MODELING THE SPATIAL AND TEMPORAL TRENDS OF WATER QUALITY IN BOREAL MANAGED WATERSHEDS

Carlos A. Gonzales Inca
MODELING THE SPATIAL AND TEMPORAL TRENDS OF WATER QUALITY IN BOREAL MANAGED WATERSHEDS

Carlos A. Gonzales Inca
In memory of my grandfather Justino Inca Arqque, who always believed in the power of the knowledge to generate wellbeing.

A la memoria de mi abuelo Justino Inca Arqque, quien siempre creyó en la fuerza del conocimiento para generar bienestar.
Abstract

Land use changes have altered natural hydrological pathways and biogeochemical cycling of carbon, nitrogen and phosphorus, among other elements, affecting the quality of aquatic ecosystems such as rivers, lakes and coastal areas. In this dissertation, the spatial and temporal trends of water quality variation in Finnish managed watersheds was studied by applying methods of multivariate statistics, time-series analysis, ecohydrological modeling and high-resolution geospatial data. The results show the complex effects of current land use, particularly agriculture, on stream water quality. New emerging trends of nutrient concentrations and loads were detected in the time-series analysis, such as an increase in the concentrations and loads of dissolved reactive phosphorus and total nitrogen, and a decrease in suspended sediment concentration in streams. This might be linked to the current erosion reduction strategy of land management for water protection. An ecohydrological modeling assessment showed an increasing downstream nutrient export from agricultural watershed under climate change scenarios. The modeling results also showed a potential nutrient export reduction by restoring potential biogeochemical hotspot areas - wet areas or areas prone to water saturation. These areas can function as nutrient sinks and enhance the watershed resiliency. High-resolution geospatial data allowed easier and more accurate mapping of wet areas as well as the extracting of their hydraulic characteristics. However, the ecohydrological models involved several sources of uncertainties, which need to be carefully addressed with extensive observational data, expert knowledge of model parameter definitions, proper modeling unit selection and empirical knowledge of the functioning of the studied watershed system. The results of this dissertation highlight the importance of combined methods for watershed management research, and the proper identification of the biophysical processes in the modeling of non-point pollutant sources; this can in turn lead to an efficient water protection measure, and restoring biogeochemical hotspot areas within the watershed.

Keywords: Watershed, modeling, water quality, ecohydrological, biogeochemical, climate change, resiliency, wet areas
Tiivistelmä


Avainsanat: Valuma-alue, mallinnus, veden laatu, ekohydrologinen, ilmaston muutos, resilienssi, kosteikkoalueet
Resumen

El cambio del uso del suelo ha alterado los procesos hidrológicos naturales y los ciclos biogeoquímicos del carbono, el nitrógeno y el fósforo, entre otros elementos, afectando directamente la calidad de los ecosistemas acuáticos como los ríos, lagos y zonas costeras. En esta disertación, las tendencias espaciales y temporales de la variación de la calidad del agua en cuencas hidrográficas finlandesas se estudiaron mediante la aplicación de métodos de estadística multivariante, análisis de series de tiempo, modelos ecohidrológicos y datos geoespaciales de alta resolución. Los resultados muestran los efectos complejos del uso actual del suelo, particularmente la agricultura, en la calidad del agua de los ríos y corrientes. Se detectaron nuevas tendencias emergentes de concentraciones y cargas de nutrientes en el análisis de series temporales, como un aumento en la concentración y carga del fósforo disuelto reactivo y nitrógeno total, y una disminución en la concentración de sedimentos en suspensión en los ríos y corrientes. Esto podría estar vinculado a la estrategia actual de manejo del suelo, orientado a la reducción de la erosión para la protección del agua. Una evaluación a través de modelización ecohidrológica mostró un aumento de la exportación de nutrientes aguas abajo de la cuenca agrícola bajo escenarios de cambio climático. Los resultados de la modelización también mostraron una posible reducción de la exportación de nutrientes mediante la restauración de posibles zonas críticas biogeoquímicas: áreas húmedas o áreas propensas a la saturación de agua. Estas áreas pueden funcionar como sumideros de nutrientes y mejorar la resiliencia de la cuenca. Los datos geoespaciales de alta resolución permitieron un fácil y más preciso cartografiado de las áreas húmedas, así como la extracción de sus características hidráulicas. Sin embargo, los modelos ecohidrológicos involucraron varias fuentes de incertidumbre, que deben abordarse cuidadosamente con bastantes datos de observación, conocimiento experto de las definiciones de los parámetros del modelo, selección adecuada de la unidad de modelado y conocimiento empírico del funcionamiento del sistema de la cuenca estudiada. Los resultados de esta disertación destacan la importancia de los métodos combinados para la investigación de gestión de cuencas hidrográficas y la identificación adecuada de los procesos biofísicos en la modelización de fuentes contaminantes difusas; esto a su vez puede conducir a una medida eficiente de protección del agua, y restauración de áreas claves de alta función biogeoquímica dentro de la cuenca.

Palabras clave: Cuenca hidrográfica, modelización, calidad del agua, ecohidrológico, biogeoquímico, cambio climático, resiliencia, áreas húmedas
## Contents

Abstract .......................................................................................................................... 4

Tiivistelmä .................................................................................................................... 5

Resumen ........................................................................................................................ 6

List of Scientific Articles ........................................................................................... 8

Abbreviations .............................................................................................................. 9

1. Introduction ......................................................................................................... 10

2. Research aims ...................................................................................................... 12

3. Theoretical Background .................................................................................. 14
   3.1. Aquatic ecosystem and water resource management research............ 14
   3.2. Spatial and temporal effect of land use on surface water quality........ 15
   3.3. Analysis scale in assessment of the effect of land use on water quality ................................................................................................................................. 17
   3.4. Ecohydrological models............................................................................... 18

4. Materials and Methods ..................................................................................... 21
   4.1. Study area ........................................................................................................... 21
   4.2. Environmental data and information.............................................................. 22
   4.3. Data analysis and modeling............................................................................. 24
       4.3.1. Statistical analysis and time-series trend detection.................... 24
       4.3.2. Topographical-based mapping of wet areas............................. 24
       4.3.3. Ecohydrological modeling......................................................... 26

5. Results and Discussion ..................................................................................... 30
   5.1. Stream water quality indicates both land-use and riparian processes................................................................................................................................. 30
   5.2. Emerging trends in stream water quality variables................................. 32
   5.3. Topographic-based wet area mapping is highly influenced by the modeling techniques ................................................................................................................ 34
   5.4. High spatial and temporal uncertainty in ecohydrological modeling.. 36

6. Conclusions ........................................................................................................... 43

7. Acknowledgements ............................................................................................ 45

8. References ............................................................................................................. 47

Original Articles ....................................................................................................... 57
List of Scientific Articles

This dissertation is supported by four scientific articles listed below:


<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>CSA</td>
<td>Critical Source Area</td>
</tr>
<tr>
<td>DRP</td>
<td>Dissolved Reactive Phosphorus</td>
</tr>
<tr>
<td>DTM</td>
<td>Digital Terrain Model</td>
</tr>
<tr>
<td>DTW</td>
<td>Depth to Water Index</td>
</tr>
<tr>
<td>EPA</td>
<td>United State Environmental Protection Agency</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>GLM</td>
<td>Generalized Linear Model</td>
</tr>
<tr>
<td>GLUE</td>
<td>Generalized Linear Uncertainty Estimation</td>
</tr>
<tr>
<td>HELCOM</td>
<td>Helsinki Commission</td>
</tr>
<tr>
<td>HRU</td>
<td>Hydrological Response Unit</td>
</tr>
<tr>
<td>IWRM</td>
<td>Integrated Water Resources Management</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Light Detection and Ranging</td>
</tr>
<tr>
<td>MMF</td>
<td>Morgan-Morgan-Finney soil erosion prediction model</td>
</tr>
<tr>
<td>N</td>
<td>Nitrogen</td>
</tr>
<tr>
<td>NSE</td>
<td>Nash-Sutcliffe Efficiency</td>
</tr>
<tr>
<td>P</td>
<td>Phosphorus</td>
</tr>
<tr>
<td>ParaSol</td>
<td>Parameters optimization and uncertainty analysis tool</td>
</tr>
<tr>
<td>PBIAS</td>
<td>Percent Bias</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>RDA</td>
<td>Redundancy Analysis</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>RSR</td>
<td>the RMSE-observations' standard deviation ratio</td>
</tr>
<tr>
<td>SLICES</td>
<td>Separated Land Use/Land Cover Information System</td>
</tr>
<tr>
<td>SUFI</td>
<td>Sequential Uncertainty Fitting</td>
</tr>
<tr>
<td>SWAT</td>
<td>Soil and Water Assessment Tool</td>
</tr>
<tr>
<td>SYKE</td>
<td>Finnish Environmental Institute</td>
</tr>
<tr>
<td>TWI</td>
<td>Topographic Wetness Index</td>
</tr>
<tr>
<td>TSS</td>
<td>Total Suspended Solids</td>
</tr>
<tr>
<td>WQ</td>
<td>Water Quality</td>
</tr>
</tbody>
</table>
1. Introduction

Water is an essential resource for maintaining life, biodiversity, ecosystem functioning and social well-being. In spite of the great importance of water for life, the quality and quantity of water resources has been rapidly impaired by human land use in the last decades (Haygarth and Jarvis 2002). In Finland, the surface area of fresh waters cover 10% of the territory, being distributed to tens of thousands of lakes of varying size and streams (Eloranta 2004). However, in spite of the large inland water area, the water volume is relatively small, because most of the water bodies are shallow (Eloranta 2004). Finnish socio-economic activities are highly dependent on water resources and most of the Finnish water bodies are highly vulnerable to pollution, particularly in the coastal area in the west, where intense agricultural practices have been developed. In Finland, point pollution sources, such as urban sewage, has been substantially reduced (Meriläinen et al. 2003), but diffuse sources, agriculture and forestry in particular, account for large amounts of nutrients (33% of total nitrogen and 56% of total phosphorus) into surface water draining to Baltic Sea (HELCOM 2004). Agricultural areas are the main diffuse pollutant sources to aquatic ecosystems globally, causing problem of toxic algae bloom, eutrophication, loss of aquatic biodiversity, etc. (Carpenter et al. 1998). The effects of agricultural practices on aquatic ecosystems and water resources are a continuous process with a long-term impact on the environment. Scientists and policy-makers are currently challenged by the question of, how to secure water quantity and quality for the next generations and maintain a healthy ecosystem, while covering the demand for food production, industry, water supply and entertainment (Pimentel et al. 2004). These issues are also acknowledged in legislative frameworks, e.g., in EU’s water framework directive (European Commission 2000), which are aimed at an ecological integrity and water resource protection.

In Finland, diffuse water pollution from agriculture has been studied in field experiments, small catchment monitoring and computer simulation models (e.g. Rekolainen & Posch 1993, Rankinen et al. 2004, Bärlund et al. 2007, Tattari et al. 2009, Puustinen et al. 2010, Warsta 2011, Kirkkala 2014, Huttunen et al. 2016). Because of the high cost, low control and low manipulability of catchment scale experiments to assess different scenarios of catchment management, computer-based catchment models are in high demand. However, early catchment scale models were mainly developed for hydrological prediction. Many of them are currently also extended to model biogeochemical and ecological processes, thus called ecohydrological models. As these models cannot capture all natural
processes, many of them stochastic, they inevitably present some shortcomings. Nevertheless, ecohydrological models do have diverse applications, such as assessing the environmental impact of infrastructure development, ecosystem services, landscape management, climate change effects and watershed resiliency (Collins and McGonigle 2008).

Methodologically, alternative ecohydrological models differ in their conceptual basis, mathematical formulation, and landscape representation (lump, semi-distributed and full spatial distributed) and the simulation result is often very different, particularly in terms of their spatial pattern (Grayson and Blöschl, 2000). For a proper representation of all hydrological, biogeochemical and ecological processes in the model, a process-oriented unit of modeling is required. In other words, ecohydrological models need to capture different zones of the landscape (e.g. saturation-excess runoff zones and riparian areas), which can be distinguished from the surrounded areas in their hydrological, biogeochemical and ecological processes and functions (Tetzlaff. et al., 2009). Having the modeling unit delineation based on a study of the landscape’s hydrogeomorphic features, the semi-distributed ecohydrological model is assumed to achieve better prediction results. The mapping of such units can be easily made using high-quality digital terrain models (Schneiderman et al. 2007, Fuka et al. 2016).

Despite the usefulness of computer modeling to study catchment ecohydrological processes, it is important to carry out empirical studies, not only to identify factors affecting stream water quality, but also to determine on what scale a certain factor is operating and affecting the quality of stream water (Buck et al. 2004). Nowadays, the availability of spatial environmental data, the development of geographic information system (GIS) analysis and multivariate statistical methods have allowed the study of the linkages between stream water quality and landscape features in a more integrated manner (e.g. Griffith et al. 2002; Varanka and Luoto 2012, Li et al. 2015, Varanka et al. 2015). Moreover, studying long-term detection of hydrological and water quality trends is important, for instance, to determine stream water quality responses to changes in land use, land management, and climate (Hirsch et al. 2010). Additionally, the new automatic water quality monitoring systems with high temporal resolution (sub-hourly) data provide an advantage in gaining a better insight into nutrient and sediment mobilization dynamics on a catchment level (Kotamäki et al. 2009).
2. Research aims

The general aim of this dissertation is to gain a better understanding of the linkages between landscape management and aquatic ecosystems in order to foster water resources protection in boreal catchments. The dissertation address the following research questions:

- What is the role of landscape spatial pattern controlling stream water quality, and which parts of the watershed are the most important in their biochemical functioning, regulating the nutrient export from terrestrial to stream ecosystems?
- What has been the effect of the current land management strategy for water protection in Finnish catchments on surface water quality?
- How the hydrological and biogeochemical processes occurring within a watershed can be better represented with the current computer modeling techniques?
- Can the restoration of biogeochemical hotspot areas reduce nutrient loading from land to stream water in the current and future climate condition?

In order to answer these questions, the dissertation has the following specific objectives:

1) To evaluate the effect of the land use pattern on surface water quality in different hydrogeomorphic functional zones in agricultural watersheds.
2) To analyze long-term trends of agriculture-related water quality variables.
3) To assess a process-oriented ecohydrological model for evaluation of the effects of land use/management and climate change on water quality, by mapping and simulating wet areas restoration.

To reach these objectives several statistical methods including multivariate analysis, temporal time series analysis, ecohydrological models, and sediment fingerprinting were applied. In addition, GIS analyses with high-resolution environmental data were used.

The dissertation outcome is based on four scientific articles (listed on page 3).
Article I provides an empirical evaluation of the relationship between land use patterns and spatial water quality variation through a multivariate statistical analysis. The study area included 16 agricultural sub-watersheds. The landscape patterns were characterized in three scales of different hydrological importance: the entire sub-watershed area, saturation-excess zones, and riparian areas.

In Article II, a long-term trend analysis of water quality was carried out in two agricultural watersheds having different hydrological responses. Three statistical trend detection methods were applied by using data from a 19 to 34 year-long time series of nutrient concentrations, loads, and river discharges.

In Article III, an evaluation of sediment transfer during the snow-melt period was conducted. The snow-melt period was assumed to be a critical period of pollutant mobilization from agricultural areas into watercourses. Automatically registered, temporal high-resolution data (sub-hourly) of turbidity, sediment concentration, water discharge and weather conditions were used. In addition, the sediment source was fingerprinted by analyzing Cesium-137 radioactivity and total phosphorus content. In addition, a snowmelt runoff and a process oriented soil erosion model were applied to study the sediment erosion pattern.

In Article IV, a process oriented Soil and Water Assessment Tool (SWAT) ecohydrological model was applied to simulate the main biogeochemical processes and to estimate the nutrient loads in the watershed. Furthermore, the watershed resilience and climate change effects were evaluated. The article also presents a model-based evaluation of the potential reduction of nutrient loads by restoring biogeochemical hotspot areas. To this end, SWAT model was set up, calibrated, and validated, and with model sensitivity and uncertainty analysis included.
3. Theoretical Background

3.1. Aquatic ecosystem and water resource management research

For sustainable development, the maintenance of healthy ecological status and the integrity of the aquatic ecosystem services are highlighted by several studies and policy programs (Abell et al. 2002, Frederiksen et al. 2007, Colling and McGonigle 2008). In the early stages, aquatic systems (lentic and lotic) were studied by mainly focusing on internal bio-physical processes. For ecological processes in rivers, the benchmark concept was the river continuum concept (Vannote et al., 1980, Lorenz et al. 1997), which assumes a mass flux gradient from headwater to downstream, to which different organisms are adapted. The headwater is assumed to have low primary productivity and dependent on allochthones detritus, but the dependence on allochthone nutrients decreases lower downstream where the sources of organic matter and stream primary production increase (Lorenz et al. 1997). Because the river continuum concept does not fully explain the ecological and biogeochemical status of rivers, several other concepts have emerged, such as the stream hydraulics concept (Statzner and Higler 1986), which gives more emphasis to the role of stream geomorphologic and hydraulic properties influencing the ecological functioning of rivers (Statzner and Higler 1986, Petts 1994). Similarly, the river–floodplain interaction concept emphasizes the floodplain hydrological and biogeochemical processes influencing river ecology; the nutrient release function from the floodplain and riparian zone nutrient sink being particularly highlighted (Pinay et al. 1990, Brunet et al. 1994). Currently, most of the studies on river ecology and river water quality protection emphasize the catchment or landscape integrated concept (Moldan and Cerny 1994), where a river is viewed as a four-dimensional connected system (Ward 1989): 1) Laterally, the river is influenced by the riparian zones and upland continuum landscape composition; 2) longitudinally, it is influenced by biophysical processes from upstream to downstream; 3) vertically, the influence is from underground geological characteristics, particularly in the hyporheic zone, where the surface and groundwater mix (Grimm et al. 2007). All these connections vary in the 4) temporal dimension, as the hydrological and biogeochemical processes vary seasonally, annually and in the long-term (Ward, 1989).
As an operational framework for an integral conservation of ecological, hydrological, and biogeochemical functions of catchment, and the ecosystem services they provide, the concept of River Basin Management (RBM) and Integrated Water Resources Management (IWRM) is often used in scientific literature and policy decision making. Although these concepts vary in definition and scope, they also overlap in recognizing the terrestrial and aquatic ecosystem as an interdependent system, the river basin as the fundamental environmental management unit, and the inclusion of socioeconomic factors in the planning process. The planning of river basin management started in Europe in the 1960s when it was mostly oriented towards flood and debris protection. Currently, it has been extended to conservation planning and sustainable use of water resources (surface water, groundwater and coastal water). For instance, the European Water Framework Directive states as a target the maintenance of a good ecological status for all form of water resources (Frederiksen et al. 2007). Integrated water resources management, in turn, attempts to provide a more holistic framework for water and land use planning and to ensure ecosystem sustainability together with social and economic welfare (Varis et al. 2014). In contrast to the traditional sector-by-sector top-down management approach, the IWRM promotes the integration of water user sectors: nature, agriculture, human population, and industries (GWP and INBO, 2009), which can extend beyond the river basin. It encourages more participatory processes and proactive decision making in order to promote the efficient management of water resources (GPW and INBO, 2009).

3.2. Spatial and temporal effect of land use on surface water quality

Surface Water Quality (WQ) is affected by multiple environmental factors (e.g. weather, geology, soil, topography, biota, and land use and land management). From all the factors influencing water quality, land biota cover has changed more rapidly than other factors due to land use change, with agricultural practices being the major driver (Riebsame et al. 1994). As all landscapes are complex by their very nature, the combination and interaction of different environmental factors with agricultural land use leads to different effects on the catchment hydrological and biogeochemical processes affecting aquatic ecosystem; most of these processes are not yet fully understood (Galloway et al 2003).
Agricultural systems are sustained by the application of synthetic fertilizers and manure, and a modification of the landscape function by artificial drainage (Haygarth and Jarvis 2002). This has changed the natural nutrient recycling of the landscape. For example, a lowering of the water table has enhanced nutrient mineralization and caused the loss of large wetlands, where nutrient sinks naturally occurred (Arango and Tank 2008). The effects of an excessive loading of nitrogen (N) and phosphorus (P) into aquatic ecosystems, causing eutrophication and toxic algae bloom, has been widely documented in the literature (e.g. Carpenter et al. 1998, Kirkkala 2014). These elements in agricultural systems have relatively different hydrological pathways of transportation into rivers and lakes. N losses generally occur as a soluble reactive form of N, ammonium (NH$_4^+$N), nitrate (NO$_3^-$N) and organic nitrogen (Galloway et al 2003). NH$_4^+$N usually represent a small part of N losses and NO$_3$-N is transported by overland flows, subsurface flows, and groundwater flows (Lepistö 1996). N is denitrified back to the non-reactive N$_2$ in areas where high NO$_3$-N or NH$_4^+$N concentrations, high organic matter and high soil water saturation coincide (Galloway et al 2003). The amount of denitrification occurring in areas prone to water saturation in agricultural land has not been thoroughly studied (Galloway et al 2003). P is delivered to watercourses as particulate P and/or dissolved reactive P (Sharpley 2006). Particulate P is linked to soil erosion, and hence the importance of soil erosion reduction in agricultural areas (Rekolainen 2006). Particulate P and dissolved reactive P are mostly transported by runoff, although several studies have demonstrated substantial loss of particulate P and dissolved reactive P by subsurface drainage (e.g. Ulén and Mattsson 2003, Uusitalo et al. 2007). Several studies have pointed out that a substantial amount of P loss occurs in small areas, defined as critical source areas. In these areas, a high P concentration, high erosivity, the transportation capacity of the runoff, and the hydrological connectivity to the stream network coincide (Pionke et al. 2000, McDowell and Srinivasan 2009). However, the flow path length, the soil moisture condition, the soil texture and soil pH are also important factors influencing P loss (McDowell and Sharpley 2002).

N and P exports increase during high flow periods (spring and autumn) and decrease during low flow periods (winter and summer). Substantial amounts of annual nutrient loss from agricultural areas may occur in only a few extreme events (Royer et al. 2006). High-resolution (sub-hourly) water quality data also show a high temporal variability in
nutrient concentrations with different forms and directions of hysteresis (House and Warwick 1998, Bowes et al. 2005). Seasonal weather conditions also affect soil redox conditions; during snowmelt and in the autumn rainy period, soils are wet and the water table is high and soil oxygen is reduced, affecting the biogeochemical processes in the soil (Creed and Sass 2011). The soil wetness condition is also affected by topography. Depressions and flat areas are more prone to water saturation and for a longer time than the rest of the area. These areas can function as nutrient sinks for elements such as N, but they can also act as sources for some other elements e.g. P, which may be extracted when the runoff increases (Sharpley et al. 2008). In general, areas with high biogeochemical dynamics are defined as biogeochemically critical areas, and the periods of major alterations in soil biochemistry as biogeochemically critical periods (Vidon et al. 2010; Pinay et al. 2015).

3.3. Analysis scale in assessment of the effect of land use on water quality

The detection of the effect of land use on water quality in rivers is scale dependent (Buck et al. 2004). However, the concept of scale can be confusing due to its various definitions, such as map scale, analysis scale and phenomenon scale (Montello 2001). In this dissertation, the analysis scale is assumed to be the size of the analysis unit at which certain phenomenon are studied (Montello 2001), whereas the phenomenon scale is referred to as the scale in which certain phenomena are best represented (Zhang et al. 2004). In practice, it is very challenging to define precise scales in hydrological and water quality studies, and most studies simply use an arbitrary scale of analysis (where the data is available).

In general, in a large analysis unit, such as a large river basin, the effects of land use on water quality might not be detected, whereas the effect signal might be stronger in a smaller catchment. For instance, Burt and Pinay (2005) show a large nitrate flux variability in catchments ranging from 5 to 500 km², and a low variability in a very large basin. Therefore, the selection of an adequate catchment size to study factors affecting stream water quality is very important. Kyllmar et al. (2014) suggest small size catchments for the monitoring and detection of the effect of land management measures on stream water quality. A large catchment involves complex processes where auxiliary sources or sinks of pollutants can also become involved. Moreover, for P loss, there is a higher variability...
in small catchments than in large ones, although the difference is rather small according to Kronvang et al. (2007). Scientific literature also includes several empirical studies relating stream water quality to catchment features and land use in very different sizes of catchments; most of these studies identified the agricultural land use as the most important factor explaining water quality variability (e.g. Ekholm et al. 2000, Mattsson et al. 2009, Ekholm et al. 2015). However, other studies in larger catchments highlight the role of other factors like physiography (Varanka and Luoto 2012; Varanka et al. 2015), and lake and in-river processes (Malve et al. 2012, Wollheim et al. 2006). Additionally, a number of studies of land use effects on stream water quality have characterized the analysis unit by contrasting landscape features in the entire catchment area and in the riparian zone (e.g. Johnson et al. 1997, Sliva and Williams 2001, Guo et al. 2010). However, the criteria defining riparian or buffer zones vary and the results of the studies are conflicting and not comparable.

Runoff is the major driver of nutrient transportation, and runoff generation in the landscape is complex and varying in time and space (Grayson and Blöschl 2000). While runoff generation on a plot scale can be dominated by infiltration-excess overland flow and perceived immediately after a rainfall event, on the watershed scale the response can be delayed and several mechanisms can be involved and become dominant, such as saturation-excess, subsurface through flow, transmissivity feedback, etc. (Lepistö 1996). In headwater catchments, stream flows are generally more responsive to precipitation and they exhibit high flow and hydro-chemical variability (McGlynn et al. 2004, Laudon et al. 2007, Vivoni et al. 2008, Dawson et al. 2011), while in large catchments the variability is low or with attenuated hydrological response (Sanford et al. 2007).

3.4. Ecohydrological models

In the last decade, several computer-based hydrological, hydraulic, and water quality models have been developed, and they continue to evolve. Nowadays, hydrological models are more integrated and have incorporated the major physical, chemical, and biological processes occurring in terrestrial and aquatic ecosystems. Consequently, they are called ecohydrological models (Krysanova and Arnold 2008), but other names are also used in the literature, e.g. catchment models, biogeochemical models, integrated models. In general, there are several computer models, but
based mostly on the same few biophysical theories. Some commonly used models are: AGricultural Non-Point Source Pollution (AGNPS), Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS), Soil and Water Assessment Tool (SWAT), Geospatial Watershed Assessment (AGWA), Integrated Catchment Assessment (INCA), Hydrologic Simulation Program-Fortran (HSPF), Gridded Surface Subsurface Hydrologic Analysis (GSSHA), Water Evaluation And Planning System (WEAP), and finally the watershed simulation and forecasting system and water quality and nutrient load model (WSFS-VEMALA). These models can be classified in several groups, depending on the criteria of classification (Fig 1A), such as the conceptual approach (physical or empirical-based), spatial representation (distributed, semi-distributed or lump/spatial average models), and a simulation time-step (continuous or event-based), but also on the nature of the data input (determinist and stochastic models) (Grayson and Blöschl 2000). Due to the spatial nature of the environmental data and information, ecohydrological models are well integrated with geographic information system. The watershed is the most recognizable feature in a landscape and watershed scale modeling is widely applied (Mulligan 2004). The watershed area is also easily delineated in geographic information software using a digital terrain model, although river catchment area does not necessarily correspond to the topographic divisor, e.g. in a karstic watershed.

The selection of a particular type of model depends on the purpose of the application and data availability. If the purpose, for example, is a prediction of the maximum river discharge in the watershed outlet or in a site of interest, a simple lump hydrological model can perform well. In such a case, the objective is just to know the quantity of water delivered at the outlet, and it no emphasis is placed on where and how the water is stored and released within the catchment (Mulligan 2004). However, if the purpose is to evaluate a given water protection/management plan, a particular land use change effect or assess the catchment’s ecosystem services, a more advanced spatially distributed model is required. Such models consider where precisely the different hydrological and biogeochemical processes within the catchment occur.

To select which appropriate model complexity to apply can sometimes be challenging. In general, two model building approaches, bottom-up and top-down, are acknowledged in ecohydrological modeling (CRC-CH 2005):

1) In a bottom-up approach the modeling starts with the simplest model to
Theoretical Background

fit the observation data, e.g. precipitation and river discharge, and the model complexity is increased by adding more components contributing to a better fit of the observed data. 2) In a top-down approach all identified, relevant components are included in the model, although some of them are omitted due to sensitivity analysis, i.e. if they do not have a significant influence on the simulation of the observed data. Each approach has advantages and disadvantages. The selection of a proper level of model complexity should also be evaluated as regards the study objectives and what the available data can ultimately support. Too simple or too complex models lead to conceptual problems, and uncertainties in the parameters and data used (EPA, 2009) (Fig 1B).

Fig. 1. A. Types of hydrological and eco-hydrological models and their data requirements (based on Grayson and Blöschl 2001). B. Links between model complexity and uncertainty. Total model uncertainty (red line) increases when the model is too simple or too complex. Data uncertainty (blue dotted line) increases when the model complexity increases, while model framework uncertainty (green dashed line) decreases when the model complexity increases. Modified from Hanna (1988) and EPA (2009).
4. Materials and Methods

4.1. Study area

The study area included the watershed of the Yläneenjoki River (231 km²) and the Pyhäjoki River (78 km²), in southwestern Finland; and the headwatershed of the Lepsämänjoki River (22.4 km²), a tributary of the Vantaanjoki River in Southern Finland (Fig. 2). These areas were selected because long-term environmentally monitored data was available and several water protection measures had been implemented in the watershed. Yläneenjoki River and the Pyhäjoki River are tributaries of Lake Säkylän Pyhäjärvi, which has suffered eutrophication in the past decades. Currently, the lake has ecologically recovered, but is still highly sensitive to eutrophication due to its shallowness, intense agricultural practices in the catchment area and climate change (Ventelä et al. 2007). The Vantaanjoki River discharges into the Baltic Sea, where eutrophication is one of the main threats to the marine environment.

The main land cover forms in these watersheds are agricultural areas, forested areas and built up areas. As in other part of the Finnish rural landscape, the studied area has changed rapidly from the 1950s as a part of the overall shift towards intense mechanized agricultural practices; for example, open ditches have been replaced by tile drainage making the agricultural landscape more homogeneous (Hietala-Koivu 1999; Hietala-Koivu 2002). Since the 1990s, nearly all farmers in the study area have committed themselves to the European Union’s (EU) agri-environmental program and implemented water protection measures such as field soil erosion control through minimum tillage practices, the installation of stream vegetation buffer zones, as well as the construction of artificial wetland and sedimentation ponds (Ventelä et al. 2011, Aakkula and Leppänen 2014). Additional measures include the installation of nutrient trapping and filters in the drainage systems (Kirkkala et al. 2012).

The study areas belong to the boreal temperate zone. The terrain was formed by land uplift 5 600 BP (Eronen et al. 1982, Kirkkala 2014). The geological units in the zone are granitic migmatites, late Svecofennian granites, and rapakivi granites (Korsman et al. 1997, Skyttä and Mänttäri 2008). The dominant soil types in the studied watersheds are marine and lacustrine clay soils, glaciofluvial sandy till, and bedrock. The climate shows strong seasonality. The average precipitation in the southwestern
study area is 630 mm, and in the south 650 mm, of which snowfall represents 10-20%. The annual mean discharge of the rivers is 2.7 m$^3$s$^{-1}$ in Yläneenjoki, 0.9 m$^3$s$^{-1}$ in Pyhäjoki and 0.2 m$^3$s$^{-1}$ in the head watershed of Lepsämänjoki, a tributary of the river Vantaanjoki. All the rivers have two periods of high flow (April and November-December) and two periods of low flow (May-October and January-March).

Fig. 2. The study area included two adjacent watersheds of the Yläneenjoki River and the Pyhäjoki River, in southwestern Finland (1); and the head watershed of the Lepsämänjoki River, a tributary of the Vantaanjoki River in Southern Finland (2). Black arrows indicate river flow direction. The land use map is the SLICES (separated Land Use/Land Cover Information System) data (Sucksdorff and Teiniranta 2001).

4.2. Environmental data and information

The hydrological and water quality data collected by the Finnish Environment Institute (SYKE) was used (available at
River discharge and the following stream water quality variables were selected: nitrite and nitrate nitrogen concentrations (NO$_2$N, NO$_3$N), total nitrogen concentration (N$_{tot}$), dissolved reactive phosphorus concentration (DRP), total phosphorus concentration (P$_{tot}$), and total suspended sediment concentration (TSS). The selected water quality variables are related to N and P, which are the key limiting elements in aquatic systems. NO$_2$N and NO$_3$N were measured together, with NO$_3$N representing the largest portion, thus, henceforth they will only be referred to as NO$_3$N. Additionally, a high temporal resolution data of river discharge and turbidity recorded by automated station installed in the River Lepsämänjoki (Fig. 3A) was used (Article III). The data was provided by the Water Protection Association of the River Vantaanjoki and the Helsinki Region. Weather data was obtained from the Finnish Meteorological Institute (Article III and IV). Additionally, climate change data was obtained from World Climate Research Programme’s CMIP3 Multi-Model Dataset (Meehl et al. 2007).

Spatial data such as the stereo-photogrammetric 10-meter digital terrain model (DTM) and the Light Detection and Ranging (LiDAR) based 2-meter DTM were obtained from the National Land Survey of Finland, and the soil map using 1: 20 000 from the Geological Survey of Finland. The SLICES (separated Land Use/Land Cover Information System) land use data (Sucksdorff and Teiniranta 2001) was also used. Most of the spatial data needed pre-processing for the purpose of the research. Land use and soil data were reclassified into broader land cover types to be used for landscape indices estimation (Article I), erosion modeling (Article III) and ecohydrological modeling (Article IV).

Field work was carried out (Article II) in the snow melt period in 2012 in the River Lepsämänjoki area, by collecting suspended sediment samples in seven peak flow events with a time-integrated sediment sampler (Fig. 3B). A soil and sediment sampling campaign from agricultural fields, river banks, and channel beds was carried out in the same period. Cesium-137 (^{137}Cs) radioactivity and total phosphorus content were analyzed from the collected samples to identify sources of suspended sediment.
4.3. Data analysis and modeling

4.3.1. Statistical analysis and time-series trend detection

The following statistical models were applied in this research: a generalized linear model (GLM) (Article I) was used to study the relationship between the individual water quality variables and landscape indices; a multivariate redundancy analysis - RDA (Zuur et al. 2007) was used to study the linear response of multiple water quality variables to multiple landscape indices simultaneously and to find out relationship pattern between the variables (Article I); and a time series trend analysis (Article II) was carried out to find trend patterns in long-term water quality data and nutrient loads. For this purpose, the univariate Mann-Kendall trend test, the multivariate Mann-Kendall (MMK) trend test (Hirsch and Slack 1984, Hirsch et al. 2010) and the MMK trend test applied to flow-normalized water quality data were used. The flow normalization of water quality and nutrient loads data were carried out by a semiparametric regression method (Wahlin 2008).

4.3.2. Topographical-based mapping of wet areas

The soil wetness condition is an important aspect affecting hydrological and biogeochemical processes (Groffman et al. 2009, Creed and Sass 2011). Therefore, the accurate mapping of areas with soil prone to water saturation or wet areas was highly relevant (Galloway et al 2003, Ågren et al. 2015). High resolution LiDAR-DTM provided good information for the prediction and mapping of wet areas (Murphy et al. 2011). However, LiDAR-DTM often contains artifact pits and spurious topographic noise, because of a greater surface roughness at finer resolutions (MacMillan et
al. 2003; Zandbergen 2006). In addition to which, the inability of LiDAR-DTM to represent manmade underpass infrastructures, for instance bridges and road culverts, hinders accurate drainage pattern modeling (Li et al. 2013, Lindsay and Dhun 2014). Therefore, for a hydrological application, LiDAR DTM needs to be pre-corrected and artificial discontinuous flows over the terrain need to be avoided (Soille 2004, Lindsay and Creed 2005, Wang and Liu, 2006). To identify the manmade drainage obstructions, a simple boolean algebra operation was carried out with a grid of roads and stream network topographic data in a geographic information system. After which, the DTM was hydrologically corrected with an ANUDEM algorithm (Hutchinson, 1989), which couples the minimization of a terrain specific roughness penalty with an automatic drainage enforcement.

Another aspect affecting DTM-based hydrological index estimation is the flow direction algorithm selection (Seibert and McGlynn 2007). There are several algorithms to compute flow direction. For example the eight flow direction (D8) (O’Callaghan and Mark 1984) multiple flow direction (Freeman 1991), digital elevation model network (DEMON) (Costa-Cabral and Burges, 1994), infinite possible flow directions (D∞) (Tarboton, 1997), and triangular multiple flow direction (MD∞) (Seibert and McGlynn 2007). A detailed revision of these algorithms is provided by Erskine et al. (2006) and by Seibert and McGlynn (2007). In general, multiple flow direction algorithms are preferred.

The most widely applied DTM-based hydrological indices for the mapping of wet areas are the topographic wetness index (TWI), and its variants, e.g. downslope TWI (Hjerdt et al. 2004), hydrological connectivity index (Lane et al. 2009), catchment area modified TWI in SAGA GIS or MTWI (Böhner and Selige 2006). TWI is also used to identify landscape groundwater recharge and discharge areas (Brydsten 2008). TWI is a component of the rainfall-runoff model called TOPMODEL and it is defined as \( \ln(a/\tan\beta) \), where \( a \) is a cumulative upslope area per unit contour length or a specific catchment area, while \( \beta \) is the slope. The natural logarithm (ln) scales this index into a linear range (Beven and Kirkby 1979, Grabs et al. 2009). Another approach for wet area mapping is by estimating the elevation above the stream network, such as the cartographic depth to water index (White et al. 2012) and height above the nearest drainage (Nobre et al. 2011). All the above mentioned indices assume steady-state conditions and spatially invariant conditions for
infiltration and transmissivity. Additionally, the combination of two or more DTM-based indices, e.g. terrain classification index for lowlands (TCI\textsubscript{low}) (Bock et al. 2007), provide further tools for terrain analysis and wet area mapping. The TCI\textsubscript{low} index combines an inverted normalized STWI with altitude above the channel network.

Mapping wet areas based on terrain hydrological indices can be classified into classes of potential wetness condition. In the simplest form, the classification is performed using threshold values (Article I and Article III). However, the selection of an appropriate value for the threshold is challenging as it depends on the characteristics of a particular terrain, and many times an arbitrary value is taken (Mengistu et al. 2014). For a more precise terrain index based wet area mapping, a machine learning algorithm such as a random decision forest algorithm (James et al. 2013) can be applied. For this purpose, a training and evaluation data was generated from the photogrammetric-based wetland map of the study area. The relative classification accuracy was evaluated by estimating the error matrix (Congalton 1991) by comparing the classification result to existing maps of wetlands.

Riparian zones are important wet areas. DTM-based mapping of riparian zones is better obtained by estimating the elevation above the stream network (Article I). This method produces riparian zones with varying width and identifies areas with high potential of hydrological connection to the stream network (Fernández et al. 2012). The riparian zones classification was performed by using threshold values of river maximum daily flow elevation corresponding to a 10 year return period, which was estimated from the time-series river discharge data. Wet area mapping based on the elevation above the stream network was also used to identify potential areas in order to construct an artificial wetland and model the potential nutrient removal (Article IV). Such areas might pose a longer hydro-period or a longer time per year in which the area remains in a water saturated condition than the surrounding areas, and therefore it is more efficient for nutrient removal.

**4.3.3. Ecohydrological modeling**

A physical-oriented soil erosion model named Modified Morgan-Morgan-Finney (MMF; Morgan and Duzant 2008) (Article III) and Soil and Water Assessment Tool (SWAT, Arnold et al. 2012a) (Article IV) were used in this study. SWAT is a semi-distributed ecohydrological model operating with daily time steps. It has integrated algorithms to simulate watershed
hydrology, soil erosion, and some major terrestrial bio-geochemical processes, including C, N and P cycling, ground water flow, plant growth and land management effect (Arnold et al. 2012a). In SWAT, an in-river biochemical transformation is simulated by coupling a one-dimensional hydrodynamic model and the enhanced water quality model (QUAL 2E) (Brown and Bernwell 1987, Arnold et al. 2012a). Spatial data input required for the SWAT model was prepared by using a SWAT geographic information system interface (ArcSWAT and QSWAT). SWAT needs a large number of parameter values, which were completed from previous modeling experiences in the study area (Bärlund et al. 2007, Tattari et al. 2009), literature reviews, and by using a pedo-transfer function to derive soil characteristics from texture data (Schaap 2005).

The ecohydrological modeling procedure included the model set up, its calibration and validation, as well as a sensitivity and uncertainty analyses for the model (Abbaspour 2015). Computation proceeds as an iterative process until an acceptable level of model performance is reached. Then, the observed data was separated into calibration data and validation data. Several methods exist to estimate the general model performance. The commonly used methods are regression ($r^2$) between the simulated and observed data, a root mean square error based (RMSE) on the value of 0 indicating a perfect fit, and the Nash-Sutcliffe efficiency (NSE), in which normalized statistics that determine the relative magnitude of the residual variance (noise) is compared to the measured data variance (information). NSE ranges from $-\infty$ to 1, and values close to 1 indicate good performance, while values <0.0 indicate a poor model performance (Moriasi et al. 2007). These methods are greatly affected by extreme values (outliers), and regression methods are insensitive to additive and proportional differences between model output and observation. Alternative methods, which reduce the limitation of the previous methods for model evaluation, include the modified NSE, percent bias (PBIAS), and the RMSE-observations’ standard deviation ratio (RSR). PBIAS estimates the relative average bias of the simulated output to the observed data. Positive values indicate model underestimation bias, negative values model overestimation bias and zero values non bias (Gupta et al. 1999, Moriasi et al. 2007). RSR standardizes the RMSE with the observations’ standard deviation, which allows comparison of different variables with different data variability. It varies from zero to large positive values, values close to zero indicating zero RMSE or residual variation and therefore no error in the simulation output.
Materials and Methods

(Moriasi et al. 2007). The advantages and disadvantages of these methods are discussed in Moriasi et al. (2007), Krause et al. (2005) and Abbaspour (2015). In this dissertation, the model performance evaluation was based on combined criteria using several statistics as in Moriasi et al. (2007) and Arnold et al (2012b). For example, in a monthly time-step simulation, satisfactory model performance in predicting discharge might reach NSE >0.5, RSR < 0.7 and PBIAS ±25%. Similarly, for nitrate and total phosphorus loads prediction: NSE of >0.5, RSR < 0.7 and PBIAS ±70%

In the modeling process, the initial parameter value was usually far from suitable, which lead to a large discrepancy between the simulation and observed data. Therefore, the initial model set up and parameter values needed to be adjusted by model calibration (Grayson and Blöschl 2001). Ecohydrological models involve a large number of parameters to be adjusted and this demands computer-based calibration procedures (Abbaspour 2015). Inverse modeling is a widely used calibration procedure technique (Vrugt et al. 2008), where the true values of the parameters are assumed to be unknown, because most of the parameter values are obtained in a small plot experiment and laboratory analysis, and might not be representative of spatially highly varying parameters (Beven 2001). Several mathematical models have been developed to identify a set of optimal values of parameters for the inverse modeling approach. An extensive review of these methods can be found in Matott et al. (2009) and Zhang et al. (2009). In this dissertation, a comparative calibration using the following parameter optimization algorithm was carried out: Parameters optimization and uncertainty analysis tool (ParaSol), generalized linear uncertainty estimation (GLUE), sequential uncertainty fitting (SUFI), particle swarm optimization (PSO) (Abbaspour 2015).

Sensitivity analysis attempts to identify which parameters and data inputs have the strongest effect on the model output. Sensitivity analysis can be separated into two general categories: local sensitive analysis and global sensitive analysis (Cibin et al. 2010). The first explores the influence of an individual parameter at one point in time. A commonly used local sensitive analysis method is the one-at-a-time (OAT) analysis, which shows the sensitivity of a variable to the change in one parameter while holding all the other parameters constant (Abbaspour 2015). In order to increase sampling efficiency, latin hypercube sampling is usually coupled with OAT. In latin hypercube the sampling process is stratified along the range of parameter value variation (Abbaspour 2015). A global sensitivity analysis
Attempts to analyze the influence of all factors simultaneously (Cibin et al. 2010). To this end, the Monte Carlo approach is commonly used. Here an intense repeated sampling is carried out over a specified range of parameters values. This incorporates the variance of the input variable on the model output (EPA 2009). The parameter sensitivity statistic is evaluated by multiple regressions of the parameters generated by latin hypercube sampling against the objective function values (Abbaspour 2015). The relative significance of each parameter is determined by t-stat, which is the coefficient of the parameters divided by its standard error. A large absolute value for a t-stat with low p-values (<0.05) indicates a sensitive parameter (Abbaspour 2015).

In the context of ecohydrological modeling, uncertainty can result from incompleteness and/or misspecification of the model structure, unknown true parameter values, poor spatial and temporal data quality, errors in computational algorithms, etc. (EPA 2009). Therefore, an uncertainty assessment is an effort to qualitatively or quantitatively evaluate model reliability (Grayson and Blöschl 2000). Model calibration and sensitivity and uncertainty analyses are interlinked, and generally the same algorithm of model calibration and parameter sensitivity analysis provides a tool for uncertainty assessment, which can be expressed for example as P-factor and R-factor. The P-factor expresses the percentage of observed data enveloped by the 95 percentage of the predictive uncertainty band (PPU95), which is on the 2.5% and 97.5% level of the accumulative distribution of the model output in a stochastic calibration. The R-factor indicates the thickness of the PPU95, ranging from zero to infinity. Values close to zero indicate low uncertainty in the model prediction (Abbaspour 2015).

The model calibration as well as the sensitivity and uncertainty analyses was carried out with SWATCUP software (Abbaspour 2015). The calibrated SWAT ecohydrological model was used to evaluate climate change impact on nutrient loads, watershed resilience, and the effect of restoration on a wet area with a long hydro-period (Article IV).
5. Results and Discussion

5.1. Stream water quality indicates both land-use and riparian processes

The results of my research (Article I) show that the spatial variation of nutrient concentrations in stream was ($N_{tot}$, NO$_3$-$N$, P$_{tot}$, and DRP) directly related to the portion of agricultural land-use in the entire watershed. When the percentage of agricultural land has increased in the watershed, the concentration of $N_{tot}$, NO$_3$-$N$, P$_{tot}$, and DRP also increased. However, the strength of the relationship varied seasonally and across the scale of analysis (watershed-wide, zones of saturation-excess runoff and riparian areas). This finding agrees with the results found by other studies claiming that agricultural land use is the main factor affecting stream nutrients variations on a watershed scale (e.g. Mattsson et al. 2009, Varanka and Luoto 2012, Ekholm et al. 2015). However, these studies present different scales of analysis, for instance, different catchment sizes. Similarly, studies considering riverine areas usually delineate stream buffer zones by using an arbitrary width (e.g. Sliva and Williams 2001, Guo et al. 2010). These arbitrary boundaries may proffer little explanatory power as regards spatial water quality variability (Clerici et al. 2013), since they may not capture the most important hydrological and biogeochemical processes in the landscape.

This dissertation claims a process-oriented scale of analysis, in which the analysis unit for water quality studies and nutrient export from terrestrial ecosystem into aquatic systems must be based on the clear representation of the hydrological and biogeochemical functional unit of the landscape. Although watershed is a functional hydrological unit of the landscape, internal watershed characteristics need to be better represented, such as variable source areas or zones of saturation-excess runoff and hydrological meaningful riparian areas (Baker et al. 2006, Yang et. al. 2011, Lane et al. 2009). The digital terrain model based definition of the scale of the analysis applied in this dissertation poses advantages in explaining spatial water quality variation, nevertheless, in a heavily altered agricultural landscape it seems to be hampered by artificial agricultural drainage, which enhances the delivering of nutrients from agricultural fields to streams (Ulén and Mattsson 2003, Uusitalo et al. 2007, Arango...
Results and Discussion

and Tank 2008). Therefore, the agricultural area on a watershed-wide scale becomes relevant for most of the studied water quality variables.

Seasonally, changes in the strengths of the relationships between landscape and water quality variables might reflect the hydrobiogeochemical seasonality in boreal climate. It was observed (Article I) that \( N_{\text{tot}} \) and \( \text{NO}_3-N \) were well predicted by the total agricultural area of the watershed in winter, spring, and autumn, when most of the watershed area tends to be well hydrologically connected. However, during the summer low flow period they were better explained by the portion of agricultural area in the saturation-excess zone of the watershed, but with much less explanatory power. In addition, the riparian zones became important nitrate sinks only in autumn. Some field studies have shown that the nutrient sink capacity of riparian zones is less efficient in the winter and spring snowmelt period, than in wet autumn conditions (Arango and Tank 2008). \( P_{\text{tot}} \) and DRP were also well predicted by watershed-wide agricultural areas in all seasons. Additionally, the saturation-excess zone was also important for DRP in spring and autumn. This might reflect additional DRP sources from organic saturated areas in high flow periods, because re-flooding organic soils increase DRP release (e.g. Reddy 1983, Sharpley et al. 2008). Suspended sediment concentration was well predicted by the portion of agricultural land located in the saturation-excess zone and clay soils in all seasons. This might be linked to the high erodibility of clay soils in agricultural lands combined with the high runoff production in saturated areas (Singh et al. 2009), although soil erosion in agricultural land is commonly linked to the slope; it is assumed that steeper slopes present high soil erosion. However, runoff is usually increased in saturated area and can cause severe erosion, particularly in wet seasons (Singh et al. 2009).

The study on the sediment transfer from agricultural lands in the snowmelt period in the Lepsämänjoki River watershed (Article III) provided a more detailed insight concerning sediment mobilization from agricultural areas into watercourses. Based on the high temporal resolution hydrometric data, hysteresis analysis of sediment concentration showed a clock-wise pattern for most of the snowmelt events, suggesting a rapid mobilization of sediments. A similar pattern has also been observed in other small catchments and not only in sediment concentrations but also for \( P_{\text{tot}} \) (House and Warwick 1998, Bowes et al. 2005). However, by analyzing only the hysteresis pattern, the origin of the sediments remains
unclear. They can originate from the nearest (agricultural) areas, from channel sediment resuspension or from bank erosion. To identify the dominant source of suspended sediments, sediment fingerprint with Cesium-137 ($^{137}$CS) was analyzed. The result showed that the agricultural fields are the dominant source of suspended sediments showing a higher $^{137}$CS radioactivity and higher content of $P_{\text{tot}}$ than the sediments from the channel bank and bed. Additionally, by modeling soil erosion with a process-oriented erosion model in GIS with LiDAR data, a high erosion rate was observed in agricultural areas with high runoff, mainly located in saturation excess zones. In addition, the spatial patterns of erosion rate for the various sizes of sediment - clay, silt and sand - was different. Fine clay and silt sediments presented highest erosion rate and broadest eroded area. Based on the modeling results, the sediment contribution of agricultural land into streams was estimated, and it was shown that more than 75% of the main stream segment is potentially affected by sediment eroded from agricultural fields (Article III). It is important to note that fine sediment enrichment can occur during the snowmelt erosion process and these sediments are highly active chemically and transport large amounts of agrochemical pollutants, including P (e.g. Uusitalo et al. 2001, Hartikainen et al. 2010, Kleinman et al. 2011).

5.2. Emerging trends in stream water quality variables

The results of the time series analysis of water quality data in the Yläneenjoki River and Pyhäjoki River showed different trends (Article II). In the Yläneenjoki River, an increasing trend was found in the concentrations and loads of $N_{\text{tot}}$, NO$_3$-N, DRP and a decreasing trend in the concentration of total suspended solids. In the Pyhäjoki River, no clear trends were detected in any of the studied water quality variables. An increasing trend of loads and concentrations of $N_{\text{tot}}$, and DRP and a decreasing trend of $P_{\text{tot}}$ have also been found in other agricultural watersheds in Finland (Ekholm et al. 2015, Rankinen et al. 2016, Tattari et al. 2017) and elsewhere (Jarvie et al. 2017).

In boreal watersheds, inter-annual climatic variations induce high variability in stream water quality confounding the land management effects (Stålnacke and Grimvall 2001). Previous studies in the Yläneenjoki River have found a relationship between increased nutrient loading and the occurrence of mild winters, which have become more frequent in the last years (Ventelä et al. 2011). Currently, the reliability of trend detection is
constrained by observed data availability. Conventional water quality monitoring data provides only temporally sparse observations, which miss important peak flows (Ekholm et al. 2015). River discharge data is available as a continuous daily observation and it is generally used as surrogate data to interpolate water quality concentrations and to improve load estimations. However, weak linear relationships were found between flow, solute and sediment concentrations in my study area, as in many other catchments (e.g. Johnes 2007). In the light of high-frequency monitoring data, the concentration of water quality variables exhibits a hysteresis of varying form and direction, depending on the antecedent conditions of nutrient flux from land to water (Article III, House and Warwick 1998, Bowes et al. 2005). Therefore, there is an unavoidable degree of uncertainty in long-term trend detection using conventional data. Some emergent methodologies, e.g. wavelet transforms (Nalley et al. 2012) and spectral fractal scaling (Kirchner and Neal 2013), could help to address this problem by combining high-frequency sampling data with conventional data and reconstructing long-term data. However, such methodologies are still under development. Currently, the availability of high-frequency monitoring data is still scarce; it exists only for short-term and only for a few water quality variables. Also, automatic station generated data is not free of errors (Kirchner et al. 2004).

The multivariate semiparametric regression model was used to flow-normalize the water quality variable concentrations and loads (Article II). This way the effect of flow variation in the water quality data was minimized. This method attempted to complete the information from several time-series providing an advantage compared to other regression-based methods of flow normalization using singular time-series (Wahling and Grinvall 2008). Consequently, the trend analysis using flow-normalized data can be related to other causes than to the variation in river flow. However, in my study area, the normalization was only effective for the loading data, but not for the concentrations, which exhibited weak correlation with the flow data. In spite of this, the trend analysis provided an important clue as regards the tendency of water quality in the study area. The decreasing trend of TSS concentration might be linked to the erosion-reducing land management practices implemented in the last decade. The increasing trends of the other solutes (TN; NO3-N and DRP) might reflect a response to the same practices, as many field studies have found an increasing mobilization of solutes in areas with minimum-tillage practice.

It was also found in this dissertation that long-term trends will be influenced by watershed characteristics (Article II). The watershed of the Yläneenjoki River, where trends for most of the water quality variables were detected, is characterized by clay soils, and shows a low base-flow index. This suggests that a lower portion of the flow in the river was sustained by groundwater contribution. In contrast, the watershed of the Pyhäjoki River, where no trends were detected for most of the water quality variables, is a sand soil dominated watershed and shows a higher base-flow index. Therefore, it seems that there is a higher contribution of groundwater to sustain the river flow. These aspects make the response of stream water quality to land management confusing, although in both watersheds similar water protection measures were implemented.

5.3. Topographic-based wet area mapping is highly influenced by the modeling techniques

DTM-based wet area mapping resulted in different spatial patterns when applying different DTM resolutions and algorithms to calculate the topographic wetness index (TWI). The standard computing of TWI was greatly influenced by small differences in elevation and flow direction, as these affected the calculation of the slope and upslope catchment area respectively (Franchiniav et al. 1996, Erskine 2006, Lin et al. 2010). A modified topographic wetness index (MTWI) corrected this fact. For example, with LiDAR DTM (2 m resolution), the wet area pattern computed by using MTWI resulted in a more homogeneous pattern in gentle slopes, and also more extreme values, than when using the standard TWI (Fig. 4). Wet area estimation based on the coarse DTM resolution (10 m resolution) also resulted in different spatial patterns when applying different algorithms for the topographic wetness index (TWI or MTWI), but both tended to estimate higher values than with the fine resolution LiDAR data, and more uniform and broader wet areas. Similar results were also obtained by Lin et al. (2010) and Buchanan et al. (2014). In spite of the differences in the spatial pattern produced by the different algorithms and digital terrain model resolutions, the overall classification accuracy of the classified wet areas by the random forest machine learning approach was comparable. The TWI and MTWI from the LiDAR data both resulted in an accuracy of 70%, and from a DTM with a 10 m resolution in 75%.
Results and Discussion

Wet area mapping by estimating the elevation above the stream network resulted in a different spatial pattern than the mapping based on TWI in my study area. This method does not take into account the flow contribution from upland areas and only identifies areas with a similar altitudinal elevation to the stream network (Murphy et al. 2011, Ågren et al. 2014). Nevertheless, this method is useful for riparian zone delineation, as it identifies wet areas surrounding the stream network. The identification of riparian zones with hydrological and biogeochemical meaning was important for a better statistical inference of land use change and land management impact on stream water quality (Article I). The application of combined indices resulted in a more accurate prediction, for example, the wet area mapping based on the TCI$_{low}$ index resulted in an accuracy of 82%. However, the evaluation of mapping accuracy solely by classification accuracy index is not sufficient, and a more detailed spatial pattern prediction screening is required. For instance, even though the accuracy index of wet areas based on the TCI$_{low}$ index was the highest, the method tended to overestimate the wet areas extension more than other methods, and it was observed that wet areas were identified even in high slope areas. Overall, using MTWI with LiDAR DTM resulted in a more consistent mapping of wet areas. The inclusion of dynamic hydrological
processes in the wet area estimation can improve the accuracy of wet area mapping (Sørensen et al. 2005, Grabs et al. 2009)

5.4. High spatial and temporal uncertainty in ecohydrological modeling

The modeling results showed that adopting different approaches to a discretized landscape unit for modeling in semi-distributed models plays an important role in spatial pattern prediction, such as runoff and nutrient sources. Figure 5 shows the predicted annual average of runoff (2000-2013) of a SWAT model with two different forms to define the modeling unit or hydrological response unit (HRU). The commonly used delineation of HRU was based on overlapping land use, soil types and slope, and the alternative approach for HRU delineation was based on a topographic wetness index instead of slope. The first approach predicted runoff occurrence in a broad area of the watershed but with a low runoff rate. In the second approach, the runoff occurrence areas were substantially reduced, but with high runoff rates, particularly in areas prone to water saturation and agricultural areas in clayey soils. Even though the prediction patterns were different, both modeling approaches showed similar statistical performance (Table 1), and both approaches were able to reproduce the hydrological seasonal pattern in an acceptable manner. Similar results were also found in others studies (Easton et al. 2008, Arnold et al. 2010). This predictive difference in the spatial pattern of hydrological processes, which also influence on biogeochemical processes, has important implications for realistic identification of nutrients source areas, particularly for the assessment of ecosystem services, environmental impact, and catchment management plans. Therefore, a better representation of the heterogeneity and complexity of the landscape and its ecohydrological process is required (McGlynn 2004, Tetzlaff et al. 2009). Currently, most of the modeling approaches geared to overcoming this problem are oriented towards fully spatially distributed, grid-based models (Gorgan et al. 2012, Rathjens et al. 2015, Nijzink et al. 2016). However, adding more details into the model does not always guarantee an improvement in the model’s prediction (Arnold et al. 2015), nor a proper representation of the internal organization of the landscape system. On the contrary, it leads to higher levels of predictive uncertainty and equifinality (Savenije 2010).
Results and Discussion

Fig. 5. Annual average runoff prediction (2000-2013) in the Yläneenjoki River watershed (Finland). Two types of landscape discretization for defining the hydrological response unit (HRU) in the Soil and Water Assessment Tool (SWAT) model is shown. A) The commonly used approach in the SWAT model to define HRU, based on an overlapping slope, land use, and soil types. B) The HRU was defined by replacing the slope by a topographic wetness index in the previous approach.

Table 1. Statistical evaluation of the SWAT model simulating the monthly flow in the Yläneenjoki River. Coefficient of determination ($r^2$) of the simulated and observed data, Nash-Sutcliffe model efficiency coefficient (NSE), percent bias (PBIAS), ratio of the mean square error to the standard of measured data (RSR), P-factor (representing the percentage of observed data enveloped in the 95% prediction uncertainty band, 95PPU), and R-factor (represent the thickness of the 95PPU band). Model A is the standard modeling approach and the model B is the topographic wetness index–based modeling.

<table>
<thead>
<tr>
<th></th>
<th>$r^2$</th>
<th>NSE</th>
<th>PBIAS (%)</th>
<th>RSR</th>
<th>P-factor</th>
<th>R-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>0.67</td>
<td>0.68</td>
<td>21.2</td>
<td>0.64</td>
<td>0.63</td>
<td>1.25</td>
</tr>
<tr>
<td>Model B</td>
<td>0.73</td>
<td>0.72</td>
<td>-0.8</td>
<td>0.53</td>
<td>0.69</td>
<td>0.65</td>
</tr>
</tbody>
</table>

A hydrogeomorphic-based modeling unit in a semi-distributed model offers an advantage for simulations by identifying more precisely those zones with different hydrologic, biogeochemical, and ecological processes and functions. Such properties are largely influenced by the geomorphic features of the landscape. Hydrogeomorphic indices, like the topographic wetness index, have been used in hydrological modeling (Beven 2006). In semi-distributed models, classifying the landscape into recharge, transition, and saturation zones allows the model, in a simpler manner and
retaining the model parsimony, to adjust or calibrate the model according to the dominant hydrological condition of the zone. For example, in the SWAT model, the runoff curve number can be adjusted according to the different soil moisture condition assumed for the hydromorphic zones. A hydro-geomorphic based SWAT model adjustment has already been adopted in some studies (Collick et al. 2015; Fuka et al. 2016,). There are also emerging methods for hydrologic simulation based on landscape bio-functional zones (Savenije 2010; Savenije and Hrachowitz 2017).

The calibration, sensitivity, and uncertainty analysis of the SWAT ecohydrological model (HRU based on a modified topographic wetness index) with different algorithms, using the same objective functions, showed comparable statistics to the model performance for discharge prediction, but with some differences in the parameter sensitivity analysis (Table 2). The common parameters identified as sensitive by all the calibration algorithms were a delayed time for aquifer recharge (GW_DELAY), the aquifer percolation coefficient (RCHRG_DP), the available water capacity of the soil layer (SOL_AWC) in forest and cropland, and the base recession constant (ALPHA_BF). Additionally, some algorithms also identified the saturated hydraulic conductivity (SOL_K), threshold water level in a shallow aquifer for the base flow (GWQMN), the minimum snow water content of the minimum depth above which there is 100% cover (SNOCOVMX.bsn), and the initial moisture condition SCS II curve value (CN2) in cropland with a hydrological soil group A as the sensitive parameters. These parameters were rather different from the identified sensitive parameters in the previous study by Tattari et al. (2009) using a SWAT model with a standard HRU delineation. For example, in this study, the soil depth (SOL_Z), the snow pack temperature lag factor (TIMP), the soil evaporation compensation factor (ESCO), and the initial moisture condition SCS II curve value CN2 for all land use classes, were identified as highly sensitive parameters. One reason for these differences might be the different ranges of parameters used in the calibration. In addition, the parameter calibration and sensitive analysis were carried out in a more detailed HRU in this study. For example, for the CN curve, the calibration was carried out for each of the land use types occurring in each hydrological soil group, while in the studies carried out by Tattari et al. (2009) a single parameter range for CN value calibration for all lands use types in all soil types was used. On the other hand, in a highly-parameterized model there are several,
if not thousands, of combinations of parameters that can lead to a similar solution – this is the non-uniqueness or equifinality problem described by Beven (2006), and is independent of the calibration methods used. When the number of parameters is high and the calibration ranges are broad, it leads to a higher level of predictive uncertainty, making the model output unreliable and difficult to interpret (Cibin et al. 2010; Zhang 2015; Houshmand Kouchi et al. 2017). Therefore, reducing model uncertainty due to the parameter definitions is one of the most important steps in the modeling process. Retaining the model parsimony and defining the proper initial parameter ranges for calibration is recommended. Moreover, as little information is usually available to validate all the models output quantitatively, soft calibration plays an important role (Yen et al. 2016). This includes the use of empirical knowledge about the system functioning, for example, the rate of denitrification in agricultural fields. Increasing the model complexity by adding more processes, does not improve the model’s statistical performance (Orth et al. 2015). However, processes considered important need to be included in the model. For example, subsurface drainage is an important component influencing agrochemical mobilization in the Finnish agro-system. Applying base-flow separation in the model calibration improves the statistics of the model performance (Zhang et al. 2013, Zhang et al. 2015).
Table 2. Sensitivity analysis of SWAT model parameters by Particle Swarm Optimization (PSO), Sequential Uncertainty Fitting (SUFI), ParaSOL: Optimization and uncertainty analysis tool, Generalized Likelihood Uncertainty Estimation (GLUE). The HGS stand for Hydrological Soil Group. Wet areas restoration shows that it may have the potential to mitigate the effects of climate change on water quality.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>PSO</th>
<th>SUFI</th>
<th>ParaSOL</th>
<th>GLUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum snow water content of the minimum depth above which there is 100%</td>
<td>-0.03</td>
<td>0.98</td>
<td>0.44</td>
<td>0.66</td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Cropland with HSG A</td>
<td>-0.15</td>
<td>0.88</td>
<td>-0.29</td>
<td>0.77</td>
</tr>
<tr>
<td>(CN2.mgt_B___SWHT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Cropland with HSG D</td>
<td>0.25</td>
<td>0.80</td>
<td>-1.80</td>
<td>0.07</td>
</tr>
<tr>
<td>(CN2.mgt_D___SWHT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth to impervious layer (DEP_IMP.hru___SWHT)</td>
<td>-0.51</td>
<td>0.61</td>
<td>0.30</td>
<td>0.77</td>
</tr>
<tr>
<td>Soil evaporative compensation coefficient (ESCO.hru)</td>
<td>-0.70</td>
<td>0.48</td>
<td>1.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Forest with HSG D</td>
<td>0.25</td>
<td>0.80</td>
<td>-1.80</td>
<td>0.07</td>
</tr>
<tr>
<td>(CN2.mgt_D___FRSE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Forest with HSG A</td>
<td>0.25</td>
<td>0.80</td>
<td>-1.80</td>
<td>0.07</td>
</tr>
<tr>
<td>(CN2.mgt_A___FRSE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Cropland with HSG A</td>
<td>0.25</td>
<td>0.80</td>
<td>-1.80</td>
<td>0.07</td>
</tr>
<tr>
<td>(CN2.mgt_A___SWHT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface runoff lag coefficient (SURLAG.bsn)</td>
<td>-1.12</td>
<td>0.26</td>
<td>-1.56</td>
<td>0.12</td>
</tr>
<tr>
<td>Initial Moisture condition SCS II curver value; in Forest with HSG B</td>
<td>-1.24</td>
<td>0.22</td>
<td>-0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>(CN2.mgt_B___FRSE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial residuo cover or Material in the residue pool for the top 10mm of</td>
<td>1.06</td>
<td>0.29</td>
<td>0.28</td>
<td>0.78</td>
</tr>
<tr>
<td>soil, in cropland (RSDIN.hru___SWHT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold water level in shallow aquifer for revap (REVAPMN.gw)</td>
<td>1.31</td>
<td>0.19</td>
<td>1.76</td>
<td>0.08</td>
</tr>
<tr>
<td>Saturated hydraulic conductivity (SOL_K(1).sol)</td>
<td>1.61</td>
<td>0.11</td>
<td>4.48</td>
<td>0.00**</td>
</tr>
<tr>
<td>Fraction of snow water equivalent that provides 50% of snow cover (SNO50COV.bsn)</td>
<td>-1.72</td>
<td>0.09</td>
<td>-0.87</td>
<td>0.38</td>
</tr>
<tr>
<td>Mannin's n value for overland flow in cropland (OV_N.hru___SWHT)</td>
<td>1.78</td>
<td>0.07</td>
<td>0.24</td>
<td>0.81</td>
</tr>
<tr>
<td>Threshold water level in shallow aquifer for base flow (GWQMN.gw)</td>
<td>-2.01</td>
<td>0.04*</td>
<td>-3.70</td>
<td>0.00**</td>
</tr>
<tr>
<td>Maximum canopy storage, in forest (CANMX.hru___FRSE)</td>
<td>-3.06</td>
<td>0.00**</td>
<td>-0.21</td>
<td>0.83</td>
</tr>
<tr>
<td>Delay time for aquifer recharge (GW_DELAY.gw)</td>
<td>-3.83</td>
<td>0.00**</td>
<td>-4.29</td>
<td>0.00**</td>
</tr>
<tr>
<td>Aquifer percolation coefficient (RCHRG_DP.gw)</td>
<td>3.89</td>
<td>0.00**</td>
<td>5.84</td>
<td>0.00**</td>
</tr>
<tr>
<td>Revap coefficient (GW_REVAP.gw)</td>
<td>-6.21</td>
<td>0.00**</td>
<td>-7.47</td>
<td>0.00**</td>
</tr>
<tr>
<td>Available water capacity of the soil layer, in cropland (SOL_AWC(1).sol___SWHT)</td>
<td>-9.47</td>
<td>0.00**</td>
<td>-13.49</td>
<td>0.00**</td>
</tr>
<tr>
<td>Available water capacity of the soil layer, in forest (SOL_AWC(1).sol___FRSE)</td>
<td>-10.90</td>
<td>0.00**</td>
<td>-1.83</td>
<td>0.07**</td>
</tr>
<tr>
<td>Base recession constant (ALPHA_BF.gw)</td>
<td>50.70</td>
<td>0.00**</td>
<td>55.61</td>
<td>0.00**</td>
</tr>
</tbody>
</table>

Results and Discussion
Finally, there are other sources of predictive uncertainty related to the data quality, both spatial and temporal. The result of my modeling approach indicated a better model performance for the continuously monitored variables, such as stream flow, but a low performance for those variables with temporally sparse observations (water quality variables) (Article IV). Even though the hydrological process are relatively more simple to predict than biogeochemical processes, some studies using temporally more frequent continuous data were able to improve the nutrient export accuracy (e.g. Jeong et al. 2010; Yang et al. 2016). Predictive uncertainty due to spatial data quality was not evaluated, but a high resolution of data for soil, land use, and topography was used in the modeling. In general, the results of studies comparing different data resolutions found that high spatial resolution data tends to obtain a better ecohydrological prediction (Geza and McCray 2008, Yang et al. 2014, Son et al. 2016, Thomas, et al. 2017, Fisher et al. 2017).

Modeling assessment of runoff, stream flow, nutrient loads, as well as soil erosion showed substantially higher values in the climate change scenarios (2046-2064) than current, and they were particularly high in winter and spring-time (Jan-May). However, there is a considerable degree of uncertainty in the modeling output, and as consequence, the results of nutrient load assessment in climate change scenarios are diverse, particularly in the quantitative estimation of future nutrient loads (Arheimer et al. 2012, Rankinen et al. 2013, Huttunen et al. 2015). However, most of the studies agree in suggesting an impairment risk to aquatic ecosystems and Baltic Sea due to excess nutrient loading. Similarly, empirical studies using current weather data have found an increasing export of nutrients in mild winter years (Puustinen et al. 2007, Ventelä et al. 2011), and there is a general assumption that nutrient export in climate change scenarios will increase (e.g. Kirkkala 2014).

The discrepancy in the quantitative assessment of future nutrient loading is caused by different methods, different periods, and the data sources adopted in the ecohydrological modeling. Therefore, a more careful and transparent modeling approach is required, such as by explicitly documenting data sources, parameter definitions, modeling sensitivity, and uncertainty analysis. Another source of uncertainty in predicting future nutrient loadings from agricultural catchments is the difficulty of properly representing the biogeochemical processes in future conditions as the growing season and land management practices change. More research is required in order to conduct a more reliable assessment of the climate
Results and Discussion

change effects on nutrient loading as well as an assessment of future abatement strategies to prevent water quality impairment.

Watershed resilience assessment based on the modeling output indicated different responses for nitrate and total phosphorus loadings (Article IV). Most of the studied sub-catchments showed low resilience to change in nitrate loads but high resilience to change in total phosphorus loading. The vulnerability index was high in all sub-catchments for both nutrients. The watershed resilience index for nitrate was positively correlated with the relative area of forestland and inversely correlated with the relative area of clayey soils in the sub-catchments. The Resilience index for total phosphorus loads was inversely correlated with the relative area of agricultural lands and clay soils of the sub-watersheds. This result implies the low resilience capacity of watersheds dominated by clay soils and agricultural land use, making them highly vulnerable to nutrient load changes. These areas might demand more effort as regards restoration. However, the estimation of the resilience index (resiliency-reliance-vulnerability index) is highly dependent on the accuracy of the nutrient loading estimation and the standard water quality threshold, where the system passes from a safe to a failure stage (Hoque et al. 2012). The methodology of the watershed resiliency assessment to biogeochemical change still needs more methodological development in order to provide consistent assessment tools (Hipsey et al. 2015).

The modeling assessment of the effect of restoring wet areas, which have potentially high biogeochemical value, showed a reduction of nutrient export from the catchment to downstream of 15 kg ha\(^{-1}\) yr\(^{-1}\) of nitrate, 2.4 kg ha\(^{-1}\) yr\(^{-1}\) of total phosphorous, and 0.4 ton ha\(^{-1}\) yr\(^{-1}\) of suspended sediment. This indicates the relative removal capacity of wet areas in the study area of 15% for nitrate loading, 35% for total phosphorus loading and 4% of suspended solid loading. Thus, by restoring wet areas as constructed wetlands contributes to an increase in watershed resiliency. However, empirical field studies report varying nutrient reduction efficiencies of restored wetlands, for example nitrogen reduction ranges from 11 to 280 kg M ha\(^{-1}\) yr\(^{-1}\) (Koskiaho et al. 2003) and total phosphorus reduction from 0-460 kg P ha\(^{-1}\) yr\(^{-1}\) (Koskiaho et al. 2003, Uusi-Kämppä et al. 2000). In my study area, the simulated values of nutrient removal were rather low compared to those found in field research. This indicates that the relative removal capacity in my study area can be even higher than simulated. Mapping potential wet areas to be restored with LiDAR DTM, as was carried out here, can identify areas with high nutrient removal capacity.
6. Conclusions

The following main conclusions are drawn from this work:

1) Landscape patterns affect water quality differently on different spatial scales. The strength of the relationships vary seasonally. Spatial variations of total N, nitrate and total P were mostly affected by the agricultural practices on the watershed level, but for nitrate the riparian zone was an important factor in the autumn. This implies that this zone functions as a potential sink for nitrate. Dissolved reactive P was also related to agricultural land use in the entire watershed, and additionally to other land covers in zones prone to water saturation, e.g., forests were also important in high flow periods. This indicates a potential contribution of dissolved reactive phosphorus from areas with high organic matter content, when they become hydrologically connected to water courses. This issue requires further investigation to identify and understand the underlying biophysical processes.

2) The dominant sources of suspended sediments appear to be linked to agricultural areas located in clay soils and saturation-excess zones, where high rates of runoff are produced. As fine clay sediments are chemically highly active and transport the pollutants bound in them, and fine sediment enrichment ratios occur during the soil erosion process, it is important to further develop strategies to control fine sediment erosion in agricultural lands. In the snowmelt period, sediment mobilization can occur rapidly as many snowmelt events show a clockwise hysteresis.

3) New emerging long-term trends of nutrient concentrations and loads were detected in a managed agricultural watershed, indicating a decrease in total suspended solids over time, but an increase in total N and dissolved reactive phosphorus in the clayey soil watershed. This phenomenon may be linked to the change in land management, however, no such cause-effect evaluation was carried out in this study. No similar trends were found for most of the water quality variables in the sand soil watershed. Therefore, it was found that there were spatially and temporally dissimilar responses in the water quality trends to similar land management in watersheds. This finding
Conclusions

suggests evaluating the side effect of the current water protection strategies in agricultural fields in particular.

4) The mapping of hydrologically and biogeochemically meaningful zones or hotspot areas, such as wet areas and riparian zones is greatly supported by high-resolution topographic data and geographic information system techniques. These allow a process-oriented spatial data analysis for hydrological and biogeochemical modeling to investigate factors affecting stream water quantity and quality, as well as the identification of potential areas for restoration.

5) The ecohydrological modeling results contained high degrees of uncertainty. Model evaluation is important not only for the model performance statistics but also for evaluating the prediction of the spatial patterns; this is particularly relevant to correctly identifying where the important hydrological and biogeochemical processes occur. Integration with empirical studies and soft calibration is an important step to reduce model predictive uncertainty. Defining a modeling unit based on more meaningful units of landscape hydrology and biogeochemistry improves model predictions.

6) An increasing nutrient export was predicted in climate change scenarios, which implies surface water resources and aquatic ecosystem impairment. Restoration of wet areas could reduce nutrient export and enhances watershed resilience, particularly in agricultural watersheds with clayey soils. However, a more accurate modeling of the biogeochemical process in hotspot areas is required.

7) Finally, the integration and complementing of different methods of data analysis and modeling is very important for watershed management research, since by its nature the environment is a complex system and environmental data is limited, and, in addition to which, none of the environmental data analysis methods and models are free of uncertainties. More consistent environmental information is very important in order to foster more efficient water protection measures and to the restoring and management of key ecosystem services in current and future climatic conditions; thus, leading to sustainable watershed management.
Acknowledgements

I am very grateful to my supervisors Dr. Risto Kalliola and Dr. Timo Huttula. Thank you very much, Risto, for supporting my project all the way since the beginning till the end, and reading and discussing on my manuscripts. Thank you also for giving me space to develop my own research ideas and respecting my independent work. Thank you, Timo, for encouraging me to carry out this research and introducing me researches in water related issues in Finland. Your course on physical limnology and lake hydrodynamic modeling was my favorite when I was a student at the University of Jyväskylä. Even though I have moved now a bit from the lake shore to the whole catchment, your teaching was fundamental for me.

I want to thank all my coauthors: Dr. Ahti Lepistö (Finnish Environmental Institute - SYKE), Dr. Teija Kirkkala (Pyhäjärvi Institute), Dr. Risto Uusitalo (Natural Resources Institute Finland - LUKE), Dr. Eila Hietaharju (Department of Geography and Geology of University of Turku), M.Sc. Pasi Valkama (The Water Protect. Assoc. of the River Vantaa & Helsinki Region), Dr. Jan-Olof Lill (Accelerator Lab., Turku PET Centre, Åbo Akademi University) and Dr. Joakim Slotte (Div. Physics, Åbo Akademi University). Your inputs were very valuable to succeed in my PhD research. Thank you, Teija, for introducing me the study area and sharing your experiences in watershed management including the fascinating trip to China.

Then, I would like to acknowledge the pre-examiners of my PhD dissertation, Dr. Katarina Kyllmar and Dr. Jukka Aroviita, for a thorough examination of my thesis and important comments to improve it. I also want to thank Prof. Jan Hjort for accepting to examine this dissertation.

Financial support to carry out this research was granted by Maj and Tor Nessling Foundation, Turku University Foundation, Maa- ja Vesitekniiikan Tuki Ry, The Emil Aaltonen Foundation, Department of Geography and Geology of University of Turku. Grants for travels and seminar attendances were also provided by VALUE - Doctoral Program in Integrated Catchment and Water Resources Management. All these funding sources are hereby acknowledged.

I also want to express my gratitude to all my colleagues and staff from the Department of Geography and Geology of University of Turku. It was very pleasant to share all these years with you. Special thanks go to Dr. Joni Mäki and Dr. Reija Hietala for inviting me to give lectures in your courses when I was a newcomer in the department. I really appreciated this and felt
Acknowledgements

welcome to the department. Thanks to Leena Laurila for kindly helping me with all complicated IT stuff, and all researchers of the fluvial group for sharing interesting seminars. Particular thanks to Claude Flener for helping in Yläneenjoki River surveying and to Jenni-Mari Vesakoski and Tua Nylen for reading and commenting the drafts of my manuscripts.

I would also like to thank all the friends I met during this project, Israel Gomez, Glenda Cardenas, Nelly Llerena, Liisa Puhakka and many others. It was very nice to go for lunch and coffees with you and talk about our cultural shocks, frustrations, and joys of course too. Thank you also for the researchers of Amazon Research Team for sharing meetings dealing with the topics close to my academic and personal interest. I also want to thank Dr. Anders Siren and Umer Alvi for a great company in sharing the office space and talking about everything. Thank you also to Ulla Helimo, Lassi Suominen and Johanna Toivonen for research cooperation and joint field work in Altomayo River basin in Peru. Even though this was not part of my PhD work, it was very interesting and enjoyable.

I reserved these lines to acknowledge my family for the immense support during these years. Thank you Andreé, Inka, Vilja and Selja, for being just you - so fantastic in all aspects. When I am with you, all work stresses that I’ve sometimes brought home blow away. Thank you, Johanna, for a huge support and encouragement during the years of my PhD, and before that. Without your support this book would never had come out. I’m very lucky to share academic and family life with you - I love you. I also want to express my grand gratitude to my extended family, my grandmother Narcisa, my aunts Nicolasa, Nery, Marcusa, Leonor, uncle Bernardino and cousins, for all your support in whatever I did - you are just incredible. I also extend my thanks to Leena and all the Toivonen family. You have really made me feel like home in Finland.
References


Guo Q.H., Ma K.M., Yang L. & He K. 2010. Testing a dynamic complex hypothesis in the analysis of References
References


estimation methodology, sampling frequency, baseflow index and catchment population density. J. Hydrol. 332: 241–258.


References


Sanford S.E., Creed I.F. Tague C.L., Beall F.D. & Buttle J.M. 2007. Scale-dependence of natural


