Assessment of sample frequency bias and precision in fluvial flux calculations – an improved low bias estimation method

F. Worrall[1][1], N.J.K.Howden[2], & T.P.Burt[3]

1. Dept of Earth Sciences, Science Laboratories, South Road, Durham, DH1 3LE, UK. Fax: 0191 374 2510, Tel: 0191 374 2535, Email: fred.worrall@durham.ac.uk.

2. Dept. of Civil Engineering, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR., UK.

3. Dept. of Geography, Science Laboratories, South Road, Durham, DH1 3LE, UK

Abstract

Despite the importance of calculating the flux of solutes and particulates through the global fluvial network the number of studies that have considered the bias and precision of any method is limited. Furthermore, no study has, on the basis of the bias of the method, proposed new methods with a lower bias nor considered the implications of the bias estimation for existing published studies. Using 3 years of high frequency data (hourly) for dissolved organic carbon (DOC) this study systematically degraded the data and recalculated the flux for varying sample frequencies and considered a range of interpolation, ratio and extrapolation methods. The results show that:

i) Interpolation and ratio methods showed a consistent, small bias for sampling frequencies up to every 14 days, but bias rapidly increased for lower sample frequencies with the flux estimates being between 40 and 45% of the "true" flux at 31 day (monthly) sampling.

ii) The best ratio method was based upon correction against an unrealistic assumption that river flow was normally distributed.

[1] Corresponding author

iii) Extrapolation methods based on fixed sampling period monitoring proved to be erratic but no better than interpolation methods.

Based upon the nature of the sources of variation within the flow and solute datasets we propose the following method for calculating the fluvial flux (F) of a solute:

$$F = KE(C_i)Q_{total}$$

Where: Q_{total} = the total flow in a year (m^3/yr); $E(C_i)$ = the expected value of the sampled concentrations (mg/l); and K= a conversion factor. This new method preserved all the available flow information and had a bias of as low as 8% for monthly sampling. When the method was applied to DOC flux from Great Britain bias correction meant a 97% increase in the national flux over previous estimates.

Keywords: dissolved organic carbon; DOC; DOC budget

Introduction

There are many methods that have been proposed for the calculation of dissolved or particulate fluxes from rivers. Methods used to estimate fluvial fluxes of a given component can be broadly classified into either interpolation methods (e.g. Webb et al., 1997) or extrapolation methods (e.g. De Vries and Klavers, 1994). Some studies (OSPAR, 1998) also consider ratio methods alongside interpolation and extrapolation methods, however, ratio methods are usually based upon a correction factor being applied to an interpolation method. A summary of current methods is given in Table 1.

Given the importance of flux estimates for studies of solute and solid material transport through the fluvial network, many studies have proposed flux calculation methods and some studies have also assessed the accuracy and precision of different approaches for different determinands and over different sampling regimes. A number of these have been used with correction factors in order to remove sampling or calculation bias (e.g. Ferguson, 1986). These flux correction factors have been considered from a number of perspectives. Cohn (2005) considered the problem of solute flux estimation when measured concentrations are close to detection limits but most studies have considered the differences between sampling frequencies and methods. The quality of methods and sampling frequencies need to be discussed in two ways. Firstly, the accuracy can be considered as the difference between the true load and estimated load and represents the systematic bias. Secondly, the precision of the method or sampling frequency which represents the spread of the load estimates about a certain value and represents the consistency of the load estimates. In many studies that discuss uncertainty in flux estimation due to changing method or sampling frequency it is the precision that is described and not the bias or accuracy. An example of this is, Webb et al. (1997) who considered 5 interpolation and 2 extrapolation methods and found that for suspended sediment flux estimation extrapolation methods gave the least biased results and that bias increased with decreased sample frequency, but they could give only precision estimates for solute fluxes because there was no "true" value for each determinand against which to compare the bias. Similarly, Littlewood et al., (1998) could only trace precision with changing sampling frequency with "indicative" curves. Littlewood (1995) and Webb et al. (2000) generated synthetic concentration time series, tested a number of flux estimation methods and suggested that interpolation methods were generally more reliable and less prone to errors than the more complex extrapolation methods. They suggested that extrapolation methods work best where a good rating curve between concentration and flow can be found, but this is not normally the case for determinands that exhibit a strong seasonal component. However, neither study could comment on bias and both studies relied on the error structure that was in the synthetic time series, i.e. if the synthetic time series did not include all the components of the variation in the data then the error estimation in the flux calculation would have itself been biased. Johnes (2007) considered 17 catchments where there was daily measurement of phosphorus but had no sub-daily data and had to assume that "method 5" (Littlewood, 1995) was the true value and only considered precision but not bias. Skarbøvik et al. (2012) considered a record of daily of suspended sediment and although they were able to suggest that the least biased were extrapolation methods compared to interpolation methods this result was not true across all sampling frequencies and years. Burt et al. (2011) using daily flow and nitrate concentration data did consider both bias and precision to show that seven-day sampling gave a high level of precision with 95% of flux estimates within 2.7% of the "true" flux (based on a complete set of daily samples). Even 28-day ("monthly") sampling had 95% of values within 6.5% of the mean: a surprisingly high level of precision given the infrequency of the sampling. In terms of accuracy, they showed that the mean flux derived from 7-day sampling was very close to the "true" flux, but the 14-day and 28-day means underestimated the "true" load. Kulasova et al. (2012) did have sub-daily measurement of nitrate and total phosphorus and although they discuss precision relative to extrapolation methods they did not consider bias of the methods they used. Cassidy and Jordan (2011), with sub-daily measurement of phosphorus, considered both bias and precision in their approach and thus showed bias with decreasing sampling frequency with bias of up to 60% upon weekly sampling (ie. 60% lower than the true value) and high uncertainty for all sampling frequencies except for near continuous monitoring.

Several studies have recommended or considered adaptive strategies. Kronvang and Bruhn (1996) suggested taking samples hydrologically rather than on a calendar basis and a number of studies (Cooper and Watts, 2002; Skarbøvik et al., 2012) have suggested including flood samples alongside regular sampling. Without a "true" load estimate the authors of studies of adaptive strategies cannot comment on their bias or precision, and indeed, Cassidy and Jordan (2011) found an over-estimation, or positive bias, when flood samples were included in flux estimation. Roberts (1997) considered the sampling frequency within the UK's Harmonised Monitoring Scheme (Simpson, 1980) and compared calculation of annual fluvial phosphate flux based on 52 weekly samples, 4 weeks of daily samples and 4 days of hourly samples, but concluded that for pragmatic and financial reasons sampling was degraded to monthly. Flux estimation methods based upon adaptive strategies will not be appropriate for many monitoring programmes designed to estimate flux simply because most national monitoring programmes are based upon regular sampling. Moater et al. (2012) considered the precision and bias of differing sampling frequencies given daily sampling at 125 sites, the study did not consider differing methods of flux estimation although it did propose an empirical approach to flux estimation based on measures of the flow duration curve.

Although the studies above have considered particulates and a range of solutes none of the studies above have considered the precision and bias of flux estimation methods relative to dissolved organic carbon (DOC). Recent interest in DOC has been based upon the observation that DOC concentrations in the surface waters across the northern hemisphere have been increasing over the last few decades (eg. Monteith et al., 2007). Studies of flux have been used to consider the causes of this rise (eg. Worrall et al., 2008) but the flux of DOC is also a vital component of the carbon cycle of the terrestrial biosphere (eg. Worrall et al., 2009a) and changes in the DOC flux across a watershed is component in the estimation of the flux of greenhouse gases to the atmosphere (Worrall et al., 2012). The aims of this study, therefore, are multiple. Firstly, to assess the bias and precision of flux calculation methods relative to DOC: a solute not previously considered. Secondly, to not only provide a measure of bias but to assess how the change in the bias of the method with changing sampling frequency could be used to direct the method of flux estimation.

Approach & Methodology

This study was able to consider sub-daily monitoring of the DOC concentrations and river flow over a 3 year period. From hourly data a "true" value of load was calculated for each of the available years. The time series of DOC concentration and river flow were then systematically degraded so that combinations of data were selected based upon a single sample being collected each day with frequency of sampling from 1 day to 31 days, i.e. from daily to monthly. Within each day, pairs of data were selected at random and 100 sets of flux calculations for each of the three years were made for each sampling frequency for each year and for each flux estimation method. Results from each sampling frequency were compared to the "true" load flux for each year for each sampling frequency for each flux estimation method in order to give an estimate of the bias, and precision, with increasing sample frequency.

Study site

Data were collected for the River Dee just upstream of the city of Chester where data could be twinned with flow records (Figure 1). The river Dee at Chester has a catchment area of 1674 km^2, with annual average rainfall (1961 – 1990) of 1143 mm with 10% of the catchment being classed as mountain, heath and bog which can be considered as the major source of the DOC considered in this study (National Riverflow Archive – http://www.ceh.ac.uk/data/nrfa/). The concentration data were collected hourly between 1st January 2009 and 31st December 2011 and the flow data every 15 minutes over the same period. Over the 3 year period the median river discharge = 15.5 m^3/s with 95% exceedence flow = 5.8 m^3/s and 5% exceedence flow = 101.4 m^3/s. For the DOC concentration over the 3 year period the median concentration = 11.2 mg C/l with 95th percentile = 21.3 mg C/l and the 5th percentile = 4.6 mg C/l. The DOC concentration data were collected using an UV absorbance probe (ABB AV400) calibrated for DOC concentration using potassium hydrogen phthalate on a regular basis. The calibration between UV absorbance and DOC concentration is not necessarily stationary and would be a source of uncertainty above and beyond that due to flux estimation method, however further discussion is beyond the scope of this study as this study concerns the error in the flux estimation method and not the error in the measurements. The error in the relationship between UV absorbance and DOC used within the measuring probe would only be of concern to this study if it had a systematic bias that was true whenever, and wherever, it was used for DOC measurement and in the UK DOC concentration is not normally made using this approach (Simpson, 1980).

Flux estimation methods

The study did not choose to compare all possible interpolation and extrapolation methods rather selected methods for contrast. We choose "method 2" (Littlewood et al., 1998) which is a simple interpolation method:

$$F = K\sum_1^n \frac{C_i Q_i}{n} \qquad \text{(i)}$$

Where: C_i = the instantaneous concentration (mg/l); Q_i = the instantaneous discharge (m^3/s); n = number of samples; and K= a unit conversion factor.

More advanced interpolation methods are interpolation methods that attempt to correct for the conditions at the time of sampling – sometimes referred to as ratio methods = and the most used or recommended is "method 5" (eg. Johnes, 2007):

$$F = K\left(\sum_1^n C_i Q_i\right)\left(\frac{\bar{Q}}{\sum_1^n Q_i}\right) \qquad \text{(ii)}$$

Where: $\bar{Q}$= the average discharge for the period (m^3/s); n = the number of samples; and K = conversion factor. The interpolation "method 5" is "method 2" with the ratio factor correction, the correction factor compares the flows at times of sampling to the other measured flows. However, "method 5" assumes a normal distribution of flow data which would be unusual for most known rivers. Therefore the general case would be:

$$F = K\left(\sum_1^n C_i Q_i\right)\left(\frac{E(\bar{Q})}{E(Q_i)}\right) \qquad \text{(iii)}$$

Where: E(x) = expected value of x. The form of the expected value would differ depending on the type of distribution that best described the distribution of the sampled and total flows. For the purpose of this study normal; log normal and gamma distributions were considered. Note that for a normal distribution the expected value would be the arithmetic mean; for the log-normal distribution the expected value would be the geometric mean; and the gamma distribution the expected value – kθ, where: k = shape factor and θ = scale factor. Equation (iii) is "method 5" when a normal distribution is considered and henceforward this study will refer to "method 6" when E(x) is based upon a log normal distribution and "method 7" when E(x) is based upon a gamma distribution. The normal and log normal distributions were fitted to sampled and total measured flows by method of moments and the gamma distribution was fitted by maximum likelihood.

A log-log rating curve approach was used as an extrapolation method for calculating flux estimates:

$$logC_i = AlogQ_i + logB \qquad \text{(iv)}$$

Where A and B are constants. Equation (iv) was fitted to the data sampled for each sampling frequency and then applied to the entire flow record for each year of the available data.

A comparison of the concentration, flow and flux cumulative distributions (Figure 2) suggests that the flux distribution was really dominated by the flow and not the distribution of the concentration and therefore, methods that do not preserve the distribution of the flow or do not maximise the information available in situations where flow is more extensively sampled than the concentration will show considerable bias. For example, when calculating the DOC flux across the UK, Worrall and Burt (2007) used the data collected as by the Harmonised Monitoring Scheme (Simpson, 1980) and for most catchments for most years only monthly samples were available while in most cases daily flows were available. Therefore, this study proposes the following method:

$$F = KE(C_i)Q_{total} \qquad \text{(v)}$$

Where: Q_{total} = the total flow in a year (m^3/yr); $E(C_i)$ = the expected value of the sampled concentrations (mg/l). As above the form of the expected value depends upon the distribution chosen to describe the sampled concentrations and as above for the purpose of this study normal, log normal and gamma distributions were chosen. Equally, if the flux is dominated by variation in the flow and sample sizes in any one year are small then taking data from previous years sampling could improve the estimation of the concentration distribution. The latter would be true as long as the concentration time series was stationary or could be made stationary. For the purposes of this study it was assumed that the DOC time series was stationary for the 3 years of the study. So $E(C_i)$ was calculated not only for normal, log normal and gamma distributions but also for each of these based upon sampling from all three years of available data or from each year of available data. To assess the best combination this was considered as a factorial experimental of four factors and their interactions. The four factors were: the year the data were collected (with three levels - 2009, 2010 & 2011); the averaging method (with three levels - normal, log-normal and gamma distributions); assessment period (with two levels based upon averaging over 1 or 3 years of data); and the frequency of sampling (with levels from 1 to 31 days). By considering as a factorial experiment we can identify not only significant effects, by analysis of variance (ANOVA), but also the size of the effect and so identify which combination of approaches would lead to lowest possible bias.

Application of results

The findings of the above studies will be applied to consideration of DOC flux from the UK. The flux of DOC from the UK has been calculated from 1975 to 2007 by Worrall et al. (2009b). However, the approach was based upon "method 5" and performed only an analysis of precision but not of bias in the method. In Worrall et al. (2009b) the flux estimates were based upon data from the harmonised monitoring network (Simpson, 1980) and only sites where sampling was at least monthly in any one year were included, although sampling frequencies were as regular as sub-weekly for some sites for some years.

Results

The comparison between the "true" load and the results from interpolation methods shows a time course in the systematic bias of the results with overestimation on daily sampling with systematic underestimation occurring between day 9 and 14, but by day 31 the methods were between 40 and 45% of the "true" load (Figure 3) an underestimate of up to 60%. Of the interpolation methods, "method 5" was the least biased overall and was always the highest estimate of the load. The reason for the better performance of "method 5" is not because of inherently better method of estimation rather because it assumes a normal distribution for flow and so it systematically overestimates the expected value of the flow distribution (Figure 4). So by using a normal distribution the calculation fortuitously skews itself towards higher values and so preserves higher estimates of the annual flux. The imprecision of the interpolation methods, as expressed as the 5^{th} to 95^{th} interpercentile range, increases with decreasing sampling frequency and by a sampling frequency of 15 days the imprecision is between ±15 and 20% which is very similar to the precision of the interpolation methods estimated by Webb et al. (1997), for a sampling frequency of 31 days the imprecision reaches ±12% (Figure 5).

For the extrapolation method the change across the sampling frequencies shows that the pattern of bias was erratic but for monthly sampling the bias of an extrapolation method is no better than interpolation with median estimates between 19 and 55% of the "true" load (Figure 6). The precision of the extrapolation shows that for all the extrapolation methods considered that imprecision rose rapidly as sample frequency dropped to every 5 days after which imprecision varied between ±20 and 40% of the median (Figure 7). The erratic nature of the response of using equation (iv) for regularly monitored data must be ascribed to the sensitivity of the method to the inclusion of high flow events within the sample. The result would appear less erratic if the study had perhaps used 500 rather than 100 sets of samples, but such an increase would not have changed the result with respect to the bias of the method.

Estimates based upon equation (v) showed that in general the method provided a less bias result for sampling frequencies of greater than 14 days (Figure 8) and indeed the analysis of variance (ANOVA) from the factorial analysis of the results found no significant differences between sampling frequencies up to 14 days for all approaches based upon equation (v). It is obvious that one of the versions of equations (v) based upon the annual estimate of the expected value from fitting a gamma distribution to the sample data was very biased with estimates being 39% of the “true” load with sample frequencies less than 2 days (Figure 8). This large bias is perhaps for a very practical reason, i.e. the difficulty of estimating a gamma distribution with limited data. Conversely, the best overall method was equation (v) based upon the expected value of gamma distribution based on the data sampled from all years of the available data with flux estimates being 92% of the “true” load even at sampling frequencies as low as every 31 days. However, when only one year of data was available the best method was Equation (v) based upon a normal distribution, again the reason for this is most likely that the normal distribution systemically overestimates the expected value compared to that of the population and thus keeping the flux estimates higher and so closer to the “true” load. The ANOVA of the factorial results showed that all factors had a significant effect upon outcome and degree of bias. Over all sampling frequencies the most accurate, least biased method, was confirmed as equation (v) based on the expected value calculated from the gamma distribution of all 3 years of available data and when the interaction between sampling frequency; averaging method; and assessment period shows that at the lowest sample frequencies the best method is still averaging over all 3 years of available data based upon a gamma distribution. However, the ANOVA showed that there was a significant difference in the bias of methods based upon equation (v) between the years of available data (i.e. between 2009, 2010 and 2011). But, even in the year which had the significantly highest bias (2009) the best method was still that based on equation (v) using the expected value from the fit of a gamma distribution to all 3 years of available data methods based on equation (v) – in this case the bias was 19% and the value would be 81% of the true value. To summarise, the true load based upon hourly data in 2009 was 16.8 ktonnes C/yr, based on equation (v) with the expected value based upon the fitting of a gamma distribution to all 3 years of data gives a value of 14.4 ktonnes C/yr whereas the best performing ratio method gave a value of 7.8 ktonnes C/yr.

The methods based upon Equation (v) were dramatically more precise than those for the previous methods with imprecision rising almost linearly with declining sample frequency but to only ±2% after 31days (Figure 9). The reason for the high precision of the method expressed in Equation (v) is that the method was deliberately chosen to be dominated by the major source of variation, i.e. the flow. The total annual flow is constant between sampling frequencies and so even though the estimation of the expected value of the concentration changes and becomes less precise its effect on the overall flux estimation is minimal.

Application of the results

Given that previously "method 5" had been used to calculate the DOC flux and so the result in Figure 3 can be used to correct the bias for each site for each year where there was at least monthly sampling, i.e. for a site in one year where there was weekly sampling the above result would suggest the flux would be overestimated by "method 5" but if in a subsequent year there was only fortnightly sampling then "method 5" would underestimate the flux of DOC from that site. Thus bias correcting each year of data from each site and then amalgamating to the national level suggests that were previously the DOC flux varied from 812 to 1920 ktonnes C/yr after bias correction the DOC flux varied from 797 to 3090 ktonnes C/yr (Figure 10) – all values are given to 3 significant figures for comparison purposes. Firstly, the average correction over the course of the entire time series was 1.965 (97%) thus showing by how much the flux of DOC has been underestimated in previous studies. Secondly, the range of DOC flux has a lower minimum value than the uncorrected values and this can be seen to occur in 1975 when the average sampling frequency was 23 per year, i.e. once every 17 days. As for many sites the sampling frequency was greater than every 14 days and so the flux was being overestimated by "method 5". The annual average sample frequency decreases across the period of the record with an average sample frequency of every 17 days in 1975 to once every 29 days in 2005 with a significant trend and average decrease in sampling frequency of 1 day less frequent every 3 years. Given the result in Figure 3 a significant shift in sample frequency represents a significant increase in the bias correction required over the course of the time series. Therefore, when the uncorrected series shows no significant trend over time and this may simply be because of decline in the sample frequency. In the case of the national annual DOC flux time series this was not the case: there is no significant trend in either the uncorrected or corrected series. Given the result above on the changing precision of time series it should be possible to better constrain the precision on the flux time series, however, as pointed out by Worrall and Burt (2007) the dominant error in the time series was the upscaling of the individual flux records for individual sites for each year to the scale of the country and that estimation error cannot be improved by any of the above results.

Discussion

The implications of this study are clear, that many flux estimates of DOC and other solutes or particulates are serious underestimates. Previous studies that considered the bias or imprecision of flux estimates (eg. Cassidy and Jordan, 2011) have shown similar inaccuracy and bias in low frequency sampling as demonstrated in this study but have not gone further to demonstrate the implications of what they found for any of the studies that used the biased or imprecise methods. The implications for many studies is obvious and has already been illustrated with reference to the authors study of DOC flux from the UK: any study that used either interpolation, or indeed extrapolation methods, on data with a sampling frequency of less than once every 2 weeks must be considered to have severely underestimated the flux. For example, the data used in the estimate of flux came from the harmonised monitoring scheme (HMS - Simpson, 1980) which was established in order to assess fluxes from Europe to its surrounding oceans. The HMS network use “method 5” and what has been observed for the UK data collection is that the sampling frequency has declined, perhaps for practical and economic reasons, to monthly from an average of fortnightly over the period of the scheme. Worrall and Burt (2007) have already pointed out that the HMS reported fluxes are an underestimate of fluxes because they do not allow for the unsampled catchments and in the UK the sampled catchments of the HMS only represent just over 60% of the UK land area and no correction for river flows from the other 40% was made. Now it would be possible to suggest that many of the fluxes calculated under these schemes are less than 50% of the true value because of sampling bias at such low frequency sampling. The length of the HMS records (back to 1974) could represent one solution to the problem of low sample frequency as this study has shown that in stationary data then the longer the period over which an expected value of the concentration distribution was calculated then the better the estimate of the flux. Of course that requires stationarity over several years if not decades of data but it is relatively trivial to make a time series stationary.

In some cases the implications of the above study will have little impact. Worrall et al., (2009a) calculated the DOC flux from a peat-covered catchment as part of estimating the carbon budget of the catchment. Worrall et al. (2009a) used “method 5” on samples taken weekly and suggested a standard error on the estimation of ±11%. The results of this study do suggest that interpolation was probably the best method for calculating a flux given such a sampling regime but at 7 days “method 5” would overestimate the flux by 10.8% with an imprecision of ±8%, i.e. the imprecision and indeed the flux had been overestimated. An over estimation on DOC flux in that study would mean that DOC varied from 11.3 to 77.3 tonnes C/km^2/yr rather than 12.5 to 85.9 tonnes C/km^2/yr as reported – again 3 significant figures are shown for comparative purposes. Since DOC flux is a loss pathway within the C budget then its overestimation means that the size of the C sink represented by the peats of this catchment has been underestimated. It should also be noted that in the study the dissolved CO_2 was also calculated by "method 5" but its flux peaked at 15.1 tonnes C/km^2/yr giving only a maximum bias of 1.5 tonnes C/km^2/yr.

There are cases where flux estimation will not have the same inaccuracy as discussed above. For example, Howden et al. (2010) examined the longest water quality record in the world and assessed the changes in nitrate flux over 120 years based upon monthly data, however, the monthly data in that case were a true average and not a monthly spot sample, i.e. the real sampling frequency was actually sub-weekly and so the bias would be on the order of +5% rather than the -60% which might be assumed if the detail of the sampling strategy were not understood.

What is not clear from this study is the transferability of this result, i.e. could the pattern of bias and precision found here be indicative, and therefore applicable in other catchments or for other determinands? Firstly, the result found here is similar in magnitude and direction to the results of Cassidy and Jordan (2011) and Moatar et al. (2012). The important result of this study is that the present problem with most flux calculation methods is not the approximation of the type of distribution, i.e. there was little improvement by changing between estimation based upon normal, log normal or gamma distributions rather that the greatest improvement was achieved when all the available flow data were used. The reason for the success of the approach outlined in Equation (v) is that the approach preserved that maximum amount of information from the variable with the greatest variation. Any flux estimation method is in essence the multiplication of a concentration and a flow variable and Goodman (1960) shows that the variance of any product of two variables will be dominated by the variable with the greater variance and it is easy to demonstrate that the variance of flow is greater than that of DOC concentration. Therefore, the result presented here would have general applicability wherever the variance in the flow dominates the estimation which for most water constituents would be the case.

One way of demonstrating the advantage of Equation (v) is to calculate how well each method represents the true variance in the original time series. Goodman (1960) proposed methods for the calculation of the variance of a product, where for independent variables the variance in the product of independent variables is estimated by:

$$V(xy) = \bar{x}^2 V(y) + \bar{y}^2 V(x) + V(x)V(y) \qquad \text{(vi)}$$

And where the variables are not independent by:

$$V(xy) = \bar{x}^2 V(y) + \bar{y}^2 V(x) + 2\bar{x}\bar{y}E_{11} + 2\bar{x}E_{12} + 2\bar{y}E_{21} + E_{22} - E_{11}^2 \qquad \text{(vii)}$$

Where: V(x) = variance in x; and $E_{ij} = E\{(\Delta x)^i(\Delta y)^j\}$ given $\Delta x = x - \bar{x}$. In 2009 the variance in the river flow was 1589 while that for the DOC concentration was only 29 while that for the flux series was 548400. Applying Equation (vii) for flux calculation by Equation (i) to 100 random samples of concentration and river flow for a 30 day sampling period gave an average estimate of the variance in the flux time series that was 75 ± 12% of the true value of the variance in the original flux series. Alternatively, considering flux calculation by Equation (v) then it is appropriate to use Equation (vi) because the data are no longer paired and so independent of each other then with 100 sets of randomly sampled concentration data and the river flow series for 2009 the average variance was 106 ± 9% of the true value of the variance. Therefore, Equation (v) better preserves the variance in the flux time series.

Conclusions

The study has shown that interpolation, ratio and extrapolation methods for calculating fluvial flux of solutes develop very considerable bias for sampling frequencies greater than every 14 days, with underestimation by 60% observed for dissolved organic carbon (DOC) flux in monthly samples. On the basis of an assessment of the source of variation within the flux calculation we can show that a simpler method based upon all the available flow data is less biased than existing methods with as little as 8% underestimation even on monthly spot sampling.

Acknowledgements

The authors would like to thank United Utilities for access to the DOC records.

References

Burt, T.P., Howden, N.J.K., Worrall, F., McDonnell, J.J., 2011. On the value of long-term, low-frequency water quality sampling: avoiding throwing the baby out with the bathwater. Hydrol. Process. 25(5), 828-830.

Cassidy, R., Jordan, P., 2011. Limitations of instantaneous water quality in surface water catchments: comparison with near-continuous phosphorus time-series data. Journal of Hydrology 405, 182-193.

Cohn, T.A., 2005. Estimating contaminant loads in rivers: an application of adjusted maximum likelihood to type 1 censored data. Water Resources Research 41, W07003.

Cooper, D.M., Watts, C.D., 2002. A comparison of river load estimation: application to dissolved organic carbon. Environmetrics 13, 7, 733-750.

De Vries, A., Klavers, H.C., 1994. Riverine fluxes of pollutants : monitoring strategies first, calculation methods second. Eur. Water Polln. Contr. 4, 12-17.

Duan, N., 1983. Smearing estimate: a non=parametric retransformation method. Journal of the American Statistical Society 78, 605-610.

Goodman, L.A., 1960. On the exact variance of products. Journal of the American Statistical Society 55, 708- 713.

Howden, N.J.K., Burt, T.P., Worrall, F., Whelan, M.J., 2010. Nitrate concentrations and fluxes in the River Thames 1868 to 2008: the long-term impact of land-use change. Hydrological Processes 24, 18, 2657-2662.

Ferguson, R.I., 1986. River loads underestimated by rating curves. Water Resources Research 22, 74-76.

Johnes, P.J., 2007. Uncertainities in annual riverine phosphorus load estimation, impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. Journal of Hydrology 332, 241-258.

Kronvang, B., Bruhn, A.J., 1996. Choice of sampling strategy and estimation method for calculating nitrogen and phosphorus transport in small lowland streams. Hydrological Processes 10, 11, 1483-1501.

Kulasova, A., Smith, P.J., Bevan, K.J., Blazkova, S.D., Blavacek, J., 2012. A method of computing uncertain nitrogen and phosphorus loads in a small stream from anagricultural catchment using continuous monitoring data. Journal of Hydrology 458-459, 1-8.

Littlewood, I.G., 1995. Hydrological regimes, sampling strategies, and assessment of errors in mass load estimates for United Kingdom rivers. Environment International 21(2), 211-220.

Littlewood, I.G., Watts, C.D., Custance, J.M., 1998. Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975 – 1994). Science of the Total Environment 210 (1-6), 21-40.

Moatar, F., Meybeck, M., Raymond, S., Birgand, F., Curie, F., 2012. River flux uncertainties predicted by hydrological variability and riverine material behaviour. Hydrological Processes DOI10.1002/hyp.9464.

Monteith, D.T., Stoddard, J.L., Evans, C.D., de Wit, H.A., Forsius, M., Hogesan, T., Wilander, A., Skjelkvale, B.L., Jefferies, D.S., Vuorenmaa, J., Keller, B., Kopacek, J., Vesely, J., 2007. Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. Nature 450 (7169), 537-541.

OSPAR, 1998. Principles of the comprehensive study on riverine inputs and direct discharges. OSPAR Commission, Paris, 17pp.

Phillips, J.M., Webb, B.W., Walling, D.E., Leeks, G.J.L., 1999. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. Hydrological Processes 13, 1035-1050.

Roberts, G., 1997. The influence of sampling frequency on streamflow chemical loads. Journal of Chartered Institution of Water and Environmental Management 11, 114–118.

Simpson, E.A., 1980. The harmonization of the monitoring of the quality of rivers in the United Kingdom. Hydrological Sciences Bulletin 25: 13-23.

Skarbøvik, E., Stånacke, P.,Bogen, J., Bønsnes, T.E., 2012. Impact of sampling frequency on mean concentrations and estimated loads of suspended sediment in a Norwegian river: implications for water treatment. Science of the Total Environment 433,462-471.

Walling, D.E., 1984. Dissolved loads and their measurement. In Hadley, R.F. and Walling, D.E. (Eds.). "Erosion and Sediment Yields: some methods of measurement". Geo Books, Norwich, pp. 111-177.

Webb, B.W., Phillips, J.M., Walling, D.E., Littlewood, I.G., Watts, C.D., Leeks, G.J.L., 1997. Load estimation methodologies for British rivers and their relevance to the LOIS RACS(R) programme. Science of the Total Environment 194-195, 379-389.

Webb, B.W., Phillips, J.M., Walling, D.E., 2000. A new approach to deriving "best-estimate" chemical fluxes for rivers draining the LOIS study area. Science of the Total Environment 251-252, 45-54.

Worrall, F., Burt, T.P., 2007. Flux of dissolved organic carbon from U.K. rivers. Global Biogeochemical Cycles 21(1): GB1013.

Worrall, F., Burt, T.P., Adamson, J.K., 2008. Long-term records of dissolved organic carbon flux from peat-covered catchments: evidence for a drought effect? Hydrological Processes 22, 3181-3193.

Worrall, F., Burt, T.P., Rowson, J.G., Warburton, J., Adamson, J.K., 2009a. The Multi-annual carbon budget of a peat-covered catchment. Science of the Total Environment 407, 13, 4084-4094.

Worrall, F., Burt, T.P., Howden, N.J.K., Whelan, M.J., 2009b. The Flux of Nitrate from Great Britain 1974 – 2005 in the context of the terrestrial nitrogen budget of Great Britain. Global Biogeochemical Cycles 23, GB3017.

Worrall, F., Davies, H., Bhogal, A., Lilly, A., Evans, M.G., Turner, E.K., Burt, T.P., Barraclough, D., Smith, P., and G. Merrington. The flux of DOC from the UK – predicting the role of soils, land use and in-stream losses. Journal of Hydrology 448-449, 149-160.

Figure 1. Location of the monitoring and gauging station for the high frequency data used in this study.

Figure 2. the comparison of distribution of the hourly values of the riverflow, concentration and fluxes for the study rescaled relative to the maximum observed value of each.

Figure 3. Comparison of the relative flux estimation for interpolation methods over sampling frequencies from daily to one sample every 31 days.

Figure 4. Comparison of fitted distributions to the observed flow data.

Figure 5. Comparison of the relative precision for interpolation methods over sampling frequencies from daily to one sample every 31 days.

Figure 6. Comparison of the relative flux estimation for extrapolation methods over sampling frequencies from daily to one sample every 31 days.

Figure 7. Comparison of the relative precision for extrapolation methods over sampling frequencies from daily to one sample every 31 days.

Figure 8. Comparison of the relative flux estimation for methods based upon equation (v) over sampling frequencies from daily to one sample every 31 days.

Figure 9. Comparison of the relative precision for methods based upon equation (v) over sampling frequencies from daily to one sample every 31 days.

Figure 10. Comparison between the flux of DOC from Great Britain between 1975 and 2005 calculated with and without bias correction.