1993

Albany Harbours sampling program - experiences, myths and the need for standards

David Weaver

Recommended Citation

Weaver DM (1993) Albany Harbours sampling program - experiences, myths and the need for standards. Presented at a workshop "Environmental Water Quality - sampling for nutrients" Department of Agriculture, Bunbury, August 27th, 1993.

Introduction

Making accurate and precise estimations of pollutant loadings is becoming more important as our waterways suffer from increased inputs of pollutants and as statutory requirements to meet target loads are established. The most studied pollutant in waterways in WA is phosphorus (P) since it is reported to be the stimulus for algal growth. Many reports publish nutrient load data without any discussion of potential errors in sampling, chemical analysis or load calculations, or any discussion of the assumptions made and conclusions drawn from those calculations. The result can be the adoption of incorrect sampling, analytical and calculation procedures, leading ultimately to incorrect conclusions. This may lead to an untimely delay in the implementation of appropriate management strategies to reduce nutrient loads.

These notes identify some potential sources of error in sampling, chemical analysis and load calculations. A major assumption in all discussions in these notes is that all flow data are 100% accurate and that all estimates of bias or accuracy do not include errors associated with flow measurements. Special thanks must go to the following people in the preparation of these notes:

Rob Donahue (Waterways Commission - Perth)	Contributed many ideas, a balanced perspective and many of the references
Steve Janicke (Albany Waterways Management Authority)	Contributed ideas and wrote the programs to estimate load bias and probabilities
Adrian Reed (Department of Agriculture - Albany)	Collected all the data

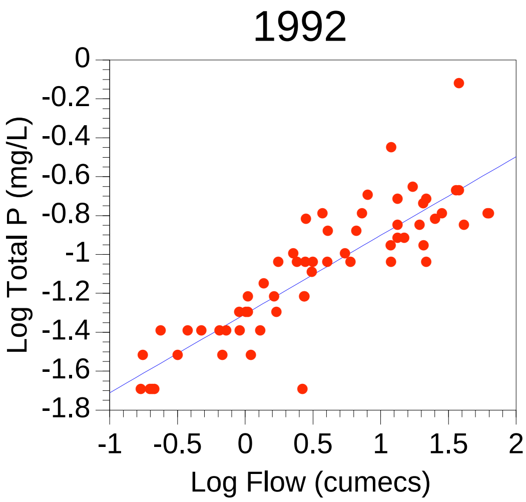
Sampling errors

Most sampling for nutrients involves the collection of surface grab samples. Martin et al (1992) compared water quality parameters in surface grab samples and cross sectionally integrated samples. Sediment concentrations and some sediment associated constituents (total phosphorus) were routinely lower in surface grab samples. Generally the percentage of fine grained material was higher in surface grab samples and these samples underestimated sediment concentrations. Finer particles are more evenly distributed vertically than coarse particles, which are more concentrated at depth. The largest differences will occur where there has been limited streamflow mixing.

Sampling strategies

Many studies have shown that nutrient concentrations increase with increasing discharge (or stage height) in river systems. Cullen and O'Loughlin (1982) and Chittleborough (1983) point out that the amount and concentration of transported substances derived from diffuse sources is linked to the occurrence of surface runoff producing events. The concentration of solutes and particulates changes most rapidly when discharge changes rapidly, so it is important to have a sampling frequency that varies with discharge. The same is true for the Kalgan River which discharges into Oyster Harbour, Albany. The figure below shows the relationship between discharge and total P concentration for the Kalgan River in 1992. It highlights the importance of sampling when peak flows occur, because the concomitant increase in discharge and concentration means that large loads of nutrients are exported in peak flow events.

Figure 1. Relationship between discharge (cubic metres per second) and Total P concentration (mg/L) for the Kalgan River in 1992



An incorrect sampling strategy could under-estimate contaminant loading by up to 85%. Figure 2 shows some sampling strategies.

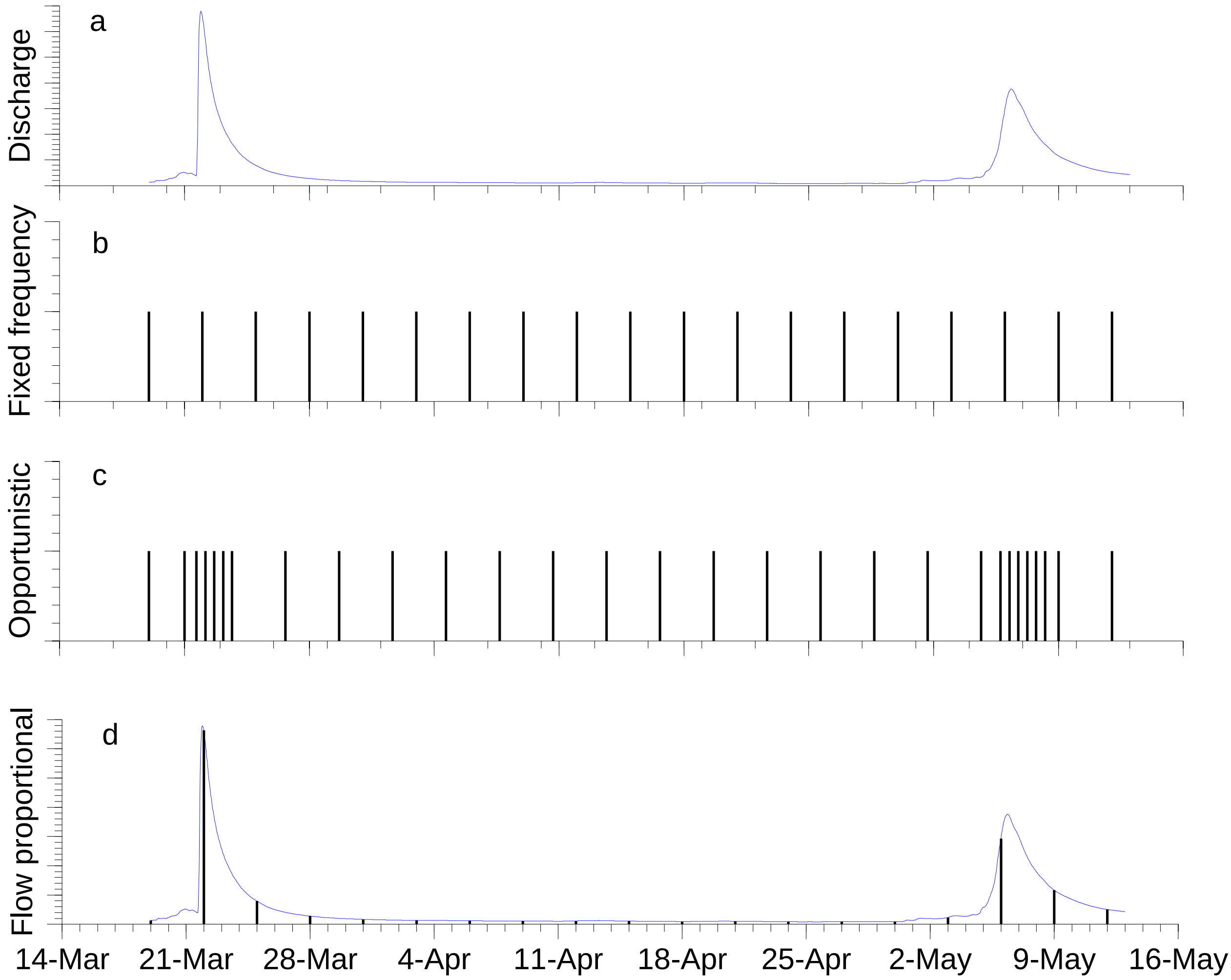
(a) Stream or river flow is shown in the first thumbnail sketch.

(b) Sampling may occur at equal time intervals (fixed frequency sampling).

(c) Sampling may occur depending on river flow or runoff events (opportunistic sampling). During low flows, a sample is collected at regular intervals. When river flow is high, more intense sampling occurs, but not necessarily at regular intervals.

(d) During flow proportional sampling, samples are usually collected at regular intervals, but a number of sub samples are mixed according to the recorded flow to give a flow weighted concentration.

Figure 2 Opportunistic, fixed frequency and flow proportional sampling strategies



An example of how opportunistic and fixed frequency sampling affects the annual P load for the Kalgan River is shown in the table below. The fixed frequency chemical dataset was generated by extracting samples from the opportunistic chemical dataset at weekly intervals. A different answer would result depending on the time and date of the first sample extracted from the opportunistic dataset.

Table 1. Comparison of load estimation based on opportunistic or fixed frequency sampling strategies

Sampling regime	Kalgan 1991 P load (tonnes)	Kalgan 1992 P load (tonnes)	Chelgiup Creek 1991 P load (tonnes)	Chelgiup Creek 1992 P load (tonnes)
Fixed frequency	5.9	8.4	1.6	0.7
Opportunistic	38.9	9.6	2.5	3.1
% difference	85	12	36	77

The importance of an adequate sampling and flow recording strategy cannot be overemphasised. During three days in July 1991 (<1% of the year) it is estimated that 36 tonnes of P passed through the Kalgan River gauging site. Only

3 more tonnes of P passed through the gauging site for the remainder of the year. If sampling and flow measurements had not occurred during this period the P load could have been grossly under-estimated.

Errors in chemical analysis

Chemical analysis varies between laboratories and according to the analytical procedure used. The following figures (3a, b and c) show the variation according to analytical procedure and laboratory.

Figure 3. Comparison of phosphorus analysis between laboratories

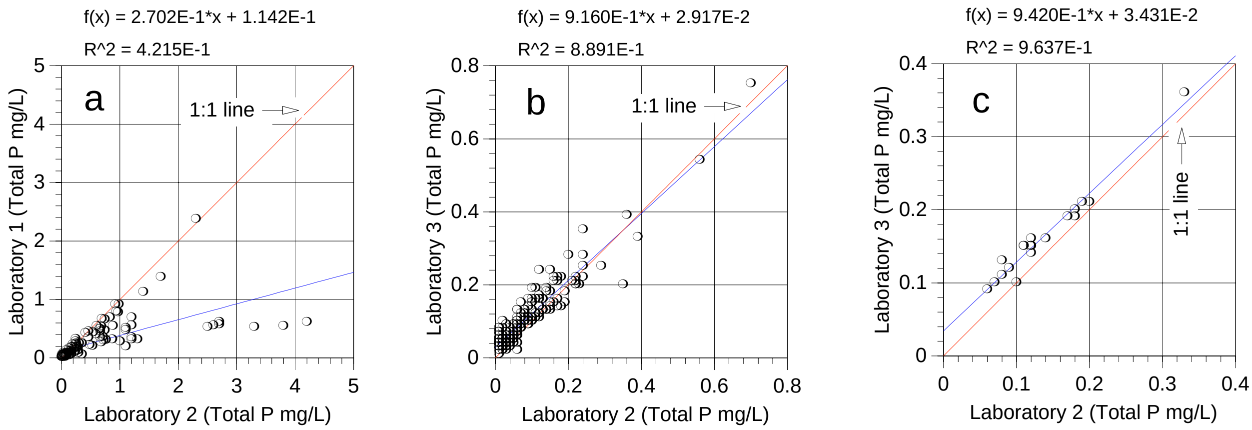


Figure 3 (a) Laboratory 1 had been informed that sediment associated nutrients were not relevant to the analysis. The analyst, in assaying for P, shook the sample and allowed any sediment to settle for up to 48 hours. After 48 hours a subsample was pipetted into a vial, reagents added, absorbance values recorded, and concentrations calculated. Laboratory 2 used standard methods for Total P analysis. The results recorded by Laboratory 1 were erratic because they were determined by the reactivity of any suspended particulates pipetted into the vial, the amount of particulates that had settled according to the temperature during the 48 hour period and how far the analyst placed the pipette tip into the sample, and the absorbance caused by any particulates remaining in the sample because no digestion step was carried out. The technical staff collecting samples had also been informed to ignore sediment collected by stage height samples and to decant the liquid into sample bottles, leaving sediment behind. Most of these problems arose because it was believed that the Albany Harbours catchments behaved exactly the same as the Peel-Harvey catchments and that sediment associated nutrients were not an issue.

Figure 3 (b) Laboratory 2 and Laboratory 3 both use standard methods for Total P analysis. There are differences, however, not as great as between Laboratory 1 and Laboratory 2. The following table shows how annual P loads in 1992 varied when chemical analysis from different laboratories is used in the calculations

Figure 3 (c) A potential source of difference between Laboratory 2 and Laboratory 3 was the delay time prior to analysis after sample collection. Figure (c) shows differences between Laboratory 2 and Laboratory 3 still occurred even when the delay time to analysis for both labs was short.

Table 2. Comparison of load estimation based on different laboratory analysis

	Kalgan 1992 P load (kg)	% difference
Laboratory 2	7179	-
Laboratory 3	9573	33

Errors in load estimation for altered chemical conditions

Total phosphorus is usually reported to an accuracy of 0.01 mg/L. Therefore, load calculations should be reported to these limits (±0.01 mg/L). The following table shows the possible range in loads for the Kalgan River in 1992 for a range of conditions of altered Total P values.

Table 3. Comparison of load estimation for a range of altered chemical conditions

Condition	Load (kg)	% difference
Raw data	7179	-
-0.01 mg/L	6425	-11
+0.01 mg/L	7934	11
-0.02 mg/L	5705	-21
+0.02 mg/L	8687	21
-0.05 mg/L	3777	-47
+0.05 mg/L	10950	53
Total P peaks removed	4501	-37
Fixed P concentration (0.04 mg/L)	3016	-58
Fixed P concentration (0.01 mg/L)	754	-89

The above table assumes that correct methods are being used to calculate loads in the first place.

Recent developments in the measurement of "soluble" and "particulate" nutrients

A number of recent papers highlight problems associated with the choice of 0.45μm filters as the arbitrary separation between "particulate" and "dissolved" fractions (Douglas et al, 1993; Oliver et al, 1993) and the tendency to ignore the <1μm in the determination of sediment loads (Hart et al, 1993).

Douglas et al (1993) states that "the common definition that material present in solutions that have been filtered through a <0.45μm filter as being 'dissolved' is not correct. Colloidal species at least two orders of magnitude smaller than 0.45μm can be fractionated from such solutions. The term 'dissolved' must be applied with far more caution in the definition of filtered material." Douglas et al (1993) separated samples into five particulate fractions (25-1000μm, 1-25μm, 0.2-1μm, 0.006-0.2μm, 0.003-0.006μm) and one dissolved fraction (<0.003μm). They also found an increase in the content of organic carbon, Mg, Ca, Na, K, Cu and Zn with decreasing particle size, indicating the importance of the colloidal fraction. Even as early as 1983, Chittleborough (1983) pointed out that much of the material that was considered to be 'dissolved' (<0.45 μm) was in fact colloidal and that most (56.5%) of the P in creek water was associated with colloids (0.008-0.2 μm)

Hart et al (1993) observed that the inclusion of coarse colloidal matter (0.1-1.0μm) increased the estimation of particulate matter transport using standard methods (particle size $>\approx 1\mu$m) by 18 tonnes to 38 tonnes. This indicated that previous estimates of the quantities of particulate matter transport in the creek system under investigation may have been underestimated by as much as 50%. Hart et al (1993) concluded that many previous studies may have therefore underestimated the particulate load and overestimated the dissolved load.

Oliver et al (1993) used tangential flow filtration (TFF) to produce three phosphorus fractions ("dissolved P" <3 nm, "colloidal P" 3 nm - 1 μm and "particulate P" >1 μm) in waters collected from the Murray and Darling Rivers. At each location where samples were analysed the colloidal fraction contained about 40% of the P. The greatest proportion of dissolved P was in the Darling River (45%) whilst the Murray River had 6-17% in this form.

Accuracy of flow event load estimation based on sampling regime

Table 1 shows how opportunistic and fixed frequency sampling strategies compare when estimating loads. However, there are an infinite number of possible sampling regimes that may occur within a flow event. Each of these possible sampling regimes will describe the chemistry of the event in a different way, depending on the number of samples collected and when each sample is collected. Even when the number of samples collected during an event is the same, an infinite number of sampling regimes could occur. This is because of the practicalities of getting to distant sites and if events occur at night, weekends or during working hours. The sampling regime will also be different for different personnel.

Figure 4. Discharge, sampled total P and a possible scenario of sampled total P

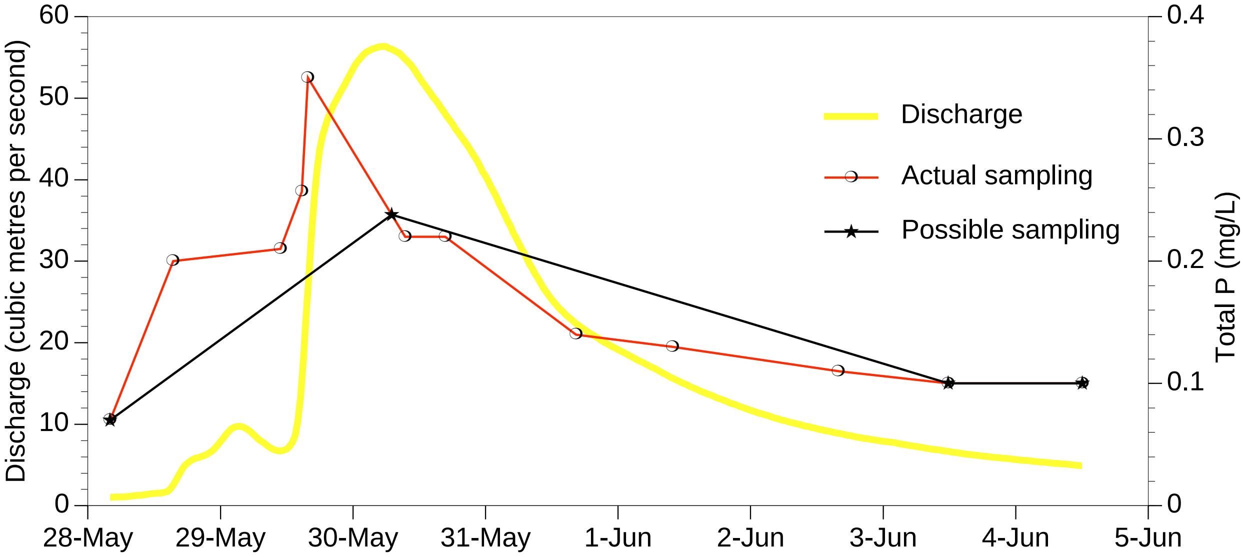


Figure 4 shows how discharge and total P (12 samples) varied over an event for the Kalgan River in 1993. It also shows a potential sampling regime of 4 samples based on the actual sampling regime. There are an infinite number of combinations of sampling regimes that include 4 samples over this event. Therefore, there are an infinite number of loads that could be calculated for all the random combinations of 4 samples during the event. By using statistical procedures we can build up a distribution of loads for any sampling regime. We can then determine the load bias (% difference from actual load) for that sampling regime, and determine what probability there is that the distribution loads are within a certain percentage of the true load.

This procedure was carried out for the data shown in Figure 4. We assume that the actual sampling represents the total P characteristics for the event. In each case 500 surveys of the linearly interpolated total P concentrations were executed to arrive at a load distribution for the sampling regime specified. The combinations of sampling regimes shown in Table 4a were tested. These included a random sampling with fixed end points, a random sampling with fixed end points and the peak always included, and a random sampling with fixed end points but with a region of peak concentrations excluded.

- The bias was calculated as the percentage difference from the true load. The probability value represents the proportion of the 500 distribution load values that are within a specified percentage of the true load.
- Random sampling tended to underestimate loads as did excluding the peak region
- Including the peak tended to overestimate loads
- Including the peak gave a narrower range of predicted loads than the other methods (ie. the frequency distribution of predicted loads had a lower standard deviation than random sampling or exclusion of the peak region).
- There is a higher probability that distribution loads are included within 10% of the true load value when the peak is included and there are more than 6 samples

Table 4a. Mean, minimum, maximum, standard deviation and bias of event load estimation and probability that distribution load will be included within 5%, 10% or 20% of the true load

Random sampling

No. of random samples	Total number of samples	Minimum load (kg)	Maximum load (kg)	Mean load (kg)	Standard Deviation	% bias	Probability		
							5%	10%	20%
10	12	1427	2892	2236	189	-4.21	0.49	0.79	0.97
8	10	1291	2871	2192	241	-6.12	0.38	0.67	0.90
6	8	1178	2901	2141	289	-8.30	0.27	0.56	0.84
4	6	1005	3199	2045	406	-12.40	0.18	0.39	0.68
2	4	989	3209	1794	522	-23.18	0.07	0.18	0.44
1	3	975	3224	1561	563	-33.13	0.03	0.10	0.29

Peak always included

No. of random samples	Total number of samples	Minimum load (kg)	Maximum load (kg)	Mean load (kg)	Standard Deviation	% bias	Probability		
							5%	10%	20%
9	12	2309	2907	2450	84	4.93	0.55	0.94	1.00
7	10	2302	3056	2488	114	6.56	0.41	0.87	0.96
5	8	2309	3157	2547	139	9.09	0.22	0.73	0.93
3	6	2339	3253	2679	205	14.74	0.08	0.42	0.73
1	4	2539	3258	2941	243	25.97	0.00	0.03	0.34

Peak region excluded

No. of random samples	Total number of samples	Minimum load (kg)	Maximum load (kg)	Mean load (kg)	Standard Deviation	% bias	Probability		
							5%	10%	20%
10	12	1280	2237	2000	148	-14.35	0.00	0.56	0.93
8	10	1179	2270	1960	186	-16.07	0.00	0.25	0.78
6	8	1052	2283	1898	233	-18.72	0.01	0.20	0.65
4	6	1052	2321	1795	298	-23.10	0.03	0.12	0.50
2	4	986	2304	1587	376	-32.01	0.03	0.08	0.32
1	3	975	2190	1416	404	-39.36	0.00	0.05	0.24

This same procedure was used to estimate bias and probability for a region of flow that was characterised by baseflow only (Table 4b). Only random sampling was tested. Similar results were found in that the load tended to be underestimated, however, the bias was much smaller for the same number of samples where a significant runoff event occurred. The probabilities were also higher, indicating that there was a greater likelihood that the correct load would be determined for a given number of samples, independent of their location in time.

Table 4b. Mean, minimum, maximum, standard deviation and bias of event load estimation and probability that distribution load will be included within 5%, 10% or 20% of the true load

Random sampling

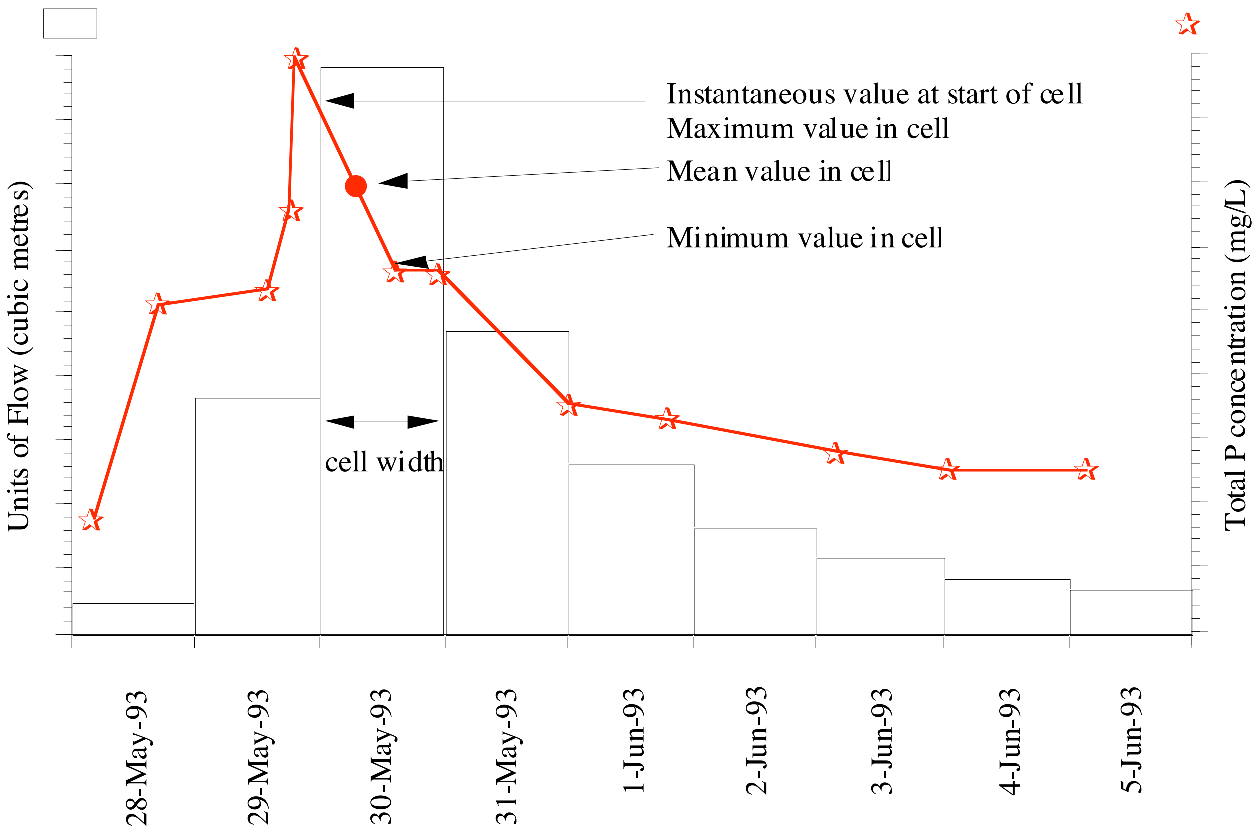
No. of random samples	Total number of samples	Minimum load (kg)	Maximum load (kg)	Mean load (kg)	Standard Deviation	% bias	Probability		
							5%	10%	20%
10	12	21.4	24.5	24.0	0.34	-0.45	0.98	1.00	1.00
5	7	19.6	24.7	23.6	0.87	-2.16	0.82	0.95	1.00
2	4	17.2	24.7	22.2	1.85	-7.80	0.46	0.66	0.92
1	3	17.1	24.7	20.7	2.26	-14.23	0.22	0.36	0.68

We can see from the data in Tables 1-4 that it would be extremely difficult to achieve a bias of <1%. It is possible using the data in Table 4 to determine how many samples would be required to achieve a bias of 1%. Thirty three random samples would be required to achieve a bias of 1%.

Accuracy of load estimation based on calculation method

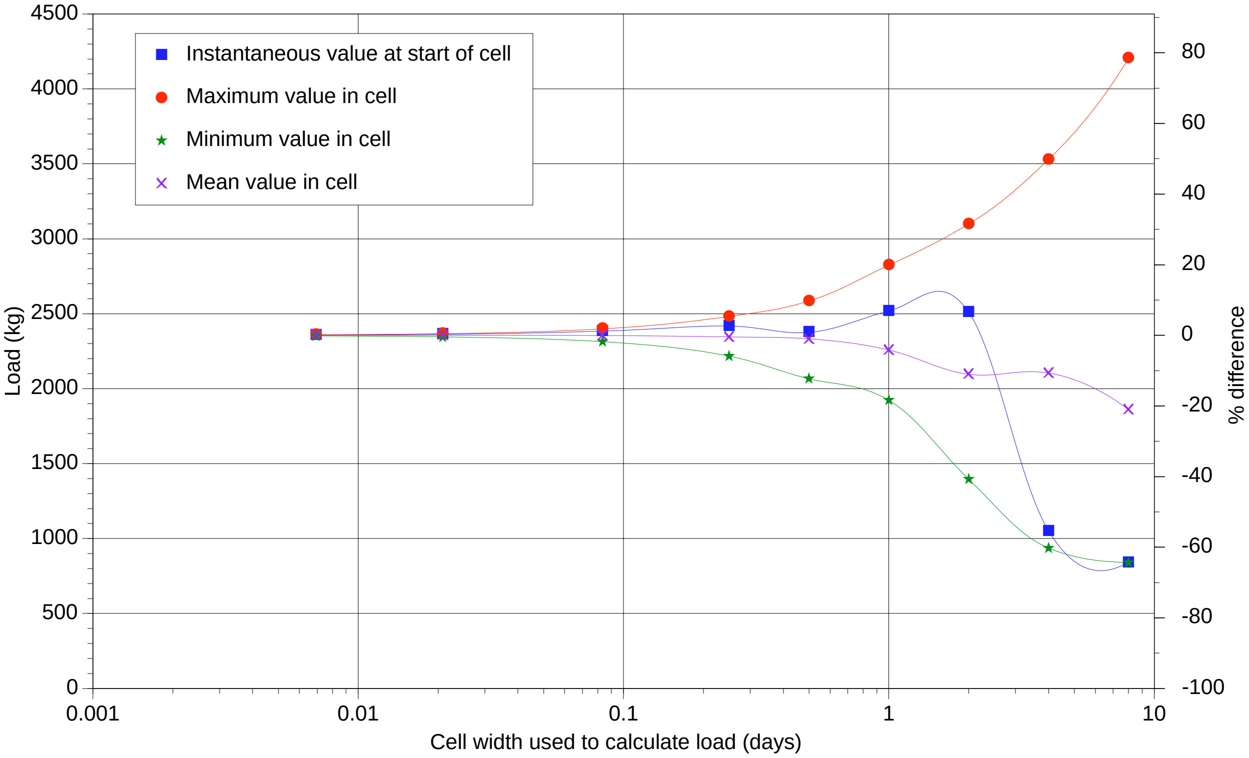
The most commonly used tools to calculate nutrient loads are spreadsheets and time series data management packages. It is common to use a daily time step in spreadsheets to calculate loads. This approach only allows for a single sample each day, unless you use much finer intervals of flow and hence concentration in a spreadsheet. In small catchments with flashy hydrographs, it may only take a few hours for the water level to reach its peak. Often when stage height samples are collected on the rising stage, more than one sample is collected in a day interval. If a spreadsheet is used to calculate loads, which sample concentration do you use to estimate the load for the chosen flow interval (1 day) when the concentration may change several orders of magnitude? Do you use the mean concentration, the maximum, the minimum or the concentration at the start of the flow interval? Figure 5 shows a possible scenario of conditions where the units of flow are daily total cubic metres and the total P concentrations are indicated by the stars. You can see that on a number of occasions there is more than one sample per day and on other occasions there are no samples per day. To calculate the load we need one value of concentration for each value of flow. Which do you use given the wide range of concentrations shown?

Figure 5. Dealing with continuous and discontinuous data



These difficulties arise because one of the variables used to calculate a load (flow) is almost continuous (has a very narrow but unequally spaced interval of 5 to 15 minutes, based on the trace from a logger) and the other variable (concentration) is discontinuous (has a wide and unequally spaced interval based on the sampling regime). Until sufficient samples are taken over peak events to characterise a hydrograph chemically, we can only linearly interpolate between the discontinuous concentrations to achieve continuous concentrations. This way we have a concentration value for each flow interval that we have chosen. There is however an effect of interval width on the calculated load, depending on whether you choose to use the mean concentration, the maximum concentration, the minimum concentration or the concentration at the start of the flow interval. This is shown in Figure 6. Note that there is almost no effect when the interval is very small (ca. <30 minutes shows <1% difference from the estimated true load).

Figure 6. The effect of choice of cell width and cell value on load estimation and difference (%) from the estimated true load



Summary and Conclusions

There are a range of potential errors associated with water quality monitoring. These include:

- sampling techniques eg. surface grab sampling versus cross sectionally integrated
- sampling procedures eg. fixed frequency sampling tends to underestimate nutrient loads more than opportunistic sampling. Including the peak value tends to overestimate nutrient loads. Collecting many samples over a peak leads to the most accurate and precise estimates of nutrient load
- chemical analysis eg. differences between laboratories can lead to substantial differences in nutrient loads
- there is some doubt as to the use of the term "soluble" for samples filtered through 0.45μm filters. Filterable may be a better term.
- the less than 0.45μm fraction can contain material that is particulate and colloidal and has nutrients associated with it
- load calculation method eg. the least error in the estimation of nutrient load using the integration method was obtained by using flow information with the best resolution possible (5-15 minutes) and by converting discrete time series chemistry into a continuous variable by linear interpolation.

If nutrient loads are to be reported, the methodology in arriving at those nutrient loads must be reported also. Sampling technique, sampling procedure, chemical analysis procedure and load calculation method (including the time step used in making calculations) should also be reported as part of a nutrient load. Some estimate of the error should also be reported. Most reported loads will be an underestimate, and it is **unlikely**, given that the discussion above includes no estimate of the errors associated with flow, that load estimates using currently established techniques are underestimated by less than 30%. Sampling procedures that best define hydrograph chemistry should be used where possible.

References and further reading

Bruton-G (1982) Water quality sampling and data analysis. In Water Quality Management: monitoring programs and diffuse runoff. Ed by Barry Hart, Water Studies Centre, Chisholm Institute of Technology, Melbourne

Burn-DH (1990) Real time sampling strategies for estimating nutrient loadings. Journal of Water Resources Planning and Management, 116(6): 727-741

Chittleborough-DJ (1983) The nutrient load in surface waters as influenced by land use patterns. In "The effects of changes in land use upon water resources" Ed JW Holmes, Water Research Foundation of Australia

Cullen-PW and O'Loughlin-EM (1982). Non-point sources of pollution. In "Prediction in water quality" Eds Cullen and O'Loughlin. Aust. Acad. of Sci., Canberra, pp 437-453

Douglas-GB, Beckett-R and Hart-BT (1993) Fractionation and concentration of suspended particulate matter in natural waters. Hydrological Processes, 7, 177-191

☞ **Ellis-JC (1989) Handbook on the design and interpretation of monitoring programmes. Water Research Centre plc, Henley Road, Medmenham, PO Box 16, Marlow, Bucks. SL7 2HD**

Hart-BT, Douglas-GB, Beckett-R, vanPut-A and van Grieken-RE (1993) Characterisation of colloidal and particulate matter transported by the Magela Creek system, northern Australia. Hydrological Processes, 7, 105-118

Hoare-RA (1982) Nitrogen and phosphorus in the Ngongotaha stream. New Zealand J. of Mar. and Freshwater Res., 16, 339-349

Martin GR, Smoot JL and White KD (1992). A comparison of surface grab and cross sectionally integrated stream water quality sampling methods. Water Environment Research, 64 (7): 866-876

Oliver-RL, Hart-BT, Douglas-GB and Beckett-R (1993) Phosphorus speciation in the Murray and Darling Rivers. Chemistry in Australia (latest issue)

Preston-SD, Bierman-VJ and Silliman-SE (1989) An evaluation of methods for the estimation of tributary mass loads. Water Resources Research, 25(6): 1379-1389

Rekolainen-S, Posch-M, Kamari-J and Ekholm-P (1991) Evaluation of the accuracy and precision of annual phosphorus load estimates from two agricultural basins in Finland. Journal of Hydrology, 128:237-255

Richards-RP (1985) Estimating the extent of reduction needed to statistically demonstrate reduced nonpoint phosphorus loading to Lake Erie. J. Great Lakes Res, 11:110-117

Richards-RP and Holloway-J (1987) Monte Carlo studies of sampling strategies for estimating tributary loads. Water Resources Research, 23(10): 1939-1948

Rosich-RS and Cullen-P (1982) Nutrient runoff. In Water Quality Management: monitoring programs and diffuse runoff. Ed by Barry Hart, Water Studies Centre, Chisholm Institute of Technology, Melbourne

Stack,-WP; Belt,-KT (1989) The selection of appropriate flow averaging periods in evaluating pollutant loadings using the flow interval method. Lake-and-Reservoir-Management. 1989, 5: 2, 67-73

Stevens-RJ and Stewart-DA (1981) The effect of sampling interval and method of calculation on the accuracy of estimated phosphorus and nitrogen loads in drainage water from two different sized catchment areas. Record of Agricultural Research, 29:29-38

Yaksich-SM and Verhoff-FH (1983) Sampling strategy for river pollutant transport. Journal of Environmental Engineering, 109(1): 219-231

Young-TC, dePinto-JV and Heidtke-TM (1988) Factors affecting the efficiency of some estimators of fluvial total phosphorus load. Water Resources Research, 24(9): 1535-1540

☞ Essential reading