



# Map-based prediction of organic carbon in headwater streams improved by downstream observations from the river outlet

J. Temnerud<sup>1,2</sup>, C. von Brömssen<sup>3</sup>, J. Fölster<sup>1</sup>, I. Buffam<sup>4</sup>, J.-O. Andersson<sup>5</sup>, L. Nyberg<sup>6</sup>, and K. Bishop<sup>1,7</sup>

<sup>1</sup>Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden

<sup>2</sup>Research Department, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

<sup>3</sup>Unit of applied statistics and mathematics, Department of Economics, Swedish University of Agricultural Sciences, Uppsala, Sweden

<sup>4</sup>Department of Biological Sciences and Department of Geography, University of Cincinnati, Cincinnati, Ohio, USA

<sup>5</sup>Department of Geography, Karlstad University, Karlstad, Sweden

<sup>6</sup>Centre for Climate and Safety, Karlstad University, Karlstad, Sweden

<sup>7</sup>Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Correspondence to: J. Temnerud (johan.temnerud@slu.se)

Received: 25 May 2015 – Published in Biogeosciences Discuss.: 16 June 2015

Revised: 15 December 2015 – Accepted: 28 December 2015 – Published: 21 January 2016

**Abstract.** In spite of the great abundance and ecological importance of headwater streams, managers are usually limited by a lack of information about water chemistry in these headwaters. In this study we test whether river outlet chemistry can be used as an additional source of information to improve the prediction of the chemistry of upstream headwaters (size  $< 2 \text{ km}^2$ ), relative to models based on map information alone. We use the concentration of total organic carbon (TOC), an important stream ecosystem parameter, as the target for our study. Between 2000 and 2008, we carried out 17 synoptic surveys in 9 mesoscale catchments (size  $32\text{--}235 \text{ km}^2$ ). Over 900 water samples were collected in total, primarily from headwater streams but also including each catchment's river outlet during every survey. First we used partial least square regression (PLS) to model the distribution (median, interquartile range (IQR)) of headwater stream TOC for a given catchment, based on a large number of candidate variables including sub-catchment characteristics from GIS, and measured river chemistry at the catchment outlet. The best candidate variables from the PLS models were then used in hierarchical linear mixed models (MM) to model TOC in individual headwater streams. Three predictor variables were consistently selected for the MM calibration sets: (1) proportion of forested wetlands in the sub-catchment (positively correlated with headwater stream TOC), (2) propor-

tion of lake surface cover in the sub-catchment (negatively correlated with headwater stream TOC), and (3) river outlet TOC (positively correlated with headwater stream TOC). Including river outlet TOC improved predictions, with 5–15 % lower prediction errors than when using map information alone. Thus, data on water chemistry measured at river outlets offer information which can complement GIS-based modelling of headwater stream chemistry.

## 1 Introduction

Headwaters make up most of the watercourse length and hence provide a large proportion of the lotic habitat in a landscape (Meyer et al., 2007). The headwaters also provide much of the water and solutes to downstream locations (Person et al., 1936; Leopold et al., 1964). It is widely known that variability in water chemistry changes with catchment size, typically with small watercourses showing the highest variability in space (Wolock et al., 1997; Temnerud and Bishop, 2005) and time (Nagorski et al., 2003; Buffam et al., 2007). Significant field sampling efforts (Hutchins et al., 1999; Smart et al., 2001; Likens and Buso, 2006; McGuire et al., 2014) have been made to quantify the variability of

headwaters in individual catchments. Readily derived GIS data from maps and satellite images have been used to model some chemical constituents in larger rivers (Alexander et al., 2007), but are seldom effective at predicting the chemistry of individual headwater streams (Strayer et al., 2003b, a; Temnerud et al., 2010). This is presumably in large part due to the greater importance of small-scale heterogeneity in headwater catchment characteristics, as compared to riverine catchments where much of the variability averages out at larger spatial scales (Gomi et al., 2002; MacDonald and Coe, 2007).

Since aquatic monitoring activities are generally located at downstream sites (Evans et al., 2010), this might provide information about the headwaters upstream from the monitoring sites. In an attempt to use environmental monitoring data to predict seldom assessed headwater streams, the chemistry at the river outlet was used by Temnerud et al. (2010) to predict the median and interquartile range (IQR) of several environmentally relevant stream chemistry parameters, including total organic carbon (TOC), acid neutralising capacity (ANC) and pH. This demonstrated that the river outlets were correlated to statistical features of the upstream population of headwaters. In that study significant relationships were found for ANC, pH and TOC between headwaters median and IQR vs the river outlets, with the strongest relationships for ANC, and the weakest for TOC. Of seven different leave-one-out attempts one model was significant for headwater median TOC and none for headwater TOC IQR (Temnerud et al., 2010). No map information was employed in that study.

In this study the goal was to test whether map information can be combined with river outlet chemistry to predict TOC in individual headwaters. More specifically, would the combination of map and river outlet chemistry give a better prediction than either one used separately. In this follow-up study we have chosen to focus solely on the prediction of TOC, for two main reasons. First, TOC is of great ecological importance for boreal and other watercourses because of its influence on pH, buffering capacity, nutrient bioavailability, metal and pollutants transport, light climate, microbial productivity, and carbon cycling (Wetzel, 2001; Schwarzenbach et al., 2003). Secondly, the statistical distribution of headwater TOC was not well predicted in the previous study using only downstream chemistry as the predictor (Temnerud et al., 2010). If the approach succeeds with TOC, then there is reason to hope that it would be even more effective in predicting other aspects of water chemistry.

An important aspect of modelling headwaters is that the spatial variation is largely dependent on temporal factors, often flow-related (Buffam et al., 2007), but also season (temperature and precipitation) and even long-term trends (Hytteborn et al., 2015). This temporal variation within a single headwater can be greater than the variation of TOC between catchments in the same biome. We want to make the reader aware that it could be easier to model headwaters in different

catchments (at the same time) than headwaters in the same catchments that are sampled at different times.

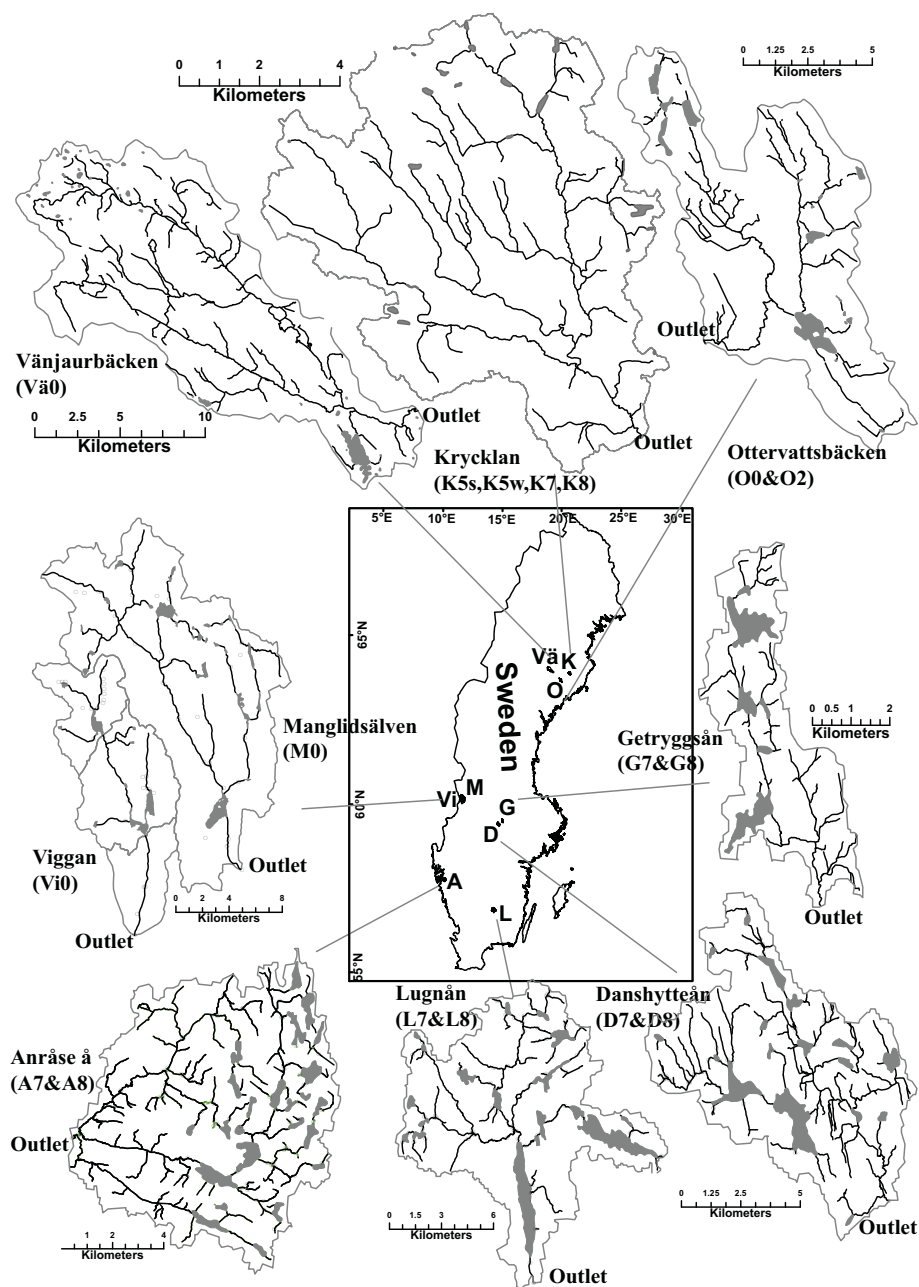
Thus we modelled headwater TOC concentration's on several occasions in different stream networks using catchment map information (Andersson and Nyberg, 2009) complemented by data on the river outlet TOC concentration.

While this may seem straight-forward, there are in fact some theoretical challenges; the method must deal with both strong correlations between observations and between explanatory variables. We used a two-step modelling approach to handle these challenges. First we used partial least square regression (PLS), which can deal with strong correlations between explanatory variables, to model headwater stream TOC median and IQR-values for the mesoscale catchments from information on catchment land cover, geology, soil type maps, vegetation and river outlet chemistry. The best explanatory variables from the PLS models were then used as candidate variables for hierarchical linear mixed models (MM). Such MM have the advantage of being able to deal with strong correlations between observations (such as within a mesoscale catchment stream network), but MM are not appropriate for large numbers of explanatory variables. So PLS modelling was used first to narrow the number of explanatory variables, and then MM was used in the next step to model individual headwater streams. Thus, MM can account for the clustered data structure of catchment properties in a drainage network (Littell et al., 2006). Mixed models have been used rather successfully in related types of data evaluation (Jager et al., 2011; Sakamaki and Richardson, 2013). Two distinguishing features of our study are that (i) we test a combination of information from maps (GIS) with direct measurements of chemistry at the river outlet to create models of individual headwaters and (ii) we test our models on data that were not used in model calibrations.

## 2 Methods

### 2.1 Sampling approach

The synoptic surveys of mesoscale catchments used in this study were designed to provide a snapshot of the water chemistry in stream networks (Table 1, Fig. 1), with a strategy of sampling most watercourse junctions, lake inlets and outlets as well as the river outlets from each mesoscale catchment. In total there were data from 17 synoptic surveys conducted between 2000 and 2008 in nine catchments distributed across Sweden, spanning a north-south gradient of 800 km through the north-temperate and boreal zones (Fig. 1). This data set amounted to 938 stream samples of which 420 were from headwaters. Headwaters are defined as first order streams with catchments smaller than 2 km<sup>2</sup> in each of the nine drainage networks sampled (Table 1). The number of sampled headwaters differs between surveys ( $n_{HW}$  in Table 2). All sampling during a given survey was carried out during a



**Figure 1.** Map of Sweden with the nine investigated catchments, see Table 1 for coordinates. Labels in brackets are the abbreviated names of the catchments and year of sampling.

1 to 3-day period (Table 2 and Fig. S1 in the Supplement), except for R. Krycklan in winter 2005, which took 2 weeks due to cold weather and difficulties finding the streams in deep snow. Stable base flow discharge was maintained throughout that winter sampling period, so that survey was still effectively a snapshot in time.

Five of the nine catchments contain at least one headwater stream site that has been monitored for runoff and chemistry regularly for a decade or more (Edström and Rystam, 1994; Temnerud et al., 2009; Köhler et al., 2008; Löfgren

et al., 2011; Laudon et al., 2013). Of the other four catchments, R. Ottervatnsbäcken was sampled twice and R. Vänjaurbäcken sampled once (the name R. Sörbäcken was used in the references, as this is a tributary in R. Vänjaurbäcken, Temnerud and Bishop, 2005; Temnerud et al., 2007) while R. Viggan and R. Manglidsälven were sampled once, for this study.

**Table 1.** Characteristics of the rivers. Cluster is the grouping of calibration and test sets used in the modelling. Air temperature ( $T$  °C) and specific discharge ( $q$  mm day<sup>-1</sup>) are the medians of daily values for 1990–2010, precipitation ( $P$  mm) is the median of the yearly sum.  $P$ ,  $T$  and  $q$  are modelled by SMHI, see Methods for more details.

Cluster	River	Lat. N & Long E	Size (km <sup>2</sup> )	$T$	$P$	$q$
c2	Anråse å <sup>a</sup> (A)	58°01'; 11°51'	74	7.8	983	0.95
c2	Danshytteån <sup>a</sup> (D)	59°42'; 15°05'	72	6.1	774	0.70
c2	Getryggsån <sup>a</sup> (G)	59°48'; 15°17'	32	5.9	800	0.84
c2, c3	Krycklan <sup>b</sup> (K)	64°14'; 19°46'	61	2.6	659	0.51
c2	Lugnån <sup>a</sup> (L)	57°06'; 14°48'	122	6.1	831	0.60
c1	Mangslidsälven (M)	60°23'; 12°54'	235	4.9	823	0.75
c1, c3	Ottervattsbäcken <sup>c</sup> (O)	64°02'; 19°06'	71	2.6	659	0.50
c1	Vänjaurbäcken <sup>c</sup> (Vä)	64°19'; 18°43'	200	2.2	649	0.47
c1	Viggen (Vi)	60°21'; 12°46'	116	3.9	870	0.85

<sup>a</sup> Temnerud et al. (2009); <sup>b</sup> Laudon et al. (2013); <sup>c</sup> Temnerud and Bishop (2005).

**Table 2.** Median concentration of total organic carbon (TOC, mg L<sup>-1</sup>), with 25th and 75th percentiles in brackets, plus weather parameters ( $T$ ,  $P$ ,  $q$ ) using the median value from the 30 days prior to sampling as modelled by SMHI (See Methods for more details). Clusters are the different groups of calibration and test sets used in the modelling. Sets are the different data sets used, with acronyms corresponding to river names (Table 1) and the last digit of the sampling year. M = Month of sampling, HW is headwaters and  $n_{HW}$  = number of sampled HW.

Cluster	Set	Year	M	$n_{HW}$	TOC <sub>HW</sub>	OutletTOC	$T$	$P$	$q$
c2	A7	2007	10	45	8.6 (5.8–12)	6.7	11	4.3	1.3
c2	A8	2008	4	45	7.7 (6.6–9.9)	6.9	3.9	2.8	4.4
c2	D7	2007	10	34	16 (13–29)	11	9.9	2.0	0.41
c2	D8	2008	4	33	12 (9.6–17)	10	1.7	1.2	1.2
c2	G7	2007	10	21	27 (18–36)	12	9.1	3.2	0.63
c2	G8	2008	4	22	18 (13–20)	8.9	1.3	0.95	1.2
c3	K5s	2005	6	24	15 (12–18)	10	13	1.8	0.69
c3	K5w	2005	2	17	12 (6.9–17)	5.0	-5.1	0.75	0.16
c2	K7	2007	7	12	12 (9.7–15)	4.3	14	2.1	0.26
c2	K8	2008	9	22	20 (16–25)	15	11	4.7	1.6
c2	L7	2007	10	26	20 (16–32)	15	7.9	1.9	0.91
c2	L8	2008	4	26	15 (10–20)	12	0.7	1.2	1.3
c1	M0	2000	8	7	19 (13–22)	14	15	1.8	1.3
c1	O0	2000	6	31	20 (16–27)	15	9.7	1.4	1.3
c3	O2	2002	8	24	20 (14–32)	15	17	3.0	0.21
c1	Vä0	2000	6	18	12 (10–15)	9.5	9.9	2.2	1.8
c1	Vi0	2000	8	13	16 (9.0–19)	15	14	2.2	1.5

## 2.2 Study sites

All sampled catchments (including headwaters, intermediate watercourses and river outlets) consisted mainly of forest (> 80 %) with a dominance of coniferous forest made up of Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) (Table S1). Mires and small humic lakes made up most of the remaining parts of the catchments, while the proportion of agricultural and developed areas were minimal (< 1 %). The mean annual air temperature in the river catchments (1990–2010) ranged from 7.8 °C in the southernmost river, R. Anråse å, to 2.6 °C in the northernmost, Krycklan. Mean annual precipitation (1990–2010) ranged from 980 mm at R. Anråse å, to 649 mm at R. Vänjaurbäcken.

Daily mean air temperature, daily precipitation and daily runoff for 1961–2010 at each river outlet were modelled by the Swedish Meteorological and Hydrological Institute (SMHI) (Johansson, 2000; Johansson and Chen, 2003), data received 20th of June 2013 from <http://luftweb.smhi.se>. Daily runoff for 1990–2010 at each river outlet was modelled based on Hydrological Predictions for the Environment program (HYPE) (Lindström et al., 2010), data received 20th of June 2013 from <http://vattenweb.smhi.se>. Typical accuracy in the HYPE modelling for small catchments (< 200 km<sup>2</sup>) is 10 % (Strömqvist et al., 2012; Arheimer and Lindström, 2013). Catchment-specific daily mean air temperature, total precipitation and mean specific discharge are illustrated

(Fig. S1) for the period of 30 days up to and including each sampling.

### 2.3 Map information

To relate headwater TOC to catchment characteristics, we began with 34 catchment parameters (Table S1, Figs. S3–S8) taken from the Swedish land cover data map (SMD), year 2000, version 2.1, which is based on the CORINE database (Bossard et al., 2000) as well as Geology and Quaternary deposits from the Geological Survey of Sweden (SGU) map, scale 1.1 million. The kNN-database of vegetation that has forestry variables estimated from LANDSAT 5 and LANDSAT 7 satellite photos (version year 2000, Reese et al., 2002, 2003), provided data about the average age and height of the forest as well as volume of the biomass for different tree species. All catchment map information uses the same version, year and scale of the maps so that the map data are commensurate between catchments (Table S1 in the Supplement).

### 2.4 Chemical analyses

After collection, all water samples were kept dark and cool until they were analysed. Total organic carbon (TOC) was measured by combustion and analysis as CO<sub>2</sub> using a Shimadzu TOC-VPCH analyser after acidification and sparging to remove inorganic carbon. Dissolved organic carbon (DOC) is the concentration of organic carbon in a filtered water sample (common cut-off is 0.45 µm). It has previously been shown that DOC and TOC differ on average by less than 5 % (Ivarsson and Jansson, 1994; Köhler, 1999), so TOC is essentially identical to DOC in the large majority of the Swedish surface waters (see also Gadmar et al., 2002; Laudon et al., 2011).

### 2.5 Statistical analysis

The main objective of this article is to model the TOC of individual headwaters based on map information and river outlet TOC. We use the following two-step approach where we first find the best map variables for predicting the average level and distribution around that level using PLS. Secondly, we use MM to predict individual headwaters within each catchments.

We are particularly interested in determining whether models based on geographical data, like lake surface coverage, forest coverage or altitude can be improved by adding data on TOC concentrations measured at the river outlet.

#### 2.5.1 Modelling headwater median and interquartile range

To model the median and interquartile range of TOC in headwaters in different catchments we use partial least squares regression (PLS). Variables included in this model are TOC

at the river outlet (OutletTOC) and a number of variables describing information derived from land cover, geology, soil type maps and vegetation (kNN) (Table S1). The main purpose of PLS was to narrow down the number of explanatory variables for subsequent use in the mixed-models approach.

All data, both explanatory and response variables, were centred by mean normalisation and weighted by dividing the variables with the standard deviation prior to PLS analysis in SIMCA for Windows v13.0 (Umetrics). PLS identifies the relationship between explanatory variables and response variables through a linear model, and is less sensitive to correlated explanatory variables (so-called multicollinearity) when compared to multiple linear regression approaches (Geladi and Kowalski, 1986), since the explanatory variables are combined to factors.

In the PLS analyses, the goodness-of-fit parameter  $Q^2$  was used to quantify the model performance, which is the average ( $n = 7$ , default value in SIMCA) explained variance of a randomly selected fraction (1/7 of the data) of the validation data not used to fit the model. In robust models,  $R^2$  and  $Q^2$  are often similar, but the latter will decline as models become increasingly over-fit. Even though PLS models work by defining factors, i.e. combinations of explanatory variables, it is also possible to compute coefficients and weights that describe the direction and relative strength of the individual explanatory variables on the response variable; weights with larger absolute values indicate greater importance to a given latent component. All PLS-models were refined by iteratively removing variables that had non-significant coefficients. This procedure served to minimise the difference between  $R^2$  and  $Q^2$  values while retaining high explanatory power, i.e. to find a model that can be generalized to new data, while retaining good explanatory power. The relative importance of each explanatory variable is ranked using variable importance on the projection (VIP) scores, derived as the sum of square of the PLS weights across all components. VIP values greater than one are considered to indicate variables that are most important to the overall model (Eriksson et al., 2006).

PLS allows for more explanatory variables than observations and gives us the possibility to include many candidate variables. Still, some of the variables available needed to be excluded for the following reasons:

- Some variables, e.g. volume of oak, had zero value for all observations or only few observations different from zero. These variables could not be included due to the lack of variation.
- Geographical variables for the river outlet were not included, since they correlate highly with the median of the corresponding variable at the headwater scale. The latter is considered to bear more information and was therefore included.

The analysis was run on three calibration data sets: one with data from year 2000 where headwaters within four mesoscale catchments (these catchments are called c1) were sampled Cal<sub>PLS00c1</sub> ( $n = 4$ ; M0, O0, Vå0, Vi0). In Cal<sub>PLS07c2</sub> the headwaters of five mesoscale catchments (catchments c2) were sampled in 2007 ( $n = 5$ ; A7, D7, G7, K7, L7) and in Cal<sub>PLS08c2</sub>, the same five mesoscale catchments as in Cal<sub>PLS07</sub>, were sampled in 2008 ( $n = 5$ ; A8, D8, G8, K8, L8). The test data sets consist of the same groupings, but were not used in the calibrations (Test<sub>PLS00c1</sub>, Test<sub>PLS07c2</sub>, Test<sub>PLS08c2</sub>).

After running the three different PLS models we determined whether the same variables appeared to be important ( $VIP > 1$ ) in the model fittings. These variables were taken as good candidates to reproduce the TOC level of the headwaters and therefore included in the mixed modelling of the individual headwaters in the next step (Sect. 2.5.2). An additional test set was created, Test<sub>PLS02&05c3</sub>, which was comprised of data from the two catchments sampled once in 2002 and twice in 2005 (catchments c3,  $n = 3$ ; O<sub>2</sub>, K5s, K5w, where s stands for summer and w for winter). No calibration was done on this data set, but was used for testing robustness of the models with respect to seasonality.

### 2.5.2 Modelling individual headwaters

When modelling individual headwaters we want to predict specific values for each of the different headwaters in all catchments. As an effort to improve these simulations, we make an assumption that headwaters within the same mesoscale catchment are more similar to each other than to headwaters from other mesoscale catchments due to subtle combinations of physiographic, weather and other factors which combine to influence the TOC levels in ways which are not readily apparent from the available map information, but might be reflected in differences between the average TOC levels in the different mesoscale catchments. This assumption leads to a new data structure. To model this data structure we use hierarchical linear mixed models (MM; Littell et al., 2006), which allow the estimation of the correlation between headwaters within the same mesoscale catchment and adjusts the analysis accordingly. A MM does not allow for highly correlated explanatory variables, so the number of explanatory variables must be substantially smaller than the number of observations. To fit these models we use candidate explanatory variables from the PLS approach described in Sect. 2.5.1. In the PLS analysis some explanatory variables were excluded due to a large number of zero values giving rise to a median sub-catchment value of zero for all catchments. One of these parameters, lake surface area, may still be important for modelling individual headwaters. If lakes have a moderately large volume (appreciable residence time) they are known to influence the organic content (Eriksson, 1929; Birge and Juday, 1926). Therefore lake cover surface is expected to have an influence on the prediction of individual

TOC<sub>HW</sub>, and thus this variable was included as a potential predictor in the MM models even though this variable was not important in the PLS modelling.

MMs were performed using package lme4 (version 1.1-7) in the software R (version 3.1.2) (R Development Core Team, 2014). Headwater data from year 2000 (numbers of headwaters = 69; M0, O0, Vå0, Vi0), 2007 ( $n_{HW} = 138$ ; A7, D7, G7, K7, L7) or 2008 ( $n_{HW} = 148$ ; A8, D8, G8, K8, L8) were used as calibration data sets (denoted Cal<sub>MM00c1</sub>, Cal<sub>MM07c2</sub> and Cal<sub>MM08c2</sub>, respectively), one set at a time. We have two objectives in this approach:

- i. To make predictions on the same headwaters but at different time points. For this observations for 2008 (Test<sub>MM08c2</sub>) were predicted by the model calibrated with data from 2007 (Cal<sub>MM07c2</sub>). The observations from 2007 (Test<sub>MM07c2</sub>) were also predicted based on a model calibrated using the 2008 data (Cal<sub>MM08c2</sub>).
- ii. To make predictions on a new set of headwaters at a different time point. For this:
  - observations from 2002 and 2005 (Test<sub>MM02&05c3</sub>), 2007 (Test<sub>MM07c2</sub>) and 2008 (Test<sub>MM08c2</sub>) respectively were predicted by models calibrated using data from 2000 on an entirely different set of catchments (Cal<sub>MM00c1</sub>) and
  - observations from 2000 (Test<sub>MM00c1</sub>), 2002 and 2005 (Test<sub>MM02&05c3</sub>) were predicted by models calibrated using data from 2007 or 2008 on an entirely different set of catchments (Cal<sub>MM07c2</sub> and Cal<sub>MM08c2</sub>, respectively).

Test<sub>MM02&05</sub> is sampled in one catchment 2002 and two times in another catchment in 2005 ( $n_{HW} = 65$ ; O<sub>2</sub>, K5s, K5w, where s stands for summer and w for winter), no calibration was done on this data set. We used the testing to see if the calibrations worked on other data sets which were not included in the calibration.

To test the impact of including TOC at river outlet (OutletTOC) on the MM performance, three model versions were made from each calibration data set:

- version Out includes OutletTOC but no map information,
- version Map includes map information but not OutletTOC while
- version OutMap includes both OutletTOC and map information.

In total nine different MM were calibrated.

For fitting of the MM, Akaike information criterion (AIC) (Akaike, 1974) and  $p$  values were used. AIC is a goodness-of fit measure, which is corrected for the complexity of the

model, similar to the adjusted  $R^2$ . The  $p$ -values in regression models determine if parameter estimates are significantly different from zero, i.e. if there is a significant relationship between an explanatory variable and the response variable. The  $p$  values were calculated according to Kenward and Roger (1997) using “krmodcomp” in R package “pbkrtest” (version 0.4-1). During the model fitting, added variables were checked to see if they increased the predictive ability of the model by computing the prediction error sum of squares (PRESS). The smaller the PRESS value the closer the prediction is to the observed values. Kenward and Roger (1997) version of  $R^2$  for predictions is called  $P^2$  (similar to  $Q^2$  for PLS). The  $P^2$  were calculated according to Méndez Mediavilla et al. (2008), with the modification that instead of leave-one-out validation we compute  $P^2$  on the evaluation test sets:  $P^2 = 1 - \text{PRESS}/\text{TSS}$  where TSS is the total sum of squared differences between modelled values and the mean of observations in the evaluation set. Median absolute errors (MedAE) and median relative errors (MedRE%) were also calculated.

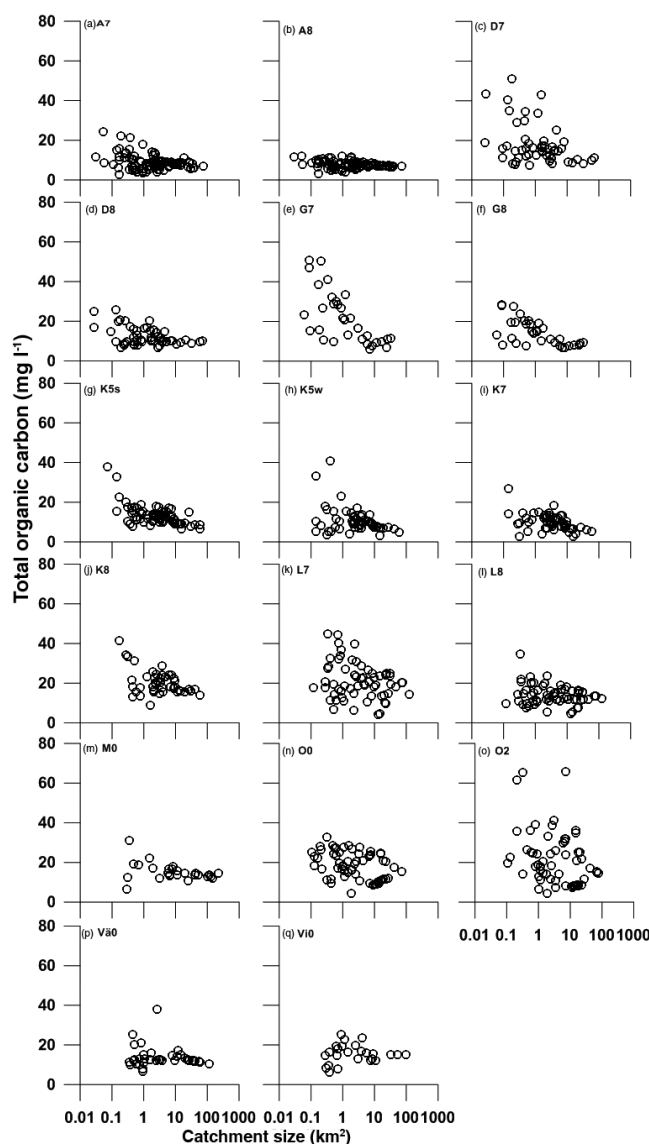
As an additional step in the evaluation of the models the most successful MM from the nine MM calibrations was tested on the sites between the headwaters and the river outlets, the intermediate sites ( $n = 501$ ).

### 3 Results

For all synoptic surveys the median  $\text{TOC}_{\text{HW}}$  values were higher than the values at the respective outlets (Table 2), with large differences ( $> 20\%$ ) for 14 of 17 synoptic surveys. The median  $\text{TOC}_{\text{HW}}$  for all surveys together ( $12 \text{ mg L}^{-1}$ ) was also higher than the median outlet TOC of  $10 \text{ mg L}^{-1}$  (Fig. 2 and Table 2). For all synoptic surveys, except A8, there was a funnel-shape in the TOC concentration with larger variation in smaller catchments that attenuates with increasing catchment size (Fig. 2). Reproducing this variation in the headwaters (Fig. S2), and assigning individual headwaters to the proper value within that large variation, is one of the challenges of modelling water chemistry in a landscape perspective.

#### 3.1 Modelling headwater median

The first principal component (PC) of the PLS-model of median headwater TOC was significant for both PLS calibration sets for the years 2007 and 2008, but not the second PC. No significant PLS-model was established using calibration set for the year 2000 ( $\text{Cal}_{\text{PLS00c1}}$ ). Calibration using data set 2007 ( $\text{Cal}_{\text{PLS07c2}}$ ) gave higher  $R^2$  and  $Q^2$  than using the data set for 2008 calibration ( $\text{Cal}_{\text{PLS08c2}}$ ) (Table 3). The calibration of the model using the  $\text{Cal}_{\text{PLS08c2}}$  data was evaluated by using the test data. This yielded a PRESS for the median TOC that was lower than when the model was calibrated using  $\text{Cal}_{\text{PLS07c2}}$  (Table 3). Based on PLS-modelling



**Figure 2.** Total organic carbon (TOC in  $\text{mg L}^{-1}$ ) as a function of catchment size (note  $\text{Log}_{10}$  scale) for 17 synoptic surveys of 9 catchments; (a) is A7, (b) is A8, (c) is D7, (d) is D8, (e) is G7, (f) is G8, (g) is K5s, (h) is K5w, (i) is K7, (j) is K8, (k) is L7, (l) is L8, (m) is M0, (n) is O0, (o) is O2, (p) is Vå0 and (q) is Vi0. See Tables 1 and 2 for description of surveys.

of median headwater TOC, suitable candidates for the mixed models (MM) of individual headwaters were altitude of sampling sites, OutletTOC, proportion of clear-felled, coniferous forest, mixed forest, wet mires, coniferous forest on mires as well as the volume of birch-, spruce- and total forest volume.

#### 3.2 Modelling headwater interquartile range (IQR)

For both PLS calibration sets year 2007 and 2008 the first principal component (PC) was significant in the PLS-models for IQR headwater TOC, but not the second PC. No signif-



**Table 3.** Partial least square regression (PLS) results indicating the goodness of fit for the prediction of median and the interquartile range (IQR) of headwater total organic carbon (TOC) concentration ( $\text{mg L}^{-1}$ ). Cal is calibrated and 00, 07 and 08 refer to sampling year (2000, 2007 and 2008) of catchment (number =  $n$ ). PRESS is the prediction error sum of squares.

Cal	Var	$n$	$R^2$	$Q^2$	PRESS
00 <sub>c1</sub>	Median	4	ns	ns	ns
	IQR	4	ns	ns	ns
07 <sub>c2</sub>	Median	5	96	94	581
	IQR	5	96	92	346
08 <sub>c2</sub>	Median	5	90	83	192
	IQR	5	54	17	530

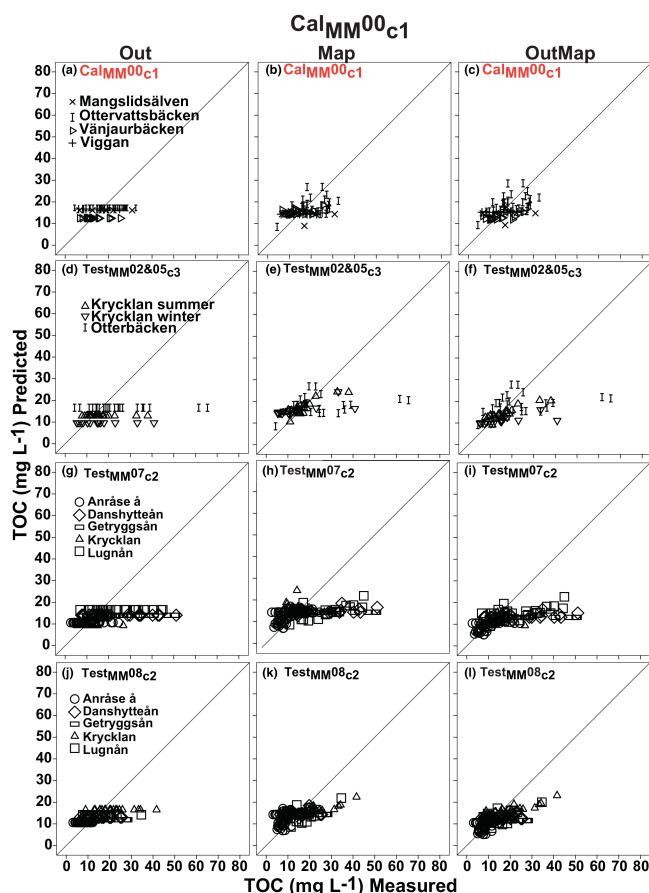
icant PLS-model was established using calibration set year 2000 (Cal<sub>PLS</sub>00<sub>c1</sub>). The calibration of the model using the Cal<sub>PLS</sub>07<sub>c2</sub> data was evaluated by using the test data. This yielded a PRESS for TOC IQR that was lower than when the model was calibrated using Cal<sub>PLS</sub>08<sub>c2</sub> (Table 3). Headwaters TOC IQR modelled by PLS indicates three variables as suitable candidates for the MM: OutletTOC, proportion of clear-felled area and birch volume.

### 3.3 Modelling individual headwaters

The PLS approach from Sect. 3.1 identified the variables that can determine the median level of TOC, and the variation around that median, on a range of different catchments. In the following analysis we seek models to predict the TOC of individual watercourses at different locations and points in time.

### 3.4 Modelling individual headwaters and predicting those same headwaters at other points in time

When we fit the models to a calibration set, e.g. headwaters measured in 2007, we start with a base model consisting of the variables identified by the PLS model for the interquartile range, i.e. OutletTOC on the catchment scale and proportion wet mires as well as volume birch, with different values for different headwaters. The base model was fitted with a MM using catchment as the random factor describing the hierarchical structure. Other variables were included in a forward selection procedure always adding the most significant variable of the remaining set of variables. Candidate variables used in this were all land use variables (including lake surface coverage) and all variables giving the volume of different tree species with exception of the volume of oak and beech, since these volumes are generally very low and zero for many headwaters. After fitting the model the ability to predict new data was tested and non-significant variables in the model were individually removed to check if their removal also worsened the predictive ability of the model.



**Figure 3.** Scatterplots of measured headwaters with total organic carbon (TOC in  $\text{mg L}^{-1}$ ) on the x axis, and the three different versions of the mixed models Cal<sub>MM</sub>00<sub>c1</sub> on the y axis: Out version on the left panel, OutMap version on the right panel and Map in-between. Data for year 2000 (Cal<sub>MM</sub>00<sub>c1</sub>) on the top row in red text, followed by Test data; second row 2002 & 2005 data, third row is 2007 data and the last row is 2008 data. R. Anråse å indicated by circles, R. Danshytteån by diamonds, R. Gettryggsån by rectangles, R. Krycklan by triangles (winter 2005 by upside-down triangles), R. Lugnån by squares, R. Mangslidsälven by multiplication sign, R. Ottervattsbäcken by upside-down triangles, R. Vänjaurbäcken by right tilted triangles and R. Viggan by plus sign. The black line is the 1:1 line.

The models created by this procedure are listed in Table 4. The models produced by Cal<sub>MM</sub>07<sub>c2</sub> and Cal<sub>MM</sub>08<sub>c2</sub> were very similar and can predict the data at the same sites quite well, i.e. the Cal<sub>MM</sub>07<sub>c2</sub> model can predict the Test<sub>MM</sub>08<sub>c2</sub> data set well and vice versa (Fig. 4–5 and Table 5).

#### 3.4.1 Modelling individual headwaters and predicting new headwaters at other time points

When we use the calibration set Cal<sub>MM</sub>00<sub>c1</sub> to fit a model, the variables selected (same procedure as in Sect. 3.3.1) were OutletTOC, lake surface coverage and coniferous forest on mires. We evaluate this model by predicting values in the test



**Table 4.** Coefficients for the best-fit hierarchical linear mixed models (MM), where  $\text{Log}_{10}(\text{TOC}_{\text{HW}})$  is the response variable. See the Method section for more details.

MM model	Version	Intercept	OutletTOC ( $\text{mg L}^{-1}$ )	Lake surface coverage*	Coniferous forest on mires*	Broad-leaved forest*	Sites altitude (m a.s.l.)
Cal00 <sub>c1</sub>	Out	0.861	0.0245				
	Map	1.156		−1.509	0.592		
	OutMap	0.885	0.0201	−1.479	0.568		
Cal07 <sub>c2</sub>	Out	0.822	0.0380				
	Map	1.074		−1.867	1.022	−0.900	0.000619
	OutMap	0.736	0.0363	−1.863	0.956	−0.970	0.000554
Cal08 <sub>c2</sub>	Out	0.671	0.0439				
	Map	0.913		−0.675	1.303	−0.697	0.000903
	OutMap	0.700	0.0199	−0.684	1.243	−0.740	0.000940

\* Is the proportion of the entire catchment area covered by this particular feature.

sets Test<sub>MM07c2</sub> and Test<sub>MM08c2</sub> (Fig. 3). The best prediction model parameters are given in Table 4, with model performance in Fig. 3 and error results in Table 5. Predictions for new sites in the test set Test<sub>MM00c1</sub> and Test<sub>MM02&05c3</sub> are less satisfactory and indicate that the models might be over fitting the data.

### 3.4.2 Evaluation on intermediate sites

The headwater models were also tested on the sites of intermediate size (i.e.  $>2\text{ km}^2$ ) but excluding the river outlet (lower parts of Table 5). In general the intermediate sites were modelled as successfully as the headwaters (Table 5). Cal<sub>MM07c2</sub> gave predictions for the intermediate sites which were not as good as for the other data sets, i.e. higher MedRE%, than Cal<sub>MM08c2</sub> and Cal<sub>MM00c1</sub>.

### 3.4.3 Evaluation of river outlets

To test the effect of including the river outlet TOC on the performance of MM predictions for individual headwaters, three versions of each calibration data set was used to create three different models, one using outlet TOC alone (Out), one using map information alone (Map), and one using both the outlet and map information (OutMap). The map variables were the same for each calibration data set but with different calibrated coefficients (Table 4). In 25 out of 27 different combinations of MM (three different calibration data sets, three versions of each calibration (Out, Map, OutMap) and three different test data sets), the OutMap version gave the best performance with the lowest PRESS, while two Map versions (map information only, no OutletTOC included) gave the lowest PRESS (Table 5 and Fig. 3–5). Similar results were observed for the intermediate sites (Table 5 and Fig. 3–5). The OutMap version gave 5–15 % better predictions than Map only.

## 4 Discussion

Twenty-five of twenty-seven tests, including river outlet chemistry (OutletTOC), resulted in lower errors in the mixed models predictions of the TOC for individual headwaters, and intermediate sites, compared to using map information alone. The measurements at the river outlet were necessary to reproduce more correct average headwater TOC levels (Table 5). Excluding the OutletTOC measurements leads to the assumption that average TOC levels in the headwaters were similar in different catchment stream networks if the map information is similar, which is not always true (cf. sampling 2007 and 2008). This is the first article to test how to include river outlet chemistry with map information for modelling headwaters, and how well the river outlet chemistry improved the models.

To predict the correct mean values for headwater TOC is still a challenge in our application of mixed models, since the calibration sets consist of 4–5 catchment systems, and these were sampled only a few times for each calibration. This clearly makes generalisations to new catchments or flow-situations difficult and uncertain. Calibration sets Cal07 and Cal08 share most of the catchments, but are measured during different years and perhaps more importantly, during different flow situations and seasons. Even if most headwaters are the same in Cal<sub>MM07c2</sub> and Cal<sub>MM08c2</sub>, and the models produced were similar, it was still not possible to account for more than about 50 % of the variation in the other set of data, indicating that there is large variability in time. Weather is a factor that varies with time (and space) and influences stream water chemistry. In our approach we did not include the weather related data (temperature, precipitation, flow) in the models (PLS and MM) since it was not available for headwaters, but only for the river outlet. Presumably discharge for each headwater could benefit the models, but measuring discharge at all individual headwaters would have been very time consuming (and was not performed). To model discharge with appropriate accuracy at all these head-

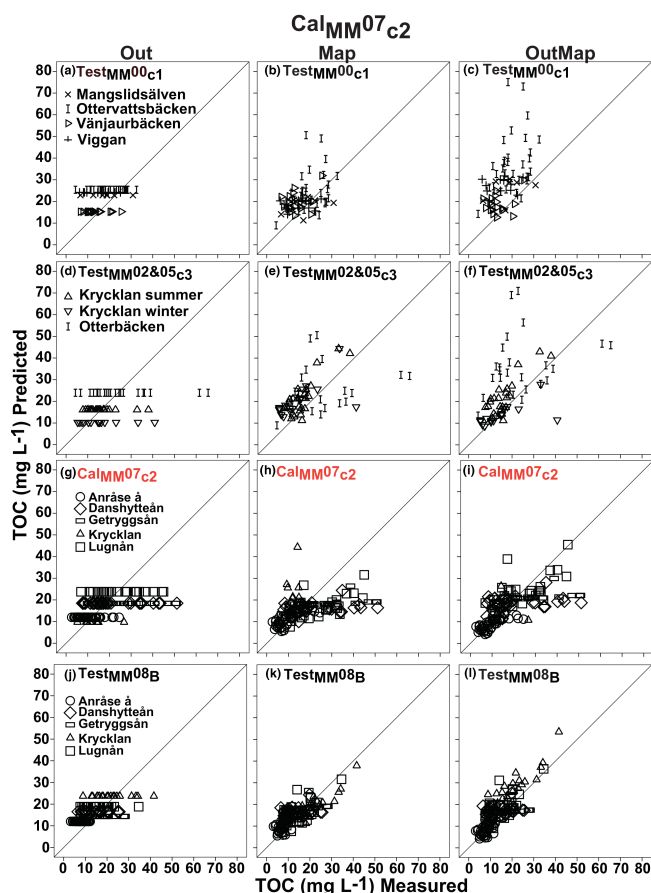
**Table 5.** Hierarchical linear mixed models (MM) results predicting headwater (HW) organic carbon (TOC) concentration in Log10. Cal is calibrated and 00, 07 and 08 refer to sampling year (2000, 2007 and 2008). The different coefficients are found in Table 4. Each calibration has three versions (intercept is always included): Out, Map and OutMap. The Out version includes OutletTOC, the Map version includes the map information, while OutMap includes both OutletTOC and map information. The prediction error sum of squares (PRESS) is the squared differences between observed and predicted values for the Y-data kept out of the model fitting in the Test sets (00 stands for year 2000, 02&05 for 2002 and 2005, 07 for 2007 and 08 for 2008). The bold values show the lowest PRESS of the three versions for that Test data, bold PRESS per test sites are the lowest for that version. Intermediate sites (is) stands for sites between headwaters and the river outlet.  $R^2$  for predictions is called  $P^2$  (similar to  $Q^2$  for PLS):  $P^2 = 1 - \text{PRESS}/\text{TSS}$  where TSS is the total sum of squared differences between modelled and the mean of observations. AE is absolute error and RE is relative error, calculated on TOC in  $\text{mg L}^{-1}$ .

Cal	Version	00 <sub>c1</sub>	00 <sub>c1</sub>	00 <sub>c1</sub>	07 <sub>c2</sub>	07 <sub>c2</sub>	07 <sub>c2</sub>	08 <sub>c2</sub>	08 <sub>c2</sub>	08 <sub>c2</sub>
Cal n <sub>HW</sub>		69			138			148		
Test <sub>HW</sub>		02&05 <sub>c3</sub>	07 <sub>c2</sub>	08 <sub>c2</sub>	00 <sub>c1</sub>	02&05 <sub>c3</sub>	08 <sub>c2</sub>	00 <sub>c1</sub>	02&05 <sub>c3</sub>	07 <sub>c2</sub>
n Test <sub>HW</sub>		65	138	148	69	65	148	69	65	138
PRESS <sub>HW</sub>	Out	3.42	8.71	4.78	2.28	3.68	5.98	3.10	4.58	7.92
	Map	2.79	8.51	5.35	2.24	2.70	3.97	3.17	3.57	7.63
	OutMap	2.47	7.82	4.00	2.26	2.88	3.92	3.01	3.45	6.26
$P^2_{\text{HW}}$	Out	13.5	19.8	42.9	9.2	8.5	20.6	5.0	3.3	31.5
	Map	29.4	21.7	36.1	10.8	32.9	47.3	2.7	24.5	34.0
	OutMap	37.6	28.0	52.0	10.1	28.3	48.0	7.5	27.1	45.9
MedAE <sub>HW</sub>	Out	6.85	8.00	4.99	3.30	4.61	4.89	6.30	6.85	7.52
	Map	4.54	7.24	5.56	7.67	5.79	4.91	7.84	5.43	7.34
	OutMap	6.32	6.59	5.03	8.29	7.51	4.57	7.14	6.84	5.80
MedRE% <sub>HW</sub>	Out	45.9	53.4	45.9	47.1	26.8	60.8	33.1	46.6	55.5
	Map	42.1	56.3	52.2	52.2	66.8	39.3	50.6	65.3	53.1
	OutMap	42.1	56.3	52.2	98.1	44.2	35.7	58.6	54.3	48.0
n <sub>is</sub>		129	155	135	82	129	135	82	129	155
PRESS <sub>is</sub>	Out	5.59	5.35	3.19	3.24	4.92	6.24	3.30	5.18	5.28
	Map	5.95	7.92	2.98	2.42	4.44	1.71	2.50	4.17	7.26
	OutMap	4.45	3.93	1.41	2.52	3.15	1.92	2.46	3.31	4.59
$P^2_{\text{is}}$	Out	20.3	47.9	52.4	−11.2	19.8	−46.8	−17.8	14.7	26.8
	Map	15.3	22.9	55.5	17.0	27.6	59.6	10.7	31.3	−0.46
	OutMap	36.6	61.8	78.9	13.6	48.6	54.9	12.1	45.4	36.4
MedAE <sub>is</sub>	Out	3.97	4.22	2.08	5.89	5.10	2.27	5.65	5.22	3.81
	Map	3.74	4.69	3.57	2.90	3.99	2.22	3.29	3.64	4.78
	OutMap	3.63	3.00	2.42	4.18	4.48	2.62	2.66	3.60	3.78
MedRE% <sub>is</sub>	Out	40.9	43.0	43.2	69.9	53.6	84.0	41.4	38.9	34.9
	Map	68.2	59.0	36.8	40.6	94.6	23.1	37.3	72.2	50.3
	OutMap	38.2	33.5	23.5	88.4	73.6	25.0	41.6	52.0	40.3

waters (size < 2 km<sup>2</sup>) is so far too difficult due to large heterogeneity at these small scales (Lyon et al., 2012), and lack of precipitation data for all these headwaters.

With the approach in this study we were able to achieve  $P^2$  values around 50 %, indicating that about 50 % of the variation in the individual headwater test sets can be explained with a model including the explanatory variables OutletTOC, lake surface coverage and proportion of coniferous forests on mires. In two of the three calibrations, the proportion of broad-leaved forest and elevation was also significant.

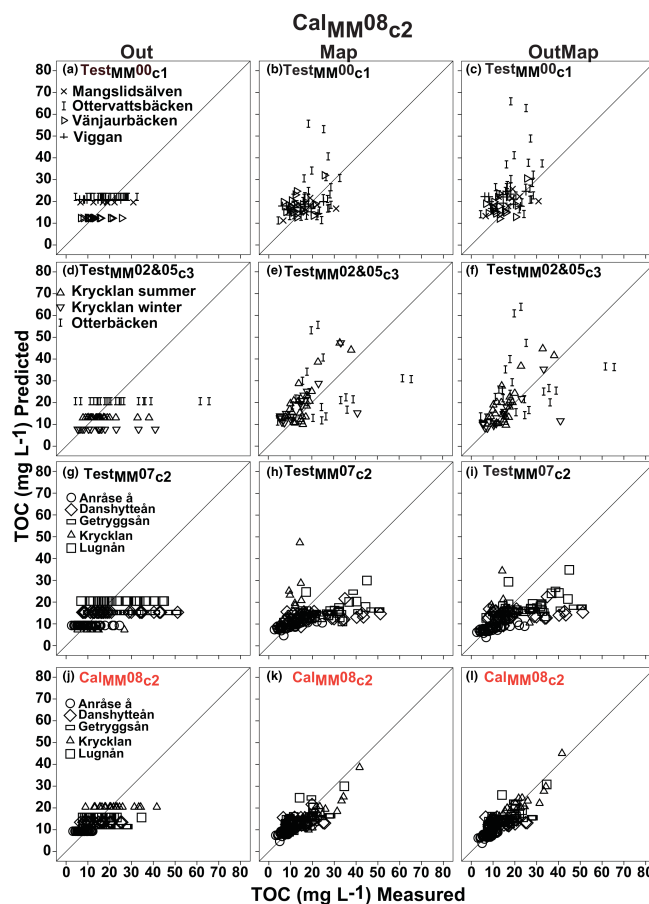
Most lakes in these catchments are dimictic (mixing of the lake from the surface to bottom twice each year). Some of the data used in this study (year 2007 and 2008) have been used to evaluate the impact of lakes on stream water chemistry and there were indications that lake influence differs as a function of season, catchment and constituent (Lyon et al., 2011). The presence of lakes had a stronger influence on stream water TOC levels in October 2007 than in April 2008. Thus the presence of lakes could influence the impact of river outlet TOC on headwater TOC in MM. Lakes are known to often decrease TOC concentration (Müller et al., 2013; Weyhenmeyer et al., 2012), although this effect is not always visible



**Figure 4.** Scatterplots of measured headwaters with total organic carbon (TOC in  $\text{mg L}^{-1}$ ) on the x axis, and the three different versions of the mixed models CalMM07c2 on the y axis: Out version on the left panel, OutMap version on the right panel and Map in-between. Data for year 2007 (CalMM07c2) on the third row in red text, followed by Test data; first row 2000 data, second row 2002 & 2005 data and the last row is 2008 data. R. Anråse å indicated by circles, R. Danshytteån by diamonds, R. Getryggsån by rectangles, R. Krycklan by triangles (winter 2005 by upside-down triangles), R. Lugnån by squares, R. Mangslidsälven by multiplication sign, R. Ottervattsbacken by upside-down triangles, R. Vänjaubacken by right tilted triangles and R. Viggan by plus sign. The black line is the 1 : 1 line.

at a landscape scale (Lottig et al., 2013) and lakes can also delay pulses of TOC within river networks, which is a complicating factor (Hyttborn et al., 2015).

The proportion of coniferous forest on mires had a positive sign and proportion of lake surface coverage had a negative sign for all calibration sets. This is plausible based on earlier studies (Andersson and Nyberg, 2009; Pers et al., 2001; Oni et al., 2013; Lottig et al., 2013; Clark et al., 2010). Ågren et al. (2014); Hope et al. (1997); Löfgren et al. (2014); Mattsson et al. (2003) and Walker et al. (2012) have also found that the amount of organic matter in the catchment soils (mire, wet- or peat land proportion) is often positively correlated with



**Figure 5.** Scatterplots of measured headwaters with total organic carbon (TOC in  $\text{mg L}^{-1}$ ) on the x axis, and the three different versions of the mixed models Cal08 on the y axis: Out version on the left panel, OutMap version on the right panel and Map in-between. Data for year 2008 (CalMM08c2) on the last row in red text, followed by Test data; first row 2000 data, second row 2002 & 2005 data and the third row is 2007 data. R. Anråse å indicated by circles, R. Danshytteån by diamonds, R. Getryggsån by rectangles, R. Krycklan by triangles (winter 2005 by upside-down triangles), R. Lugnån by squares, R. Mangslidsälven by multiplication sign, R. Ottervattsbacken by upside-down triangles, R. Vänjaubacken by right tilted triangles and R. Viggan by plus sign. The black line is the 1 : 1 line.

stream TOC concentration (e.g. Mulholland et al., 2001), even if the extent of organic soils can be hard to estimate from maps (Creed et al., 2003; Johnston et al., 2008).

That broad-leaved forest had a negative coefficient for Cal07c2 and Cal08c2 (Table 4–5) could be related to several factors. In these systems most of the broad-leaved forest is made up of birch (*Betula pendula*). The negative coefficient of broad-leaved forest/birch could be a direct effect of birch on water chemistry; i.e. more birch in a coniferous landscape could give runoff with lower organic carbon (Brandtberg et al., 2000; Fröberg et al., 2011). However, in a set of explanatory variables like this, with a large amount of geographical

information, many of the variables are correlated. This results in the fact that similarly good models could be found with other sets of explanatory variables. For instance, in our data set we found high correlations among volume of various tree species, furthermore volume of pine also had high correlation with these variables: proportion of coniferous forest, site altitude, volume of birch, broad-leaved forest and volume of spruce.

In this work we use calibration sets and test sets rather than the popular leave-one-out cross-validation method, since we have dependent, clustered data. Shao (1993) showed that leave-one-out cross-validation tends to select unnecessarily large models if observations are correlated. Libiseller and Grimvall (2003) showed that this is true for data that are serially correlated, since a single removed observation can be reproduced easily by observations in the temporal vicinity of the left-out observation. The same should hold for clustered data, where a single left-out headwater would be reasonably predicted by other headwaters in the same catchment.

## 5 Conclusion

Our modelling approach, using both river outlet TOC and map information from the headwater catchments, could explain up to 52 % of the variance in TOC among individual headwater streams. This is much better performance than an earlier attempt using river outlet TOC without map information (Temnerud et al., 2010), in which only one of seven models were significant for predicting headwater median TOC and none were significant for predicting headwater TOC IQR. The key factor in our approach here is the use of mixed models which allow the same headwaters to have different TOC depending on weather and flow etc. Since MM cannot use large numbers of correlated explanatory variables, PLS was used to identify a set of candidate explanatory variables for the MM. Since our combined approach increased the predictability for TOC, it would be interesting to evaluate whether the method could improve prediction of headwater pH and ANC, for which models using outlet catchment chemistry were already fairly successful (Temnerud et al., 2010).

In order to have the same map resolution for all catchments, due to lack of universal availability of fine-scale data, a rather coarse resolution was used (e.g. 50 m grid data for altitude and soil map of scale 1:1 million). The Swedish authorities are LiDAR scanning all of Sweden to build a 2 m grid digital elevation model and are generating maps connecting all watercourses up to the headwaters (through lakes and wetlands), which by 2017 will provide new data that could help in modelling the headwaters. This improved map information might further improve the mixed model approach demonstrated here which includes river outlet chemistry. This will hopefully get us closer to the ability to predict

individual headwaters that are such vital building blocks of aquatic ecosystems, but remain so very difficult to model.

**The Supplement related to this article is available online at doi:10.5194/bg-13-399-2016-supplement.**

*Acknowledgements.* Financial support for this research was provided by the Swedish University of Agricultural Sciences, the Swedish Environmental Protection Agency project Climate Change and Environmental Objectives (CLEO), MISTRA Future Forests and, for Johan Temnerud, the Knowledge Foundation. We would like to thank all samplers and laboratories for excellent work. Mats Fröberg is acknowledged for help on birch impact on organic carbon in soils and streams.

Edited by: T. J. Battin

## References

- Ågren, A. M., Buffam, I., Cooper, D. M., Tiwari, T., Evans, C. D., and Laudon, H.: Can the heterogeneity in stream dissolved organic carbon be explained by contributing landscape elements?, *Biogeosciences*, 11, 1199–1213, doi:10.5194/bg-11-1199-2014, 2014.
- Akaike, H.: A new look at the statistical model identification, *IEEE T. Automat. Contr.*, 19, 716–723, doi:10.1109/TAC.1974.1100705, 1974.
- Alexander, R. B., Boyer, E. W., Smith, R. A., Schwarz, G. E., and Moore, R. B.: The role of headwater streams in downstream water quality, *J. Am. Water Resour. As.*, 43, 41–59, doi:10.1111/j.1752-1688.2007.00005.x, 2007.
- Andersson, J.-O. and Nyberg, L.: Using official map data on topography, wetlands and vegetation cover for prediction of stream water chemistry in boreal headwater catchments, *Hydrol. Earth Syst. Sci.*, 13, 537–549, doi:10.5194/hess-13-537-2009, 2009.
- Arheimer, B. and Lindström, G.: Implementing the EU Water Framework Directive in Sweden, in: *Runoff Predictions in Ungauged Basins – Synthesis across Processes, Places and Scales*, edited by: Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H., Cambridge University Press, Cambridge, 353–359, 2013.
- Birge, E. A. and Juday, C.: The organic content of lake water, *P. Natl. Acad. Sci.*, 12, 515–519, 1926.
- Bossard, M., Feranec, J., and Otahel, J.: CORINE land cover technical guide – Addendum 2000, European Environment Agency, Copenhagen, 105 pp., <http://www.eea.europa.eu/publications/COR0-landcover> (last access: 8 November 2012), 2000.
- Brandtberg, P. O., Lundkvist, H., and Bengtsson, J.: Changes in forest-floor chemistry caused by a birch admixture in Norway spruce stands, *Forest Ecol. Manag.*, 130, 253–264, doi:10.1016/S0378-1127(99)00183-8, 2000.
- Buffam, I., Laudon, H., Temnerud, J., Mörtz, C.-M., and Bishop, K.: Landscape-scale variability of acidity and dissolved organic

- carbon during spring flood in a boreal stream network, *J. Geophys. Res.-Biogeo.*, 112, G01022, doi:10.1029/2006JG000218, 2007.
- Clark, J. M., Bottrell, S. H., Evans, C. D., Monteith, D. T., Bartlett, R., Rose, R., Newton, R. J., and Chapman, P. J.: The importance of the relationship between scale and process in understanding long-term DOC dynamics, *Sci. Total Environ.*, 408, 2768–2775, doi:10.1016/j.scitotenv.2010.02.046, 2010.
- Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A., and Dillon, P. J.: Cryptic wetlands: Integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes, *Hydrol. Process.*, 17, 3629–3648, doi:10.1002/hyp.1357, 2003.
- Edström, M. and Rystam, P.: Network of stations for field research area (in Swedish: FFO – Stationsnät för fältforskningsområden), Swedish Meteorological and Hydrological Institute, Norrköping, SMHI hydrology report 53, 53 pp., 1994.
- Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., and Wold, S.: Multi- and megavariable data analysis – Part 1: Basic Principles and Applications, Umetrics AB, Umeå, 2006.
- Eriksson, J. V.: Den kemiska denudationen i Sverige (The chemical denudation of Sweden, in Swedish with a French summary), Swedish Meteorological and Hydrological Institute, Stockholm, Report Band 5, nr. 3, 95 pp., 1929.
- Evans, C., Cooper, D., Monteith, D., Helliwell, R., Moldan, F., Hall, J., Rowe, E., and Cosby, B.: Linking monitoring and modelling: Can long-term datasets be used more effectively as a basis for large-scale prediction?, *Biogeochemistry*, 101, 211–227, doi:10.1007/s10533-010-9413-x, 2010.
- Fröberg, M., Hansson, K., Kleja Berggren, D., and Alavi, G.: Dissolved organic carbon and nitrogen leaching from Scots pine, Norway spruce and silver birch stands in southern Sweden, *Forest Ecol. Manag.*, 262, 1742–1747, doi:10.1016/j.foreco.2011.07.033, 2011.
- Gadmar, T. C., Vogt, R. D., and Østerhus, B.: The merits of the high temperature combustion method for determining the amount of natural organic carbon in surface freshwater samples, *Int. J. Environ. An. Ch.*, 82, 451–461, doi:10.1080/0306731021000018099, 2002.
- Geladi, P. and Kowalski, B. R.: Partial least-squares regression – a tutorial, *Anal. Chim. Acta*, 185, 1–17, doi:10.1016/0003-2670(86)80028-9, 1986.
- Gomi, T., Sidle, R. C., and Richardson, J. S.: Understanding processes and downstream linkages of headwater systems, *BioScience*, 52, 905–916, doi:10.1641/0006-3568(2002)052[0905:UPADLO]2.0.CO;2, 2002.
- Hope, D., Billett, M. F., and Cresser, M. S.: Exports of organic carbon in two river systems in NE Scotland, *J. Hydrol.*, 193, 61–82, doi:10.1016/S0022-1694(96)03150-2, 1997.
- Hutchins, M. G., Smith, B., Rawlins, B. G., and Lister, T. R.: Temporal and spatial variability of stream waters in Wales, the Welsh borders and part of the West Midlands, UK. 1. Major ion concentrations, *Water Res.*, 33, 3479–3491, doi:10.1016/S0043-1354(99)00057-3, 1999.
- Hytteborn, J. K., Temnerud, J., Alexander, R. B., Boyer, E. W., Futter, M. N., Fröberg, M., Dahné, J., and Bishop, K. H.: Patterns and predictability in the intra-annual organic carbon variability across the boreal and hemiboreal landscape, *Sci. Total Environ.*, 520, 260–269, doi:10.1016/j.scitotenv.2015.03.041, 2015.
- Ivarsson, H. and Jansson, M.: Temporal variations in the concentration and character of dissolved organic matter in a highly colored stream in the coastal zone of northern Sweden, *Arch. Hydrobiol.*, 132, 45–55, 1994.
- Jager, H. I., Bevelhimer, M. S., King, R. L., and Smith, K. A.: Landscape influences on headwater streams on Fort Stewart, Georgia, USA, *Environ. Manage.*, 48, 795–807, doi:10.1007/s00267-011-9722-4, 2011.
- Johansson, B.: Areal precipitation and temperature in the Swedish Mountains – An evaluation from a hydrological perspective, *Nord. Hydrol.*, 31, 207–228, 2000.
- Johansson, B. and Chen, D. L.: The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling, *Int. J. Climatol.*, 23, 1523–1535, doi:10.1002/joc.951, 2003.
- Johnston, C. A., Shmagin, B. A., Frost, P. C., Cherrier, C., Larson, J. H., Lamberti, G. A., and Bridgman, S. D.: Wetland types and wetland maps differ in ability to predict dissolved organic carbon concentrations in streams, *Sci. Total Environ.*, 404, 326–334, doi:10.1016/j.scitotenv.2007.11.005, 2008.
- Kenward, M. G. and Roger, J. H.: Small sample inference for fixed effects from restricted maximum likelihood, *Biometrics*, 53, 983–997, doi:10.2307/2533558, 1997.
- Köhler, S.: Quantifying the role of natural organic acids on pH and buffering in Swedish surface waters, Doctoral Thesis, Forest Ecology, Swedish University of Agricultural Sciences, Umeå, 163 pp., 1999.
- Köhler, S. J., Buffam, I., Laudon, H., and Bishop, K. H.: Climate's control of intra-annual and interannual variability of total organic carbon concentration and flux in two contrasting boreal landscape elements, *J. Geophys. Res.-Biogeo.*, 113, G03012, doi:10.1029/2007JG000629, 2008.
- Laudon, H., Berggren, M., Ågren, A., Buffam, I., Bishop, K., Grabs, T., Jansson, M., and Köhler, S.: Patterns and dynamics of dissolved organic carbon (DOC) in boreal streams: The role of processes, connectivity, and scaling, *Ecosystems*, 14, 880–893, doi:10.1007/s10021-011-9452-8, 2011.
- Laudon, H., Taberman, I., Ågren, A., Futter, M., Ottosson-Löfvenius, M., and Bishop, K.: The Krycklan Catchment study – A flagship infrastructure for hydrology, biogeochemistry and climate research in the boreal landscape, *Water Resour. Res.*, 49, 7154–7158, doi:10.1002/wrcr.20520, 2013.
- Leopold, L. B., Wolman, M. G., and Miller, J. P.: Fluvial Processes in Geomorphology, A Series of Books in Geology, edited by: Gilluly, J. and Woodford, A. O., W. H. Freeman and Company, San Francisco, 522 pp., 1964.
- Libiseller, C. and Grimvall, A.: Model selection for local and regional meteorological normalisation of background concentrations of tropospheric ozone, *Atmos. Environ.*, 37, 3923–3931, doi:10.1016/S1352-2310(03)00502-8, 2003.
- Likens, G. E. and Buso, D. C.: Variation in streamwater chemistry throughout the Hubbard Brook Valley, *Biogeochemistry*, 78, 1–30, doi:10.1007/s10533-005-2024-2, 2006.
- Lindström, G., Pers, C., Rosberg, J., Strömquist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spa-

- tial scales, *Hydrol. Res.*, 41, 295–319, doi:10.2166/nh.2010.007, 2010.
- Littell, R., Milliken, G., Stroup, W., Wolfinger, R., and Schabenberger, O.: *SAS for Mixed Models*, 2nd Edn., SAS Institute, Cary, 828 pp., 2006.
- Löfgren, S., Aastrup, M., Bringmark, L., Hultberg, H., Lewin-Pihlblad, L., Lundin, L., Karlsson, G. P., and Thunholm, B.: Recovery of soil water, groundwater, and streamwater from acidification at the Swedish Integrated Monitoring catchments, *AMBIO*, 40, 836–856, doi:10.1007/s13280-011-0207-8, 2011.
- Löfgren, S., Fröberg, M., Yu, J., Nisell, J., and Ranneby, B.: Water chemistry in 179 randomly selected Swedish headwater streams related to forest production, clear-felling and climate, *Environ. Monit. Assess.*, 186, 8907–8928, doi:10.1007/s10661-014-4054-5, 2014.
- Lottig, N. R., Buffam, I., and Stanley, E. H.: Comparisons of wetland and drainage lake influences on stream dissolved carbon concentrations and yields in a north temperate lake-rich region, *Aquat. Sci.*, 75, 1–12, doi:10.1007/s00027-013-0305-8, 2013.
- Lyon, S. W., Braun, H., and Temnerud, J.: Accounting for in-stream lakes when interpolating stream water chemistry observations, *JoSH*, 11, 1–20, <http://www.spatialhydrology.net/index.php/JOSH/article/view/102>, 2011.
- Lyon, S. W., Nathanson, M., Spans, A., Grabs, T., Laudon, H., Temnerud, J., Bishop, K. H., and Seibert, J.: Specific discharge variability in a boreal landscape, *Water Resour. Res.*, 48, W08506, doi:10.1029/2011WR011073, 2012.
- MacDonald, L. H. and Coe, D.: Influence of headwater streams on downstream reaches in forested areas, *Forest Sci.*, 53, 148–168, 2007.
- Mattsson, T., Finér, L., Kortelainen, P., and Sallantausta, T.: Brook water quality and background leaching from unmanaged forested catchments in Finland, *Water Air Soil Poll.*, 147, 275–297, doi:10.1023/A:1024525328220, 2003.
- McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., and Bailey, S. W.: Network analysis reveals multiscale controls on streamwater chemistry, *P. Natl. Acad. Sci. USA*, 111, 7030–7035, doi:10.1073/pnas.1404820111, 2014.
- Méndez Mediavilla, F. A., Landrum, F., and Shah, V.: A comparison of the coefficient of predictive power, the coefficient of determination and AIC for linear regression, *Proceedings of the 39th Annual Meeting of the Decision Sciences Institute*, Atlanta, 2008.
- Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., and Leonard, N. E.: The contribution of headwater streams to biodiversity in river networks, *J. Am. Water Resour. As.*, 43, 86–103, doi:10.1111/j.1752-1688.2007.00008.x, 2007.
- Mulholland, P. J., Fellows, C. S., Tank, J. L., Grimm, N. B., Webster, J. R., Hamilton, S. K., Martí, E., Ashkenas, L., Bowden, W. B., Dodds, W. K., McDowell, W. H., Paul, M. J., and Peterson, B. J.: Inter-biome comparison of factors controlling stream metabolism, *Freshwater Biol.*, 46, 1503–1517, doi:10.1046/j.1365-2427.2001.00773.x, 2001.
- Müller, R. A., Futter, M. N., Sobek, S., Nisell, J., Bishop, K., and Weyhenmeyer, G. A.: Water renewal along the aquatic continuum offsets cumulative retention by lakes: Implications for the character of organic carbon in boreal lakes, *Aquat. Sci.*, 75, 535–545, doi:10.1007/s00027-013-0298-3, 2013.
- Nagorski, S. A., Moore, J. N., McKinnon, T. E., and Smith, D. B.: Scale-dependent temporal variations in stream water geochemistry, *Environ. Sci. Technol.*, 37, 859–864, doi:10.1021/es025983+, 2003.
- Oni, S. K., Futter, M. N., Bishop, K., Köhler, S. J., Ottosson-Löfvenius, M., and Laudon, H.: Long-term patterns in dissolved organic carbon, major elements and trace metals in boreal headwater catchments: trends, mechanisms and heterogeneity, *Biogeosciences*, 10, 2315–2330, doi:10.5194/bg-10-2315-2013, 2013.
- Pers, C., Rahm, L., Jonsson, A., Bergström, A.-K., and Jansson, M.: Modelling dissolved organic carbon turnover in humic Lake Örträsket, Sweden, *Environ. Model. Assess.*, 6, 159–172, doi:10.1023/A:1011953730983, 2001.
- Person, H. S., Johnston, C. E., and Beall, R. T.: Little waters: A study of headwater streams & other little waters, their use and relations to the land, *The Upstream Engineering Conference*, Washington D.C., 1–82, <http://www.archive.org/details/littlewatersstud00persrich> (last access: 16 January 2016), 1936.
- Reese, H., Nilsson, M., Sandström, P., and Olsson, H.: Applications using estimates of forest parameters derived from satellite and forest inventory data, *Comput. Electron. Agr.*, 37, 37–55, doi:10.1016/S0168-1699(02)00118-7, 2002.
- Development Core Team: *R: A Language and Environment for Statistical Computing*, Vienna, 2014.
- Reese, H., Nilsson, M., Granqvist Pahlén, T., Hagner, O., Joyce, S., Tingelöf, U., Egberth, M., and Olsson, H.: Countrywide estimates of forest variables using satellite data and field data from the national forest inventory, *AMBIO*, 33, 542–548, doi:10.1579/0044-7447-32.8.542, 2003.
- Sakamaki, T. and Richardson, J. S.: Nonlinear variation of stream–forest linkage along a stream-size gradient: An assessment using biogeochemical proxies of in-stream fine particulate organic matter, *J. Appl. Ecol.*, 50, 1019–1027, doi:10.1111/1365-2664.12099, 2013.
- Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M.: *Environmental Organic Chemistry*, 2 Edn., John Wiley, Hoboken, 1313 pp., 2003.
- Shao, J.: Linear model selection by cross-validation, *J. Am. Stat. Assoc.*, 88, 486–494, doi:10.1080/01621459.1993.10476299, 1993.
- Smart, R. P., Soulsby, C., Cresser, M. S., Wade, A. J., Townend, J., Billett, M. F., and Langan, S.: Riparian zone influence on stream water chemistry at different spatial scales: A GIS-based modelling approach, an example for the Dee, NE Scotland, *Sci. Total Environ.*, 280, 173–193, doi:10.1016/S0048-9697(01)00824-5, 2001.
- Strayer, D. L., Beighley, R. E., Thompson, L. C., Brooks, S., Nilsson, C., Pinay, G., and Naiman, R. J.: Effects of land cover on stream ecosystems: Roles of empirical models and scaling issues, *Ecosystems*, 6, 407–423, doi:10.1007/PL00021506, 2003a.
- Strayer, D. L., Ewing, H. A., and Bigelow, S.: What kind of spatial and temporal details are required in models of heterogeneous systems?, *Oikos*, 102, 654–662, doi:10.1034/j.1600-0706.2003.12184.x, 2003b.
- Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and Lindström, G.: Water and nutrient predictions in ungauged basins: Set-up and evaluation of a model at the national scale, *Hydrolog. Sci. J.*, 57, 229–247, doi:10.1080/02626667.2011.637497, 2012.
- Temnerud, J. and Bishop, K.: Spatial variation of streamwater chemistry in two Swedish boreal catchments: Implications for

- environmental assessment, *Environ. Sci Technol.*, 39, 1463–1469, doi:10.1021/es040045q, 2005.
- Temnerud, J., Seibert, J., Jansson, M., and Bishop, K.: Spatial variation in discharge and concentrations of organic carbon in a catchment network of boreal streams in Northern Sweden, *J. Hydrol.*, 342, 72–87, doi:10.1016/j.jhydrol.2007.05.015, 2007.
- Temnerud, J., Fölster, J., Pilström, F., and Bishop, K.: Synoptisk provtagning av små vattendrag i södra Sverige oktober 2007 (Synoptical sampling of small watercourses in southern Sweden in October 2007, in Swedish), Swedish University of Agricultural Sciences, Uppsala, Report 2009:6, 56 pp., <http://publikationer.slu.se/Filer/Rapport2009-6.pdf> (last access: 16 January 2016), 2009.
- Temnerud, J., Fölster, J., Buffam, I., Laudon, H., Erlandsson, M., and Bishop, K.: Can the distribution of headwater stream chemistry be predicted from downstream observations?, *Hydrol. Process.*, 24, 2269–2276, doi:10.1002/hyp.7615, 2010.
- Walker, C. M., King, R. S., Whigham, D. F., and Baird, S. J.: Landscape and wetland influences on headwater stream chemistry in the Kenai Lowlands, Alaska, *Wetlands*, 32, 301–310, doi:10.1007/s13157-011-0260-x, 2012.
- Wetzel, R. G.: *Limnology. Lake and River Ecosystems*, 3 Edn., Academic Press, London, 1006 pp., 2001.
- Weyhenmeyer, G. A., Fröberg, M., Karlun, E., Khalili, M., Kothawala, D., Temnerud, J., and Tranvik, L. J.: Selective decay of terrestrial organic carbon during transport from land to sea, *Glob. Change Biol.*, 18, 349–355, doi:10.1111/j.1365-2486.2011.02544.x, 2012.
- Wolock, D. M., Fan, J., and Lawrence, G. B.: Effects on basin size on low-flow stream chemistry and subsurface contact time in the Neversink river watershed, New York, *Hydrol. Process.*, 11, 1273–1286, doi:10.1002/(SICI)1099-1085(199707)11:9<1273::AID-HYP557>3.0.CO;2-S, 1997.