (Table II).

We expected to see greater disparity between the actual and estimated annual loads with decreasing sampling frequency used to calculate the annual exports. While the average differences using quarterly samples (methods QF and QR) were greater than those for the monthly values (methods MF and MR) (Table II), there was no clear trend or pattern in the individual differences by sampling frequency.

3.3. Maximum Daily NO_3 Loads

Actual maximum daily NO_3 loads from WS4 ranged from about 0.7 to 2.8 kg ha^{-1} d^{-1} (Table III). As with the annual loads, there were no trends of decreasing ability to estimate maximum daily loads with decreasing sampling frequency and no significant differences among the means of the estimates or actual maximum daily exports (Table III). Again, maximum daily loads almost always were underestimated; only two of the 30 individual estimates were greater than the actual maximum daily load, and both of these estimates were very close to the actual maximum daily load (0.894 kg ha^{-1} d^{-1} estimated by methods W and MF vs. the actual maximum daily load of 0.871 kg ha^{-1} d^{-1}).

Maximum daily loads were not estimated as accurately as annual loads. Percent errors between the actual and estimated values (i.e., [(actual-estimated)/actual] $\times$ 100) for daily values primarily ranged from 20–55%, whereas the percent errors for annual loads generally were less than 13%.

3.4. Annual Suspended Sediment Exports

Actual annual suspended sediment loads were quite variable from year to year in both the unharvested and harvested watersheds (Table IV), but each year's actual exports were consistently greater in the harvested watershed than the unharvested watershed. In the harvested watershed actual exports ranged from 45 to 725 kg ha^{-1} yr^{-1}, whereas on the unharvested watershed actual suspended sediment exports ranged from 19 to 425 kg ha^{-1} yr^{-1} (Table IV).

Suspended sediment loads and differences between actual and estimated suspended sediment loads (Table IV) for each of the five methods were much greater for all years than those for NO_3 exports (Table II). Only three of the suspended sediment estimates were greater than the actual annual exports, and all of these were from the unharvested watershed (Table IV).

All of the mean estimates were significantly less than the actual annual exports on both watersheds (Table IV). There were no statistically significant differences among the 5 mean suspended sediment export estimates. However, with few exceptions, annual estimates of suspended sediment exports using weekly sampling (method W) better approximated the actual annual exports than any of the less frequent sampling schemes (methods MF, MR, QF, and QR) (Table IV). Differences across methods within a year were not dramatic, but year-to-year variation in the difference values was much more marked (Table IV). Regardless of the sampling frequency used, estimates of annual suspended sediment yields were poorer in 1994

TABLE III

Difference between actual maximum daily NO_3 loads and estimated maximum daily loads using the five methods described in the text for WS4

Water year	Actual daily maximum[a]	Method									
		W		MF		MR		QF		QR	
		Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference
	(kg ha^{-1} d^{-1})										
1988	0.871	0.894	−0.023	0.894	−0.023	0.745	0.126	0.791	0.080	0.676	0.195
1989	1.329	0.990	0.339	0.888	0.441	0.689	0.640	0.593	0.736	0.593	0.736
1990	1.260	0.871	0.389	0.970	0.290	1.026	0.234	0.908	0.352	0.920	0.340
1991	0.709	0.496	0.213	0.610	0.099	0.645	0.064	0.610	0.099	0.561	0.148
1992	1.303	0.896	0.407	0.907	0.396	0.896	0.407	0.863	0.440	0.896	0.407
1993	2.767	2.321	0.446	2.383	0.384	2.692	0.075	2.383	0.384	1.795	0.972
Mean[b]	1.373a	1.078a	0.295	1.109a	0.265	1.116a	0.258	1.025a	0.349	0.907a	0.466

[a] The actual maximum daily load was determined using all stormflow and routine sample concentrations from WS4.

[b] Means of actual values and estimates with different letters are significantly different from one another.

TABLE IV

Difference between actual annual suspended sediment loads and estimated annual loads using the five methods described in the text for the unharvested and harvested watersheds. Harvesting occurred June–November 1995

Water year	Actual annual export[a]	Method									
		W		MF		MR		QF		QR	
		Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference
	(kg ha^{-1} yr^{-1})										
Unharvested watershed											
1994	424.89	43.02	381.87	15.62	409.27	9.63	415.26	35.89	389.00	41.20	383.69
1995	19.12	4.82	14.30	1.60	17.52	35.11	−15.99	1.19	17.93	1.95	17.17
1996	342.34	62.93	279.41	13.04	329.30	62.52	279.82	4.78	337.56	14.43	327.91
1997	20.82	10.21	10.61	9.13	11.69	9.61	11.21	5.57	15.25	5.34	15.48
1998	26.34	14.19	12.15	56.35	−30.01	14.56	11.78	121.92	−95,58	8.32	18.02
Mean[b]	166.70a	27.03b	139.67	19.15b	147.55	26.29b	140.42	33.87b	132.83	14.25b	152.45
Harvested watershed											
1994	318.53	25.28	293.25	13.63	304.90	29.72	288.81	5.77	312.76	6.25	312.28
1995	44.87	33.84	11.03	2.07	42.80	2.04	42.83	1.20	43.67	4.03	40.84
1996	724.55	80.52	644.03	25.26	699.29	60.55	664.00	17.87	706.67	40.15	684.40
1997	50.88	19.36	31.52	15.36	35.52	11.51	39.37	36.68	14.20	4.51	46.37
1998	58.75	26.06	32.69	14.37	44.38	12.06	46.69	22.31	36.44	20.43	38.32
Mean[b]	239.52a	37.01b	202.50	14.14b	225.37	23.18b	216.34	16.77b	222.75	15.07b	224.44

[a] The actual annual export was determined using all stormflow and routine sample concentrations from the unharvested or harvested watershed.
[b] Means of actual values and estimates with different letters are significantly different from one another.

and 1996 than in other years. The sum of the actual annual suspended sediment exports for 1995, 1997, and 1998 were less than the individual suspended sediment exports for either 1994 or 1996 on either catchment. Reasons for the significant suspended sediment exports in 1994 and 1996 are presented in the Discussion section.

3.5. Maximum Daily Suspended Sediment Loads

Year-to-year variation in actual maximum daily loads of suspended sediment exports for both the harvested and unharvested watersheds was very large (Table V). Actual maximum daily values on the unharvested watershed ranged from approximately 2 to 312 kg ha^{-1} d^{-1}, and values on the harvested watershed ranged from approximately 8 to 336 kg ha^{-1} d^{-1}. Mean actual maximum daily loads for the two watersheds were nearly identical, approximately 96 and 106 kg ha^{-1} d^{-1}. The highest actual daily loads for 1994 and 1996 were many times greater than the highest actual daily loads for the other years (Table V). These individual days' exports constituted 45–73% of the total annual exports in 1994 and 1996 (Table IV).

All five estimates of maximum daily loads were significantly different than the actual daily loads (Table V), and almost all of the estimates of maximum daily suspended sediment loads were underestimated by all methods for all years. The maximum daily load was overestimated only twice, both times on the unharvested watershed in 1998. The closest estimated maximum daily load to the actual maximum daily load was underestimated by a factor of approximately 1.7 on both watersheds (method MR, 1995 unharvested watershed, and method W, 1995 harvested watershed), which translates to approximately 40 percent error on both watersheds. The average difference between the actual maximum daily load and estimated maximum daily loads was 92.5 and 103.9 kg ha^{-1} d^{-1}, respectively, for the unharvested and harvested watersheds, representing average percent errors of 96.8 and 97.4%, respectively. Estimates of maximum daily loads across the five methods were similar to one another. The largest maximum daily export of suspended sediment estimated by any of the methods on both watersheds was approximately 18 kg ha^{-1} d^{-1}, and only two estimates on each watershed were greater than 10 kg ha^{-1} d^{-1} for all the years.

4. Discussion

4.1. NO_3 Loads

Almost all of the estimates of annual exports of NO_3 differed from the actual annual exports by less than 2 kg ha^{-1} yr^{-1} and were not significantly different than the actual annual exports (Table II). Thus, in situations where concentrations do not fluctuate substantially with flow or time-related conditions, as supported

TABLE V

Difference between actual maximum daily suspended sediment loads and estimated maximum daily loads using the five methods described in the text for the unharvested and harvested watersheds. Harvesting occurred June–November 1995

Water year	Actual daily maximum[a]	Method									
		W		MF		MR		QF		QR	
		Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference	Estimate	Difference
	($kg\ ha^{-1}\ d^{-1}$)										
Unharvested watershed											
1994	312.24	17.19	295.05	2.06	310.18	0.92	311.32	9.69	302.55	5.90	306.34
1995	10.94	0.32	10.62	0.11	10.83	6.59	4.35	0.06	10.88	0.11	10.83
1996	150.18	5.58	144.60	1.11	149.07	11.40	138.78	0.16	150.02	1.49	148.69
1997	2.14	0.45	1.69	0.30	1.84	0.46	1.68	0.17	1.97	0.14	2.00
1998	2.31	1.17	1.14	4.68	−2.37	0.80	1.51	5.23	−2.92	0.26	2.05
Mean[b]	95.56a	4.94b	90.62	1.65b	93.91	4.03b	91.53	3.06b	92.50	1.58b	93.98
Harvested watershed											
1994	149.27	7.87	141.40	2.17	147.10	2.62	146.65	0.28	148.99	0.43	148.84
1995	30.78	18.19	12.59	0.16	30.62	0.25	30.53	0.11	30.67	0.20	30.58
1996	335.98	8.87	327.11	1.00	334.98	11.26	324.72	1.00	334.98	2.12	333.86
1997	8.13	1.46	6.67	1.06	7.07	0.42	7.71	1.47	6.66	0.36	7.77
1998	9.03	2.05	6.98	0.67	8.36	0.77	8.26	0.67	8.36	2.02	7.01
Mean[b]	106.64a	7.69b	98.95	1.01b	105.63	3.06b	103.57	0.71b	105.93	1.03b	105.61

[a] The actual maximum daily load was determined using all stormflow and routine sample concentrations from the unharvested or harvested watershed.
[b] Means of actual values and estimates with different letters are significantly different from one another.

by the small differences in minimum and maximum concentrations (Table I), frequent water quality sampling, involving storm-events, appears to be unnecessary. Furthermore, the greater difficulty and expense associated with deploying a stormflow sampling strategy compared to less intensive sampling, makes a less intensive sampling strategy more desirable.

Quarterly sampling of stream water NO_3 was sufficient to estimate annual NO_3 exports fairly accurately as long as continuous and accurate measurements of streamflow were applied to the available concentrations. When concentrations vary little with discharge, streamflow dominates load calculations (Swistock *et al.*, 1997). Even NO_3, which is one of the most seasonally varying inorganic nutrients of concern in TMDLs, did not vary substantially enough in this forested watershed to significantly alter estimates of annual totals using only quarterly sampling. Given the relatively high concentrations of NO_3 in stream water in this mid-Appalachian region (Williard *et al.*, 1997) compared to other places in the United States (Stoddard, 1994), low-intensity sampling may be adequate for many other small forested or predominantly forested watersheds in the Appalachian region and elsewhere. However, NO_3 concentrations in WS4 may be less variable during storms and seasonally than many other forested watersheds due to WS4's reported state of advanced nitrogen saturation (Stoddard, 1994; Peterjohn *et al.*, 1996). Relatively high NO_3 concentrations during both the growing and dormant season are theorized to result from nitrogen saturation, compared to more varying concentrations within-storms and among-seasons in ecosystems where nitrogen is still limiting (Stoddard, 1994). But even with the greater variability in NO_3 concentrations that might be expected in the absence of nitrogen saturation, low-intensity sampling probably would be applicable in other forested watersheds because the range of reported NO_3 (and total nitrogen) concentrations is fairly small from both undisturbed forested watersheds (Clark *et al.*, 2000; Martin *et al.*, 1986) and managed watersheds (harvests, including clearcuts). This small range of variability applies even to watersheds in the White Mountains of New Hampshire, which have the highest reported changes in NO_3 concentrations from clearcutting in North America (Vitousek and Melillo, 1979). Furthermore, because streamflow values are much larger and more important in the loading calculation than even the largest NO_3 concentrations reported from small, forested watersheds, the lack or presence of nitrogen saturation is largely a non-issue.

Non-intensive sampling also may be appropriate for estimating annual NO_3 exports from large watersheds. Review of NO_3 concentrations from 67 watersheds comprised of predominantly undeveloped land (i.e., forests, wetlands, or rangeland) in the US Geologic Survey's National Water-Quality Assessment Program (NAWQA) database (http://water.usgs.gov/nawqa/nutrients/) shows that the highest concentration observed at any site sampled was 28.78 mg NO_3 L^{-1} from a 18 904 km^2 watershed. As this concentration was fairly high, it is likely that other land uses, such as agriculture, also were important in this predominately-forested watershed. Given that the instantaneous flow of 0.16 m^3 s^{-1} on WS4 dominated the loading calculation when the highest NO_3 concentration was 7.08 mg NO_3 L^{-1}, the flow of 9.20 m^3 s^{-1} from this large NAWQA watershed should overwhelmingly dominate the loading calculation, even with a concentration that is about 4.5 times greater than that observed for the smaller WS4.

However, the assumption that non-intensive sampling is adequate for NO_3 on a wider scale and on any specific larger watershed should be verified. In watersheds containing substantial portions of agricultural or developed land, this assumption may not be true, since concentration changes with time or runoff events could be substantial. For example, NAWQA data for 89 predominantly agricultural and 27 mixed agricultural/urban watersheds, respectively, show much higher maximum concentrations of approximately 160 and 206 mg NO_3 L^{-1} with minimum concentrations of 0 mg NO_3 L^{-1}. The corresponding instantaneous flows associated with these highest concentrations were 1.02 and 0.14 m^3 s^{-1}. The wide range of concentrations coupled with the disproportionately larger maximum concentrations and similar or only slightly higher flows on these large non-forested watersheds may mean that more intensive sampling is necessary to adequately characterize annual exports from different land uses.

Even if historical or current water quality sampling shows little variation in concentrations of ohe or more chemical constituents with time or flow, this behavior should be verified periodically. Changes in concentrations can occur from land use changes, natural ecosystem disturbances, and other factors, and re-verification will assure that assumptions being employed in TMDL assessment remain properly based.

4.2. Suspended Sediment Loads

In contrast to NO_3 concentrations, suspended sediment concentrations on the unharvested and harvested watersheds varied greatly (Table I), and none of the methods closely estimated the actual annual or daily maximum suspended sediment loads. One might make the argument that employment of the same random sample in the annual and daily load calculations using monthly and quarterly samples biased the results and a Monte Carlo type of simulation might show that other selected samples could have improved the results. However, any improvement would have been minimal, given that large differences still were found between the actual loads and estimated loads when 52 weekly samples were used. Thus, running the analysis multiple times using different random samples would have resulted in little improvement in the monthly and quarterly estimates since only 12 and 4 samples, respectively, would have been used in the calculations.

The harvesting treatment seems to have had little effect on either the actual annual or maximum daily suspended sediment exports. This could be attributable to the significant natural variability in suspended sediment concentrations that may overshadow any treatment effects. The only year in which harvesting and road construction may have increased suspended sediment availability was 1996, the first water year after harvesting was completed (Tables IV and V). Suspended sediment exports also were very high on the unharvested watershed in 1996, but the yields from the unharvested catchment in 1996 were less than those in 1994, while those from the harvested watershed in 1996 were more than double those in 1994. Very large storms occurred in both years. Potential harvesting and road effects, however, were not expressed during postharvest years that did not have large runoff events.

To a large degree, suspended sediment concentrations and yields on these watersheds, as well as most forested watersheds, are controlled predominately by the energy available in discharge. Suspended sediment concentrations and yields increase dramatically in response to individual high stormflows (Figure 1). For example, Kochenderfer *et al.* (1997) reported that 84% of suspended sediment yield during one year from a 39 ha forested watershed was attributed to a single storm with a 47 yr return period. Total suspended sediment exported during this individual storm exceeded cumulative suspended sediment losses over the subsequent 5 yr.

The extremely high suspended sediment loads, both annually and daily in 1994 and 1996 were principally the result of two very high flows associated with regional flood events. Flow duration and partial-duration flood series values (Table VI) were calculated for the dates and corresponding flows used to determine exports. Streamflow data from WS4 were used for the calculation of these parameters to illustrate the extreme nature of these flows because 50 yr of continuous streamflow data were available from WS4 compared to only 5 yr for the unharvested and harvested watersheds. Recurrence intervals, or return periods were calculated for the annual-maximum series instantaneous peakflows (n = 50). The partial duration series values corresponding to those annual-maximum series values were determined from the mathematical relationships between the two developed by Langbein (1960).

The large annual and daily exports for 1994 (Tables IV and V) were due predominantly to a single event, and those in 1996 were due to multiple events. The flows used in the calculation of the single maximum daily load in 1994 and 1996 corresponded to events that each had a return period of about 7.5 yr and the average daily flow was in the top approximate 0.1 percent of all recorded average daily flows (Table VI). The high flows for 1995, 1997, and 1998 had return periods associated with the average annual flood flow or bankfull stage (assuming a return period of 1.2–2.5 yr) or less for this region.

High streamflow occurred throughout 1996, resulting in contributions by three additional high flow events (Table VI) to the high total annual suspended sediment exports for 1996 on the unharvested and harvested watersheds (Table IV). The 19 July event had the highest return period and smallest percentage of time flow was equaled or exceeded (exceedence time) of any of the events occurring during sediment monitoring, but it did not result in the maximum daily load of suspended sediment. Lower suspended sediment export probably resulted from lower sediment availability in July compared to earlier in the year due to prior

TABLE VI

Exceedence times for the highest flows each water year that were used to calculate exports, and associated instantaneous peakflows and return periods for those flows[a]

Water year	Date	Mean daily flow ($m^3 s^{-1}$)	Total flow for the day (m^3)	Percent of time flow was exceeded	Associated instantaneous peakflow ($m^3 s^{-1}$)	Peakflow return period[b] (yr)
1988	6 March 1989	0.116	10060	0.140	0.162	1.91
1989	31 December 1989	0.087	7544	0.323	0.168	1.98
1990	26 May 1990	0.097	8414	0.230	0.234	2.80
1991	16 February 1992	0.074	6358	0.558	0.090	1.23
1992	24 March 1993	0.117	10101	0.139	0.201	2.37
1993	9 February 1994	0.331	28568	0.072	0.521	25.0
1994	7 May 1994	0.119	10291	0.133	0.441	7.27
1995	14 May 1995	0.026	2206	7.564	0.042	0.93
1996	19 January 1996	0.226	19485	0.073	0.450	7.66
1997	1 March 1997	0.044	3823	2.391	0.092	1.25
1998	19 June 1998	0.064	5557	0.849	0.108	1.38
Additional 1996 high flows						
1996	6 May 1996	0.088	7566	0.320	0.135	1.63
1996	19 July 1996	0.248	21457	0.073	0.498	13.1
1996	17 September 1996	0.047	4054	2.062	0.057	1.00

[a] All flows and corresponding flow duration and flood frequency calculations were made using WS4 discharges (1951–2000), because the unharvested and harvested sediment watersheds were monitored for streamflow for only five years.

[b] Return period values were determined as a partial-duration flood series because there were multiple high flow values in 1996.

repeated flushing in the January and May events as well as other frequent storms in 1996. Year 1996 was the wettest year on record (beginning in 1931) based on historical precipitation data from the National Climate Data Center (Parsons, WV site; http://lwf.ncdc.noaa.gov/oa/climate/online/coop-precip.html). Suspended sediment availability can decrease with repeated flushing, even if additional energy is provided. Thus, within-storm suspended sediment concentrations often peak prior to peakflow (Figure 1).

Overestimations of suspended sediment, by both the annual and daily calculation methods were rare. When they did occur, they were in years with low exports when no extreme, dominating stormflow events occurred. In the years with large streamflow events and accompanying high exports, estimation of maximum daily loads employing any sampling frequency that did not specifically include storm sampling always underestimated exports. Walling and Webb (1981) found estimates of annual suspended sediment yields were less than 20% of the actual annual load when only weekly sampling was employed. Including stormflow event samples improved annual suspended sediment load estimates, but even after adding hundreds of stormflow concentrations and corresponding discharge values, annual yields were still less than 80% of the actual annual loads. Meade *et al.* (1990) found that even single daily suspended sediment concentrations applied to summation methods still resulted in approximately 20–25% deviations from actual annual loads based on 15 yr of data.

Even though most states rely upon non-intensive sampling for establishment of baseline water quality levels, non-intensive sampling is clearly not a viable approach for estimating suspended sediment exports or maximum concentrations from forested watersheds. This is true even when continuous records of discharge are available. The substantial variation in suspended sediment with streamflow is too large to ignore, and it is difficult to estimate suspended sediment concentrations accurately simply based on flow. Bunte and MacDonald (1998) found that approximately 10 yr of suspended sediment measurements are needed to estimate annual suspended sediment yields to within 50% of their true value with 95% confidence. Many other factors, such as time since last storm event, precipitation characteristics, steepness of the rising limb of the hydrograph, antecedent flow, the locations and types of sediment sources, sediment storage and delivery conditions, seasonal variations in protection of hillside sediment by vegetation, the hysteresis associated with rising and falling sediment limbs, etc. (Walling, 1977; Rieger and Olive, 1986; Beschta, 1987; Goodwin and Denton, 1991; Bunte and MacDonald, 1998), result in unique sediment responses for each storm (Bunte and MacDonald, 1998) that are difficult to predict or model accurately. The substantial natural variability in suspended sediment concentrations may make it difficult, if not impossible, to determine if improvements in water quality are achieved in impaired streams by additional best management practices (BMP) implementation during silvicultural activities; ranges in suspended sediment concentrations are simply so large that they may overshadow any improvement of nonpoint silvicultural sources.

Suspended sediment concentrations from large watersheds appear to be equally variable with time and flow, and thus, one would assume just as difficult from which to estimate exports. The five undeveloped NAWQA watersheds that also were monitored for suspended sediment concentrations showed substantial variation in the observed concentrations. Concentrations ranged from 0–3870 mg L^{-1} for a 259 km^2 watershed, 1–614 mg L^{-1} for a 89 km^2 watershed, 1–824 mg L^{-1} for a 1720 km^2 watershed, 0–643 mg L^{-1} for a 502 km^2 watershed, 1–1180 mg L^{-1} for a 583 km^2 watershed, and 0–46 mg L^{-1} for a 14 km^2 watershed. The latter had minimal variation, but a large number of missing values, so its actual range was probably greater than 46 mg L^{-1}.

Due to the large error involved in estimating suspended sediment loads, some states are reverting to surrogate measures in place of or in addition to water column measurements of suspended sediment. These surrogates include size of streambed substrate, channel geomorphological metrics, and biological (fish and invertebrates) population and diversity measures (U.S. Environmental Protection Agency, 1999). These metrics range in their sensitivity for detecting changes in water quality. However, if the TMDL is aimed specifically at decreasing sediment in the water column (e.g., for drinking water, etc.), only direct measurements of turbidity or suspended sediment are applicable and should be used to determine impairment levels; surrogates may not be able to identify if water column impairment exists during critical times, such as storm events, or the surrogates may be affected by some other condition not related to the TMDL constituent.

Developing a better understanding of the processes controlling suspended sediment within storms in different types of systems (e.g., high gradient headwater streams, lower gradient rivers, etc.) may go a long way toward development of process-driven models in which suspended sediment during various stormflows could be estimated accurately. Without such understanding and the resulting process-oriented tools, subsequent source assessment estimates of suspended sediment via modeling likely will be inaccurate, which in turn will make realistic TMDL allocation to different land uses or land owners difficult if not impossible. Until such tools become available, high frequency sampling that specifically includes intensive episodic sampling and high-resolution streamflow records are the needed parameters for accurately estimating suspended sediment loads (Walling and Webb, 1981, 1982, 1987, 1988; Thomas, 1985; Thomas and Lewis, 1993).

5. Conclusion

Establishing accurate estimates of baseline loads in impaired streams for TMDL implementation will be necessary to achieve water quality objectives. Data from a small, forested watershed showed that quarterly sampling of NO_3 concentrations was sufficient to obtain estimates of baseline conditions that were not statistically different than actual maximum daily loads or total annual loads. This success was due to the small variability in NO_3 concentrations over time, including during stormflow, and the availability of accurate and continuous streamflow data. Evaluation of NAWQA data suggest that maximum daily and total annual NO_3 loads also can be estimated relatively accurately from infrequently collected data on larger watersheds dominated by undeveloped land. Hence, any pollutant that does not have wide concentration variations with time and discharge probably can be estimated accurately with quarterly sampling and possibly less frequent sampling, if streamflow measurements are available.

By contrast, establishment of accurate baseline loads for suspended sediment in forested watersheds will be unlikely without intensive monitoring that includes stormflow sampling. Even weekly sampling was insufficient to provide statistically similar maximum daily or total annual loads compared to the actual values. The presence of harvesting and road construction activities may have had some short-term negative effects on already poor estimation capabilities using non-intensive sampling, though it is impossible to separate the silvicultural influences from flood-related responses during the first year after harvesting. There were no effects attributable to harvesting on suspended sediment load calculations after the first year. The large variability in suspended sediment concentrations that occurs naturally within and among years may make it impossible to measure any improvement achieved by additional BMP implementation from nonpoint silvicultural pollution sources.

References

Beschta, R. L.: 1987, 'Conceptual models of sediment transport in streams', in *Sediment Transport in Gravel-bed Rivers*, John Wiley, New York, pp. 387–419.

Bunte, K. and MacDonald, L. H.: 1998, 'Scale Considerations and the Detectability of Sedimentary Cumulative Watershed Effects', *Report* submitted to the U.S. Forest Service and the National Council for Air and Stream Improvement of the Paper Industry Council, Colorado State University, Fort Collins, CO, p. 327.

Clark, G. M., Mueller, D. K. and Mast, M. A.: 2000, 'Nutrient concentrations and yields in undeveloped stream basins of the United States', *J. Amer. Water Resour. Assoc.* **36**, 849–860.

Edwards, P. J., Kochenderfer, J. N. and Wood, F.: 2001, 'Using Specific Conductance To Predict Stream Water Chemistry During Storms', in *AWRA Spring Specialty Conference, Water Quality Monitoring and Modeling*, 30 April–2 May 2001, San Antonio, TX, pp. 25–30.

Edwards, P. J. and Wood, F.: 1993, 'Field and Laboratory Quality Assurance/Quality Control Protocols and Accomplishments for the Fernow Experimental Forest Watershed Acidification Study', USDA For. Serv. NEFES, *Gen. Tech. Rep. NE-177*, Radnor, PA, p. 15.

Eschleman, K. N., Morgan II, R. P., Webb, J. R., Deviney, F. A. and Galloway, J. N.: 1998, 'Temporal patterns of nitrogen leakage from mid-Appalachian forested watersheds: Role of insect defoliation', *Water Resour. Res.* **34**, 2005–2016.

Goodwin, P. and Denton, R. A.: 1991, 'Seasonal Influences on the Sediment Transport Characteristics of the Sacramento River, California', in *Proceedings of the Institute of Civil Engineers, Part 2 – Research and Theory* **91**, pp. 163–172.

Houck, O. A.: 1999, *The Clean Water Act TMDL Program: Law, Policy, and Implementation*, Environmental Law Institute, Washington, DC, p. 564.

Keith, L. H. (ed.): 1991, *Compilation of E.P.A.'s Sampling and Analysis Methods*, Lewis Publishers, Inc., Boca Raton, FL, p. 803.

Kochenderfer, J. N. and Edwards, P. J.: 1990, 'Design and Construction of a Low-Cost Stream-Monitoring Shelter', USDA For. Serv. NEFES, *Gen. Tech. Rep. NE-135*, Broomall, PA, p. 10.

Kochenderfer, J. N., Edwards, P. J. and Wood, F.: 1997, 'Hydrologic impacts of logging an Appalachian watershed using West Virginia's best management practices', *North. J. Appl. For.* **14**, 207–218.

Langbein, W. B.: 1960, 'Plotting Positions in Frequency Analysis', USGS Water Supply Pap. 1543-A, pp. A48–A51.

Martin, C. W., Hornbeck, J. W., Likens, G. E. and Buso, D. C.: 2000, 'Impacts of intensive harvesting on hydrology and nutrient dynamics of northern hardwood forests', *Can. J. Fish. Aquat. Sci.* **57**, 19–29.

Martin, C. W., Pierce, R. S., Likens, G. E. and Bormann, F. H.: 1986, 'Clearcutting Affects Stream Chemistry in the White Mountains of New Hampshire', USDA For. Serv. Northeast. For. Exp. Sta. Res. Pap. NE-579, p. 12.

Meade, R. H., Yuzyk, T. R. and Day, T. J.: 1990, 'Movement and Storage of Sediment in Rivers of the United States and Canada', in *The Geology of North America. Vol. 0-1, Surface Water Hydrology*, The Geological Society of America, Boulder, CO, pp. 255–280.

Peterjohn, W. T., Adams, M. B. and Gilliam, F. S.: 1996, 'Symptoms of nitrogen saturation in two central Appalachian hardwood forest ecosystems', *Biogeochem.* **35**, 507–522.

Reinhart, K. G., Eschner, A. G. and Trimble Jr., G. R.: 1963, 'Effect on Streamflow of Four Forest Practices in the Mountains of West Virginia', USDA For. Serv. Res. Pap. NE-1, p. 79.

Rieger, W. A. and Olive, L. J.: 1986, 'Sediment Responses During Storm Events in Small Forested Watersheds', in *Statistical Aspects of Water Quality Monitoring*, Elsevier, Oxford, England, pp. 490–498.

SAS Institute, Inc.: 1988, *SAS/STAT User's Guide, Release 6.03 Edition*, SAS Institute, Inc. Cary, North Carolina, p. 1028.

Stoddard, J. L.: 1994, 'Long-Term Changes in Watershed Retention of Nitrogen: Its Causes and Consequences', in *Environmental Chemistry of Lakes and Reservoirs. Advances in Chemistry Series Number,* **237**, American Chemical Society, Washington, DC, pp. 223–284.

Swistock, B. R., Edwards, P. J., Wood, F. and DeWalle, D. R.: 1997, 'Comparison of methods for calculating solute exports from six forested Appalachian watersheds', *Hydrol. Proc.* **11**, 655–669.

Thomas, R. B.: 1985, 'Measuring Suspended Sediment in Small Mountain Streams', USDA For. Serv., Pacific Southwest Forest and Range Experiment Station, *Gen. Tech. Rep. PSW-83*, p. 9.

Thomas, R. B.: 1991, 'Systematic Sampling for Suspended Sediment', in *Proceedings of the Fifth Federal Interagency Sedimentation Conference. Las Vegas, NV*, Interagency Advisory Committee on Water Data, Subcommittee on Sedimentation, pp. 2.17–2.24.

Thomas, R. B. and Lewis, J.: 1993, 'A comparison of selection at list time and time-stratified sampling for estimating suspended sediment loads', *Water Resour. Res.* **29**, 1247–1256.

U.S. Environmental Protection Agency: 1999, 'Protocol for Developing Sediment TMDLs', United States Environmental Protection Agency, Washington, DC, EPA 841-B-99-004, p. 102 + appendices.

Vitousek, P. M. and Melillo, J. M.: 1979, 'Nitrate losses from disturbed forests: Patterns and mechanisms', *For. Sci.* **25**, 605–619.

Walling, D. E.: 1977, 'Limitations of the Rating Curve Technique for Estimating Suspended Sediment Loads, with Particular Reference to British Rivers', in *Erosion and Solid Matter Transport in Inland Waters*, IAHS Publ. No. 122, pp. 34–48.

Walling, D. E. and Webb, B. W.: 1981, 'The Reliability of Suspended Sediment Load Data', in *Erosion and Sediment Transport Measurement*, IAHS Publ. No. 133, pp. 177–194.

Walling, D. E. and Webb, B. W.: 1982, 'Sediment Availability and the Prediction of Storm Period Sediment Yield', in *Recent Developments in the Explanation and Prediction of Erosion and Sediment Yield*, IAHS Publ. No. 137, pp. 327–337.

Walling, D. E. and Webb, B. W.: 1987, 'Suspended Load in Gravel-bed Rivers: U.K. Experience', in *Sediment Transport in Gravel-bed Rivers*, John Wiley, New York, pp. 691–732.

Walling, D. E. and Webb, B. W.: 1988, 'The Reliability of Rating-Curve Estimates of Suspended Sediment Load: Some Further Comments', in *Sediment Budgets*, IAHS Publ. No. 174, pp. 337–350.

Williard, K. W. J., DeWalle, D. R., Edwards, P. J. and Schnabel, R. R.: 1997, 'Indicators of nitrate export from forested watersheds of the mid-Appalachians, United States of America', *Global Biogeochem. Cycles.* **11**, 649–656.