

# Evaluation of Regression Methodology with Low-Frequency Water Quality Sampling to Estimate Constituent Loads for Ephemeral Watersheds in Texas

**Gurpal S. Toor\*** University of Florida

**R. Daren Harmel** USDA-ARS

**Brian E. Haggard** University of Arkansas

**Gerd Schmidt** Martin Luther University

Water quality regulation and litigation have elevated the awareness and need for quantifying water quality and source contributions in watersheds across the USA. In the present study, the regression method, which is typically applied to large (perennial) rivers, was evaluated in its ability to estimate constituent loads ( $\text{NO}_3\text{-N}$ , total N,  $\text{PO}_4\text{-P}$ , total P, sediment) on three small (ephemeral) watersheds with different land uses in Texas. Specifically, regression methodology was applied with daily flow data collected with bubbler stage recorders in hydraulic structures and with water quality data collected with four low-frequency sampling strategies: random, rise and fall, peak, and single stage. Estimated loads were compared with measured loads determined in 2001–2004 with an autosampler and high-frequency sampling strategies. Although annual rainfall and runoff volumes were relatively consistent within watersheds during the study period, measured annual nutrient and sediment concentrations and loads varied considerably for the cultivated and mixed watersheds but not for the pasture watershed. Likewise, estimated loads were much better for the pasture watershed than the cultivated and mixed landuse watersheds because of more consistent land management and vegetation type in the pasture watershed, which produced stronger correlations between constituent loads and mean daily flow rates. Load estimates for  $\text{PO}_4\text{-P}$  were better than for other constituents possibly because  $\text{PO}_4\text{-P}$  concentrations were less variable within storm events. Correlations between constituent concentrations and mean daily flow rate were poor and not significant for all watersheds, which is different than typically observed in large rivers. The regression method was quite variable in its ability to accurately estimate annual nutrient loads from the study watersheds; however, constituent load estimates were much more accurate for the combined 3-yr period. Thus, it is suggested that for small watersheds, regression-based annual load estimates should be used with caution, whereas long-term estimates can be much more accurate when multiple years of concentration data are available. The predictive ability of the regression method was similar for all of the low-frequency sampling strategies studied; therefore, single-stage or random strategies are recommended for low-frequency storm sampling on small watersheds because of their simplicity.

FEDERAL requirements and lawsuits involving water quality issues have elevated the awareness and need for monitoring water quality and quantifying the contribution of various watershed sources. Section 303(d) of the Clean Water Act requires each state to develop a list of impaired water bodies and total maximum daily loads for listed waters. The total maximum daily loads determines the amount of a constituent that a water body can receive while maintaining its beneficial uses and divides the allowable constituent load among the various sources (agricultural, wastewater treatment plants, etc.) in the watershed. In relatively small watersheds, the best method for determining the contribution of these sources is likely a high-frequency sampling strategy using an autosampler (Harmel et al., 2006b; Harmel et al., 2003), which provides adequate data to calibrate watershed models and evaluate management scenarios. Less intensive sampling programs are often initiated because high-frequency water quality sampling requires substantial financial and personnel resources. Furthermore, low-frequency sampling programs may be appropriate in some larger watersheds when paired with statistical techniques to estimate constituent loads (Haggard et al., 2003).

The alternatives to intensive monitoring may involve using regression between daily flow and constituent concentration to estimate daily loads. The transport of water quality constituents, such as nutrients, sediment, and pesticides, is commonly reported in terms of loads, which are a function of the volumetric rate of water passing a fixed monitoring point in a stream and the constituent concentration in that water. Several software programs, such as GCLAS (Koltun et al., 2006), ESTIMATOR (Cohn et al., 1992; Cohn et al., 1989), and LOADEST (Crawford, 1991; Runkel et al., 2004), incorporate regression methodology to estimate constituent loads when intensive water quality sampling data are not available. Originally, the regression method used simple linear relation, but this method has been modified to account for seasonality, nonlinearity, serial correlation, and other complications (Robertson and Roerish, 1999).

Copyright © 2008 by the American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America. All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Published in *J. Environ. Qual.* 37:1847–1854 (2008).

doi:10.2134/jeq2007.0232

Received 8 May 2007.

\*Corresponding author (gstoorto@ufl.edu).

© ASA, CSSA, SSSA

677 S. Segoe Rd., Madison, WI 53711 USA

G.S. Toor, Soil and Water Science Dep., Center for Landscape Conservation and Ecology, Gulf Coast Research and Education Center, 14625 C.R. 672, Univ. of Florida Institute of Food and Agricultural Sciences, Wimauma, FL 33598. R.D. Harmel, USDA-ARS, Temple, TX. B.E. Haggard, Biological and Agricultural Engineering Dep., Univ. of Arkansas, Fayetteville, AR 72701. G. Schmidt, Martin Luther Univ., Halle-Wittenberg, Germany.

**Abbreviations:**  $R^2_{NS}$ , Nash-Sutcliffe coefficient.

Table 1. Selected characteristics of study watersheds.

	Cultivated watershed (Y13)	Pasture watershed (W10)	Mixed watershed (Y)
Area, ha	4.6	8.0	125.1
Slope, %	2.3	2.5	2.4
Landuse/crop	Corn (2001–02); Corn (2002–03); Wheat (2003–04)	coastal bermudagrass	33% cultivated; 36% improved pasture; 31% rangeland
Poultry litter rate, Mg ha <sup>-1</sup> yr <sup>-1</sup>	4.5	6.7	–
Mean N rate, kg ha <sup>-1</sup> yr <sup>-1</sup>	237	172	60
Mean P rate, kg ha <sup>-1</sup> yr <sup>-1</sup>	122	180	37
	Soil test P Mehlich-3 P by ICP, ‡ mg kg <sup>-1</sup>		
2001–02	50.4	44.1	na
2002–03	52.1	43.0	na
2003–04	43.8	20.3	na
Rainfall, mm	1070 ± 87†		
Runoff, mm	287 ± 57	138 ± 18	226 ± 12
Total number of runoff events	11 ± 4	4 ± 1	10 ± 2
Total number of discrete samples collected	108 ± 24	48 ± 14	107 ± 48

† Annual mean ± SD (*n* = 3).

‡ ICP, inductively coupled plasma-optical emission spectroscopy.

The use of daily data is common in regression methodology because many databases and water quality models use daily time steps. This is considered a “big river” approach, but the regression method has also been applied to smaller perennial systems (Robertson and Roerish, 1999). Although regression methods are scientifically accepted for determining annual constituent loads for large-scale (perennial) streams and rivers (Haggard et al., 2003; Robertson and Roerish, 1999), their applicability to small watershed characterized by ephemeral flow conditions has not been evaluated.

In ephemeral streams and field-scale sites, surface runoff during precipitation events is the dominant transport mechanism, and baseflow contribution is minimal. Therefore, data are collected not on periodic intervals but in response to rainfall–runoff events. Thus, the objective of this study was to evaluate the performance of the regression method with flow data and water quality data collected with several low-frequency water quality sampling strategies to estimate constituent loads from small (ephemeral) watersheds. If the regression method performs well (i.e., can accurately estimate loads relative to measured loads), then regression methodology can be used with confidence to assess loads in other ephemeral watersheds. This result would not only reduce the need for intensive water quality sampling but would also expand the ability of existing programs to assess various sources that contribute to water quality impairments in a watershed or region.

## Methods

### Site Description and Data Collection

Three watersheds at the USDA-ARS Grassland Soil and Water Research Laboratory near Riesel, Texas were used in this study (Table 1). Watershed Y13 (hereafter cultivated watershed) is 4.6 ha, and watershed W10 (hereafter pasture watershed) is 8.0 ha. Both of these are “edge-of-field” sites that receive runoff from a homogeneous land use area. Watershed Y (hereafter mixed watershed) is a 125.1-ha mixed land use site with contribution from cultivated and pasture areas. All of

these small watersheds are ephemeral, with flow dominated by surface runoff and a minor contribution of lateral subsurface return flow (Allen et al., 2005; Harmel et al., 2006c). Inputs of N and P from inorganic fertilizer and poultry litter application to these watersheds from 2001 to 2004 are included in Table 1.

A flow control structure with a well established, reliable stage discharge relationship is located at each watershed outlet (Y13: 5:1 broad-crested v-notch weir; W10: combination Columbus Shallow Notch Weir and 1.83 m Parshall Flume; Y: Columbus A-1 deep notch weir). The flow control structures provided reliable stage measurements for accurate discharge determination. Each flow control structure was equipped with an automated ISCO 6700 sampler (ISCO, Inc., Lincoln, NE) to capture storm water quality samples. From 2001 to 2002, the samplers were programmed to collect samples on variable time intervals based on historical knowledge of runoff durations, and the samplers were programmed to take flow-interval (1.32 mm volumetric depth) discrete samples in 2003.

Three years of data were used for this study (August 2001 through July 2004). Continuous discharge data were collected on 5- to 15-min intervals. Frequent discrete storm samples, along with the corresponding discharge, were collected during runoff events. With these discrete samples, alternative sampling scenarios were generated. The storm samples were analyzed to determine concentrations of dissolved NO<sub>3</sub>-N and PO<sub>4</sub>-P, total N and P, and sediment.

Collected runoff samples were acidified with concentrated HCl, transported to the laboratory for analyses, and stored at 4°C before analysis. Dissolved NO<sub>3</sub>-N and PO<sub>4</sub>-P concentrations were determined with a Technicon Autoanalyzer IIC (Bran-Luebbe, Roselle, IL). Sediment concentrations were determined by allowing the sample to settle for 3 to 5 d and decanting off a majority of the solution. The sediment slurry was dried at 116°C for 18 to 24 h, and the mass of sediment was determined. This mass divided by the measured volume of collected sample represented the sediment concentration. Particulate total Kjeldahl N and P content in this sediment was deter-

Table 2. Low-frequency sampling strategies analyzed in this study.

Sampling strategy	Notation	Description
Single stage sampling—usually accomplished by a siphon sampler (Graczyk et al., 2000) or similar device	single stage	first sample collected at or above stage corresponding to 99.9% exceedence for average daily flow (0.03 m <sup>3</sup> s <sup>-1</sup> for Y13 and W10; 0.4 m <sup>3</sup> s <sup>-1</sup> for Y)
Peak flow sampling	peak	one sample collected nearest the peak flow (if multiple peaks occurred only the maximum peak was considered)
One sample per storm—usually a grab sample taken by on-call personnel	random	one sample taken at a random time within the storm
One sample from the rising limb and one sample in the falling limb	rise fall	one random sample taken before the peak and one random sample taken after peak (if the peak flow occurred on the first sample, the first sample was used for the rise; if only one sample was taken, that sample was used for the rise and the fall)

mined by a salicylic acid modification of a semimicro-Kjeldahl digestion procedure (Technicon Industrial Systems, 1976).

### Load Calculation for Measured Data

Measured daily, monthly, and annual loads were determined from intensive storm sampling and the mass accumulation integration method for individual runoff events (Eq. [1]). Dissolved NO<sub>3</sub>-N, PO<sub>4</sub>-P, and sediment loads were determined by multiplying the concentration in each discrete sample by the corresponding flow volume and summing these incremental loads for the runoff event duration. Total N and P loads were determined by summing the dissolved and particulate concentrations, multiplying the corresponding flow volume, and summing for the event.

$$L = 10^{-6} \times \sum_{i=1}^n Q_i C_i \quad [1]$$

where  $L$  is load (kg d<sup>-1</sup>),  $Q_i$  is flow volume for the  $i$ th sample (L d<sup>-1</sup>),  $C_i$  is concentration for  $i$ th sample (mg L<sup>-1</sup>), and  $n$  is number of samples.

### Load Estimation with the Regression Method

Four common, low-frequency water quality sampling strategies were produced by selecting specific discrete samples from the complete set of samples collected for each storm at the three watersheds. The analyzed strategies were (i) one sample per storm at a single stage (99.9% exceedence for average daily flow), (ii) one sample per storm at the peak flow, (iii) one sample per storm at a random time, and (iv) one sample in rising flow and one sample in decreasing flow (see Table 2 for detailed descriptions).

For each of the four sampling strategies, selected samples with NO<sub>3</sub>-N, PO<sub>4</sub>-P, sediment, total N, and total P concentrations (mg L<sup>-1</sup>) and corresponding mean daily discharge rates (ft<sup>3</sup> s<sup>-1</sup>) were input into MS Excel for regression analysis. Loads (kg d<sup>-1</sup>) of constituents were computed by multiplying constituent concentration and mean daily discharge. Regression method, which utilizes the relationship between the natural logarithm (ln) transformed constituent concentration and mean daily discharge (Eq. [2]), yielded slope and intercept for each of the analyzed low-frequency sampling strategies.

$$\ln(L) = \beta_0 + \beta_1 \ln(Q) \quad [2]$$

where  $L$  is load (kg d<sup>-1</sup>),  $Q$  is mean daily discharge (m<sup>3</sup> s<sup>-1</sup>),  $\beta_0$  is intercept, and  $\beta_1$  is slope.

Mean daily discharge data ( $Q$ ) available for August 2001 to June 2004 were used in Eq. 2 along with the slope and intercept values calculated from measured data (see above) to estimate loads of constituents. The estimated loads were corrected for logarithm transformation bias by adding half of the standard error of the regression as recommended by Cohn (1995). The estimated loads for each day were summed to calculate monthly and annual loads.

### Comparison between Measured and Estimated Loads

The estimated daily, monthly, and annual loads were then compared with the measured loads. Tests of significant correlation between mean daily flow rate and constituent concentrations and loads were conducted at a priori  $\alpha = 0.05$  probability level. Relative errors (RE) were also calculated (Eq. [3]) for all estimated constituent loads.

$$RE(\%) = ([ML - EL]/ML) \times 100 \quad [3]$$

where ML is measured load (kg ha<sup>-1</sup>), and EL is estimated load (kg ha<sup>-1</sup>).

A modification by Harmel and Smith (2007) of the Nash-Sutcliffe (Nash and Sutcliffe, 1970) coefficient of efficiency ( $R^2_{NS}$ ) was also calculated to evaluate goodness-of-fit between measured and estimated constituent loads considering measurement uncertainty. The average uncertainties for typical scenarios were assigned to measured NO<sub>3</sub>-N, PO<sub>4</sub>-P, total N, total P, and sediment loads according to Harmel et al. (2006a). This comparison is important because the uncertainty in measured hydrologic and water quality data can be substantial and thus should be considered in the evaluation of estimated loads (Beven, 2006; Harmel et al., 2006a; Pappenberger and Beven, 2006).

## Results and Discussion

### Rainfall and Runoff Characteristics

Annual rainfall totals were relatively consistent for the three watersheds (997–1166 mm) throughout the three study years. Annual rainfall for all sites averaged 1070 mm with a SD of 87 mm (Table 1) compared with a long-term (1939–1999) average of 898 mm with a standard deviation of 235 mm for the study area. Annual runoff was relatively consistent within each watershed, but runoff over 3-yr period was less for pasture (126–159 mm, <13% of rainfall) compared with the cultivated (252–353 mm, <27% of rainfall)

Table 3. Mean annual measured nutrient and sediment concentrations and loads in surface runoff.

Year	Concentrations			Loads		
	Cultivated watershed (Y13)	Pasture watershed (W10)	Mixed watershed (Y)	Cultivated watershed (Y13)	Pasture watershed (W10)	Mixed watershed (Y)
	mg L <sup>-1</sup>			kg ha <sup>-1</sup>		
				<u>NO<sub>3</sub>-N</u>		
2001-02	10.3 ± 16.2†	0.24 ± 0.32	4.18 ± 4.5	28.10	0.15	3.90
2002-03	18.7 ± 25.1	0.12 ± 0.12	4.78 ± 4.9	28.90	0.09	7.30
2003-04	3.1 ± 5.8	0.12 ± 0.14	0.93 ± 2.6	5.60	0.22	1.70
				<u>PO<sub>4</sub>-P</u>		
2001-02	0.40 ± 0.19	0.66 ± 0.11	0.32 ± 0.16	1.35	0.54	0.33
2002-03	0.44 ± 0.16	1.11 ± 0.24	0.30 ± 0.08	0.88	1.14	0.58
2003-04	0.46 ± 0.18	0.61 ± 0.24	0.24 ± 0.16	0.96	1.01	0.46
				<u>Total N</u>		
2001-02	17.3 ± 16.7	0.41 ± 0.33	6.81 ± 4.5	50.50	0.26	6.30
2002-03	20.8 ± 24.8	0.29 ± 0.22	5.51 ± 4.9	32.60	0.21	8.70
2003-04	3.6 ± 6.4	0.32 ± 0.16	1.32 ± 3.1	6.60	0.53	2.50
				<u>Total P</u>		
2001-02	3.18 ± 3.54	0.68 ± 0.11	1.31 ± 1.10	10.6	0.56	1.25
2002-03	1.00 ± 0.57	1.13 ± 0.22	0.44 ± 0.19	1.86	1.15	0.84
2003-04	0.57 ± 0.30	0.65 ± 0.24	0.29 ± 0.20	1.16	1.07	0.54
				<u>Sediment</u>		
2001-02	3744 ± 5227	41 ± 13	1072 ± 1260	12,191	38	966
2002-03	727 ± 782	36 ± 44	196 ± 251	1252	19	356
2003-04	120 ± 239	49 ± 36	86 ± 90	227	75	161

† Mean ± SD (n = 3).

and mixed (212–235 mm, <21% of rainfall) watersheds. As a result, fewer runoff events occurred and fewer samples were collected in the pasture watershed than other watersheds during study period (Table 1).

### Measured Constituent Concentrations and Loads

Although annual rainfall and runoff volumes were relatively consistent within watersheds during the study period, annual nutrient (N, P) and sediment concentrations and loads varied considerably for the cultivated and mixed watersheds (Table 3). In contrast, concentrations and loads of nutrients and sediment were relatively consistent for the pasture watershed during three study years. For example, mean annual concentrations of NO<sub>3</sub>-N were greater and more variable for the cultivated (3.1–18.7 mg L<sup>-1</sup>) and mixed (0.93–4.78 mg L<sup>-1</sup>) watersheds compared with the pasture watershed (0.12–0.24 mg L<sup>-1</sup>) during all study years. A similar pattern was observed for the total N concentrations. However, mean annual PO<sub>4</sub>-P concentrations were greater at the pasture (0.61–1.11 mg L<sup>-1</sup>) than cultivated (0.40–0.46 mg L<sup>-1</sup>) and mixed (0.24–0.32 mg L<sup>-1</sup>) watersheds. The greater PO<sub>4</sub>-P concentrations for the pasture watershed could be attributed to multitude of factors, such as (i) surface application of litter in pasture compared with the incorporation in croplands, (ii) tillage in croplands may have enhanced sorption of PO<sub>4</sub>-P on sediments, or (iii) lesser dilution in pasture as runoff was lower (13% of rainfall) than croplands (21–27% of rainfall). Mean annual total P concentrations across all watersheds during the 3 yr typically ranged from 0.29 to 1.31 mg L<sup>-1</sup> except for the cultivated watershed during 2001–2002 (3.18 mg L<sup>-1</sup>). Sediment concentrations

were highly variable from 1 yr to another with greater values at the cultivated (3744 mg L<sup>-1</sup>) and mixed (1072 mg L<sup>-1</sup>) watersheds during 2001–2002, and considerably lower values at the pasture watershed (36–41 mg L<sup>-1</sup>) in all years (Table 3).

The variability in concentrations and loads of nutrients and sediment in these watersheds may be attributed to differing vegetative cover and crop types during the study period. The pasture watershed land use remained consistent (coastal bermudagrass) throughout the study, but cropping patterns and vegetative cover changed in the cultivated and mixed watersheds. For example, the cropping pattern in the cultivated watershed changed from corn production in 2001–2002 and 2002–2003 to wheat production in 2003–2004. Additionally, in 2003–2004 wheat provided surface coverage in the cultivated and mixed watersheds during both the fall and spring rainy periods, which resulted in reduced sediment losses (161–227 kg ha<sup>-1</sup>) than previous years (356–12,191 kg ha<sup>-1</sup>). The inter-annual changes in vegetative cover were also more pronounced in the cultivated and mixed watershed because of the crop production cycles. The timing of precipitation in relation to fertilizer or manure application also contributed to differences between the watersheds, as runoff soon after nutrient application generated higher nutrient concentrations.

### Relation of Daily Flow with Measured Concentrations and Loads

Correlations between NO<sub>3</sub>-N, PO<sub>4</sub>-P, total N, total P, and sediment concentrations (mg L<sup>-1</sup>) and mean daily flow rate were poor and not significant (*P* > 0.05) for all watersheds for individual years and for the entire 3-yr period (data not reported). This illustrates an important difference compared

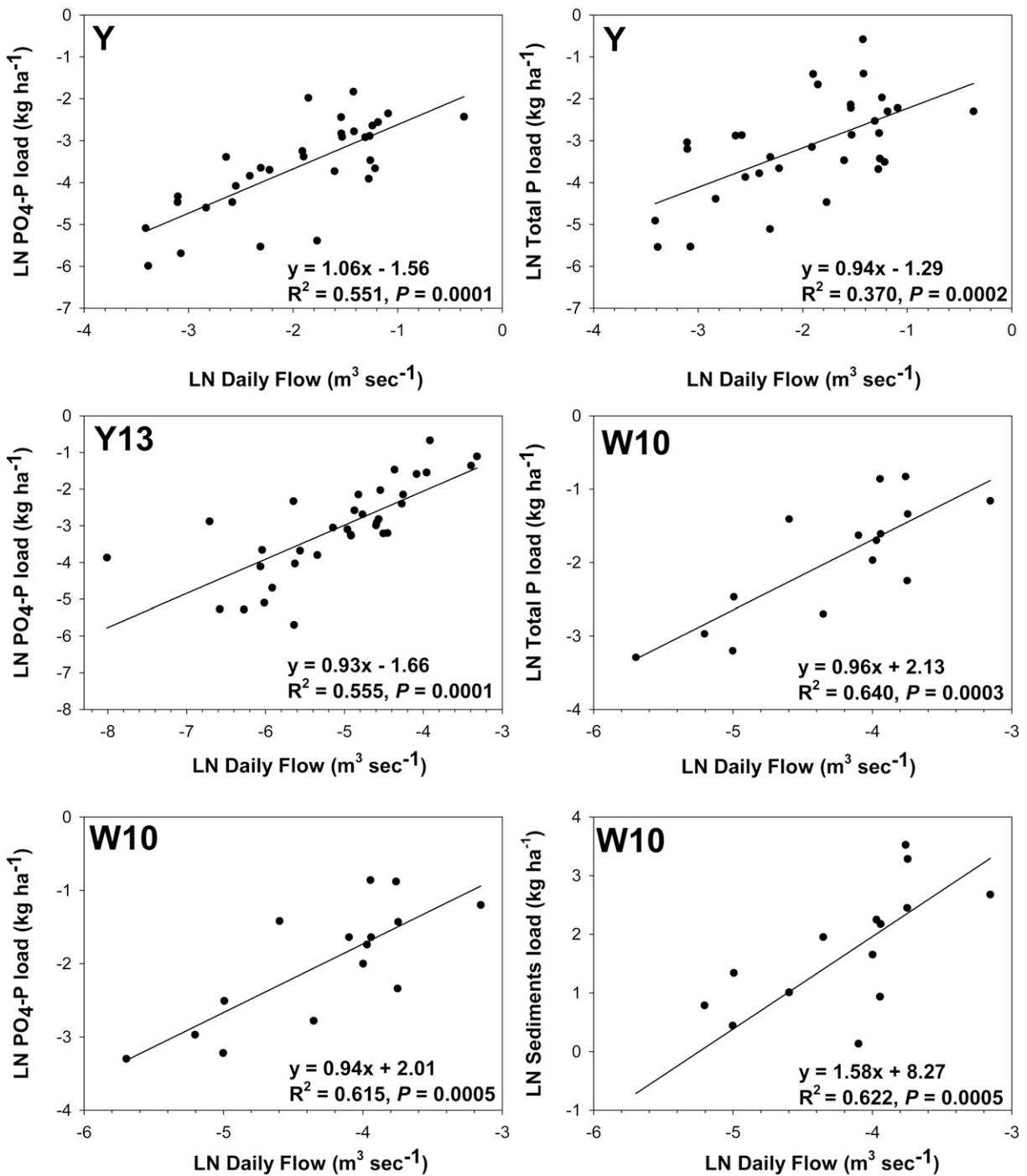


Fig. 1. Relation between daily flow and loads of dissolved P (Y, Y13, W10), total P (Y, W10), and sediment (W10) (2001–2004).

with large rivers, which commonly exhibit significant correlations between flow and constituent concentrations (Cohn, 1995; Ferguson, 1986). For large watersheds, flow and sediment concentrations are typically correlated because flow rate drives bank erosion and sediment resuspension, which contribute a substantial portion of the total sediment load. In contrast, sediment transport from small ephemeral watersheds is largely controlled by overland flow and gully erosion (Edwards and Owens, 1991; Haan et al., 1994). Thus, vegetative cover, which changes substantially throughout the

year in cultivated watersheds but minimally in pasture and rangeland watersheds, can reduce sediment detachment even under adequate transport conditions, thereby reducing sediment transport. In large rivers, the concentrations of dissolved nutrients are typically inversely related with flow because of the dilution.

In contrast to constituent concentrations, several significant correlations were observed between mean daily flow rates and constituent loads (PO<sub>4</sub>-P for all watersheds, total P for mixed and pasture watersheds, sediment for pasture water-

shed) (Fig. 1). The correlations between mean daily flow and  $\text{NO}_3\text{-N}$  and total N loads were not significant for any watershed (data not reported). Dissolved  $\text{PO}_4\text{-P}$  loads and daily flow were significantly correlated for each of the watersheds with  $R^2$  values ranging from 0.55 to 0.62. Similarly, total P load and daily flow were significantly correlated ( $P = 0.0003$ ;  $R^2 = 0.64$ ) for the pasture watershed in part because  $\text{PO}_4\text{-P}$  was a large fraction (more than 94%) of the total P load. In addition, only the pasture watershed exhibited a significant correlation ( $P = 0.0005$ ;  $R^2 = 0.62$ ) between sediment and daily flow. The strength of correlations (pasture > mixed > cultivated) can be attributed to the decreasing trend in temporal changes in vegetative cover, which magnifies the variability in nutrient and sediment transport at these sites (Harmel and King, 2005).

Significant correlations ( $r^2 = 0.36\text{--}0.84$ ) were observed between sediment loads and  $\text{NO}_3\text{-N}$ , total N, and  $\text{PO}_4\text{-P}$  loads for the cultivated and mixed watershed, with the highest correlation between total P and sediment ( $r^2 = 0.84$ ). These significant correlations indicate the important role of sediments in transporting nutrients to water bodies (Bechmann and Stalnacke, 2005; Daniel et al., 1982, Toor et al., 2005).

### Estimation of Annual Nutrient and Sediment Loads

The regression method with concentration data from each of the sampling strategies (single stage, peak, random, rise fall) was quite variable in its ability to accurately estimate annual nutrient and sediment loads for three watersheds (Fig. 2). The varying performance occurred in spite of relatively consistent annual rainfall and runoff volumes (Table 1). The varying performance is probably due to differing timing and intensity of precipitation in relation to fertilizer application, tillage operations, vegetative cover condition, and soil moisture status. Compared with annual load estimates, combined estimates for the 3-yr period were much more accurate. The pattern of decreased relative error for the 3-yr period was also evident when data were grouped by constituent (Fig. 2a), watershed (Fig. 2b), and sampling scheme (Fig. 2c).

Among the constituents, the relative error was generally lowest for  $\text{PO}_4\text{-P}$  (Fig. 2a), which follows from the significant correlations between  $\text{PO}_4\text{-P}$  loads and daily flow for each of the watersheds. The stronger correlations may be attributed to reduced within-event variation of  $\text{PO}_4\text{-P}$  concentrations, as discussed in Harmel and King (2005). The relative errors for estimated annual  $\text{PO}_4\text{-P}$  loads ranged from  $-48\%$  to  $+43\%$ , whereas for the 3-yr period relative errors ranged from 0 to  $+24\%$ , with most values between 0% and  $+12\%$ . Total N, total P, and  $\text{NO}_3\text{-N}$  load estimates were less accurate than  $\text{PO}_4\text{-P}$ . Most of the load estimates had relative errors for individual years ranging from  $-235\%$  to  $+70\%$  and for the 3-yr period from  $-55\%$  to  $+40\%$ . Sediment load estimates were even poorer, with relative errors typically ranging from  $-715\%$  to  $+80\%$  for individual years and  $-60\%$  and  $+55\%$  for the study period (Fig. 2a).

Among watersheds, the pasture watershed exhibited lower relative errors in annual load estimation (Fig. 2b). This result

is attributed to more consistent management and vegetation type in the pasture watershed, which produced stronger correlations between constituent loads and mean daily flow rates. In the mixed and cultivated watersheds, erosion and nutrient loss varied in response to changes in vegetation type and cover within years and between years, which resulted in poorer load predictions using regression methods at this small scale.

Using performance ratings for percent bias (Moriassi et al., 2007), the relative errors for load estimates were described qualitatively. According to this criterion, our relative errors for  $\text{PO}_4\text{-P}$  loads indicated mostly "good" to "very good" estimation for individual years. In contrast, relative error values for  $\text{NO}_3\text{-N}$ , total N, total P, and sediment ranged from "unsatisfactory" to "very good" for individual years. The relative errors for the overall 3-yr study period were mostly "good" to "very good" for  $\text{PO}_4\text{-P}$ ,  $\text{NO}_3\text{-N}$ , total N, and total P but ranged from "unsatisfactory" to "very good" for sediment loads.

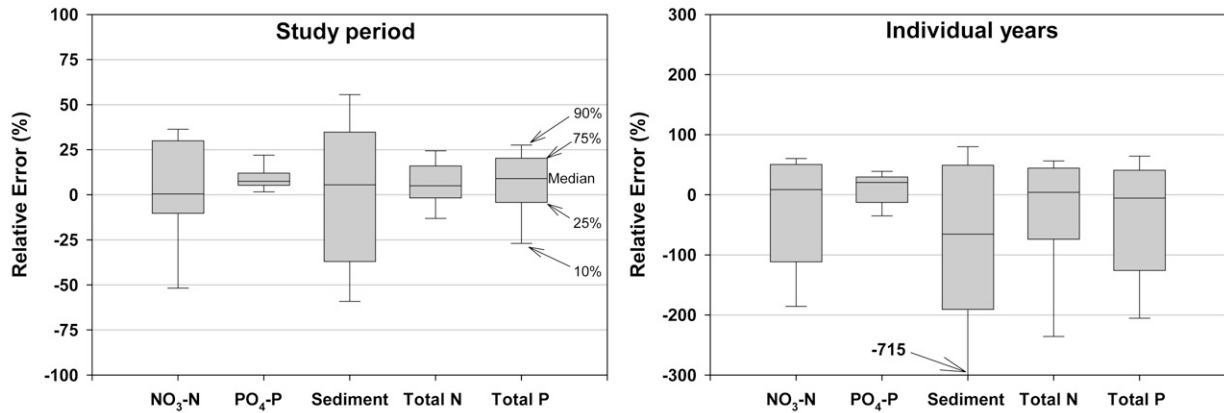
A modification of the Nash-Sutcliffe coefficient ( $R^2_{\text{NS}}$ ) by Harmel and Smith (2007) was used to compare annual constituent load estimates (grouped across watershed type and sampling scheme) with measured loads considering the inherent measurement uncertainty. Average uncertainties for arbitrary "typical" scenarios according to Harmel et al. (2006a) were assigned to measured  $\text{NO}_3\text{-N}$  ( $\pm 17\%$ ),  $\text{PO}_4\text{-P}$  ( $\pm 23\%$ ), total N ( $\pm 29\%$ ), total P ( $\pm 30\%$ ), and sediment ( $\pm 18\%$ ) loads. The modified  $R^2_{\text{NS}}$  values yielded similar conclusions as the correlation and relative error calculations. The modified  $R^2_{\text{NS}}$  values for total N (0.65) and  $\text{PO}_4\text{-P}$  (0.53) loads are "good" and "satisfactory" according to the performance ratings of Moriassi et al. (2007). In contrast, the  $R^2_{\text{NS}}$  values for  $\text{NO}_3\text{-N}$  (0.37), sediment (0.28), and total P ( $-0.02$ ) loads are "unsatisfactory."

### Evaluating the Best Sampling Strategies to Estimate Loads

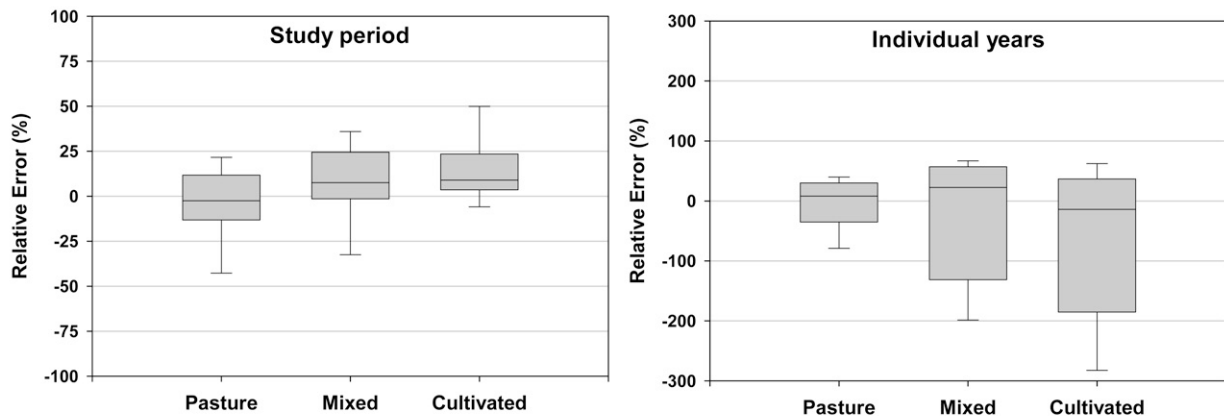
Results of the present study showed little difference in regression-based load predictions between the evaluated low-frequency sampling strategies (Fig. 2c). Most of the relative errors were between  $-260\%$  and  $+70\%$  for individual years and between  $-55\%$  and  $+45\%$  for the study period. Because predictive abilities were similar for all of the low-frequency sampling strategies studied, the choice between sampling strategy depends on other factors, such as time, effort, and cost.

Based on these factors, single-stage and random sampling are recommended to estimate constituent loads with the regression method in typical projects. Single-stage sampling can be conducted with basic supplies and does not require personnel traveling to the site during storm events, which can be difficult in short-duration events typical of ephemeral watersheds (Graczyk et al., 2000). Personnel, however, need to visit the site following runoff events to retrieve collected samples within the quality assurance protocol. Random grab sampling requires personnel traveling to the site during runoff, but this strategy allows immediate preservation of collected samples and collection of a sample in the centroid of flow, which is the recommended protocol (USGS, 1999). The other sampling strategy alternatives (peak, rise fall) require automated sampling equipment with at least a simple data logger program to initiate

### (a) Constituent



### (b) Watersheds



### (c) Sampling scheme

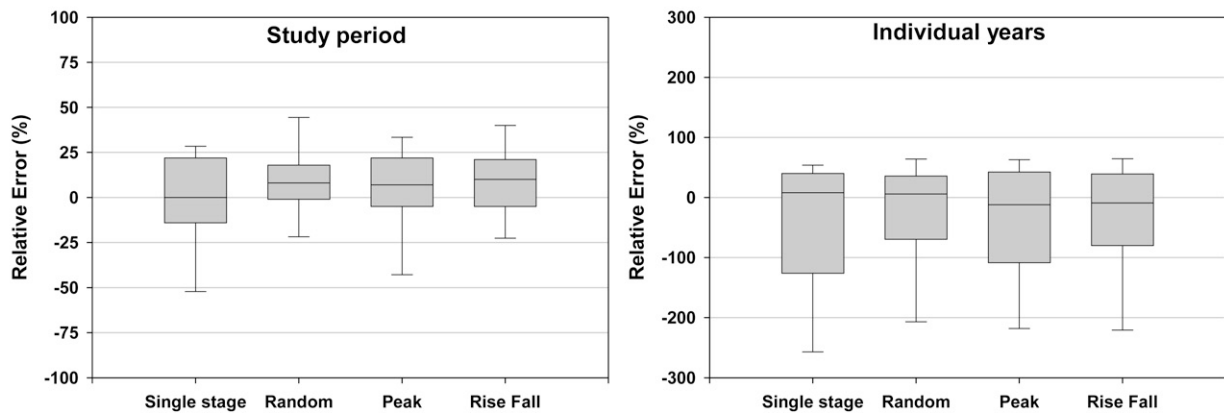


Fig. 2. Relative errors in constituent loads for the entire study period (2001–2004) and for individual years (notice the increased scale for individual years). (A) Constituent (data grouped across sampling scheme and watershed). (b) Watershed (data grouped across sampling scheme and constituent). (c) Sampling scheme (data grouped across watershed and constituent type). Box plots represent the 10th, 25th, 50th, 75th, and 90th percentiles.

sampling. However, with the investment in automated equipment, it is advisable to sample intensively and directly determine constituent loads. The peak and rising-falling sampling strategies could also be performed by personnel, but this would require extended visits to the site during runoff to wait for the

appropriate sampling time, which would be especially difficult for these specific sampling timeframes. Whichever sampling strategy is used, accurate stream flow measurement and appropriate sample handling and analysis must be performed to limit the uncertainty in load estimates (Harmel et al., 2006a).

## Conclusions

Compared with measured loads from three ephemeral watersheds in Texas, regression methodology varied considerably in its ability to estimate nutrient and sediment loads with concentration data from four low-frequency sampling strategies (single stage, random, peak, rise fall). Load estimates were often poor for individual years but improved substantially for the combined 3-yr period. This difference is attributed to the annual variability in constituent source and transport mechanisms that vary based on vegetative cover condition, fertilizer inputs and application timing, and precipitation timing and intensity, all of which affect the correlation between daily flow and concentration. In general, load estimates were better for the pasture watershed than the cultivated and mixed watersheds because of more consistent land management and vegetation type, which produced stronger correlations between constituent loads and mean daily flow rates. It is expected that other watersheds with consistent land use and management, both inter- and intra-annually, would exhibit stronger correlations and accurate load estimates. Load estimates for  $\text{PO}_4\text{-P}$  were consistently better than for the other constituents possibly because  $\text{PO}_4\text{-P}$  concentrations were less variable within storm events. Based on these results, regression estimates methods of annual loads for small watersheds should be used with caution for short periods (annual or less); however, regression methodology can produce reasonably accurate long-term loads where multiple years of water quality data are available. The predictive ability of regression methodology was similar for all of the low-frequency sampling strategies studied; therefore, single-stage or random storm sampling strategies are recommended because of their simplicity for use with regression methods on small watersheds. However, for small watershed projects that require intensive sampling, especially those of short duration, intensive automated sampling remains the recommended alternative.

## References

- Allen, P.M., R.D. Harmel, J.G. Arnold, B. Plant, J. Yelder, and K.W. King. 2005. Field data and flow system response in clay (Vertisol) shale terrain, north central Texas, USA. *Hydrol. Processes* 19:2719–2736.
- Bechmann, M., and P. Stalacke. 2005. Effect of policy-induced measures on suspended sediments and total phosphorus concentrations from three Norwegian agricultural catchments. *Sci. Total Environ.* 344:129–142.
- Beven, K. 2006. On undermining the science? *Hydrol. Processes* 20:3141–3146.
- Cohn, T.A. 1995. Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers. U.S. National report to the International Union Geodesy and Geophysics 1991–1994. *Rev. Geophys. Supplement* 33:1117–1123.
- Cohn, T.A., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers. 1992. The validity of a simple log-linear model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. *Water Resour. Res.* 28:2353–2363.
- Cohn, T.A., L.L. DeLong, E.J. Gilroy, R.M. Hirsh, and D.K. Wells. 1989. Estimating constituent loads. *Water Resour. Res.* 25:937–942.
- Crawford, C.G. 1991. Estimation of suspended-sediment rating curves and mean suspended-sediment loads. *J. Hydrol.* 129:331–348.
- Daniel, T.C., R.C. Wendt, P.E. McGuire, and D. Stoffel. 1982. Nonpoint source loading rates from selected land uses. *Water Resour. Bull.* 18:117–120.
- Edwards, W.M., and L.B. Owens. 1991. Large storm effects on total soil erosion. *J. Soil Water Conserv.* 46:75–78.
- Ferguson, R.I. 1986. River loads understanding rating curves. *J. Water Resour. Res.* 22:74–76.
- Graczyk, D.J., D.M. Robertson, W.J. Rose, and J.J. Steuer. 2000. Comparison of water-quality samples collected by siphon samplers and automatic samplers in Wisconsin. USGS Fact Sheet FS-067-00. USGS, Washington, DC.
- Haan, C.T., B.J. Barfield, and J.C. Hayes. 1994. Design hydrology and sedimentology for small catchments. Academic Press, New York.
- 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. *Appl. Eng. Agric.* 19:187–194.
- Harmel, R.D., R.J. Cooper, R.M. Slade, R.L. Haney, and J.G. Arnold. 2006a. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASAE* 49:689–701.
- 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, B.E. Haggard, D.G. Wren, and J.M. Sheridan. 2006b. Practical guidance for discharge and water quality data collection on small watersheds. *Trans. ASAE* 49:937–948.
- Harmel, R.D., K.W. King, and R. Slade. 2003. Automated storm water sampling on small watersheds. *Appl. Eng. Agric.* 19:667–674.
- Harmel, R.D., C.W. Richardson, K.W. King, and P.M. Allen. 2006c. Runoff and soil loss relationships for the Texas Blackland Prairies Ecoregion. *J. Hydrol.* 331:471–483.
- Harmel, R.D., and P.K. Smith. 2007. Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *J. Hydrol.* 337:326–336.
- Koltun, G.F., M. Eberle, J.R. Gray, and G.D. Glysson. 2006. User's manual for the Graphical Constituent Loading Analysis System (GCLAS). U.S. Geological Survey Techniques and Methods, 4-C1, 51 p. USGS, Washington, DC.
- Moriassi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASAE* 50:885–900.
- Nash, J.E., and J.V. Sutcliffe. 1970. River flow forecasting through conceptual models, Part I: A discussion of principles. *J. Hydrol.* 10:282–290.
- Pappenberger, F., and K.J. Beven. 2006. Ignorance is bliss: 7 reasons not to use uncertainty analysis. *Water Resour. Res.* 42:W05302.
- 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.
- Runkel, R.L., C.G. Crawford, and T.A. Cohn. 2004. Load estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. USGS, Reston, VA.
- Technicon Industrial Systems. 1976. Individual/simultaneous determination of nitrogen and/or phosphorus in BD acid digest. *Industrial Method* 344-74a. Bran-Luebbe, Roselle, IL.
- Toor, G.S., L.M. Condron, B.J. Cade-Menun, H.J. Di, and K.C. Cameron. 2005. Preferential phosphorus leaching from an irrigated grassland soil. *Eur. J. Soil Sci.* 56:155–167.
- USGS. 1999. Section A: National field manual for collection of water-quality data. Techniques of Water-Resources Investigations of the U.S. Geological Survey, Book 9. USGS, Washington, DC.