

University of New Hampshire

University of New Hampshire Scholars' Repository

New Hampshire EPSCoR

Research Institutes, Centers and Programs

6-9-2014

Sources of uncertainty in estimating stream solute export from headwater catchments at three sites

Ruth D. Yanai

SUNY College of Environmental Science and Forestry

Naoko Tokuchi

Kyoto University

John L. Campbell

USDA Forest Service

Mark B. Green

Plymouth State University

Eiji Matsuzaki

University of Northern British Columbia

See next page for additional authors https://scholars.unh.edu/nh_epscor

Comments

This is an Open Access article published by Wiley in Hydrological Processes in 2014, available online:

<https://dx.doi.org/10.1002/hyp.10265>

Recommended Citation

Yanai, R. D., Tokuchi, N., Campbell, J. L., Green, M. B., Matsuzaki, E., Laseter, S. N., Brown, C. L., Bailey, A. S., Lyons, P., Levine, C. R., Buso, D. C., Likens, G. E., Knoepp, J. D. and Fukushima, K. (2015), Sources of uncertainty in estimating stream solute export from headwater catchments at three sites. *Hydrological Processes*, 29: 1793– 1805. doi: 10.1002/hyp.10265.

This Article is brought to you for free and open access by the Research Institutes, Centers and Programs at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in New Hampshire EPSCoR by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.

Authors

Ruth D. Yanai, Naoko Tokuchi, John L. Campbell, Mark B. Green, Eiji Matsuzaki, Stephanie N. Laseter, Cindi L. Brown, Amey S. Bailey, Pilar Lyons, Carrie R. Levine, Donald C. Buso, Gene E. Likens, Jennifer D. Knoepp, and Keitaro Fukushima

Sources of uncertainty in estimating stream solute export from headwater catchments at three sites

Ruth D. Yanai,^{1*} Naoko Tokuchi,² John L. Campbell,³ Mark B. Green,⁴ Eiji Matsuzaki,⁵ Stephanie N. Laseter,⁶ Cindi L. Brown,⁶ Amey S. Bailey,⁷ Pilar Lyons,^{1,‡} Carrie R. Levine,^{1,†} Donald C. Buso,⁸ Gene E. Likens,⁹ Jennifer D. Knoepp⁶ and Keitaro Fukushima¹⁰

¹ Department of Forest and Natural Resources Management, SUNY College of Environmental Science and Forestry, Syracuse, NY, USA

² Division of Forest Biosphere, Kyoto University, Kyoto, Japan

³ Northern Research Station, USDA Forest Service, Durham, NH, USA

⁴ Center for the Environment, Plymouth State University, Plymouth, NH, USA

⁵ Natural Resources and Environmental Studies Graduate Program, University of Northern British Columbia, Prince George, British Columbia, Canada

⁶ Coweeta Hydrologic Laboratory, USDA Forest Service, Southern Research Station, Otto, NC, USA

⁷ Hubbard Brook Experimental Forest, USDA Forest Service, Northern Research Station, North Woodstock, NH, USA

⁸ Hubbard Brook Experimental Forest, Cary Institute of Ecosystem Studies, North Woodstock, NH, USA

⁹ Cary Institute, Millbrook, NY, USA

¹⁰ Graduate School of Agriculture, Kyoto University, Kyoto, Japan

Abstract:

Uncertainty in the estimation of hydrologic export of solutes has never been fully evaluated at the scale of a small-watershed ecosystem. We used data from the Gomadansan Experimental Forest, Japan, Hubbard Brook Experimental Forest, USA, and Coweeta Hydrologic Laboratory, USA, to evaluate many sources of uncertainty, including the precision and accuracy of measurements, selection of models, and spatial and temporal variation. Uncertainty in the analysis of stream chemistry samples was generally small but could be large in relative terms for solutes near detection limits, as is common for ammonium and phosphate in forested catchments. Instantaneous flow deviated from the theoretical curve relating height to discharge by up to 10% at Hubbard Brook, but the resulting corrections to the theoretical curve generally amounted to <0.5% of annual flows. Calibrations were limited to low flows; uncertainties at high flows were not evaluated because of the difficulties in performing calibrations during events. However, high flows likely contribute more uncertainty to annual flows because of the greater volume of water that is exported during these events. Uncertainty in catchment area was as much as 5%, based on a comparison of digital elevation maps with ground surveys. Three different interpolation methods are used at the three sites to combine periodic chemistry samples with streamflow to calculate fluxes. The three methods differed by <5% in annual export calculations for calcium, but up to 12% for nitrate exports, when applied to a stream at Hubbard Brook for 1997–2008; nitrate has higher weekly variation at this site. Natural variation was larger than most other sources of uncertainty. Specifically, coefficients of variation across streams or across years, within site, for runoff and weighted annual concentrations of calcium, magnesium, potassium, sodium, sulphate, chloride, and silicate ranged from 5 to 50% and were even higher for nitrate. Uncertainty analysis can be used to guide efforts to improve confidence in estimated stream fluxes and also to optimize design of monitoring programmes. © 2014 The Authors. *Hydrological Processes* published John Wiley & Sons, Ltd.

KEY WORDS Gomadansan; Hubbard Brook; Coweeta Experimental Forest; nutrient flux; model selection; gap filling; water chemistry; measurement error

Received 1 July 2013; Accepted 3 June 2014

INTRODUCTION

The accurate estimation of hydrologic solute export is essential to understanding nutrient budgets in forested

ecosystems. The total uncertainty in estimated hydrologic fluxes of nutrients is difficult to evaluate fully at watershed scales (Harmel *et al.* 2006). There have been many analyses comparing methods for calculating stream solute export (Bukaveckas *et al.*, 1998; LaBaugh *et al.*, 2009; Birgand *et al.*, 2010; Ullrich and Volk, 2010; Verma *et al.*, 2012). However, sources of uncertainty also include the precision and accuracy of measurements and natural variation in space and time (Rode and Suhr, 2007; Harmel *et al.*, 2009).

Uncertainty in streamwater nutrient export involves both the chemical analysis of solutes and the measure-

*Correspondence to: Ruth D. Yanai, Department of Forest and Natural Resources Management, SUNY College of Environmental Science and Forestry, Syracuse, NY, USA.
E-mail: rdyantai@syr.edu

†Current Address: Department of Environmental Science, Policy, and Management, University of California-Berkeley, Berkeley, CA, USA

‡Current Address: CDH Energy Corp., Cazenovia, NY, USA

ment of water fluxes. While uncertainty in chemical analysis is commonly reported and generally small, measuring stream discharge is more complicated. Discharge is usually estimated from stage–discharge relationships in weirs or in natural channels, relying on water level recorders (analogue or digital) for stage estimates at fine time scales. The stage–discharge relationship can be validated using volumetric measurements of discharge at low flow (Hornbeck, 1965), but validation data sets at very high flows are problematic (Di Baldassarre and Montanari, 2009; McMillan *et al.*, 2010). For low flows, at the Hubbard Brook Experimental Forest and Coweeta Hydrological Laboratory, theoretical rating curves have been adjusted based on validation measurements, and while the magnitudes of the corrections are probably small, they have never been reported.

Another source of measurement uncertainty is missing values in the discharge record, which can be filled using simulation models or statistical relationships with other streams (Sauer, 2002). The uncertainty introduced by these gap-filling techniques can be estimated by filling artificial gaps and comparing the predicted to the actual discharge.

The area of the catchment at the point of the discharge measurement is used to calculate streamflow on an areal basis. The delineation of the catchment is thus another source of measurement uncertainty in estimating both streamflow and solute export. Methods of estimating catchment area include ground surveys and topographic maps, including digital elevation modelling at various resolutions. Such multiple approaches can be compared to indicate uncertainty in delineating the topographic boundaries of a catchment. Water movement across topographic divides in deep groundwater is another source of uncertainty in dividing stream exports by the contributing area (Winter *et al.*, 2003).

While discharge is measured continuously with chart recorders or at scales of minutes or hours using data loggers, solute concentrations are collected much less frequently in long-term studies, usually at intervals of a week to a month. The calculation of solute export as the product of concentration and discharge is thus subject to uncertainty in the temporal pattern of concentration between sampling dates (Johnson *et al.*, 1969; Harmel *et al.*, 2009). The simplest models use the measured concentration and cumulative discharge for the interval or a linear interpolation between sampling dates. More complicated models use other factors (discharge, for example) to predict concentration between sampling times (Aulenbach and Hooper, 2006). Uncertainty in these techniques for temporal interpolation is another source of uncertainty in the resulting export values.

We selected three sites that monitor streams from multiple headwater catchments, namely Gomadansan

Experimental Forest, Japan, Hubbard Brook Experimental Forest, USA, and Coweeta Hydrologic Laboratory, USA, for a comprehensive survey of sources of uncertainty. At these sites, we could compare streams to evaluate spatial variation in stream export of water and solutes across the landscape. We evaluated interannual variation and compared this with spatial variation across the three sites for 2000–2009, except for Gomadansan, for which data are available beginning in 2003. We estimated the uncertainties due to measurements of concentration, discharge, and catchment area and the uncertainties in the choice of models describing flux as a function of concentration and discharge. We tested the hypothesis that uncertainties due to measurement error and model selection are small compared with natural variation in space and time, as has been implicitly assumed in studies that do not consider other sources of uncertainty.

Sites and monitoring methods

Gomadansan, Hubbard Brook, and Coweeta (Figure 1) differ in climate, vegetation, hydrology, and soils (Table I). The most important climatic differences are the high rainfall at Gomadansan and the long period of snow cover at Hubbard Brook. There are 5 gauged streams at Gomadansan, 9 at Hubbard Brook, and 15 at Coweeta (Table II).

All three sites have permanent stream gauges to measure continuous head or stage height (Table II). Coweeta has V-notch and trapezoidal (Cippoletti) weirs, Gomadansan has V-notch and rectangular weirs, and Hubbard Brook has V-notch weirs augmented in some cases by San Dimas flumes or compound weirs (Sauer

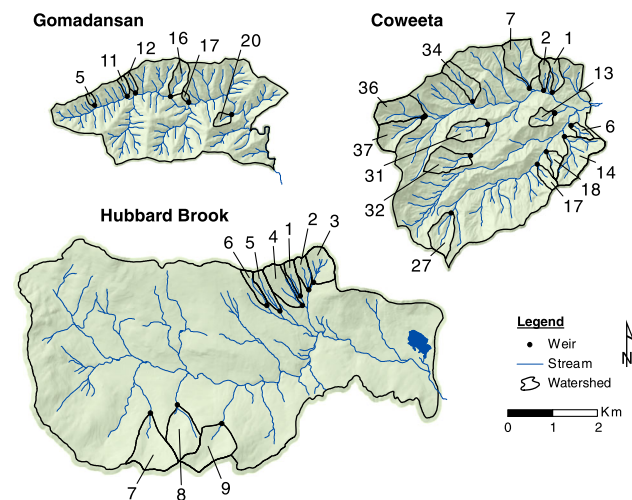


Figure 1. Maps of the Gomadansan Experimental Forest, Coweeta Hydrologic Laboratory, and Hubbard Brook Experimental Forest, showing the monitored catchments at each site

Table I. Sites included in the uncertainty analysis of solute export in headwater streams

	Location	Elevation range	Climate	Average annual precipitation	Vegetation	Soils
Coweeta	35.03N, 83.25W	675–1592 m	Marine humid temperate Mean annual average temperature: 13 °C 3 °C in January 22 °C in July	1800 to 2500 mm	Mixed hardwood and pine	Inceptisols (Dystrochrepts and Haplumbrepts) Ultisols (Hapludults and Hapludults)
Hubbard Brook	43.94N, 71.75W	222–1015 m	Continental Mean annual average temperature: 6 °C −9 °C in January 18 °C in July	1400 mm	Deciduous northern hardwoods and conifers	Spodosols (Haplorhods and Dystrochrepts)
Gomadansan	34.04N, 135.35E	860–1370 m	Mean annual average temperature: 12 °C Average annual 12.3 0.8 °C in January 23.5 °C in July	2520 mm	<i>Cryptomeria japonica</i> plantation	Inceptisols (Dystrochrepts)

and Turnipseed, 2010). To record stage height, Gomadansan uses a pressure transducer and data logger. Coweeta uses a shaft encoder (Design Analysis H-330) connected to a float and pulley, contained within a stilling well. As the float rises or falls, the stage height is encoded and recorded to a data logger (Design Analysis X500L). Hubbard Brook has a similar system but uses Leupold–Stevens A-35 strip chart recorders with a 7-jeweled Chelsea Marine clock for the V-notches and Belfort FW-1 recorders for the flumes. Hubbard Brook is transitioning to electronic data collection with Campbell Scientific CS410 shaft encoders and CR1000 data loggers.

We analysed data from 2000–2009, which was the most recent 10-year period for which data were available. Annual fluxes are reported for a water year beginning on 1 June for Gomadansan and Hubbard Brook and 1 November for Coweeta. These dates were chosen to minimize change in water storage (in snowpack and soils) from one year to the next (Likens, 2013).

At Hubbard Brook and Coweeta, samples are collected for chemical analysis on a regular weekly schedule, regardless of the weather, but occasionally adjusted for holidays to a 6- or 8-day week. At Gomadansan, samples are collected on a less regular schedule, at most biweekly but often at intervals of a month or more; rainy days are avoided because of dangerous conditions and high variability of solute chemistry.

Samples are analysed within 2 days of collection for Coweeta and within 2 months for Gomadansan. For Hubbard Brook, samples are normally analysed for nitrate (NO₃) and ammonium (NH₄) within 2 weeks of collection, and the remaining solutes are analysed over about 3 months (Buso *et al.*, 2000). Samples are analysed by methods described separately for Coweeta (Swank and Waide, 1988), Gomadansan (Fukushima and Tokuchi, 2009), and Hubbard Brook (Buso *et al.*, 2000).

Uncertainty in analysis of water chemistry

Uncertainty in the chemical analysis of streamwater depends on the accuracy, precision, and detection limits of the analytical methods. Accuracy, as reported here, is based on the difference between measured and certified concentrations of external quality control standards (Table III). Inaccuracies can be positive or negative; we report the average of the absolute values as an indication of the error expected for a single sample. For example, the inaccuracy in concentration of a single sample averaged 4% for Hubbard Brook and Coweeta and 12% for Gomadansan for the base cations calcium (Ca), magnesium (Mg), potassium (K), and sodium (Na; Table III). It is also important to know whether there is a bias, indicated by a non-zero average of the accuracies. For example, when the inaccuracies in base cations were

Table II. Descriptions of catchments within the three sites

Number and Name	Area (ha)	Year installed	Notch type	Elevation range (m)		Aspect	Description
<i>Coweeta</i>							
1 Copper Branch	16	1934	90° V	705	988	S	Burned and herbicided 1942–1957; white pine planted in 1957
2 Shope Branch	12	1934	90° V	709	1004	SSE	Mixed hardwoods undisturbed since 1927
6 Sawmill Branch	9	1934	90° V	696	905	NW	Clearcut in 1958, planted to grass, last herbicided 1967; successional vegetation
7 Big Hurricane	59	1934	90° V	722	1077	S	Clearcut and cable logged in 1977
8 Shope Fork No. 1	760	1934	12' Cipolletti	702	1600	Multiple	Fourth-order stream draining watersheds 1, 2, 7, 13, 32, 34,36, and 37
13 Carpenter Branch	16	1936	12' Cipolletti	725	912	ENE	Clearcut in 1939 and 1962; no products removed in either treatment
14 Hugh White Branch	61	1936	120° V	707	992	NW	Mixed hardwoods undisturbed since 1927
17 Hertzler Branch	13	1936	90° V	760	1021	NW	All woody vegetation cut annually from 1940 until 1955; no products removed. White pine planted in 1956
18 Grady Branch	13	1936	120° V	726	993	NW	Mixed hardwoods undisturbed since 1927
27 Hard Luck Creek	39	1946	120° V	1061	1454	NNE	Mixed hardwoods undisturbed since 1927
31 Mill Branch	34	1981	120° V	869	1146	ENE	Mixed hardwoods undisturbed since 1927
32 Cunningham Creek number 2	41	1941	120° V	920	1236	ESE	Mixed hardwoods undisturbed since 1927
34 Bee Branch	33	1938	120° V	866	1184	SE	Mixed hardwoods undisturbed since 1927
36 Pinnacle Branch	49	1943	120° V	1021	1542	ESE	Mixed hardwoods undisturbed since 1927
37 Albert Branch	44	1942	120° V	1033	1592	ENE	Clearcut in 1963; no products removed.
<i>Hubbard Brook</i>							
1	12	1956	90° V	488	747	SE	Wollastonite (CaSiO ₃) added in 1999
2	16	1957	120° V	503	716	SE	Clearcut in 1965, herbicided until 1968; no products removed
3	42	1957	120° V	527	732	SW	Northern hardwood forest undisturbed since 1917
4	36	1960	120° V	442	747	SE	Strip clearcut from 1970 to 1974
5	22	1962	90° V, 3' flume	488	762	SE	Whole tree harvested in 1983–1984
6	13	1963	90° V, 2' flume	549	792	SE	Northern hardwood forest undisturbed since 1917
7	77	1965	120° V, 4' flume	619	899	NW	Northern hardwood forest undisturbed since 1917
8	59	1968	120° V, 4' flume	610	905	NW	Northern hardwood forest undisturbed since 1917
9	68	1995	120° V	685	910	NE	Spruce-fir forest undisturbed since 1917. Wetlands in headwaters
<i>Gomadansan</i>							
5	4	1998	90° V	1070	1241	SE	Clearcut in 1978. Planted to Japanese cedar
11	6	1998	90° V	1001	1228	SE	Clearcut in 1981. Planted to Japanese cedar
12	7	1998	Rectangular	981	1231	SE	Clearcut in 1993. Planted to Japanese cedar
16	24	1998	Rectangular	959	1304	S	Clearcut in 2003. Planted to Japanese cedar
17	3	1998	90° V	954	1129	SW	Clearcut in 2007. Planted to Japanese cedar
20	9	1998	90° V	897	1196	NE	Clearcut in 1967. Planted to Japanese cedar

averaged over long time periods, the bias was only 1% for Hubbard Brook and 2% for Coweeta (Table III). Some reported uncertainties were high at Gomadansan because quality control samples were not routinely run.

Precision, at Hubbard Brook, is determined from repeated analysis of the same streamwater sample: one sample of every 40 is analysed four times. For Coweeta,

precision is determined quarterly from the variation in certified quality control standards obtained from an independent source. We report precision as the average standard deviation of replicate samples in units of concentration and as the coefficient of variation (CV) or standard deviation divided by the mean (Table III). It is more useful to report precision in units of concentration for dilute samples and in units relative to the mean at

Table III. Uncertainty in the chemical analysis of solutes in streamwater for Gomadansan (2003–2007), Hubbard Brook (1999–2011) and Coweeta (2003–2008). Accuracy is the difference between the analysis and the certified concentration, reported as the average of the absolute values of the errors (absolute accuracy) and as the average of the positive and negative errors (bias). Precision describes the variation in replicate analysis of the same sample. Detection limits were determined from the precision of low-concentration or zero-concentration samples. Values below detection were not used to calculate accuracy and precision

	Ca	Mg	K	Na	NH ₄	NO ₃	SO ₄	PO ₄	Cl	SiO ₂
	Absolute Accuracy (%)									
Gomadazan	12.4	18.6	11.8	4.2	3.8	8.4	15.7	ND	14.4	ND
Hubbard Brook	1.5	5.5	3.4	4.3	6.5	5.6	2.4	6.3	5.3	3.5
Coweeta	4.1	3.4	5.8	4.0	10.6	6.7	6.1	11.2	5.1	5.6
	Bias (%)									
Gomadazan	-12.4	-18.6	11.8	4.1	-3.8	8.4	15.7	ND	14.4	ND
Hubbard Brook	0.0	-2.0	0.3	1.8	-4.5	1.4	-1.7	2.0	2.2	2.3
Coweeta	-0.7	0.9	-4.2	-0.5	-0.9	-4.2	-4.5	0.1	-1.4	4.6
	Precision (%)									
Gomadazan	0.8	2.1	1.4	0.3	0.3	1.5	0.8	ND	0.6	ND
Hubbard Brook	1.0	1.0	2.3	1.2	3.6	6.8	0.6	18.1	1.2	1.6
Coweeta	1.9	0.9	1.0	0.8	3.3	0.7	1.5	15.9	2.0	0.2
	Precision (mg l ⁻¹)									
Gomadazan	0.051	0.025	0.010	0.015	0.014	0.032	0.035	ND	0.032	ND
Hubbard Brook	0.008	0.002	0.002	0.019	0.001	0.004	0.020	0.002	0.013	0.064
Coweeta	0.022	0.005	0.005	0.009	0.000	0.001	0.005	0.002	0.010	0.071
	Detection limit (mg l ⁻¹)									
Gomadazan	0.123	0.036	0.068	0.020	0.009	0.012	0.069	0.283	0.006	ND
Hubbard Brook	0.011	0.001	0.004	0.003	0.006	0.003	0.009	0.002	0.008	0.044
Coweeta	0.036	0.003	0.017	0.007	0.003	0.001	0.002	0.009	0.009	0.007

higher concentrations. For example, precision in phosphate (PO₄) analysis was poor relative to the mean (16–17%) but not in units of concentration (0.002 mg l⁻¹), compared with other solutes (Table III).

For Hubbard Brook, the method detection limit is reported as the 99% confidence interval of ten analytical blanks, about six times per year. At Coweeta, detection limits are determined quarterly as the 99% confidence interval of ten replicates using the lowest external quality control standards. For Gomadansan, replicates of the lowest available check calibrant run on the same day were used to calculate the detection limit at 99% confidence. Solutes that are near or below the detection limit have high uncertainty if reported in units relative to the mean. For example, NH₄ at Hubbard Brook was consistently below the detection limit of 0.006 mg N l⁻¹ in 2005–2006, and NH₄ exports were calculated based on a concentration of half the detection limit. Fortunately, for the purpose of calculating ecosystem budgets, it is not important to know very small fluxes with high confidence. We can be extremely confident that the important inorganic solute for N export at Hubbard Brook is NO₃, not NH₄, as streamwater concentrations of NO₃ averaged 0.08 mg N l⁻¹ for 1997–2007. In cases where knowing small concentrations is important, laboratory methods can be selected to provide lower detection limits.

In summary, uncertainties associated with chemical analysis were generally <5% for the dominant solutes. For low-concentration solutes, uncertainties may be higher in units of % but are low in units of concentration or discharge. Biases in concentration will not affect comparisons over time or among sites for samples analysed consistently but will complicate comparisons with samples analysed by other labs.

Uncertainty in stage height measurements

At Hubbard Brook and Coweeta, there is a weekly stage height calibration, which involves comparing a direct reading of stage height to the recording pen or data logger. When there is a discrepancy of more than 1 mm, there is an adjustment made to the data logger, for Coweeta, or to the chart in the gage house, at Hubbard Brook. At Hubbard Brook but not at Coweeta, the correction is pro-rated back to the previous check during data processing.

For Hubbard Brook W6, we had 11 years for which these stage height calibrations have been digitized (1997–2007) and for W1, we have 4 years (1999–2002); for Coweeta, we had 1 year (2005) for all 15 catchments. In some cases, there was a consistent bias between the gage reading and the recorder: for five streams, the hook gauge was significantly lower than the recorder, and for two streams, it was significantly higher (Figure 2). The average discrepancy ranged from -0.5 to +0.5 mm, depending on the stream; the

average across streams was not significantly different from zero ($P=0.30$).

We conducted a sensitivity analysis to quantify the effect of a stage height correction, using the driest and wettest years of the recent record at Hubbard Brook (2001 and 2003) and Coweeta (2000 and 2009). For Hubbard Brook W3, which has a 120° V-notch weir, the effect of a 1-mm adjustment was 1.3–1.7% of annual flux. For W1, which has a 90° V-notch weir, the effect of a height

adjustment of 1 mm was 1.7–2.3% of annual flux, and the effect of a 3-mm adjustment was 6–7%. At Coweeta W2, also a 120° V-notch weir, the effect of a 1-mm adjustment was 2.3–3.4% of annual flux, and the effect of a 3-mm adjustment was 8–11%. This exercise illustrates the effect of a consistent bias in stage height; random errors would tend to cancel out over time and have less effect.

Finally, at Coweeta and Hubbard Brook, a level and a survey rod are used to detect change in the position of the hook-gauge bar relative to the V-notch (Hornbeck 1965). When the reading deviates by more than about 1 mm from the previous record, the correction factor is adjusted (Figure 3). This correction is required because the weekly calibration (Figure 2) is based on a comparison to the hook gauge or the inside gauge, not the V-notch. These surveys are conducted annually on each weir within the Coweeta basin and less regularly at Hubbard Brook (Figure 3).

Uncertainty in the height–discharge relationship

At Hubbard Brook and Coweeta, rating curves were developed during the calibration period of each weir, and corrected curves are used in the estimation of discharge. At Gomadansan, the theoretical curve is used. At Hubbard Brook, we have found documentation for the development of some of these corrections. Measurements were made at low flows (up to 6 cm of stage height) using buckets, and a curve was drawn to describe the deviation of the observed from the theoretical curve (e.g. Figure 4). Only at W1 are measurements available from higher flows, and even there, the height–discharge relationship was calibrated only up to 12 cm of stage height.

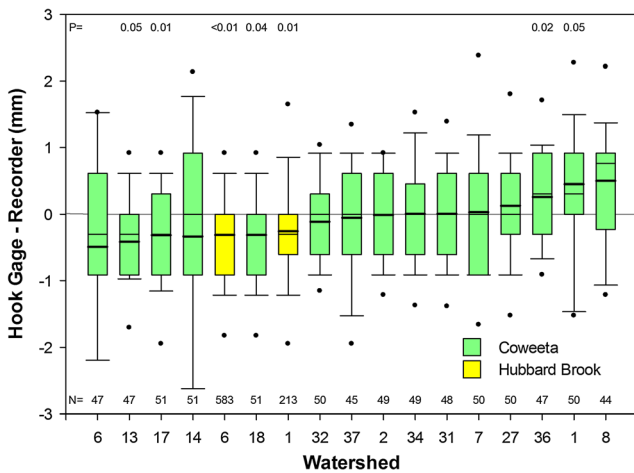


Figure 2. Weekly hook-gauge readings compared with the recording pen (Hubbard Brook) or data logger (Coweeta) for two streams at Hubbard Brook and all the streams at Coweeta, in order of the mean discrepancy, shown by a heavy line. Boxes show interquartile ranges, whiskers show the 10th and 90th percentiles, and dots show the 5th and 95th percentiles. The number of samples is given at the bottom of the chart, and the P value (for those with significant differences) is given at the top of the chart

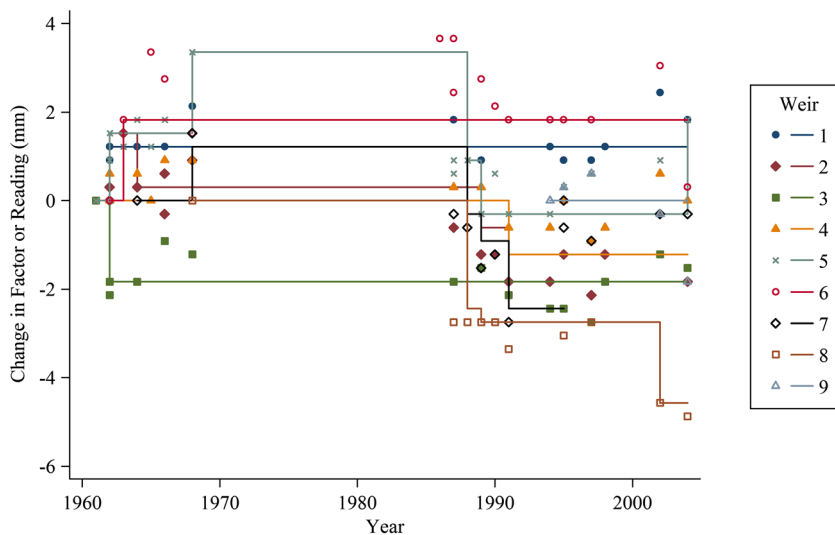


Figure 3. Surveys of the hook gauge in relation to the V-notch at Hubbard Brook, showing readings over time (symbols) and adjustments to the height factor (lines), displayed relative to the heights at the beginning of the record. Actual heights are arbitrary relative to a fixed point on the weir. In principle, corrections are made if the readings differ by more than 1 mm from the previous correction

We found data for 79 measurements of flow at stage heights from 0.6 to 6 cm at Hubbard Brook W4, made in 1963 (Figure 5). The measured flow was consistently greater than predicted by the theoretical curve ($P=0.001$), but the magnitude of the difference was small in units of stage height (<0.3 mm).

We quantified the effect of using the calibrated curve for each of the streams at Hubbard Brook by comparing annual discharge estimated by the theoretical curves to that using the corrected curves for each stream, for stage height data collected from 1994 to 2007. Although the corrections to the curves were greater than 10% at some conditions (Figure 4), the annual fluxes differed very little. Discharge estimated with the calibrated curves was lower than the theoretical calculation for W1 and W6, on average, while for W2–W5, the calibration resulted in higher calculated discharge, and for W7 and W8, some years were higher and some lower. Except for W1, the annual discrepancies were always less than 0.5%, and the average annual discrepancy was 0.08%, or about 0.6 mm of an annual average of 843 mm. For W1, the effect of the calibration

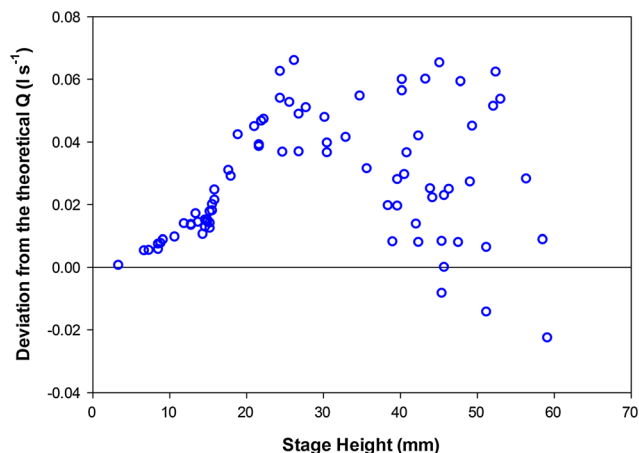


Figure 5. Calibration measurements made at weir 4 at Hubbard Brook in 1963

was greater, ranging from 1.6 to 2% of the total, because the curve was adjusted up to 15 cm of stage height.

Note that the calibration of the height–discharge relationship pertains only to relatively low flows; none of our sites has adjusted discharge estimates at stage heights >15 cm. Discharge rates higher than the calibration (12 cm in Figure 4; 6 cm at the other weirs) are important, accounting for 95% of the water flux for W6 and 98% for W3 at Hubbard Brook. For this reason, uncertainty in discharge at high flows is potentially significant but difficult to quantify (Hersch, 1995).

Uncertainty in filling gaps in the discharge record

Although the record of discharge, unlike that for chemistry, is nearly continuous, there are occasions when discharge information is not available. Weirs are cleaned annually, which requires that the ponding basins be drained, but during times of low or constant flow, this introduces minimal uncertainty in the record. Similarly, annual maintenance of the flumes at Hubbard Brook can be scheduled when flow is low and the flumes are not in use. Repairs to the weirs are needed less often; these may require hours or days. The longest gaps at Coweeta were for weir repair, with six weirs down for durations of 3–5 months in 2003–2004.

In addition to routine maintenance causing gaps in the record, accidents can happen. The longest gap in the discharge record at Hubbard Brook was caused by a gas explosion at W9 in December 2010, which damaged the insulation to the well and required a month to be repaired. The gas heaters that maintain ice-free basins and weirs at Hubbard Brook are checked twice a week; thus, only short gaps normally result from mechanical problems or empty fuel tanks, and streamflows are often

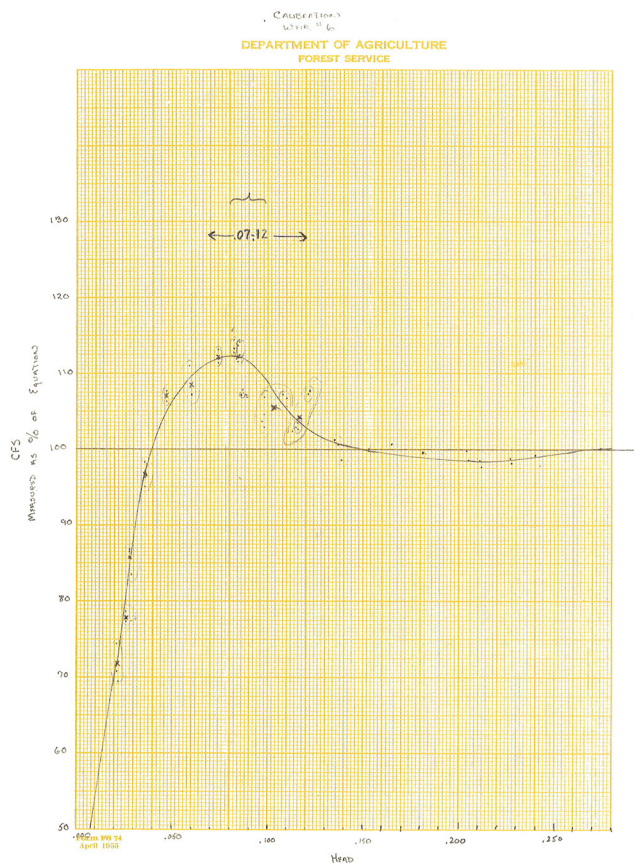


Figure 4. Rating curve for weir 6 at Hubbard Brook developed in 1965. The hand-drawn curve was used to create a look-up table to replace the theoretical height–discharge relationship at low flow

uniform and low during the coldest periods. When the heaters fail in the V-notch, it is possible for the ice to freeze in a form that siphons the water out of the ponding basin and artificially lowers the basin stage. Gaps in the flume data result primarily from debris blocking the throat of the flume, including ice during spring thaw.

We characterized the length of gaps in the discharge record at all three sites. At Hubbard Brook, averaging across nine streams, the best year (2006) had only 2.3 days of gaps per stream; the worst year (2008) had 25 days per stream with missing data. Routine maintenance accounted for 1.7 days per stream per year, on average. Incidents related to ice accounted for 1.7 additional days per year, while equipment failures averaged almost 12 days per year. At Hubbard Brook, more than half of the gaps were caused by problems with the chart recorders, usually the clocks, which are over 50 years old and are no longer manufactured. In 2010, Campbell Scientific shaft encoders were installed, which are monitored using a data logger with an internal clock, and these have proven to be more reliable. Because these transmit data hourly, problems can be detected more quickly than by relying on the weekly rounds of sample collection.

For Coweeta, we described the period from 2005 to 2009, which avoids the gaps of several months due to weir repair in 2003–2004. The average frequency of gaps in the discharge record from 2005 to 2009 was only 1.5 days per stream per year. During that period, maintenance accounted for 0.8 days per stream per year, and equipment failure averaged 0.5 days, usually caused by data logger or battery failures. Ice accounted for gaps of only 0.2 days per stream per year.

At Gomadansan, from 2003 to 2009, the average number of days per year without discharge data ranged from 6 per year in S17 to 166 per year in S5. The longest gaps are caused by landslides, which can damage the weirs and the data loggers. Gaps also result when the batteries run down between visits, which happens most often during the winter.

At all three sites, gaps in the discharge record are filled by comparison with other weirs. At Hubbard Brook, small gaps (minutes or hours) are filled during the editing process, by overlaying the hydrograph with the gap onto the hydrograph of a similar stream. In the case of missing peaks, a regression relationship is used; the r^2 of these regressions is typically very good (0.98).

We estimated the uncertainty associated with filling gaps by regression in the discharge record at Gomadansan. We used 5 years of observations (2003–2007) to create regressions of pairs of streams

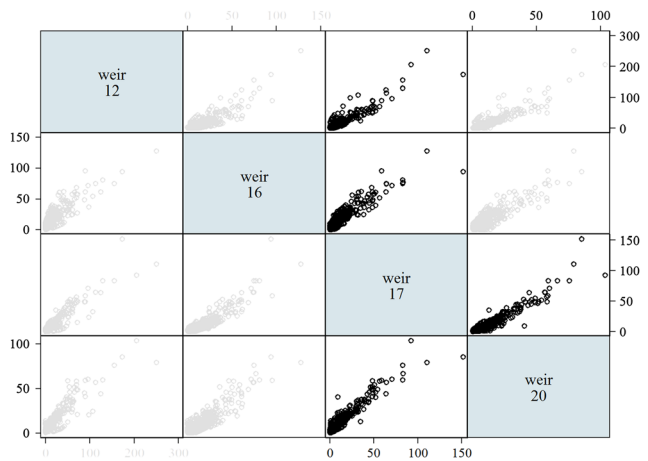


Figure 6. Regressions comparing stream pairs at Gomadansan, for use in filling gaps in the discharge record. The best regression for each stream is shown in black; those in grey were not used in the gap-filling exercise

and chose the pairs with the best r^2 to estimate missing observations (Figure 6). We created artificial gaps of the duration and time of year of the actual gaps but in a different year. We then filled the gaps by predictions based on regressions (excluding the observations from the artificial gaps) and compared the estimated with the observed discharge. The average absolute value of the error in daily discharge was 1.5 mm. Gaps of 1–3 days resulted in less than 0.5% error in the annual estimate of flow (Figure 7). Gaps of 1–2 weeks gave an average error of 1% of annual flow. Longer gaps still resulted in <2% error, except for two long gaps of 2 or 3 months that gave errors of 7–8% (Figure 7).

An alternative method of filling gaps is simulation modelling based on precipitation and catchment characteristics. At Gomadansan, there were two periods in 2008 (21–31 March and 3–11 November) when data were not available from any of the weirs. A hydrologic model (Sato *et al.*, 2008) previously validated at this site was used to simulate discharge for these time periods. We found the model predictions to differ from observations by an average of 5–9 mm for any 10-day period in 2008, depending on the stream.

Uncertainty in catchment area

At all sites, the volume of water passing the weir is divided by the area of the catchment to calculate flows per unit area, as this allows comparisons of streams draining areas of different sizes. Unlike measurements of solute concentration and stage height, the catchment area is assumed not to change over time, and areas initially assigned to each catchment have been used at all our sites without further scrutiny.

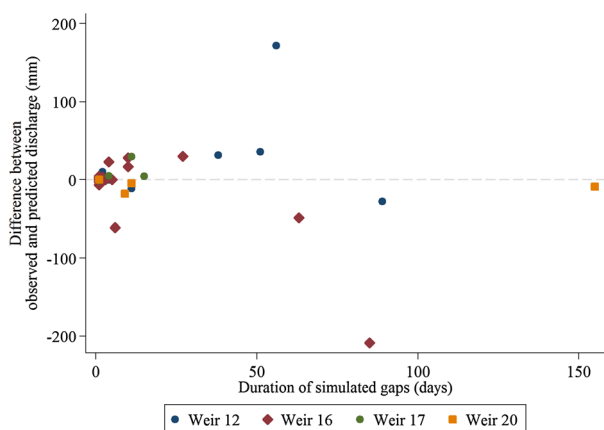


Figure 7. Error introduced by filling gaps in the discharge record at Gomadansan, based on filling artificial gaps and comparing the predicted to actual discharge

At Gomadansan, the catchment areas were estimated by drawing topographic boundaries on a map in ArcGIS. At Coweeta and Hubbard Brook, the topographic boundaries were surveyed on the ground beginning in the 1930s and 1950s respectively; aerial photography was used to assist in boundary location at Hubbard Brook. Closure error in the Hubbard Brook surveys ranged from 5 to 11 m, which was 0.2 to 0.4% of the boundary length.

High-resolution digital topographic maps have become available through the application of remote-sensing technology. We used a LIDAR digital elevation map with 1-m resolution to estimate the boundaries and areas of the catchments at Hubbard Brook. These differed by up to 5% from the surveyed areas with an average error of 2.5% and no significant bias ($P=0.69$).

Note that these watershed delineation approaches are based on the surface topography of a catchment, which may not coincide with the subsurface topography; hydrologic divides may also change with the height of the groundwater (Winter *et al.*, 2003). Discrepancies between the surface and subsurface hydrologic gradients may be large in karst terrain (Genereux *et al.*, 1993), but in our sites, there is thought to be little water flow through the bedrock. Clearly, lateral water movement beneath surface topographic divides would contribute uncertainty to rates of stream fluxes reported per unit area of the catchment. Similarly, any leakage that causes flow to bypass the weir could also contribute to uncertainty in the flux (Dresel *et al.* 2012).

Uncertainty in model selection

To calculate annual export of solutes requires multiplying concentration by discharge at some time

scale and summing over the year. Slightly different models have been used to multiply solute concentration by discharge at the three sites. Before the use of computers, this was done using a constant estimate of concentration between sampling dates. At Hubbard Brook, the measured concentration is used on the day of sample collection, and for other days, the concentration used is the average of the preceding and following samples. At Coweeta, weekly discharge is multiplied by the concentration measured at the end of the week. At Gomadansan, fluxes were calculated using linear interpolation of solute concentration between sampling dates.

We compared the effect of these three different interpolation methods for NO_3 and Ca, which have contrasting behaviour, at Hubbard Brook W6 for 1997–2008. The differences in estimates of annual exports ranged from 2 to 12% for NO_3 ((maximum – minimum)/median) but were much smaller for Ca, ranging from 0.3 to 4%, depending on the year. Calcium concentrations were more constant over time (averaging 1% change between weekly observations), so there was less uncertainty in interpolation than for NO_3 , which was more variable (19% change weekly) at this site (based on weekly observations in W6 for 1997–2008).

The importance of the selection of interpolation method depends on the frequency of sampling (Birgand *et al.* 2010); there would be no uncertainty in interpolation if concentrations were measured continuously. We simulated monthly sampling of Hubbard Brook W6 and found that model selection uncertainty increased to 5 to 30% for NO_3 but was almost unchanged for Ca (0.3 to 4%). Regression models, such as those describing concentration as a function of discharge, provide another approach to interpolation, which might better describe variation between sampling dates (Johnson *et al.*, 1969; Birgand *et al.*, 2010; Verma *et al.*, 2012). Uncertainties in these models are difficult to assess because they depend on the autocorrelation of solute concentrations over time.

We evaluated the uncertainty in stream load estimates associated with the frequency of sampling, using the Hubbard Brook weekly data for NO_3 and Ca at W6 and simulating monthly sampling as a subset of the weekly data. Using monthly samples resulted in estimates of annual export that differed from the sum of weekly observations by 1 – 66% for NO_3 , averaging 26% for the years 1997–2008. For Ca, the difference was 1 – 39%, averaging 6%. Uncertainties would be greater for solutes that are more variable in concentration over time.

Natural variation across streams and years

Natural variation in space and time is another source of uncertainty in estimating stream discharge and nutrient

export. Unlike measurement uncertainty, natural variation cannot be reduced by improved measurements, although the magnitude of the variation can be described with greater confidence.

In many disciplines, natural variation is characterized by sampling multiple experimental units, and the resulting uncertainty is described as sampling error. There has been some resistance to use of streams as replicates in ecosystem science, where manipulations of small-watershed ecosystems are difficult to replicate, and each one is considered unique. Within the sites, the streams have different treatment histories (Table II), including planting to different species at Coweeta, a Ca addition at Hubbard Brook, and clearcutting at all three sites.

These streams can be treated as sampling units, if the population of interest is the landscape represented at each site, and variation within the sites can be compared with variation across the three sites. Variation from year to year is important because it affects confidence in annual ecosystem budgets. We characterized the variation among streams and years within the sites using the CV (Table IV).

Hubbard Brook had the most consistent runoff per unit area (Figure 8), with a CV across streams averaging only 5% (3 to 9%, depending on the year). At Gomadansan, variation in runoff across streams averaged 28% and at Coweeta 36% (Table IV). The catchments at Coweeta vary more in size and cover type, while the streams at

Hubbard Brook and Gomadansan all drain small headwater catchments (Figure 1, Table II).

Variation in annual runoff across years depends on climate characteristics as well as the hydrology of the catchments. For variation over time, Hubbard Brook had the lowest variation in runoff over the period 2000–2009, averaging 23%, with streams in very close agreement (22–25%). At Coweeta, interannual variation averaged 36% with a range from 20 to 50%. At Gomadansan, for 2003–2009, the range was 20–36%, averaging 28% (Table IV).

Variation among streams in solute concentrations was highest for NO_3 (Figure 8), likely reflecting differences in biological processing of nitrogen. At Gomadansan, NO_3 was high in streams following clearcutting. At Hubbard Brook, some catchments were more affected than others by an ice storm in 1998 (Bernhardt et al., 2003). Differences in vegetation are also reflected in stream pH (data not shown); for example, Hubbard Brook W9, which has the lowest pH, is predominantly coniferous with areas of wetlands.

At Hubbard Brook, sulphate (SO_4) and chloride (Cl), which are derived from atmospheric inputs, varied little across streams, averaging 8% CV (Table IV, Figure 8). In contrast, the concentration of solutes derived from weathering varied more, presumably reflecting variation in parent material or hydrologic contact time. At Hubbard Brook, the CVs for Ca, Mg, K, and silicate (SiO_2) averaged 20%. At Gomadansan, SO_4 was most variable

Table IV. Magnitudes of uncertainty from various sources. Sampling uncertainty is reported for runoff and annual weighted concentration of solutes, rather than for annual solute export, which is the product of runoff and concentration

Measurement uncertainty

Laboratory analysis of solutes: 1–5% for dominant solutes (Table III)

Stage height measurement: 1% of annual runoff

Gap filling: 1% for gaps of 1–2 weeks at Gomadansan

Catchment area: 0–5% at Hubbard Brook

Frequency of solute sampling: 1–66% for NO_3 and 1–39% for Ca at Hubbard Brook

Sampling uncertainty

Coefficient of variation (%)	runoff	Ca	Mg	K	Na	NH_4	NO_3	SO_4	PO_4	Cl	SiO_2
Natural variation across streams (average for all years)											
Gomadansan	28	22	26	11	10	38	118	33	—	19	20
Hubbard Brook	5	29	17	24	15	9	65	8	—	7	16
Coweeta	36	33	29	24	28	22	188	53	39	23	22
Natural variation over time (average for all streams)											
Gomadansan	25	26	16	7	5	42	50	9	—	7	19
Hubbard Brook	23	10	9	19	7	9	59	11	—	17	7
Coweeta	35	8	6	8	4	35	41	5	51	6	4

Model uncertainty

Height–discharge calibration: 0.1–2% at low flow at Hubbard Brook

Height–discharge calibration at high flow: unknown

Model selection error: 2–12% for NO_3 and 0.3–4% for Ca at Hubbard Brook

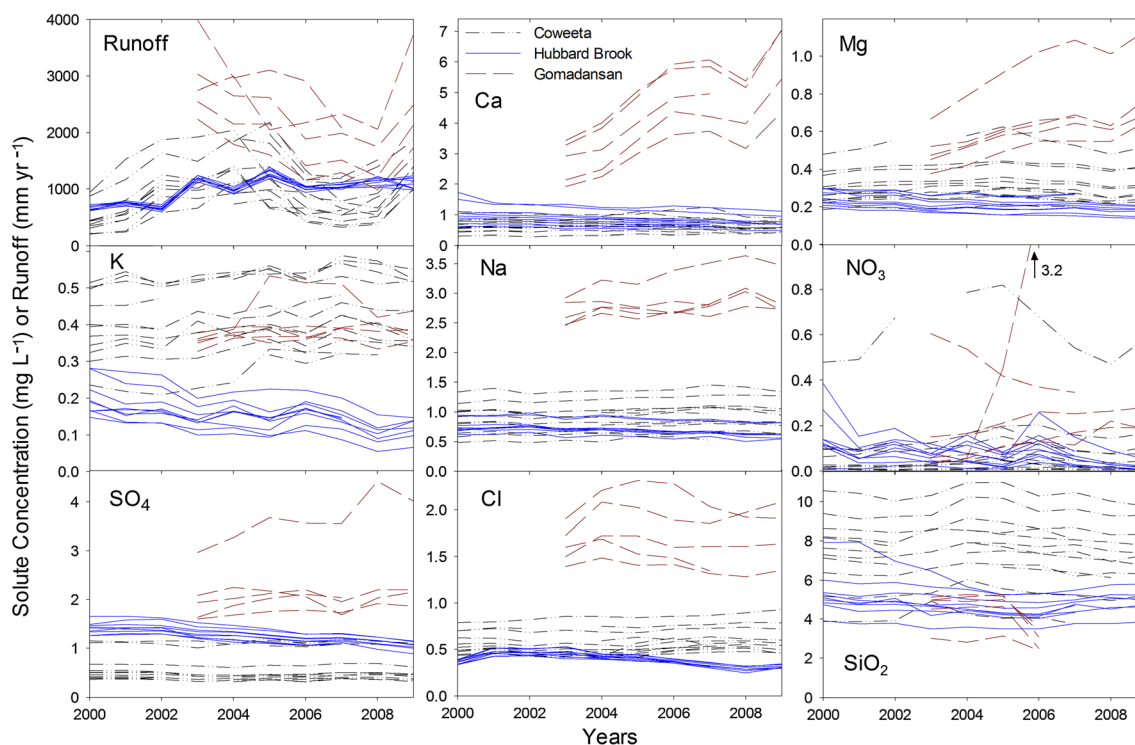


Figure 8. Annual runoff and volume-weighted average annual concentrations of solutes for multiple watersheds and multiple years at Gomadansan, Hubbard Brook, and Coweeta. Solute export is the product of runoff and concentration; interannual variation in solute export (not shown) follows the pattern in runoff

across streams (after NO_3 and NH_4), Na was least variable, and the base cations averaged 17%. At Coweeta, variation across streams in solute concentration was higher than at the other sites for most elements (Table IV), perhaps reflecting local differences in parent material (Swank and Waide, 1988).

Interannual variation in solute concentration was lowest at Coweeta for most solutes (Table IV, Figure 8), with coefficients of variation ranging from 4% (for Na) to 41% (for NO_3 ; average across streams). Variation over time was highest for NO_3 at all three sites (Figure 8). The stream with the highest NO_3 concentrations was the catchment that was clearcut in 2003 at Gomadansan. At Coweeta, high-elevation catchments tended to have greater NO_3 concentrations compared with low-elevation catchments. At Hubbard Brook, the streams with high NO_3 in water year 2000 were those that were affected by the 1998 ice storm (Houlton *et al.*, 2003; Judd *et al.*, 2007). Also visible at Hubbard Brook is the effect of a catchment-scale wollastonite (CaSiO_3) addition in 1999, which resulted in long-term changes in several major solutes (Ca, SiO_2 , H^+ and SO_4 ; Peters *et al.*, 2004).

Phosphate concentrations are not shown because PO_4 was not analysed at Gomadansan, and concentrations at Hubbard Brook were usually below

detection limits. For Coweeta, where detection limits are lower, PO_4 concentrations varied by 51% from year to year and 39% from stream to stream (Table IV). However, variation was small in units of concentration or annual export. Uncertainties in small values need to be viewed in perspective; if they are combined with larger values, for example, in ecosystem nutrient budgets, they may contribute little additional uncertainty to the total.

For most solutes, the sites were quite distinct, in spite of the variation within sites (Figure 8). Gomadansan had the highest concentrations of Ca, Na, SO_4 , and Cl. For K and Cl, Hubbard Brook streams were lower than the other sites, especially in the more recent years. Silica was highest at Coweeta, which is underlain by saprolite (Swank and Waide, 1988); only the Hubbard Brook stream that had a wollastonite addition (CaSiO_2) was in the range of the Coweeta data.

DISCUSSION

In some ecosystem studies, uncertainties can be so large as to make the results difficult to interpret. For example, in a salt marsh in the northern Gulf of Mexico, the

uncertainties in inputs and outputs were larger than the estimated fluxes (Lehrter and Cebrian, 2010). In forests, high uncertainty in soil nutrient pools can make it impossible to establish confidence in ecosystem sources and sinks (Yanai *et al.*, 2012). For stream fluxes of nutrients, however, at least for the sites and solutes that we considered in this paper, uncertainties in measurements and models were generally small compared with natural variation over time and across sites. Therefore, comparisons of catchments, or comparisons of time periods within catchments, are not likely to be at risk of making spurious conclusions based on random error.

There is a need for continued development of uncertainty analysis applied to hydrologic studies. Although many sources of uncertainty have been identified and quantified (Harmel *et al.* 2006, McMillan *et al.* 2012), there has yet to be a comprehensive analysis of uncertainty in solute export from a small headwater catchment. Some important components of the uncertainty in stream fluxes have yet to be quantified, such as the uncertainty in the height–discharge relationship at high flow. For models, assigning uncertainty to the interpolation of stream chemistry between sampling dates requires information on temporal autocorrelation at time steps finer than the existing measurements. In other words, the uncertainty of stream concentration at a time not sampled depends on how close it is in time to a sampled point and how quickly over time the concentrations become independent. In addition, to propagate all the sources of uncertainty through the final calculation of stream exports requires not only estimates of uncertainty in all the component measurements and models but also the relationship among these sources. When they are not independent, uncertainties can be amplified (in the case of correlated errors) or they may tend to cancel out (if negatively correlated). This information may be more difficult to obtain than the magnitudes of the individual uncertainties.

Although the complexities can be daunting, it is important to pursue closer approximations to the true uncertainty. Even crude estimates of uncertainty can be extremely useful. Knowing which aspects of a problem contribute the greatest uncertainty makes it possible to direct attention to improvement where it can have the greatest benefit. For example, we found that great attention had been given to describing the height–discharge relationship at low flows, which have very little impact on annual export; finding a way to validate this relationship at high flows would have higher payoff. In the design of monitoring programmes, as in research, uncertainty analysis can help to improve allocation of effort (Harmel *et al.*, 2009; Lindenmayer and Likens, 2010, Levine *et al.*, 2014). For stream monitoring, there is a trade-off between the number of streams sampled and the frequency of sampling; the optimal allocation of effort

could be defined as that which produces the least uncertainty in the results. Quantifying uncertainty improves the value of environmental monitoring for policy purposes and ultimately allows us to monitor progress in improving scientific knowledge.

ACKNOWLEDGEMENTS

This paper is a product of Quantifying Uncertainty in Ecosystem Studies (QUEST), www.quantifyinguncertainty.org, which was supported by the National Science Foundation (NSF) and the NSF Long-Term Ecological Research (LTER) Network Office. Many of the ideas implemented in this paper were developed in a meeting of a QUEST Working Group in March 2011. In June 2011, the Japanese Society for the Promotion of Science awarded a fellowship to Ruth Yanai, hosted by Naoko Tokuchi, to work on this project. Masanori Katsuyama and Nobu Ohte applied the hydrological model at Gomadansan. Tom Luther interpolated the precipitation volumes. Zak Krug entered data from the Coweeta archives handwritten in decades past, and Bill O'Neill and Adam Wild helped produce the figures. Wayne Martin, Jamie Shanley, and Doug Burns provided helpful reviews of the manuscript.

The research at Gomadansan is supported by the staff of the Wakayama Experimental Forest of Kyoto University. The Hubbard Brook Experimental Forest is operated by the US Forest Service, and it forms part of the NSF LTER network. Research at Coweeta Hydrologic Laboratory was funded by the US Forest Service, Department of Agriculture, Southern Research Station, and the NSF LTER (DEB-9632854 and DEB-0218001). Chemistry data for Hubbard Brook were obtained through funding to G.E. Likens from the A.W. Mellon Foundation and the NSF, including the LTER and Long-Term Research in Environmental Biology programmes.

REFERENCES

- Aulenbach BT, Hooper RP. 2006. The composite method: an improved method for stream-water solute load estimation. *Hydrological Processes* **20**: 3029–3047. DOI: 10.1002/hyp.6147.
- Bernhardt ES, Likens GE, Buso DC, Driscoll CT. 2003. In-stream uptake dampens effects of major forest disturbance on watershed nitrogen export. *Proceedings of the National Academy of Sciences* **100**: 10304–10308. DOI: 10.1073/pnas.1233676100.
- Birgand F, Fauchoux C, Gruau G, Augéard B, Moatar F, Bordenave P. 2010. Uncertainties in assessing annual nitrate loads and concentration indicators: Part 1. Impact of sampling frequency and load estimation algorithms. *Transactions of the ASABE* **53**(2): 437–446.
- Bukaveckas PA, Likens GE, Winter TC, Buso DC. 1998. A comparison of methods for deriving solute flux rates using long-term data from streams in the Mirror Lake watershed. *Water, Air, and Soil Pollution* **105**: 277–293.
- Buso DC, Likens GE, Eaton JS. 2000. Chemistry of precipitation, streamwater, and lakewater from the Hubbard Brook Ecosystem Study: A record of sampling protocols and analytical procedures. Gen. Tech.

- Rep. NE-275, U.S. Department of Agriculture, Forest Service, Northeastern Research Station: Newtown Square PA; 52.
- Di Baldassarre G, Montanari A. 2009. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences* **13**: 913–921.
- Dresel PE, Hekmeijer P, Dean JF, Harvey W, Webb JA, Cook P. 2012. Use of laser-scan technology to analyse topography and flow in a weir pool. *Hydrology and Earth System Sciences Discussions* **9**: 3721–3738.
- Fukushima K, Tokuchi N. 2009. Factors controlling the acid-neutralizing capacity of Japanese cedar forest watersheds in stands of various ages and topographic characteristics. *Hydrological Processes* **23**: 259–271.
- Genereux DP, Hemond HF, Mulholland PJ. 1993. Spatial and temporal variability in streamflow generation on the West Fork of Walker Branch Watershed. *Journal of Hydrology* **142**: 137–166.
- Harmel RD, Cooper RJ, Slade RM, Haney RL, Arnold JG. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of ASABE* **49**: 689–701.
- Harmel, RD, Smith DR, King KW, Slade RM. 2009. Estimating storm discharge and water quality data uncertainty: A software tool for monitoring and modeling applications. *Environmental Modelling and Software* **24**(7): 832–842.
- Herschey RW. 1995. *Streamflow Measurement*, 2nd edition. Elsevier Applied Science: London; 524.
- Hornbeck JW. 1965. Accuracy in streamflow measurements on the Fernow Experimental Forest. Res. Note NE-29. U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station: Broomall, PA; 8.
- Houlton BZ, Driscoll CT, Fahey TJ, Likens GE, Groffman PM, Bernhardt ES, Buso DC. 2003. Nitrogen dynamics in ice storm-damaged forest ecosystems: implications for nitrogen limitation theory. *Ecosystems* **6**: 431–443.
- Johnson NM, Likens GE, Bormann FH, Fisher DW, Pierce RS. 1969. A working model for the variation in stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire. *Water Resources Research* **5**: 1353–1363. DOI: 10.1029/WR005i006p01353.
- Judd K, Likens GE, Groffman PM. 2007. High nitrate retention during winter in soils of the Hubbard Brook Experimental Forest. *Ecosystems* **10**: 217–225.
- Labagh JW, Buso DC, Likens GE. 2009. Evaluation of methods and uncertainties in the chemical budget. In *Mirror Lake: Interactions Among Air, Land and Water*, Winter TC, Likens GE (eds), University of California Press: Los Angeles; 225–299. <http://www.jstor.org/stable/10.1525/j.ctt1ppvb5>
- Lehrter JC, Cebrian J. 2010. Uncertainty propagation in an ecosystem nutrient budget. *Ecological Applications* **20**(2): 508–24.
- Levine CR, Yanai RD, Lampman GG, Burns DA, Driscoll CT, Lawrence GB, Lynch JA, Schoch N. 2014. Evaluating the efficiency of environmental monitoring programs. *Ecological Indicators* **39**: 94–101. DOI: 10.1016/j.ecolind.2013.12.010.
- Likens GE. 2013. *Biogeochemistry of a Forested Ecosystem*, 3rd edition. Springer: New York; 208.
- Lindenmayer DB, Likens GE. 2010. *Effective Ecological Monitoring*. CSIRO Publishing and Earthscan: Melbourne and London; 170.
- McMillan H, Freer J, Pappenberger F, Krueger T, Clark M. 2010. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrological Processes* **24**: 1270–1284.
- McMillan H, Krueger T, Freer J. 2012. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes* **26**: 4078–4111.
- Peters SC, Blum JD, Driscoll CT, Likens GE. 2004. Dissolution of wollastonite during the experimental manipulation of Hubbard Brook Watershed 1. *Biogeochemistry* **67**: 309–329.
- Rode M, Suhr U. 2007. Uncertainties in selected river water quality data. *Hydrology and Earth System Sciences* **11**: 863–874.
- Sato Y, Ma X, Xu J, Matsuoka M, Zheng H, Liu C, Fukushima Y. 2008. Analysis of long-term water balance in the source area of the Yellow River basin. *Hydrological Processes* **22**: 1618–1629.
- Sauer VB, Turnipseed DP. 2010. Stage measurement at gaging stations. In *U.S. Geological Survey Techniques and Methods*, Book 3, chap. A7; 45.
- Sauer VB. 2002. Standards for the analysis and processing of surface-water data and information using electronic methods. *USGS Water-Resources Investigations Report: Water Resources Investigation*; number 2001–4044.
- Swank WT, Waide JB. 1988. Characterization of baseline precipitation and stream chemistry and nutrient budgets for control watersheds. In *Ecological Studies, Vol.66: Forest Hydrology and Ecology at Coweeta*, Swank WT, Crossley Jr DA (eds). Springer-Verlag: New York; 57–79. DOI: 10.1007/978-1-4612-3732-7_4.
- Ullrich A, Volk M. 2010. Influence of different nitrate–N monitoring strategies on load estimation as a base for model calibration and evaluation. *Environmental Monitoring and Assessment* **171**: 513–527 DOI: 10.1007/s10661-009-1296-8.
- Verma S, Markus M, Cooke RA. 2012. Development of error correction techniques for nitrate-N load estimation methods. *Journal of Hydrology* **432–433**: 12–25.
- Winter TC, Rosenberry DO, LaBaugh JW. 2003. Where does the ground water in small watersheds come from? *Ground Water* **41**: 989–1000.
- Yanai RD, Levine CR, Green MB, Campbell JL. 2012. Quantifying uncertainty in forest nutrient budgets. *Journal of Forestry* **110**(8): 448–456 DOI: 10.5849/jof.11-08.