# IMPROVING STREAM SOLUTE LOAD ESTIMATION BY THE COMPOSITE METHOD: A COMPARATIVE ANALYSIS USING DATA FROM THE PANOLA MOUNTAIN RESEARCH WATERSHED

Brent T. Aulenbach[/1] and Richard P. Hooper[/2]

AUTHORS: [1/]Hydrologist, U.S. Geological Survey, 3039 Amwiler Road, Suite 130, Peachtree Business Center, Atlanta, Georgia 30360-2824, btaulenb@usgs.gov; and [2/]Hydrologist, CUAHSI, 2000 Florida Avenue, NW, Washington, DC, 20009, rhooper@cuahsi.org.
REFERENCE: *Proceedings of the 2005 Georgia Water Resources Conference*, held April 25–27, 2005, at the University of Georgia. Kathryn J. Hatcher, editor, Institute Ecology, The University of Georgia, Athens, Georgia.

**Abstract.** The setting and attainment of total maximum daily load (TMDL) has become an important criterion for protecting the water quality of surface waters. Accurately estimating streamwater solute loads during short-time intervals with limited resources is crucial to the implementation of TMDLs. The two most common techniques for estimating streamwater loads have shortcomings. The period-weighted approach requires comprehensive sampling to obtain unbiased load estimates, whereas the regression-model method does not accurately estimate loads during short-time intervals and typically does not model unusual events such as combined sewer overflows (a critical input for the TMDL approach). An improved load estimation technique known as the composite method, which combines aspects of both the period-weighted approach and the regression model method, can more accurately estimate loads during short-time intervals and with fewer sampling requirements. This paper demonstrates the usefulness of the composite method for estimating streamwater loads using data from the Panola Mountain Research Watershed near Atlanta, Georgia.

## INTRODUCTION

Streamwater load, often referred to as mass flux, is the mass of chemical solutes or sediment transported across a stream cross-section during a specific time period. In watershed studies, mass flux serves as an integrated measure of all processes within a watershed that affect water quality (Semkin and others, 1994). With increased emphasis on watershed-based strategies for the control of nonpoint-source pollutants, reliable measures of mass flux are needed. In the United States, stream reaches that do not meet water-quality standards are subject to waste-load allocation schemes based on the total maximum daily load (TMDL). A TMDL is defined as the maximum amount of a pollutant that a water body can receive and still meet water-quality standards (U.S. Environmental Protection Agency, 2004).

Loads typically are estimated using either the period-weighted approach or the regression-model (or rating-curve) method. In the period-weighted approach, measured solute concentrations in streamwater samples are assumed to approximate the streamwater concentration around the period the sample is collected (e.g., Likens and others, 1977; Larson and others, 1995). The period-weighted approach requires extensive sampling to avoid biases in loads. The regression-model method involves developing a regression model for concentration as a function of other continuous variables and then determining loads through time by multiplying the nearly continuous streamflow measurements by the model concentration (e.g., Johnson, 1979; Cohn and others, 1992). The regression-model method typically requires fewer data than the period-weighted approach.

The composite method is a new approach to estimating loads that combines the period-weighted approach and the regression-model method. In the composite method, a regression model is used to predict concentration variations between samples due to changing hydrologic conditions and season. A period-weighted approach is then used to adjust the predicted regression model concentrations to the actual sample concentrations when a sample is collected, and applies the residuals (the difference between the model predicted and observed concentrations) to the concentration model in a piecewise linear manner to periods between sample collections (Aulenbach and Hooper, 2001).

## METHODS

The data set analyzed in this study consists of streamwater chemistry (from weekly grab sampling augmented by more frequent automatic sampling during rainstorms) and flow measured at the outlet of the Panola Mountain Research Watershed (PMRW), Georgia. PMRW is a 41-hectare forested experimental basin, which is 25 kilometers southeast of Atlanta in the southern Piedmont physiographic province (Huntington and others, 1993). The data set spans a 13-year period from water year (WY) 86 through WY98 and includes 2,790 samples analyzed for major anions (sulfate, chloride, and bicarbonate [as alkalinity]), cations (calcium, magnesium, and sodium), and dissolved silica using standard methods. The PMRW data set was chosen because the analysis required an extensively sampled stream to accurately estimate loads and assess sampling strategies with various subsampling scenarios.

Loads were calculated using the period-weighted approach, the regression-model method, and the composite method. For the period-weighted approach, measured concentrations were linearly interpolated through time between samples (Larson and others, 1995). Regression models for each solute were developed for the regression-model method and the composite method in which solute concentrations were estimated from a hyperbolic function of streamflow, seasonal sinusoidal terms, and a dummy variable that indicates whether a sample was collected during the rising limb of a storm hydrograph.

A bootstrap experiment was designed to determine and compare the accuracy and precision of the load estimates for the composite method and the period-weighted approach using a variety of different sampling designs. Results of this bootstrap experiment are presented only for alkalinity. Two water years, WY97 and WY98, were chosen for the bootstrap experiment because these years had the most complete large-storm sample coverage with 78 percent (%) (WY97) and 77% (WY98) of the large storms sampled (large storms are defined herein as storms with peak flows greater than 10 liters per second [0.21 centimeters per day]). The percentage of storms used in the models was varied such that 0, 10, 20, 40, 60, and 77–78% of the large storms were included on an annual basis. Load estimates were simulated 100 times each for the 10, 20, 40, and 60% large-storm test cases using a different set of randomly selected storms for each simulation. The various storm-sampling scenarios were simulated in combination with two different routine fixed-interval sampling designs: weekly and monthly.

## RESULTS

Average load estimates from the regression-model method for the 13-year study period were similar to the load estimates from the composite method for most solutes (Fig. 1). Four of the seven solutes differed by 0.6% or less, whereas alkalinity (–2.9%) and sodium (1.4%) had larger differences, and sulfate (16%) had the largest difference (Table 1). The regression-model method load estimates differed more from the composite method load estimates for shorter time periods, with the differences increasing from annual to monthly time scales. The range in the standard deviation of the percent differences between the regression-model and composite methods for the seven solutes is from 4.1 to 23% on an annual basis, and from 7.0 to 41% on a monthly basis (Table 1).

**Table 1. Summary of comparison of regression-model method and composite method load esimated on a period of study (from water year 86 to water year 98), annual, quarterly, and monthly basis.**

[Regression models were fit to entire 13-year period; %, percent]

| Solute | Difference between regression and composite loads, period (%) | Standard deviation of annual percent differences | Standard deviation of quarterly percent differences | Standard deviation of monthly percent differences |
|---|---|---|---|---|
| Alkalinity | –2.9 | 8.5 | 11.0 | 12.0 |
| Calcium | 0.6 | 4.2 | 5.2 | 7 |
| Magnesium | –0.3 | 4.1 | 6 | 7.6 |
| Sodium | 1.4 | 4.8 | 6.7 | 7.5 |
| Sulfate | 16 | 23 | 35 | 41 |
| Chloride | 0.2 | 4.8 | 6.4 | 7.8 |
| Dissolved silicia | –0.4 | 6.6 | 8.2 | 8.9 |

Accuracies in load estimates from the bootstrap experiment for alkalinity were determined by comparing them to the best-test case load estimates—i.e., from the composite method with all samples included. Errors in load estimates are summarized on an annual and monthly basis in Table 2 and Figure 2. Loads estimated using the period-weighted approach considerably overestimate loads when no or few storms were included. In test cases for which 20% or fewer storms were included, errors in loads on an annual basis range from 13 to 24% (Table 2). This bias is the result of the combination of two factors. First, there is an inverse relation between concentration and discharge (dilution) for alkalinity. Second, with the inclusion of few storms, the estimates of loads during most storms were determined using concentrations from the fixed-interval sampling representing samples collected primarily during baseflow conditions. As a consequence, higher baseflow concentrations were incorporated into the load estimates during storm periods, resulting in overestimation of overall loads.

Loads estimated using the period-weighted approach in test cases for which all storm samples were included were somewhat underestimated for weekly fixed-interval sampling test case, –1.6% during WY97 and –0.8% during WY98, and were more understated for monthly fixed-interval sampling test case, –6.8% during WY97 and –6.5% during WY98 (Table 2). These biases are likely an artifact of storm-sampling design. The first storm sample was not collected until after the initial rise in streamflow, and the last storm sample was collected before flow completely returned to baseflow. Consequently, storm samples collected at the beginning of an event and on recession prior to receding to baseflow have lower concentrations than during adjacent baseflow periods, resulting in underestimated loads during these baseflow periods. This effect is more pronounced for monthly fixed-interval sampling test cases because storm concentrations were applied to longer periods of baseflow.

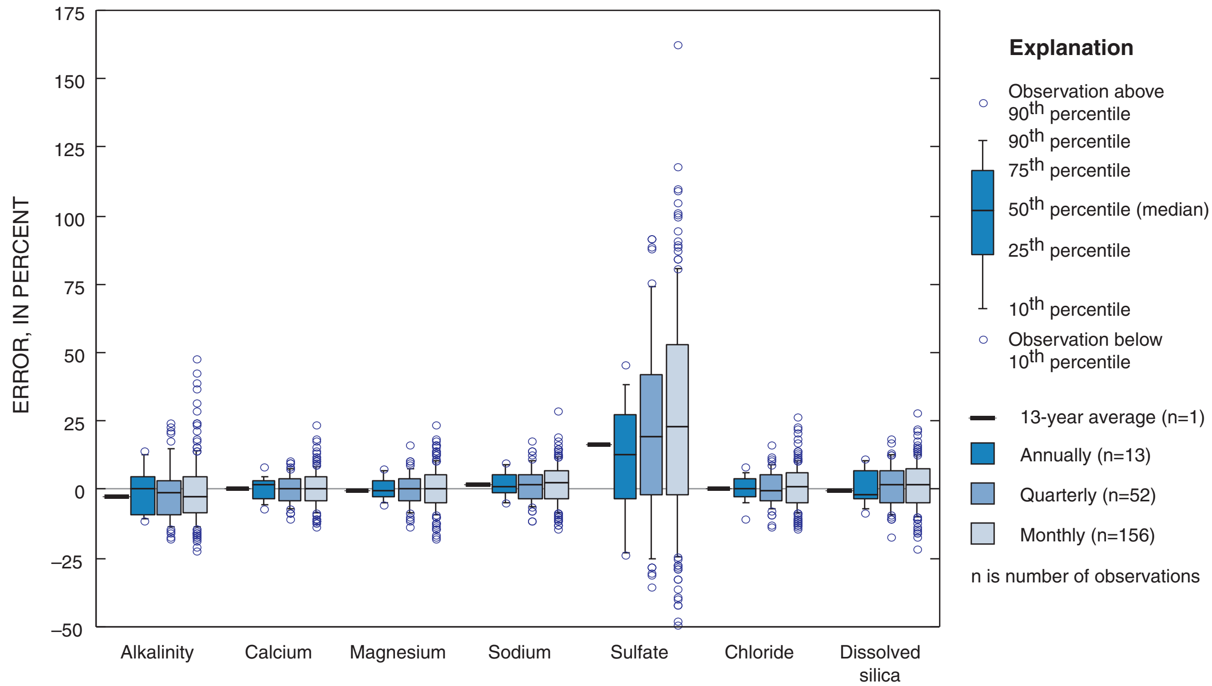
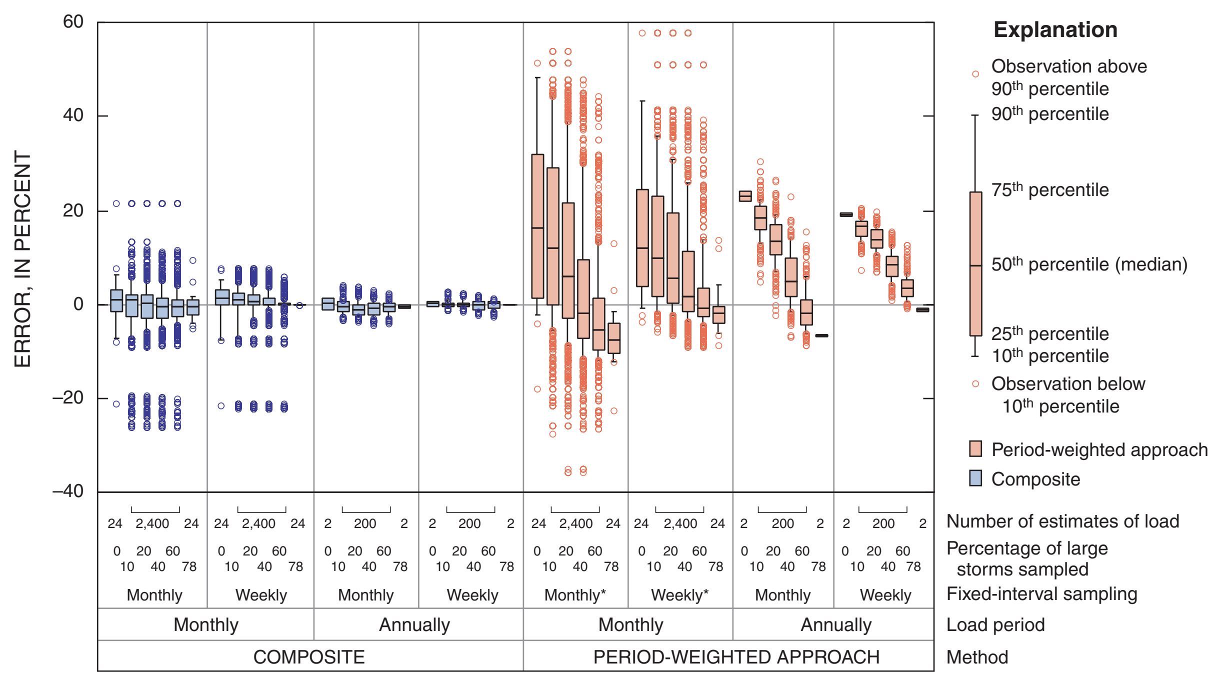



**Figure 1. Boxplot summary of percent errors between regression-model method and best composite method load estimates for estimation periods of annual, quarterly, and monthly for the period of study (from water year 86 to water year 98).**




*Test cases have model load estimates with greater than 60 percent error.
Not shown in order to expand scale to show detail.

**Figure 2. Boxplot summary of the bootstrap experiment for alkalinity for water year (WY) 97–98 on a monthly and annual basis using composite method and period-weighted approach, weekly and monthly fixed-interval sampling, and various percentages of large storms sampled. Error in load estimates is the percent difference between the test case and the best test-case load estimate. Bootstrap experiment run 100 times for each WY, randomly selecting large-storms for each test case. There is only one estimate of load per year for test cases with no (0%) or all sampled (77–78%) large storms selected, because there is only one possible selection of storms for these test cases.**

**Table 2. Results of the bootstrap experiment for alkalinity using the composite method and period-weighted approach, weekly and monthly fixed-interval sampling, and different percentages of large storms sampled.** [NA, not applicable; WY, water year; %, percent]

| Period | % Storms | Composite Weekly | | Composite Monthly | | Period-Weighted Weekly | | Period-Weighted Monthly | |
|---|---|---|---|---|---|---|---|---|---|
| | | Error (%) | Variance (%) | Error (%) | Variance (%) | Error (%) | Variance (%) | Error (%) | Variance (%) |
| Annual WY97 | 78 | NA | NA | –0.7 | NA | –1.6 | NA | –6.8 | NA |
| | 60 | –0.1 | 0.6 | –0.9 | 1.1 | 3.2 | 2.1 | –1.7 | 3.3 |
| | 40 | –0.2 | 0.8 | –1.1 | 1.5 | 7.9 | 2.4 | 3.9 | 3.3 |
| | 20 | –0.4 | 0.7 | –1.5 | 1.3 | 13 | 1.7 | 13 | 2.7 |
| | 10 | –0.5 | 0.6 | –1.5 | 1.2 | 16 | 1.5 | 17 | 2.1 |
| | 0 | –0.6 | NA | –1.3 | NA | 19 | NA | 22 | NA |
| Annual WY98 | 77 | NA | NA | 0.0 | NA | –0.8 | NA | –6.5 | NA |
| | 60 | –0.2 | 1.1 | –0.6 | 1.6 | 4 | 2.9 | –0.4 | 5.8 |
| | 40 | –0.2 | 1.2 | –0.8 | 1.9 | 8.5 | 3.3 | 7.4 | 6.4 |
| | 20 | 0.3 | 1 | –0.4 | 1.7 | 14 | 2.7 | 14 | 5.6 |
| | 10 | 0.5 | 0.9 | 0.4 | 1.2 | 17 | 2.3 | 19 | 4.1 |
| | 0 | 0.7 | NA | 1.4 | NA | 20 | NA | 24 | NA |
| Monthly WY97–98 | All | NA | NA | 2.1 | 3.1 | 3.8 | 5.1 | 8.3 | 6.5 |
| | 60 | 0.9 | 1.7 | 2.2 | 2.7 | 5.2 | 8.4 | 6.5 | 9.1 |
| | 40 | 1.7 | 3.1 | 2.7 | 3.9 | 8.8 | 13 | 11 | 20 |
| | 20 | 2.7 | 4.4 | 3.2 | 5.3 | 14 | 20 | 16 | 26 |
| | 10 | 3.1 | 5 | 3.6 | 6.3 | 16 | 22 | 21 | 33 |
| | 0 | 3.6 | 5.9 | 4.5 | 7.3 | 19 | 25 | 26 | 38 |

**Note:** Error for annual loads is the percent difference between the test case and the best test-case load estimates (composite method using all samples). Error for monthly loads is the average of monthly absolute percent differences between the test case and the best test-case load estimate. In test cases in which 100 runs of random storms are sampled, the average load estimate of the test case is used for comparison. For the annual loads, variance is the coefficient of variation of the load estimates for test cases that have 100 random runs. For monthly loads, variance is the standard deviation of the monthly percent differences (not the absolute value of these differences).

Average monthly absolute percent differences for loads estimated using the period-weighted approach ranged from 3.8 to 26% (Table 2) and are similar to the range observed for the magnitude of the average annual differences. The number of storms included in the test case was more important than the frequency of fixed-interval sampling, with monthly load estimates being more accurate for test cases with the greater number of storms included and weekly fixed-interval sampling.

The composite method had much smaller errors in loads than the period-weighted approach. Errors for the composite method test cases on an annual basis range from –1.5 to 1.4% (Table 2). Increasing the percentage of storms sampled from 0% to all storms sampled (77–78%) was about equally important as increasing the frequency of fixed-interval sampling from monthly to weekly in improving the accuracy of the load estimates. Average monthly absolute percent differences for the composite method test cases range from 0.9 to 4.5% and are consistently larger than the magnitude of the annual error for the same test case. Increasing the number of storms included in the test case from 0% to all storms sampled improved accuracy of the monthly load estimates somewhat more than increasing the frequency of the fixed-interval sampling from monthly to weekly.

The precision (variance) of the annual load estimates associated with the random selection of storms for each test case was determined by calculating the coefficient of variation (the standard deviation divided by the mean, expressed as a percentage) of the 100 load estimates. There are no calculations of precision for test cases in which either all or none of the storms were sampled, because there is only one possible set of samples and, therefore, only one estimate of load. Precision of the composite method loads was better than that of the period-weighted approach loads, with the coefficient of variation ranging from 0.6 to 1.9% for the composite method and from 1.5 to 6.4% for the period-weighted approach (Table 2). The precision across different storm-sampling test case percentages was relatively constant, but varied with water year and fixed-interval sampling frequency.

The precision of monthly load estimates was assessed from the standard deviation of the 24 monthly percent differences (actual differences were used, not the magnitude of the differences) during the 2-year bootstrap experiment. For test cases with 100 runs, the average monthly load of the 100 runs was used to calculate the monthly difference. The composite method is more precise than the period-weighted approach, with the standard deviation in the monthly percent differences ranging from 1.7 to 7.3% for the composite method and 5.1 to 38% for the period-weighted approach (Table 2). Lower variability was observed for test cases with higher frequency fixed-interval sampling and more storm sampling.

## DISCUSSION

When considering the acceptable level of error associated with estimating streamwater loads, the errors associated with the accuracy of flow measurements, errors associated with the representativeness of the water samples collected, and the accuracy of the chemical analyses also need to be considered. Each of these errors typically can be on the order of from 5 to 10%, even when care is taken in the field and laboratory.

When comparing the load differences between the composite and the regression model method, the composite method is assumed to produce the best load estimate because it improves the regression-model results by adjusting the model concentrations toward the observed sample concentrations through time. For long time periods, the regression-model method generally is adequate for estimating loads, such as shown herein for the 13-year study period, with the exception of sulfate. Inaccuracies for sulfate are likely the result of low predictability of the sulfate regression model—relations between concentration and the independent variables (flow, season, and hydrologic condition) were weakest for sulfate among the seven solutes modeled. The regression-model method has shortcomings when estimating loads during shorter time periods (e.g., annually to monthly) because the model predicts the average concentration response for the entire period, not the response for the specific shorter timeframes. If residual concentrations varied randomly over time, smaller errors in the shorter-period regression-model method load estimates would be expected; however, residuals were observed to vary systematically through time.

The period-weighted approach requires extensive sampling to avoid bias. If sampling is based on a fixed-time interval, the majority of the samples collected will represent baseflow, the predominant flow condition. However, a significant portion of the load may occur during storms for which sampling is underrepresented with respect to the proportion of total runoff. This storm-sampling underrepresentation, in conjunction with flow-related variations in streamwater concentrations during storms, results in biases in load estimates. Biases in load estimates can be reduced by adequately sampling storms, along with appropriately designed sampling, to capture the relatively abrupt concentration changes that occur at the beginning of storms and during the return to baseflow conditions near the end-of-storm recessions. The period-weighted approach load estimates are relatively accurate on an annual basis when all storm and fixed-interval samples are used, with errors of –1.6% and –0.8% for WY97 and WY98, respectively. Load estimates were less accurate and were imprecise on a monthly basis.

The composite method requires many fewer samples than the period-weighted approach to achieve the same accuracy and precision. The composite method is less sensitive to sampling frequency than the period-weighted approach because it relies on the regression model to predict solute concentration responses to changing hydrologic conditions between collected samples. Whereas the regression model method develops a model of average chemical response and then ignores the remaining information contained in the residual concentrations, the composite method retains the residual concentrations to adjust the regression-model-predicted concentrations to the observed sample concentrations and applies the residual error between samples using a period-weighted approach. The composite method thereby can incorporate short-term deviations from the regression model and can include unusual events, such as combined sewer overflows, as long as the event is adequately sampled. Therefore, the composite method should be the method of choice for studies requiring streamwater load estimates, providing more accurate and precise loads, especially during short timeframes, and often with fewer sampling requirements.

## LITERATURE CITED


Aulenbach, B.A., and R.P. Hooper. 2001. Removing climatic effects from trends in streamwater load estimates. In *AWRA Annual Spring Specialty Conference Proceedings*. “Water quality monitoring and modeling,” April 30–May 2, 2001, J.J. Warwick (ed.). American Water Resources Association, Middleburg, Va., pp. 47–52.

Cohn, T.A., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers. 1992. The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay. *Water Resources Research*, v. 28, no. 9, pp. 2353–2363.

Huntington, T.G., R.P. Hooper, N.E. Peters, T.D. Bullen, and Carol Kendall. 1993. Water, energy, and biogeochemical budgets investigation at Panola Mountain Research Watershed, Stockbridge, Georgia–A research plan. U.S. Geological Survey Open-File Report 93-55, 39 pp.

Johnson, A.H. 1979. Estimating solute transport in streams from grab samples. *Water Resources Research*, v. 15, no. 5, pp. 1224–1228.

Larson, S.J., P.D. Capel, D.A. Goolsby, S.D. Zaugg, and M.W. Sandstrom. 1995. Relations between pesticide use and riverine flux in the Mississippi River basin. *Chemosphere*, v. 31, no. 5, pp. 3305–3321.

Likens, G.E., F.H. Bormann, R.S. Pierce, J.S. Eaton, and N.M. Johnson. 1977. Biogeochemistry of a forested ecosystem, Springer–Verlag, New York, 146 pp.

Semkin, R.G., D.S. Jefferies, and T.A. Clair. 1994. Hydrochemical methods and relationships for study of stream output from small catchments. In *Biogeochemistry of small catchments: A tool for environmental research*. Bedřich Moldan and Jiří Černý (eds.), John Wiley and Sons, New York, pp. 163–187.

U.S. Environmental Protection Agency. 2004. Introduction to TMDLs. Accessed December 15, 2004, at *http://www.epa.gov/owow/tmdl/intro.html*