1 **Map-based prediction of organic carbon in headwater streams improved by**
2 **downstream observations from the river outlet**
3
4 J. Temnerud[1,2,*], C. von Brömssen[3], J. Fölster[1], I. Buffam[4], J.-O. Andersson[5], L. Nyberg[6], K.
5 Bishop[1,7]
6
7 [1] Department of Aquatic Sciences and Assessment, Swedish University of Agricultural
8 Sciences, Uppsala, Sweden.
9 [2] Research Department, Swedish Meteorological and Hydrological Institute, Norrköping,
10 Sweden.
11 [3] Unit of applied statistics and mathematics, Department of Economics, Swedish University of
12 Agricultural Sciences, Uppsala, Sweden.
13 [4] Department of Biological Sciences and Department of Geography, University of Cincinnati,
14 Cincinnati, USA.
15 [5] Department of Geography, Karlstad University, Karlstad, Sweden.
16 [6] Centre for Climate and Safety, Karlstad University, Karlstad, Sweden.
17 [7] Department of Earth Sciences, Uppsala University, Uppsala, Sweden.
18
19 [*] Corresponding author present address: Department of Aquatic Sciences and Assessment,
20 Swedish University of Agricultural Sciences, P.O. Box 7050, SE-75007 Uppsala, Sweden.
21 Phone: +46-70-3514840, fax: +46-18-673156, e-mail: Johan.Temnerud@slu.se
22
23 Journal: Biogeosciences
24 Running title: Modelling TOC in headwater streams
25
26

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 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 32-235 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) proportion 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 offers information which can complement GIS-based modelling of headwater stream chemistry.

Keywords: mixed model, partial least square, organic carbon, headwater, stream

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; Strayer et al., 2003a; 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 neutralizing 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.

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

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

2 Methods

2.1 Sampling approach

123 The synoptic surveys of mesoscale catchments used in this study were designed to provide a
124 snapshot of the water chemistry in stream networks (Table 1, Fig. 1), with a strategy of
125 sampling most watercourse junctions, lake inlets and outlets as well as the river outlets from
126 each mesoscale catchment. In total there were data from 17 synoptic surveys conducted
127 between 2000 and 2008 in nine catchments distributed across Sweden, spanning a north-south
128 gradient of 800 km through the north-temperate and boreal zones (Fig. 1). This data set
129 amounted to 938 stream samples of which 420 were from headwaters. Headwaters are defined
130 as first order streams with catchments smaller than 2 km^2 in each of the nine drainage
131 networks sampled (Table 1). The number of sampled headwaters differs between surveys
132 (n_{HW} in Table 2). All sampling during a given survey was carried out during a one to three
133 day period (Table 2 and Fig. S1), except for R. Krycklan in winter 2005, which took two
134 weeks due to cold weather and difficulties finding the streams in deep snow. Stable base flow
135 discharge was maintained throughout that winter sampling period, so that survey was still
136 effectively a snapshot in time.

137 Five of the nine catchments contain at least one headwater stream site that has been monitored
138 for runoff and chemistry regularly for a decade or more (Edström and Rystam, 1994;
139 Temnerud et al., 2009; Köhler et al., 2008; Löfgren et al., 2011; Laudon et al., 2013). Of the
140 other four catchments, R. Ottervattsbäcken was sampled twice and R. Vänjaurbäcken sampled
141 once (the name R. Sörbäcken was used in the references, as this is a tributary in R.
142 Vänjaurbäcken, Temnerud and Bishop, 2005; Temnerud et al., 2007) while R. Viggan and R.
143 Mangslidsälven were sampled once, for this study.

4

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 silvestris*) (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^2) 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, Fig. 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., 2003; Reese et al., 2002), 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_2 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.

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

2.5.1 Modelling headwater median and interquartile range

192 To model the median and interquartile range of TOC in headwaters in different catchments
193 we use partial least squares regression (PLS). Variables included in this model are TOC at the
194 river outlet (OutletTOC) and a number of variables describing information derived from land
195 cover, geology, soil type maps and vegetation (kNN) (Table S1). The main purpose of PLS
196 was to narrow down the number of explanatory variables for subsequent use in the mixed
197 models approach.

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

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

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

224   - Some variables, e.g. volume of oak, had zero value for all observations or only few
225   observations different from zero. These variables could not be included due to the lack
226   of variation.

227   - Geographical variables for the river outlet were not included, since they correlate
228   highly with the median of the corresponding variable at the headwater scale. The latter
229   is considered to bear more information and was therefore included.

230 The analysis was run on three calibration data sets: One with data from year 2000 where
231 headwaters within four mesoscale catchments (these catchments are called c1) were sampled
232 $Cal_{PLS}00_{c1}$ (n = 4; M0, O0, Vä0, Vi0). In $Cal_{PLS}07_{c2}$ the headwaters of five mesoscale
233 catchments (catchments c2) were sampled in 2007 (n = 5; A7, D7, G7, K7, L7) and in
234 $Cal_{PLS}08_{c2}$, the same five mesoscale catchments as in $Cal_{PLS}07$, were sampled in 2008 (n = 5;

235 A8, D8, G8, K8, L8). The test data sets consist of the same groupings, but were not used in
236 the calibrations (Test$_{PLS}$00$_{c1}$, Test$_{PLS}$07$_{c2}$, Test$_{PLS}$08$_{c2}$).

237 After running the three different PLS models we determined whether the same variables
238 appeared to be important (VIP>1) in the model fittings. These variables were taken as good
239 candidates to reproduce the TOC level of the headwaters and therefore included in the mixed
240 modelling of the individual headwaters in the next step (Section 2.5.2). An additional test set
241 was created, Test$_{PLS}$02&05$_{c3}$, which was comprised of data from the two catchments sampled
242 once in 2002 and twice in 2005 (catchments c3, n=3; O2, K5s, K5w, where s stands for
243 summer and w for winter). No calibration was done on this data set, but was used for testing
244 robustness of the models with respect to seasonality.

2.5.2 Modelling individual headwaters

246 When modelling individual headwaters we want to predict specific values for each of the
247 different headwaters in all catchments. As an effort to improve these simulations, we make an
248 assumption that headwaters within the same mesoscale catchment are more similar to each
249 other than to headwaters from other mesoscale catchments due to subtle combinations of
250 physiographic, weather and other factors which combine to influence the TOC levels in ways
251 which are not readily apparent from the available map information, but might be reflected in
252 differences between the average TOC levels in the different mesoscale catchments. This
253 assumption leads to a new data structure. To model this data structure we use hierarchical
254 linear mixed models (MM; Littell et al., 2006), which allow the estimation of the correlation
255 between headwaters within the same mesoscale catchment and adjusts the analysis
256 accordingly. A MM does not allow for highly correlated explanatory variables, so the number
257 of explanatory variables must be substantially smaller than the number of observations. To fit
258 these models we use candidate explanatory variables from the PLS approach described in
259 2.5.1. In the PLS analysis some explanatory variables were excluded due to a large number of
260 zero values giving rise to a median sub-catchment value of zero for all catchments. One of
261 these parameters, lake surface area, may still be important for modelling individual
262 headwaters. If lakes have a moderately large volume (appreciable residence time) they are
263 known to influence the organic content (Eriksson, 1929; Birge and Juday, 1926). Therefore
264 lake cover surface is expected to have an influence on the prediction of individual TOC$_{HW}$,
265 and thus this variable was included as a potential predictor in the MM models even though
266 this variable was not important in the PLS modelling.

267 MMs were performed using package lme4 (version 1.1-7) in the software R (version 3.1.2) (R
268 Development Core Team, 2014). Headwater data from year 2000 (numbers of headwaters =
269 69; M0, O0, Vä0, Vi0), 2007 (n_{HW} = 138; A7, D7, G7, K7, L7) or 2008 (n_{HW} = 148; A8, D8,
270 G8, K8, L8) were used as calibration data sets (denoted Cal$_{MM}$00$_{c1}$, Cal$_{MM}$07$_{c2}$ and
271 Cal$_{MM}$08$_{c2}$, respectively), one set at a time. We have two objectives in this approach:

272   i)   To make predictions on the same headwaters but at different time points. For
273   this observations for 2008 (Test$_{MM}$08$_{c2}$) were predicted by the model
274   calibrated with data from 2007 (Cal$_{MM}$07$_{c2}$). The observations from 2007
275   (Test$_{MM}$07$_{c2}$) were also predicted based on a model calibrated using the 2008
276   data (Cal$_{MM}$08$_{c2}$).

277   ii)   To make predictions on a new set of headwaters at a different time point. For
278   this:

279   • observations from 2002 and 2005 (Test$_{MM}$02&05$_{c3}$), 2007 (Test$_{MM}$07$_{c2}$) and
280   2008 (Test$_{MM}$08$_{c2}$) respectively were predicted by models calibrated using data
281   from 2000 on an entirely different set of catchments (Cal$_{MM}$00$_{c1}$) and

282   • observations from 2000 ($Test_{MM}00_{c1}$), 2002 and 2005 ($Test_{MM}02\&05_{c3}$) were
283   predicted by models calibrated using data from 2007 or 2008 on an entirely
284   different set of catchments ($Cal_{MM}07_{c2}$ and $Cal_{MM}08_{c2}$, respectively).

285 $Test_{MM}02\&05$ is sampled in one catchment 2002 and two times in another catchment in 2005
286 ($n_{HW} = 65$; O2, K5s, K5w, where s stands for summer and w for winter), no calibration was
287 done on this data set. We used the testing to see if the calibrations worked on other data sets
288 which were not included in the calibration.

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

291   • version Out includes OutletTOC but no map information,

292   • version Map includes map information but not OutletTOC while

293   • version OutMap includes both OutletTOC and map information.

294 In total nine different MM were calibrated.

295 For fitting of the MM, Akaike information criterion (AIC) (Akaike, 1974) and p-values were
296 used. AIC is a goodness-of fit measure, which is corrected for the complexity of the model,
297 similar to the adjusted R^2. The p-values in regression models determine if parameter estimates
298 are significantly different from zero, i.e. if there is a significant relationship between an
299 explanatory variable and the response variable. The p-values were calculated according to
300 Kenward and Roger (1997) using 'krmodcomp' in R package 'pbkrtest' (version 0.4-1).
301 During the model fitting, added variables were checked to see if they increased the predictive
302 ability of the model by computing the prediction error sum of squares (PRESS). The smaller
303 the PRESS value the closer the prediction is to the observed values. Kenward and Roger
304 (1997) version of R^2 for predictions is called P^2 (similar to Q^2 for PLS). The P^2 were
305 calculated according to Méndez Mediavilla et al. (2008), with the modification that instead of
306 leave-one-out validation we compute P^2 on the evaluation test sets: $P^2 = 1-PRESS/TSS$ where
307 TSS is the total sum of squared differences between modelled values and the mean of
308 observations in the evaluation set. Median absolute errors (MedAE) and median relative
309 errors (MedRE%) were also calculated.

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

3   Results

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

3.1 Modelling headwater median

324 The first principal component (PC) of the PLS-model of median headwater TOC was
325 significant for both PLS calibration sets year 2007 and 2008, but not the second PC. No

326   significant PLS-model was established using calibration set year 2000 ($Cal_{PLS}00_{c1}$).
327   Calibration using data set 2007 ($Cal_{PLS}07_{c2}$) gave higher R^2 and Q^2 than using the data set for
328   2008 calibration ($Cal_{PLS}08_{c2}$) (Table 3). The calibration of the model using the $Cal_{PLS}08_{c2}$
329   data was evaluated by using the test data. This yielded a PRESS for the median TOC that was
330   lower than when the model was calibrated using $Cal_{PLS}07_{c2}$ (Table 3). Based on PLS-
331   modelling of median headwater TOC, suitable candidates for the mixed models (MM) of
332   individual headwaters were: altitude of sampling sites, OutletTOC, proportion of clear-felled,
333   coniferous forest, mixed forest, wet mires, coniferous forest on mires as well as the volume of
334   birch-, spruce- and total forest volume.

3.2   Modelling headwater interquartile range (IQR)

336   For both PLS calibration sets year 2007 and 2008 the first principal component (PC) was
337   significant in the PLS-models for IQR headwater TOC, but not the second PC. No significant
338   PLS-model was established using calibration set year 2000 ($Cal_{PLS}00_{c1}$). The calibration of
339   the model using the $Cal_{PLS}07_{c2}$ data was evaluated by using the test data. This yielded a
340   PRESS for TOC IQR that was lower than when the model was calibrated using $Cal_{PLS}08_{c2}$
341   (Table 3). Headwaters TOC IQR modelled by PLS indicates three variables as suitable
342   candidates for the MM: OutletTOC, proportion of clear-felled area and birch volume.

3.3   Modelling individual headwaters

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

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

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

362   The models created by this procedure are listed in Table 4. The models produced by
363   $Cal_{MM}07_{c2}$ and $Cal_{MM}08_{c2}$ were very similar and can predict the data at the same sites quite
364   well, i.e. the $Cal_{MM}07_{c2}$ model can predict the $Test_{MM}08_{c2}$ data set well and the other way
365   round (Fig. 4-5 and Table 5).

3.3.2   Modelling individual headwaters and predicting new headwaters at other time points

368   When we use the calibration set $Cal_{MM}00_{c1}$ to fit a model, the variables selected (same
369   procedure as in 3.3.1) were OutletTOC, lake surface coverage and coniferous forest on mires.
370   We evaluate this model by predicting values in the test sets $Test_{MM}07_{c2}$ and $Test_{MM}08_{c2}$ (Fig.
371   3). The best prediction model parameters are given in Table 4, with model performance in

372   Fig. 3 and error results in Table 5. Predictions for new sites in the test set $Test_{MM}00_{c1}$ and
373     $Test_{MM}02\&05_{c3}$ are less satisfactory and indicate that the models might be over fitting the
374   data.

3.3.3   Evaluation on intermediate sites

376   The headwater models were also tested on the sites of intermediate size (i.e. >2 km^2) but
377   excluding the river outlet (lower parts of Table 5). In general the intermediate sites were
378   modelled as successfully as the headwaters (Table 5). $Cal_{MM}07_{c2}$ gave predictions for the
379   intermediate sites which were not as good as for the other data sets, i.e. higher MedRE%, than
380     $Cal_{MM}08_{c2}$ and $Cal_{MM}00_{c1}$.

3.3.4   Evaluation of river outlets

382   To test the effect of including the river outlet TOC on the performance of MM predictions for
383   individual headwaters, three versions of each calibration data set was used to create three
384   different models, one using outlet TOC alone (Out), one using map information alone (Map),
385   and one using both the outlet and map information (OutMap). The map variables were the
386   same for each calibration data set but with different calibrated coefficients (Table 4). In 25 out
387   of 27 different combinations of MM (three different calibration data sets, three versions of
388   each calibration (Out, Map, OutMap) and three different test data sets), the OutMap version
389   gave the best performance with the lowest PRESS, while two Map versions (map information
390   only, no OutletTOC included) gave the lowest PRESS (Table 5 and Fig. 3-5). Similar results
391   were observed for the intermediate sites (Table 5 and Fig. 3-5). The OutMap version gave 5-
392   15% better predictions than Map only.

4   Discussion

394   In 25 of 27 tests, including river outlet chemistry (OutletTOC) resulted in lower errors in the
395   mixed models predictions of the TOC for individual headwaters, and intermediate sites,
396   compared to using map information alone. The measurements at the river outlet were
397   necessary to reproduce more correct average headwater TOC levels (Table 5). Excluding the
398   OutletTOC measurements leads to the assumption that average TOC levels in the headwaters
399   were similar in different catchment stream networks if the map information is similar, which
400   is not always true (cf. sampling 2007 and 2008). This is the first article to test how to include
401   river outlet chemistry with map information for modelling headwaters, and how well the river
402   outlet chemistry improved the models.

403   To predict the correct mean values for headwater TOC is still a challenge in our application of
404   mixed models, since the calibration sets consist of 4-5 catchment systems, and these were
405   sampled only a few times for each calibration. This clearly makes generalisations to new
406   catchments or flow-situations difficult and uncertain. Calibration sets Cal07 and Cal08 share
407   most of the catchments, but are measured during different years and perhaps more
408   importantly, during different flow situations and seasons. Even if most headwaters are the
409   same in $Cal_{MM}07_{c2}$ and $Cal_{MM}08_{c2}$, and the models produced were similar, it was still not
410   possible to account for more than about 50% of the variation in the other set of data,
411   indicating that there is large variability in time. Weather is a factor that varies with time (and
412   space) and influences stream water chemistry. In our approach we did not include the weather
413   related data (temperature, precipitation, flow) in the models (PLS and MM) since it was not
414   available for headwaters, but only for the river outlet. Presumably discharge for each
415   headwater could benefit the models, but measuring discharge at all individual headwaters
416   would have been very time consuming (and was not performed). To model discharge with
417   appropriate accuracy at all these headwaters (size<2 km^2) is so far too difficult to perform 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 were 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) has 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 at a landscape scale (Lottig et al., 2013) and lakes can also delay pulses of TOC within river networks, which is a complicating factor (Hytteborn 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 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 $Cal07_{c2}$ and $Cal08_{c2}$ (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/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

465 headwater streams. This is much better performance than an earlier attempt using river outlet
466 TOC without map information (Temnerud et al., 2010) , in which only one of seven models
467 were significant for predicting headwater median TOC and none were significant for
468 predicting headwater TOC IQR. The key factor in our approach here is the use of mixed
469 models which allow the same headwaters to have different TOC depending on weather and
470 flow etc. Since MM cannot use large numbers of correlated explanatory variables, PLS was
471 used to identify a set of candidate explanatory variables for the MM. Since our combined
472 approach increased the predictability for TOC, it would be interesting to evaluate whether the
473 method could improve prediction of headwater pH and ANC, for which models using outlet
474 catchment chemistry were already fairly successful (Temnerud et al., 2010).
475
476 In order to have the same map resolution for all catchments, due to lack of universal
477 availability of fine-scale data, a rather coarse resolution was used (e.g. 50 meter grid data for
478 altitude and soil map of scale 1.1 million). The Swedish authorities are LiDAR scanning all of
479 Sweden to build a 2 meter grid digital elevation model and are generating maps connecting all
480 watercourses up to the headwaters (through lakes and wetlands), which by 2017 will provide
481 new data that could help in modelling the headwaters. This improved map information might
482 further improve the mixed model approach demonstrated here which includes river outlet
483 chemistry. This will hopefully get us closer to the ability to predict individual headwaters that
484 are such vital building blocks of aquatic ecosystems, but remain so very difficult to model.

Acknowledgements

492

References

494 Ågren, A. M., Buffam, I., Cooper, D. M., Tiwari, T., Evans, C. D., and Laudon, H.: Can the
495   heterogeneity in stream dissolved organic carbon be explained by contributing
496   landscape elements?, Biogeosciences, 11, 1199-1213, doi:10.5194/bg-11-1199-2014,
497   2014.

498 Akaike, H.: A new look at the statistical model identification, IEEE Transactions on
499   Automatic Control, 19, 716 - 723, doi:10.1109/TAC.1974.1100705, 1974.

500 Alexander, R. B., Boyer, E. W., Smith, R. A., Schwarz, G. E., and Moore, R. B.: The role of
501   headwater streams in downstream water quality, JAWRA, 43, 41-59,
502   doi:10.1111/j.1752-1688.2007.00005.x, 2007.

503 Andersson, J.-O., and Nyberg, L.: Using official map data on topography, wetlands and
504   vegetation cover for prediction of stream water chemistry in boreal headwater
505   catchments, HESS, 13, 537-549, doi:10.5194/hess-13-537-2009, 2009.

506 Arheimer, B., and Lindström, G.: Implementing the EU Water Framework Directive in
507   Sweden, in: Runoff Predictions in Ungauged Basins – Synthesis across Processes,
508   Places and Scales, edited by: Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and
509   Savenije, H., Cambridge University Press, Cambridge, 353-359, 2013.

510 Birge, E. A., and Juday, C.: The organic content of lake water, PNAS, 12, 515-519, 1926.

511 Bossard, M., Feranec, J., and Otahel, J.: CORINE land cover technical guide – Addendum
512   2000, European Environment Agency, Copenhagen, 105 pp.,
513   http://www.eea.europa.eu/publications/COR0-landcover, 2000.

514 Brandtberg, P. O., Lundkvist, H., and Bengtsson, J.: Changes in forest-floor chemistry caused
515   by a birch admixture in Norway spruce stands, For. Ecol. Manage., 130, 253-264,
516   doi:10.1016/S0378-1127(99)00183-8, 2000.

517 Buffam, I., Laudon, H., Temnerud, J., Mörth, C.-M., and Bishop, K.: Landscape-scale
518   variability of acidity and dissolved organic carbon during spring flood in a boreal
519   stream network, J. Geophys. Res. Biogeosci., 112, G01022,
520   doi:10.1029/2006JG000218, 2007.

521 Clark, J. M., Bottrell, S. H., Evans, C. D., Monteith, D. T., Bartlett, R., Rose, R., Newton, R.
522   J., and Chapman, P. J.: The importance of the relationship between scale and process in
523   understanding long-term DOC dynamics, Sci. Total Environ., 408, 2768-2775,
524   doi:10.1016/j.scitotenv.2010.02.046, 2010.

525 Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A., and Dillon, P. J.: Cryptic wetlands:
526   Integrating hidden wetlands in regression models of the export of dissolved organic
527   carbon from forested landscapes, Hydrol. Process., 17, 3629-3648,
528   doi:10.1002/hyp.1357, 2003.

529 Edström, M., and Rystam, P.: Network of stations for field research area (in Swedish: FFO -
530   Stationsnät för fältforskningsområden), Swedish Meteorological and Hydrological
531   Institute, Norrköping, SMHI hydrology report 53, 53 pp., 1994.

532   Eriksson, J. V.: Den kemiska denudationen i Sverige (The chemical denudation of Sweden, in
533   Swedish with a French summary), Swedish Meteorological and Hydrological Institute,
534   Stockholm, Report Band 5, nr 3, 95 pp., 1929.

535   Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., and Wold, S.: Multi-
536   and megavariate data analysis – Part 1: Basic Principles and Applications, Umetrics AB,
537   Umeå, 2006.

538   Evans, C., Cooper, D., Monteith, D., Helliwell, R., Moldan, F., Hall, J., Rowe, E., and Cosby,
539   B.: Linking monitoring and modelling: Can long-term datasets be used more effectively
540   as a basis for large-scale prediction?, Biogeochemistry, 101, 211-227,
541   doi:10.1007/s10533-010-9413-x, 2010.

542   Fröberg, M., Hansson, K., Kleja Berggren, D., and Alavi, G.: Dissolved organic carbon and
543   nitrogen leaching from Scots pine, Norway spruce and silver birch stands in southern
544   Sweden, For. Ecol. Manage., 262, 1742-1747, doi:10.1016/j.foreco.2011.07.033, 2011.

545   Gadmar, T. C., Vogt, R. D., and Østerhus, B.: The merits of the high temperature combustion
546   method for determining the amount of natural organic carbon in surface freshwater
547   samples, Intern. J. Environ. Anal. Chem., 82, 451-461,
548   doi:10.1080/0306731021000018099, 2002.

549   Geladi, P., and Kowalski, B. R.: Partial least-squares regression - a tutorial, Anal. Chim. Acta,
550   185, 1-17, doi:10.1016/0003-2670(86)80028-9, 1986.

551   Gomi, T., Sidle, R. C., and Richardson, J. S.: Understanding processes and downstream
552   linkages of headwater systems, Bioscience, 52, 905-916, doi:10.1641/0006-
553   3568(2002)052[0905:UPADLO]2.0.CO;2, 2002.

554   Hope, D., Billett, M. F., and Cresser, M. S.: Exports of organic carbon in two river systems in
555   NE Scotland, J. Hydrol., 193, 61-82, doi:10.1016/S0022-1694(96)03150-2, 1997.

556   Hutchins, M. G., Smith, B., Rawlins, B. G., and Lister, T. R.: Temporal and spatial variability
557   of stream waters in Wales, the Welsh borders and part of the West Midlands, UK. 1.
558   Major ion concentrations, Wat. Res., 33, 3479-3491, doi:10.1016/S0043-
559   1354(99)00057-3, 1999.

560   Hytteborn, J. K., Temnerud, J., Alexander, R. B., Boyer, E. W., Futter, M. N., Fröberg, M.,
561   Dahné, J., and Bishop, K. H.: Patterns and predictability in the intra-annual organic
562   carbon variability across the boreal and hemiboreal landscape, Sci. Total Environ., 520,
563   260–269, doi:10.1016/j.scitotenv.2015.03.041, 2015.

564   Ivarsson, H., and Jansson, M.: Temporal variations in the concentration and character of
565   dissolved organic matter in a highly colored stream in the coastal zone of northern
566   Sweden, Arch. Hydrobiol., 132, 45-55, 1994.

567   Jager, H. I., Bevelhimer, M. S., King, R. L., and Smith, K. A.: Landscape influences on
568   headwater streams on Fort Stewart, Georgia, USA, Environ. Manage., 48, 795-807,
569   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, doi:10.2166/nh.2000.013, 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 Bridgham, 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. Biogeosci., 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., Strömqvist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, Hydrol. Res., 41, 295–319, doi:10.2166/nh.2010.007, 2010.

609 Littell, R., Milliken, G., Stroup, W., Wolfinger, R., and Schabenberger, O.: SAS for Mixed
610   Models, 2nd ed., SAS Institute, Cary, 828 pp., 2006.

611 Löfgren, S., Aastrup, M., Bringmark, L., Hultberg, H., Lewin-Pihlblad, L., Lundin, L.,
612   Karlsson, G. P., and Thunholm, B.: Recovery of soil water, groundwater, and
613   streamwater from acidification at the Swedish Integrated Monitoring catchments,
614   AMBIO, 40, 836-856, doi:10.1007/s13280-011-0207-8, 2011.

615 Löfgren, S., Fröberg, M., Yu, J., Nisell, J., and Ranneby, B.: Water chemistry in 179
616   randomly selected Swedish headwater streams related to forest production, clear-felling
617   and climate, Environ. Monit. Assess., 186, 8907-8928, doi:10.1007/s10661-014-4054-5,
618   2014.

619 Lottig, N. R., Buffam, I., and Stanley, E. H.: Comparisons of wetland and drainage lake
620   influences on stream dissolved carbon concentrations and yields in a north temperate
621   lake-rich region, Aquat. Sci., 75, 1-12, doi:10.1007/s00027-013-0305-8, 2013.

622 Lyon, S. W., Braun, H., and Temnerud, J.: Accounting for instream lakes when interpolating
623   stream water chemistry observations, JoSH, 11, 1-20,
624   http://www.spatialhydrology.net/index.php/JOSH/article/view/102, 2011.

625 Lyon, S. W., Nathanson, M., Spans, A., Grabs, T., Laudon, H., Temnerud, J., Bishop, K. H.,
626   and Seibert, J.: Specific discharge variability in a boreal landscape, Water Resour. Res.,
627   48, W08506, doi:10.1029/2011WR011073, 2012.

628 MacDonald, L. H., and Coe, D.: Influence of headwater streams on downstream reaches in
629   forested areas, Forest Science, 53, 148-168, 2007.

630 Mattsson, T., Finér, L., Kortelainen, P., and Sallantaus, T.: Brook water quality and
631   background leaching from unmanaged forested catchments in Finland, Water Air Soil
632   Pollut., 147, 275-297, doi:10.1023/A:1024525328220, 2003.

633 McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., and Bailey, S. W.:
634   Network analysis reveals multiscale controls on streamwater chemistry, PNAS, 111,
635   7030-7035, doi:10.1073/pnas.1404820111, 2014.

636 Méndez Mediavilla, F. A., Landrum, F., and Shah, V.: A comparison of the coefficient of
637   predictive power, the coefficient of determination and AIC for linear regression,
638   Proceedings of the 39th Annual Meeting of the Decision Sciences Institute, Atlanta,
639   2008.

640 Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., and Leonard, N. E.:
641   The contribution of headwater streams to biodiversity in river networks, JAWRA, 43,
642   86-103, doi:10.1111/j.1752-1688.2007.00008.x, 2007.

643 Mulholland, P. J., Fellows, C. S., Tank, J. L., Grimm, N. B., Webster, J. R., Hamilton, S. K.,
644   Martí, E., Ashkenas, L., Bowden, W. B., Dodds, W. K., McDowell, W. H., Paul, M. J.,
645   and Peterson, B. J.: Inter-biome comparison of factors controlling stream metabolism,
646   Freshw. Biol., 46, 1503-1517, doi:10.1046/j.1365-2427.2001.00773.x, 2001.

647 Müller, R. A., Futter, M. N., Sobek, S., Nisell, J., Bishop, K., and Weyhenmeyer, G. A.:
648   Water renewal along the aquatic continuum offsets cumulative retention by lakes:

649   Implications for the character of organic carbon in boreal lakes, Aquat. Sci., 75, 535-
650   545, doi:10.1007/s00027-013-0298-3, 2013.

651   Nagorski, S. A., Moore, J. N., McKinnon, T. E., and Smith, D. B.: Scale-dependent temporal
652   variations in stream water geochemistry, Environ. Sci Technol., 37, 859-864,
653   doi:10.1021/es025983+, 2003.

654   Oni, S. K., Futter, M. N., Bishop, K., Köhler, S. J., Ottosson-Löfvenius, M., and Laudon, H.:
655   Long-term patterns in dissolved organic carbon, major elements and trace metals in
656   boreal headwater catchments: Trends, mechanisms and heterogeneity, Biogeosciences,
657   10, 2315-2330, doi:10.5194/bg-10-2315-2013, 2013.

658   Pers, C., Rahm, L., Jonsson, A., Bergström, A.-K., and Jansson, M.: Modelling dissolved
659   organic carbon turnover in humic Lake Örträsket, Sweden, Environ. Model. Assess., 6,
660   159-172, doi:10.1023/A:1011953730983, 2001.

661   Person, H. S., Johnston, C. E., and Beall, R. T.: Little waters; A study of headwater streams &
662   other little waters, their use and relations to the land, The Upstream Engineering
663   Conference, Washington D.C., 1-82,
664   http://www.archive.org/details/littlewatersstud00persrich, 1936.

665   Reese, H., Nilsson, M., Sandström, P., and Olsson, H.: Applications using estimates of forest
666   parameters derived from satellite and forest inventory data, Comput Electron Agri, 37,
667   37-55, doi:10.1016/S0168-1699(02)00118-7, 2002.

668   Reese, H., Nilsson, M., Granqvist Pahlén, T., Hagner, O., Joyce, S., Tingelöf, U., Egberth,
669   M., and Olsson, H.: Countrywide estimates of forest variables using satellite data and
670   field data from the national forest inventory, AMBIO, 33, 542-548, doi:10.1579/0044-
671   7447-32.8.542, 2003.

672   Sakamaki, T., and Richardson, J. S.: Nonlinear variation of stream–forest linkage along a
673   stream-size gradient: An assessment using biogeochemical proxies of in-stream fine
674   particulate organic matter, J. Appl. Ecol., 50, 1019-1027, doi:10.1111/1365-
675   2664.12099, 2013.

676   Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M.: Environmental Organic
677   Chemistry, 2 ed., John Wiley, Hoboken, 1313 pp., 2003.

678   Shao, J.: Linear model selection by cross-validation, JASA, 88, 486-494,
679   doi:10.1080/01621459.1993.10476299, 1993.

680   Smart, R. P., Soulsby, C., Cresser, M. S., Wade, A. J., Townend, J., Billett, M. F., and
681   Langan, S.: Riparian zone influence on stream water chemistry at different spatial
682   scales: A GIS-based modelling approach, an example for the Dee, NE Scotland, Sci.
683   Total Environ., 280, 173-193, doi:10.1016/S0048-9697(01)00824-5, 2001.

684   Strayer, D. L., Beighley, R. E., Thompson, L. C., Brooks, S., Nilsson, C., Pinay, G., and
685   Naiman, R. J.: Effects of land cover on stream ecosystems: Roles of empirical models
686   and scaling issues, Ecosystems, 6, 407-423, doi:10.1007/PL00021506, 2003a.

687 Strayer, D. L., Ewing, H. A., and Bigelow, S.: What kind of spatial and temporal details are
688   required in models of heterogeneous systems?, Oikos, 102, 654-662,
689   doi:10.1034/j.1600-0706.2003.12184.x, 2003b.

690 Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and Lindström, G.: Water and nutrient
691   predictions in ungauged basins: Set-up and evaluation of a model at the national scale,
692   Hydrolog. Sci. J., 57, 229-247, doi:10.1080/02626667.2011.637497, 2012.

693 Temnerud, J., and Bishop, K.: Spatial variation of streamwater chemistry in two Swedish
694   boreal catchments: Implications for environmental assessment, Environ. Sci Technol.,
695   39, 1463-1469, doi:10.1021/es040045q, 2005.

696 Temnerud, J., Seibert, J., Jansson, M., and Bishop, K.: Spatial variation in discharge and
697   concentrations of organic carbon in a catchment network of boreal streams in Northern
698   Sweden, J. Hydrol., 342, 72-87, doi:10.1016/j.jhydrol.2007.05.015, 2007.

699 Temnerud, J., Fölster, J., Pilström, F., and Bishop, K.: Synoptisk provtagning av små
700   vattendrag i södra Sverige oktober 2007 (Synoptical sampling of small watercourses in
701   southern Sweden in October 2007, in Swedish), Swedish University of Agricultural
702   Sciences, Uppsala, Report 2009:6, 56 pp., http://publikationer.slu.se/Filer/Rapport2009-
703   6.pdf, 2009.

704 Temnerud, J., Fölster, J., Buffam, I., Laudon, H., Erlandsson, M., and Bishop, K.: Can the
705   distribution of headwater stream chemistry be predicted from downstream
706   observations?, Hydrol. Process., 24, 2269-2276, doi:10.1002/hyp.7615, 2010.

707 Walker, C. M., King, R. S., Whigham, D. F., and Baird, S. J.: Landscape and wetland
708   influences on headwater stream chemistry in the Kenai Lowlands, Alaska, Wetlands,
709   32, 301-310, doi:10.1007/s13157-011-0260-x, 2012.

710 Wetzel, R. G.: Limnology. Lake and River Ecosystems, 3 ed., Academic Press, London, 1006
711   pp., 2001.

712 Weyhenmeyer, G. A., Fröberg, M., Karltun, E., Khalili, M., Kothawala, D., Temnerud, J., and
713   Tranvik, L. J.: Selective decay of terrestrial organic carbon during transport from land to
714   sea, Glob. Change Biol., 18, 349-355, doi:10.1111/j.1365-2486.2011.02544.x, 2012.

715 Wolock, D. M., Fan, J., and Lawrence, G. B.: Effects on basin size on low-flow stream
716   chemistry and subsurface contact time in the Neversink river watershed, New York,
717   Hydrol. Process., 11, 1273-1286, doi:10.1002/(SICI)1099-
718   1085(199707)11:9<1273::AID-HYP557>3.0.CO;2-S, 1997.
719
720

Tables

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[-1]) 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 (km²)	T	P	q
c2	Anråse å[a] (A)	58º01'; 11º51'	74	7.8	983	0.95
c2	Danshytteån[a] (D)	59º42'; 15º05'	72	6.1	774	0.70
c2	Getryggsån[a] (G)	59º48'; 15º17'	32	5.9	800	0.84
c2, c3	Krycklan[b] (K)	64°14'; 19°46'	61	2.6	659	0.51
c2	Lugnån[a] (L)	57º06'; 14º48'	122	6.1	831	0.60
c1	Mangslidsälven (M)	60°23', 12°54'	235	4.9	823	0.75
c1, c3	Ottervattsbäcken[c] (O)	64°02'; 19°06'	71	2.6	659	0.50
c1	Vänjaurbäcken[c] (Vä)	64°19', 18°43'	200	2.2	649	0.47
c1	Viggan (Vi)	60°21', 12°46'	116	3.9	870	0.85

[a] Temnerud *et al.* (2009)
[b] Laudon et al. (2013)
[c] Temnerud and Bishop (2005)

732 Table 2. Median concentration of total organic carbon (TOC, mg L^{-1}), with 25th and 75thpercentiles in
733 brackets, plus weather parameters (T, P, q) using the median value from the 30 days prior to sampling as
734 modelled by SMHI (See Methods for more details). Clusters are the different groups of calibration and test
735 sets used in the modelling. Sets are the different data sets used, with acronyms corresponding to river
736 names (Table **Error! Reference source not found.**) and the last digit of the sampling year. M = Month of
737 sampling, HW is headwaters and n_{HW} = number of sampled HW.

Cluster	Set	Year	M	n_{HW}	TOC_{HW}	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

738
739

740 Table 3. Partial least square regression (PLS) results
741 indicating the goodness of fit for the prediction of
742 median and the interquartile range (IQR) of headwater
743 total organic carbon (TOC) concentration (mg L^{-1}). Cal
744 is calibrated and 00, 07 and 08 refer to sampling year
745 (2000, 2007 and 2008) of catchment (number=n).
746 PRESS is the prediction error sum of squares.

Cal	Var	n	R^2	Q^2	PRESS
00_{c1}	Median	4	ns	ns	ns
	IQR	4	ns	ns	ns
07_{c2}	Median	5	96	94	581
	IQR	5	96	92	346
08_{c2}	Median	5	90	83	192
	IQR	5	54	17	530

747
748

749 Table 4. Coefficients for the best-fit hierarchical linear mixed models (MM), where Log_{10} (TOC_{HW}) is the
750 response variable. See the Method section for more details.

MM model	Version	Intercept	OutletTOC (mg L^{-1})	Lake surface coverage[p]	Coniferous forest on mires[p]	Broad-leaved forest[p]	Sites altitude (m a.s.l.)
Cal00_{c1}	Out	0.861	0.0245				
	Map	1.156		-1.509	0.592		
	OutMap	0.885	0.0201	-1.479	0.568		
Cal07_{c2}	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_{c2}	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

751 [p] is the proportion of the entire catchment area covered by this particular feature.
752
753

Cal	Version	00_{c1}	00_{c1}	00_{c1}	07_{c2}	07_{c2}	07_{c2}	08_{c2}	08_{c2}	08_{c2}
Cal n_{HW}			69			138			148	
$Test_{HW}$		$02\&05_c$	07_{c2}	08_{c2}	00_{c1}	$02\&05_c$	08_{c2}	00_{c1}	$02\&05_c$	07_{c2}
n $Test_{HW}$		65(3)	138	148	69	65(3)	148	69	65(3)	138
$PRESS_{HW}$	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_{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
$MedAE_{HW}$	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
$MedRE\%_{HW}$	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_{is}		129	155	135	82	129	135	82	129	155
$PRESS_{is}$	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_{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
$MedAE_{is}$	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
$MedRE\%_{is}$	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

766

767

Figures legend

Fig. 1. Map of Sweden with the nine investigated catchments, see Table **Error! Reference source not found.** for coordinates. Labels in brackets are the abbreviated names of the catchments and year of sampling.

Fig. 2. Total organic carbon (TOC in mg L^{-1}) as a function of catchment size (note 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 Vä0 and q) is Vi0. See Table 1 and Table 2 for description of surveys.

Fig. 3. Scatterplots of measured headwaters with total organic carbon (TOC in mg L^{-1}) on the x-axis, and the three different versions of the mixed models Cal$_{MM}$00$_{c1}$ 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$_{MM}$00$_{c1}$) 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. Getryggsån by rectangles, R. Krycklan by triangles (winter 2005 by upside-down triangles), R. Lugnån by squares, R. Mangslidsälven by multiplication sign, R. Ottervattsbäcken by up-side-down triangles, R. Vänjaurbäcken by right tilted triangles and R. Viggan by plus sign. The black line is the 1:1-line.

Fig. 4. Scatterplots of measured headwaters with total organic carbon (TOC in mg L^{-1}) on the x-axis, and the three different versions of the mixed models Cal$_{MM}$07$_{c2}$ on the y-axis: Out version on the left panel, OutMap version on the right panel and Map in-between. Data for year 2007 (Cal$_{MM}$07$_{c2}$) on the third row in red text, followed by Test data; first row 2000 data, second row 2002 & 2005 data and the last row is 2008 data. R. Anråse å indicated by circles, R. Danshytteån by diamonds, R. Getryggsån by rectangles, R. Krycklan by triangles (winter 2005 by upside-down triangles), R. Lugnån by squares, R. Mangslidsälven by multiplication sign, R. Ottervattsbäcken by up-side-down triangles, R. Vänjaurbäcken by right tilted triangles and R. Viggan by plus sign. The black line is the 1:1-line.
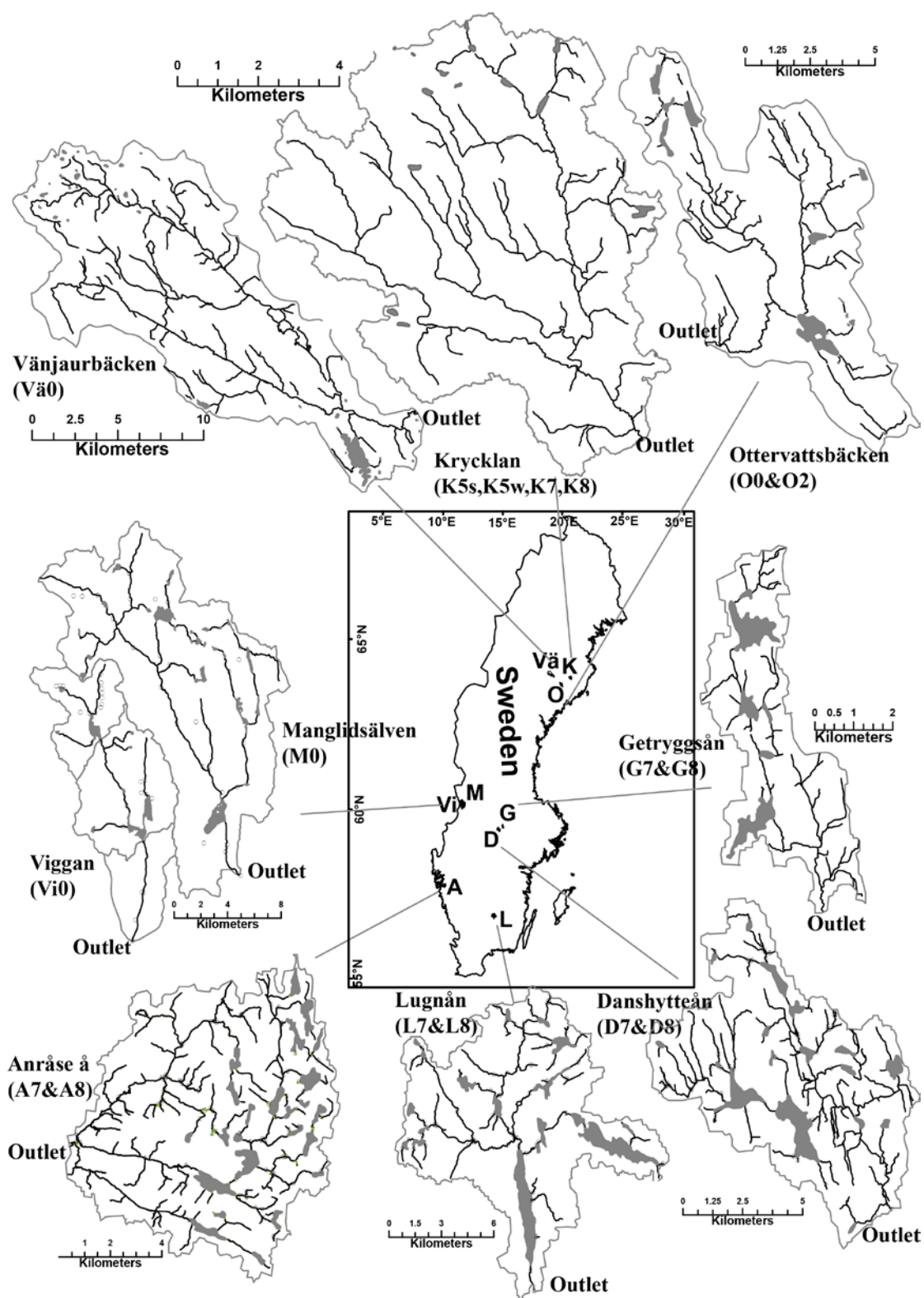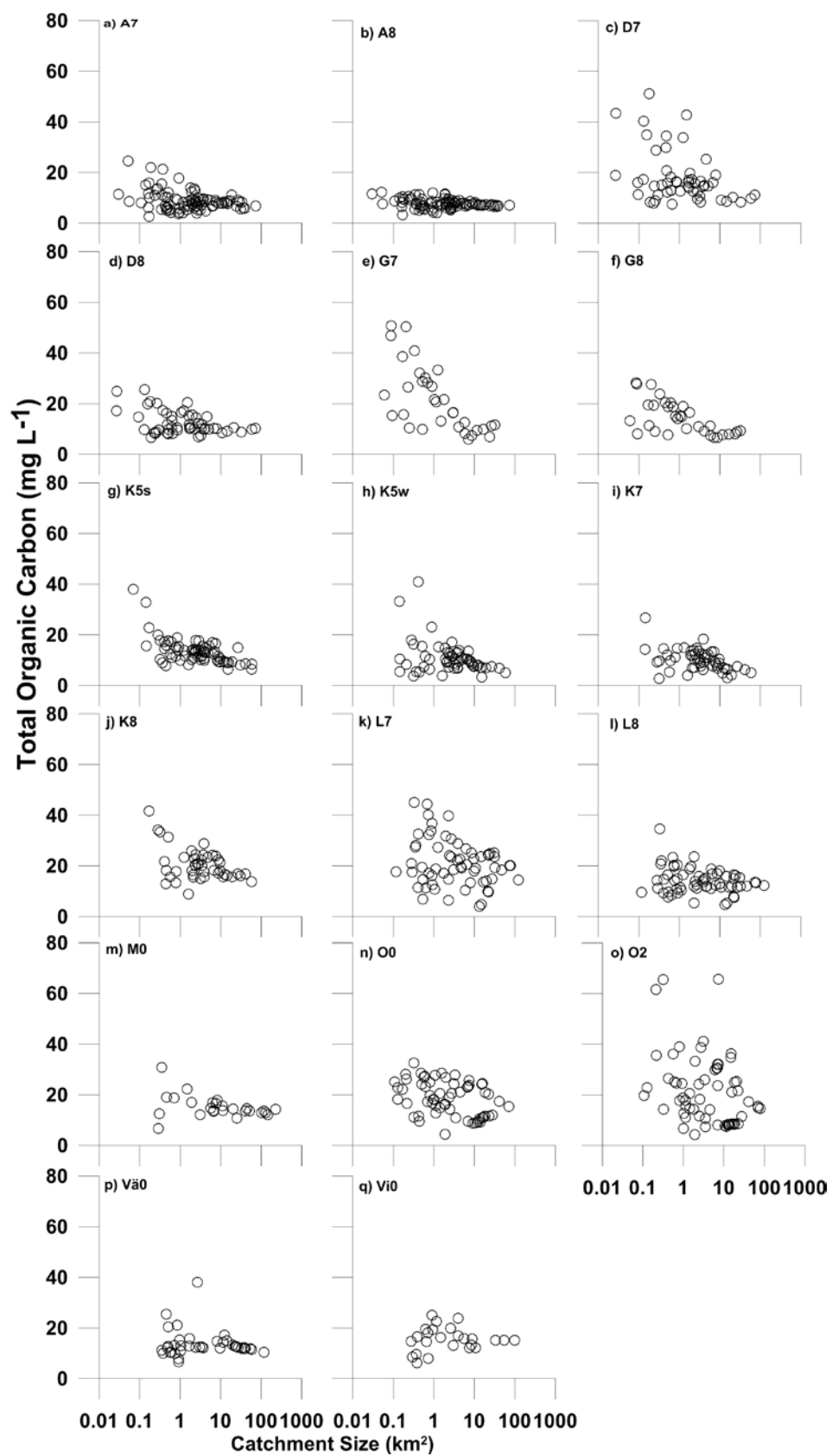
Fig. 5. Scatterplots of measured headwaters with total organic carbon (TOC in mg L^{-1}) on the x-axis, and the three different versions of the mixed models Cal08 on the y-axis: Out version on the left panel, OutMap version on the right panel and Map in-between. Data for year 2008 (Cal$_{MM}$08$_{c2}$) 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. Ottervattsbäcken by up-side-down triangles, R. Vänjaurbäcken by right tilted triangles and R. Viggan by plus sign. The black line is the 1:1-line.
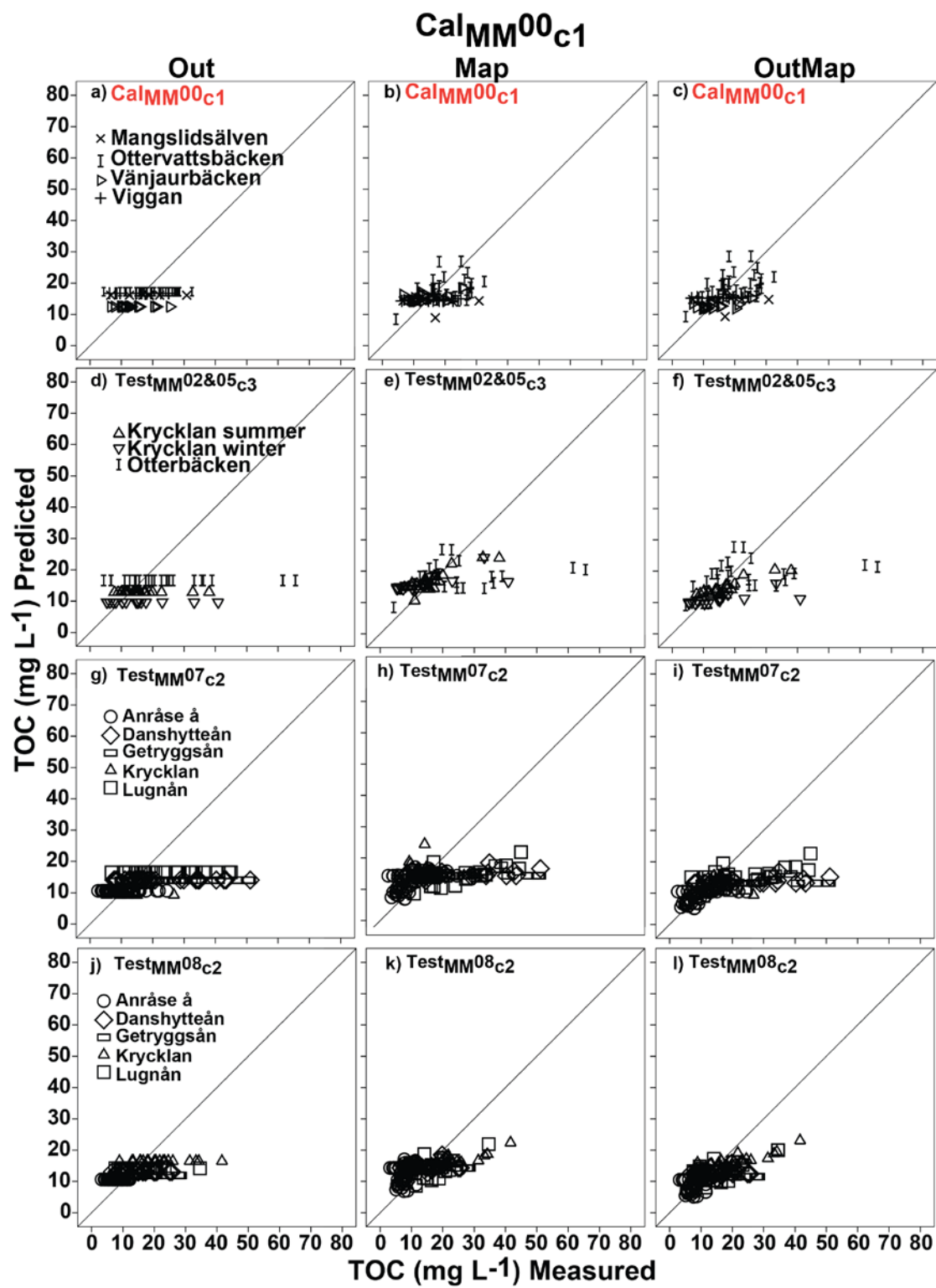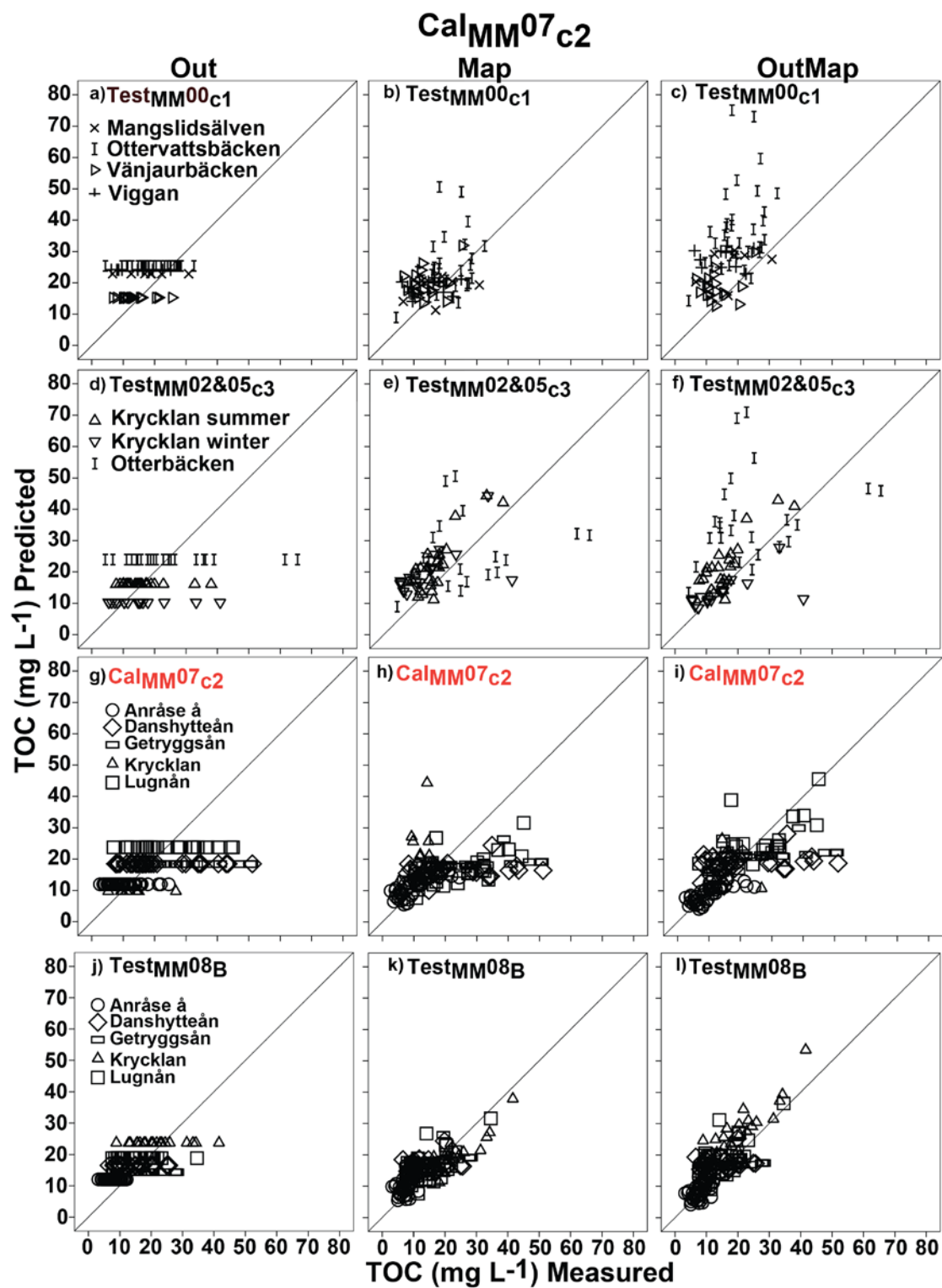
Fig. 1

809

Fig. 2

810
811

Fig. 3

812
813

Fig. 4

Fig. 5