Graduate Thesis in Soil Science 2005

Predicting Dissolved Organic Carbon Concentrations in Swedish Boreal Streams from Map Information

Anders Lindsjö
Department of Environmental Assessment
Swedish University of Agricultural Sciences, SLU

Supervisors: Prof. Kevin Bishop, Department of environmental Assesment, SLU
Johan Temnerud, Man-Technology-Environment Research centre, Örebro University
Martin Erlandsson, Department of environmental Assesment, SLU

2005 Rapport 2005:10

Predicting Dissolved Organic Carbon Concentrations in Swedish Boreal Streams from Map Information

ISSN 1403-97710

"Streams are gutters down which flow the ruins of continents."

Luna Leopold

Abstract

When sampling watercourses for environmental assessment it is not feasible to sample all the headwaters or at least not a large fraction of them. This because of the great number of headwaters and their variability in both flow and water chemistry. So if map data could help to predict the water chemistry found on specific headwaters this would be a great asset for environmental assessment of streams.

This study addresses the possibility of predicting the dissolved organic carbon (DOC) concentrations of headwaters during low flow conditions in boreal catchment networks from map information (land-use, soil and vegetation data). DOC was chosen as the test variable to model in this study because of its importance for aquatic ecosystems and strong couplings to other chemistry data (implicit in this is that if DOC is predicted well other chemistry could also be predicted well). Statistical models relating water chemistry to map information were derived from DOC measurements on different catchments in the boreal zone of Sweden using multivariate techniques. The sizes of the studied catchments ranged between 0.01-346 km^2.

Up to 62% of the spatial variation in DOC-concentrations could be explained by the information on any single map and a combination of soil and land-use map variables explained up to 90% of the variation in DOC. As other studies have found, the coverage of peat showed a positive correlation with DOC. The scale of the map material used in modelling did not seam to matter, with the 1:100 000 land-use map working as well as the 1:20 000 scale map, if not better. Little success was however found in making predictive models of DOC, but the likelihood of a model being close to acceptable was greater for transfers in time (predicting concentrations during a different year on the sites used for calibration) than in space (predicting concentrations on a nearby catchment.) Relationships between map information and DOC from downstream sites could not model data upstream on the same catchment. The difficulty in predicting low flow DOC from map variables is a complication when trying to develop tools for the environmental assessment of headwaters.

Sammanfattning

Vid provtagning av vattendrag för bedömning av deras miljömässiga tillstånd är det inte genomförbart att ta prover vid varje enskilt källflöde eller ens vid ett större antal av dem. Trots att den största sträckan av ett vattendrag utgörs av källflöden och att den största variabiliteten i både vattenflöde och vattenkemi hittas i källflödena. Därför vore det till stor fördel om kartinformation kunde användas till att förutsäga vattenkemin i vattendrag, främst källflödena, under granskningen av deras miljömässiga tillstånd.

Föreliggande examensarbetete har utvärderat möjligheten att utifrån kartinformation (markanvändning, jord och vegetationsdata) förutsäga halter av löst organiskt kol (DOC, engelsk term: dissolved organic carbon) i bäckar inom den boreala delen av Sverige. Att DOC valdes som variabel är på grund av dess stora betydelse för de akvatiska ekosystemen och dess starka kopplingar till annan vattenkemi (underförstått i detta är att om det är möjligt att förutsäga DOC så borde det vara möjligt för andra kemiska parametrar). Statistiska modeller som relaterar DOC-koncentrationer med kartmaterial från olika avrinningsområden i den boreala zonen i Sverige gjordes med multivariata metoder. Storleken på de studerade delavrinningsområdena var mellan 0,01 och 346 km^2.

Upp till 62% av den rumsliga variationen av DOC-koncentrationer kunde förklaras med information från en enskild karta och en kombination av kartmaterial kunde förklara upp till 90% av variationen av DOC. Andelen torvmark inom ett område var positivt korrelerat till koncentrationen av DOC, vilket är i linje med tidigare forskning. Den blå kartan (skala 1:100 000) gav något bättre resultat än den gula kartan i skala 1:20 000. Dock var inte någon av förutsägelserna av erforderlig precision. Emellertid fungerade modellerna bättre vid överföring i tid (förutsäga DOC under ett annat år men på samma område som modellen var byggd på) än för rumslig överföring (förutsäga DOC-koncentrationer på ett närliggande område). Modeller baserade på stora bäckar kunde inte förutsäga källflödenas DOC-koncentrationer, och vice versa. Svårigheten i att förutsäga DOC utifrån kartmaterial är en komplikation i arbetet med att utveckla verktyg för bedömningen av miljötillståndet i vattendrag och framförallt i källflödena.

Introduction

When sampling water chemistry in watercourses for environmental assessment, only a limited number of sites can be sampled in space and on a limited number of occasions during the year. The sampling sites are often somewhere downstream from the headwaters, thus aggregating a large part of the catchment stream network into one sample. This gives general values of the characteristics of the whole catchment. A problem with this is that the assessment of the upstream subcatchments will be based upon the downstream sample.

Recent research has highlighted the great variability among headwaters, Temnerud and Bishop (2005) found that in parts of boreal Sweden during low flow conditions that in streams with catchment areas larger than 15 km^2 an relatively stable water chemistry was found. But in catchments smaller than 15 km^2, water chemistry was much more variable between neighbouring streams (Temnerud and Bishop, 2005).The spatial variability of water quality is therefore a challenge for the assessment and management of water resources and aquatic ecosystems, in so far as headwaters are of interest for assessment.

The focus on water resources further downstream has meant that the significance of the headwaters is not well documented. But there is reason to believe that these headwaters are critical to downstream habitats, not least because most of the stream length, and therefore much of the aquatic habitat is found in headwaters (Doppelt et al., 1993; Shreve, 1969). Furthermore, even though many of the resource values are found well downstream from headwaters (fisheries, water supplies), human intervention (e.g. silviculture, infrastructure, agriculture) often takes place in headwaters.

The great number of headwaters and their variability in both flow and chemistry, however, means that it is not feasible to sample all, or even a large fraction of the headwaters. If map data could help predict the water chemistry found on specific headwaters, this would be a great asset for environmental assessment. The main goal of this study is to see the extent to which this is possible for dissolved organic carbon (DOC), an important water chemistry parameter in the boreal environment (Wetzel, 2001). This will be done by building statistical models of water chemistry from map material and DOC measurements from four different catchments in the boreal zone of Sweden. Dissolved organic carbon was chosen as test variable because of its strong couplings to other water chemistry variables. (Implicit in this is that the prediction of DOC would be an important and necessary step towards a complete prediction of water chemistry in headwaters.)

The amount of organic matter is considered as the most important factor in determining the biotic function of a riverine ecosystem, with the dissolved fraction being a key source of energy (McDonald et al, 2004). DOC also influences stream chemistry through its ability to complex many metals, organic pollutants and nutrients (Schwarzenbach et al., 2003; Stumm and Morgan 1996; Wetzel, 2001). It also affects the acidity and buffering capacity of streams (Dillon and Molot, 1997; Köhler et al., 1999). Furthermore the humic substances in DOC influence light penetration because humic substances are coloured (Eckhardt and Moore, 1990).

The DOC in Swedish boreal streams represents about 95% of the total organic carbon in stream water (Bishop and Pettersson, 1996). The concentration of DOC in natural freshwaters is in the range 0.5 to 50 mg/l, but values around 100 mg/l can be observed (Thurman, 1985). The sources of carbon in watercourses can be divided into three categories: allochtonous, autochthonous and anthropogenic (Industrial, agricultural and domestic sources and sinks) (Eatherall et al, 1998).

Allochtonous sources include all the organic matter that the stream receives from production that occurred outside the stream channel e g terrestrial production (Allan, 1995). The water that enters the stream either comes from groundwater discharge or surface runoff. This means that the amounts of carbon that enters the stream will vary both with the soil characteristics and seasonal changes in the hydrology of the catchment. Organic carbon enters the catchment both through precipitation and through leaching and decomposition of vegetation and soil organic horizons (Eckhardt and Moore, 1990, Eatherall et al, 1998). The carbon in the soil water can be removed through adsorption to soil particles and microbial decomposition (Elder et al, 2000).

Autochthonous organic matter is produced within the stream. Primary production by periphyton, macrophytes and phytoplankton are the most important sources of autochthonous matter. However leaf shredding invertebrates can increase the release of DOC, especially during low flow periods. DOC can also be removed from the stream water by both abiotic and biotic processes. The most important biotic process is uptake by micro organisms, assimilation of organic carbon into microbial biomass and remineralization to carbon dioxide by respiration. The abiotic factors are adsorption and flocculation, which also makes it more available to organisms. (Allan, 1995)

A number of studies have found that somewhere between 50-75% of the variance in DOC concentration in surface waters can be predicted from map information, with some measure of peat soils or wetlands being the most important (positively correlated) factor in environments similar to Swedish conditions. (Aitkenhead et al., 1999; Eatherall et al., 1998; Eckhardt and Moore, 1990; Elder et al, 2000; Gergel et al., 1999; Kortelainen, 1993; Kortelainen and Saukonen, 1998) Wetlands are important both as large stores of carbon and when connected hydrologically to streams, a large source of DOC to streams (Bengtsson and Törneman, 2004; Creed et al., 2003; Hope et al., 1994; Laudon et al., 2004; Mullholland and Kuenzler, 1979). Landscape elements known to decrease the DOC concentrations in streams are lakes of moderate volume (Meili, 1992; Pers et al., 2001) Most of these studies, however, have been conducted on larger catchments. Fewer have tried to predict the chemistry for the headwaters, and none have had data to test the prediction of all the headwaters within a single basin. This study seeks to address this issue of predicting DOC in headwaters in an entire boreal catchment network from map information; in the hope that this may lead to a better understanding of how to address the spatial variability of headwaters in environmental assessment.

The aim of this study is to predict the DOC-concentrations in boreal catchments from landscape information derived from maps and remotely sensed data with a focus on the low flow conditions that prevail during most of the year. The approach was to build statistical models, using multivariate techniques that predict DOC from using landscape information and synoptic water chemistry measurements at different locations within two different catchments in the boreal zone of Sweden. This forested landscape is relatively uniform in many regards, with little overt human influence. While it may seem intuitive that a contiguous, relatively homogeneous landscape should be easier to model than a heterogeneous landscape, some studies has pointed out that it is actually more difficulty to predict the variability of water chemistry in a homogeneous landscape due to the similarity in the landscape factors that often are associated with differences in water chemistry. (Herlihy et al., 1998; Rapp et al., 1985)

Another concern in modelling an entire catchment network is that sites downstream from the headwaters are not independent values from the point of linear regression calibration and validation i.e. when sampling a whole catchment you are sampling the same water. So care was

taken in this study to either use multivariate approaches were covariance is accepted or only using independent data in regression models.

Digital databases in Sweden provide a large number of map variables at varying scales of resolution. The large range of explanatory variables means that the paper begins with multivariate, non-parametric methods PCA and PLS to identify the most important clusters of variables. Such "models" cannot be readily transferred to different sites for making predictions, so multiple linear regression (MLR) is then used on indicator variables from the major independent clusters of variables to arrive at predictive models that might be transferable to other catchments. Several data sets of catchment networks for low flow conditions are available for calibration and testing of these models. The size of these catchment networks ranges between 0.01 and 346 km^2, containing several dozen nodes in each stream network that was sampled. These catchments are located in the Swedish boreal forests, which are mostly coniferous with very little agriculture or other overt human impact beyond forestry.

In creating the models for DOC from map information, important questions will be which types of maps and spatial information that are most useful (soil maps or land cover maps), and whether higher resolution map information (1:20 000 vs. 1:100 000) improves the predictions (since such data are significantly more expensive for the user given current pricing structures).The study will also consider whether predictive models made from headwaters work on higher order streams in the same landscape, and vice versa, to see if the more commonly available data from high order streams is of value in making predictive models of headwaters.

Site description

Study areas

The two catchment systems studied in this paper, the River Öre and Stream Krycklan which drains into the Vindeln River, are located in northern part of Sweden (Figure 1). The sampled subcatchments ranged in size from 0.01-346 km^2. Boreal, coniferous forest dominates the landscape, with very little agriculture or other major human impact beyond forestry. Some 10-30% of the catchment size is comprised of wetlands. As for all unregulated boreal rivers, the runoff is characterized by a high runoff during spring due to snow melt, and a low base flow condition during the rest of the year, interrupted by higher flows caused by rainfall events (Ivarsson and Jansson, 1994). Overland flow is rare due to the high infiltration capacity of the till soils. The mean precipitation ranges between 500 and 700 mm per year (Alexandersson et al., 1991), of which one-third falls as snow (Ottosson Löfvenius et al., 2003). Mean annual temperature for the region of River Öre ranges between -2 °C and -4 °C (Alexandersson et al., 1991). The highest postglacial coastline is situated between 230 and 260 meters above sea level (m a.s.l.) (Ivarsson and Karlsson, 1992).

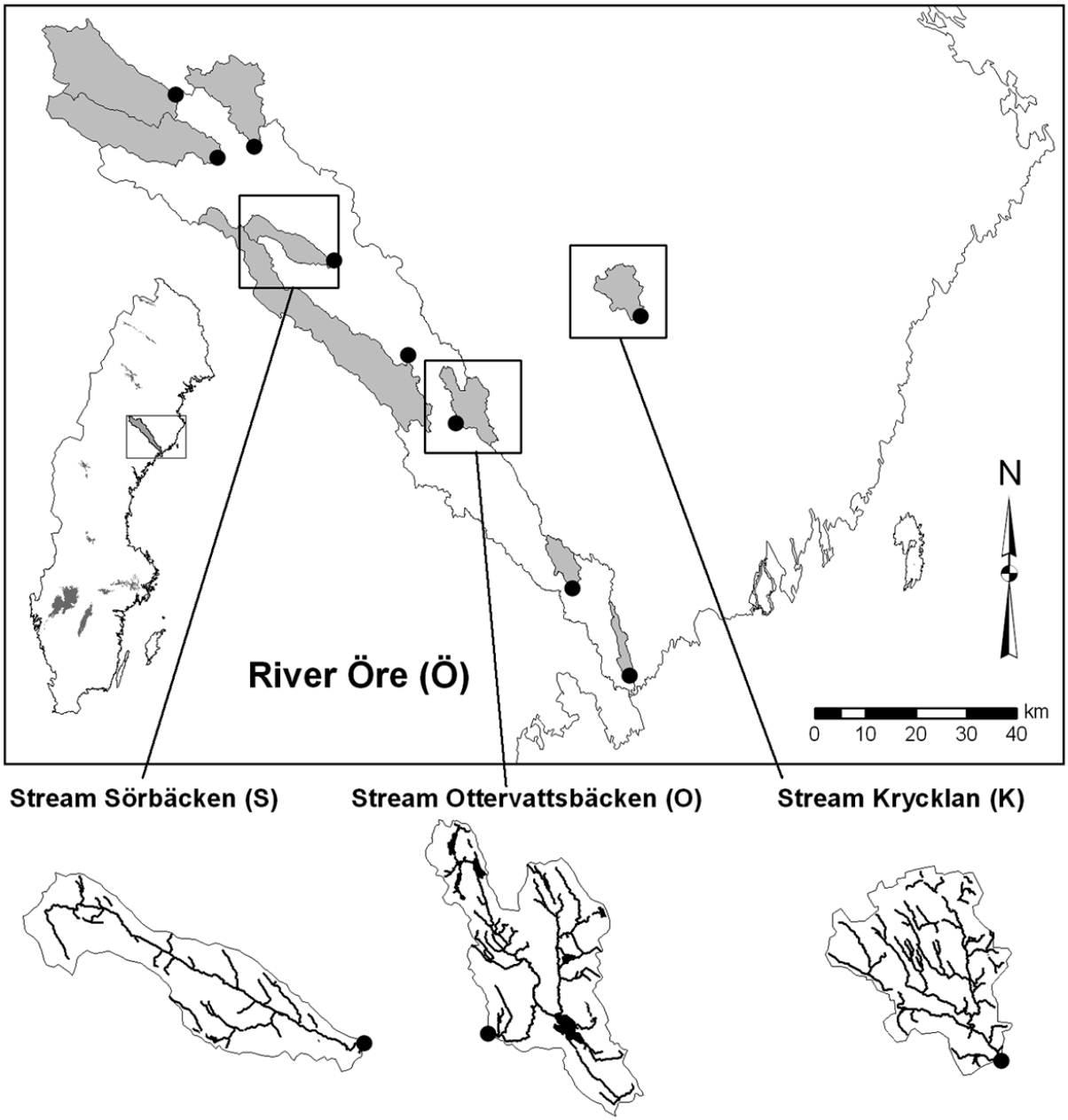

Figure 1 Map of Sweden with the River Öre catchment, including subcatchments, and the Krycklan catchment. Black dots are outlets. In S, O and K, solid black lines are surface water. These catchments were further divided into more subcatchments.

Catchments in the River Öre basin

Eight subcatchments are located in the River Öre basin (the main outlet at 63°42'N, 19°36'E). These subcatchments, hereafter denoted as Ö-91, were between 24 to 346 km^2 large (median 100 km^2;Table 1). The bedrock consists mainly of biotite-rich gneisses of sedimentary origin in the eastern part while the western part consists of granites. The tills are mostly silty-sandy to sandy-silty (Ivarsson and Karlsson, 1992). The median lake percentage for Ö-91 was 2.4%, with 5^{th} and 95^{th} percentiles of 0.15 and 5.7. Percentiles are hereafter given within brackets (5^{th} / 95^{th}). Wetlands has a median coverage of 22% (9.1 / 34). The percentage of peat soils within the catchments were similar to the values for the wetlands with a median of 25% (12 / 39). The drainage density for the Ö-91 catchments had a median of 0.81 (0.44 / 1.4).

Table 1. Median, with 5th and 95th percentiles in brackets, for some variables in the six datasets. m.d. is missing data. Map variables are from different maps, thus the sum of all presented data exceeds 100%.

Dataset	O-00	O-02	K-03	K-04	S-00	Ö-91
n	61	66	15	85	41	8
Area (km^2)	2.0 (0.20 / 28)	5.2 (0.24 / 62)	2.4 (0.12 / 32)	2.9 (0.25 / 19)	3.1 (0.45 / 58)	100 (28 / 320)
DOC (mg l^{-1})	18 (9.0 / 28)	18 (7.6 / 41)	16 (8.9 / 54)	18 (13 / 25)	12 (9.8 / 21)	10 (8.3 / 14)
Lake surface area %	0.5 (0 / 11)	2.2 (0 / 10)	0 (0 / 2.4)	0.2 (0 / 3.1)	0 (0 / 1.7)	2.4 (0.15 / 5.7)
Forest %	70 (54 / 95)	68 (52 / 84)	83 (55 / 92)	80 (60 / 93)	81 (72 / 96)	73 (60 / 87)
Arable land %	0 (0 / 1.0)	0 (0 / 1.6)	0 (0 / 2.4)	0 (0 /3.1)	0 (0 / 0.36)	m.d.
Wetland total %	19 (3.1 / 40)	19 (3.4 / 39)	15 (1.8 / 43)	14 (1.3 / 30)	19 (2.2 / 36)	22 (9.1/ 34)
Peat%	22 (2.7 / 46)	21 (3.3 / 46)	16 (0 / 55)	14 (0.05 / 30)	20 (2.2 / 36)	25 (12 / 39)
Till%	66 (46 / 94)	65 (46 / 85)	60 (35 / 89)	63 (46 / 85)	74 (57 / 93)	57 (34 / 74)
Drainage density	1.1 (0.40 / 2.9)	1.0 (0.56 / 2.8)	1.6 (0.26 / 4.7)	1.5 (0.93 / 3.6)	1.0 (0.59 / 2.8)	0.81 (0.44 / 1.4)

In two of the subcatchments in the River Öre basin, stream Ottervattsbäcken (O) and stream Sörbäcken (S), a larger number of subcatchments were sampled on two occasions. Stream Ottervattsbäcken (O) has its outlet at N64°02' and E19°06' and stream Sörbäcken has its outlet at N64°19' and E18°38'. The catchment size of O is 78 km^2, and the median size of the 61 subcatchments sampled in 2000 (O-00) was 2.0 km^2 (0.2 / 28), whilst in 2002 (O-02) the median size (n = 66) was 5.2 km^2 (0.24 / 62) (Table 1). The elevation of the catchment ranges from 196 to 370 (m a.s.l.). Approximately 30% of the catchment is situated above the highest post-glacial coastline, which in O is located at approximately 232 m a.s.l. The median lake percentage for O-00 was 0.5% (0 / 11) and for O-02, 2.2% (0 / 10) (Table 1). The median wetland coverage for both datasets was 18% (3 / 40). A similar pattern was found for the peat soil coverage, with a median of 22% (2.7 / 46) on O-00, and 21% (3.3 / 46) on O-02. Median drainage density for O-00 was 1.1 and for O-02 1.0. For 90% of the subcatchments, the values were between 0.4 and 2.9.

The catchment size of S is 63 km^2 with a median subcatchment size of 3.1 km^2 (0.45 to 58 km^2) (n = 41). The elevation of catchment Sörbäcken ranges from 255 to 570 m a.s.l., and the whole catchment is situated above the highest post-glacial coastline. The lake percentage for S was low compared with the other catchments, the median of all subcatchments was 0%, and the 95th percentile was 1.7% (Table 1). Median percentage of both wetland and peat soils was close to 20% (wetland 19% and peat 20%), and the 5th and 95th percentiles were identical (2.2% / 36%).

The Catchment Krycklan

Stream Krycklan (K) is situated in the basin of the River Vindelälven (64°14´N, 10°46´N). The Krycklan catchment includes the Vindeln Experimental Forests, where climate and runoff have been monitored at the Svartberget Research Station (64º14´N, 19º46´E) since 1980. The climate of catchment Krycklan is similar to that of O and S. A glacial till covers the bedrock and the soils are mainly well developed iron podzols. Further downstream, the glacial till gives way to glacio-fluvial sediments consisting mainly of sand. Some strips of agricultural land are also found further down in the valley. The catchment size of K is 61 km^2. As was the case for O, different sampling occasions have been used in the study. For K-03 (sampled in 2003) fewer

subcatchments were sampled compared with K-04 (sampled in 2004). The median size of K-03 (n = 15) was 2.4 km^2 (0.12 / 32 km^2) and for K-04 (n = 85) 2.9 km^2 (0.25 / 19 km^2) (Table 1). For the lake percentage both the median and 5th percentile was close to zero. The 95th percentile however was 2.4% (K-03) and 3.1% for K-04. Both the median percentages of wetlands and peat soils were around 15% for both datasets. A fairly high drainage density was found for K with a value of 1.6 (K-03) and 1.5 for K-04.

Material and methods

Sampling

In the River Öre basin, the outlets of eight catchments, Ö-91, were sampled in August 1991 (Ivarsson and Jansson, 1994). The additional sampling of S and O took place during low flow conditions 14th-20th June 2000 (S and O-00). During 19th-22nd August 2002, the O catchment (O-02) was sampled again (Temnerud and Bishop, 2005). For S and O, all samples were collected approximately 10 meters upstream, and in 2002, also downstream from every stream junction. For O, a total of 61 stream junctions were sampled in 2000. In 2002, 45 sites were resampled (the other sites were dry) plus 21 additional sites (mostly downstream from the headwater sampling sites of 2000) making for a total of 66 sites. For S a total of 41 sites were sampled in 2000, and no samples were taken in 2002. Data for Krycklan are from 2 sampling occasions, one sampling of 15 sites during low flow on August 6th 2003 (K-03) and one sampling of 85 sites on the falling limb of the spring flood on April 22nd 2004 (K-04) (Buffam unpublished data). The K-03 sites were sampled above every stream junction, and the 85 K-04 sites were from above and below stream junctions. The samples where collected either 5-10 m above or 25-50 m below every junction, depending upon the size of the streams. In total, six "sets" of stream chemistry data were compiled for use in the modelling (Table 1). All samples were taken under low flow conditions, with the exception of K-04.

DOC concentrations

The median value for the whole dataset was 16 mg l^{-1} (8.3 / 33). The median DOC-concentration for Ö-91 data was 10 mg l^{-1} (8.3 / 14) (Table 1). For analytical methods see Ivarsson and Jansson (1994). The median concentrations for S, O-00, O-02, K-03 and K-04 were between 16 and 18 mg l^{-1}. The 95th percentiles for O-02 (41 mg l^{-1}) and K-03 (54 mg l^{-1}) were higher than for O-00 (28 mg l^{-1}) and K-04 (25 mg l^{-1}). DOC concentrations from both K, O and S were analysed according to Temnerud and Bishop (2005).

Map materials

The goal of this study is to see if DOC concentrations in watercourses can be predicted from map data. The map information (digital maps) that has been used is the soil map, land-use maps and the kNN database (forestry estimates). The maps used have different scales, and hence the map variables differ in resolution.

The land-use information was available in to different scales, the blue map in scale 1:100 000 (X_{blue}) and the yellow map in scale 1:50 000 (X_{yellow}). The **blue map** (road map) is produced by Lantmäteriet (SNLS, 2002), and is mainly developed for displaying the network of roads in Sweden. However some land use classification is available, which are; built up area, water surfaces, open land, forest and wetland land divided up into three classes. The **yellow map** (economic map) has also been produced by Lantmäteriet 2002 (SNLS, 2002) This map shows land use and other variables such as roads and buildings. Both these maps are produced from

aerial photos and field observations. For X_{blue} a total of 10 land-use variables were available and for X_{yellow} a total of 11 variables, Table 2. The yellow map was only available for O and K, however the blue map was available for all catchments. The main reason for using maps of different scale was to study the difference in predictions by using map information in different scales.

The **soil map** (X_{soil}) is produced by the Geological survey of Sweden (SGU, 2001) and it gives an overview of soil types within or close to the soil surface; the soil type is classified from its formation and particle size distribution. The map has a scale of 1:50 000 and an aerial photo has been used to produce the soil map together with some field controls. The reliability of the map is therefore better around the road network. In areas with dense forests the reliability of the map is affected due to the high canopy closure of the forest. The map used in this study is a digital map containing three different layers; one containing soil type surfaces, one with the occurrence of boulders and one with amount of wave washed deposits. The soil type surface has been simplified to make it easier to compare areas with differing soil classifications. The X_{soil} provides an additional 16 variables, all described in Table 2. The wave washed layer is divided into four different classes; lightly, moderately, extensively wave washed and unsorted deposits. The variable "organic soil type" is referred to as peat in this study. The percentage of each variable for each subcatchment from the X_{blue}, X_{yellow} and X_{soil} was used in the modelling.

This **kNN-database** is a geographic database containing forestry variables estimated from Landsat 2000 satellite photos (Granqvist Pahlén et al., 2004; Reese et al., 2003),. The database was developed by the Swedish University of Agricultural Sciences. The areas that have been estimated are the areas that according the blue map are covered by forest. The variables available are total standing volume broken down by different tree species, as well as the height and age of the stands. In this study average height, stand age and the volume (m^3 $hectare^{-1}$) for the different tree species; birch (*Betula spp.*), Norway spruce (*Picea abies*), Scots pine (*Pinus silvestris*) and Lodge pole pine (*Pinus contorta*) have been used. The kNN-data were available for S, O and K. The kNN-database provides another 6 variables to be used in the modelling, Table 2.

In addition to the data read directly from the map, six variables were derived and calculated from the map information (i.e. constructed variables), this because there was reason to believe that they might capture important aspects of the relationship between the landscape and surface water chemistry. (X_{cons}): stream length (km), drainage density (m^1), sinuosity, slope (%), percentage above the highest coastline (%) and elevation (m) (Table 2). Drainage density is the ratio of the total length of streams within a catchment to the total area of the catchment. Drainage density gives a measure of the average lateral flow path length through soil to the stream network. The average drainage density in Sweden is roughly 1 km km^{-2} (Dahlström, 2005). Sinuosity is determined by dividing the stream length with the shortest distance between two sampling sites. Sinuosity is a measure of a stream's crookedness e.g. if it meanders or is a "straight" line. A stream with a sinuosity of 1 is completely straight with no bends, and with a ratio over 1.4 the stream has numerous closely spaced bends i.e. very meandering. Sinuosity is related to slope. Natural streams with steep slopes have low sinuosity, and vice versa.

Table 2. Map variables used, their explanation and abbreviation. For the land-use maps, two maps of different scale were used, X_{yellow} scale 1:20 000 and X_{blue} 1:100 000.

Map variables	Abbreviation	Explanation	1:20 000	1:100 000
Land-use maps			X_{yellow}	X_{blue}
Arable field	Arable%	Arable field	x	x
Forest	Forest%	Coniferous and mixed forest	x	x
Forest clear cut	Clear%	Clear cut	x	
Open land	Open%	Other open land	x	x
Pasture	Pasture%	Pasture		x
Water	Lake%	Water surfaces e g lake surface	x	x
Wetland forest	WF%	Wetland normal to wetland forest	x	x
Wetland impassable	WI%	Wetlands that are hard to cross	x	x
Wetland open	WO%	Wetland open land to wetland normal	x	x
Wetland total	WT%	W forest + W impassable + W open	x	x
Wetland total + water	WTL%	Wetland total + water (e g lake surface)	x	x
Wetland 10 meter%	W10m%	Wetland within 10 m of the stream channel	x	
Soil map, 1:50 000, X_{soil}				
Organic soil type	Peat%	Wetland, peat bog and mud (gyttja)		
Clay	Clay%	Clay, glacial/post glacial clay,		
Silt	Silt%	Silt glacial/postglacial/unspecified		
Sand	Sand%	Sand, not esker		
Gravel	Gravel%	Gravel, not esker		
Stone-Boulder	Boulder%	Stone to boulders and shingle		
Esker alluvium	Alluv%	Esker alluvium in general (Esker -gravel/-sand)		
Boulder clay	BClay%	Boulder clay (till clay)		
Till	Till%	Till with clay content less than 15%		
Thin soil	ThinS%	Thin or incoherent soil surface		
Mountain	Rock%	Archaean rock		
Lightly sorted deposits %	SDL	Not sorted or lightly sorted - No extensive wave-washing		
Moderately sorted deposits	SDM	Moderately sorted - Thinner and less widespread sorted deposits		
Unsorted deposits%	SDU	Areas where no wave-washing has taken place		
Extensively sorted deposits	SDE	Extensive sorting of deposits		
kNN-database, X_{kNN}				
Age of stand	StandAge	The age of the forest stand (years)		
Average height kNN	kNN_H	Average height of the forest (m)		
Volume birch	Birch	The volume of birch within the subcatchment, (m^3/ha)		
Volume lodge pole pine	Lodge	The volume of Lodgepole pine within the subcatchment, (m^3/ha)		
Volume Norwegian spruce	Spruce	The volume of Norwegian spruce within the subcatchment, (m^3/ha)		
Volume Scots pine	Pine	The volume of Scots pine within the subcatchment, (m^3/ha)		
Constructed variables, X_{cons}				
% above the highest postglacial coastline	%HC	Percentage of catchment above the highest postglacial coastline		
Stream length	SL	Stream length for the sub catchment (km)		
Lake length	LL	Distance between inlet and outlet (km)		
Total stream length	T_SL	The total stream length (km)		
Drainage density	Dens	Stream length (T_SL) divided by area (1/m)		
Sinuosity	Sin	Stream length divided by the shortest distance between sites		
Slope	Slope%	Slope across the actual stream length (%)		
Slope bird distance	SlopeB%	Slope across the shortest distance between sites (%)		
Average height	AvH	Average elevation for the subcatchment (m a.s.l.)		
Highest height	HiH	Highest elevation for the subcatchment (m a.s.l.)		
Lowest Height	LowH	Lowest elevation for the subcatchment (m a.s.l.)		

All the delineation of the catchments and sub catchments has been made by Jakob Nisell and Jan Seibert (both at the department of environmental assessment, Swedish University of Agricultural Sciences). This delineation has been done through a raster with altitude, flow directions for each pixel and the coordinates for each of the sampling sites. The delineation has taken into account the highest pixels and the flow direction of the soil water of each pixel from the topography of the area. The result of this process is a shape file for each sub catchment. These shape files have been further used by Jakob Nisell in ArcGIS8.0 (ESRI) to calculate the variables for the maps used. Most of the result from the calculations made by Jakob Nisell is in area (km^2), which has been recalculated into percentage coverage for each subcatchment (by dividing with the area in question). (Seibert, pers. comm. 2005)

Data analysis and modelling DOC

The aim in this study is to see if various landscape variables can be used to predict the chemistry in streams. One of the problems in doing this, is that many of the variables are intercorrelated. The other problem is the autocorrelation of water sampling when they are taken from nested subcatchments, i.e., if a sample is taken upstream from another sampling site you are in fact taking samples of the same water. Because there were many correlations between landscape variables, different multivariate methods were used to analyse the relation between landscape variables and DOC in this study: principal components analysis (PCA), partial least square regression (PLS) and multiple linear regression (MLR). PCA transforms many predictive variables, which can be intercorrelated, into a few components, each of which is a combination of many variables. Usually the 2-5 most important components are considered. The importance of individual variables is visualized more clearly in the correlation loading plot compared to the standard PCA loading plot. Correlation loadings are computed for each variable for the displayed principal components (PC). To complement the PCA, Spearman rank correlation coefficients, r_s, between DOC and the landscape variables and their significance (*** $p<0.001$, ** $p<0.01$ and * $p<0.05$) were computed (JMP v 4.0, SAS Institute Inc,).

While PCA primarily is used to detect structures in the data, PLS is appropriate for making predictions. PLS has the advantage of being able to work with covarying data in both a large predictor dataset (X-matrix) as well as in the response (Y-matrix). PLS is also well suited for detecting outliers (Geladi and Kowalski, 1986). MLR can be interpreted as a special case of PLS. MLR provides models that might be less susceptible to over fitting and, thus, easier to use on other catchments. Because MLR works best with few independent, uncorrelated, variables, the PCA/PLS results, together with Spearman ranking, were used to guide the selection of a few, relatively uncorrelated predictor variables.

The problem of autocorrelated water chemistry variables along the stream network was dealt with when making predictive models of downstream sites from headwaters by selecting sampling sites that are independent of each other for use in model building. Stream order 1, headwaters (Strahler, 1957), are those streams that were on 1:50 000 scale land-use maps. Sampling sites of SO1 were assumed to be independent of each other. For the downstream sampling sites with larger catchment size, there were many nested subcatchments and, thus, only few sampling sites could be classified as independent. Selection of independent stream order 2 sites (ISO2) was done by selecting sites so that no site had another SO2 sampling site upstream. There were too few SO3 sites to allow for selection of an independent SO3 data set. The SO3 values from O-02 or S were used to calibrate MLR and PLS models for O-02 or S. The models were either validated on SO1 or ISO2 within the same catchment. The procedure was also reversed, using SO1 to calibrate models that were validated against SO3 and ISO2. Furthermore ISO2 were used for calibrating models that were validated against SO1 and SO3.

The predictive modelling of DOC from map variables was done using both PLS and MLR (Unscrambler v 9.1.2, CAMO, Norway). For PCA and PLS (but not for MLR), all data were centred by mean normalization and weighted by dividing the variables with the standard deviation. To prevent negative predictions of DOC and to minimize the effect of extreme outliers in regression analyses, all the variables were transformed (for PCA, PLS and MLR). Different transformations were tested, and the square-root transformation was found to be most suitable. No outliers were removed from the datasets. All map variables means $X_{all} = X_{blue}+X_{soil}+X_{kNN}+X_{cons}$.

Criteria for a good model

The criteria for testing the models were the differences between measured DOC-concentrations and modelled DOC. The root mean square error (RMSE) is a measure of the average difference between measured and predicted values. The maximum RMSE for an acceptable model was set to the value of the 5th percentile for the whole data matrix, which was 8.3 mg l^{-1}. Besides the RMSE, we also required an acceptable model to reproduce the observed variability. This was evaluated by the slope of the regression line between measured and modelled values. Values between 0.7 and 1.3 were considered to be acceptable. The Pearson correlation coefficient, r, between measured and modelled values is another measure of model performance. Values above 0.6 were considered to indicate a good model. The Pearson correlation for the calibration is denoted as R^2, and for the validation Q^2. For the models derived using PLS and the calibration datasets, the variance in the data matrix explained by the 1st and 2nd principal component (PC) can be used as measure of how good the model is. A value above 60% was set as limit for an acceptable model.

Result

PCA

In the PCA for O-02 done upon all of the map variables (X_{all}) and DOC, the first and second principal component explained 40% of the variance in the data matrix (Figure 2). The variables between the ellipses explain between 50% and 100% of the 40%-variation in the data matrix. PCA loadings for the variables explaining more than 50% of the PCA variation results, and the Spearman rank correlation coefficients (in brackets) showed similar patterns. The variables positively correlated with DOC in both PC1 and PC2 were peat% (r_s 0.63***), WT% (r_s 0.60***), WO% (r_s 0.54***), WTL% (r_s 0.41***) and LowH (r_s 0.34***). The negatively correlated variables with DOC (both PC1 and PC2) were silt (r_s -0.42***) and forest (r_s -0.30*). The variables that had negative correlations with DOC in PC1 and positive in PC2 were; ThinS% (r_s -0.75***), rock% (r_s -0.76***), lake% (r_s -0.54***), HC% and HiH. Most of the kNN variables explained more than 50% of the PCA variation (all but Lodge pole pine), however Spearman rank correlation did not support these results. No variables were positively correlated with DOC in the first PC and negatively in the second PC. Peat% and wetland total% are found close to each other, indicating a strong relationship between these, which is supported by Spearman ranking (r_s 0.96***).

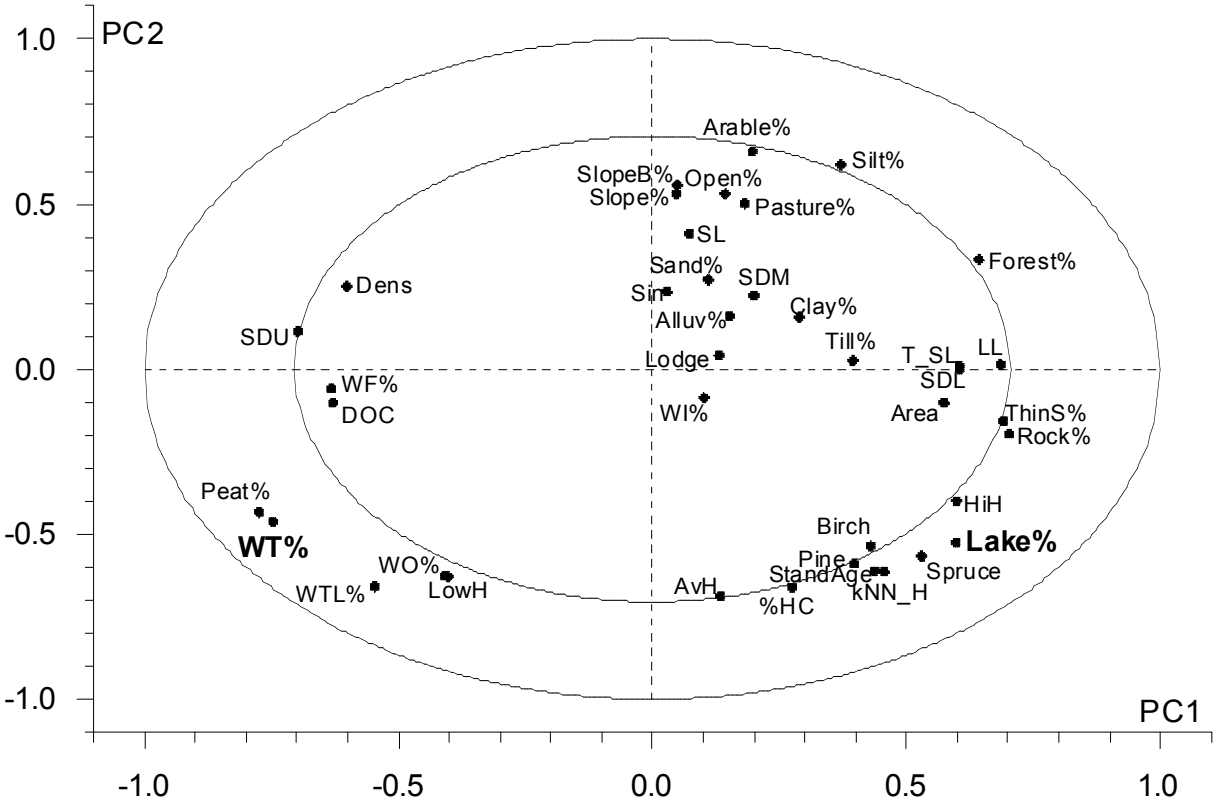

Figure 2. Correlation loading of PCA for O-02 with variables X_{blue}, X_{soil}, X_{kNN} and X_{cons}. See Table 2 for abbreviations. The variables explain 40% of the variation. 4 variables were constant and not included in the PCA (gravel%, boulder%, boulder clay and extensively sorted deposits). The outer ellipse is the unit-circle and indicates 100% of the variance explained by the components displayed. The inner ellipse indicates 50% of explained variance. Bold marked variables were used in MLR.

PLS

No single map had the information that could provide an acceptable model within the acceptable limits, as exemplified with the O-02 data (Table 3). However, the validation of PLS models on O-00 with either the blue map or the soil map variables gave results that were close to being acceptable. Slopes for both models were 0.98, with a Q^2 of 0.54 and an RMSE of 6.7 mg/l (blue map) or 6.6 mg/l (soil map), i.e. within or close to the acceptable limits (Table 3).

Table 3. PLS results from using different maps (X_{yellow}, X_{blue}, X_{soil}, X_{kNN}, X_{cons}). See Table 2 for abbreviations. O-02 was used as calibration set (n = 66) and validated against O-00 (n = 61). n var is the number of variables in each map.

Map variables	n var	slope	Q^2 or R^2	RMSE	PC1	PC2	SUM PC
Calibration; O-02							
X_{yellow}	11	0.31	0.33	11	35	27	62
X_{blue}	10	0.49	0.55	9.2	40	22	62
X_{soil}	11	0.45	0.46	10	58	4	62
X_{kNN}	6	0.09	0.10	13	5	7	12
X_{cons}	11	0.22	0.25	12	25	6	31
Validation; O-00							
X_{yellow}	11	0.59	0.30	6.9			
X_{blue}	10	0.98	0.54	6.7			
X_{soil}	11	0.98	0.54	6.6			
X_{kNN}	6	0.21	0.10	6.8			
X_{cons}	11	0.47	0.31	8.7			

Somewhat surprisingly, the models based on the 1:100 000 scale blue map land-use variables gave a better validation on O-00 than those based on the land-use variables from the 1:20 000 scale yellow map. The slope for X_{yellow} was 0.59 versus 0.98 for the blue map.

Table 4. PLS explained variance (%) for calibrations of the datasets using X_{all}.

Catchment	n	PC1	PC2	Sum
O-00	61	58	10	68
O-02	66	52	20	72
K-03	15	72	18	90
K-04	85	46	15	61
S-00	41	36	11	47
Ö-91	8	41	40	81

PLS based on X_{kNN} or X_{cons} did not predict DOC within the set limits. Using all the map variables together (X_{all}) gave PLS models where the calibrations explained between 68% and 90% of the variation in DOC (Table 4). In the PLS calibration for the different datasets in Table 5, four of the datasets had an explanation of variance in the first and second PC above the set limit value for the variation of DOC, with all but S-00 (41%) above 61% (Table 5). Efforts were made to find PLS calibrations based on datasets from one catchment that could be validated successfully against datasets from another catchment, however, all but one PLS did not meet the set limits. The exception was the calibration of O-00 validated against K-04, with both slope (0.78) and RMSE (6.6) within set limits but with a Q^2 of 0.50, outside but still close to the set limit.

A few more cases of semi-successful validations were found when the validation and calibration data came from the same catchment, but different years. The validation of O-00 from the O-02 calibration worked reasonably well, slope (0.89) and RMSE (7.8) within set limits but with Q^2 (0.42) outside of the limit. The model calibrated on K-03 validated on K-04 worked acceptably looking at both slope (1.2) and RMSE (7.0) but with Q^2 (0.37) outside the limit value. To summarize, we had one model out of 17 that almost worked across two different catchments, and two models of four that were almost acceptable for validation within the same catchment but for a different year. In each of the three semi-acceptable models, the results were much worse when the validation and calibrations data were reversed e.g. when calibrating a model on K-04 and validating on O-00.

Table 5. PLS regressions results using X_{all} variables. Calibration was performed on O-02, K-03, K-04 and S-00 and validation was made on the remaining datasets. For K-03 and K-04 calibrations some variables (SL, LL, Sin and slope) from X_{cons} were not available. The same situation was found for S-00 but the variables not available were HiH, LowH (X_{cons}) and Pine (X_{kNN}). Regarding Ö-91 either X_{kNN} or all the X_{cons} variables were available so therefore Ö-91 has been kept out.

		n	slope	Q^2 or R^2	RMSE	PC1	PC2	SUM PC
O-02	cal	66	0.42	0.42	10	52	19	71
O-00	val	61	0.89	0.47	7.8			
K-03	val	15	0.58	0.60	9.6			
K-04	val	85	1.6	0.60	7.7			
S-00	val	41	0.47	0.12	8.9			
K-03	cal	15	0.66	0.79	6.9	75	14	89
O-00	val	61	0.17	0.0079	17			
O-02	val	66	0.12	0.012	20			
K-04	val	85	1.2	0.37	7.0			
S-00	val	41	-2.4	0.05	111			
K-04	cal	85	0.47	0.48	2.9	47	14	61
O-00	val	61	0.064	0.0042	19			
O-02	val	66	-0.024	0.0011	24			
K-03	val	15	0.16	0.56	12			
S-00	val	41	0.10	0.086	13			
O-00	cal	61	0.55	0.51	4.7	58	11	69
O-02	val	66	0.22	0.34	12			
K-03	val	15	0.29	0.63	12			
K-04	val	85	0.78	0.50	6.6			
S-00	val	41	0.19	0.041	8.9			
S-00	cal	41	0.28	0.32	4.2	33	10	43
O-00	val	61	0.13	0.055	15			
O-02	val	66	0.082	0.077	30			
K-03	val	15	0.15	0.25	14			
K-04	val	85	0.27	0.12	5.1			

MLR

The variables that were used in the MLR were wetland total% (WT%) and lake surface area% (lake%). These variables were selected on the basis of the PCA plot (marked bold in Figure 2). Both PC1 and PC2 had positive loading values for WT%, and lake% was chosen as a variable from PC2. Spearman ranking between WT%, and lake% was not significant, r_s -0.22 at p = 0.080. The O-02 data was used for calibration. Two validations were close to the set limits. The validation on O-00 had a slope of 0.90, and an RMSE 6.8 within limits, but the Q^2 (0.48) was outside the limits (Table 6). The K-04 validation also had a slope (0.71) and RMSE (5.6) within the set limits, but Q^2 was even lower (0.24). It is also worth noting that the validations of S-00 and Ö-91 had a slope for S-00 of 0.06 and a Q^2 for Ö-91 of 0.02.

Table 6. MLR regressions with the variables WT% and lake% calibrated on O-02 and validated against the other datasets

	n	slope	Q^2 or R^2	RMSE
Calibration				
O-02	66	0.40	0.42	10
Validation				
O-00	61	0.90	0.48	6.8
K-03	15	0.49	0.71	8.8
K-04	85	0.71	0.24	5.6
S-00	41	0.058	0.003	12
Ö-91	8	0.19	0.015	9.9

PLS and MLR models of downstream and upstream data

To investigate the ability of downstream sites to predict upstream sites, and vice versa, a set of PLS was performed using X_{all} and MLR with the variables WT% and lake%. In these model tests, SO1, independent SO2 or SO3 data from specific datasets (e.g. O-02 or S-00) were used in calibrations and validations (Table 1). None of the MLR models gave acceptable models within the set limits (Table 7). But some of the PLS models based on SO3 or independent SO2 could be calibrated, and validated acceptably well against each other within the same catchment. For instance, taking O-02 calibration on SO3 and then validating against independent SO2 (O-02) had a slope of 0.75, a Q^2 of 0.42 and an RMSE of 9.5, just outside of the limit range. When calibrating on independent SO2, only validation against SO3 for O-02 gave close to acceptable results. None of the models calibrated on SO3 or independent SO2 data, however, could predict SO1 acceptably. And the calibrations on SO1 datasets neither calibrated well nor validated well against the downstream data.

Table 7. PLS (X_{all}) and MLR (WT% and lake%) using stream orders 3 (SO3), independent stream order 2 (ISO2) and stream order 1 (SO1) data for calibration, c, (validation, v, on the other datasets). See Methods for more details.

MLR		n	slope	Q^2or R^2	RMSE	**PLS**		n	slope	Q^2or R^2	RMSE	PC1	PC2	PC sum
O-02						**O-02**								
SO3	c	14	0.83	0.87	3.5	SO3	c	14	0.87	0.91	3.0	93	3	96
ISO2	v	10	0.82	0.32	13	ISO2	v	10	0.75	0.42	9.5			
SO1	v	21	0.35	0.12	18	SO1	v	21	0.19	0.05	19			
SO1	c	21	0.20	0.22	14	SO1	c	21	0.34	0.33	13	41	20	61
ISO2	v	10	0.43	0.36	8.8	ISO2	v	10	0.38	0.21	10			
SO3	v	14	0.47	0.86	5.5	SO3	v	14	0.25	0.34	8.2			
ISO2	c	10	0.38	0.37	8.3	ISO2	c	10	0.58	0.75	5.6	68	25	93
SO1	v	21	0.17	0.20	15	SO1	v	21	0.13	0.19	17			
SO3	v	14	0.40	0.86	6.3	SO3	v	14	0.61	0.82	8.3			
S-00						**S-00**								
SO3	c	13	0.46	0.47	0.67	SO3	c	13	0.47	0.47	0.7	48	10	58
ISO2	v	5	0.74	0.37	2.5	ISO2	v	5	0.50	0.52	1.5			
SO1	v	20	0.24	0.29	7.2	SO1	v	20	0.12	0.21	6.9			
SO1	c	20	0.68	0.69	3.9	SO1	c	20	0.39	0.44	5.2	41	24	65
ISO2	v	5	-0.18	0.00	9.7	ISO2	v	5	0.89	0.23	5.3			
SO3	v	13	0.88	0.05	5.2	SO3	v	13	-0.51	0.10	8.9			
ISO2	c	5	0.65	0.69	1.2	ISO2	c	5	0.71	0.75	1.1	75	23	98
SO1	v	20	0.08	0.06	7.0	SO1	v	20	0.08	0.11	7.2			
SO3	v	13	0.20	0.06	1.7	SO3	v	13	-0.37	0.13	3.6			

Discussion

Map variables were correlated to some of the variability of DOC in streams, Table 4. Percentage peat for example had a positive correlation, whilst lake surface area had a negative correlation with DOC, which has been indicated by previous research on lakes (e.g. Kortelainen and Saukonen, 1998). And by combining soil and land-use variables together with vegetation and constructed variables more of the variation in the data was explained. One unexpected result was the fact that the 1:100 000 blue land-use map gave slightly better predictions than the 1:20 000 yellow map, the blue map models with a slope closer to 1 and a lower RMSE than for the yellow map, albeit with a similar correlation. A possible explanation of this fact is that the map information is based upon the same or similar underlying data. Another explanation could be the reasoning of Herhily et al. (1998) and Rapp et al. (1985) that a homogeneous landscape is more difficult to use in prediction models.

Surprisingly also was the fact that models based on O worked poorly by validating them on S. It was assumed that S should have been better predicted than K-04, due to the fact that O and S both were sampled during low water flow and K-04 at the falling limb of the spring flood. Comparing the structure of the areas could be an explanation in the outcome. Both K and O has more lakes than S and for K and O the sinuosity and drainage density also differed compared with S, which might explain the outcome of the models. Another possible explanation could be that S is situated completely above the highest postglacial coastline, whilst both O and K are situated more or less on the highest postglacial coastline. And above this line the Quaternary deposits are thicker and coarser grained than for the deposits below the former coastline (Ivarsson and Karlsson, 1992). This probably affects the DOC on its way to the stream through these deposits, and therefore making O and K more compatible.

The efforts in making predictive models of DOC from landscape information did not work well. For the attempts in finding a model that worked for the same catchment during low flow but for

another year almost half of 5 trials worked rather well. However only one of over 20 attempts to transfer a model from one catchment into a nearby catchment gave close to an acceptable prediction, which could be explained by the homogenous landscape in the catchments studied. Maybe the wishful thinking of a more heterogeneous landscape could lead to better predictions (Herlihy et al., 1998; Rapp et al., 1985). And finally the trials of making PLS and MLR models built upon either stream order 3 or independent stream order 2 data were not able to predict the DOC concentrations in the headwaters or vice versa for that part too. This indicates that although downstream sites are influenced and linked to upstream/headwater streams, the relationship between map information and DOC varies with catchment scale.

Conclusion

The difficulty in predicting DOC concentrations during low water flow from information derived from maps is a problem for the development of tools for assessing the environment of streams and their headwaters mainly. It is possible that in using some kind of annual mean DOC measurements during one or several years could give rise to better predictions. Another approach could be to add specific discharge as a variable in the models, this to have a more hydrological variable, which might lead to better predictions. It also might be so that high water flow conditions would be easier to model, somewhat indicated by that stream Krycklan almost was accepted as a fairly good validation. For the monitoring of environmental assessment of streams, consideration needs to be given to how the temporal and spatial variation in chemistry should be addressed.

Acknowledgements

First of all I want to thank my most excellent supervisors for their support whenever I needed it, especially to Johan for all the hours spent over the phone and to Kevin who always saw the bigger picture and solved problems along the way. I also thank Martin Erlandsson for supporting me in the world of multivariate statistics. I would also like to thank Jakob Nisell for his excellent GIS work.

And a deep gratefulness towards my parents, Bengt and Belinda, who has always given me precious support and encouragement during all of my years of studying. And to my lovely girlfriend for always being there, whenever I need it!

References

Aitkenhead, J.A., Hope, D. and Billett, M.F., 1999. The relationship between dissolved organic carbon in stream water and soil organic carbon pools at different spatial scales. Hydrological Processes, 13(8): 1289-1302.

Alexandersson, H., Karlström, C. and Larsson-McCann, S., 1991. Temperaturen och nederbörden i Sverige. Referensnormaler. (Temperature and precipitation in Sweden 1961-90, Reference normals, in Swedish with English summary). Report Meteorologi nr 81, Swedish Meteorological and Hydrological Institute, Norrköping.

Allan, J.D. 1995. Stream ecology: Structure and function of running waters. Kluwer Academic Publishers.Dordrecht-Netherlands, 388pp.

Bengtsson, G. and Törneman, N., 2004. Dissolved organic carbon dynamics in the peat-streamwater interface. Biogeochemistry, 70(1): 93-116.

Bishop, K. and Pettersson, C. 1996. Organic carbon in the boreal spring flood form adjacent subcatchments. Environ. Int. 22: 535-540

Creed, I.F., Sanford, S.E., Beall, F.D., Molot, L.A. and Dillon, P.J., 2003. Cryptic wetlands: integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes. Hydrological Processes, 17(18): 3629-3648.

Dahlström, N., 2005. Function and dynamics of woody debris in boreal forest streams. Doctoral Thesis, Umeå University, Umeå, 21 pp.

Doppelt, B., Scurlock, M., Frissell, C. and Karr, J., 1993. Entering the Watershed: A New Approach to Save America's River Ecosystems. Island Press, Washington D.C., 462 pp.

Dillon, P.J. and Molot, L.A., 1997. Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. Water Resources Research, 33(11): 2591-2600.

Eatherall, A., Naden, P.S. and Cooper, D.M., 1998. Simulating carbon flux to the estuary: The first step. Science of the Total Environment, 210(1-6): 519-533.

Eckhardt, B.W. and Moore, T.R., 1990. Controls on dissolved organic-carbon concentrations in streams, Southern Quebec. Canadian Journal of Fisheries and Aquatic Sciences, 47(8): 1537-1544.

Elder, J.F., Rybicki, N. B., Carter, V., Weintraub, V., 2000. Sources and yields of dissolved carbon in northern Wisconsin stream catchments with differing amounts of peatland. Wetlands, 20(1): 113-125.

Geladi, P. and Kowalski, B.R., 1986. Partial least-squares regression - a tutorial. Analytica Chimica Acta, 185: 1-17.

Gergel, S.E., Turner, M.G. and Kratz, T.K., 1999. Dissolved organic carbon as an indicator of the scale of watershed influence on lakes and rivers. Ecological Applications, 9(4): 1377-1390.

Granqvist Pahlén, T., Nilsson, M., Egberth, M., Hagner, O. and Olsson, H., 2004. kNN-Sverige: Aktuella kartdata över skogsmarken (kNN-Sweden: Actual mapdata for the forests, in Swedish). Fakta Skog, 12: 1-4.

Herlihy, A.T., Stoddard, J.L. and Johnson, C.B., 1998. The relationship between stream chemistry and watershed land cover data in the mid-Atlantic region, US. Water, Air and Soil Pollution, 105(1-2): 377-386.

Hope, D., Billett, M.F. and Cresser, M.S., 1994. A review of the export of carbon in river water - fluxes and processes. Environmental Pollution, 84(3): 301-324.

Ivarsson, H. and Jansson, M., 1994. Regional variation of dissolved organic matter in running waters in central northern Sweden. Hydrobiologia, 286(1): 37-51.

Ivarsson, H. and Karlsson, L.-I., 1992. Geological and geochemical conditions in the River Öre drainage basin, northern Sweden. Gerum report nr.15, Geography Department, Umeå University, Umeå.

Kortelainen, P., 1993. Content of total organic carbon in Finnish lakes and its relationship to catchment relationship. Canadian Journal of Fisheries and Aquatic Sciences, 50: 1477-1483.
Kortelainen, P. and Saukkonen, S., 1998. Leaching of nutrients, organic carbon and iron from Finnish forestry land. Water, Air and Soil Pollution, 105(1-2): 239-250.
Köhler, S., Hruška, J. and Bishop, K., 1999. Influence of organic acid site density on pH modelling of Swedish lakes. Canadian Journal of Fisheries and Aquatic Sciences, 56(8): 1461-1470.
Laudon, H., Köhler, S. and Buffam, I., 2004. Seasonal TOC exports from seven boreal catchments in northern Sweden. Aquatic Sciences, 66(2): 223-230.
McDonald, S., G., B.A., D., P.P. and Robards, K., 2004. Analytical chemistry of freshwater humic substances. Analytica Chimica Acta, 527: 105–124.
Meili, M., 1992. Sources, concentrations and characteristics of organic matter in softwater lakes and streams of the Swedish forest region. Hydrobiologia, 229: 23-41.
Mulholland, P.J. and Kuenzler, E.J., 1979. Organic carbon exports from upland and forested wetland catchments. Limnology and Oceanography, 24(5): 960-967.
Ottosson Löfvenius, M., Kluge, M. and Lundmark, T., 2003. Snow and soil frost depth in two types of shelterwood and a clear-cut area. Scandinavian Journal of Forest Research, 18(1): 54-63.
Pers, C., Rahm, L., Jonsson, A., Bergström, A.-K. and Jansson, M., 2001. Modelling dissolved organic carbon turnover in humic Lake Örträsket, Sweden. Environmental Modelling and Assessment, 6: 159-172.
Rapp, G. et al., 1985. Acid deposition and watershed characteristics in relation to lake chemistry in northeastern Minnesota. Environment International, 11(5): 425-440.
Reese, H. et al., 2003. Countrywide estimates of forest variables using satellite data and field data from the national forest inventory. Ambio, 33(8): 542-548.
Schwarzenbach, R.P., Gschwend, P.M. and Imboden, D.M., 2003. Environmental Organic Chemistry. John Wiley, Hoboken, 1313 pp.
Seibert, J. Personal commentary. Department of Environmental assessment, Swedish university of agricultural sciences, Uppsala, Sweden.
SGU, 2001. Soil map, 1:50 000. The Geological Survey of Sweden, Uppsala.
Shreve, R.L., 1969. Stream lengths and basin area in topographically random channel networks. Journal of Geology, 77(4): 397-414.
SNLS, 2002. Land-use map, scale 1:20000 and scale 1:100000. The Swedish National Land Survey, Gävle.
Strahler, A.N., 1957. Quantitative analysis of watershed geomorfology. Transactions of the American Geophysical Union, 38(6): 913-920.
Stumm, W. and Morgan, J.J., 1996. Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters. John Wiley, New York, 1022 pp.
Temnerud, J. and Bishop, K., 2005. Spatial variation of streamwater chemistry in two swedish boreal catchments: Implications for environmental assessment. Environmental Science and Technology, 39(6): 1463-1469.
Thurman, E. M., 1985. Organic Geochemistry of Natural Waters. Martin Nijhoff/ Dr W Junk Publishers, Dordrecht, 497 pp.
Wetzel, R.G., 2001. Limnology. Lake and River Ecosystems. Academic Press, London, 1006 pp.