MINIMUM FLOW CONSIDERATIONS FOR AUTOMATED STORM SAMPLING ON SMALL WATERSHEDS

R. Daren Harmel, Kevin W. King, June E. Wolfe* and H. Allen Torbert**

USDA-ARS, 808 E. Blackland Road, Temple, Texas 76502
**TAES, 720 E. Blackland Road, Temple, Texas 76502 and*
***USDA-ARS, P.O. Box 3439, Auburn, Alabama 36830*

Abstract.—The issue of a minimum flow threshold (also referred to as enable level) above which to trigger sampling plays an important role in water quality sampling projects; however, guidance on developing appropriate storm sampling strategies for small streams is limited. As a result, arbitrary strategies are used that may not accurately characterize pollutant flux. Therefore, the objectives of this study were to: (1) compare measured nutrient flux data to hypothetical results collected under several alternative minimum flow threshold or enable level scenarios and (2) publish initial guidance on setting minimum flow thresholds for automated storm sampling in small watersheds. Comparison of measured nutrient fluxes for various enable level scenarios illustrated that substantial error is introduced even with relatively small enable level increases. Based on these results, minimum flow thresholds for automated sampling equipment should be set such that even small storms with small increases in flow depth are sampled. In order to manage the number of samples collected, enable levels should be raised only after careful consideration of the resulting consequences. Alternatives for decreasing the number of samples in nutrient flux measurements, such as increasing the time or flow volume between samples or compositing several samples into one collection bottle, introduce substantially less error than does increasing minimum flow thresholds.

Monitoring water quality during storm events is becoming increasingly important in characterization of pollutant loading to water bodies, especially as National Water Quality Inventories (USEPA 1995; USEPA 2000) continue to report that nonpoint source (NPS) pollution adversely impacts rivers, lakes and coastal waters. NPS pollution includes runoff from diffuse sources such as urban areas, farms, and silvicultural operations. Excessive anthropogenic NPS inputs of the macro-nutrients, nitrogen (N) and phosphorus (P) or "cultural eutrophication" can create accelerated algal growth which degrades aquatic ecosystem health, increases water treatment costs and diminishes recreational and aesthetic values (Kolbe & Luedke 1993).

The traditional monitoring focus on periodic grab sampling of low flows to characterize point source pollution (discharged from specific locations such as factories and waste water treatment plants) is now often coupled with automated storm flow monitoring to characterize NPS pollution. Most commercially available automated samplers contain similar components, including: programmable operation and memory, water level recorder, sample collection pump and sample bottles. Typical storm sampling operation involves setting a minimum flow threshold or enable level to start and finish sampling (either a flow depth or a rainfall depth per specified time) and setting a time or flow interval on which to collect samples after the sampler is triggered. This type of automated storm monitoring is often the cornerstone of small watersheds projects whose objectives are to compare water quality impacts of various land management activities, evaluate water quality improvement following implementation of best management practices and determine annual pollutant fluxes for Total Maximum Daily Load (TMDL) projects (Tate et al. 1999; Robertson & Roerish 1999).

On small watershed monitoring projects, however, sampling and funding considerations, along with NPS variability, often make it difficult to achieve project objectives (Tate et al. 1999). Budget determination is generally the first step in monitoring projects (Shih et al. 1994). Most sampling proposals specify a maximum number of storms that will be sampled or a maximum number of samples that will be collected, so that a reasonable sampling expectation can be met. Service and maintenance of automated sampling equipment is labor intensive and expensive, and cost considerations often limit the number of samples that can be collected and analyzed (Robertson & Roerish 1999; Dissmeyer 1994). Another consideration in developing a sampling scheme is the number of samples that can be collected and analyzed by a laboratory in a reasonable time frame (Novotny & Olem 1994). Since a large portion of the cost of a monitoring program is directly related to the number of samples, determination of a proper minimum flow threshold and sample frequency is important in achieving objectives within budget limitations. A high minimum flow threshold and/or low frequency sampling bypasses important information and may lengthen the project duration (Novotny & Olem 1994; Shih et al. 1994). However, a low minimum flow threshold and/or high frequency sampling may be inhibited by available financial and laboratory resources.

Guidance on developing storm sampling strategies for small streams is limited, but examples for larger perennial streams and rivers are presented by Robertson & Roerish (1999). The United States Geological Survey NPS program in Wisconsin collects 100 to 200 fixed interval grab samples and storm flow samples per year for small streams (watersheds less than 100 km^2). The typical National Water Quality Assessment strategy collects monthly samples supplemented by four to eight storm samples per year for about 2.5 yr. For larger streams and rivers, precision and accuracy increase with sampling frequency in almost all cases. In smaller watersheds, which are typically more variable in their response than larger ones, more intensive sampling is generally needed to achieve precise and accurate load estimates (Richards & Holloway 1987).

Comparisons of specific automated sampling alternatives are also limited. However, issues of discrete (one sample per bottle) versus composite sampling (several samples per bottle) and flow-weighted (based on flow volume) versus time-weighted sampling (based on time intervals) have been addressed by King & Harmel (2001); Shih et al. (1994); Miller et al. (2000) and others. One important question that has not received attention is what storm size should be sampled, which translates into how many storms are sampled. As stated earlier, this issue of a minimum flow threshold above which to trigger sampling plays an important role in developing sampling strategies. However, without published studies on the impact of setting enable levels, arbitrary decisions are made. General guidance on this issue indicated that for determination of annual storm loads, storms with rain exceeding 25 mm/hr or runoff exceeding 13 mm should be sampled and that generally three to five storms per year create about 75% of the annual runoff (Slade pers. comm.). Tate et al. (1999) state that a majority of annual flow and NPS loading occurs during four to six storms per year on California rangelands. For large rivers, commonly as much as 80% of annual NPS load is contributed by 20% of flows (Richards & Holloway 1987).

Richards & Holloway (1987) indicated that assessment of the adequacy of sampling programs for large rivers is needed. That need also exists for small streams, especially since numerous small watershed monitoring programs are underway with limited assessment of sampling program adequacy. No published guidance is available on setting minimum flow thresholds. If they are set too low, samples will be taken on every runoff event even though no significant NPS load is transported. In this case, analysis cost and personnel time will be wasted. If enable levels are set too high, substantial portions of runoff events and

Table 1. Characteristics of watershed study sites.

	Traditional	Precision	Airport	Mixed Urban
Area	5.7 ha	9.1 ha	37.5 ha	66.5 ha
Slope	1 - 5%	1 - 5%	1 - 4%	1 - 8%
Soil texture	Clay	Clay	70% Impervious, silty clay to sandy clay loam	12% Impervious, silty clay to sandy clay loam
Landuse	Corn	Corn	Airport	Airport, golf course, residential
Land management	Conventionally-applied fertilizer, terraces, residue management	Precision applied fertilizer, terraces, residue management	Mowing, limited fertilizer and pesticide use	Mowing, aeration, moderate fertilizer and pesticide use, irrigation
Flow channel	Ephemeral - grass waterway	Ephemeral - grass waterway	Small perennial stream	Irrigation return flow supplements small perennial stream

possibly entire events will not be sampled, thus valuable information will be missed. Therefore, the objectives of this study were to: (1) compare measured NO_3+NO_2-N load data to hypothetical load data collected under various enable level scenarios and (2) produce initial guidance on setting minimum flow thresholds for automated storm sampling in small watersheds.

Materials and Methods

Study site description.—Runoff and water quality data from two nutrient load studies on four watersheds ranging from 5.7 to 66.5 ha in central Texas were used in this analysis (Table 1). Two were agricultural watersheds located 3 km east of Temple, Texas, and two were urban watersheds in Austin, Texas. The Austin/Temple area receives 813 to 889 mm normal annual precipitation, has an average of 273 growing season days per year, and average maximum daily temperatures from 15°C in January to 35°C in August (NOAA 1999).

Flow measurement and water quality sampling.—To monitor surface runoff on the agricultural watersheds near Temple, Texas, a 0.61 m H-flume equipped with an ISCO 4230/3700 flow meter and sampler system was installed at the "outlet" of each field. An ISCO 674 rain gauge and two HOBO rain event recorders were also installed on site to record rainfall data. From February 1999 through January 2001, flow rates were recorded every five minutes during runoff events. Time-weighted, composite samples with four 200 mL samples per bottle were collected automatically during runoff events. Samplers were programmed to sample all runoff events with adequate flow depth to submerge the sampler intake (approximately 38 mm water depth) and allow sample collection. To provide adequate resolution in short duration events and adequate sampling capacity for longer events, samples were taken in five min intervals for 65 min, 15 min intervals for the next 660 min, and 30 min intervals for the final 1200 min.

Similar monitoring strategies were used to measure surface runoff on the urban sites in Austin, Texas. An ISCO 6700 automatic sampler, an ISCO 4150 area velocity flow logger, and an ISCO 674 rain gauge were installed at each site. Two round culverts drain the airport site, and a box culvert drains the mixed urban site. From April 1998 through March 2000, flow rates were recorded every 15 minutes during runoff events. Time-weighted composite samples with six 150 mL samples per bottle were collected automatically during runoff events. As with the agricultural sites, samplers were programmed to sample runoff events with adequate flow depth to submerge the sample intake (38 mm water depth) and allow sample collection. Samples were taken at five min intervals for 120 min, 15 min intervals for the next 720 min, 30 min intervals for the next 1440 min, and 60 min intervals for the next 1440 min.

Samples were collected within 48 hr of runoff events, acidified, iced and transported to the laboratory where they were stored at 4°C prior to analysis. Samples were analyzed for dissolved nitrate plus nitrite nitrogen (NO_3+NO_2-N) concentrations using a Technicon Autoanalyzer IIC (Technicon Instruments Corp., Tarrytown, New York) and colorimetric methods published by Technicon Industrial Systems (1973).

For each of the four watersheds, measured dissolved NO_3+NO_2-N loads were determined by multiplying measured nutrient concentrations by corresponding flow volumes and summing these incremental loads for the duration of the runoff event. This measured load was then compared to loads that would have been measured for increased enable levels. For

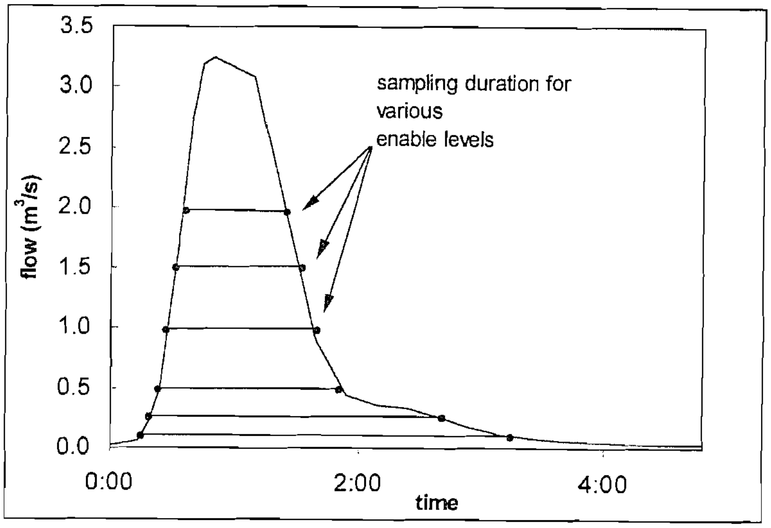

Figure 1. Sample storm illustrating duration of sample collection for various enable levels.

the 0.61 m H-flumes on the agricultural watersheds, increased enable levels ranged from 38 to 305 mm (0.001 to 0.06 m^3/s). Increased enable levels ranged from 137 to 762 mm (0.02 to 0.49 m^3/s) for the airport site and from 519 to 1067 mm (0.04 to 1.06 m^3/s) for the mixed urban site. An example storm is presented in Figure 1 to illustrate the duration of sample collection for various hypothetical enable levels.

Results and Discussion

Runoff events.—Dissolved NO_3+NO_2-N loads for each site were analyzed for a total of 122 measured runoff events over two years. A summary of rainfall and runoff data for events in which samples and flow rate data were collected from both sites in the urban and agricultural watersheds is presented in Table 2. A wide range of rainfall depths and intensities, runoff volumes, and peak flow rates occurred during the study period.

Results from this study match well with information provided by Slade (pers. comm.) and other studies such as Tate et al. (1999) that generally report that three to six events per year create about 75% of the annual storm runoff and NPS load. Our results for these study sites

Table 2. Properties of rainfall and runoff events.

	Traditional	Precision	Airport	Mixed Urban
Number of runoff events	24	18	40	40
Peak flows (m^3/s)	0.00 - 0.32	0.00 - 0.39	0.02 - 4.91	0.03 - 9.82
Runoff volumes (m^3)	0.14 - 946	0.15 - 2260	39 - 77000	109 - 89000
Runoff depths (mm)	0.00 - 36	0.00 - 25	0.10 - 205	0.20 - 134
Rainfall (mm)	8 - 63	9 - 63	5 - 227	4 - 187
Max 15 min rainfall (mm)	19	19	26	26

showed that three to six events per year produced on average from 74 to 87% of the NO_3+NO_2-N load and that between 64 and 100% of the annual load could have been captured by sampling only the largest six storms each year (Table 3).

As enable levels increase, an increasing amount of pollutant flux is not captured; therefore, increasing enable levels results in increased error compared to the true or total load. To quantify these increases, relative errors (percent deviation from the total measured load) and absolute errors (magnitude of deviation from the total measured load) were calculated. Figures 2 and 3 illustrate that errors increase rapidly as enable levels increase, especially for the smaller watersheds. Errors for the smaller agricultural watersheds were substantial even for small increases in enable levels because small increases in enable level resulted in relatively large increases in flow rates and because large changes in nutrient concentration occurred during the storms events; therefore, substantial flow volume and nutrient flux were not sampled with increased minimum flow thresholds.

In most water quality sampling projects, appropriate sampling to adequately measure loads must be conducted within the constraint of limited project resources. To reduce analysis costs and overcome laboratory time and personnel limitations, the number of samples can be managed by raising enable levels, increasing duration or flow volume between samples and/or compositing several samples together. However, when each of these adjustments are made errors in pollutant flux measurements increase. Based on the results of this study and comparisons to King & Harmel (2001), enable levels should be raised only after

Table 3. Annual NO_3+NO_2-N loads determined by measuring the largest storm events.

Measure only the Largest (Number of events)	Percent of Measured Annual Load		
	(Average)	(Standard deviation)	(Range)
1	49	± 25	16 - 89
2	64	± 26	27 - 97
3	74	± 23	38 - 98
4	80	± 20	49 - 100
5	84	± 17	58 - 100
6	98	± 15	64 - 100

careful consideration of the resulting consequences, since small increases in enable levels resulted in large errors. King & Harmel (2001) showed that increasing the duration between samples from 5 min to 15 min, which reduced the number of samples by 66%, resulted in less than 1% average increases in relative error. Even when samples were composited up to six samples per bottle, which further reduced samples numbers by 83%, less than 20% average increases in relative error occurred. In contrast to relatively small increases in relative error for increased duration and flow volume presented by King & Harmel (2001), relative errors increased rapidly when minimum flow thresholds were raised for the watersheds in this study. Figure 4 illustrates that less error is introduced with corresponding reduction in sample numbers by increasing duration or flow volume between samples, with further reduction possible with composite sampling. This figure presents the most valuable result of these analyses: alternative strategies are recommended over raising minimum flow thresholds. Minimum flow thresholds should be set at low levels, such that even small storms with small increases in flow depth are sampled. On watersheds of the size studied (6 to 67 ha), minimum flow thresholds of 0.001 to 0.04 m^3/s are recommended.

CONCLUSIONS

As human population grows and water resources increase in value from a water supply and an aquatic ecosystem standpoint, accurate characterization of water quality will become more important. In order to correctly quantify total water quality constituent fluxes, the traditional methodology of periodic low flow grab sampling to characterize point sources must be coupled with storm flow monitoring to characterize

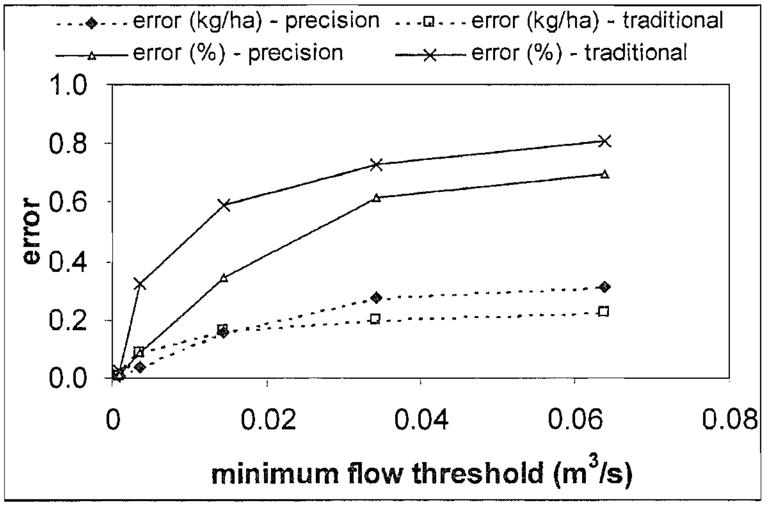

Figure 2. Relative and absolute errors the small agricultural watersheds for various minimum flow thresholds.

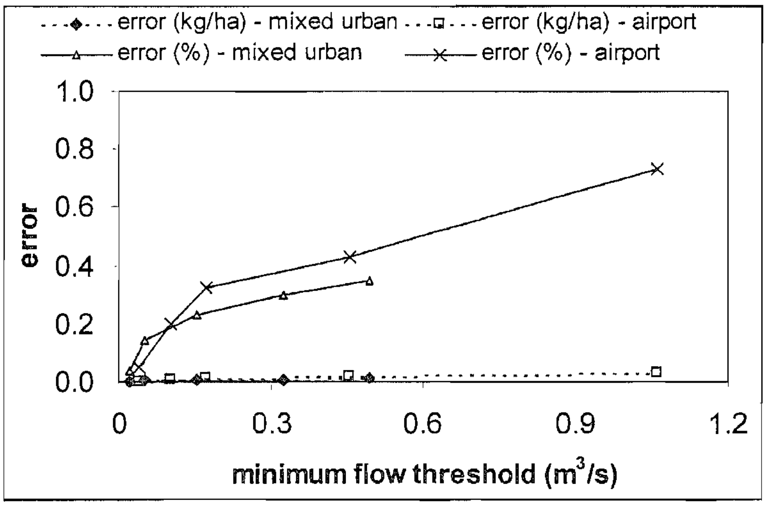

Figure 3. Relative and absolute errors the larger urban watersheds for various minimum flow thresholds.

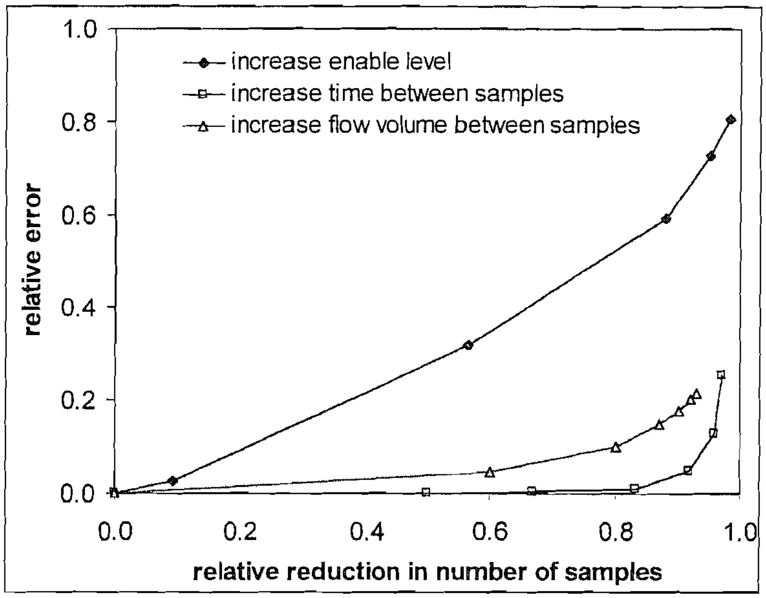

Figure 4. Comparison of relative errors for various sampling strategies to manage the number of samples.

nonpoint sources. Since guidance on developing storm sampling strategies for small streams is limited especially in light of resource constraints in most monitoring projects, appropriate guidance is needed to develop sampling strategies that accurately characterize pollutant flux within budget resources. Guidance such as presented in this study should assist monitoring program developers in setting minimum flow thresholds for automated storm sampling in small watersheds.

Comparison of measured nutrient fluxes to hypothetical fluxes collected under various enable level scenarios in this study showed that substantial error is introduced as minimum flow thresholds are increased. Based on this comparison, minimum flow thresholds for automated sampling equipment should be programmed such that even small storms with small increases in flow depth are sampled. On smaller watersheds, minimum flow thresholds of 0.001 to 0.04 m^3/s are recommended. In order to manage the number of samples collected, enable levels should not be raised above these levels without careful consideration of consequences. Alternatives for managing sample numbers, such as increasing the time or flow volume between samples, or compositing several samples, introduce substantially less error in nutrient flux measurements for the watersheds studied.

ACKNOWLEDGMENTS

We would like to thank Kirk Dean, PhD, Principal Scientist, Parsons Engineering Science; Christine Kolbe, Aquatic Scientist, Texas Natural Resource Conservation Commission (TNRCC); and Roger Miranda, Geochemist, TNRCC for review of an earlier version of this manuscript. Raymond M. Slade, Jr., hydrologist and Texas District Surface-Water Specialist for the U.S Geological Survey, also deserves credit for providing valuable insight and information on the subject of storm water sampling. Trade names in this manuscript are included for the benefit of the reader and do not imply endorsement by USDA.

LITERATURE CITED

Dissmeyer, G. E. 1994. Evaluating the effectiveness of forestry best management practices in meeting water quality goals or standards. USDA Forest Service, Southern Region, Atlanta, Georgia Misc. Pub. 1520.

King, K. W. & R. D. Harmel. 2001. Considerations in selecting a water quality sampling strategy. ASAE Paper No. 01-2134. St. Joseph, Michigan: ASAE.

Kolbe, C. M. & M. W. Luedke. 1993. A guide to freshwater ecology. Texas Natural Resource Conservation Commission. GI-34.

Miller, P. S., B. A. Engel & R. H. Mohtar. 2000. Sampling theory and mass load estimation from watershed water quality data. ASAE Paper No. 00-3050. St. Joseph, Michigan: ASAE.

National Oceanic and Atmospheric Administration. 1999. Climatological Data Annual Summary Texas 1999. Vol 104. No. 13.

Novotny, V. & H. Olem. 1994. Water quality: prevention, identification, and management of diffuse pollution. New York, Van Nostrand Reinhold, 1072 pp.

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

Robertson, D. M. & E. D. Roerish. 1999. Influence of various water quality sampling strategies on load estimates for small streams. Water Resources Research, 35(12):3747-3759.

Shih, G., W. Abtew & J. Obeysekera. 1994. Accuracy of nutrient runoff load calculations using time-composite sampling. Transactions of ASAE, 37(2):419-429.

Tate, K. W., R. A. Dahlgren, M. J. Singer, B. Allen-Diaz & E. R. Atwill. 1999. Timing, frequency of sampling affect accuracy of water-quality monitoring. California Agriculture, 53(6):44-49.

Technicon Industrial Systems. 1973. Nitrate and nitrite in water and waste water. Industrial method no. 100-70w. Technicon Instruments Corp., Tarrytown, New York, 3 pp.

United States Environmental Protection Agency. 1995. National water quality inventory 1994 report to Congress. USEPA 841-R-95-005. USEPA, Office of Water, Washington,

D.C., 572 pp.
United States Environmental Protection Agency. 2000. National water quality inventory 1998 report to Congress. USEPA 841-R-00-001. USEPA, Office of Water, Washington, D.C., 413 pp.

RDH at: dharmel@brc.tamus.edu