Hydrol. Earth Syst. Sci., 18, 4721–4731, 2014
www.hydrol-earth-syst-sci.net/18/4721/2014/
doi:10.5194/hess-18-4721-2014

Hydrology and Earth System Sciences
Open Access

Comparison of sampling methodologies for nutrient monitoring in streams: uncertainties, costs and implications for mitigation

J. Audet[1,4], L. Martinsen[2], B. Hasler[2], H. de Jonge[3], E. Karydi[1], N. B. Ovesen[1], and B. Kronvang[1]

[1]Department of Bioscience, Aarhus University, Vejlsøvej 25, 8600 Silkeborg, Denmark
[2]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark
[3]Sorbisense A/S, P.O. Box 10, 8830 Tjele, Denmark
[4]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050, 750 07 Uppsala, Sweden

Correspondence to: J. Audet (joachim.audet@slu.se)

Received: 27 May 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 4 July 2014
Revised: 16 October 2014 – Accepted: 29 October 2014 – Published: 28 November 2014

Abstract. Eutrophication of aquatic ecosystems caused by excess concentrations of nitrogen and phosphorus may have harmful consequences for biodiversity and poses a health risk to humans via water supplies. Reduction of nitrogen and phosphorus losses to aquatic ecosystems involves implementation of costly measures, and reliable monitoring methods are therefore essential to select appropriate mitigation strategies and to evaluate their effects. Here, we compare the performances and costs of three methodologies for the monitoring of nutrients in rivers: grab sampling; time-proportional sampling; and passive sampling using flow-proportional samplers. Assuming hourly time-proportional sampling to be the best estimate of the "true" nutrient load, our results showed that the risk of obtaining wrong total nutrient load estimates by passive samplers is high despite similar costs as the time-proportional sampling. Our conclusion is that for passive samplers to provide a reliable monitoring alternative, further development is needed. Grab sampling was the cheapest of the three methods and was more precise and accurate than passive sampling. We conclude that although monitoring employing time-proportional sampling is costly, its reliability precludes unnecessarily high implementation expenses.

1 Introduction

Rivers act as a major transport route for particulate and dissolved matter at catchment scale (Meybeck, 1982; Seitzinger et al., 2002). Information on the geochemical composition of the transported substances is valuable to improve our knowledge of these and quantify the erosive processes affecting the continental surface, as well as to estimate nutrient fluxes towards aquatic recipients such as lakes, estuaries, fjords, seas and oceans (Meybeck, 1982).

In recent decades, the transport of nitrogen (N) and phosphorus (P) has attracted particular attention. Anthropogenic activities, such as the increasing use of fertilisers for agricultural purposes or poor wastewater treatment capacity, have greatly affected the nutrient cycle, causing enhanced release towards aquatic ecosystems (Vitousek et al., 1997; Smith et al., 1999).

Excessive concentrations of N and P are responsible for the eutrophication of aquatic ecosystems (Carpenter et al., 1998; Birgand et al., 2007; Moss, 2008), which may lead to hypoxia and loss of biodiversity, therefore posing a health risk to humans via drinking water supplies (Smith et al., 1999). In consequence of this, in 2000 the European Union adopted the Water Framework Directive (WFD) to mitigate nutrient pollution of aquatic ecosystems. The WFD requires member states to establish at least "good" ecological status in their water bodies and requires that mitigation strategies – i.e. the chosen measures and implementation – should be cost effective. As an example, in Denmark, fulfilment of WFD requirements on a national scale involves reduction of N and phosphorus loads by 19 000 t N and 210 t P, respectively (http://naturstyrelsen.dk/vandmiljoe/vandplaner/), and Jensen et al. (2013) estimates the total cost of achieving these reduction targets to be around EUR 218 million. Therefore, reliable monitoring estimates of nutrient transport in rivers are required to select appropriate mitigation strategies and evaluate their effects.

In Denmark, the monitoring of nutrients in streams and rivers includes fortnightly or monthly sampling (Kronvang et al., 1993). The same is true for many monitoring programs in Europe; for example, in France, 80 % of water quality surveys since 1971 are based on monthly samplings (Moatar and Meybeck, 2005). Most water quality monitoring programmes are based on grab sampling involving collection of a small volume of water, generally 1–2 L, in the river. The sample is stored in a cooling box and sent directly for laboratory analysis. This method is quick and simple but has some disadvantages in that it only reveals the geochemical composition of the water at the precise moment of sampling and does not take into account that the composition may change rapidly over time (Kronvang and Bruhn, 1996; Jordan and Cassidy, 2011). Consequently, to obtain water samples depicting the temporal variability of nutrient concentrations in streams, continuous monitoring methods have been developed. These rely on flow-proportional sampling, time-proportional sampling or high frequency sampling and in situ analyses (Kronvang and Bruhn, 1996; Jordan and Cassidy, 2011). These methods are, however, costly because they require an on-site station with a power supply and perhaps also a cooling device to refrigerate and preserve the samples. Also, in areas with winter temperatures below zero, a heating device may be necessary to prevent freezing of the sampling system.

Passive samplers enabling in situ continuous sampling over time may be an alternative to the above methods as they do not require a power supply or storage and refrigeration equipment (Rozemeijer et al., 2010). However, tests to confirm the reliability of passive samplers, such as the flow-proportional SorbiCell sampler (SC-sampler) (de Jonge and Rothenberg, 2004), should be conducted under different flow conditions in streams (Jordan et al., 2013).

The objectives of this study were: (1) to test the SC-sampler under controlled conditions in flumes and in two different natural lowland streams; (2) to compare the reliability of utilising SC-samplers, grab sampling and time-proportional composite sampling to estimate nitrate and P concentrations; (3) to compare the costs of the SC-sampler, grab sampling and time-proportional composited sampling, and (4) to compare monitoring costs with the costs of implementing river basin management plans under the WFD.

2 Material and methods

2.1 Sampling methodologies

Three sampling methodologies were tested in this study – passive samplers, grab sampling and automated time-proportional sampling. The flow-proportional passive sampler SorbiCell (SC-samplers; de Jonge and Rothenberg, 2004) manufactured by Sorbisense A/S, Tjele, Denmark, which is capable of measuring average concentrations of nutrients and other substances over time (weeks–months), was applied. The sampler contains an adsorbent that captures nutrients and a soluble tracer salt (calcium-citrate) that dissolves when water passes through the sampler. The flow of water through the SC-sampler is estimated from the dissolution of the salt tracer. SC-samplers are equipped with a filter (mesh size 40–100 µm) to prevent entry of large particles to the cartridge. Average solute concentration for the installation period is calculated based on the mass of solute adsorbed and on the mass of tracer salt lost. Further details on SC-samplers are provided in de Jonge and Rothenberg (2004). Grab sampling involved filling a 2000 mL bottle with stream water collected in running water in the middle of the stream. Automatic time composited samples were taken on an hourly basis using an ISCO Glacier® Sampler (Teledyne ISCO, Lincoln NE, USA). The collected samples were kept refrigerated in the sampler until recollection and home transport for analysis.

2.2 Nutrient analysis

The SC-sampler samples were analysed for nitrate and P (a detailed description of the analysis of nitrate is provided in Rozemeijer et al., 2010). Phosphorus was determined as molybdate reactive P (without filtration) after extraction with 2 M HCl and was designated as SC-P. Tracer was extracted in 0.2 M HCl and measured as Ca in solution by atomic absorption spectroscopy. Nitrate in the water samples collected by grab sampling and continuous sampling was analysed on a Dionex ICS-1500 IC system (Dionex corp.; Sunnyvale, USA) after filtration at 0.22 µm (nylon membrane SNY 2225; Frisenette, Denmark), and total P (non-filtrated; TP), total dissolved P (0.45 µm filtration; TDP) and dissolved inorganic P (DIP) were analysed following the standard method DS/EN ISO 6878 (2004).

2.3 Flume experiment

The main aim of this first experiment was to determine the flow conditions suitable for use of SC-samplers. The passive samplers were tested in six flumes (12 m long and 0.6 m wide) having constant flow velocity (0.05, 0.08, 0.13, 0.15, 0.18 and 0.25 $m\,s^{-1}$), representing well the normal velocities and flow conditions of smaller lowland streams (Ovesen et al., 2000). The substrate was identical in all the flumes and consisted of a mixture of gravel and sand, mimicking the substrate commonly encountered in Danish streams. The flumes received water pumped from a nearby stream and therefore the water chemistry was the same in the six flumes. The experiment was conducted in late summer, during base-flow condition of the stream and therefore nutrient concentrations were relatively stables. Two to four SC-samplers were deployed on the same day in the six flumes (Fig. 1) and retrieved after 7 days; at the same time flow velocity was measured with a current meter OTT-Kleinflügel at the different SC-sampler positions in the flumes. During the deployment, water samples were collected using time-proportional sampling method and the samples were analysed for nitrate, TP, TDP and DIP.

2.4 In situ stream experiment

Nutrients were monitored at two stations located in two differently shaded lowland streams located in Jutland, Denmark; one in the open Odderbaek stream and one in the more shaded Gelbaek stream. The Odderbaek stream is a second order stream (Strahler, 1957) and has a catchment size of 27.6 km^2, of which 68 % is used for agricultural purposes. The monitoring station at Odderbaek was placed near the mouth of the stream before it flows into Lake Kulsø (latitude 55.932° N, longitude 9.310° E). Upstream of the station, a 1 km stretch was restored in December 2010 by raising the stream bed, creating meanders and disconnecting tile drains (Audet et al., 2013). The Gelbaek station was positioned at Lyngby Bridge (lat. 56.225° N, long. 9.881° E). The Gelbaek stream is a first order stream draining 11.6 km^2 of intensively farmed (> 95 % arable land) catchment with a corridor of trees in the buffer strip along the lower 2 km of the stream channel (Kronvang et al., 1997).

2.5 Sampling strategy and hydrology

The streams were visited at approx. 2–4 week intervals during June 2010–May 2011 at Odderbaek and November 2010–October 2011 at Gelbaek. On each occasion, grab sampling was performed, SC-samplers (triplicates) were collected, new passive samplers were installed and the composite sample from the automatic ISCO sampler was collected.The position of the passive samplers in the water column was adjusted at every deployment to be set at approx. 0.6 × water height to ensure comparable position in the velocity gradient of the stream cross-section. At Gelbaek, automatic sampling was interrupted between December and February due to freezing of the sampling system. Water samples obtained from grab and automatic sampling were analysed for nitrate and TP, whereas the SC-samples were analysed for nitrate and P (SC-P).

For both streams, water discharge was calculated from continuous measurements of stage utilising a vented pressure transducer and establishing a stage-discharge relationship at different water stages to cover the entire hydrological regime.

The monthly transport of nutrients was estimated by multiplying the daily discharge with the daily concentration derived from the three methods. For the grab sampling, the concentrations were linearly interpolated between sampling dates, while for the passive and continuous samplers the average concentrations obtained were used for each measurement period (ca. 2 week periods).

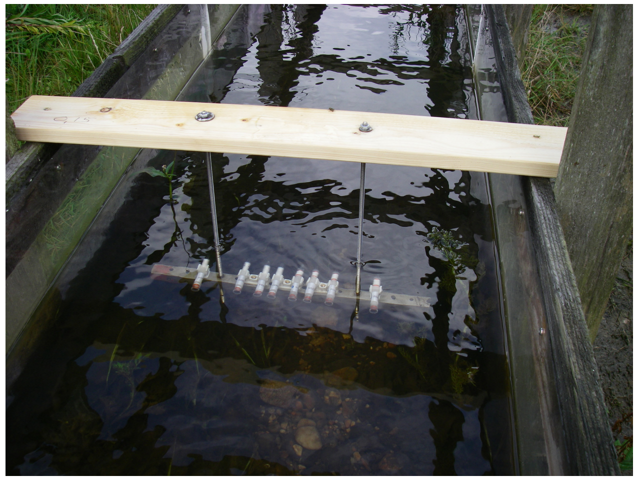

Figure 1. Picture of SorbiCell passive samplers installed in a study flume.

2.6 Statistics

Accuracy (bias) and precision were used to compare the results derived from the three sampling methods, the results from the time-proportional composited sampling being regarded as our best estimate of the "true" concentration (see Sect. 4.2). Accuracy ($\overline{\varepsilon}$) was evaluated by calculating the mean of the relative errors (ε), and the standard deviation (s) gave a measure of the precision of ε. Root-mean-square error (RMSE) was also used as it combines these two concepts (Dolan et al., 1981).

$$\mathrm{RMSE} = \sqrt{\overline{\varepsilon}^2 + s^2}$$

To check if the concentrations obtained from the SC-cells and the grab samples differed significantly from those of the time-proportional composited samples (i.e. the "true" concentrations), we used paired student's t test.

2.7 Measuring the costs of the sampling methods

The total costs of implementing the different sampling methods can be divided into different categories such as: investment costs; operational costs; and maintenance costs, whose relative weight varies. Investment costs refer to one-time costs for equipment and facilities, operational costs include salary and other input and service costs, for instance sampling bottles and analyses, and maintenance costs refer to costs associated with maintenance of equipment, in our case only relevant for water level measurements. The costs are assessed in welfare-economic prices (Johansson, 1993) and thus reflect the welfare-economic costs of implementation. Assessing the costs in welfare-economic prices rather than factor prices allows comparisons to be made with mitigation

Table 1. Nitrate and phosphorus concentration in the flume experiments determined by SC-samplers and time-proportional sampling. The SC-samplers were installed in the flumes for 1 week ($n = 14$). Concentrations ± standard error.

	SC-sampler		Time-proportional sampling			
	Nitrate mg N L^{-1}	SC-P mg P L^{-1}	Nitrate mg N L^{-1}	TP mg P L^{-1}	TDP mg P L^{-1}	DIP mg P L^{-1}
Week 1	0.95 ± 0.05	0.034 ± 0.002	0.97	0.056	0.018	0.011

costs. As investment costs are one-time costs, they should be spread over the lifetime of the investment. In our study, investment costs were converted into annual costs using a discount rate of 4 % and assuming a life time of 5 years. Costs for laboratory analyses are an important component for all monitoring methods and are assessed using list prices including transport of the samples to the lab, salary, materials and equipment costs. In the present case, cost calculations of the SC-sampling method were based on duplicate measurements, i.e. simultaneous use of two SC-samplers. Salary costs were calculated using an average salary of EUR 37 h^{-1} (average salary for laboratory staff at Aarhus University, Denmark). Common for all methods is that sampling requires visits to the monitoring site with ensuing salary and transport costs. Assuming that a technician was responsible for the sampling (salary EUR 37 h^{-1}) and that the study sites were located at an average distance of 50 km from the laboratory, the time requirement and transport were assessed following the unit costs provided by the Danish Ministry of Energy (EUR 0.2 km^{-1}). The need for transport was considered identical for all three monitoring methods as was the need for conducting water level and water flow measurements, implying that the three cost components only affected total monitoring costs and not the absolute difference in costs between the three methods.

3 Results

3.1 Testing of passive samplers in flumes

The testing of the passive samplers (SC-samplers) for a range of flow velocities revealed that the flow-through volume estimated from the dissolution of the tracer salt contained in the SC-samplers was directly proportional to the measured flow velocity in the flumes (Fig. 2a). This result demonstrates that the SC-samplers work at a flow-proportional rate when installed in running waters to estimate nutrient concentration. This was confirmed by the linear relationship traced between P accumulated in the passive samplers and the volume of tracer salt dissolved during the 1-week monitoring period (Fig. 2b). Similarly, a linear relationship was found between the accumulated nitrate and the volume of salt dissolved (Fig. 2c). Nitrate concentrations obtained from the SC-samplers installed for 1 week compared well with the

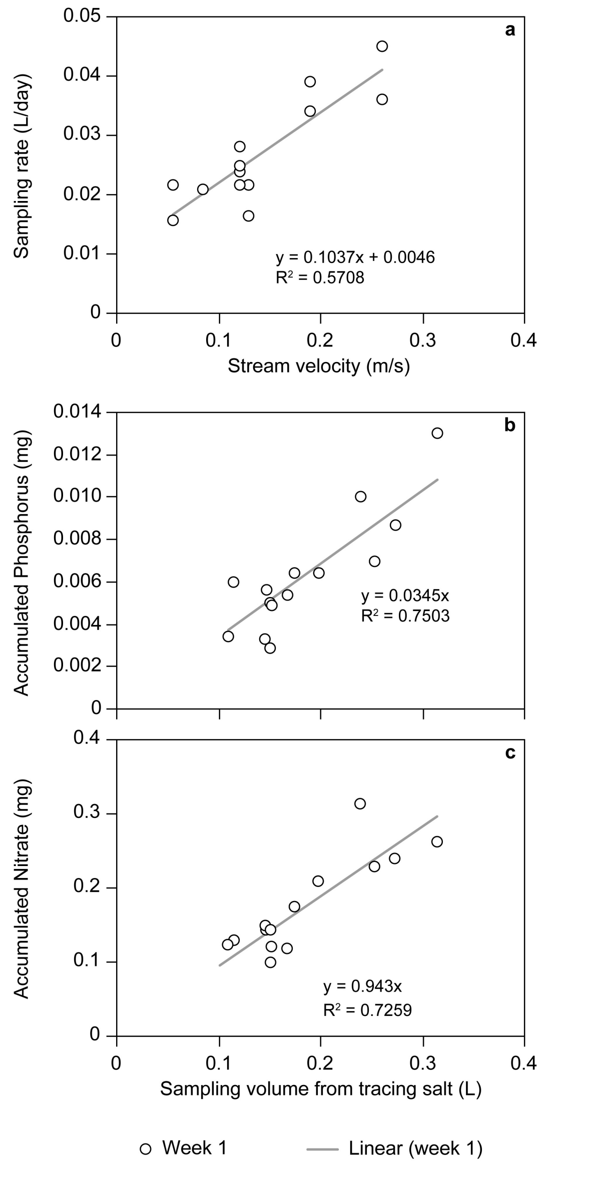

Figure 2. Relationships between **(a)** stream flow velocity in the flumes and sampling rate of the passive samplers, **(b)** between dissolved tracer salt and accumulated phosphorus in the passive samplers, and **(c)** between dissolved tracer salt and accumulated nitrate in the passive samplers. The passive samplers were installed for 1 week in the flumes.

Table 2. Accuracy (mean relative error), precision (standard deviation of the relative error) and root-mean-squared error (RMSE) of SorbiCells passive samplers and grab sampling compared to time-proportional sampling for monitoring nitrate and phosphorus in streams.

Stream	Nitrate						Phosphorus					
	SorbiCells			Grab sampling			SorbiCells			Grab sampling		
	Accuracy	Precision	RMSE	Accuracy	Precision	RMSE	Accuracy	Precision	RMSE	Accuracy	Precision	RMSE
Odderbaek	0.76	0.81	1.11	0.17	0.16	0.23	0.62	0.37	0.72	0.49	0.20	0.53
Gelbaek	0.51	0.29	0.59	0.15	0.08	0.17	0.67	0.83	1.07	0.40	0.37	0.54

results of time-proportional sampling (Table 1). Regarding P, the results from the SC-samplers were slightly lower than the TP results from the grab sampling. However, SC-P was higher than TDP and DIP (Table 1).

3.2 In situ stream sampling methods

The comparison of nitrate and P concentrations obtained by the three different sampling methods employed in two streams showed some contrasting results. For both streams, the nitrate concentrations obtained from grab and time-proportional sampling were comparable (Fig. 3) and did not differ significantly ($p > 0.05$; t test), whereas the nitrate concentration from the SC-samplers exhibited marked differences (Fig. 3). Assuming that the time-proportional method yielded the best estimate of the "true" concentration over the sampling period, the nitrate concentrations determined from the passive sampler samples were almost always overestimated at Odderbaek and generally underestimated at Gelbaek. However, the difference between the automatic samplers and the SC-samplers was only significant at Odderbæk ($p < 0.01$). For P, the concentration results from grab sampling and passive samplers were generally lower than for time-proportional sampling (Fig. 4). The P concentrations obtained from SC-samplers and grab sampling differed significantly from those of the time-proportional method at Odderbaek ($p < 0.05$ and $p < 0.01$, respectively), whereas no significant differences appeared at Gelbaek although the discharge was more "flashy" than at Odderbaek (ratio q_5 : q_{95}; 0.07 at Gelbaek and 0.21 at Odderbaek). However, pronouncedly large differences between the SC-sampler and grab sampling results were observed for TP concentrations in Odderbaek during the months of December and January (Fig. 4), which might be due to an increase in the transport of particulate P derived from erosion following heavy precipitation event as well as the restoration activities affecting the stream bed upstream of the monitoring station.

We found significant positive relationships between stream velocity and SC-sampler sampling rates ($p = 0.02$ at Gelbaek and $p = 0.02$ at Odderbaek), but variability was high as illustrated by the low R^2 (Fig. 5). Accuracy, precision and RMSE of the SC-samplers and grab sampling methods compared to the time-proportional method are presented in Table 2. For both streams, grab sampling concentrations gave a better estimate of the reference concentrations than

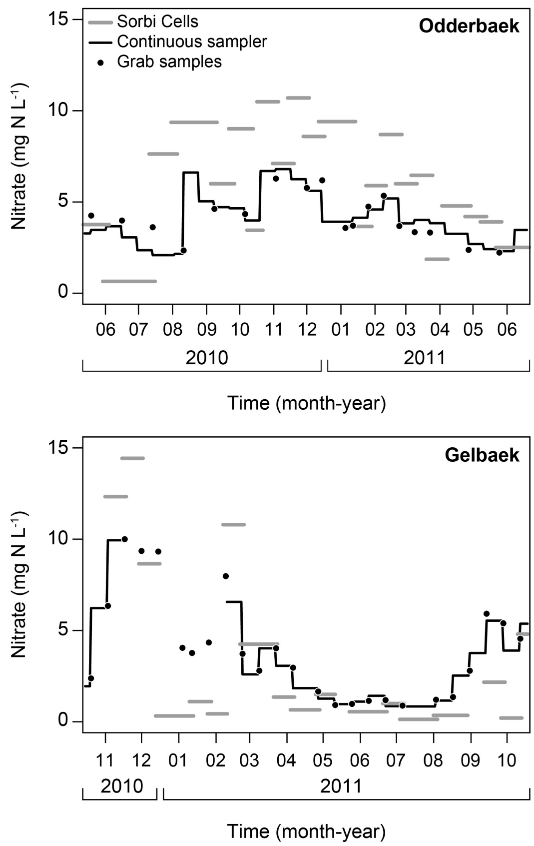

Figure 3. Nitrate concentrations at Odderbaek and Gelbaek determined by passive samplers, grab sampling and time-proportional sampler. The monitoring by the time-proportional sampler at Gelbaek was interrupted in winter because of freezing.

the SC-sampler concentrations. Nevertheless, grab sampling performance was still relatively poor for nitrate (RMSE: 23 and 17 % at Odderbaek and Gelbaek, respectively) and even poorer for P (RMSE: 53 and 54 % at Odderbaek and Gelbaek, respectively).

The results obtained for the annual transport of nitrate calculated from passive sampler concentrations showed an overestimation of 47 % at Odderbaek and an underestimation of 32 % at Gelbaek relative to the reference load (i.e. time-proportional sampling) (Table 3). For TP, the annual transport was underestimated by 43 and 23 % at Odderbaek and Gelbaek, respectively. The transport derived from

Table 3. Nitrate and phosphorus loads for three sampling methods in two streams. Deviation from the reference is given as as a percentage.

Sampling method	Odderbaek*				Gelbaek*			
	N Load		P load		N Load		P load	
	t N		kg P		t N		kg P	
SorbiCells	55.5	47 %	524	−43 %	2.3	−32 %	99	−23 %
Grab sampling	35.4	−6 %	420	−54 %	3.6	6 %	84	−35 %
Time-proportional sampling (reference)	37.7	–	915	–	3.4	–	129	–

* Load measured for the period 1 June 2010 to 31 May 2011 at Odderbaek and for the period 10 February to 31 October 2011 at Gelbaek.

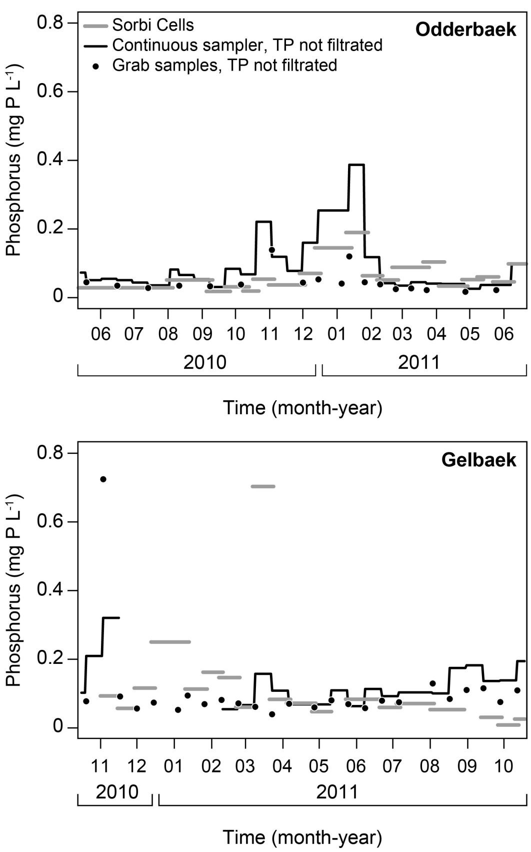

Figure 4. Phosphorus concentrations at Odderbaek and Gelbaek determined passive samplers, grab sampling and time-proportional sampler. The monitoring by the time-proportional sampler at Gelbaek was interrupted in winter because of freezing.

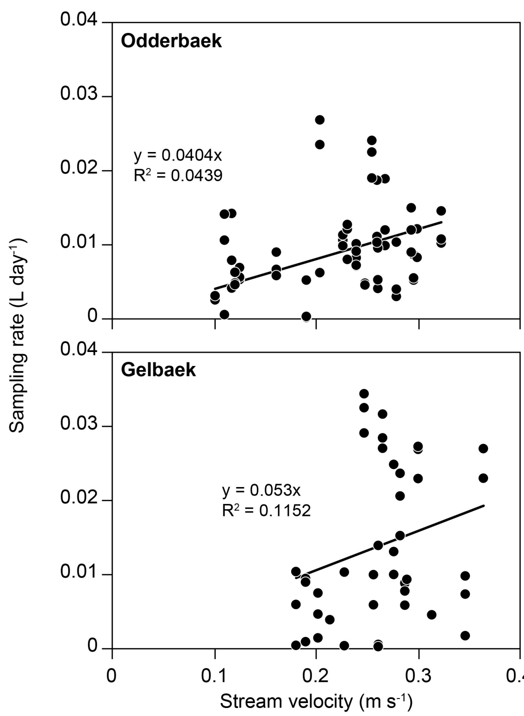

Figure 5. Relationships between stream flow velocity at Odderbaek and Gelbaek and the sampling rate of the passive samplers.

the grab sampling compared well with the reference transport for nitrate (−6 and 6 % at Odderbaek and Gelbaek, respectively) but clearly underestimated the TP transport (−54 and −35 %) in both streams (Table 2).

3.3 Costs of the sampling methods

Estimated welfare-economic costs of the sampling methods are given in Fig. 6a, which shows that SC sampling costs are nearly identical with those of time-proportional sampling, i.e. approximately EUR 3700 annually per site. From an economic point of view, the SC method has the advantages of not requiring any significant investments and allowing greater flexibility for change of sampling site. In contrast, the cost of grab sampling is much lower (EUR 2000 year^{-1} per site) as no investments are required and samples are collected using a minimum of equipment. Figure 6b shows the costs of the three methods, including water level measurement, flow level measurement and transport to and from the sampling site. Despite a substantial cost increase, this does not influence the relative cost ranking of the methods. As can be seen, non-method specific costs account for a large proportion of the total costs, revealing that monitoring is expensive irrespective of method used (Fig. 6b).

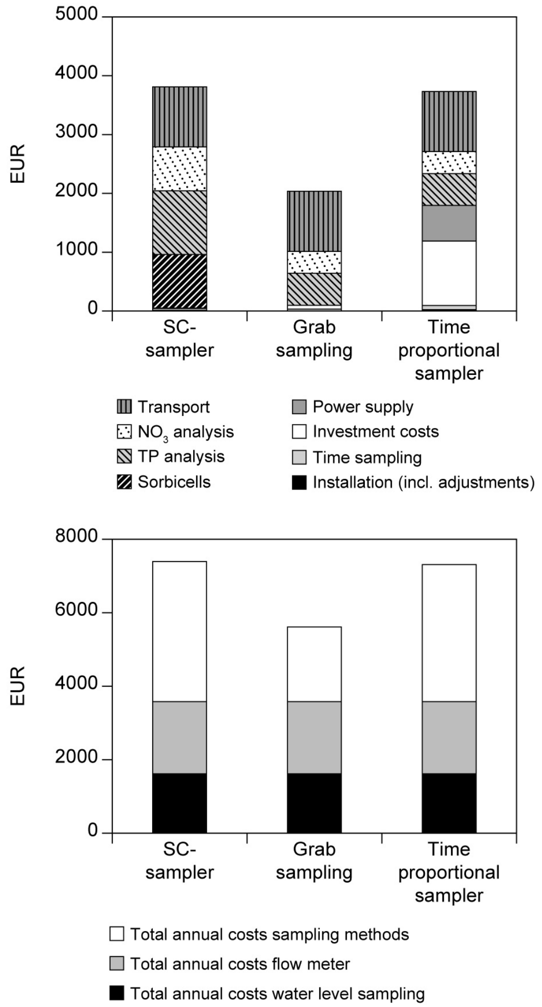

Figure 6. Comparison of costs and cost distribution per sample site and year (EUR) (top panel) and total annual costs of the sampling methods per site, including water level and flow metering (bottom panel).

4 Discussion

4.1 Monitoring of nutrients in stream waters using passive samplers

SorbiCell passive samplers (SC-samplers) have been shown to be capable of reproducing the nitrate concentration level and seasonal pattern in a stream, ditch and three tile drains in The Netherlands (Rozemeijer et al., 2010). In the Dutch study, although the SC-sampler underestimated the nitrate concentration in the summer months, calculated loads based on the SC samples were nearly similar to the loads derived from continuous measurements of nitrate concentrations using Hydrion sensor equipment (Rozemeijer et al., 2010). To some extent these findings corroborate the results of our controlled flume experiment but are not supported by our field results in the two streams. The SC-sampler had a high RMSE for both nitrate-N measurements (111 and 59 %) and P measurements (72 and 107 %) for both study streams (Table 2), these results being much inferior to those of a fortnightly grab sampling procedure for nitrate-N (23 and 17 %) and P (53 and 54 %).

In our stream experiment, flow velocities in the channel were not correctly mimicked by the SC-sampler, which in turn influences the capability of the SC-sampler to measure nutrient concentrations. This is evidenced from a comparison of data sets on in situ measured flow velocities, where the SC-samplers were mounted in the cross-sectional profile and the flow through the SC-samplers was measured from the loss of tracer salt. The linear relationships linking sampling rate to flow velocity had a slope of 0.04 in the Odderbaek stream and 0.05 in the Gelbaek stream. This is much lower than the rates recorded in our initial flume experiment where the slope was 0.10 and less water flowed through the SC-samplers in the streams than in the flumes at comparable stream velocities. We believe that the responsible factor is the physical blocking of the SC-samplers with vegetation detritus and periodically fine sediments from the stream bed and banks in Odderbaek due to restoration activities involving heavy machinery. Furthermore, Jordan et al. (2013) questioned the assumption of a linear increase in SC-sampler flow through at enhanced flow velocities produced by the increasing risk of recirculating wakes developing downstream of the cartridge. In addition, it is unclear which fraction of P is recovered in the passive samplers, which complicates comparison with standard methods. In particular, the recovery of particulate P may be poor because of the sampler's filter and potential occurrence of desorption processes in the sampler cartridge (Jordan et al., 2013). In our flume study, the P fraction recovered from the passive samplers comprised between total P and total dissolved P. Finally, the deployment duration may also have influenced the results from the SC-samplers as they were deployed for 1 week in the flumes in contrast with the streams where the sampling time was 2 weeks. This may have affected the performance of the passive samplers because of possible clogging and desorption processes.

4.2 Evaluation of time-proportional and grab sampling strategies

In the present study, time-proportional composite sampling was used as the best estimate of the true load. However, in dynamic systems such as streams, flow-proportional composite sampling is conceptually a better approach to estimate nutrient fluxes (Abtew and Powell, 2004; Ort et al., 2010). The advantages of flow-proportional composite sampling versus time-proportional composite sampling were compared in a study conducted in three small-sized streams in Norway (Haraldsen and Stålnacke, 2006). The results showed that annual nitrate-N loads were highest when calculated from flow-proportional sampling in two of the streams (0.4–7.2 %) but lower in the third stream (20.4 %) compared to time-proportional sampling. For total P, one of the streams had a higher annual transport for flow-proportional than for time-proportional sampling (38.4 %), whereas transport was lower for flow-proportional sampling in two of the streams (8.2–9.6 %). The use of time-proportional composite sampling (hourly sampling) against a flow-proportional sampling programme has been evaluated in a smaller Danish stream based on a Monte Carlo evaluation of the bias and precision (standard deviation) of the two methods utilising a 1-year sampling effort with 2300 single measurements of the concentration of total phosphorus (Kronvang and Iversen, 2002). The estimated annual total P load from time-proportional sampling had a higher bias (−12.2 %), than the annual load calculated based on flow-proportional sampling (−0.2 %). Both sampling methods showed, however, a nearly similar precision (standard deviation: 0.8 % for time-proportional and 0.3 % for flow-proportional sampling). Therefore, flow-proportional sampling is superior to time-proportional sampling in delivering more unbiased load estimates of total P. A similar conclusion is, however, not to be expected in the case of total N because of the more smoothed concentration pattern during the year and the absence of spikes (Kronvang and Bruhn, 1996).

We therefore find it safe to conclude that time-proportional composite sampling in the case of both total N and P yields precise load estimates, but with a lower accuracy (more bias) in the case of total P than flow-proportional sampling. The accuracy of the load estimates of especially total P will, however, be strongly dependent on the stream monitored regarding its hydrological regime and P pathways (Kronvang and Bruhn, 1996; Haraldsen and Stålnacke, 2006; Jordan and Cassidy, 2011).

In a study of two smaller streams in Denmark, Kronvang and Bruhn (1996) found an RMSE of 1.1–5.4 % for total N and 10.5–20.2 % for total P using fortnightly grab sampling, increasing to 4.4–5.3 % for total N and 16.9–28.7 for total P with monthly grab sampling when compared to high frequency sampling (4 to 24 h interval). In another study of the River Loire in France, Moatar and Meybeck (2005) compared monthly grab sampling against high frequency sampling (1–4 day intervals) and found the RMSE of the annual load to be 6 % for nitrate and 9 % for TP. These values were much lower than in our study showing RMSE values of 17–23 % for total N and 53–54 % for total P for fortnightly grab sampling. A likely explanation may be that the small streams investigated in our study exhibited a more dynamic pattern in nutrient concentrations than larger rivers such as the River Loire.

Some common features emerge from our study and those previously conducted on sampling methodology and transport estimation: (1) grab sampling nearly always underestimates the “true” annual loads of total P and has high RMSE values (Tables 2 and 3); (2) grab sampling may both underestimate and overestimate “true” annual loads of N (Tables 2 and 3); and (3) use of SC-samplers did not improve the annual load estimates for either N or P in our two investigated streams.

4.3 Method costs

The per-site monitoring costs of the three different sampling methods reveal almost identical costs per year for use of SC-samplers and time-proportional sampling. Hence, economic considerations do not change the conclusion that time-proportional sampling seems preferable to passive sampling. This may, though, change in the future if the passive sampling method can be improved to enhance measurement accuracy, rendering duplicate measurements unnecessary. Important advantages of the passive sampling method are the absence of investment costs and its flexibility in allowing easy relocation of monitoring stations. Comparison of time-proportional sampling with grab sampling provides a less clear choice – time-proportional sampling was still the most reliable method, but the difference was not as pronounced as for the passive sampling method. There is, though, a substantial difference in costs, and this – combined with the other advantages of grab-sampling in terms of investment costs and flexibility – suggests that grab sampling may, in some situations, be the best choice.

4.4 Implications for the costs of river management plans and the implementation of the WFD

Both over- and underestimation of nutrient concentrations may have serious implications for the magnitude of the costs involved in meeting the load reduction targets specified by the WFD. Regarding method measurement certainty, we have previously mentioned that passive sampling overestimated the annual N load by 47 % and underestimated the annual P load by 43 % at Odderbaek using the time-proportional method as reference. For Gelbaek, both N and P were underestimated by the passive sampling method. These over- and/or underestimations of the true nutrient concentrations by passive samplers may have significant – both economic and environmental – implications if the passive sampling method is used as the base for WFD implementation.

If nutrient concentrations are overestimated (i.e. the measured concentrations exceed the true concentration), the need for reduction of nutrient emissions will be overestimated too; the current status will thus appear to be farther away from the target of good ecological status than actually is the case. This may lead to over-implementation of mitigation measures. Seen from a strictly environmental point of view this would be positive in that the ecological status would be improved to a status even better than “good”, but from a welfare-economic point of view this would be a wasteful expenditure of society’s resources. In contrast, if nutrient concentrations are underestimated, also the need for additional mitigation measures will be underestimated, likely leading to non-compliance with the requirements of the WFD. Seen from an

Table 4. Advantages and disadvantages of the three nutrient monitoring methods tested in the present study.

Method	Advantages	Disadvantages
Passive sampler	– Flow integrated (i.e. continuous sampling over time relative to flow conditions in the stream)	– Lack of documentation – Reliability (still under development) – Difficult to compare P results with other international standards for filtration and analysis – Costs – Malfunctions with loss of data
Grab sampling	– Fast – Simple – Inexpensive (only a bottle + analysis)	– Representative only for the conditions at the time of sampling; thus, short-lasting peak flow events are most often not represented. If they are a false signal for a too long/for a prolonged period is obtained when utilising linear interpolation between each grab sample in time.
Time proportional sampling	– Time integrated	– Equipment costs – Power supply required – Maintenance – Malfunctions with loss of data

ecological point of view this is not desirable, as the ecological condition will not be sufficiently improved; seen from an economic point of view, costs will be reduced, which may seem advantageous from a farmer's perspective; but from a welfare-economic (society's) aspect, assuming that the set target reflects society's preferences, this will entail damage (or resource) costs and inefficient use of society's resources.

If mitigation efforts are based on erroneous estimates of nutrient concentrations, implications may be severe and vary significantly from case to case depending on the required reduction (i.e. the current state) and the availability or feasibility of employing different mitigation measures. To illustrate the extent of the costs, for Ringkoebing catchment, the recipient for Odderbaek, the average cost of N reduction is estimated to EUR 5 per kilogramme N (Jacobsen, 2012) and the total costs of achieving the required reduction are estimated to EUR 2.2 million per year (Jacobsen, 2012). If N loads at all monitoring sites in the catchment are assumed to be overestimated by the 47 % observed at Odderbaek, total annual costs for attaining the N target for Ringkoebing fjord would increase to EUR 2.9 million per year. The more specific consequences will vary between catchments, as the load reduction targets and the costs per kilogramme reduction are dissimilar due to differences in loads and production and in the feasibility of implementing low-cost measures. For another Danish catchment, the Limfjorden catchment, for which load reductions requirements are higher and the estimated costs per kilogramme N almost twice as high, the costs would increase from EUR 40 to 57 million per year. The fact that the mitigation costs are not linear but most often marginally increasing supports our conclusion (Hasler et al., 2014). Although a 47 % overestimation of N loads cannot be assumed for all sites, the example shows that significant costs may arise from basing WFD implementation on incorrect measurement results. As illustrated above, the costs of overestimation are fairly straightforward to assess as they may be expressed in terms of the costs of the measures that are implemented in excess of the measures necessary for meeting the target. In contrast, the costs of underestimation are more difficult to calculate, as there are no readily available prices of the damage costs of 1 kg N (as the damage of 1 kg N varies between recipients). Underestimation results in failure to meet the ecological target, and this may be seen as equivalent to failure to obtain the level of environmental quality given by the difference between the set target and the achieved target. Valuation studies assessing the value of achieving different levels of ecological status, including "good", are available (Jørgensen et al., 2013) where the value may be interpreted as the value lost, or damage cost, incurred if good ecological status is not achieved. The results of these studies cannot, however, be readily transferred to ours, and since no valuation studies have been performed for Ringkoebing fjord, we will not attempt to estimate the potential costs associated with underestimation of nutrient concentrations.

5 Conclusions

No definite conclusions can be drawn regarding best measurement practices based on the cost assessments made in this study, but several important points have arisen that are worth contemplating. Thus, we found that monitoring costs vary significantly between methods but that there was no clear relationship between costs and quality. When comparing passive sampling with time-proportional sampling, the superiority of time-proportional sampling is fairly obvious, whereas the differences between passive sampling and grab sampling are less clear – which method is the best depends on the specific situation. More importantly, our analysis illustrates that monitoring costs are likely much lower than mitigation costs. Consequently, one should be careful not to put much focus on monitoring-related cost savings, particularly if these entail decreased measurement certainty. Hence, the welfare-economic costs incurred by basing mitigation efforts on erroneous measuring results probably greatly exceed monitoring cost savings.

To synthesise our findings, we present a summary table of the advantages and disadvantages associated with the three sampling methodologies studied (Table 4). As can be seen, if time-proportional sampling is not feasible, for instance due to the relatively high costs, grab sampling should be favoured over passive samplers, as further development is required to make them a reliable nutrient sampling alternative. The resources spent on increasing the reliability and certainty of monitoring results save implementation costs that are far higher than the monitoring costs.

Acknowledgements. The authors are grateful for assistance from the technicians at Aarhus University. This work is part of the MONITECH project supported by the Danish Council for Strategic Research (grant 2104-08-0050). J. Audet received funding from the Carlsberg Foundation during his stay at SLU, Uppsala.

Edited by: F. van Gee

References

Abtew, W. and Powell, B.: Water quality sampling schemes for variable flow canals at remote sites, J. Am. Water Resour. Assoc., 40, 1197–1204, doi:10.1111/j.1752-1688.2004.tb01579.x, 2004.

Audet, J., Elsgaard, L., Kjaergaard, C., Larsen, S. E., and Hoffmann, C. C.: Greenhouse gas emissions from a Danish riparian wetland before and after restoration, Ecol. Eng., 57, 170–182, doi:10.1016/j.ecoleng.2013.04.021, 2013.

Birgand, F., Skaggs, R. W., Chescheir, G. M., and Gilliam, J. W.: Nitrogen removal in streams of agricultural catchments - A literature review, Crit. Rev. Environ. Sci. Technol., 37, 381–487, doi:10.1080/10643380600966426, 2007.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H.: Nonpoint pollution of surface waters with phosphorus and nitrogen, Ecol. Appl., 8, 559–568, 1998.

de Jonge, H. and Rothenberg, G.: New Device and Method for Flux-Proportional Sampling of Mobile Solutes in Soil and Groundwater, Environ. Sci. Technol., 39, 274–282, doi:10.1021/es049698x, 2004.

Dolan, D. M., Yui, A. K., and Geist, R. D.: Evaluation of River Load Estimation Methods for Total Phosphorus, J. Great Lakes Res., 7, 207–214, doi:10.1016/S0380-1330(81)72047-1, 1981.

Haraldsen, T. K. and Stålnacke, P.: Methods for water quality sampling and load estimation in monitoring of Norwegian agricultural catchments, Nord. Hydrol., 37, 81–92, 2006.

Hasler, B., Smart, J. C. R., Fonnesbech-Wulff, A., Andersen, H. E., Thodsen, H., Mathiesen, G., Smedberg, E., Göke, C., Czajkowski, M., Was, A., Elofsson, K., Humborg, C., Wolfsberg, A., and Wulff, F.: Hydro-economic modelling of cost-effective transboundary water quality management in the Baltic Sea, Water Resour. Econ., 5, 1–23, doi:10.1016/j.wre.2014.05.001, 2014.

Jacobsen, B. H.: Foreløbig analyse af landbrugets omkostninger ved en yderligere reduktion på 10.000 tons N (Preliminary analysis of the costs related to a further reduction of 10000 tonne N), Note for the N-committee under the Ministry of Finance, Copenhagen University, 2012.

Jensen, C. L., Jacobsen, B. H., Olsen, S. B., Dubgaard, A., and Hasler, B.: A practical CBA-based screening procedure for identification of river basins where the costs of fulfilling the WFD requirements may be disproportionate – applied to the case of Denmark, J. Environ. Econ. Policy, 2, 164–200, doi:10.1080/21606544.2013.785676, 2013.

Johansson, P. O.: Cost-Benefit Analysis of Environmental Change, Cambridge University Press, Cambridge, UK, 1993.

Jordan, P. and Cassidy, R.: Technical Note: Assessing a 24/7 solution for monitoring water quality loads in small river catchments, Hydrol. Earth Syst. Sci., 15, 3093–3100, doi:10.5194/hess-15-3093-2011, 2011.

Jordan, P., Cassidy, R., Macintosh, K. A., and Arnscheidt, J.: Field and Laboratory Tests of Flow-Proportional Passive Samplers for Determining Average Phosphorus and Nitrogen Concentration in Rivers, Environ. Sci. Technol., 47, 2331–2338, doi:10.1021/es304108e, 2013.

Jørgensen, S. L., Olsen, S. B., Ladenburg, J., Martinsen, L., Svenningsen, S. R., and Hasler, B.: Spatially induced disparities in users' and non-users' WTP for water quality improvements – Testing the effect of multiple substitutes and distance decay, Ecol. Econ., 92, 58–66, doi:10.1016/j.ecolecon.2012.07.015, 2013.

Kronvang, B. and Bruhn, A. J.: Choice of sampling strategy and estimation method for calculating nitrogen and phosphorus transport in small lowland streams, Hydrol. Process., 10, 1483–1501, 1996.

Kronvang, B., Ærtebjerg, G., Grant, R., Kristensen, P., Hovmand, M., and Kirkegaard, J.: Nationwide Monitoring of Nutrients and Their Ecological Effects: State of the Danish Aquatic Environment, Ambio, 22, 176–187, 1993.

Kronvang, B., Grant, R., and Laubel, A. L.: Sediment and phosphorus export from a lowland catchment: Quantification of sources, Water Air Soil Poll., 99, 465–476, 1997.

Kronvang, B. and Iversen, H. L.: Experiences with different sampling procedures in Denmark, in: Monitoring of agricultural catchment in the Nordic countries –Methods for water quality sampling and load estimation, edited by: Haraldsen, T. K. and Stålnacke, P., Nordic Council of Ministers – Environment, 749, 43–51, 2002.

Meybeck, M.: Carbon, Nitrogen, and Phosphorus Transport by World Rivers, Am. J. Sci., 282, 401–450, 1982.

Moatar, F. and Meybeck, M.: Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire, Hydrol. Process., 19, 429–444, doi:10.1002/hyp.5541, 2005.

Moss, B.: Water pollution by agriculture, Philos. T. Roy. Soc. Lond. B, 363, 659–666, doi:10.1098/rstb.2007.2176, 2008.

Ort, C., Lawrence, M. G., Rierckermann, J., and Joss, A.: Sampling for Pharmaceuticals and Personal Care Products (PPCPs) and Illicit Drugs in Wastewater Systems: Are Your Conclusions Valid? A Critical Review, Environ. Sci. Technol., 44, 6024–6035, doi:10.1021/es100779n, 2010.

Ovesen, N. B., Iversen, H. I., Larsen, S. E., Müller-Wohlfeil, D. I., Svendsen, L. M., Blicher, A. S., and Jensen, P. M.: Afstrømningsforhold i danske vandløb, National Environmental Research Institute, Aarhus University, Silkeborg, Denmark, 238 pp., 2000.

Rozemeijer, J., van der Velde, Y. P., de Jonge, H., van Geer, F., Broers, H. P., and Bierkens, M.: Application and evaluation of a new passive sampler for measuring average solute concentrations in a catchment scale water quality monitoring study, Environ. Sci. Technol., 44, 1353–1359, doi:10.1021/es903068h, 2010.

Seitzinger, S. P., Kroeze, C., Bouwman, A. F., Caraco, N., Dentener, F., and Styles, R. V.: Global patterns of dissolved inorganic and particulate nitrogen inputs to coastal systems: Recent conditions and future projections, Estuaries, 25, 640–655, doi:10.1007/BF02804897, 2002.

Smith, V. H., Tilman, G. D., and Nekola, J. C.: Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems, Environ. Pollut., 100, 179–196, doi:10.1016/s0269-7491(99)00091-3, 1999.

Strahler, A. N.: Quantitative analysis of watershed geomorphology, Transactions American Geophysical Union, 38, 913–920, 1957.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H., and Tilman, D. G.: Human alteration of the global nitrogen cycle: Sources and consequences, Ecol. Appl., 7, 737–750, 1997.