3-2003

The Development of Relationships Between Constituent Concentrations and Generic Hydrological Variables

Carmen T. Agouridis
University of Kentucky, carmen.agouridis@uky.edu

Dwayne R. Edwards
University of Kentucky, dwayne.edwards@uky.edu

The Development of Relationships Between Constituent Concentrations and Generic Hydrological Variables

Digital Object Identifier (DOI)
https://doi.org/10.13031/2013.12975

Notes/Citation Information
Published in *Transactions of the ASAE*, v. 46, issue 2, p. 245-256.

THE DEVELOPMENT OF RELATIONSHIPS BETWEEN CONSTITUENT CONCENTRATIONS AND GENERIC HYDROLOGICAL VARIABLES

C. T. Agouridis, D. R. Edwards

ABSTRACT. *The collection and analysis of samples from storm events constitutes a large portion of the effort associated with water quality research. Estimating concentrations or loads from these events is often difficult. The equipment necessary to analyze the samples and the required laboratory resources are typically significant expenses incurred by the researcher. One potential method to reduce these costs is through the development of generic relationships between concentrations and easily measured variables such as dimensionless flow rate or time. The benefits recognized from such an effort include a reduction in the number of required samples, resulting in a reduction in cost. Using data collected from an Arkansas stream near Fayetteville, relationships between the generic variables (time and flow) and several constituents (nitrate–N, orthophosphate, total phosphorus, ammonia–N, total Kjeldahl nitrogen, chemical oxygen demand, total suspended solids, fecal coliforms, and fecal streptococci) were examined. Results of the analyses indicated that a form of the gamma function could be used to estimate the flow–weighted mean concentrations and loads of the constituents at a significant cost savings to the user, assuming that single–peak hydrograph data were readily available. By using a single sample collected at the peak of the storm along with information pertaining to the time of sample collection, time of the peak of the storm hydrograph, and the constituent concentration of the sample, the flow–weighted mean concentration or load could be determined. Results of the analysis indicate that the method performed reasonably well. Since the analysis of only one sample is required to determine the flow–weighted mean concentration or load, instead of several samples, this method is quite appealing to users on a limited budget.*

Keywords. *Cost benefit, Flow, Flow–weighted mean concentration, Loads, Model, Water quality.*

Concentrations of water quality constituents are closely related to stream discharge. Certain constituent concentrations, such as total metals, total nutrients, and suspended sediments, increase with increasing discharge, while other constituent concentrations, such as dissolved ions and total dissolved solids, experience a dilution effect with increased stream discharge (Rickert, 1985). Researchers often calculate the flow–weighted mean concentration (FWMC) for each measured constituent to relate the instream constituent load (mass) to the mean discharge for a defined period (volume) (Hoos et al., 2000). This is done to reduce the data into a single value that meaningfully represents the behavior of concentration throughout the entire storm event. The use of the FWMC provides two significant advantages: (1) a portion of the constituent concentration variation associated with the discharge is removed, resulting in better identification of time trends, and (2) sites may be compared with less bias than is often associated with differing sample volume and frequency, flows, or time periods (Cooke et al., 2000; Hoos et al., 2000; Mueller and Stoner, 1998; Rickert, 1985).

While the development of an appropriate sampling strategy to accurately represent the hydrograph is a difficult task, the actual calculation of the FWMC ($\bar{c}$) is quite intuitive. Cooke et al. (2000) and Huber (1993) defined the FWMC as the total mass load (M) divided by the total stream flow volume (V). Knowledge of the time of sample collection with respect to the start of the storm, the discharge at the sampled point, and the concentration of the constituent of interest are the only requirements to calculating the FWMC. Equation 1 illustrates this concept:

$$\bar{c} = \frac{\int M(t)dt}{\int Q(t)dt} = \frac{\int c(t)Q(t)dt}{\int Q(t)dt} \cong \frac{\sum_{i=1}^{n}\left(\frac{c_i + c_{i+1}}{2}\right)\left(\frac{Q_i + Q_{i+1}}{2}\right)}{\sum_{i=1}^{n}\left(\frac{Q_i + Q_{i+1}}{2}\right)} \quad (1)$$

The variable c represents the time–variable concentration, while Q represents the time–variable flow, both measured during the storm event (Cooke et al., 2000; Huber, 1993). The total number of samples during the event is represented by the variable n. Each storm has one $\bar{c}$ for each of the analyzed constituents.

The accuracy of the FWMC for the constituent will depend largely on the implemented water quality sampling strategy. According to the U.S. EPA (1973), periodic

Article was submitted for review in March 2002; approved for publication by the Soil & Water Division of ASAE in December 2002. ASAE Meeting Paper No. 012113.

The authors are **Carmen T. Agouridis, ASAE Member,** Graduate Research Assistant,and **Dwayne R. Edwards, ASAE Member Engineer,** Professor,Department of Biosystems and Agricultural Engineering, University of Kentucky, Lexington,Kentucky. **Corresponding author:** Dwayne R. Edwards, Department of Biosystems and Agricultural Engineering, 128 C. E. Barnhart Building, University of Kentucky, Lexington, KY 40546; phone: 859–257–3000; fax: 859–257–5671; e–mail: dedwards@bae.uky.edu.

collection of samples throughout the storm event is required. This statement was echoed by Mueller and Stoner (1998), who noted that sampling must occur evenly over the hydrograph at both high and low flows to minimize the chance of biased results. Stone et al. (2000) compared the four sampling strategies of "time–composite sampling with continuous flow measurements (TC), flow–proportional sampling with independent measurement of flow (FP), grab sampling with instantaneous flow measurements (IG), and grab sampling for quality assurance/quality control checks using daily USGS flow measurements (UG)." Samples were analyzed for nitrate–N, ammonia–N, and TKN. The researchers noted substantial differences among the sampling techniques, with the FP method resulting in greater calculated or estimated mass loading rates than the other methods (IG, UG, and TC did not differ significantly). The elevated FWMCs for the sampled constituents resulted from the intensive level of sampling during high flows. The FP method over–represented the rising limb of the hydrograph and under–represented the falling limb of the hydrograph, biasing results.

Substantial volumes of research in the area of water quality, especially with regards to streams, involve the determination of FWMCs (Kuykendall et al., 1999; Guillard et al., 1999; Turtola and Yli–Halla, 1999; Wood et al., 1999; McFarland and Hauck, 1999; Moorman et al., 1999). The extensive usage of FWMCs coupled with the difficulty associated with determining this parameter has created a need for defining a more efficient method for estimating concentrations or loads from runoff events. In an effort to reduce the cost and labor involved in collecting and analyzing the multiple samples required to calculate FWMCs, an attempt was made to define relationships between easily measured dimensionless flow rate or time and constituent concentrations. The goal was to estimate the constituent FWMC for a storm event from a single sample. Data from two Arkansas streams located near Fayetteville (Moores Creek and Beatty Branch) were used to develop and verify the method. Constituents examined in this process included nitrate–N (NO_3–N), orthophosphate (PO_4–P), total phosphorus (TP), ammonia–N (NH_3–N), total Kjeldahl nitrogen (TKN), chemical oxygen demand (COD), total suspended solids (TSS), fecal coliforms (FC), and fecal streptococci (FS). The goal was to estimate the FWMC using only one concentration value and the knowledge of associated flow rate or time into the storm event. The one assumption made in the development of the method is that single–peak hydrograph data will be readily available. A benefit of this method is that cost savings will be incurred with a reduced number of samples requiring chemical and/or microbiological analyses.

Methods

Data used in the method development and verification processes were collected over an approximately three–year period from Moores Creek and Beatty Branch, two streams located in the Lincoln Lake basin approximately 19 km west of Fayetteville, Arkansas (Edwards et al., 1994; Edwards et al., 1996). The drainage area contributing to the Moores Creek sampling site was 1800 ha and that of the Beatty Branch sampling site was 795 ha. Land use within the Moores Creek and Beatty Branch watersheds was notably different (table 1). On a relative basis, the Moores Creek watershed had a greater percentage of land use designated as pasture and urban, while Beatty Branch had more land classified as forest. Flow in the streams is typically perennial, although periods of drought have resulted in no flow. The streams are comprised of stony beds (7.6 to 15.2 cm diameters), with small pools forming during low flow periods.

Instrumentation was installed at each site to collect water samples during storm events, and the stream stage was monitored. A pressure transducer (model PCDR950, Druck, Inc.) attached to a flagstone was placed in the stream to measure water height, and stage data were recorded at 5 min intervals with a Campbell Scientific CR10 datalogger. Rating curves were developed in accordance with procedures outlined by the USGS (1969, 1984). An American Sigma automatic water sampler (model 800SL portable liquid sampler) collected samples during the storm events. Samples (1 L sample volume) were collected at the beginning of a storm event and continued at 2 h intervals until the conclusion of the storm. Transport and analyses of the water samples occurred no later than 24 h following each storm event.

Water quality analyses of a subset of each set of runoff samples occurred at the Arkansas Water Resources Center Water Quality Laboratory. The subsets of runoff samples were chosen to ensure adequate representation of the rising and falling limbs of the hydrograph along with the peak. Resource constraints prohibited the analyses of all collected runoff samples. Constituent analyses included NO_3–N (ion chromatography), PO_4–P (ion chromatography), TP (ascorbic acid colormetric method following sulfuric acid digestion), NH_3–N (ammonia–selective electrode method), TKN (macro–Kjeldahl method), COD (closed–reflux, colormetric method), TSS, FC (membrane filter technique), and FS (membrane filter technique). All analyses were conducted in accordance with standard methods (Greenberg et al., 1992).

Selection of Storm Events and Calculation of FWMC

The initial step was the identification of suitable storms through the development of storm hydrographs using data from Moores Creek. These hydrographs were examined, and hydrographs containing multiple peaks or other gross irregularities were eliminated from the process (figs. 1a and 1b). A total of 25 storm events with a single–peak hydrograph were selected from a possible 44 (i.e., the remaining 19 storms had multiple peaks or irregularities) and were used to investigate methods of estimating FWMC from a single concentration value. The size and duration of the peak flow were not factors used in the process of selecting storms. Using a Visual Basic program, the necessary data for each sampled point (constituent concentrations, flow, time, peak flow, and peak time) were extracted from the data set for each storm

Table 1. Land use proportions for Moores Creek and Beatty Branch watersheds, Arkansas.

Land Use	Moores Creek (% Area)	Beatty Branch (% Area)
Pasture	61.8	56.5
Agriculture	4.2	2.3
Urban	7.1	0.0
Forest	25.9	39.5
Other	1.0	1.7

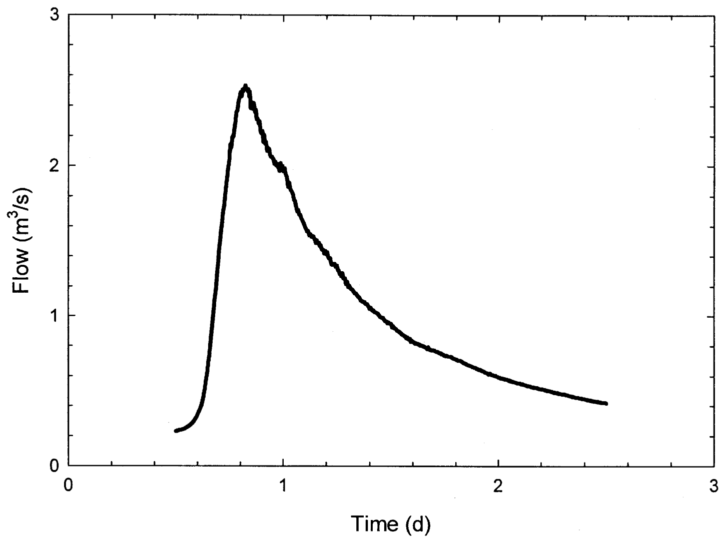

Figure 1a. Example of a suitable storm (5 February 1993 from Beatty Branch).

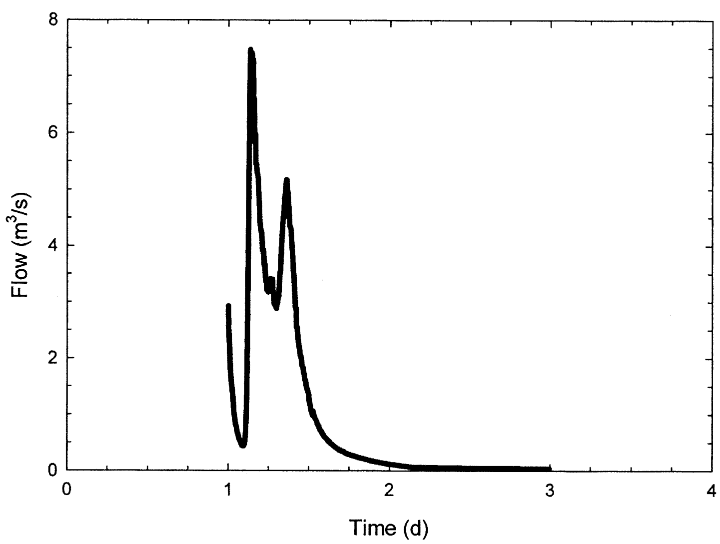

Figure 1b. Example of an unsuitable storm (13 September 1993 from Beatty Branch).

event. Once the suitable storms were identified, the FWMC ($\bar{c}$) was calculated using equation 1 for each parameter and each storm.

Selection of Generic Hydrological Parameters

Flow rate and time since the beginning of storm flow, two easily measured hydrological parameters, were selected for comparison with the FWMC with the goal of identifying relationships between the various $\bar{c}$ values and either time or flow rate. To convert all of the parameters to dimensionless quantities, the constituent concentrations at each sampled time ($c_{i,j}$) were divided by the FWMC ($\bar{c}_j$), while the associated flow (Q_i) and time (t_i) at each sampled point were divided by their peak flow rate (Q_p) and time to peak (t_p) of the storm hydrograph, respectively. The data set for each storm thus consisted of a set of values of $c_{i,j}/\bar{c}_j$, t_i/t_p, and Q_i/Q_p, where i is the index for the number of samples, and j is the index for the number of parameters. Graphs of $c_{i,j}/\bar{c}_j$ versus Q_i/Q_p and $c_{i,j}/\bar{c}_j$ versus t_i/t_p were constructed for use in determining which hydrologic parameter ratio (t_i/t_p or Q_i/Q_p) exhibited the strongest relationship with $c_{i,j}/\bar{c}_j$ (fig. 2). The initial expectation for the flow ratio parameter (Q_i/Q_p) was that the concentration ratios generated from the rising and falling limbs of the storm hydrograph would travel along the same path. The higher constituent concentration values produced in the initial stages of the storm in comparison with those generated in the latter stages of the storm resulted in hysteresis. Based on the amount of hysteresis present, the flow parameter ratio (Q_i/Q_p) was removed from further consideration in the modeling process. The time parameter ratio (t_i/t_p) continually increased throughout the duration of the storm, so hysteresis was not possible. Subsequent efforts were therefore focused on relating the measured $c_{i,j}/\bar{c}_j$ ratios to the time parameter ratio (t_i/t_p).

Examination for Seasonality

The suitable storm events were separated into seasons consisting of three months each: January through March (winter), April through June (spring), July through September (summer), and October through December (fall). Fol– lowing the division of the 25 storm events into seasonal

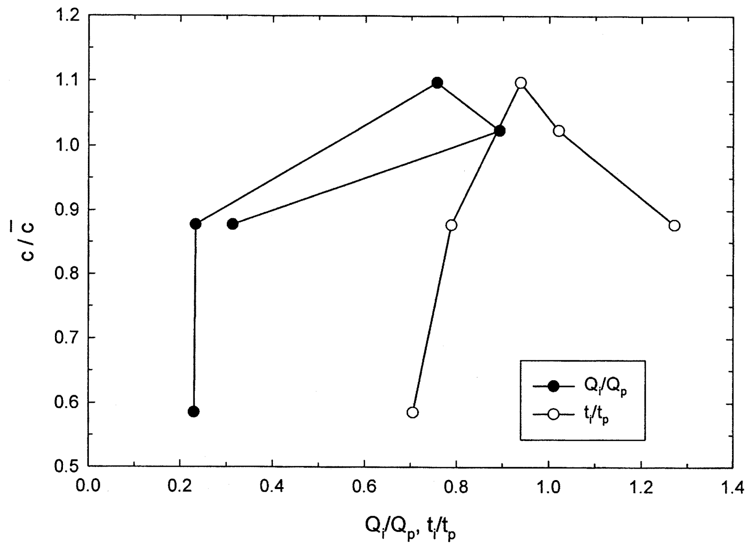

Figure 2. Relationship between dimensionless concentration ($c / \bar{c}$) and dimensionless time (t_i/t_p), dimensionless flow (Q_i/Q_p) for PO_4–P for the 1 April 1994 storm on the Moores Creek.

categories, the constituent ratios ($c_{i,j}/\bar{c}_j$ versus t_i/t_p) from the appropriate storms for each season were individually graphed using the same format as described in the previous section. The resultant graphs were visually examined to determine if seasonal trends were present in the data (i.e., whether relationships between $c_{i,j}/\bar{c}_j$ and t_i/t_p were different for different seasons), potentially highlighting the need to develop seasonal models relating $c_{i,j}/\bar{c}_j$ to t_i/t_p.

Model Choice

The graphs of $c_{i,j}/\bar{c}_j$ versus t_i/t_p were examined for each storm and each parameter to identify a family of curves that best represented the relationships. Based on the characteristics displayed by the relationships between the concentration ratio ($c_{i,j}/\bar{c}_j$) and the time parameter (t_i/t_p), a form of the gamma distribution function was chosen to describe the relationship (Haan, 1977):

$$\frac{c_{i,j}}{\bar{c}_j} = \lambda^{\eta}\left(\frac{t_i}{t_p}\right)^{\eta-1} e^{-\left(\frac{t_i}{t_p}\right)\lambda} \tag{2}$$

Since no requirement existed for the function to integrate to one, the denominator of the gamma function as reported by Haan (1977) was eliminated. The gamma distribution is often used in hydrology for determining rainfall probabilities (Barger and Thom, 1949; Barger et al., 1959; Friendman and Janes, 1957, Mooley and Crutcher, 1968) and annual runoff probabilities (Markovic, 1965).

The constants for the gamma function (λ and η) were calculated using Solver in Microsoft Excel® such that the sum of squared errors for $c_{i,j}/\bar{c}_j$ (i.e., the squared difference between the measured and the predicted values for $c_{i,j}/\bar{c}_j$) was minimized. These gamma constants were determined for each storm and parameter. The gamma constants for the individual storms were then arithmetically averaged across all storms to determine the resulting model constants for each parameter. After identifying λ and η, the storm $\bar{c}_j$ can be estimated from any concentration value $c_{i,j}$:

$$\bar{c}_j = \frac{c_{i,j}}{\lambda^{\eta}\left(\frac{t_i}{t_p}\right)^{\eta-1} e^{-\left(\frac{t_i}{t_p}\right)\lambda}} \tag{3}$$

Selection of Appropriate Sampling Period

Since the goal of the project was to develop a method to determine constituent FWMC for a storm event from only one sample point, the next question was related to determining when the optimum point to collect a sample during the storm occurred, so that the error in estimating $\bar{c}_j$ was minimized. The relative error (E_r) and absolute error (E_a) for each constituent for each sampled point during each storm were calculated. Note that within a storm having n samples, there will be n calculations of E_r: one for each sample, since the concentration value of each sample can be used to estimate $\bar{c}_j$.

$$E_r = \frac{\sqrt{\left(\bar{c}_{j,o} - \bar{c}_{i,j,p}\right)^2}}{\bar{c}_{j,o}} \tag{4}$$

$$E_a = \left(\frac{\bar{c}_{j,o} - \bar{c}_{i,j,p}}{\bar{c}_{j,o}}\right) \tag{5}$$

The subscript o denotes values estimated from observed flow and concentration data (eq. 1), while the subscript p indicates values predicted by equation 3. Graphs of relative error and absolute error versus t_i/t_p were constructed for each constituent (figs. 3 and 4). Note that the values of t_i/t_p are those associated with all sampling times for a particular constituent. The optimum sampling point was selected as the time that produced the minimal values of E_r and E_a. The optimal times were the same for both E_r and E_a, since the only difference between the two graph types was that the E_r was a forced positive value, while the E_a provided information regarding under or over estimation.

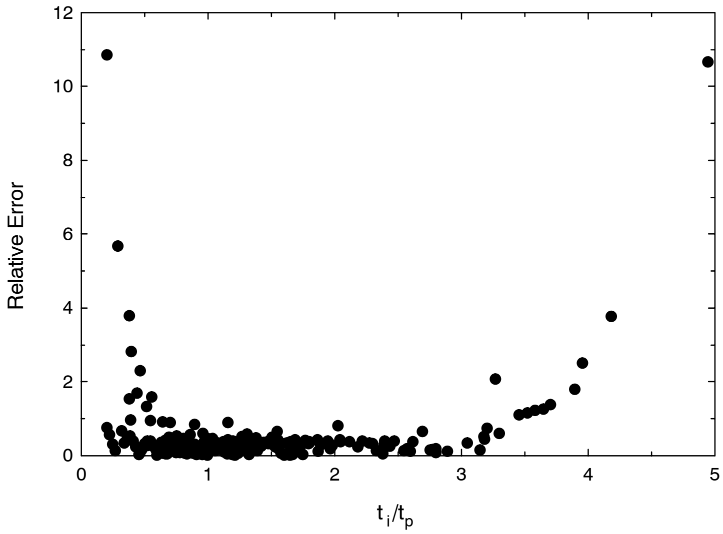

Figure 3. Relative error as a function of dimensionless time (t_i/t_p) for NO_3–N for the Moores Creek data set.

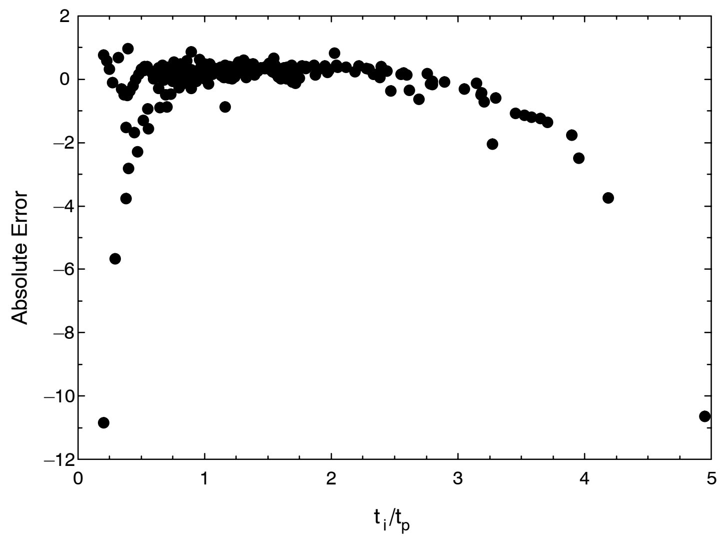

Figure 4. Absolute error as a function of dimensionless time (t_i/t_p) for NO_3–N for the Moores Creek data set.

METHOD VERIFICATION

Storm event data from the Beatty Branch monitoring station were used to verify the method by the previously described procedures. From approximately three years of data, 20 storm events out of 63 possible storm events were selected for use in the verification process (i.e., 43 storms with multiple peaks or irregularities were eliminated). The same criteria for selecting storms that were used in the development of the method were also used to select storms for the verification of the method. Only storms with a single–peak hydrograph were selected; storms that contained multiple peaks or other gross irregularities were eliminated. Storms were not eliminated based on the size of the peak flow or the duration. In accordance with the method development, the dimensionless hydrologic parameter t_i/t_p was used. The predicted FWMC was calculated from equation 3 using the concentration value nearest to the peak of the storm (i.e., the value of $c_{i,j}$ used to determine $\overline{c}_j$ was that occurring nearest the peak) because in the method development, the relative error (E_r) and the absolute error (E_a) were minimum, indicating that the peak of the storm was the optimum sampling point for the method (figs. 3 and 4). The observed FWMC, based on all available samples, was calculated from equation 1. These values were used in the method verification process. If a measured sample was not taken exactly at the peak of the storm, then the next closest point was selected. For storms that contained multiple points near the peak, the data point for use in the statistical analyses was randomly chosen. Since seasonality trends were not detected with the data set used in the model development (i.e., seasonal models were not developed), these trends were not examined with the verification data set. The flow–weighted mean loads for each constituent were also examined because of the current movement toward developing managerial practices based on load data.

To measure the success of the method in predicting the FWMC and loads, a linear regression using SAS® was performed for each constituent (SAS, 1996). The measured values of $\overline{c}_j$ were compared to the predicted values of $\overline{c}_j$ using the equation $y = x$ where the predicted values were the dependent variables and the observed values were the independent variables. The flow–weighted mean loads were also examined using the same procedure. Since the linear regression program used in SAS® tested the null hypothesis that the slope equals zero, Student's t–test was performed to test the desired null hypothesis that the slope equaled one. The SAS® output for the intercept was appropriate since the tested null hypothesis was that the intercept equaled zero. Examination of the slopes and intercepts would allow for the identification of systematic error (i.e., whether or not the user would need to adjust the predicted results). Assessment of the capability of the method to successfully predict the FWMC and loads was in part based upon the coefficient of determination, the mean square error, the coefficient of variation, the slope, and the intercept of the regression analyses.

To understand why the method performed better with certain constituents, calculations were performed to determine the mean squared relative errors ($MSRE_j$) between the observed $\bar{c}$ and the predicted $\bar{c}$, and then to compare these errors among the different constituents. The success of the gamma function in modeling a constituent's concentration over time was reflected in the ability of the method to predict the $\bar{c}$ values for one constituent versus another. The amount of variation in the data from the gamma function would be reflected in higher MSRE values. These calculations were performed for each sampled point (*N* samples, *n* values of SE_j) on the constituent data from Moores Creek where:

$$\mathrm{MSRE}_j = \mathrm{E}_r^{\,2} \tag{6}$$

Once the $MSRE_j$ was determined, the value was summed to provide a total MSRE for each storm. To compare all of the constituents, the $MSRE_{all}$ was calculated:

$$\mathrm{MSRE}_{all} = \frac{\sum \mathrm{MSRE}_j}{N} \tag{7}$$

where *N* represents the number of storms in the data set. Using only the $MSRE_j$ values at the peak of the storm (i.e., $t_i/t_p = 1$), the $MSRE_{peak}$ was determined to accertain how well the model performed when only data at the peak were selected, as opposed to the selection of a random point as seen with $MSRE_{all}$.

Results

Potential seasonality trends were examined by graphing storms into groups of four seasons (seasons were based on whole months that followed the calendar seasonal divisions as closely as possible). Graphs of the seasonal storm events were examined for potential trends related to t_i/t_p versus $c_{i,j}/\bar{c}_j$. Seasonal trends related to Q_i/Q_p versus $c_{i,j}/\bar{c}_j$ were not examined since the parameter Q_i/Q_p had been eliminated from the method development process. Aspects that were examined included multiple storms within a season exhibiting the same pattern, time to peak, variation in constituent concentration levels across seasons, and constituent associations. The examined constituent associations included changes in the nitrogen species (NO_3–N, TKN, and NH_3–N), changes in phosphorus (PO_4–P and TP), and solids (TSS) and bacteriological changes related to FC and FS. Seasonality trends, assessed via visual examination, were not noted with any of the sampled constituents.

With the selection of a form of the gamma distribution as the appropriate model for describing the relationship between t_i/t_p and $c_{i,j}/\bar{c}_j$, constants for the model were determined. The calculated model gamma constants, λ and η, were similar for each constituent (table 2). Mean values for λ ranged between 1.52 and 2.40, while mean values for η ranged between 2.92 to 3.64. The coefficient of variation for each constituent and for each constant was examined to determine the amount of variability associated with the constants relative to the mean. According to Saxton (1999), the coefficient of variation associated with biological data is generally 10% to 30%. For the gamma constant λ, the largest coefficient of variation was 39%, with an average value of 24%. As for the gamma constant η, the largest coefficient of variation was 40%, with an average of 18%. These results indicate that the variability of the calculated gamma constants was in the acceptable range for the majority of the constituents. For all of the constituents except COD and TSS, the variability associated with λ fell within the acceptable range. Only FC and NO_3–N did not meet the acceptability criteria for the constant η.

The successful application of the method required that the user have knowledge pertaining to the accuracy of the method at various points during the storm. In other words, did the method exhibit a greater level of accuracy at the onset, peak, or conclusion of the storm, or was the level of accuracy fairly constant throughout the duration of the storm event? Based on the graphs of relative error (E_r) and absolute error (E_a), the peak of the storm (i.e., $t_i/t_p = 1$) was determined to be the optimum sampling point for the model. Examination of figures 3 and 4, however, revealed that a high range of indifference surrounded the peak, indicating that the user did not have to collect a sample exactly at the peak for the successful performance of the method. A sufficient level of

Table 2. Average gamma model parameters for Moores Creek data set (25 storms).

	Constituent								
Parameter	COD	FS	FC	NH_3–N	NO_3–N	PO_4–P	TKN	TP	TSS
λ_{mean}	2.16	1.66	1.68	1.52	1.69	1.57	1.83	1.79	2.40
CV[a]	0.33	0.20	0.24	0.20	0.28	0.17	0.21	0.19	0.39
λ_{max}	4.58	2.43	2.56	2.11	3.13	2.26	2.64	2.62	5.79
λ_{min}	1.39	0.85	0.98	0.97	1.08	1.26	0.93	1.20	0.76
η_{mean}	3.13	3.20	3.44	3.50	3.64	3.35	3.18	3.09	2.92
CV[a]	0.10	0.18	0.33	0.17	0.40	0.09	0.12	0.12	0.14
η_{max}	3.64	4.91	6.77	5.17	10.54	3.88	4.16	3.85	3.76
η_{min}	2.30	2.17	2.40	2.69	2.70	2.73	2.41	2.47	1.97

[a] Denotes coefficient of variation.

accuracy was obtainable for samples collected near the peak, although the goal remained to collect the sample at the peak.

While the design of the sampling procedure is the responsibility of the user, the authors suggest two sampling strategies. The first option requires the user to interface an automatic sampler and stage recorder with a datalogger or other controller. A simple program could be written such that the controller would examine the stage to detect the point on the hydrograph when the flow began to decrease. Upon the detection of the decreasing flow (i.e., the passing of the peak of the storm), the automatic sampler would collect the desired amount of water. To ensure that the measured lower stage value actually signified the falling limb of the hydrograph and not a fluctuation resulting from a disturbance, the authors suggest establishing the moment of sample collection as a percentage of the previous stage. To handle multi–peak hydrographs, additional programming logic would be required to prevent the failure of this option.

The second sampling strategy requires less input than the first, although it too requires the interfacing of an automatic sampler and stage recorder to a controller. However, the collection of samples would ensue after a defined stage was reached. Sample collection would then continue on a time–incremental basis until the stage fell below the defined stage threshold value. The data used to develop and verify this method were collected in accordance with the second sampling strategy (see discussion under Methods). However, this method was developed only for single–peak hydrographs and was not designed for use with multi–peak hydrographs or those with gross irregularities. Determination of the FWMC would require another technique than that method presented here for non–single–peaked hydrographs. With either the above suggested sampling strategies or any other, the authors suggest testing of the sampling strategy prior to the actual data collection to ensure proper function.

Verification of the method was performed for the flow–weighted mean concentrations and loads for each constituent parameter. Ideally, the observed and predicted values for $\bar{c}$ would be equal, resulting in a line with the equation $y = x$. The simple linear regression model provided an understanding of the strength of the relationship between the observed and the predicted $\bar{c}$ for each constituent (figs. 5a and 5b). The performance of the model was assessed based on a coefficient of determination (R^2) of 0.9 or greater for successful performance, 0.8 to 0.89 for moderately successful performance, and 0.5 to 0.79 for fair performance. These values equate to a minimum correlation of 0.95 for a successful model, 0.89 for a moderately successful model, and 0.71 for a fairly successful model. While these ranges are subjective in the categorical determination of the method, they do not attribute success to a model that did not display at least a 70% linear relationship (correlation) between the observed and predicted FWMC and load values (Saxton, 1999).

In accordance with the stated criteria, the model was successful in predicting the flow–weighted mean concentration for several of the examined constituents (table 3). Based on the results for the flow–weighted mean concentration, the model performance was successful for the constituents NO_3–N, NH_3–N, TKN, and FS; moderately successful for the constituents COD, TSS, and TP; and only fair for PO_4–P and FC. However, the results from statistical analyses indicated that the null hypotheses that the slope (FWMC: NO_3–N and FS; FWML: NO_3–N, COD, TP) equaled one and the intercept (FWMC: FS) equaled zero were true only for a portion of the concentrations and loads. The user would need to perform an adjustment to the resultant FWMC and FWML to correct for systematic error (table 3). One technique of adjustment is to subtract the intercept of the line and multiply by the reciprocal of the slope, as illustrated in equations 8 and 9. This procedure is demonstrated with NH_3–N_p having a slope of 1.153 and an intercept of 0.004:

$$NH_3 - N_p = 1.15 NH_3 - N_o + 0.004 \quad (8)$$

where the subscript p denotes the predicted value, while o represents the observed value. The corrected predicted value for NH_3–N_o can be obtained by rearranging equation 8 as:

$$NH_3 - N_o = \frac{NH_3 - N_p - 0.004}{1.153} \quad (9)$$

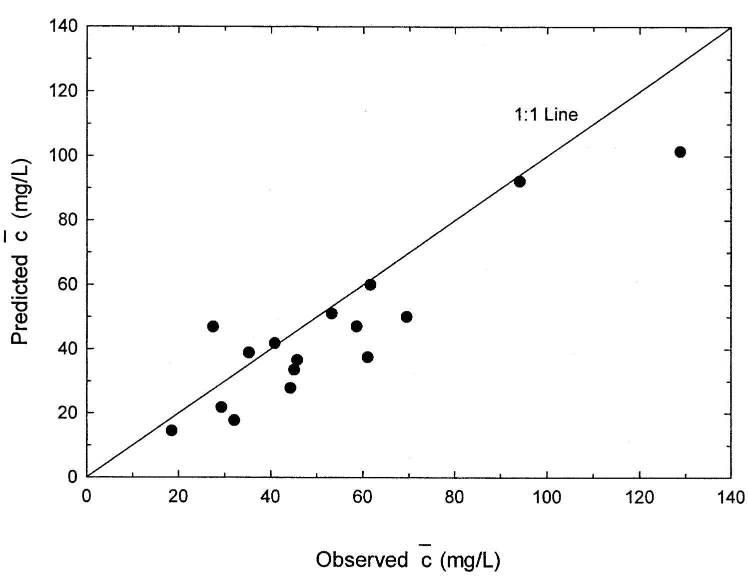

Figure 5a. Prediction of flow–weighted mean concentration for chemical oxygen demand from the Beatty Branch data set.

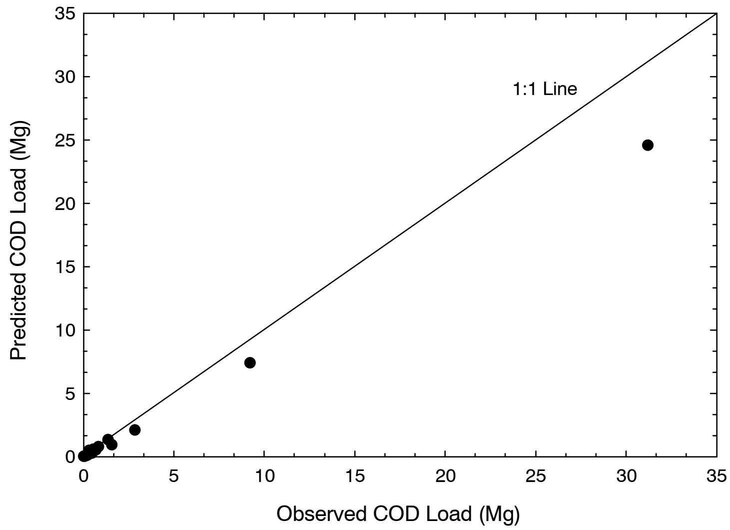

Figure 5b. Prediction of flow–weighted mean load for chemical oxygen demand from the Beatty Branch data set.

Table 3. Results of regressing predicted flow–weighted mean concentration (FWMC) versus observed flow–weighted mean concentration and predicted flow–weighted mean load (FWML) versus observed flow–weighted mean load for the Beatty Branch data set.

	FWMC[a]				FWML[a]			
Parameter	Slope (mg/L)	Intercept (mg/L)	R^2	MSE (mg/L)	Slope (mg)	Intercept (mg)	R^2	MSE (mg)
NO_3–N	0.803	0.004^R	0.99	0.05	0.777	–4.9e5^R	0.99	4.1e6
NH_3–N	1.153^R	0.004^R	0.98	0.05	1.190^R	–2.0e5^R	0.90	2.9e6
COD	0.977^R	–4.227^R	0.83	13.85	0.790	4.9e7^R	0.99	1.4e8
PO_4–P	0.804^R	0.044^R	0.66	0.09	0.756^R	9.8e5^R	0.71	3.7e6
TKN	0.802^R	0.280^R	0.95	1.60	0.949^R	6.3e6^R	0.99	2.5e7
TP	1.050^R	0.028^R	0.84	0.08	1.150	–2.2e5^R	0.97	6.9e6
TSS	1.090^R	11.25^R	0.81	25.65	1.200^R	3.8e8^R	0.98	2.0e9
	(cfu/100 mL)	(cfu/100 mL)		(cfu/100 mL)	(cfu/100 mL)	(cfu/100 mL)		(cfu/100 mL)
FC	1.430^R	–1.1e3^R	0.55	1.2e4	—	—	—	—
FS	1.890	–1.1e4	0.99	1.0e4	—	—	—	—

[a] The null hypotheses were that the slope equaled one and the intercept equaled zero for the regression of the observed $\bar{c}$ versus the predicted $\bar{c}$. The superscript R indicates that the null hypothesis was rejected at the 5% level of significance.

To understand why the method performed better for certain constituents than for others, two items were examined: plots of the observed $\bar{c}$ versus predicted $\bar{c}$ for the constituents, and the $MSRE_{all}$ and $MSRE_{peak}$ values for the constituents (table 4). The plots indicated the presence of possible outliers for the constituents PO_4–P, TKN, FC, NH_3–N, and FS (figs. 6a through 6e). For PO_4–P, FC, and TKN, these influential points resulted in lower R^2 values, while the opposite was true for NH_3–N and FS. The constituents COD, TSS, TP, and NO_3–N demonstrated a wider range of fluctuation or variation (figs. 5a, 6f, 6g, and 6h). Examination of the $MSRE_{all}$ in table 4 for Moores Creek revealed a different order of predicted constituent performance than that indicated by the linear regression in table 3 (worst to best: TSS, TKN, COD, TP, NO_3–N, FC, FS, PO_4–P, and NH_3–N). For $MSRE_{peak}$, the results were (worst to best) FS, TSS, NH_3–N, FC, TP, PO_4–P, NO_3–N, COD, and TKN. While the mean squared relative error calculations were performed on a different data set than that used in the linear regression analysis, the results exhibited similarity. These analyses do not answer the question of why one constituent performed better than another, but they do shed light on the relative error one can expect from this method when a sample is randomly selected versus selection at the peak of the hydrograph. In some constituents, such as TKN, TSS, and COD, over a 100–fold decrease in the MSRE was achieved by sampling at the peak. Others, such as NO_3–N and TP, had a moderate reduction (approximately a 20–fold decrease). For FS, NH_3–N, FC, and OP, sampling at the peak only slightly improved the MSRE.

As for the flow–weighted mean loads, the model performed successfully for all of the examined constituents except PO_4–P, whose fit was fair. Interestingly, the model

Table 4. Mean squared relative error (MSRE) results for the Moores Creek data set.

Constituent	$MSRE_{all}$[a]	$MSRE_{peak}$[b]
NO_3–N	1.86	0.10
NH_3–N	0.70	0.28
COD	13.38	0.08
PO_4–P	0.97	0.14
TKN	18.50	0.05
TP	4.70	0.18
TSS	237.85	0.33
FC	1.51	0.28
FS	1.41	0.43

[a] For all samples.
[b] For samples only at t_i/t_p.

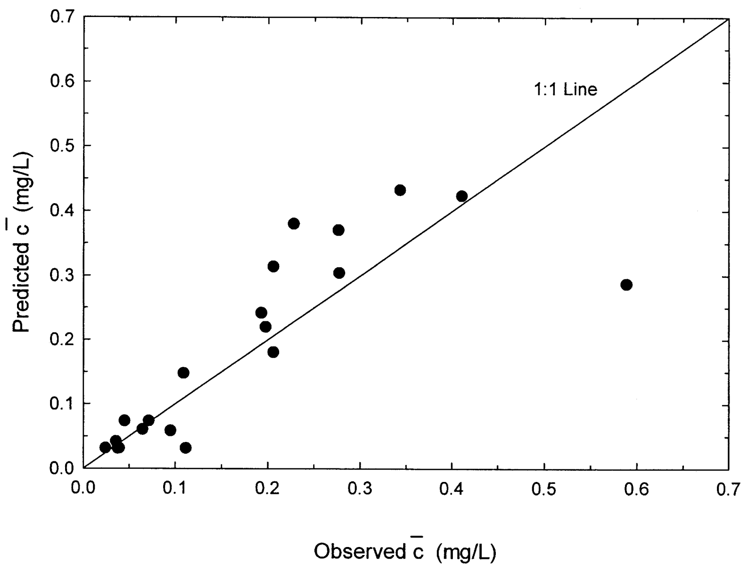

Figure 6a. Prediction of flow–weighted mean concentration for PO_4–P from the Beatty Branch data set.

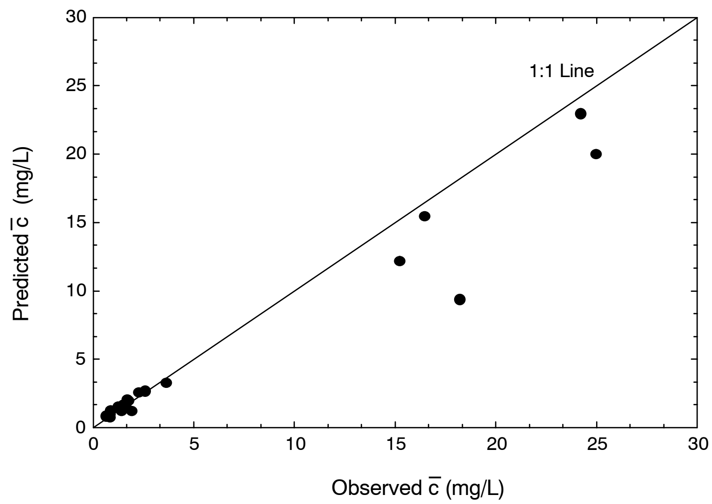

Figure 6b. Prediction of flow–weighted mean concentration for total Kjeldahl nitrogen from the Beatty Branch data set.

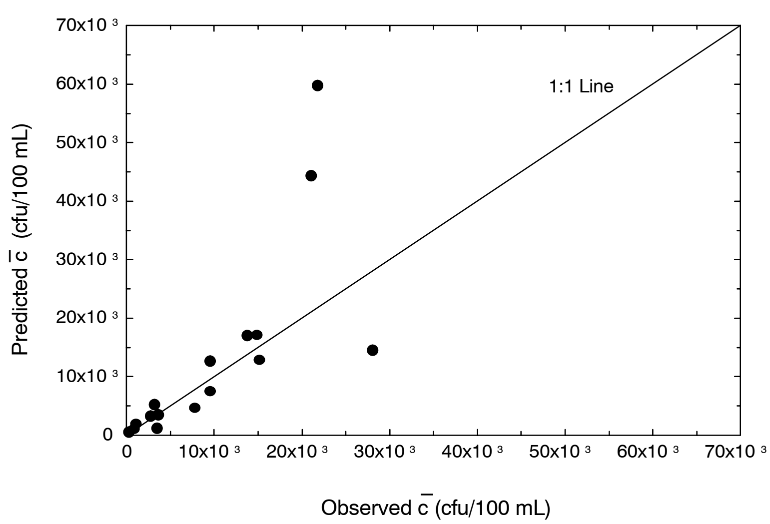

Figure 6c. Prediction of flow–weighted mean concentration for fecal coliforms from the Beatty Branch data set.

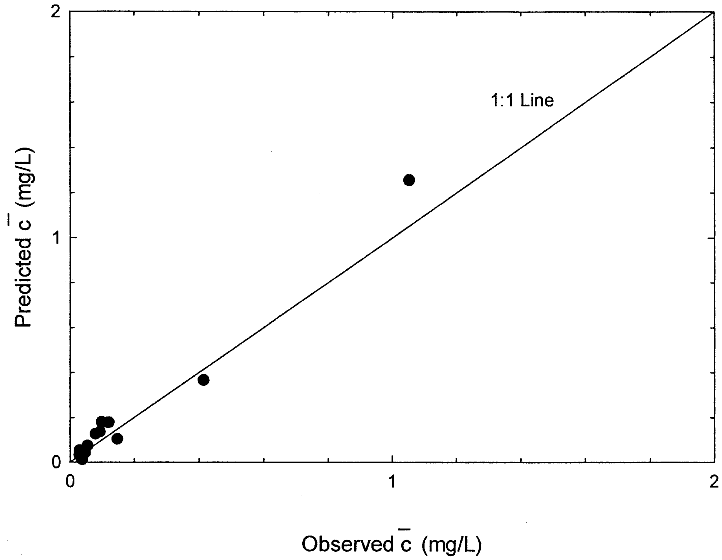

Figure 6d. Prediction of flow–weighted mean concentration for NH_3–N from the Beatty Branch data set.

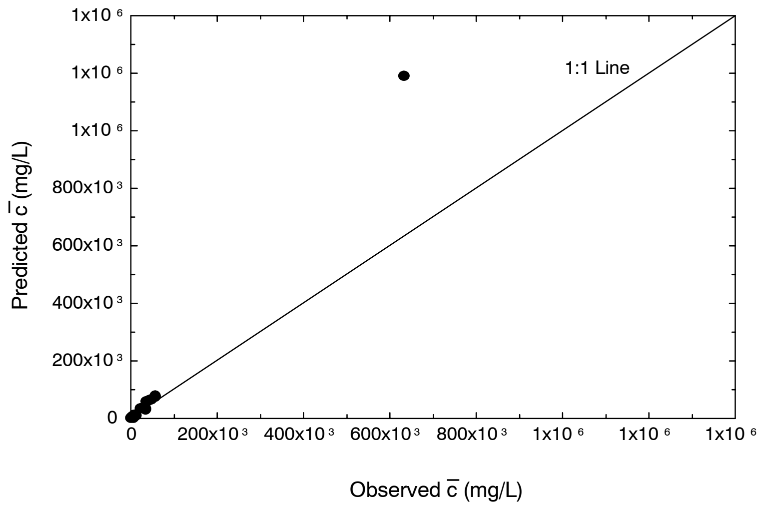

Figure 6e. Prediction of flow–weighted mean concentration for fecal streptococci from the Beatty Branch data set.

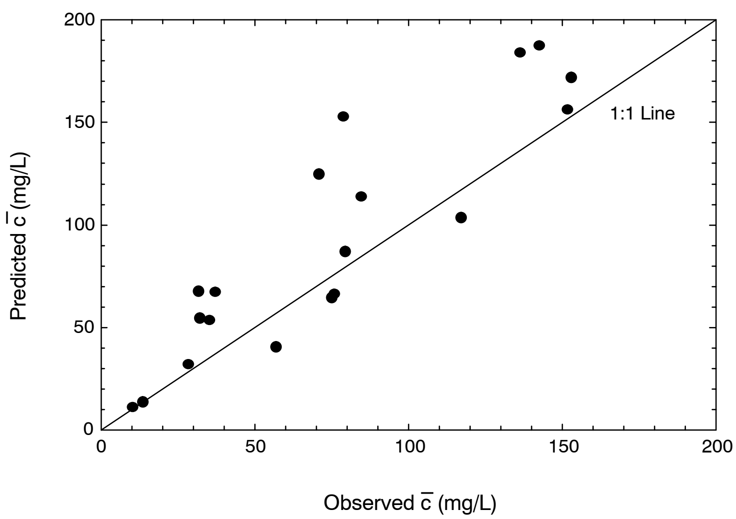

Figure 6f. Prediction of flow–weighted mean concentration for total suspended solids from the Beatty Branch data set.

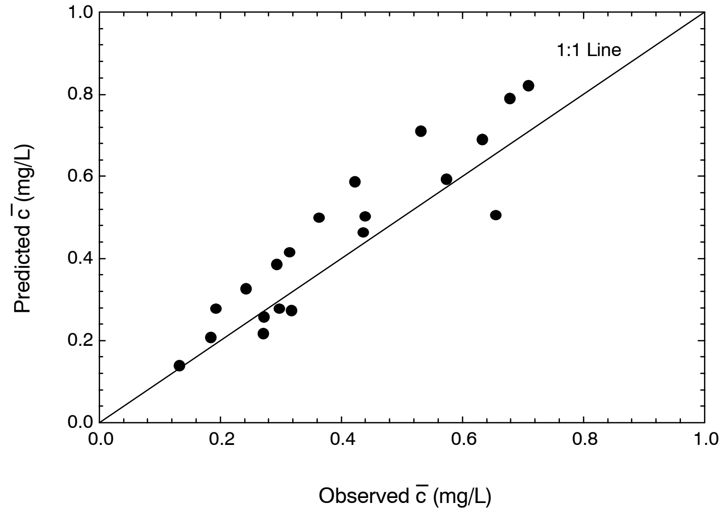

Figure 6g. Prediction of flow–weighted m ean concentration for total phosphorus from the Beatty Branch data set.

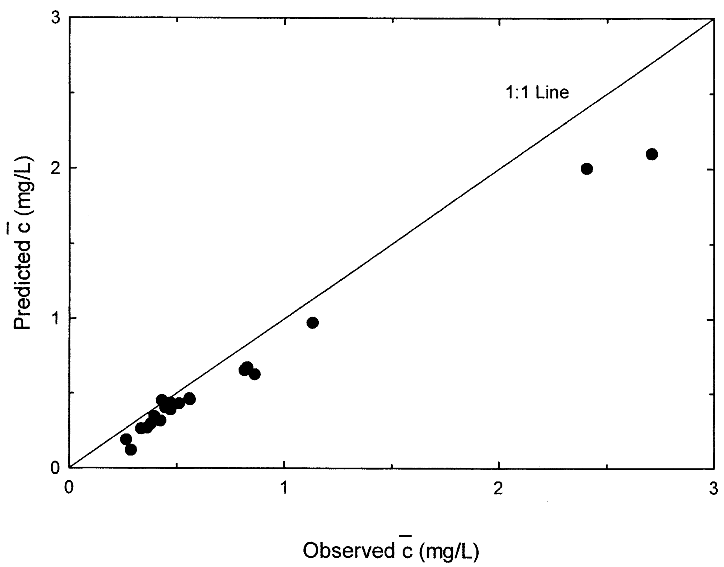

Figure 6h. Prediction of flow–weighted mean concentration for NO_3–N from the Beatty Branch data set.

appeared to predict the flow–weighted mean load better than the flow–weighted mean concentration (table 3). The hydrographs used in the development and verification of the method had shapes that were characteristically represented by the gamma function. By multiplying the constituent concentrations by the corresponding flow volumes, the concentration data were essentially formed to the gamma function, as the magnitude of the flow was much greater than that of the concentration. The small fluctuations or variabilities in the observed and predicted concentrations became less influential as their effects were masked by the magnitudes of the flow volumes.

Conclusions

The development of this method highlighted the possibility of determining both the flow–weighted mean concentration and load for a storm by collecting a single sample at or near the peak of the storm hydrograph. By knowing information such as the time that the sample was collected, the timing of the peak of the storm hydrograph, and the constituent concentration of the sample, the flow–weighted mean concentration can be easily and accurately calculated. For the constituents examined in this process, the method worked reasonably well, especially when the sample was taken at the peak of the hydrograph. However, a conclusive reason was not determined as to why the method performed better for certain constituents than others. Important to note was the success of the method for water quality data from two streams that experienced different land uses within their respective watersheds. This point emphasizes the applicability of the method to different watersheds, assuming that the appropriate hydrograph data are readily available.

One of the more attractive features of this method is that it allows for the computation of the FWMC at a cost savings to the user while maintaining accuracy. The typical cost for analyzing one sample for the nine constituents examined in this method was $206.25, according to McCoy and McCoy Laboratories, Inc. (personal communication, 15 June 2001) in Lexington, Kentucky. The number of samples analyzed in the data sets used to develop the method was 244, while 119 samples were analyzed in the verification data, resulting in a total laboratory cost of $74,868.75. By using this method to compute the FWMC for each of the 25 storms of the development data set and 20 storms of the verification data set, the total cost to the user would have been $9,281.25, resulting in a savings of $65,587.50. For users with a limited budget, the appeal of this method is apparent. Caution should be exercised when using this method, as it was developed for single–peak hydrographs. Additionally, the accuracy of the method is dependent on the constituent.

ACKNOWLEGEMENTS

The authors would like to express thanks to Mr. John D. Agouridis, Developer, at Diamond Manufacturing Commerce Systems, Inc., for his assistance in developing the Visual Basic program used in the project.

REFERENCES

Barger, G. L., and H. C. S. Thom. 1949. Evaluation of drought hazard. *Agronomy. J*. 41(11): 519–526.

Barger, G. L., R. H. Shaw, and R. F. Dale. 1959. Gamma distribution parameters from 2– and 3–week precipitation totals in the north central region of the United States. Ames, Iowa: Iowa State University, Agricultural and Home Economics Experiment Station.

Cooke, S. E., S. M. Ahmed, and N. D. MacApline. 2000. Introductory guide to surface water quality monitoring in agriculture. Edmonton, Alberta, Canada: Alberta Agriculture, Food, and Rural Development, Conservation and Development Branch.

Edwards, D. R., T. C. Daniel, J. F. Murdoch, P. F. Vendrell, and D. J. Nichols. 1994. The Moores Creek monitoring project: Final report submitted to the Arkansas Soil and Water Conservation Commission. Fayetteville, Ark.: Arkansas Water Resources Center.

Edwards, D. R., T. C. Daniel, H. D. Scott, J. F. Murdoch, M. J. Habiger, and H. M. Burks. 1996. Stream quality impacts of best management practices in a northwestern Arkansas basin. *Water Res. Bull*. 32(3): 499–509.

Friendman, D. G., and B. F. Janes. 1957. Estimation of Rainfall Probabilities. Connecticut Agricultural Experiment Station Bulletin No. 332. Storrs, Conn.: University of Connecticut.

Greenberg, A. E., L. S. Clesceri, and A. D. Eaton, eds. 1992. *Standard Methods for the Examination of Water and Wastewater*. 18th ed. Washington, D.C.: American Public Health Association.

Guillard, K., T. F. Morris, and K. L. Kopp. 1999. The pre–sidedress soil nitrate test and nitrate leaching from corn. *J. Environ. Qual*. 28(6): 1845–1852.

Haan, C. T. 1977. *Statistical Methods in Hydrology*. Ames, Iowa: Iowa State University Press.

Hoos, A. B., J. A. Robinson, R. A. Aycock, R. R. Knight, and M. D. Woodside. 2000. Sources, instream transport, and trends of nitrogen, phosphorus, and sediment in the lower Tennessee River basin, 1980–96. USGS Report No. 99–4139. Reston, Va.: U.S. Geological Survey.

Huber, D. R. 1993. Contaminant transport in surface water. Chapter 14 in *Handbook of Hydrology*. D. R. Maidment, ed. New York, N.Y.: McGraw–Hill.

Kuykendall, H. A., M. L. Cabrera, C. S. Hoveland, M. A. McCann, and L. T. West. 1999. Stocking method effects on nutrient runoff from pastures fertilized with broiler litter. *J. Environ. Qual*. 28(6): 1886–1890.

Markovic, R. D. 1965. Probability functions of best fit to distributions of annual precipitation and runoff. Hydrology Paper No. 8. Fort Collins, Colo.: Colorado State University.

McFarland, A. M. S., and L. M. Hauck. 1999. Relating agricultural land uses to in–stream stormwater quality. *J. Environ. Qual*. 28(3): 836–844.

Mooley, D. A., and H. L. Crutcher. 1968. An application of the gamma distribution function to Indian rainfall. Environmental Sciences Services Administration Technical Report EDS–5. Silver Springs, Md.: U.S. Department of Commerce.

Moorman, T. B., D. B. Jaynes, C. A. Cambardella, J. L. Hatfield, R. L. Pfieffer, and A. J. Morrow. 1999. Water quality in Walnut Creek watershed: Herbicides in soils, subsurface drainage, and groundwater. *J. Environ. Qual*. 28(1): 35–45.

Mueller, D. K., and J. D. Stoner. 1998. Identifying the potential for nitrate contamination of streams in agricultural areas of the United States. Reston, Va.: U.S. Geological Survey, National Water–Quality Monitoring Council.

Rickert, D. A. 1985. Water quality — Compensation for discharge in detecting trends in water quality data. Quality of Water Branch Memorandum No. 85.17. Reston, Va.: U.S. Geological Survey.

SAS. 1996. Statistical Analysis Software. Cary, N.C.: SAS Institute, Inc.

Saxton, A. M. 1999. Design and analysis for biological research with SAS® Software. Knoxville, Tenn.: University of Tennessee, Department of Animal Science.

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(2): 281–290.

Turtola, E., and M. Yli–Halla. 1999. Fate of phosphorus applied in slurry and mineral fertilizer: Accumulation in soil and release into surface runoff water. *Nutr. Cycl. Agroecosyst.* 55(2): 165–174.

U.S. EPA. 1973. Methods for identifying and evaluating the nature and extent of nonpoint sources of pollutants. Paper No. 430/9–73–014. Washington, D.C.: U.S. Environmental Protection Agency, Office of Air and Water Programs.

USGS. 1969. Discharge measurements at gaging stations. Chapter A8 in *Techniques of Water–Resources Investigations*, Book 3. Washington, D.C.: U.S. Government Printing Office.

USGS. 1984. Discharge ratings at gaging stations. Chapter A10 in *Techniques of Water–Resources Investigations*, Book 3. Washington, D.C.: U.S. Government Printing Office.

Wood, B. H., C. W. Wood, K. H. Yoo, K. S. Yoon, and D. P. Delaney. 1999. Seasonal surface runoff losses of nutrients and metals from soils fertilized with broiler litter and commercial fertilizer. *J. Environ. Qual*. 28(4): 1210–1218.