16 Small Watershed Data Collection to Support Phosphorus Modeling

Daren Harmel
U.S. Department of Agriculture-Agricultural Research Service, Temple, TX

Brian E. Haggard
University of Arkansas, Fayetteville, AR

CONTENTS

16.1 Introduction 383
16.2 Project Design Factors 384
    16.2.1 Monitoring Resources 385
    16.2.2 Flow Characterization 386
    16.2.3 Water-Quality Characterization 388
        16.2.3.1 Base Flow and Low Flow 388
        16.2.3.2 Storm Flow 389
    16.2.4 Automated Storm Sampling Settings 390
        16.2.4.1 Storm Sampling Threshold 391
        16.2.4.2 Sampling Interval 391
        16.2.4.3 Discrete vs. Composite Sample Collection 395
    16.2.5 Alternative Procedures (Regression Methods) 395
16.3 Uncertainty in P Transport Measurement 396
16.4 Summary 398
References 399

16.1 INTRODUCTION

Research and development of improved phosphorus (P) modeling methods is often hampered by the lack of adequate data on P transported in runoff from various soil and land use conditions. These data are needed to enhance model representation of soil P cycling, off-site transport, and linkages to downstream impacts. Such enhancements are necessary because models are increasingly used to guide legal, regulatory, and programmatic decisions, which directly affect farm income, water-supply protection, and ecological sustainability. Because of these implications, modelers must incorporate state-of-the-art science to accurately represent P mechanisms and to provide corresponding uncertainty estimates, both of which require appropriate P transport data for model calibration and evaluation (Sharpley et al. 2002).

The relative lack of water quality and corresponding flow data is attributed to collection difficulties involving natural rainfall variation, adverse weather conditions, travel time, field personnel requirements, and equipment maintenance (Beaulac and Reckhow 1982; Gilley and Risse 2000; Harmel et al. 2003). The resource requirements of discharge data collection and water quality sampling and analysis also limit availability of transport data (Agouridis and Edwards 2003; McFarland and Hauck 2001; Robertson and Roerish 1999; Shih et al. 1994). As a result, few researchers have made the commitment needed to adequately monitor P transport.

Many monitoring projects have been recently initiated, or existing projects modified, to provide targeted water resource data in response to water-quality concerns. The paramount objective in typical project design and modification is to accurately characterize runoff and water quality within resource constraints. The major considerations that affect accomplishment of this objective are discussed in this chapter. Automated storm water quality sampling techniques receive particular attention because most monitoring projects designed to support P modeling are assumed to utilize automated samplers to characterize P transport in surface runoff (for discussion see Section 16.2.3). Other topics addressed include: monitoring resources, flow measurement, manual base-flow and storm sampling, and alternative methods.

The influence of scale on P transport mechanisms is well established (Sharpley et al. 2002), but the categorization of watershed scales is difficult due to variable sizes, as determined by hydroclimatic setting and arbitrary selection of watershed outlet. With this variability in mind, the methods discussed are generally applicable for data collection at field-scale (< 50 ha) and small watershed (< 10,000 ha) sites. Discharge and water-quality characterization at the basin-scale are not addressed because most agencies and projects are not adequately staffed or equipped to collect data at that scale.*

16.2 PROJECT DESIGN FACTORS

The following Sections 16.2.1–16.2.5 discuss project design factors that directly affect the tradeoff between accurate transport data and monitoring resources. Specifically, they determine the quality and quantity of collected P transport data and supporting flow characterization data.

This chapter integrates and summarizes the extensive, well-established information on preferred methods of discharge and manual water-quality data collection with more recent research and guidance on automated storm sampling. Extensive guidance is available on certain aspects of hydrologic and water-quality data collection. Preferred methods in discharge data collection, developed by the U.S.

* The U.S. Geologic Survey (USGS) is an exception, as they have the expertise and personnel to collect data on larger watersheds.

Department of Agriculture (USDA) and the U.S. Geological Survey (USGS) scientists, appear in the *Field Manual for Research in Agricultural Hydrology* (Brakensiek et al. 1979) and in selected *Techniques of Water-Resources Investigations of the USGS* (e.g., Buchanan and Somers 1976, 1982; Carter and Davidian 1989; Kennedy 1984). Chow et al. (1988), Haan et al. (1994), and Maidment (1993) also provide comprehensive information on applied hydrology. In its *National Field Manual for Collection of Water Quality Data* (USGS 1999), the USGS provides guidance for its personnel on preferred methods for manual collection of water quality samples. Other publications provide extensive guidance on manual field measurements in terms of sample collection techniques and quality control (e.g., Wells et al. 1990) and general information on quality assurance (QA), sample collection, and statistical analysis procedures (e.g., Dissmeyer 1994; USDA-NRCS 1996; U.S. EPA 1997). Much of the information on preferred methods for hydrologic (Section 16.2.2) and water-quality data collection (Section 16.2.3) was compiled from these sources.

The previously available sources do not, however, provide the much-needed information on design and implementation of automated storm sampling projects to achieve monitoring goals within resource constraints. Because little such practical guidance has been developed, project design is commonly based on field experience (in the best case) or with no knowledge of design factors and potential consequences (in the worst case). Research results (e.g., King and Harmel 2003; McFarland and Hauck 2001; Miller et al. 2000; Robertson and Roerish 1999; Tate et al. 1999) and practical guidance (e.g., Behrens et al. 2004; Harmel et al. 2003; McFarland and Hauck 2001) on storm sampling on small watersheds have only recently been published.

16.2.1 Monitoring Resources

Most projects are faced with resource constraints, and monitoring resource requirements are often underestimated by project designers. Agourdis and Edwards (2003) emphasized that the collection and analysis of water quality samples is a difficult, time-consuming, and expensive task; however, this simple truth is commonly not appreciated. Personnel needs, travel time, equipment purchase and maintenance, site location, sites numbers, and laboratory analysis costs should all be carefully examined prior to project initiation.

Committed, well-trained field personnel are essential for water monitoring projects. Personnel must be on call and willing to make frequent trips to remote sites for data collection and sample retrieval, whether or not samples are collected automatically (Section 16.2.3). This travel is often necessary with little advance warning and under adverse weather conditions. These trips can also consume considerable time for conducting necessary equipment inspection, maintenance, and repair. In spite of expense and time required, maintenance of flow and water-quality monitoring equipment is an essential step in producing meaningful data. A commitment to proper maintenance limits loss of data and equipment malfunctions, which, if allowed to occur, increases the uncertainty in measured data affecting model calibration and evaluation. Back-up equipment should be purchased and made ready for rapid replacement of malfunctioning components. Site visits should be made weekly or in alternating weeks to

- check power sources, stage recorders, pumps, sample tubes, sample intakes, dessicant levels
- calibrate stage recorders to assure flow measurement accuracy
- retrieve data to limit loss caused by power failures or other malfunctions
- perform required maintenance and equipment replacement

Personnel should also visit all sampling sites as soon as possible during or after sampling events to collect or retrieve samples, check stage recorder and automated sampler function, and make necessary repairs. Delay in retrieving water quality samples and transporting them to the lab can result in substantial changes in their chemical composition. The acceptable time frame is constituent specific and should be included in project QA guidelines.

Decisions regarding project resource allocation should also consider the number and location of sampling sites and the analysis costs of collected water quality samples. Ideally, data collection sites should be established at a range of scales to adequately assess specific land-management impacts and integrated downstream effects. For best results, field-scale sampling sites should be located at the boundaries of homogeneous land use areas in the natural drainage way. Berm construction may be necessary to direct runoff to a single well-defined outlet. Downstream sampling sites should, if feasible, be established at existing flow gauges or hydraulic control structures (Section 16.2.2) with an historical flow record and a current stage-discharge relationship (rating curve).

The cost and travel time required to establish and maintain multiple sites must, however, be considered. The number of samples that can be collected and analyzed by a laboratory in a reasonable time frame as determined by project QA guidelines is another important consideration (Novotny and Olem 1994). It is prudent to estimate the number of samples that will be collected to meet reasonable sampling expectations within the project resources. For flow interval sampling strategies (Section 16.2.4.2), the mean annual number of samples can be estimated from historical runoff data. Selection of base-flow and storm sampling methodology (Sections 16.2.3 and 16.2.4) also affects the number of samples collected, which directly influences sample analysis costs.

16.2.2 Flow Characterization

Collection of adequate flow data is vital in monitoring projects designed to support P modeling efforts because runoff and associated sediment is the dominant overland P transport process. Discharge (flow rate) data, along with corresponding dissolved and particulate P concentrations, are needed to determine the mass transport values and differentiate between transport mechanisms. Typically, discharge is determined with the relation between stage (water surface level or flow depth) and discharge. A general description of stage-discharge relationships and their development is provided in most applied hydrology texts (e.g., Brakensiek et al. 1979; Maidment 1993).

With this method, stage data are recorded and translated to discharge with the stage-discharge relationship. A stage-discharge relationship alleviates the difficult task of measuring actual flow rates and instead uses stage, which is relatively easy to measure, to determine discharge.

Bubblers, pressure transducers, floats, and noncontact sensors are commonly used to provide continuous stage data. Bubblers and pressure transducers are submerged devices that measure stage by sensing the pressure head created by water depth. Noncontact sensors are suspended above the water surface and use ultrasonic or radar technology to measure water level. All of these devices are typically used in connection with an electronic data logger to store a continuous stage record. Float sensors actually float on the water surface and, in conjunction with a stage recorder, produce a graphical or electronic record of stage. Installation of a permanent staff gauge with which to calibrate stage devices is also recommended, but a surveyed reference elevation point should be established at a minimum.

The most reliable stage-discharge relationships are associated with hydraulic control structures, such as flumes or weirs, which can provide stable and accurate flow data for a number of years with minimal maintenance. These structures are often precalibrated and thus do not require development of a stage-discharge relationship. This is an important benefit because stage-discharge relationship development is a time-consuming, long-term task requiring measurement of stage, cross-sectional flow area, and flow velocity for a range of stages. Selection of an appropriate structure for local conditions should be based on the following factors: (1) expected flow range and existing headwater-tailwater effects on structure calibration; (2) floating or suspended debris and transported sediment; (3) construction and maintenance costs in relation to expected project life; and (4) need for flow measurement standardization at sites within the project. Detailed selection criteria for hydraulic control structures are provided in Bos (1976) and Brakensiek et al. (1979).

For small watershed sites, pre-calibrated hydraulic control structures are highly recommended in spite of the high cost of purchase and installation. These structures are, however, limited in the discharge they can support, which limits their use on many large watersheds. If installation of a structure is not feasible, location of sampling sites at or near established gauge stations with available data is recommended. Other preferred sampling site locations are culverts or concrete channels, which often provide reliable, consistent stage-discharge relationships. Establishing monitoring sites in natural channels subject to morphological shifts in channel geometry or in locations with limited data can create considerable difficulty in maintaining reliable stage-discharge relationships. An important consideration, regardless of channel type or measurement technique, is assurance that measurement can be made for the complete range of expected flow rates.

Another method for determining discharge utilizes measurements of cross-sectional flow area and flow velocity. This is the typical method for determining or adjusting stage-discharge relationships for sites in natural channels and for uncalibrated structures. With this method, the flow is divided into vertical sections, and mean velocity and cross-sectional flow area are determined for each section. The total discharge for that stage is the sum of discharges for each section. This procedure must be repeated for the range of expected discharges. Several portable devices are available to measure flow velocities. Velocity meters may use revolving cups that spin at a rate proportional to the velocity, or they may use Doppler, electromagnetic, or radar technology to determine flow velocity. When using each of these meters, care must be taken to determine the mean flow velocity within the vertical section of interest. Permanent in-stream velocity meters are also available that provide continuous stage and velocity measurements. In theory, these instruments use corresponding stage and velocity measurements with cross-sectional survey data to produce continuous discharge measurements; however, the flow velocity values may not adequately represent the mean velocity of the entire flow cross-section.

If a stage-discharge relationship is not established for a monitoring site and if in-stream velocity measurement is not feasible, mean velocity can be estimated using a derivative of Manning's equation. Then, cross-sectional survey data can be used with the mean velocity to estimate discharge. Manning's equation was developed for uniform flow, which is much more likely to occur in constructed channels with uniform perimeters than irregular natural channels. Therefore, Manning's equation introduces substantial uncertainty into discharge data when applied to natural channels and thus should only be used as a final option.

16.2.3 Water-Quality Characterization

Depending on watershed scale and discharge characteristics, base flow and storm runoff sampling may be needed to adequately characterize various P transport mechanisms. At small watershed sites characterized by perennial flow, base flow sampling is needed to evaluate P transport as affected by in-stream processes, direct deposition from wildlife and livestock, groundwater inflow, and point source contribution. Base flow sampling is generally unnecessary at field-scale or ephemeral small watershed sites where P transport occurs predominately in runoff events. Storm sampling is needed at each of these scales to capture the nonpoint source contribution of dissolved and particulate P and potential resuspension of P associated with in-stream sediment.

16.2.3.1 Base Flow and Low Flow

Manual grab sampling is typically used to characterize base flow and low flow water-quality. To provide the most beneficial data to support P modeling, base flow water-quality samples should be taken as often as possible and at regular time intervals not less than once per month. Samples can be taken at a single point in the flow, generally in the centroid of flow, because dissolved constituent concentrations typically are assumed to be uniform across the cross-section unless the site is located immediately downstream of a significant point source contribution (Martin et al. 1992, Slade, 2004, Ging 1999). This assumption is discussed in more detail in Section 16.2.3.2.

16.2.3.2 Storm Flow

Characterization of storm water quality is much more difficult. Storm events occur with little advance warning often outside the conventional work hours and by definition accompany adverse weather. As a result, automated water-quality sampling equipment is often used so that personnel are not required to travel to multiple sites during runoff events. In contrast, manual storm sampling requires personnel to travel to each sampling site and manually collect samples during storm events (Table 16.1). The USGS Equal-Width-Increment (EWI) and Equal-Discharge-Increment (EDI) procedures are widely accepted as proper manual storm sampling methods (USGS 1999; Wells et al. 1990). With these procedures, multiple depth-integrated, flow-proportional samples are obtained across the stream cross-section and produce accurate dissolved and particulate P concentration measurements even in large streams. Despite this advantage, manual techniques require substantial collection time for each sample, which creates difficulty in collecting multiple samples at numerous sites. Less intensive manual sampling, such as grab sampling at random times or locations during storm events, provides much less useful data compared to intensive manual or automated sampling. Regardless of the manual sampling technique utilized, samples should be collected throughout the entire range of observed flow to adequately characterize P transport.

The major advantage of automated samplers is their ability to use consistent sampling procedures to take multiple samples at multiple sites throughout complete runoff events of various durations. This is especially important at remote and/or small-scale sites because of the difficulty that field personnel have in traveling to sites and collecting adequate data within event durations. Automated samplers, however, are quite expensive to purchase and maintain and thus require considerable

TABLE 16.1
Advantages and Disadvantages of Automated and Intensive Manual Storm Sampling

Automated Storm Sampling		Manual Storm Sampling	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
Reduced on-call travel	Large investment in equipment	Low equipment cost	Large investment in personnel
Multiple samples collected automatically	Single intake (samples taken at one point)	Integrated samples throughout profile and cross-section	Frequent on-call travel often in adverse weather and dangerous conditions
Numerous sites feasible	Difficult to secure intake in the centroid of flow		Time-consuming sample collection
Avoid work in dangerous conditions	Considerable maintenance and repair requirement		Numerous sites difficult to manage

financial investment. Another potential disadvantage of automated samplers is their utilization of a single intake point, which is discussed in detail in the following paragraph. It is assumed that most monitoring projects designed to support P modeling will utilize automated sampling. This assumption is based primarily on the ability of automated samplers to take multiple samples at multiple sites with a consistent sampling procedure and on the realization that most monitoring projects will not have the resources to maintain an adequate on-call field staff* to conduct intensive manual storm sampling at multiple sites (Table 16.1).

An important difference between automated and intensive manual storm sampling is that automated samplers typically utilize a single intake while the manual EWI and EDI procedures collect integrated samples across the stream cross-section. Thus, the uniformity of water quality across the flow cross-section and within the water profile deserves consideration. It is generally assumed that dissolved constituents can be adequately sampled at a single intake point in small streams because of well-mixed conditions and in larger streams unless located immediately downstream from significant point sources prior to complete mixing (Martin et al. 1992, Slade, 2004, Ging 1999). If doubt arises as to whether dissolved constituents are uniformly distributed, this can be easily evaluated with a hand-held conductivity probe. If conductivity measurements are relatively uniform throughout the cross-section, then the assumption of well-mixed conditions is supported.

This assumption is often invalid for sediment and particulate P because their concentrations typically vary within the vertical profile and across the channel. In spite of this variability, a single sample intake is generally adequate at most field-scale sites because of shallow flow depths and well-mixed conditions. In larger streams, however, EWI or EDI sampling is needed to adequately capture the variability of sediment concentrations within the flow profile and across the channel. To use automated samplers in large streams with constituent concentration variability, single intake samples should be supplemented by manual integrated sampling (e.g., Ging 1999). With both types of samples taken at a range of discharges, the relation between concentrations at the sampler intake and the mean cross-sectional concentrations can be established and used to determine mean concentrations from single intake measurements.

16.2.4 Automated Storm Sampling Settings

Three settings are critical in programming automated samplers to collect storm water quality samples. Decisions regarding the following settings determine the number, frequency, and collection method of water-quality samples and, therefore ultimately determine the uncertainty of transport measurement (Section 16.3):

- Threshold to start and finish sampling (Section 2.4.1)
- Sampling interval on which to collect samples after sampling begins (Section 2.4.2)
- Discrete or composite sample collection (Section 2.4.3)

* The USGS, however, is one agency with the expertise and personnel to conduct proper manual storm sampling.

Most commercially available automated samplers contain the following components: programmable electronic operation and memory, water level (stage) recorder, sample collection pump, and sample bottles. Typical bottle arrangements allow from 1 to 24 sample bottles. These electronic samplers evolved from automated, mechanical samplers that were initiated with a float-activated water level switch. Alternative mechanical automated sampling procedures have been designed to provide reliable, low-cost operation for small scale monitoring, but these are not used as frequently as electronic automated samplers. Examples are the Low-Impact Flow Event (LIFE) sampler (Franklin et al. 2001; Sheridan et al. 1996) and modifications of the Coshocton Wheel sampler (Bonta 2002; Edwards et al. 1976; Malone et al. 2003; Parsons 1954, 1955). Both of these can be used for indirect measurement of runoff volume from small watersheds.

16.2.4.1 Storm Sampling Threshold

The first critical program setting for automated samplers is selecting a threshold to initiate sampling. For runoff-driven storm sampling, a minimum stage or discharge threshold is typically set, but an additional rainfall criterion is commonly included for larger watersheds. When flow depth or rate exceeds this threshold, sampling begins and typically continues as long as flow remains above this threshold; therefore, setting the minimum flow threshold directly affects the number of samples taken and the proportion of the total discharge sampled (Figure 16.1).

Results from Harmel et al. (2002) suggest that substantial sampling error is introduced as minimum flow thresholds are increased. Therefore, thresholds should be set so that as much of the storm duration as possible is sampled. To prevent pump malfunction, the sampler intake should be placed so that it is completely submerged at the minimum flow threshold. Ideally, the sampler intake should be located in the center of the channel in well-mixed flow not a pool or immediately upstream below the crest of the hydraulic control structure. The programming option to sample each time flow rises and/or falls past the threshold (i.e., as sampling is initiated and completed) should be avoided because flow fluctuations near the threshold will override the specified sampling interval and result in unnecessary samples.

16.2.4.2 Sampling Interval

The second important setting is the interval on which to sample once the sampling threshold is reached. There are two options for determining the sampling interval: time and flow (Figure 16.2). Time-interval sampling is also referred to as time-weighted, time-proportional, or fixed frequency sampling, and flow-interval sampling can be referred to as flow-weighted or flow-proportional sampling.

With time-interval sampling, samples are typically taken at equal time increments (such as every 30 min). Variable time intervals (typically with more frequent samples initially, then less frequently as the storm proceeds) can be beneficial, however, if based on adequate knowledge of site hydrology. Time-interval sampling is a simple and reliable procedure since accurate time intervals are easy to measure

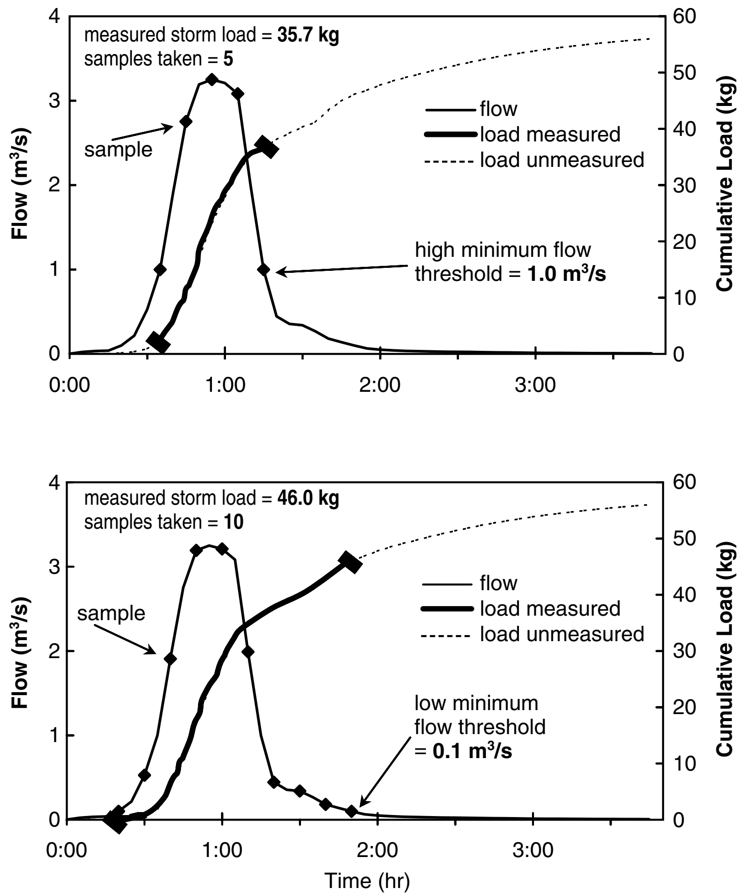

FIGURE 16.1 Loads measured with different minimum flow thresholds (1.0 and 0.1 m^3/s) with a time-interval (10 min) sampling strategy. The bold lines represent the measured portion of total storm load.

and clock failures are rare. However, if small time intervals are used, frequent sampling will quickly produce numerous samples, exceed sampler capacity, and not adequately characterize the entire runoff event (Table 16.2). Time-interval sampling does not eliminate the need for flow measurement, as flow data are necessary for load determination.

With flow-interval sampling, samples are collected on flow volume increments, such as every 2000 m^3 or 2.5 mm volumetric depth*. Flow-interval sampling requires continuous flow monitoring to determine loads and to determine sampling intervals.

* Referring to discharge intervals in volumetric depth units such as mm, which represent mean runoff depth over the entire watershed, as opposed to volume units such as m^3, normalizes discharge over various watershed sizes. This notation allows a consistent transfer of methods and results to watersheds of differing size.

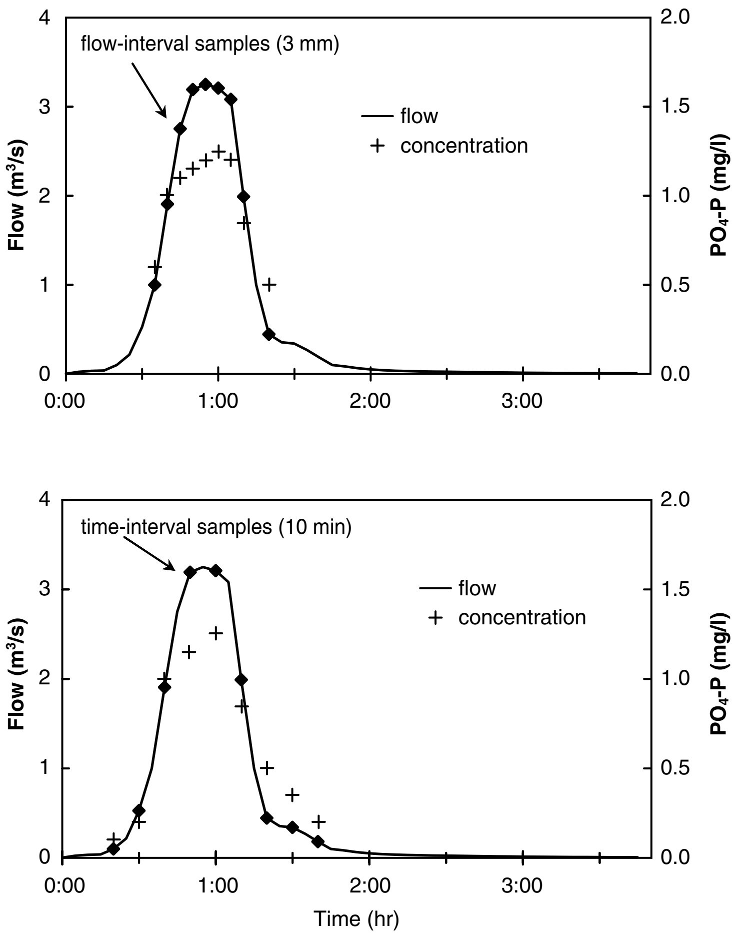

FIGURE 16.2 Example hydrograph illustrating differences in sample timing for time- and flow-interval strategies. Although both strategies collect nine samples, flow-interval sampling is most frequent at high flow rates, whereas the frequency of time-interval sampling is consistent throughout the event. Thus, the concentrations measured can be quite different.

Flow-interval sampling readily produces the Event Mean Concentration (EMC), a common method for reporting constituent concentrations defined as the arithmetic mean of individual sample concentrations collected on equal discharge intervals. The EMC multiplied by the total flow volume represents the storm load.

Statistical sampling theory indicates that the smaller the sampling interval (the more samples taken), the better actual population characteristics are estimated (Haan 2002). Several recent studies confirm this theory regarding storm monitoring (Harmel and King 2005; King and Harmel 2003, 2004; Richards and Holloway 1987; Miller et al. 2000; Shih et al. 1994). Thus, frequent sampling intervals should be used to accurately characterize storm water quality. However, intervals should

TABLE 16.2
The Number of Samples Taken Estimated for Watersheds (0.1 to 6300 ha) and the Sampling Capacity Based on a 24-Bottle Configuration for Selected Strategies

Sampling Strategy	Number of Samples			Maximum Duration (min)	
Time-Interval Discrete (min)	Range	Mean	Median	Discrete	2/Bottle
5	8 – 1237	234	164	120	240
10	4 – 619	117	82	240	480
15	3 – 413	78	55	360	720
30	2 – 207	39	28	720	1440
60	0 – 104	20	14	1440	2880
120	0 – 52	10	7	2880	5760
180	0 – 35	6	5	4320	8640
300	0 – 21	4	3	7200	14400
Sampling Strategy	**Number of Samples**			**Maximum Volume (mm)**	
Flow-Interval Discrete (mm)	**Range**	**Mean**	**Median**	**Discrete**	**2/Bottle**
1.0	0 – 132	30	25	24	48
2.5	0 – 53	12	10	60	120
5.0	0 – 26	6	5	120	240
7.5	0 – 17	3	3	180	360
10.0	0 – 13	2	2	240	480
12.5	0 – 10	2	2	300	600
15.0	0 – 8	1	1	360	720

Source: K.W. King and R.D. Harmel, *Trans. ASAE,* 46, 63, 2003. With permission.

not be set at such a high frequency that prevents complete sampling of various duration runoff events and introduces substantial uncertainty. Harmel et al. (2003) and Harmel and King (2005) provide guidance on selecting time and flow intervals for sampling small watersheds. For time-interval sampling, intervals must be adjusted based on watershed size; however, a consistent flow interval can be used on small watersheds regardless of size if volumetric depth intervals that normalize runoff volume by watershed area are used.

Several studies have concluded that flow-interval sampling better represents storm loads than time-interval sampling because a greater proportion of the samples are taken at higher flow and transport periods (Abtew and Powell 2004; Claridge 1975; Harmel and King 2005; King and Harmel 2003; Izuno et al. 1998; McFarland and Hauck 2001; Miller et al. 2000; Rekolainen et al. 1991; Richards and Holloway 1987; Shih et al. 1994). In practical terms, it is difficult to choose time-intervals that can completely sample various duration events with adequate frequency to capture constituent concentration fluctuation without exceeding sampler capacity. It is much easier for flow-interval sampling to provide intensive sampling throughout entire events of various magnitudes (Table 16.2) when runoff is expressed in volumetric depth intervals.

16.2.4.3 Discrete vs. Composite Sample Collection

Another important program setting for automated samplers is the option to collect discrete samples (one sample per bottle) or composite samples (more than one sample per bottle). Discrete sampling strategies provide the best representation of temporal variability of constituent concentrations; however, discrete sampling can produce substantial error even with frequent sampling intervals. This effect is evident in large volume/duration runoff events in which sampling capacity is reached prior to the completion of the event. As shown in Table 16.2, numerous samples can be generated, especially by time-interval sampling, but the common 24-bottle limitation allows only a fraction of the samples to be collected. Composite sampling is a powerful alternative because it increases sampler capacity by collecting more than one sample in each sample bottle. Composite sampling with two, three, or four samples per bottle reduces sample numbers to 50%, 33%, and 25% of that collected by discrete strategies. Composite sampling does, however, reduce data on the distribution of within-event constituent behavior, which limits quantification of various transport mechanisms (McFarland and Hauck 2001).

An alternative to using the sampler to composite samples involves manual compositing. Samples collected on equal time intervals can be manually composited following the sampling event by combining subsamples proportional to flow volume during each time interval. This technique does produce a meaningful estimate of the EMC but requires considerable postprocessing.

Several recent studies have concluded that composite sampling introduces less error than raising minimum flow thresholds or increasing sampling intervals, especially for flow-interval sampling (Harmel and King 2005; Harmel et al. 2000; King and Harmel 2003; Miller et al. 2000). Therefore, composite sampling is recommended for management of the number of samples collected (Section 16.3). For monitoring projects whose primary goal is load determination, not examination of within-event constituent behavior, single-bottle composite flow-interval sampling is a powerful option that reduces analysis costs while intensively sampling complete events of various durations (Shih et al. 1994; Harmel and King 2003). With this strategy, 80 to 160 flow-interval samples of 100 to 200 ml can be composited into a single sample (16 L bottle capacity) to produce the EMC. Another appropriate option involves collecting discrete samples until an adequate understanding of constituent behavior is gained and then converting to composite sampling.

16.2.5 Alternative Procedures (Regression Methods)

The previous sections discussed achieving an appropriate balance between measured P transport data quality and project resources under the assumption that resources are sufficient to intensively sample storm water quality (and base flow if significant). In situations with inadequate resources to conduct intensive water quality sampling, regression methods can be used to estimate P transport (Cohn 1995; Cohn et al. 1989). Regression methods, which typically utilize the relation between flow and constituent concentrations to estimate constituent loads, are commonly applied to large watersheds (Haggard et al. 2003; Robertson and Roerish 1999). Current versions have been modified from the original simple linear regression approach to account for nonlinearity, seasonality, and other complicating factors (Cohn 1995; Robertson and Roerish 1999). The statistical relation among discharge, concentrations, and other complicating factors is used to estimate missing daily constituent concentrations, which are then summed to give monthly, seasonal, or annual load estimates. A benefit of this statistical approach is its ability to place confidence limits on resulting load estimates.

Regression methods can be applied to relatively small water-quality datasets collected over many years; however, water quality sampling designed to provide input to regression methods must adequately describe the relation between discharge and constituent concentration. To achieve that, dissolved and particulate P concentration data must be collected throughout the range of discharge observed at that location. Sampling strategies should target both base-flow and storm events, as fixed interval sampling (e.g., monthly sampling) may not adequately represent the range of discharge. Monthly sampling strategies targeting base flow may underestimate constituent loads by over 40% (Haggard et al. 2003). Robertson and Roerish (1999) suggested that the collection of water samples during storm events may positively bias annual load estimates in smaller streams because storm concentrations are typically larger than average daily concentrations. Constituents associated with sediment transport often exhibit hysteresis across the storm event, with greater concentrations on the rising portion of the hydrograph than the corresponding discharge on the falling portion (Richards and Holloway 1987; Richards et al. 2001; Thomas 1988); therefore, samples should be collected during both the rising and falling portions.

This methodology has been widely used, particularly by the USGS, in relatively large streams and rivers across the U.S. as an effective and economical alternative to provide constituent load data (e.g., Green and Haggard 2001; Pickup et al. 2003). The application of this alternative to field-scale and small watersheds has received limited evaluation except for Robertson and Roerish (1999), who concluded that regression methods are relatively imprecise in small watersheds. Thus, application to small watersheds should be conducted with caution.

16.3 UNCERTAINTY IN P TRANSPORT MEASUREMENT

An emphasis on the need for uncertainty estimates associated with model outputs has recently re-emerged, as water quality models are increasingly used to guide natural resource decision-making and legislation (Beck 1987; Haggard et al. 2003; Hession et al. 1996; Sharpley et al. 2002). One obstacle to properly estimating model output uncertainty is the lack of understanding related to uncertain calibration and evaluation data. The uncertainty* related to nutrient transport measurement is poorly understood at best. As a result, the effects of uncertain P transport data on calibration and evaluation of P models have historically been ignored. However, if the water resource community is truly serious about uncertainty and its impact on water quality modeling, uncertainty in constituent transport measurements desperately needs intense research attention.

* It is important to note that sampling error is defined as sampling variability or sampling uncertainty and does not include mistakes in data collection and processing (Haan 2002).

This issue of uncertainty also affects monitoring project design, as most projects are faced with balancing resources and accurate discharge and water quality characterization (Agourdis and Edwards 2003; Harmel et al. 2003; Preston et al. 1992; Shih et al. 1994; Tate et al. 1999). The relative differences in uncertainty between discharge measurement alternatives (Section 16.2.2) have been known for some time. However, only recently have research and practical guidance on automated storm water quality sampling been published. Thus, projects utilizing this methodology are often implemented without regard for the effects of sampler settings on data uncertainty.

Each of the important automated storm sampling settings (storm sampling threshold, sampling interval, discrete/composite sampling) directly affects the uncertainty of storm water quality data. These settings determine whether constituent behavior (such as first flush and concentration hysteresis) is adequately characterized without exceeding sampler capacity in events of various durations. Recent research has produced the following conclusions regarding uncertainty in storm water quality data:

- Raising the minimum flow threshold decreases the proportion of the storm duration that is sampled and increases uncertainty (Harmel et al. 2002).
- Increasing the sampling interval increases uncertainty (Richards and Holloway 1987; Shih et al. 1994; Miller et al. 2000; King and Harmel 2003, 2004; Harmel and King 2005).
- Composite sampling may increase uncertainty for time-interval sampling (Miller et al. 2000; King and Harmel 2003) but by a lesser amount than corresponding increases in sampling interval.
- Composite flow-interval sampling has little effect on uncertainty (King and Harmel 2003; Harmel and King 2005).

A majority of the previous research on uncertainty was conducted by comparing various estimates of constituent flux and thus addressed relative differences (precision) in error without regard to possible deviation from the true flux (accuracy). This approach is attributed to the cost and commitment required to make true flux measurements. As a result, relative comparisons of various storm sampling strategies can be made, but little is known about the true uncertainty of each. In theory, a measurement of the true flux must be made to determine the uncertainty produced by sampling strategy estimates. However, because it is impractical in field studies to capture the entire runoff volume for actual load measurement (Parsons 1954), a true flux must be assumed. For the assumption to be valid, a sufficiently frequent sampling intensity or appropriate subsampling scheme is required. In practical terms, either an automated sampler with a flow-interval sampling strategy with a 1.5 mm or less volumetric depth interval (Harmel and King 2005) or a flow-proportional sampler (Bonta 2002; Edwards et al. 1976; Franklin et al. 2001; Malone et al. 2003; Parsons 1954, 1955; Sheridan et al. 1996) can produce true dissolved constituent loads on small watersheds. In conditions involving large watersheds or extreme in-channel concentration gradients, frequent cross-sectionally integrated sampling throughout the event duration is required to establish the true flux. Conducting such data collection is a difficult task that is beyond the resources and expertise of typical monitoring projects.

Research by Harmel and King (2005) was initiated to address uncertainty estimates in measured storm water-quality data from small agricultural watersheds. All

15 of the flow-interval strategies evaluated (sampling intervals up to 5.28 mm volumetric depth with discrete and composite sampling 2 to 5 samples per bottle) produced cumulative load error magnitudes less than ±10%. The ranking of absolute errors in individual event and cumulative load estimation (sediment > NO_3-N > PO_4-P) is attributed to differences in within-event concentration variability as measured by the coefficient of variation, *CV*, which was also noted by Claridge (1975). The mean *CV* across sites for within-event concentrations was 0.61 for sediment, 0.39 for NO_3-N, and 0.19 for PO_4-P. The authors concluded that sampling intervals up to 6 mm should produce similar load accuracy in other locations for constituents that vary relatively little within runoff events, but smaller intervals (1 to 3 mm) should be used to sample widely varying constituents.

The focus of this discussion is the uncertainty effects of various automated water quality sampling strategies; however, uncertainty is also introduced into reported nutrient data by uncertainties in discharge measurement, sample preservation/storage, and laboratory analysis (Harmel et al. 2006). Although relative differences between discharge measurement alternatives are well established, few studies on errors associated with sample preservation, storage, and analysis have been published (e.g., Kotlash and Chessman 1998; Jarvie et al. 2002). Research on the relative differences in uncertainty contributed by each of these potential sources has only recently become available (e.g., Harmel et al. 2006).

16.4 SUMMARY

In recent years, many monitoring projects have been initiated or modified to provide water quality and discharge data needed to support water resource management. Water-quality modeling in particular relies on measured nutrient transport data for calibration and evaluation. These data are needed to improve model dynamics to more accurately represent soil P cycling and transport and to improve linkages between field-scale losses and downstream transport.

In monitoring project design and modification, it is important to utilize data collection methods that accurately characterize runoff and water quality within resource constraints. The design factors that directly impact this balance are monitoring resources, flow characterization, automated vs. manual water quality sampling, and settings for automated storm sampling. Although relatively little information is available on measurement uncertainty effects for certain project design factors, the following guidelines have been established to improve data quality and reduce sampling uncertainty:

- Assemble a well-trained field staff willing to be on-call and make frequent site visits.
- Commit to proper maintenance of monitoring equipment in spite of time and expense.
- Install pre-calibrated hydraulic control structures or develop reliable stage-discharge relationships.

- At field-scale and small watershed sites, use automated sampling equipment programmed with a low minimum flow threshold to collect frequent, flow-interval composite samples.
- At larger watershed sites where particulate P concentrations are not uniform, manually collect frequent EWI or EDI samples throughout complete events, and supplement automated sample collection with EWI or EDI sampling to correct single intake concentrations to represent mean concentrations within the entire channel.

As modeling increasingly impacts water resource policy, management, and regulation, the water resource community needs to direct research attention to the issue of uncertain nutrient transport data and its impact on water-quality model output. Additional data are needed to provide scientifically sound information to decision makers about the uncertainties in calibration and evaluation data and the uncertain nature of model output.

REFERENCES

Abtew, W. and B. Powell. 2004. Water quality sampling schemes for variable flow canals at remote sites. *J. Am. Water Resour. Assoc.* 40:1197–1204.

Agouridis, C.T. and D.R. Edwards. 2003. The development of relationships between constituent concentrations and generic hydrologic variables. *Trans. ASAE* 46:245–256.

Beaulac, M.N. and K.H. Reckhow. 1982. An examination of land use-nutrient export relationships. *Water Resour. Bull.,* 18:1013–1024.

Beck, M.B. 1987. Water quality modeling: a review of the analysis of uncertainty. *Water Resour. Res.* 23:1393–1442.

Behrens, B., J.H. Riddle, and J. Gillespie. 2004. Tips to improve wet weather monitoring. *Am. Publ. Works Assoc. Report.* Also presented by Holbrook, K.R., Maximizing reliability of storm water monitoring data, *North American Surface Water Quality Conference*, Palm Desert, CA, 2004.

Bonta, J.V. 2002. Modification and performance of the Coshocton wheel with the modified drop-box weir, *J. Soil Water Cons.,* 57:364–373.

Bos, M.G. 1976. Discharge measurement structures, Publication 20, International Institute for Land Reclamation and Improvement, Wageningen, the Netherlands.

Brakensiek, D.L., H.B. Osborn, and W.J. Rawls. 1979. Field manual for research in agricultural hydrology, *Agriculture Handbook*, U.S. Dept. Agriculture, Washington, D.C., 224.

Buchanan, T.J. and W.P. Somers. 1976. Discharge measurements at gaging stations, in *Techniques of Water-Resources Investigations of the U.S. Geological Survey,* book 3, chap. A8.

Buchanan, T.J. and W.P. Somers. 1982. Stage measurement at gaging stations, in *Techniques of Water-Resources Investigations of the U.S. Geological Survey,* book 3, chap. A7.

Carter, R.W. and J. Davidian. 1989. General procedure for gaging streams, in *Techniques of Water-Resources Investigations of the U.S. Geological Survey,* book 3, chap. A6.

Chow, V.T., D.R. Maidment, and L.W. Mays. 1988. *Applied Hydrology.* New York: McGraw-Hill.

Claridge, G.G.C. 1975. Automated system for collecting water samples in proportion to stream flow rate. *New Zealand J. Sci.* 18:289–296.

Cohn, T.A., L.L. DeLong, E.J. Gilroy, R.M. Hirsch, and D.K. Wells. 1989. Estimating constituent loads. *Water Resour. Res.* 25:937–942.

Cohn, T.A. 1995. Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers. *U.S. Natl. Rep. Int. Union Geod. Geophys. 1991-1994, Rev. Geophys.* 33:1117–1123.

Dissmeyer, G.E. 1994. Evaluating the effectiveness of forestry best management practices in meeting water quality goals or standards, U.S. Department of Agriculture, Miscellaneous Publication 1520, Forest Service, Southern Region, Atlanta, GA.

Edwards, W.M., H.E. Frank, T.E. King, and D.R. Gallwitz. 1976. *Runoff Sampling: Coshocton Vane Proportional Sampler,* ARS-NC-50.

Franklin, D.H., M.L. Cabrera, J.L. Steiner, D.M. Endale, and W.P. Miller. 2001. Evaluation of percent flow captured by a small in-field runoff collector. *Trans. ASAE* 44:551–554.

Gilley, J.E. and M.L. Risse. 2000. Runoff and soil loss as affected by the application of manure. *Trans. ASAE* 43:1583–1588.

Ging, P. Water-quality assessment of south-central Texas — comparison of water quality in surface-water samples collected manually and by automated samplers, U.S. Geological Survey, Fact Sheet FS-172-99.

Haan C.T., B.J. Barfield, and J.C. Hayes. 1994. *Design Hydrology and Sedimentology for Small Catchments*. New York: Academic Press.

Haan, C.T. 2002. *Statistical Methods in Hydrology*, 2nd ed. Ames: Iowa State Press.

Haggard, B.E., T.S. Soerens, W.R. Green, and R.P. Richards. 2003. Using regression methods to estimate stream phosphorus loads at the Illinois River, Arkansas. *Applied Eng. Agric.* 19:187–194.

Harmel, R.D., R.J. Cooper, R.M. Slade, R.L. Haney, and J.G. Arnold. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *American Society of Agricultural and Biological Engineers* 49(3):689–701.

Harmel, R.D. and K.W. King. 2005. Uncertainty in measured sediment and nutrient flux in runoff from small agricultural watersheds. *Trans. ASAE* 48:1713–1721.

Harmel R.D., K.W. King, and R.M. Slade. 2003. Automated storm water sampling on small watersheds. *Applied Eng. Agric.* 19:667–674.

Harmel, R.D., K.W. King, J.E. Wolfe, and H.A. Torbert. 2002. Minimum flow considerations for automated storm sampling on small watersheds. *Texas J. Sci*. 54:177–188.

Helsel, D.R. and R.M. Hirsch. 1993. *Statistical Methods in Water Resources*. New York: Elsevier.

Hession, W.C., D.E. Storm, C.T. Haan, K.H. Reckhow, and M.D. Smolen. 1996. Risk analysis of total maximum daily loads in an uncertain environment using EUTROMOD. *J. Lake Reservoir Manage.* 12:331–347.

Izuno, F.T., R.W. Rice, R.M. Garcia, L.T. Capone, and D. Downey. 1998. Time vs. flow composite water sampling for regulatory purposes in the Everglades Agricultural Area. *Appl. Eng. Agric.* 14:257–266.

Jarvie, H.P., P.J.A. Withers, and C. Neal. 2002. Review of robust measurement of phosphorus in river water: sampling, storage, fractionation, and sensitivity. *Hydrol. Earth Sys. Sci.* 6:113–132.

Kennedy, E.J. 1984. Discharge ratings at gaging stations, in *Techniques of Water-Resources Investigations of the U.S. Geological Survey*, book 3, chap. A10.

King, K.W. and R.D. Harmel. 2003. Considerations in selecting a water quality sampling strategy. *Trans. ASAE* 46:63–73.

King, K.W. and R.D. Harmel. 2004. Comparison of time-based sampling strategies to determine nitrogen loading in plot-scale runoff. *Trans. ASAE* 47:1457–1463.

Kotlash, A.R. and B.C. Chessman. 1998. Effects of water sample preservation and storage on nitrogen and phosphorus determinations: implications for the use of automated sampling equipment. *Water Res.* 32:3731–3737.

Maidment, D.R. 1993. *Handbook of Hydrology.* New York: McGraw-Hill.

Malone, R.W., J.V. Bonta, and D.R. Lightell. 2003. A low-cost composite water sampler for drip and stream flow. *Appl. Eng. Agric.* 19:59–61.

Martin, G.R., J.L. Smoot, and K.D. White. 1992. A comparison of surface-grab and cross sectionally integrated stream-water-quality sampling methods, *Water Environ. Res.* 64:866–876.

McFarland, A. and L. Hauck. 2001. Strategies for monitoring nonpoint source runoff, TIAER Report 0115.

Miller, P.S., B.A. Engel, and R.H. Mohtar. 2000. Sampling theory and mass load estimation from watershed water quality data, American Society of Agricultural Engineers, Paper 00-3050, St. Joseph, MI.

Novotny, V. and H. Olem. 1994. *Water Quality: Prevention, Identification, and Management of Diffuse Pollution*. New York: Van Nostrand Reinhold.

Parsons, D.A. 1955. *Coshocton-Type Runoff Samplers*. ARS-41-2.

Parsons, D.A. 1954. *Coshocton-Type Runoff Samplers, Laboratory Investigations*, SCS-TP-124.

Preston, S.D., V.J. Bierman, and S.E. Silliman. 1992. Impact of flow variability on error in estimation of tributary mass loads. *J. Environ. Eng.* 118:402–419.

Rekolainen, S., M. Posch, J. Kamari, and P. Ekholm. 1991. Evaluation of the accuracy and precision of annual phosphorus load estimates from two agricultural basins in Finland. *J. Hydrol.* 128:237–255.

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

Robertson, D.M. and E.D. Roerish. 1999. Influence of various water quality sampling strategies on load estimates for small streams, *Water Resour. Res.* 35:3747–3759.

Sharpley, A.N., P.J.A. Kleinman, R.W. McDowell, M. Gitau, and R.B. Bryant. 2002. Modeling phosphorus transport in agricultural watersheds: processes and possibilities. *J. Soil Water Conserv.* 57:425–439.

Sheridan, J.M., R.R. Lowrance, and H.H. Henry. 1996. Surface flow sampler for riparian studies. *Appl. Eng. Agric.* 12:183–188.

Shih, G., W. Abtew, and J. Obeysekera. 1994. Accuracy of nutrient runoff load calculations using time-composite sampling, *Trans. ASAE* 37:419–429.

Slade, R. 2004. personal communication.

Stone, K.C., P.G. Hunt, J.M. Novak, M.H. Johnson, and D.W. Watts. 2000. Flow-proportional, time-composited, and grab sample estimation of nitrogen export from an Eastern coastal plain watershed. *Trans. ASAE* 43:281.

Tate, K.W., R.A. Dahlgren, M.J. Singer, B. Allen-Diaz, and E.R. Atwill. 1999. Timing, frequency of sampling affect accuracy of water-quality monitoring. *California Agric.* 53:44–49.

Thomas, R.B. 1988. Monitoring baseline suspended sediment in forested basins: the effects of sampling on suspended sediment rating curves. *Hydrol. Sci.* 33:499.

U.S. Department of Agriculture Natural Resource Conservation Service (USDA-NRCS). 1996. *National Water Quality Handbook: National Handbook of Water Quality Monitoring*, Part 600.

U.S. Environmental Protection Agency (EPA). 1997. *Monitoring Guidance for Determining the Effectiveness of Nonpoint Source Controls*. EPA 841-B-96-004, Washington, D.C.

U.S. Geological Survey (USGS). 1999. Handbooks for water-resources investigations, Section A., in *National Field Manual for Collection of Water-Quality Data*.

Wells, F.C., W.J. Gibbons, and M.E. Dorsey. 1990. Guidelines for collection and field analysis of water-quality samples from streams in Texas, U.S. Geological Survey, Report 90-127, Open-File.