Effects of sampling frequency on estimation accuracies of annual loadings for water quality parameters in different sized watersheds

Lin Gao, Junyu Qi, Sheng Li, Glenn Benoy, Zisheng Xing and Fan-Rui Meng

ABSTRACT

Potential errors or uncertainties of annual loading estimations for water quality parameters such as suspended solids (SS), nitrate-nitrogen (NO_3-N), ortho-phosphorus (Ortho-P), potassium (K), calcium (Ca), and magnesium (Mg) can be greatly affected by sampling frequencies. In this study, annual loading estimation errors were assessed in terms of the coefficient of variation, relative bias, and probability of potential errors that were estimated with statistical samples taken at a series of sampling frequencies for a watershed in northwestern New Brunswick, Canada, and one of its sub-watersheds. Results indicate that annual loading estimation errors increased with decreasing sampling frequency for all water quality parameters. At the same sampling frequencies, the estimation errors were several times greater for the smaller watershed than those for the larger watershed, possibly due to the flushing nature of streamflows in the smaller watershed. We also found that low sampling frequency tended to underestimate the annual loadings of water quality parameters dominated by stormflow events (SS and K) and overestimate water quality parameters dominated by baseflow (Mg and Ca). These results can be used by hydrologists and water quality managers to determine sampling frequencies that minimize costs while providing acceptable estimation errors. This study also demonstrates a novel approach to assess potential errors when analyzing existing water quality data.

Key words | annual loading, estimation accuracy, sampling errors, sampling frequency, water quality, water quality parameters

Lin Gao
Junyu Qi
Sheng Li (corresponding author)
Glenn Benoy
Zisheng Xing
Fan-Rui Meng
Faculty of Forestry and Environmental Management,
University of New Brunswick,
P.O. Box 44555, 28 Dineen Drive, Fredericton, NB E3B 5A3,
Canada
E-mail: *sheng.li@canada.ca*

Junyu Qi
Earth System Science Interdisciplinary Center,
University of Maryland at College Park, University of Maryland,
College Park, MD 20742,
USA

Sheng Li
Fredericton Research and Development Centre,
Agriculture and Agri-Food Canada,
P.O. Box 20280, 850 Lincoln Road, Fredericton, NB E3B 4Z7,
Canada

Glenn Benoy
Canadian Rivers Institute,
University of New Brunswick,
P.O. Box 4400, 10 Bailey Drive, Fredericton, NB E3B 5A3,
Canada

Zisheng Xing
Brandon Research and Development Centre,
Portage la Prairie, MB R1N 3V6,
Canada

INTRODUCTION

Nonpoint source pollution and water quality degradation associated with intensive agricultural and other land-use activities are global concerns (Cartwright *et al.* 1991; Chen *et al.* 2003; Schroder *et al.* 2004; Chow *et al.* 2010). Excessive nutrients and sediments carried by water can be transported into surface and groundwater systems, causing eutrophication (Maticic 1999) and degradation of drinking water quality (Carstea *et al.* 2016). The chemical, physical, and biological characteristics of streams and other water bodies are often used as water quality indicators to evaluate the conditions of aquatic ecosystems and to track changes in water quality over time. Annual loading (AL), defined as the total mass of a pollutant passing through a cross-section of a river per year, is regarded as an effective water quality indicator (Richards 1998). As it is time-integrated, AL reflects the long-term attributes of watersheds and land uses and can be used to estimate total nutrient and soil losses from upland areas as well as total pollutant input into connecting aquatic

doi: 10.2166/wqrj.2020.012

ecosystems (Cartwright *et al.* 1991; Chen *et al.* 2003; Schroder *et al.* 2004). For example, AL of suspended solids (SS) is normally used to estimate the severity of soil erosion from agricultural fields and assess the habitat suitability of aquatic ecosystems (Chow *et al.* 2010).

Pollutant production and transport within watersheds are affected by land-use activities such as tillage, fertilization, and biophysical factors such as soil type, vegetation, topography, and geology (Bahar & Yamamuro 2008). As a result, concentrations of water pollutants vary greatly with time (Bahar & Yamamuro 2008). For example, concentrations of SS and potassium (K) of stream water are typically higher during the high-flow snow-melting seasons than other seasons (Li *et al.* 2014). On the other hand, concentrations of calcium (Ca) and magnesium (Mg) are typically lower during high-flow events than between-events baseflow periods (Li *et al.* 2014). Halliday *et al.* (2013) found strong seasonal and diurnal patterns of nitrate concentration in upland rivers. Wade *et al.* (2012) also reported that in stream water, phosphorus concentration displayed a highly complex temporal pattern under storm conditions at sub-daily time steps. Traditionally, AL is estimated from the product of average pollutant concentration and corresponding stream discharge rate for a given sampling interval. However, dynamics in both temporal variations of flow rates and pollutant concentrations make it difficult to obtain reliable and accurate estimates of AL. Given these dynamics, it is a logical assumption that increasing sampling frequency would increase AL estimation accuracy (Quimpo & Yang 1970). In particular, new instruments such as multi-parameter sondes are available for measuring certain water quality parameters such as pH, conductivity, and turbidity at high frequencies with reasonable accuracy (Blaen *et al.* 2016; Carstea *et al.* 2016; Reynolds *et al.* 2016). However, the laboratory analysis is often required for many other water quality parameters in order to obtain acceptable measurement accuracy and reliability. High-frequency water sampling and intensive laboratory analysis pose challenges as they are labour-intensive and expensive.

A common need for many hydrologists, water quality specialists, and water resource managers is to find the minimum sampling frequency to obtain a reasonable accuracy at acceptable costs. Previous studies have compared estimation errors or uncertainties of AL with different sampling frequencies for a number of water quality parameters (Johnes 2007; Jones *et al.* 2012; Cassidy & Jordan 2011; Elwan *et al.* 2014). Rozemeijer *et al.* (2010) noted potential large estimation errors of AL of nitrate-nitrogen (NO_3-N) and P in streams related to low sampling frequencies.

Since concentrations of many pollutants are strongly affected by flow events (i.e., rainfall and snowmelt), the minimum sampling frequency needs to be able to capture these events. However, watershed biogeochemical responses to rainfall events can vary greatly with precipitation characteristics, antecedent soil moisture, and seasonal variations in biophysical and geochemical conditions. The chemical composition of rainwater as well as rainfall–runoff processes also have great impacts on nutrients entering into streams (Kirchner *et al.* 2004; Malve *et al.* 2012) and indirectly affect loading (Rode *et al.* 2016). During rainfall events, chemicals can be transported by different pathways, either through surface runoff or together with the baseflow. Low sample frequency has a risk of missing major events and can lead to systematic bias for AL estimation. There is a need to quantify the effect of sampling frequencies on the uncertainties of AL as well as the potential systematic biases of AL estimation related to low sampling frequencies.

The objective of this study was to analyze and quantify the effects of sampling frequencies on the uncertainties of AL and the potential biases of AL estimation with low-frequency sampling for SS and agricultural nutrients (NO_3-N, ortho-phosphorus (Ortho-P), K, Ca, and Mg) in streams.

METHODS AND MATERIALS

Study sites

Catchment size and land use have significant impacts on nutrient and sediment transport, thus providing estimates for AL are challenging (Malve *et al.* 2012). We used a relatively large watershed in northwestern New Brunswick, Canada, Little River Watershed (LRW) and one of its sub-basins, Black Brook Watershed (BBW) as study sites (Figure 1). The area of the LRW is approximately 389 km^2 with approximately 16.2% agricultural land, 77% forests, and 6.8% residential area (Xing *et al.* 2013). Elevation in the LRW ranges from 122 to 477 m above

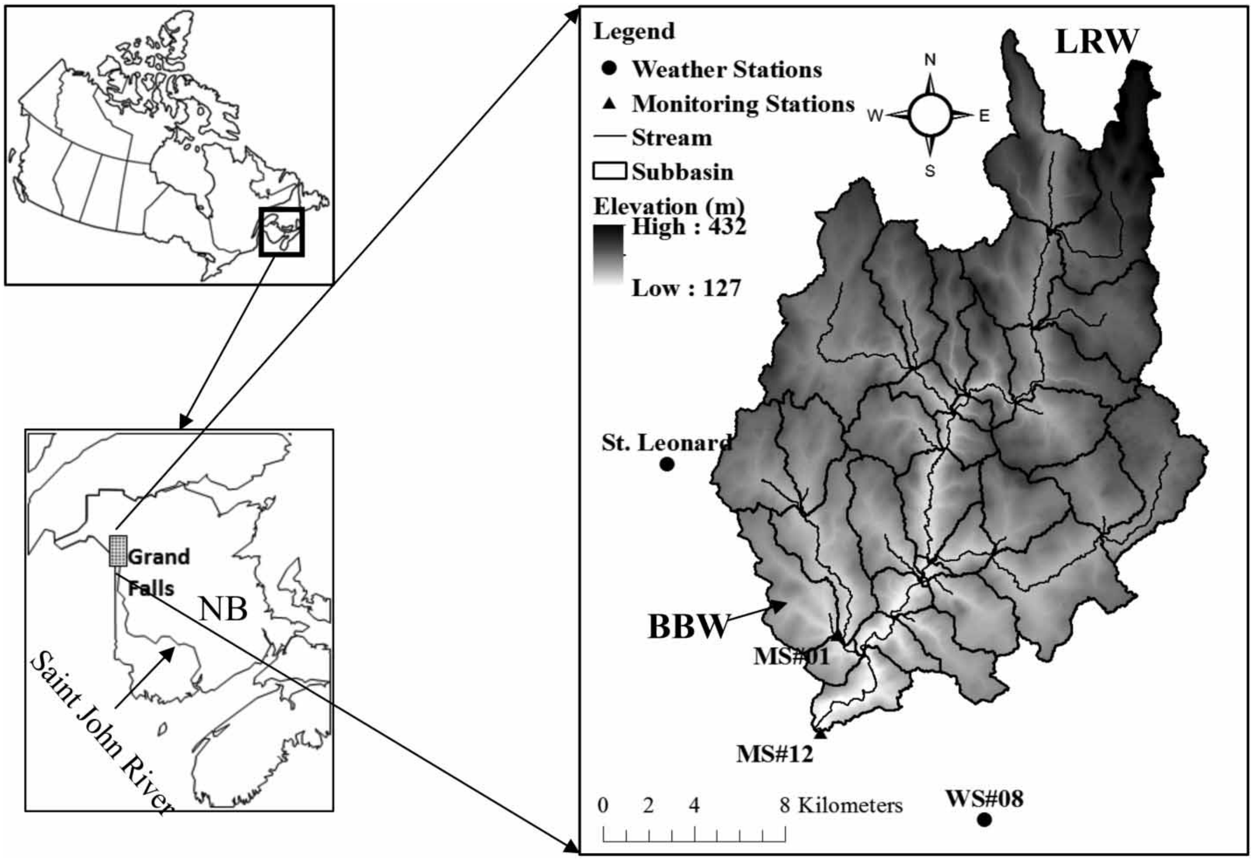

Figure 1 | Description of study sites: LRW and BBW.

sea level (Chow *et al.* 2010). The area of the BBW is 14.5 km^2 with approximately 65% agricultural land, 21% forests, and the remainder residential areas (Chow *et al.* 2010). Elevation in the BBW ranges from 180 to 260 m above sea level. The average annual precipitation (based on 30 years data) in this area is 1,134 mm, and about one-third of the precipitation is in the form of snow (Qi *et al.* 2017b). The maximum streamflow and groundwater recharge occur during the snowmelt season from March to May (Chow & Rees 2006). In 1989, the BBW was established as an Agriculture and Agri-Food Canada (AAFC) experimental watershed and used as a national benchmark for monitoring agricultural effects on stream water quality (Yang *et al.* 2009; Qi *et al.* 2017a). The watershed was intensively instrumented to study soil erosion in croplands, primarily potato fields, and the impacts of beneficial management practices on water quality. The contrasting sizes and land uses of these two watersheds and the existence of long-term historical water quality data make them the ideal sites for addressing questions about the estimation accuracy of ALs.

Raw data collection and sample analyses

High-frequency sampling data records from 2003 to 2008 were obtained from two long-term gauging stations at the outlets of LRW and BBW (Figure 2). At each station, the water level stage was monitored in a stilling well connected to the stream. Stage heights were converted to flow rates based on locally calibrated rating curves (fitted polynomial functions, Chow *et al.* 2010). In order to capture variation during storm events, stage heights were recorded using a datalogger at a fixed interval of 1 h without significant changes of the water level (<2 cm); additional records were made if the stage height changed more than 2 cm in 5 min. Thus, the water level data can be considered 5-min interval

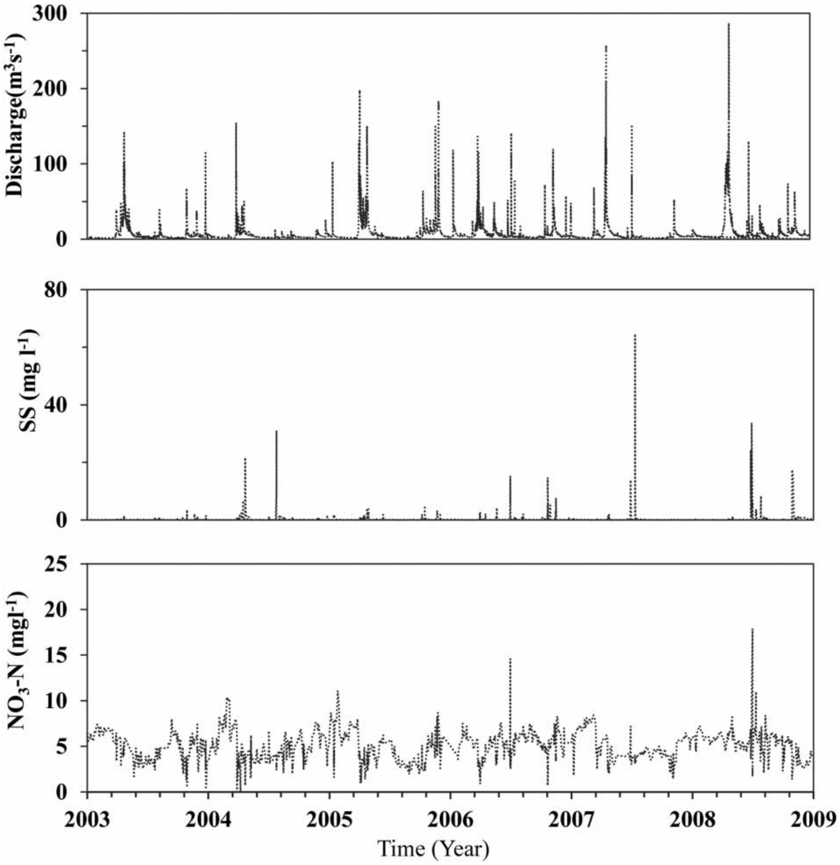

Figure 2 | Stream discharge, concentrations of SS, and NO_3-N from the year 2003 to 2008 in BBW, NB, Canada.

continuous data with a resolution of 2 cm. The sampling protocol was designed based on previous observations of the dynamics of pollutant concentrations in the system (Chow *et al.* 2010). Water samples were collected using an ISCO (Teledyne ISCO, Lincoln, NB, USA) automatic sampler controlled by the same datalogger (Chow *et al.* 2010). Sampling frequency was set at one sample every 72 h when there was no runoff event (water level changes <5 cm). During runoff events, sampling frequency increased with one sample for every 5 cm change in the water level. Water samples were packed with ice, stored in a cooler, and transported to the Soil Hydrology Lab at the AAFC Fredericton Research and Development Centre (FRDC). Before the chemical analysis, water samples were filtered through a 0.45 μm semi-permeable membrane (Millipore Crop., Billerica, MA, USA) to remove solid particles. The filtered water samples were then stored in the refrigerator and used for determining water quality parameters, including nitrate-nitrogen (NO_3-N), ortho-phosphorus (Ortho-P), K, Ca, and Mg, as well as pH and electrical conductivity (Chow *et al.* 2010).

Resampling method and sampling frequency

A 5-min interval dataset of discharge, SS, and nutrient concentrations was generated by interpolating the raw measurement data for each of the two watersheds from 2003 to 2008 using Model Maker (Cherwell Scientific Ltd). This high-frequency dataset was then resampled to create 17 sub-datasets of a series of sampling intervals: 5, 10, 15, 30, 60, and 120 min; 0.5, 1, 2, 3, 5, 7, and 14 days; and 1, 3, 6, and 12 months. For sampling intervals ≤120 min, resampling selection is limited and, therefore, all possible resampling samples are used. For sampling intervals ≥0.5 day, 50 random samples were generated using randomly selected starting points determined using Monte Carlo random seed generators.

Annual loading calculation

Following Ferguson (1986), AL of a parameter was calculated by the following equation:

$$\mathrm{AL} = \sum_{i=1}^{n} C_i \cdot Q_i \cdot \Delta T_i \tag{1}$$

where AL is the annual loading (ton year^{-1} or kg year^{-1}); n is the total number of sampling points in a year; C_i is the parameter concentration of the ith sampling point (ppm or ppb); Q_i is the flow rate at the time of the ith sampling point (m^3 s^{-1}); ΔT is the time interval (s).

In the present study, concentrations of SS, NO_3-N, Ca, Mg, and K were in ppm, and that of Ortho-P was in ppb; ALs of NO_3-N, Ortho-P, Ca, Mg, and K were in kg year^{-1}, and that of SS was in ton year^{-1}. The annual flow discharge was in m^3 year^{-1}.

Accuracy assessment

Many statistical indicators can be used to measure AL estimation errors or accuracy (Walther & Moore 2005). In this study, we used three indicators to measure AL estimation errors: coefficient of variation (CV), relative bias (RB), and probability of potential errors (PPE). The CV reflects the relative dispersion of samples at different sampling frequencies (Rus et al. 2012), while RB measures systematic errors at different sampling frequencies. The PPE measures the probability of errors to exceed selected thresholds (1%, 5%, and 10% in this study) for corresponding sampling frequencies.

Following Sokal & Rohlf (1995) and Kozak et al. (2013), the CV of AL (CV_AL) was calculated from the standard deviation σ by the following equations:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}} \tag{2}$$

$$\mathrm{CV\_AL} = \frac{\sigma}{\mu} \cdot 100 \tag{3}$$

where N is the total number of samples for a given sampling frequency; x_i is the AL of the ith sample for a given sampling frequency; μ is the average AL for the given sampling frequency.

CV_AL was modelled as a function of the sampling frequencies as shown in the following equation:

$$\mathrm{CV\_AL} = a \cdot F^{-b} \tag{4}$$

where F is the sampling frequency (# of samples year^{-1}); parameters a and b define the extent and shape of the curve fitting CV_AL with sampling frequency F.

We hypothesized that the variations of different water quality parameters should also be related to the variations of stream discharge. Therefore, the CV_AL of water quality parameters can be expressed as a function of the CV_AL of discharge as shown in the following equation:

$$\mathrm{CV\_AL_{ion}} = k_1 \cdot \mathrm{CV\_AL_{dis}}^{k_2} \tag{5}$$

where $\mathrm{CV\_AL_{ion}}$ is the CV_AL of the water quality parameters; $\mathrm{CV\_AL_{dis}}$ is the CV_AL of stream discharge. Parameters k_1 and k_2 define the shape of the curve for different water quality parameters.

The RB of AL (RB_AL) represents systematic errors. From the statistical point of view, the bias could cause more concern to scientists. In this study, RB_AL was calculated by the following equation:

$$\mathrm{RB\_AL} = \sum_{k=1}^{N} \frac{\mathrm{EA}_k - \mathrm{TA}}{\mathrm{TA}} \cdot 100\% \tag{6}$$

where TA is the AL for the baseline sampling frequency, in this case, the highest sampling frequency; EA_k is the AL for the kth statistical sample for a given sampling frequency; N is the total number of statistical samples for a given sampling frequency.

The PPE of AL (PPE_AL) was defined as the occurring frequency of the RB value greater than selected thresholds. In this study, thresholds were set at ±1%, ±5%, and ±10%, and PPE_AL was calculated as follows:

$$\mathrm{PPE\_AL}\ (1) = n(\mathrm{RB} \le -1\% \text{ or } \mathrm{RB} \ge +1\%)/N \cdot 100\% \tag{7}$$

$$\mathrm{PPE\_AL}\ (5) = n(\mathrm{RB} \le -5\% \text{ or } \mathrm{RB} \ge +5\%)/N \cdot 100\% \tag{8}$$

$$\mathrm{PPE\_AL}\ (10) = n(\mathrm{RB} \le -10\% \text{ or } \mathrm{RB} \ge +10\%)/N \cdot 100\% \tag{9}$$

where n is the number of statistical samples with RB greater than the thresholds indicated in the brackets; N is the total number of statistical samples for a given sampling frequency. The units of CV_AL, RB_AL, and PPE_AL were converted to percentages.

RESULTS AND DISCUSSION

Coefficients of variation of AL

There was a consistent trend of decreasing CV_AL values with increasing sampling frequency (Figure 3 and Table 1). The highest CV_AL values typically occurred at the low sampling frequencies (large sampling interval, such as seasonally to yearly), whereas the lowest CV_AL values typically occurred at the high sampling frequencies (small sampling interval, such as hourly). Above certain sampling frequencies, the decrease of CV_AL became less obvious. This trend agrees with general expectations as with higher sampling frequencies, there was less chance to miss significant flow events.

There were substantial differences between the two watersheds. The CV_AL values for the large watershed (LRW) were consistently many times lower than that of the smaller sub-watershed (BBW) for all parameters and all sampling frequencies (Table 1 and Figure 3). For example, with daily sampling, the CV_AL values for the larger (LRW) were all less than 1% except SS (5.33%). In contrast, the CV_AL values were greater than 4% in all cases for the smaller sub-watershed (BBW, Table 1). In particular, the CV_AL for total stream discharge was 5.58% for BBW, more than 10 times of the 0.55% for LRW. For SS, the CV_AL was 22.21% for BBW, four times the 5.33% for LRW. The CV_AL of Ca was 4.08% for BBW, about seven times higher than the 0.52% for LRW. The lower mean CV_AL values for the larger sized LRW than those for the smaller sized BBW can be attributed to the lower buffering capability of a small watershed to short-duration rainfall events, resulting higher and more peaks in the hydrograph for a smaller watershed, which caused significant variations in loadings of SS and other water quality parameters. Whereas for the larger watershed, the hydrograph was flattened and flow events lasted longer, resulting in less variation in loadings of SS and other water quality

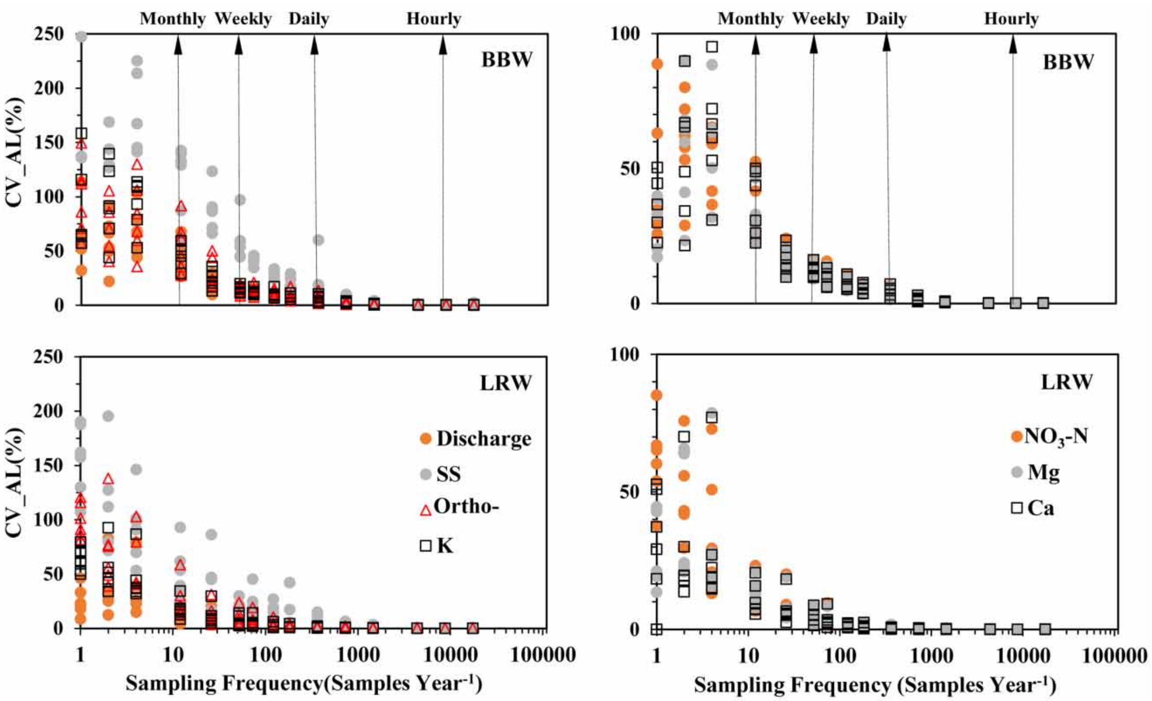

Figure 3 | The coefficients of variations of AL (CV_AL) of Stream Discharge, SS, Ortho-P, K, NO_3-N, Mg, and Ca in BBW (upper panels) and LRW (lower panels) from the year 2003 to 2008 with different sampling frequencies.

Table 1 | The mean of the CV_AL of stream discharge (Disch), SS, NO_3-N, Ortho-P, K, Mg, and Ca with different sampling frequencies in BBW and in LRW from 2003 to 2008

	Disch (%)		SS (%)		NO_3-N (%)		Ortho-P (%)		K (%)		Mg (%)		Ca (%)	
Frequency	BBW	LRW	BBW	LRW	BBW	LRW	BBW	LRW	BBW	LRW	BBW	LRW	BBW	LRW
1 (Yearly)	61.64	24.7	249.68	155.68	63.73	61.53	98.44	99.47	127.96	63.98	44.64	49.21	47.88	34.64
2	59.16	35.49	185.73	152.55	59.03	52.21	71.14	72.62	92.97	53.74	57.82	37.14	54.47	29.33
4 (Seasonally)	68.43	34.28	184.36	93.2	63.51	35.9	80.64	57.17	92.86	45.08	60.55	31.25	63.17	29.27
12 (Monthly)	41.7	13.74	115.31	51.01	40.66	15.52	54.67	26.51	47.04	17.59	34.45	12.4	36.99	11.09
26 (Biweekly)	19.32	7.56	87.91	36.93	17.57	9.29	29.1	12.85	24.01	11.46	15.81	8.22	16.31	7.33
52 (Weekly)	13.76	4.27	57.42	16.26	13.68	4.95	14.66	9.89	15.03	6.57	12.13	4.15	12.05	4.03
73	10.59	3.63	40.68	18.49	10.88	3.63	13.31	7.06	12.34	4.98	9.2	3.59	9.38	3.36
122	8.49	1.27	28.69	12.48	8.47	1.38	9.96	3.96	9.91	2.24	6.8	1.44	7.31	1.23
183	6.74	1.17	19.7	12.06	5.96	1.35	9.2	1.86	7.96	1.56	4.91	1.35	5.51	1.02
365 (Daily)	5.58	0.55	22.21	5.33	4.42	0.61	6.85	1.29	6.53	0.82	3.82	0.77	4.08	0.52
730	2.08	0.26	6.83	2.42	1.69	0.33	2.35	0.81	2.48	0.38	1.49	0.33	1.58	0.25
4,380	0.48	0.11	2.58	1.15	0.49	0.17	0.62	0.29	0.69	0.19	0.42	0.11	0.43	0.11
8,760 (Hourly)	0.09	0.03	0.6	0.23	0.11	0.03	0.13	0.06	0.15	0.04	0.1	0.03	0.1	0.03
17,520	0.05	0.02	0.27	0.12	0.05	0.02	0.07	0.01	0.08	0.02	0.05	0.01	0.05	0.01
35,040	0.08	0.03	0.71	0.19	0.08	0.03	0.11	0.04	0.14	0.04	0.08	0.03	0.08	0.03
52,560	0	0	0.08	0.02	0	0	0.01	0	0.01	0.01	0	0	0	0

parameters. These results are in agreement with previous studies (Cassidy & Jordan 2011; Jones *et al.* 2012; Miller *et al.* 2016).

The CV_AL values also varied with different water quality parameters. With only a few exceptions, the CV_AL values followed a consistent order of SS > Ortho-P ≈ K > NO_3-N ≈ Discharge > Mg ≈ Ca for all frequencies in the smaller sub-watershed. Although, at high frequencies (i.e., more than daily), the differences between different parameters decrease, such order still maintained. This pattern between different water quality parameters can be explained as follows. SS is mainly generated by soil erosion, which only occurs in intensive precipitation events. The durations of such events are usually short so that the calculated AL value was strongly affected by whether or not these intensive events were captured in the sampling. When the random time point selected for calculating the AL captured these events, the AL value will be overestimated, whereas when these events were missed, the AL will be underestimated. In contrast, Mg and Ca are mainly associated with baseflow and their concentration usually does not change significantly during intensive flow events and, therefore, are more stable. As a result, AL values calculated from different random selected time points did not markedly differ from each other. Phosphorus is well known for transporting with sediments. K is highly soluble but can also be absorbed by sediments. Therefore, K and Ortho-P were similar to SS. NO_3-N is abundant in the surface flow but can also leach into shallow subsurface flow and baseflow, so its trend was in between the two extremes and similar to the overall discharge.

For the larger watershed LRW, the CV_AL values followed a consistent order of SS > Ortho-P + K ≈ NO_3-N > Discharge ≈ Mg ≈ Ca for all frequencies. Although similar to that for BBW, there were some minor differences observed, which could have been due to differences in watershed size and land use. As watershed size increases, the share of surface flow versus that of shallow subsurface flow and baseflow decreases. Also, in BBW, agricultural fields dominated the landscape and soil erosion rates were much higher than that in LRW, where the forest was the predominant land use. A greater forested surface cover in LRW enhanced infiltration, hence reducing the share of surface flow relative to shallow subsurface flow and baseflow. As a result, discharge in LRW was more strongly affected by subsurface flow and baseflow, thus similar to Mg and Ca. The main source of soluble nutrients K and NO_3-N was the soil surface, so their transportation was determined by the surface flow and less affected by sediments or baseflow, and thus they were close to Ortho-P.

Coefficients of regression models of CV_AL with sampling frequency

The CV_AL values were significantly correlated with sampling frequency with $R^2 > 0.85$, and all the regression coefficients of b are close to 1 (Table 2). This implies that CV_AL values linearly increase with increasing sampling intervals (decreasing sampling frequencies).

It is also noted that the coefficient a for SS, Ortho-P, K, NO_3-N, Ca, and Mg in BBW were 9.73, 4.21, 4.74, 3.19, 2.90, 2.96, and 3.31, respectively, which were 1.64–3.26 times higher than the values for LRW (4.29, 2.56, 1.67, 1.56, 0.89, 1.09, and 0.91, respectively), indicating that the smaller watershed BBW had 1.64–3.26 times larger estimation errors in AL compared with that of the larger watershed LRW (Table 2). This is consistent with the conclusion drawn in the previous section based on the raw CV_AL data. The values of coefficient a for different parameters also follow the same orders of CV_AL values as summarized in the previous section. Overall, the regression analysis suggested that increasing sampling frequency could be more efficient in reducing uncertainties of AL estimation under the following conditions: (1) in a smaller watershed than in a large watershed and (2) for a sediment or surface flow-dominated water quality parameters than for a baseflow-dominated parameter.

Relationship between CV_AL of water quality parameters with the CV_AL of discharge

The CV_AL values of water quality parameters were significantly correlated with the CV_AL of discharge (Table 3) with $R^2 > 0.91$ and the regression coefficient k_2 close to 1. The regression coefficient k_1 for SS, Ortho-P, K, NO_3-N, Ca, and Mg in the BBW were consistently but only slightly lower than the values for LRW. This result indicated that the differences in CV_AL for the two watersheds were mainly related to variation in stream discharge. Lower k_1 values in the large watershed under the same sampling frequency versus with the small watershed can be attributed to less variation in stream discharge in the large watershed as well as the stronger impact of baseflow on water quality parameters in the large watershed. Among the water quality parameters, the accuracy of estimated CV_AL for SS and Ortho-P was more sensitive to the accuracy of stream

Table 2 | The mean of scaling factor (a) of Equation (8) and the ratio of scaling factor (a) of SS, NO_3-N, Ortho-P, K, Mg, and Ca with parameter (a) of discharge (F = sampling frequency #samples per year)

Equation	*a* for	SS	Ortho-P	K	NO_3-N	Ca	Mg	Discharge	*b*	R^2
$CV\_AL = aF^{-b}$	BBW	9.73	4.21	4.74	3.19	2.9	2.96	3.31	≈1	>0.85
F = Frequency	LRW	4.29	2.56	1.67	1.56	0.89	1.09	0.91	≈1	>0.89
Ratio of *a*	BBW/LRW	2.27	1.64	2.84	2.04	3.26	2.72	3.64		

Table 3 | The relationship between CV_AL of SS, NO_3-N, Ortho-P, K, Mg, and Ca with CV_AL of discharge in the BBW and LRW from the year 2003 to 2008 with different sampling frequencies

Equation	k_1 for	SS	Ortho-P	K	NO_3-N	Ca	Mg	k_2	R^2
$CV\_AL_{ion} = k_1X^{k2}$	BBW	3.07	1.23	1.37	0.94	0.87	0.87	≈1	>0.93
$X = CV\_AL(dis)$	LRW	4.01	2.34	1.67	1.55	0.89	1.08	≈1	>0.91
Ratio of k_1	BBW/LRW	0.77	0.53	0.82	0.61	0.98	0.81		

discharge compared to the other water quality parameters (i.e., K, NO_3-N, Ca, and Mg) as indicated by much greater k_1 values for SS and Ortho-P.

Relative bias of annual loading

As expected, RB_AL values were higher at low sampling frequency (seasonally to yearly) and decreased with increasing sampling frequency in both watersheds (Figure 4 and Table 4). When the sampling interval was shorter than hourly, the RB ranged from −0.04% to +0.04%. At this high sampling frequency, the magnitude of bias was very low, and the differences between watersheds as well as between water quality parameters were negligible. With daily sampling, the magnitude of RB_AL was less than 1% for all water quality parameters except for SS which had the RB_AL of 7.69% in BBW and 1.14% in the LRW. These results suggest that the systematic biases for AL calculation were marginally lower once the sampling interval was shorter than 1 day (i.e., daily sampling). This could explain why the RB of AL estimate received less attention in previous studies. However, the results also suggest that even with daily sampling, there could be substantial biases in the estimation of ALs for SS, especially for small watersheds such as BBW. When the sampling interval was longer than a week, RBs for the AL estimate could be substantial. For example, the RB_AL for SS was −11.68% and −9.87% in BBW and LRW, respectively. The RB_AL for NO_3-N, Ortho-P, K, Mg, and Ca in BBW was 1.62%, 4.47%, −3.88%, 3.05%, and 3.81%, respectively, and much higher than the corresponding values in LRW (0.17%, −0.99%, −2.7%, 0.91%, and 1.83%).

In addition, low sampling frequency (i.e., seasonally and yearly) tended to underestimate ALs for SS and water quality parameters associated with the surface flow (e.g., K and SS), and overestimate those associated with baseflow (e.g., Mg and Ca). To further examine this trend, RB_AL versus sampling frequency for SS (a typical surface runoff-related pollutant) and Ca (a typical baseflow-related pollutant) for both LRW and BBW are plotted in Figure 5(a). As shown,

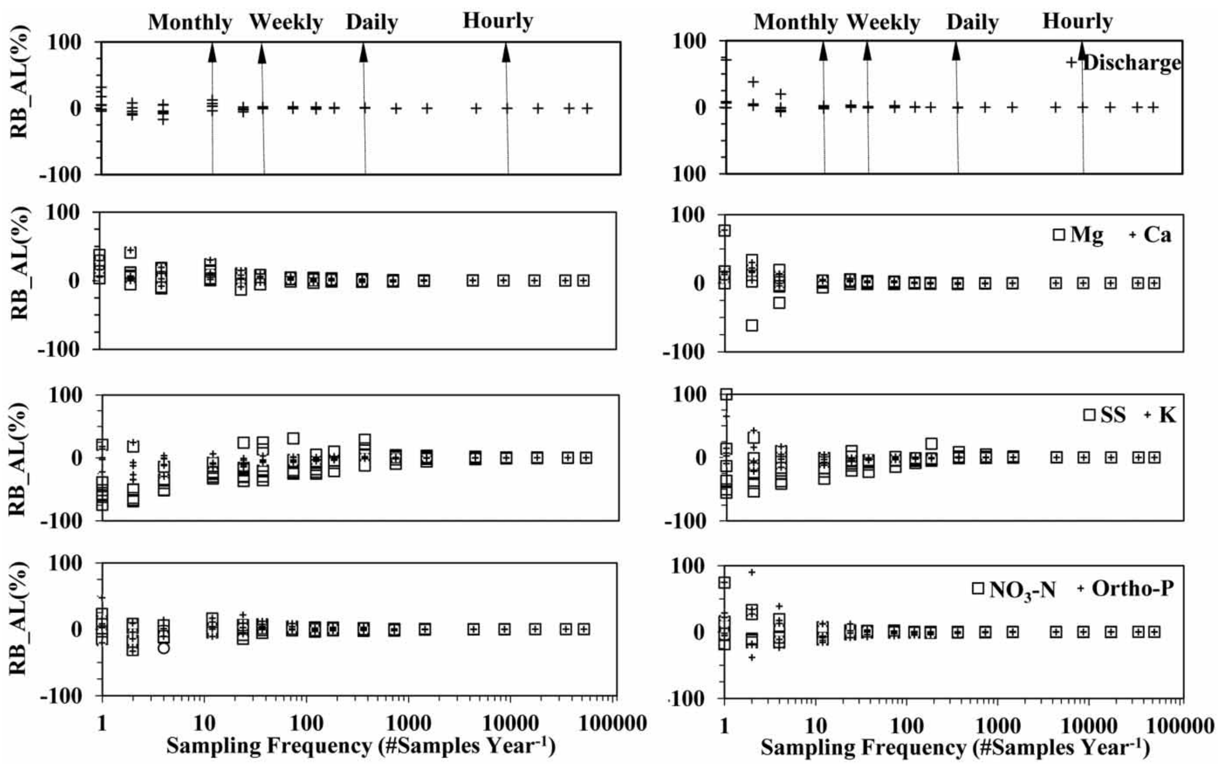

Figure 4 | RBs of estimated AL (RB_AL) for Discharge, Mg, Ca, SS, K, NO_3-N and Ortho-P from years 2003 to 2008 with different sampling frequencies in the BBW (left panel) and LRW (right panel).

Table 4 | The RB_AL of stream discharge (Disch), SS, NO_3-N, Ortho-P, K, Mg, and Ca under different sampling frequencies in BBW and in LRW from the year 2003 to 2008

	Disch (%)		SS (%)		NO_3-N (%)		Ortho-P (%)		K (%)		Mg (%)		Ca (%)	
Frequency	**BBW**	**LRW**	**BBW**	**LRW**	**BBW**	**LRW**	**BBW**	**LRW**	**BBW**	**LRW**	**BBW**	**LRW**	**BBW**	**LRW**
1 (Yearly)	8.86	15.56	−44.4	−7.21	7.86	12.18	9.73	20.54	21.9	12.26	22.38	20.68	19	23.2
2	−3.94	9.3	−47.26	−17.06	−8.87	1.69	−10.06	12.54	−15.31	0.31	11.62	0.91	9.77	17.49
4 (Seasonally)	−4.06	0.92	−38	−18.78	−8.91	−4.31	−0.62	2.96	−8.6	−0.52	5.59	−1.43	5.7	5.64
12 (Monthly)	3.04	0.04	−24.07	−13.84	1.13	−2.1	3.85	−3.22	−8.15	−4.59	9.34	1.07	10.52	3.73
26 (Biweekly)	−1.45	0.91	−17.76	−6.45	−2.15	−0.35	3.05	−0.62	−8.25	−2.78	2.37	2.32	3.39	3.73
52 (Weekly)	0.91	0.12	−11.68	−9.87	1.62	0.17	4.47	−0.99	−3.88	−2.7	3.05	0.91	3.81	1.83
73	0.38	0.49	−10.65	−5.77	0.88	0.36	2.07	−1.18	−3.09	−1.77	2.1	0.6	2.5	1.48
122	−0.15	0.09	−13.19	−4.21	0.19	−0.35	−0.35	−2.33	−2.5	−1.5	0.83	−0.06	1.06	0.37
183	0.33	0.04	−5.49	0.69	−0.04	−0.4	0.43	−1.6	−0.77	−1.14	0.46	−0.22	0.7	−0.06
365 (Daily)	0.75	−0.13	7.69	1.14	−0.55	−0.32	0.84	−0.75	0.55	−0.32	−0.32	−0.49	−0.19	−0.36
730	−0.12	−0.03	−0.44	0.68	−0.63	−0.14	−0.48	−0.23	−0.13	−0.14	−0.47	−0.24	−0.44	−0.2
4,380	0.01	0	−0.56	0.25	−0.23	−0.06	−0.21	−0.05	−0.06	−0.03	−0.2	−0.11	−0.17	−0.1
8,760 (Hourly)	0	0	0.03	0.08	−0.09	−0.01	−0.08	0	−0.02	0	−0.08	−0.04	−0.07	−0.03
17,520	0	0	−0.01	0.04	−0.04	0	−0.03	0	0	0	−0.04	−0.02	−0.04	−0.01
35,040	0	0	0	0.02	−0.02	0	−0.02	0	0	0	−0.02	−0.01	−0.02	−0.01
52,560	0	0	0.02	0.01	−0.01	0	−0.01	0	0	0	−0.01	0	−0.01	0
1 (Yearly)	0	0	−0.02	0	0	0	0	0	0	0	0	0	0	0

RB_AL of SS for both watersheds was negative when sampling frequency was low (<52 or weekly), and RB_AL for Ca was all positive at the corresponding sampling frequencies (Figure 5(a)). We also plotted the RB_AL of Ca and SS with the RB_AL of stream discharge for both watersheds in Figure 5(b). The biases of SS and Ca had no systematic relation with the bias of stream discharge, which were different from those of CV_AL. RB_AL of Ca showed a similar trend as Mg, and K showed a similar trend as SS, while NO_3-N and Ortho-P did not demonstrate a clear directional bias pattern (Table 4).

Based on our observation, the high concentration of SS only occurs in a very short time during major storm events in small watersheds. As such, there is a great chance to miss the periods of high SS concentrations when the sampling frequency is low. This could explain the observed underestimation of AL for SS. Due to the easy adsorption of K to soil particles (Chapman 1996), the underestimation of SS could have led to the underestimation of K at lower sampling frequency as well. In contrast, Mg and Ca are more closely associated with baseflow and tend to have relatively lower concentrations when there are high flows (diluted with higher discharge). Under-sampling during a higher flow period with lower Mg and Ca concentrations and higher discharge rates could cause overestimation of the concentrations of these ions (Lehmann & Schroth 2003). NO_3-N and Othro-P did not show clear patterns, which could be attributed to the interactions between sampling frequency and the timing of fertilizer applications in the watershed, particularly for the BBW.

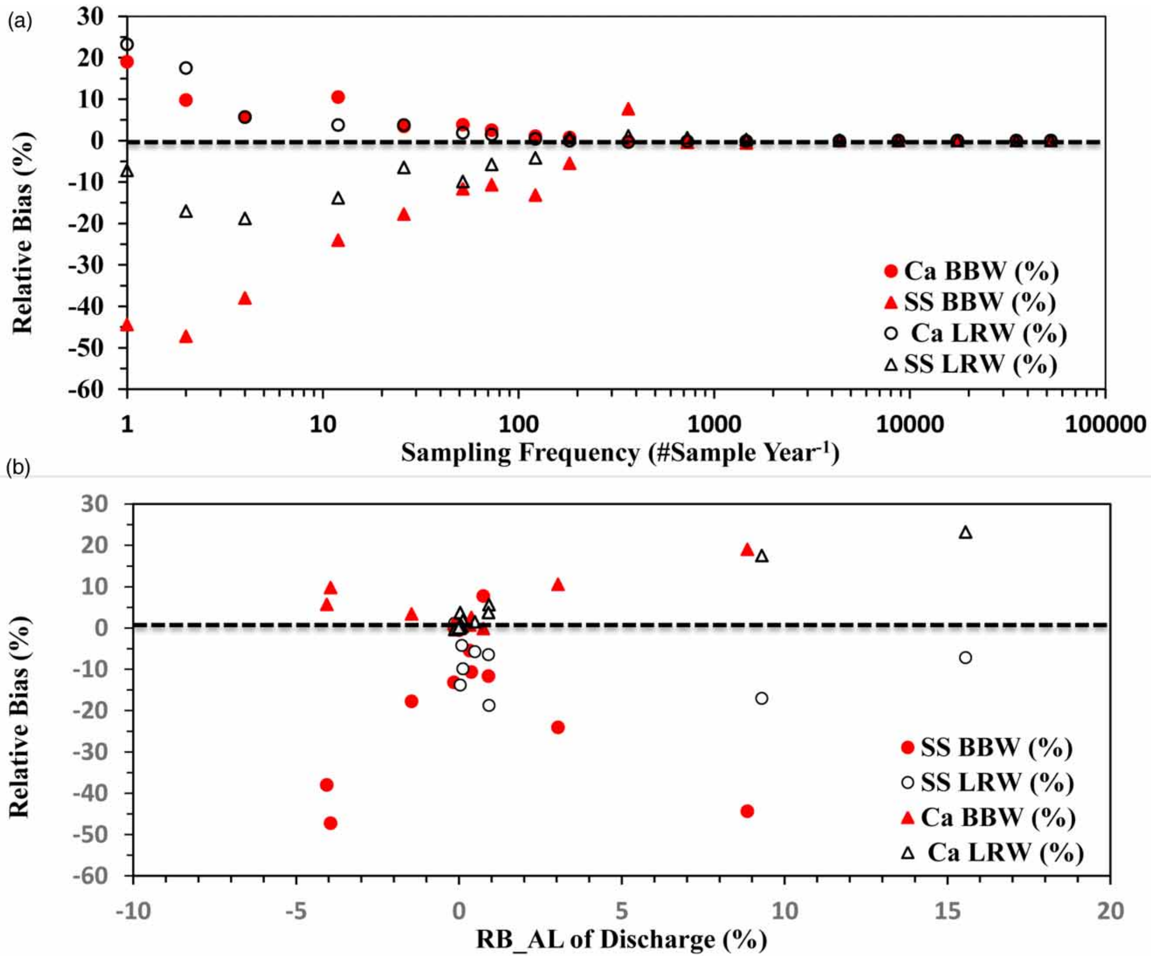

Figure 5 | RBs of estimated AL (RB_AL) for Ca, SS with different sampling frequencies (a) and RB_AL of stream discharge (b) in both BBW and LRW.

Probability of potential error in annual loading

The PPE_AL decreased with increasing sampling frequency following a reversed S curve: decreasing slowly when sampling frequency was less than 12 samples per year (monthly), decreasing rapidly until sampling frequency was higher than 1,000–8,000 samples per year (hourly) and then stabilized afterwards (Figure 6; Tables 5 and 6).

For different threshold values, the corresponding curves systematically shifted downward to the left from Cr = 1% to Cr = 10% (Figure 6). These values can be used to assess the estimation accuracy of AL. For example, for monthly sampling of SS, the PPE_AL for BBW was 99.3% for Cr = 1% and 94% for Cr = 10% and for LRW was 98.3% for Cr = 1% and 84.7% for Cr = 10% (Table 6). This implies that with monthly sampling, there will be a 99.3% and 94% chance for the estimation error of annual SS loading to be greater than 1% and 10%, respectively, for BBW. For LRW, there will be a 98.3% and 84.7% chance for annual SS loading estimation errors to be greater than 1% and 10%, respectively. Even with daily sampling, the probability of relative errors greater than 10% in the estimation of AL of SS was still as high as 13.3% for LRW (Table 6) and 57.3% for BBW (Table 5).

These results clearly indicate that to achieve the same estimation accuracy, smaller watersheds would require more frequent sampling than larger watersheds and that the required sampling frequencies are different for different water quality parameters. For example, to achieve less than 10% error (PPE_AL with Cr = 10%) for all assessed water quality parameters except for SS, 730 samples per year would be required for the BBW (Table 5) and 184 samples per year for the LRW (Table 6). These sampling frequencies are equivalent to a sampling interval of 12 and 48 h for BBW and LRW, respectively. Similarly, to achieve less than 1% error (PPE_AL with Cr = 1%) for all water quality parameters except for SS, sampling frequencies requirement would be 1,460 and 4,380 samples per year for BBW and LRW, respectively. With the same sampling frequencies, the SS could only achieve less than 5% errors. In other words, for the estimation of AL for SS in BBW, sampling frequency needs to be higher than 4,380 samples per year (one sample every 2 h) in order to achieve a relative error less than 5%. Following the same

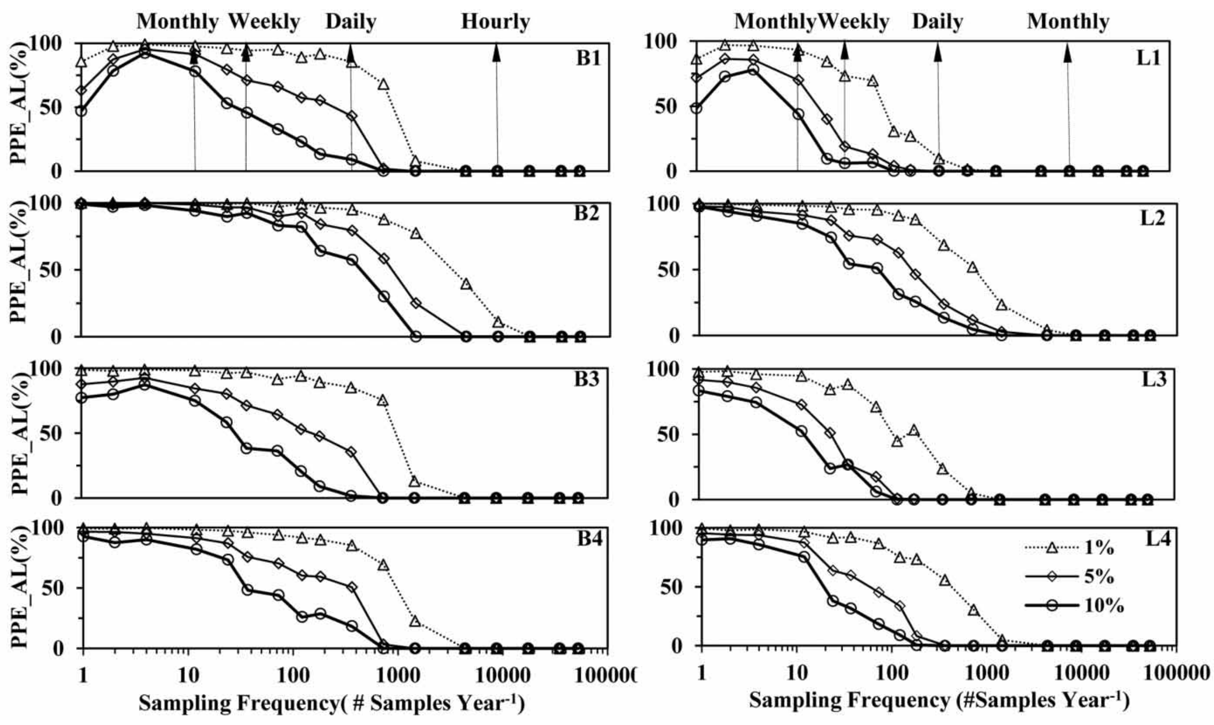

Figure 6 | The PPE of estimated AL (PPE_AL) of Stream Discharge (panels B1& L1), SS (panels B2 & L2), NO_3-N (panels B3 & L3) and Ortho-P (panels B4 & L4) in BBW (panels B1–B4) and in LRW (panels L1–L4) from the year 2003 to 2008 with different sampling frequencies.

Table 5 | The PPE_AL of stream discharge, SS, NO_3-N, Ortho-P, K, and Ca for the BBW from the year 2003 to 2008 with different sampling frequencies (# of samples per year)

	PPE_AL of discharge (%)			SS (%)			NO_3-N (%)			Ortho-P (%)			K (%)			Mg (%)			Ca (%)		
Frequency	Cr = 1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
1 (Yearly)	85.7	63	47	100	100	99	98.7	87.7	77.3	99.3	96.7	92.7	100	97	91.7	96.7	83	70	96.7	87	79
2	97.7	87.3	78.3	100	98.7	97	98	89.7	80	99	96.3	87.7	97.7	93.3	87.3	99	94	83	98.3	93.3	85.7
4 (Seasonally)	99	95.3	92.3	100	100	98.3	99	92.7	87.3	99.7	95	90	99	96.3	91.3	99.7	92.3	87	99	92.3	87.7
12 (Monthly)	97.7	91.3	78	99.3	98.7	94	98.3	84.3	75	98.3	91.3	82	99	93	86	95	83	68.3	98	86	68.7
26 (Biweekly)	96	79.3	53	99	96.7	89.7	96.3	80.3	58.3	97.3	87	73.3	96.3	85.3	73	92.7	70	45	93.7	68.3	42
52 (Weekly)	94.3	71	45.7	99.7	96.7	92.3	97	71.3	38.3	96	75.7	48.3	97.3	83.3	58.7	94.3	68.3	35.7	93	65	34.3
73	95	66	32.7	97	90	83	91.7	64.3	36.3	94	70.3	44	96	75.3	45	91.7	52.7	27	93.7	57.7	28
122	89	57.3	23	99.3	92.3	82	94.3	53	20.7	91.7	60.3	26	93.3	66.7	36.3	88.3	45.3	16	87.3	51	18
183	91.7	55.3	13.3	96.3	84	64	89.3	47.7	9	90	59.3	28.7	89.7	55	29.3	89	37.7	5.3	91.3	38	7.7
365 (Daily)	85.3	43	9	95	79.3	57.3	85.3	35.7	1.7	85.3	50.7	18.3	87.3	49.3	19.3	83	25	1	84.3	29.7	2.7
730	68	2	0	87.7	58.3	30	75.7	0	0	69.3	3.3	0	79.7	3.3	0.3	58.7	1.3	0	57.7	1.3	0
1,460	8	0	0	77.3	25	0	13	0	0	22.7	0	0	16.7	0	0	8	0	0	11.7	0	0
4,380	0	0	0	39.6	0	0	0	0	0	0	0	0	1.4	0	0	0	0	0	0	0	0
8,760 (Hourly)	0	0	0	11.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17,520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35,040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52,560	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 6 | The PPE_AL of stream discharge, SS, NO_3-N, Ortho-P, K, and Ca for the LRW from the year 2003 to 2008 with different sampling frequencies (# of samples per year)

	PPE_AL of discharge (%)			SS (%)			NO_3-N (%)			Ortho-P (%)			K (%)			Mg (%)			Ca (%)		
Frequency	Cr = 1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
1 (Yearly)	86.3	71.7	48.3	100	98	97.7	97.7	91.7	83.3	99.3	95.3	89.7	98.3	92.3	87	98	82.3	70	96.7	90.3	78
2	97	86.3	72.7	99.3	97.3	94	98.3	90	79	97.7	94	90.7	97.7	95.7	85.3	97.7	90.7	81	98.7	91.7	77.7
4 (Seasonally)	96.7	85.7	78	99	94	90.7	96	85.7	74.3	98.7	93.7	85.7	99	93.7	84.7	96.7	85.3	70.7	96.3	80	62.7
12 (Monthly)	93.3	70	43.7	98.3	91.3	84.7	94.7	72.7	52.3	96.7	87.3	75.3	98.3	87.3	58	94.7	71	42.3	89.7	65	34
26 (Biweekly)	84.3	40	9.3	97.7	87.3	74.3	84.3	51	23.7	91.3	63.7	38	94.7	71	34.7	87.7	48	14.3	84.7	37	14
52 (Weekly)	73.3	19	6	95.7	75.7	54.3	88.3	26.7	26.7	92	59.7	31.7	92.3	50.3	15.7	75.7	18	5.3	73	18.3	5.7
73	69.7	13	6.7	95.3	72.7	51	71	17.3	6	86.7	45.3	18.3	82	30.7	11	72.7	20.7	4.3	69.3	18.3	4.3
122	30.7	4	0	91	62.7	31.3	44.7	0.7	0	75	33.7	9	77.3	12.7	0.3	51	0.3	0	48.3	0.3	0
183	27	1	0	88	46.3	25.7	53.3	0.3	0	73.3	8.3	0.7	67	5.3	0	45.3	0	0	30.7	1	0
365 (Daily)	9.7	0	0	68.7	23.7	13.3	23.7	0	0	55.7	0.3	0	32	0	0	37.3	0	0	23.7	0	0
730	1.3	0	0	52	11.7	4.7	5	0	0	30.7	0	0	4.7	0	0	7.3	0	0	7.7	0	0
1,460	0	0	0	23.3	2.7	0	0	0	0	5	0	0	2.3	0	0	0	0	0	0	0	0
4,380	0	0	0	4.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8,760 (Hourly)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17,520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35,040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52,560	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

numerical methodology, sub-hourly samples would be needed in order to achieve a relative error of less than 1% for the estimation of AL for SS in BBW. For the estimation of AL of SS in the larger watershed LRW, sub-daily samples are required to achieve relative error less than 5%, and hourly samples would be required to achieve a relative error of less than 1%.

CONCLUSION

In this study, we evaluated relationships between sampling frequency and estimation errors of AL for water quality parameters (SS, NO_3-N, Ortho-P, K, Ca, and Mg) in two watersheds in Northern New Brunswick, Canada, a large forest dominated watershed and one of its sub-watersheds, a small agricultural field dominated watershed. Estimation errors of AL were measured by three statistical indicators: coefficient of variance (CV_AL), RB (RB_AL), and PPE (PPE_AL). Based on the results, we have drawn the following conclusions:

1. The magnitudes of the two measures for AL estimation errors, CV_AL and RB_AL, decreased with sampling frequency for both watersheds, LRW and BBW. At the same sampling frequency, estimation errors were higher for water quality parameters associated with surface runoff (i.e., SS, K, NO_3-N, and Ortho-P), compared with those associated with baseflow (i.e., Mg and Ca).
2. At the same sampling frequency, the smaller watershed (BBW) had estimation errors in AL several times larger than the larger watershed LRW, suggesting that relatively smaller watersheds are more strongly affected by the event-based flow, leading to higher AL estimation errors compared to larger watersheds.
3. Increasing sampling frequency in smaller watersheds is more effective in reducing AL variations than in large watersheds. The differences in CV_AL between the two watersheds were mainly related to the variation of stream discharge, as the larger watershed had a lower CV_AL under the same sampling frequency because it had lower variation in stream discharge.
4. Low sampling frequency tended to systematically underestimate the AL of water quality parameters that are associated with the surface flow (e.g., SS and K) and overestimate those associated with baseflow (e.g., Mg and Ca). These systematic biases were not related to the estimation error of stream discharge.
5. The PPE_AL is useful in determining the sampling frequency requirements for given water quality parameters and could potentially be used as a tool to determine the trade-off between achieving expected accuracy and efforts needed for sampling. We found that SS required higher sampling frequency compared with other water quality parameters to meet the same accuracy standards for both large and small watersheds. For water quality parameters other than SS, sub-hour sampling intervals would be required in order to achieve a RB of less than 1%, and sub-daily samples would be required to achieve a RB less than 5%.

ACKNOWLEDGEMENTS

Funds for this study were provided through Agriculture and Agri-Food Canada (AAFC) A-base projects ‘Reducing sediment, N and P loading from arable cropping systems to receiving waters in eastern Canada (PEI, NB, NS, QC)’ and ‘Landscape Integrated Soil and Water Conservation (LISWC) on sloping fields under potato production in Atlantic Canada’ as well as through a Canadian Natural Science and Engineering Research Council Discovery Grant.

REFERENCES

Bahar, M. M. & Yamamuro, M. 2008 Assessing the influence of watershed land use patterns on the major ion chemistry of river waters in the Shimousa Upland, Japan. *Chemistry and Ecology* **24** (5), 341–355.

Blaen, P. J., Khamis, K., Lloyd, C. E. M., Bradley, C., Hannah, D. & Krause, S. 2016 Realtime monitoring of nutrients and dissolved organic matter in rivers: capturing event dynamics, technological opportunities and future directions. *Science of the Total Environment* **569–570**, 647–660.

Carstea, E. M., Bridgeman, J., Baker, A. & Reynolds, D. M. 2016 Fluorescence spectroscopy for wastewater monitoring: a review. *Water Research* **95**, 205–219.

Cartwright, N., Clark, L. & Bird, P. 1991 The impact of agricultural on water quality. *Outlook Agricultural* **20** (3), 145–152.

Cassidy, R. & Jordan, P. 2011 Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. *Journal of Hydrology* **405** (1–2), 182–193.

Chapman, D. 1996 *Water Quality Assessments – A Guide to Use of Biota, Sediments and Water in Environmental Monitoring*, Vol. 2. University Press, Cambridge, p. 609.

Chen, J., He, D. & Cui, S. 2003 The response of river water quality and quantity to the development of irrigated agricultural in the last 4 decades in the Yellow River Basin, China. *Water Resources Research* **39** (3), 1–11.

Chow, T. L. & Rees, H. W. 2006 *Impacts of Intensive Potato Production on Water Yield and Sediment Load -Black Brook Experimental Watershed: 1992–2002 Summary*. Potato Research Centre, AAFC, Fredericton, p. 26.

Chow, L., Xing, Z., Benoy, G., Rees, H. W., Meng, F., Jiang, Y. & Daigle, J. L. 2010 Hydrology and water quality across gradients of agricultural intensity in the Little River Watershed area, New Brunswick, Canada. *Journal of Soil and Water Conservation* **66** (1), 71–84.

Elwan, A., Clark, M., Roygard, J., Singh, R., Horne, D. & Clothier, B. 2014 Effects of sampling frequency and calculation methods on estimation of annual nutrient loads: a case study of Manawatu River, New Zealand. *Improving Water Quality and the Environment* **11**. http://doi.org/10.13031/wtcw.2014-014.

Ferguson, R. I. 1986 River loads underestimated by rating curves. *Water Resources Research* **22** (1), 74–76.

Halliday, S. J., Skeffington, R. A., Wade, A. J., Neal, C., Reynolds, N., Norris, D. & Kirchner, J. W. 2013 Upland streamwater nitrate dynamics across decadal to sub-daily timescales: a case study of Plynlimon, Wales. *Biogeosciences* **10**, 8013–8038.

Johnes, P. J. 2007 Uncertainties in annual riverine phosphorus load estimation: impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology* **332** (1–2), 241–258.

Jones, A. S., Horsburgh, J. S., Mesner, N. O., Ryel, R. J. & Stevens, D. K. 2012 Influence of sampling frequency on estimation of annual total phosphorus and total suspended solids loads. *Journal of the American Water Resources Association* **48** (6), 1258–1275.

Kirchner, J. W., Neal, C. & Robson, A. J. 2004 The fine structure of water-quality dynamics: the (high-frequency) wave of the future. *Hydrological Process* **18**, 1353–1359.

Kozak, M., Bocianowski, J. & RybiNski, W. 2013 Note on the use of coefficient of variation for data from agricultural factorial experiments. *Bulgarian Journal of Agricultural Science* **19**, 644–646.

Lehmann, J. & Schroth, G. 2003 Nutrient leaching. In: *Trees, Crops, and Soil Fertility*, (G. Schroth & F. L. Sinclair, eds) CABI Publishing, Wallingford, UK, pp. 151–166.

Li, Q., Xing, Z., Danielescu, S., Li, S., Jiang, Y. & Meng, F.-R. 2014 Data requirements for using combined conductivity mass balance and recursive digital filter method to estimate groundwater recharge in a small watershed, New Brunswick, Canada. *Journal of Hydrology* **511**, 658–664.

Malve, O., Tattari, S., Riihimäki, J., Jaakkola, E., Vo, A., Williams, R. & Bärlund, I. 2012 Estimation of diffuse pollution loads in Europe for continental scale modelling of loads and in-stream river water quality. *Hydrological Process* **2394**, 2385–2394.

Maticic, B. 1999 The impact of agricultural on ground water quality in Slovenia: standards and strategy. *Agricultural Water Management* **40** (2–3), 235–247.

Miller, M. P., Tesoriero, A. J., Capel, P. D., Pellerin, B. a., Hyer, K. E., Burns, D. & Al, M. E. T. 2016 Quantifying watershed-scale groundwater loading and in-stream fate of nitrate using high-frequency water quality data. *Water Resources Bulletin* **52**, 330–347.

Qi, J., Li, S., Jamieson, R., Hebb, D., Xing, Z. & Meng, F.-R. 2017a Modifying SWAT with an energy balance module to simulate snowmelt for maritime regions. *Environmental Modelling & Software* **93**, 146–160.

Qi, J., Li, S., Yang, Q., Xing, Z. & Meng, F.-R. 2017b SWAT setup with long-term detailed landuse and management records and modification for a micro-watershed influenced by Freeze-Thaw cycles. *Water Resources Management* **31**, 3953–3974.

Quimpo, R. G. & Yang, J.-Y. 1970 Sampling considerations in stream discharge and temperature measurements. *Water Resources Research* **6** (6), 6–9.

Reynolds, K. N., Loecke, T. D., Burgin, A. J., Davis, C. A., Riveros-Iregui, D., Thomas, S. A., St Clair, M. A. & Ward, A. S. 2016 Optimizing sampling strategies for riverine nitrate using high-frequency data in agricultural watersheds. *Environmental Science & Technology* **50**, 6406–6414.

Richards, R. P. 1998 *Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs*. Project Report Prepared under Grant X, 998397,108.

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., Arhonditsis, G. B., Jordan, P., Kronvang, B., Halliday, S. J., Skeffington, R. A., Rozemeijer, J. C., Aubert, A. H., Rinke, K. & Jomaa, S. 2016 Sensors in the stream: the high-frequency wave of the present. *Environmental Science and Technology* **50** (19), 10297–10307. ISSN 1520-5851.

Rozemeijer, J. C., van der Velde, Y., van Geer, F. C., de Rooij, G. H., Torfs, P. J. J. F. & Broers, H. P. 2010 Improved load estimates for NO_3 and P in surface waters by characterizing the concentration response to rainfall events. *Environmental Science & Technology* **44**, 6305–6312.

Rus, D. L., Patton, C. J., Mueller, D. K. & Crawford, C. G. 2012 *Assessing Total Nitrogen in Surface-Water Samples – Precision and Bias of Analytical and Computational Methods*. Scientific Investigations Report 2012-5281. U.S. Geological Survey, Reston, Virginia, p. 48.

Schroder, J. J., Scholefield, D., Cabral, F. & Hofman, G. 2004 The effects of nutrient losses from agricultural on ground and surface water quality: the position of science in developing indicators for regulation. *Environmental Science & Policy* **7** (1), 15–23.

Sokal, R. R. & Rohlf, F. J. 1995 *Biometry: the Principles and Practice of Statistics in Biological Research*. Freeman, New York.

Wade, A. J., Palmer-Felgate, E. J., Halliday, S. J., Skeffington, R. A., Loewenthal, M., Jarvie, H. P., Bowes, M. J., Greenway, G. M., Haswell, S. J., Bell, I. M., Joly, E., Fallatah, A., Neal, C., Williams,

R. J., Gozzard, E. & Newman, J. R. 2012 Hydrochemical processes in lowland rivers: insights from *in situ*, high-resolution monitoring. *Hydrology and Earth System Sciences* **16**, 4323–4342.
Walther, B. A. & Moore, J. L. 2005 The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography* **28**, 815–829.
Xing, Z., Chow, L., Rees, H., Meng, F., Li, S., Ernst, B., Benoy, G. & Hewitt, L. M. 2013 Influences of sampling methodologies on pesticide-residue detection in stream water. *Archives of Environmental Contamination and Toxicology* **64** (2), 208–218.
Yang, Q., Meng, F. R., Zhao, Z., Chow, T. L., Benoy, G., Rees, H. W. & Bourque, C. P. A. 2009 Assessing the impacts of flow diversion terraces on stream water and sediment yields at a watershed level using SWAT model. *Agriculture, Ecosystems and Environment* **132** (1–2), 23–31. http://doi.org/10.1016/j.agee.2009.02.012.

First received 31 July 2019; accepted in revised form 6 January 2020. Available online 2 April 2020