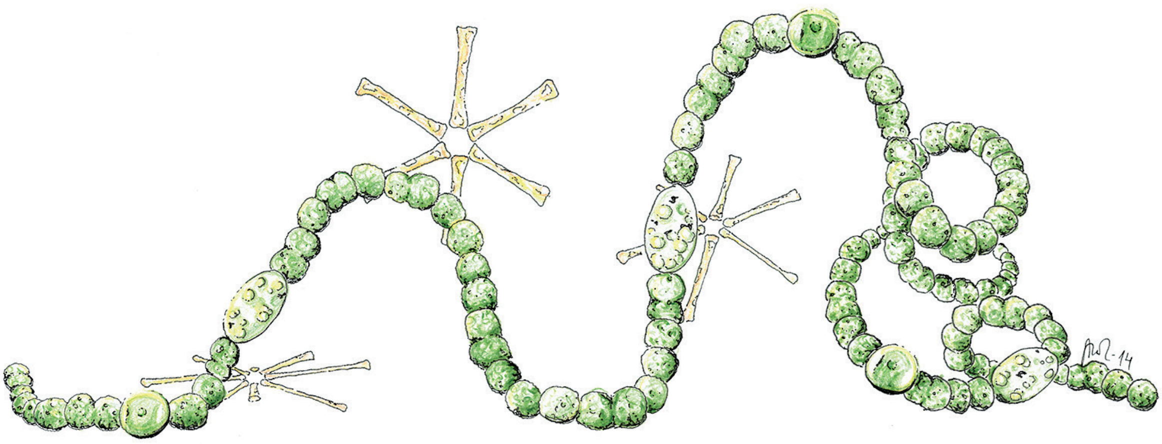
JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE
282

Anita Pätynen

Modelling Phytoplankton in Boreal Lakes



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 282

Anita Pätynen

Modelling Phytoplankton in Boreal Lakes

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella julkisesti tarkastettavaksi yliopiston Ambiotica-rakennuksen salissa YAA303, kesäkuun 6. päivänä 2014 kello 12.

Academic dissertation to be publicly discussed, by permission of the Faculty of Mathematics and Science of the University of Jyväskylä, in building Ambiotica, hall YAA303 on June 6, 2014 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

Modelling Phytoplankton in Boreal Lakes

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 282

Anita Pätynen

Modelling Phytoplankton in Boreal Lakes

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

Editors
Timo Marjomäki
Department of Biological and Environmental Science, University of Jyväskylä
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

Cover picture by Anita Pätynen.

URN:ISBN:978-951-39-5701-8
ISBN 978-951-39-5701-8 (PDF)

ISBN 978-951-39-5700-1 (nid.)
ISSN 1456-9701

Jyväskylä University Printing House, Jyväskylä 2014

ABSTRACT

Pätynen, Anita
Modelling phytoplankton in boreal lakes
Jyväskylä: University of Jyväskylä, 2014, 46 p.
(Jyväskylä Studies in Biological and Environmental Science
ISSN 1456-9701; 282)
ISBN 978-951-39-5700-1 (nid.)
ISBN 978-951-39-5701-8 (PDF)
Yhteenveto: Mallien hyödyntäminen boreaalisten järvien kasviplanktonin runsauteen ja kasviplanktonyhteisön dynamiikkaan vaikuttavien tekijöiden tarkastelussa
Diss.

Three different approaches to modelling phytoplankton production and dynamics in Finnish lakes were evaluated. The first was deterministic model, PROTECH, which simulated the growth of 8 different phytoplankton species at a daily time step in a large, shallow lake. The model was used to estimate how higher temperatures might promote the growth of cyanobacteria and alter the phytoplankton dynamics in the lake. Insufficient data introduced uncertainty to the model outputs, but could not be quantitatively estimated. The lack of routines in the model for special features of boreal lakes was also limiting. These issues are restricting for the use of such complex models. Meanwhile lake management needs modelling techniques that can be applied with the often restricted data. For this, two statistical methods, structural equation modelling (SEM) and linear mixed effects (hierarchical) modelling (LMM) were examined, and their potential advantages and earlier applications were reviewed. In the second modelling study SEM was tested with data from a small humic lake. SEM proved to be effective for examining causal relationships between phytoplankton and some basic variables: nutrients, grazing, temperature and water colour. Yet more detailed study questions were restricted by the available data. The inclusion of Bayesian analysis improved the model and allowed examination of the underlying uncertainties. Bayesian analysis was also applied in the third modelling study with LMM. With the hierarchical approach the monitoring data from over 2000 Finnish lakes was more efficiently utilised to estimate the effect of total phosphorus and total nitrogen on chlorophyll *a* concentrations in a single lake. The hierarchical chlorophyll *a* model was used as a basis for a simple model tool developed to support the lake management. Explanatory modelling studies, more frequent data and data analysis are needed to enhance the understanding of phytoplankton development in boreal lakes and for further model development. The purpose of modelling determines the desirable approach, thus different methods need to be critically evaluated.

Keywords: Boreal lakes; ecological modelling; phytoplankton; predictions; uncertainty; water quality.

Anita Pätynen, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Author's address Anita Pätynen
Department of Biological and Environmental Science
P.O. Box 35
40014 University of Jyväskylä
Finland
anita.i.patynen@jyu.fi

Supervisors Prof. Roger I. Jones
Department of Biological and Environmental Science
P.O. Box 35
40014 University of Jyväskylä
Finland

Research Manager, Adjunct Prof. Timo Huttula
Finnish Environment Institute
Survontie 9A
40500 Jyväskylä
Finland

Dr Olli Malve
Finnish Environment Institute
P.O Box 140
00251 Helsinki
Finland

Reviewers Research Prof., PhD Tom Frisk
Centre for Economic Development, Transport
and the Environment for Pirkanmaa
P.O. Box 297
33101 Tampere
Finland

Prof. Frank Peeters
University of Konstanz, Limnological Institute
78464 Konstanz
Germany

Opponent Prof. Wolf Mooij
Netherlands Institute of Ecology
P.O. Box 50
6700 AB Wageningen
The Netherlands

CONTENTS

ABSTRACT
LIST OF ORIGINAL PUBLICATIONS

1 INTRODUCTION 7
  1.1 Different approaches to modelling phytoplankton 10
  1.2 Phytoplankton-related issues in lakes 13
  1.3 Models in public policy making and lake management 14
  1.4 Aims of the study 16

2 MATERIALS AND METHODS 18
  2.1 Study lakes 18
    2.1.1 Lake Pyhäjärvi 18
    2.1.2 Lake Valkea-Kotinen 19
    2.1.3 Lake Kuortaneenjärvi 19
  2.2 Modelling 20
    2.2.1 PROTECH (I) 20
    2.2.2 Structural Equation Modelling (II, III) 23
    2.2.3 Linear Mixed Effects Modelling (II, IV) 25

3 RESULTS AND DISCUSSION 27
  3.1 Effect of higher temperature on the phytoplankton community in Lake Pyhäjärvi 27
  3.2 Factors affecting the development of phytoplankton in Lake Valkea-Kotinen 29
  3.3 Estimates of target nutrient loading for Lake Kuortaneenjärvi 30

4 CONCLUSIONS 33

Acknowledgements 35

YHTEENVETO (RÉSUMÉ IN FINNISH) 36

REFERENCES 39

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original articles, which will be referred to in the text by their Roman numerals I-IV.

I compiled and prepared the data for modelling in I and III and also performed the literature search and review in II. Anne-Mari Ventelä (I) and Tiina Tulonen (II) represented the institutions responsible for the actual sampling of the data. Alex Elliott was responsible for setting up the PROTECH model (I) and for introducing me to the application, after which I continued with the application. Petri Kiuru was responsible for the MyLake simulations and their documentation (I). I did most of the modelling in III, with statistical and methodological support from Niina Kotamäki. Ms Kotamäki and Olli Malve also checked the accuracy of statistical aspects in II. In IV, I examined the ecological aspects of the modelling, interpreted the outputs and helped in the development of the model tool. Niina Kotamäki and Olli Malve were responsible for the modelling *per se*. I was responsible for the first draft for I–III and in collaboration with Niina Kotamäki for IV. Especially Alex Elliott, Jouko Sarvala, Roger Jones, Lauri Arvola, Niina Kotamäki and Olli Malve made significant contributions as co-authors in planning and improving the articles they were involved in.

I Pätynen A., Elliott J.A., Kiuru P., Sarvala J., Ventelä A.-M. & Jones R.I. 2014. Modelling the impact of higher temperature on the phytoplankton of a boreal lake. *Boreal Environment Research* 19: 66–78.

II Pätynen A. Kotamäki N. & Malve O. 2013. Alternative approaches to modelling lake ecosystems. *Freshwater Reviews* 6: 63–74.

III Pätynen A., Kotamäki N., Arvola L., Tulonen T. & Malve O. 2014. Causal analysis of phytoplankton development in a small humic lake using structural equation modelling. Submitted manuscript.

IV Kotamäki N., Pätynen A., Taskinen A., Huttula T. & Malve O. 2014. Enhancing lake management with model based estimates of target nutrient load reduction. Manuscript.

1 INTRODUCTION

Perhaps due to the strong physical and mathematical background of modelling, and due to its pervasive presence, ecological modelling is often also viewed as a mathematical discipline. Modelling is seen as an efficient way to predict the environment and aid cost-effective management. A non-modeller utilising the model outputs may find it difficult to comprehend the reasons why this often may not be the case, as the underlying problems are sometimes overlooked by modellers themselves (Jakeman *et al*. 2006). One problem is that the often random, unpredictable and constantly evolving biological systems do not behave as systematically as the functions trying to describe them. Besides gradual succession, nonlinearity and sometimes very drastic, unforeseen shifts are common features within ecosystems (Scheffer and Carpenter 2003). In addition, the openness of these systems makes it challenging to cover all the important variables and their interactions for realistic description of the system function. Thus modelling biological systems is very much reliant on adequate and continuous data from the systems so that the mathematical equations derived are biologically robust.

Usually the main aim in ecological modelling is to predict the growth and distribution of some organisms, populations or communities, which is largely reliant on the various interactions between species and on their abiotic environment. In ecological lake modelling, predicting harmful phytoplankton blooms has been a major international focus (Robson 2014), and when examining the function of aquatic ecosystems in general, phytoplankton as primary producers plays a key role. In Finland water quality models, which primarily focus on phosphorus balances and concentrations, have been in intensive use since the late 1970s (aquatic model efforts in Finland are summarized in Virtanen 2009). Inclusion of chlorophyll *a* and phytoplankton biomass naturally followed and applications of some first ecological lake models date back to the late 1970s (e.g. Niemi 1979). The dynamics of phosphorus and phytoplankton and the eutrophication of Lake Kuortaneenjärvi (IV) were already examined with models in the 1980s (Kettunen *et al*. 1987, 1989). Nevertheless, model applications for phytoplankton have been rather sporadic and have not led to consistent development of sophisticated techniques which could be more widely applied. In addition, the description of biological processes has remained rather simple. However, as ecological issues have gained increasing attention in national and international legislation and policy making, the interest in model development and more efficient utilization of models to examine ecological questions has also risen. For some of these questions more complex models are needed. Some straightforward biological models can be derived from the results of laboratory measurements. In a controlled environment very detailed questions about reproduction rates, optimal growth conditions or sinking velocities of single phytoplankton species can be studied. Yet, when several species are put together into enclosures the species interactions, chance and even chaotic fluctuations have been reported to make the final population composition unpredictable within 15–30 days (Benincà *et al*. 2008). Acknowledging this aspect, modelling the development of phytoplankton in lakes (dynamic, open systems which are closely connected with the surrounding catchment and air) becomes challenging. The number of variables and their interactions multiply, so growth curves determined *in vitro* do not hold in open water, yet are difficult to determine *in situ* (Reynolds 2006). With phytoplankton the unsystematic and unpredictable behaviour increases when the productivity of the system increases (Soininen *et al*. 2005), and yet it is particularly this state of increasing productivity, eutrophication and its consequences, that raises the interest in better understanding of aquatic ecosystems.

Despite the fundamental challenges with phytoplankton modelling, models do serve a way to understand better the function of ecosystems and organisms by elucidating consistent patterns from observations and experimental datasets. More precisely, models can help to guide intuition about how various processes interact, they can highlight logical flaws in an argument and they can identify testable hypotheses, generate key predictions and suggest appropriate experiments. They can also be viewed as an experimental system and reshape fields by providing new ways of thinking about a problem (Peck 2004, Otto and Day 2007). On the other hand, an ability to make predictions with ecological models could increase the credibility of the whole field of ecology (Sutherland 2006). Eventually, if some clear patterns can be detected from the data, like the succession of phytoplankton during a growing season, models can offer a possibility for predictions. However, it should be stressed that ecological models are at root heuristic, not literal (Oreskes *et al*. 1994, Otto and Day 2007).

Depending on the objective of modelling and the study questions, various techniques can be applied with varying complexity and different space and time scales. A model can be based on the simple linear relationship between chlorophyll *a* and total phosphorus concentration (Vollenweider 1968) or it may try to cover the whole biogeochemical cycle within a lake (for instance CAEDYM, see Romero *et al*. 2004). However, there is always a trade-off between the realism (complexity) and the generality (simplicity) of a model (Clark 2005). The requisite complexity has been debated for many years (e.g.

Scheffer and Beets 1994, Flynn 2003, Gudimov *et al*. 2012). Phytoplankton modelling is usually performed with a 1D application due to its lesser computational requirements and because data are usually gathered from a single point of the lake. However, utilizing remote sensing data to examine the spatial distribution of phytoplankton (chlorophyll *a*) is increasing, accompanied by 3D modelling approaches for larger water areas. Biological compartments are designed and combined into hydrodynamic 3D models (see e.g. Delft3D-Wag application, Chen and Mynett 2006). If the modeller is not content with a static description of the system, the dynamics of the system can be simulated using a time step of one day, or even less if the computational time and power is not restricting, or the description can be averaged over longer periods like one year.

Before constructing a model or starting to use an existing model, careful consideration is needed to determine the appropriate modelling method, complexity of the model and the scale of modelling for the question of interest and for the data that exist or can be collected (see Johnson and Omland 2004, Otto and Day 2007, Robson *et al*. 2008). The model itself may not provide any information whether choices are valid, or whether the outputs are valid in real life (often described by 'rubbish in rubbish out'). Also the impact of biological randomness on the model outputs and on the uncertainty of model predictions need to be considered, especially when examining organisms like phytoplankton which, with zooplankton and bacteria, have been shown to be rather unpredictable variables in the models (Arhonditsis and Brett 2004, Benincà *et al*. 2008). The uncertainty derives from many sources: from the measurements made in uncontrolled environments, from the model structure and parameters, and also from technical problems in the modelling (Rode *et al.* 2010). There may also be problems in the phrasing of the study questions and determining the context of the model (Refsgaard *et al*. 2007). In addition, when many processes are combined in a model, the small flaws in each equation may accumulate the error and uncertainty the further the modelling proceeds. Particularly, the more detailed the interactions to be studied, the more complex the model eventually becomes making proper uncertainty analysis difficult to perform (Doherty and Christensen 2011). Put another way; the validity of the model, especially in a new environment, cannot be confirmed (Rykiel 1996). Thus an ideal phytoplankton model is not necessarily one that tries to capture every possible factor impacting on the phytoplankton development (which is anyway impossible). As important are models that are as simple as possible, but that catch the most important processes and hold some level of generalization. Such models can be better applied to new lakes, modified for the study questions and combined with methods providing uncertainty analysis. On the other hand, due to that simplicity the uncertainty is always present, so there is a danger that some important processes on which the uncertainty may depend are omitted from the model (Doherty and Christensen 2011).

The main abiotic components for phytoplankton growth are nutrients, light and temperature; of which nutrients promote eutrophication when in abundance. These abiotic factors directly affect phytoplankton growth and are most often included into the models. In addition, currents and other movement of water, or alternatively the stability of water layers, affect the establishment of denser phytoplankton communities, phytoplankton dispersal and the species composition. Death of cells, their sinking below the euphotic zone, and also grazing are factors removing phytoplankton from the water column (all summarised in Reynolds 2006). These factors are already much harder to include into the models, and the biotic factors and their interactions are especially difficult to measure. The grazers (aka the zooplankton) together with bacteria also enhance recycling of nutrients in the water, which is significant for phytoplankton growth (Elser and Urabe 1999). It is known from many biomanipulation studies, that fish and their control over zooplankton greatly impacts on the phytoplankton, while the sediment underlies many features of the nutrient cycle. Thus, the complexity of the system ever increases and, depending on the study questions and the available data, it is reasonable to curtail the inclusion of variables at some point yet carefully consider what information is then left out. This is especially true when models are intended for use in public policy making and in lake management tackling water quality problems. With the limited resources of many management processes or tight time schedules for decisions, the data from a single lake will inevitably be limited if the lake has not been within a comprehensive monitoring program. That makes utilising complex, data-hungry models impossible; but then again too general models may be uninformative and equally useless.

1.1 Different approaches to modelling phytoplankton

Deterministic models assume that the fate of the system can be entirely predicted. As the effect of different variables on the organisms is determined (for instance models built on differential equations), the chain of interactions leads to one outcome at a time or to one specific state of the system at any given time (Otto and Day 2007). Deterministic models are thus usually dynamic (c.f. static), describing how the system changes over time and how different forces affect it or the organisms in it. This makes them very appealing for biological studies as insights can be gained into the function of the system—and perhaps how to change it. A thorough introduction to different kinds of deterministic models with their pros and cons is provided by Jørgensen (2008). As only one kind of deterministic model was used in this thesis, the discussion is more general.

The many process-based biogeochemical models including phytoplankton are deterministic (see Trolle *et al*. 2012), although numerous others with varying complexity exist. Besides being able to represent the internal dynamics and function of the system, other advantages of deterministic modelling is the possibility for quantitative simulations and, with caution, making predictions beyond the historical variability of the system (Robson *et al*. 2008). This assumes that the model is valid for the system, numerically stable and that the studied environment does not go through drastic changes (Rykiel 1996, Clark *et al*. 2001, Sutherland 2006, Jørgensen 2008). Unfortunately even many of the most ambitious model equations are not able to capture the biological processes, simply because the ecology of phytoplankton species and the system interactions are not (yet) known thoroughly or observed adequately (Anderson 2005). Thus using the models for future predictions or their ability to describe the function of new systems in other lakes becomes questionable (discussed by Oreskes *et al*. 1994, Rykiel 1996, Clark *et al*. 2001, Arhonditsis and Brett 2004). In addition, biogeochemical models are typically embedded within hydrodynamic models, which provide adequate description of the physical environment the organisms live in. However, if the original purpose of those models has not been to describe ecological processes, it sometimes leads to oversimplification of biology and also of the biogeochemical processes. It may be purely a technical issue, as the time scales of different processes may drastically differ, and sometimes the simulation of physics is already so involved that compromises with other processes are needed; only some general processes are included or the large biological units like fish may be simplified as concentrations of phosphorus (Robson 2014).

One of the most criticized issues with deterministic models is the use of parameters that cannot be unambiguously determined (Gudimov *et al*. 2012) and the way models are calibrated by tuning these parameters until the model simulations fit the observations (Scheffer and Beets 1994, Rykiel 1996). This is not considered correct, because different sets of parameters can yield a similar outcome. In addition, if parameters are merely selected for their ability to fit the model to the observations it does not guarantee the model is able to describe the system function correctly or that the selected parameter values are reasonable for the system (Oreskes *et al*. 1994). Other possible disadvantages include high data requirements, high complexity, high computational costs and difficulty of quantitatively estimating the prediction uncertainty (Robson *et al*. 2008, Jørgensen 2008). However, these issues are acknowledged by many modellers and the importance of joint-development of the (strongest) existing mechanistic biogeochemical models is stressed (Jørgensen 2010, Mooij *et al*. 2010, Trolle *et al*. 2012).

Unlike deterministic models, statistical models provide outputs as probabilities, and thus readily take into account the unpredictability which affects the interactions within a biological system. Also, instead of one there can be several different outcomes, of which some are more likely than others. A common problem with statistical models is their subjectivity, and the possibility (unintentionally) 'to lie with statistics', because of the freedom to choose what to include in the model. Also, statistical significance can be gained without the model having biological meaningfulness, which needs to be evaluated separately (Johnson and Omland 2004) Nevertheless, the uncertainty in model results and model error are fairly easy to take into account and quantify in statistical modelling, especially when using Bayesian techniques (Clark 2005). However, the benefits of Bayesian posterior predictive inference and Marcov Chain Monte Carlo (MCMC) methods in water quality predictions and decision making has only recently been acknowledged (Malve 2007). Statistical methods have in general received surprisingly little attention in ecological, and especially in aquatic modelling (Arhonditsis *et al*. 2006, Kruk and Segura 2012). Of course that applies only to what is thought of as 'actual modelling', because statistical methods are most widely used by ecologist for analysing observational data. However, they are also widely misused. Many of the 'basic' methods have limitations and assumptions, including normality and homoscedasticity, which are often not met with ecological observation data, and often ignored (Whittingham *et al*. 2006, Zuur *et al*. 2009). These same assumptions put some limits on the use of statistical methods in ecological modelling, for example when simple regression models are drawn. The assumptions can sometimes be met by transforming the data (e.g. using natural logarithms). However, instead of transforming the data to meet the assumptions, it would be preferable to select some more appropriate method (Bolker *et al*. 2009), which often means methods that are less familiar and perhaps more difficult to adopt.

The increase in computing power and development of statistical software has only fairly recently brought some such methods, like structural ecological modelling and hierarchical modelling as well as the Bayesian analysis, closer to everyday use (Hershberger *et al*. 2003, Zuur *et al*. 2010). Perhaps because of this, in some instances statistics is considered only as a tool that provides better process description for deterministic ecological modelling. 'If a model is based entirely on statistics, it is a so-called black box model, because it has no causality. Black box models can hardly be considered ecological models, because they cannot be used to uncover new ecological knowledge' (Jørgensen 2008). However, the onset of the work of Judea Pearl with probabilistic causalities, structural equations and Bayesian networks date back decades, and his article about causality in empirical studies was published in Biometrica in 1995 (Pearl 1995). Perhaps 'entirely statistical' models can be understood in different ways, as Jørgensen (2008) briefly acknowledged structural equations as stochastic methods, but the stain on statistical modelling is unfortunate. Structural Equation Modelling and Bayesian inference have also been used in this thesis (III, IV), and there are earlier aquatic applications for Structural Equation Modelling that study interdependencies between phytoplankton and some biotic and abiotic factors (e.g. Arhonditsis *et al*. 2006, Arhonditsis *et al*. 2007a, b). Arguably, the use of statistical methods in most cases aims to 'uncover new ecological knowledge', even if they are not considered to be ecological models. Even though black box models cannot be used for this purpose, in lake management planning they can still be helpful model tools (IV).

The strong juxtaposition between deterministic and statistical methods could be set aside to consider how the best features of both methods could be utilized. In addition, whatever the technique, it is often forgotten that modelling should be an iterative process. After the proper modelling technique is agreed and the model built with existing data, its performance is tested with experiments and/or new data. The data are then used to improve the model further, and the cycle is repeated until the model provides adequate estimations (e.g. Jakeman *et al*. 2006).

1.2 Phytoplankton-related issues in lakes

The strongest motivation for lake model development in Finland arises from the Water Framework Directive (WFD, European Parliament and Council 2000) and securing good water quality status for lakes, but gaining better understanding of boreal lake ecosystems with strong seasonality is also important. Eutrophication and phytoplankton-related problems are a major factor leading to moderate or bad water quality. However, phytoplankton can also cause problems for other organisms and the whole lake ecosystem, or cause a nuisance to people. Most visible problems are due to cyanobacterial blooms of varying severity, but mass occurrences of other groups like Chrysophytes, Diatoms and the Raphidophyte *Gonyostomum semen* are also problematic. In freshwaters the species producing toxins are mostly cyanobacteria (Landsberg 2010). Toxins are harmful for people and animals, and phytoplankton blooms in general can cause aesthetic and odour problems that hinder the recreational value of lakes and the usability of the water (Paterson *et al*. 2004, Codd *et al*. 2005, White *et al*. 2005). Toxic cyanobacteria are the most widely acknowledged group causing harmful algal blooms (HABs), but other metabolic by products of phytoplankton, structures such as spines, shading effect of dense blooms and oxygen depletion due to the decomposition of the bloom scum, can also affect the health and growth of other organisms (Landsberg 2010). Of the other nuisances caused to people, diatoms are responsible for the fouling of fishing gear in cooler waters (Vuorio *et al*. 2013), while *G. semen* can cause a slimy, irritating coating on swimmers' skin (Lepistö *et al*. 1994).

Cyanobacteria have been shown to thrive in eutrophic lakes, utilising their nitrogen-fixing capability if the phosphorus:nitrogen ratio grows unfavourable due to excessive phosphorus supply. They further benefit from the temperature rise due to higher growth optima compared to other species. Also the absence of mixing forces such as wind is favourable for some cyanobacteria, because their gas vacuoles provide buoyancy (Reynolds 2006, Paerl and Huisman 2008, Wagner and Adrian 2009). As mentioned, diatoms may cause problems especially when the waters are cold. That is due to their high growth rates at low temperatures and the absence of effective grazing (Shatwell *et al*. 2008), but also because the cold waters are usually thoroughly mixed and this, together with the higher viscosity of water, is beneficial for the relatively heavy diatoms (Reynolds 2006). *G. semen* is an especially interesting species in boreal lakes, as it often occurs in lakes with high water colour and low pH, both being typical of lakes with peatland and coniferous forested catchments and high concentration of dissolved organic carbon. In addition, it may benefit from warm spring temperatures although its growth optima otherwise is under 19 °C (Trigal *et al*. 2013).

Consequently all of those species may benefit from climate change in one way or another. Cyanobacteria are expected to benefit from the ever higher occurrence of summer heat waves (Jöhnk *et al*. 2008, Paerl and Huisman 2008). Milder winters and a shorter ice cover period (Magnuson *et al*. 2000) provide opportunities for longer lasting and more pronounced diatom blooms, especially when mild winters lead to colder lake waters in spring (Keller 2007, Shatwell *et al*. 2008). The brownification of boreal lakes, that has also been predicted as one possible outcome of climate change (Naden *et al*. 2010), may benefit the growth of *G. semen*. Climate change, or any other change in the environment, may have unexpected consequences as well, and any hints of the possible outcomes would be valuable. For instance, an increase in the diatom spring bloom can alter the nutrient availability in a lake for the rest of the summer, or some new species may emerge when conditions become favourable. Surely, modelling these phenomena is a task, where the study questions need to determine the data to be gathered and the model to be used, rather than the opposite.

The principal requirement for excessive phytoplankton growth and eutrophication is the high availability of phosphorus and nitrogen (Mason 2002). Lake management and controlling the eutrophication process does not necessarily have to (and often cannot) focus on the detailed interactions of species, but on reducing the load of phosphorus and nitrogen into lakes. Many restoration methods like fish removal, dredging and altering the hydrological properties can, however, be used to support the basic management. Because the external load in developed countries increasingly originates from diffuse sources, mostly from agriculture that involves large land areas, the task of reducing it is demanding (Carpenter *et al*. 1998, Antikainen *et al*. 2008). The study questions also turn from purely ecological towards economic and social. Models can help to evaluate the tolerable amount of external loading and the possible changes in the lake ecosystem and phytoplankton community that follow reduced nutrient concentrations. Further, the (cost-) effectiveness of different management actions, how they should be targeted, and also the risk of failure can be evaluated, with speculation as to how other changes in the environment may interfere with or benefit the objectives.

1.3 Models in public policy making and lake management

Simplicity, costs, accuracy and the possibility to evaluate the reliability of model outputs are some basic features that determine whether a model is sufficient to be used in public policy making (Caminiti 2004, Malve 2007). In an ideal situation, it could be possible to benefit from the good properties of different kinds of models, by comparing the outputs and eventually aiming to some kind of model with medium complexity (discussed also in Doherty and Christensen 2011). Unfortunately that is almost never (financially) possible. The most inappropriate way for utilizing models in policy making and lake management would probably include applications with complex phytoplankton models based on sparse data sets covering only part of the year (Flynn 2003) and with too long sampling intervals (Lawson *et al*. 1995). Equally, decisions based solely on (linear) regression equations introduced in the literature or drawn from few in-lake observations are uninformative and misleading.

Deterministic models and statistical methods providing single values (accompanied by p values) have been long favoured, because modelling results presented as probability distributions are more difficult to adopt than simple values. Terms relating to uncertainty, such as error or risk can also convey negative connotations to the public (Mowrer 2000). Still, policy-makers have as a general concern how to assess the reliability of scientific information, including in peer reviewed literature (Holmes and Clark 2008). This concern is justified, especially when models are used for open systems and/or for long-term predictions. Criticism of the ways models are utilized as a basis for public policy decision is also raised in more scientific debate (Oreskes *et al*. 1994, Malve 2007, Ramin *et al*. 2012). Presenting only single values as model outputs creates the false assumption that they can be considered as exact (Mowrer 2000, Sutherland 2006). This easily leads to poor decisions and ineffective management actions. Instead of leaving out significant information and ignoring the stochastic features of nature, the communication between scientists and policy-makers should be enhanced, for instance with the help of interpreters (Holmes and Clark 2008). As a whole, translating scientific knowledge into reliable models of aquatic ecological processes has not been as successful as desired. Moreover, the use of scientific knowledge in environmental policy-making and regulation has been inefficient, the problems varying from establishment of relevant research questions to proper communication between scientists and policymakers, and agreeing, for instance, on the aims of modelling (Caminiti 2004, Holmes and Clark 2008). The risk of misuse of models increases when they are proposed as tools for lake managers, who do not have the necessary background in modelling for critical review of the modelling process (Jakeman *et al*. 2006).

Adaptive management (Holling 1978) offers a means to combine models into a wider concept, and hence emphasises the iterative modelling process that is an important aspect in reducing the uncertainty and advancing the whole modelling process. As described by Sutherland (2006) the steps in adaptive management are:

1. Using available information to create models, incorporating the uncertainty in both biological understanding and parameter estimates.
2. Determining where greater certainty would lead to improved management.
3. Performing management experiments that lead to better understanding and even better management of the system.

4. Improving models with information gained and re-evaluating the management practices and the need for further experiments.

The process takes time and, if planned too ambitiously, it easily becomes too expensive and laborious to carry out. But as also pointed out (Sutherland 2006), sometimes evaluating the effectiveness of activities performed routinely, and targeting the effort on simple but informative ones can be one step forward (e.g. before-and-after comparisons or a single treatment and control). The scientific theory needs to form a basis for model development, but it should be considered whether utilizing models in lake management needs to be strictly a scientific process. Besides science and practise not always being readily compatible, there are also many benefits from relaxing the standards. For instance, when the main aim is to tackle eutrophication citizen monitoring could be more efficiently utilized as a way to gain more data and information about lakes. That would also offer a way to collect time-series that, besides from modelling, offer means to track changes in the environment.

1.4 Aims of the study

In this thesis three different modelling techniques were adopted to examine their suitability for studying phytoplankton dynamics in Finnish lakes and for tackling the problems linked to extensive phytoplankton production and deterioration of lake water quality. For instance the number of Finnish lakes not fulfilling the good ecological quality requirements of the Water Framework Directive is over 700 (HERTTA database of Finnish Environmental Administration), which calls for some efficient modelling techniques to support the management work. The biggest problem is that majority of the lakes are monitored only a few times a year, which not only creates a need for model utilization, but makes it challenging. There is also a need to gain better insight into the various processes that independently or together affect the development of phytoplankton communities in boreal lakes and to estimate the likely impact of ongoing changes in the climate and land use on phytoplankton.

One established detailed phytoplankton model, PROTECH, was applied to a large, low-humic Finnish lake to evaluate its performance (I). The aim was to estimate the impact of higher water temperatures, expected as an outcome of climate change, on the phytoplankton community and especially on the cyanobacterial abundance. A shift in the diatom spring bloom was also expected if the ice-break in spring took place earlier. The PROTECH application was promising, but many of the problems with deterministic models became concrete, and the inadequate data from the study lake was problematic for deeper evaluation. Because of this other potential approaches with lesser data requirements were subsequently investigated, and finally structural equation modelling (SEM) and linear mixed effects (or hierarchical) modelling (LMM) were selected as alternative methods. To gain better insight into how they could be applied in phytoplankton studies, a review of earlier studies was made (II). To test SEM, it was then applied to data from a typical small boreal forest lake and the factors affecting phytoplankton development as well as the simple causal network of the selected variables was examined (III). The large monitoring dataset from Finnish lakes was utilized to improve the hierarchical model of Malve (2007) for phosphorus/nitrogen-chlorophyll *a* relationships (IV). The hierarchical chlorophyll *a* model was then used as a basis for the LLR-model tool together with a simple mass balance equation. The final aim was to create an easy-to-use tool for lake management that provides estimates of the tolerable nutrient loading levels for good water quality and, due to its Bayesian approach, provides sufficient estimates of the model error and of the uncertainties in the model outputs. Application of the tool was demonstrated with a case study of a medium-sized highly humic boreal lake.

2 MATERIALS AND METHODS

2.1 Study lakes

The three study lakes, Lake Pyhäjärvi (I), Lake Valkea-Kotinen (III) and Lake Kuortaneenjärvi (IV), were selected mainly because of the existing data available for each of them and because of some earlier studies providing information to support the interpretation of model outputs. However, after the PROTECH model application for Lake Pyhäjärvi, Lake Valkea-Kotinen was chosen as the next study lake for the Structural Equation Modelling application, because it better represents the characteristics of a boreal lake (Table 1). For instance, cyanobacteria have been shown to have a significantly weaker response to eutrophication in humic than in clearwater lakes (Ptacnik *et al*. 2008). Besides the studies of individual lakes, constructing the hierarchical chlorophyll *a* model applied to Lake Kuortaneenjärvi utilised the available observational data from 2246 Finnish lakes, retrieved from the HERTTA database of the Finnish Environmental Administration.

2.1.1 Lake Pyhäjärvi

Lake Pyhäjärvi is the largest lake in southwest Finland. Despite its size, this lake does not develop persistent summer stratification, because wind is able to mix the water column quite easily. The drainage area of Lake Pyhäjärvi is relatively small (Table1) and about half of it is forest. One fifth is in intensive agricultural use. The Rivers Yläneenjoki and Pyhäjoki are the main inflows to Lake Pyhäjärvi and the outflow is via the River Eurajoki. Municipal waste waters were discharged into the lake for a short period during the 1960s but the water quality remained good through the 1970s. Increasing phosphorus levels since the early 1980s, and the more prominent role of cyanobacteria in the 1990s (Ventelä *et al*. 2011), aroused concerns about the gradual eutrophication of the lake. An intensive restoration program started in 1995 when the Pyhäjärvi Protection Fund was created. More recently, actions have continued to maintain the good condition of Lake Pyhäjärvi (Ventelä *et al*. 2007). These include implementing protection measures like wetlands and filtering systems in the catchment area, and biomanipulation through fish removal from the lake.

2.1.2 Lake Valkea-Kotinen

Lake Valkea-Kotinen is a small headwater lake. There is no distinct inflow to the lake and the outflow is through a small stream. The lake is surrounded by forested catchment and can be considered as a reference site due to low anthropogenic influence. However, the organic carbon load from the catchment is high and gives the lake a noticeably brown colour. Because of its small size and sheltered position, the lake is dimictic and produces a steep thermal and oxygen stratification in the summer, with a 1.5–2 m thick epilimnion. The depth of the euphotic zone is approximately the same (Peltomaa and Ojala 2010). The lake and its catchment have been studied intensively since 1990 as a part of the International Co-operative Programme on Integrated Monitoring of Air Pollution Effects on Ecosystems (UNECE ICP IM) and also as a part of the Finnish Long-Term Socio-Ecological Research network (FinLTSER). In 1994 the Valkea-Kotinen region was designated as a nature reserve, as it has been more or less in its natural state from the early years of twentieth century (Vuorenmaa *et al*. 2011). Hence the lake is also an important reference site for climate change related studies.

2.1.3 Lake Kuortaneenjärvi

Lake Kuortaneenjärvi is a medium-sized lake in Western-Finland. The catchment area of the lake consists mostly of peatland and forest and the lake is highly humic. The River Lapuanjoki runs through the lake, and because the whole River Lapuanjoki catchment experiences heavy nutrient loading from agriculture and forestry it affects the condition of Lake Kuortaneenjärvi. The lake acts as a natural sedimentation basin within the catchment, enhancing the water quality of the lower parts, whereas the lake itself has become hypereutrophic, and is classified as only in moderate ecological condition. Periodic oxygen depletion causes phosphorus release from the sediment and cyanobacterial blooms are common (Väisänen 2013). However, the situation has improved markedly since the 1980s because of loading reductions and fish removals (Rautio and Aaltonen 2006).

TABLE 1 Some characteristics of Lake Pyhäjärvi, Lake Valkea-Kotinen and Lake Kuortaneenjärvi. Colour and concentrations of total phosphorus, total nitrogen and chlorophyll *a* are averages of the observations from the main basins in 1990–2010 (0–5 m), taken from the HERTTA database of Finnish Environmental Administration.

	Pyhäjärvi	Valkea-Kotinen	Kuortaneenjärvi
Lake type	Large low humic lakes	Very humic lakes	Very humic lakes
Coordinates	61° 00.486' N 22° 17.973' E	61° 14.535' N 25° 03.775' E	62° 49.499' N 23° 28.570' E
Drainage area (km^2)	615	0.3	1265
Area (km^2)	154	0.042	14.8
Mean depth (m)	5.4	2.5	3.3
Max. depth (m)	26	7	16
Retention time	~3 years	~2 years	few days
Colour (mg Pt l^{-1})	17.5	118.5	168.9
Total phosphorus (µg l^{-1})	18.2	19.3	65.6
Total nitrogen (µg l^{-1})	452	554	1115
Chlorophyll *a* (µg l^{-1})	7.3	16.4	23.1

2.2 Modelling

2.2.1 PROTECH (I)

PROTECH (Phytoplankton Responses to Environmental Change) is a family of models that has the ability to simulate the simultaneous growth of up to 10 different phytoplankton species, or functional groups, at daily resolution. PROTECH is one dimensional and vertically layered. The published equations describing the maximal replication rates of phytoplankton cells, and the way these rates respond to the physical environment are as follows (Reynolds *et al*. 2001):

The daily (d) maximum species-specific replication rate at 20 °C (r'_{20}) depends on the surface area (s) and the volume (v) of the algal 'unit' (single cell or coenobium)

$$r'_{20} = 1.142(s/v)^{0.325}d^{-1}. \tag{1}$$

In the model, this is employed in the form

$$\log r'_{20} = a + b\log(s/v) \tag{2}$$

where the standard value of the regression intercept a = log(1.142) and the standard value of the slope b = 0.325. The replication rate is adjusted to the ambient water temperature (r'_θ) according to the alga-specific temperature sensitivity (β)

$$\log r'_\theta = \log r'_{20} + \beta[1000/(273+20) - 1000/(273+\theta)]d^{-1}. \quad (3)$$

Temperature sensitivity is evaluated as

$$\beta = 3.378 - 2.505\log(s/v). \quad (4)$$

The replication rate is further adjusted for photoperiod. The fastest light-compensated daily replication rate is thought to be

$$r'_{(\theta,I)} = r'_0 \sum t_p/24 \quad (5)$$

where $\sum t_p$ is the aggregate of the daily photoperiods in h and which is proportional to

$$\sum t_p = \Gamma h_p/h_m \quad (6)$$

where h_m is the depth of the surface mixed layer, h_p the light-compensated column height and Γ the daylight period. The light-compensated column height is defined by

$$h_p = \ln(I'_0/0.5I_k)\varepsilon^{-1} \quad (7)$$

where ε is the coefficient of vertical light attenuation (m^{-1}), I'_0 the daytime-averaged photon flux across the water surface (mol photons m^{-2} s^{-1}) and I_k the photon flux necessary to saturate the instantaneous growth rate. The point where light saturates growth rate is determined as

$$I_k = r'_\theta/\alpha_r. \quad (8)$$

The slope (α_r) of light-dependent growth is a correlative of algal unit maximum dimension (m), surface area (s) and volume (v)

$$\alpha_r = 0.257(ms/v)^{0.236}. \quad (9)$$

In the model it is solved by

$$\log\alpha_r = \log(0.257) + 0.236\log(ms/v). \quad (10)$$

Hence the equation (5) can be solved by substitution in two situations:
when $h_m > h_p$

$$r'_{(\theta,I)} = [r'_\theta \Gamma (24h_m)^{-1}]\ln[2I'_0 0.257(ms/v)^{0.236}\,(r'_0)^{-1}]\varepsilon^{-1} \tag{11}$$

when $h_p >> h_m$

$$r'_{(\theta,I)} = r'_\theta \Gamma (24)^{-1}. \tag{12}$$

Finally, the replication rate (r) is responsible for the change in the number of individuals (N) from the starting population (N_0) in a time step (t)

$$N_t = N_0 e^{rt}. \tag{13}$$

In addition, each modelled phytoplankton taxon has a specific description inside the model which includes information about the cell/colony size, motility, ability to fix nitrogen, requirement for silica and whether the taxon is grazed or not. These qualities additionally determine the fate of the taxa through possible loss processes or increased opportunities to utilize available resources. As input data PROTECH requires daily values for discharge, phosphate (PO_4), nitrate (NO_3) and silica concentrations, outflow, wind speed, cloud cover, air temperature and air humidity. Because PROTECH does not have a routine for ice, the thermodynamic part of the MyLake model (Saloranta and Andersen 2007) was employed to create daily thermal profiles and ice break and formation dates. For running MyLake the additional meteorological variables required were air pressure, precipitation and global radiation. With very little calibration the model was able to simulate the seasonal lake water temperature changes, congruent to observations, and thus provided appropriate thermal input for PROTECH.

From the available phytoplankton data for Lake Pyhäjärvi, 8 taxa were selected for modelling based on their recorded prevalences: 3 cyanobacteria, *Anabaena, Oscillatoria* and *Gloeotrichia*; 2 diatoms, *Asterionella* and *Aulacoseira;* a chrysophyte, *Mallomonas;* a dinophyte, *Gymnodinium*; and a cryptophyte, *Cryptomonas*. 2001, 2005 and 2008 were selected as study years, and although these had the most comprehensive input datasets of all variables, some values still had to be interpolated or given average estimates to obtain the daily figures required by the model. For the inflow and most of the meteorological values from the Jokioinen observatory 100 km east from the lake there were daily values. PO_4 and NO_3 were measured weekly in summer, but only monthly during winter, and for silica there were only 1–2 yearly values. In the calibration process no parameters were changed, but instead the composition of the phytoplankton community, proportion of different species at the beginning and the base chlorophyll *a* level (input from inflow etc.) in different seasons was slightly tuned. During the calibration process it was noticed that inclusion of an internal nutrient loading factor was required, because the external loading alone was not able to maintain the growth of phytoplankton. The existence of an internal loading in Lake Pyhäjärvi has previously been calculated from mass balance equations (Ekholm *et al*. 1997, Nürnberg *et al.* 2012) and demonstrated in laboratory experiments (Lehtoranta and Gran 2002). Elevated phosphorus concentrations near the bottom were also evident in the monitoring data.

PROTECH outputs presented the upper 5 m chlorophyll *a* values that were compared to the corresponding biweekly May–October observations. To compare the outputs for different groups, the observed total chlorophyll *a* was divided based on the fresh weight biomass proportions (g m^{-3}) of each group. After adequate model fit was achieved, it was examined (using MyLake) how rising temperature and increased windiness would affect the physical properties of the lake and (using PROTECH) eventually the phytoplankton growth. The higher daily air temperatures applied were the averages from estimations of 19 global models (sub group of the 23 models used in the IPCC (2007) fourth assessment report) with emission scenario A1B. First an average for each day from the period 2040–2069, and then from 2070–2099 was taken and used as new daily temperatures for one year. According to some estimations by the Finnish Meteorological Institute, the windiness in southwest Finland could increase by as much as 2–4 % by the end of the century, mostly for the windiest period in September–April. Because wind-induced changes to the mixing and water column stability of Lake Pyhäjärvi could have a significant impact on the phytoplankton community, the values for wind speed were also increased by 2–4 % after first making the simulations with higher air temperatures alone.

2.2.2 Structural Equation Modelling (II, III)

Structural Equation Modelling (SEM) is a statistical technique for studying the network of causal relations between variables within a system (Shipley 2002, Hershberger *et al*. 2003). Because of the causality, as well as lower data requirements and complexity, it offers a valuable alternative to deterministic models. Still, it has received quite limited attention among aquatic modellers; for the review it was possible to identify only 13 fairly recent papers that introduced a SEM application for phytoplankton studies. In most of those the impact of different biotic and abiotic variables on phytoplankton development and on the phytoplankton community were examined (Arhonditsis *et al*. 2006, 2007a, b, Liu *et al*. 2010, Salmaso 2011, Gudimov *et al*. 2012) and especially in biodiversity studies the regional and latitudinal effect was included (Korhonen *et al*. 2011, Stomp *et al*. 2011, Chou *et al*. 2012, Matthews and Pomati, 2012). In this study SEM was used to examine the factors affecting the development of phytoplankton in Lake Valkea-Kotinen. For building up the SEM the measured values from approximately the euphotic zone for concentrations of chlorophyll *a*, nutrients (total phosphorus and nitrogen, PO_4, combined nitrite-nitrate and ammonium), water temperature and colour, from years 1990–1995 were used. Because of the strong seasonality and distinct differences between summer and winter conditions (the sampling frequency also decreasing for the winter), only the measurements from the ice-free period were included into the study. Further, only the period between the water column overturns in spring and autumn was considered; thus from each year the first and the last included sampling date was the one when the temperature difference between the surface and the bottom layer had been > 2 °C (roughly from the beginning of May to the end of October). Some sampling occasions had to be left aside, because two of the methods applied for parameter estimations (generalized least squares and asymptotically distribution-free estimates) do not allow missing values. The counts for 3 zooplankton groups, cladocerans, copepods and rotifers from the pooled samples of the upper 5 m were also available. Because the distribution of zooplankton data was strongly skewed, they were transformed using natural logarithms. More detailed descriptions about the sampling and analyses can be found from Peltomaa *et al*. (2013) and Lehtovaara *et al*. (2014).

AMOS software of SPSS was applied for the model building process. A conceptual model describing the main factors affecting the phytoplankton development in Lake Valkea-Kotinen was created (Fig. 1 A). Based on previous knowledge about the function of the Lake Valkea-Kotinen ecosystem and after performing principal component and regression analyses for the data, the final SEM was constructed (Fig. 1 B). The paths (parameters) between the different variables were calculated using maximum likelihood (ML) but also using general least squares (GLS) and asymptotically distribution free (ADF) methods for comparison. This was because the assumption of multivariate normal data for the ML was not fulfilled. To test the goodness of our model that is the congruence between observed and modelled covariance matrices, the χ^2-test was applied. Achieving a good model fit with the available data put some limitations on the variables to be included in the model. To improve the model further, Bayesian analysis was performed to gain better error estimates for some measured variables and information about the underlying uncertainties.

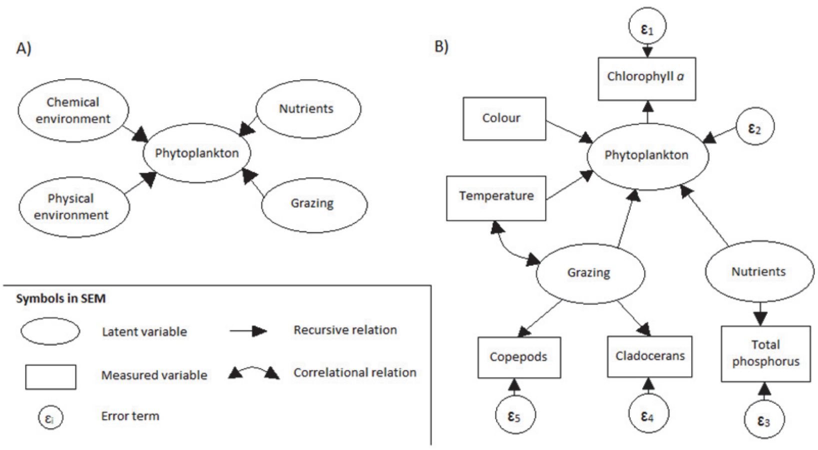


FIGURE 1 Conceptual model (A) of the possible factors affecting phytoplankton development in Lake Valkea-Kotinen and the final Structural Equation Model (SEM) (B) for Lake Valkea-Kotinen.

2.2.3 Linear Mixed Effects Modelling (II, IV)

Linear Mixed Effects (or hierarchical) Modelling (LMM) allows the nested and hierarchical data structure of many environmental variables to be acknowledged (Zuur *et al*. 2009). For instance, the different characteristics (e.g. area, depth, water colour, residence time) of a lake partly determine its water quality and response to elevated nutrient levels (Malve and Qian 2006, Jackson *et al*. 2007, Nõges 2009). Besides lake characteristics, the different ecoregions (Lamon and Qian, 2008), landscapes (Wagner *et al*. 2011) have also been used as hierarchy levels. When using large datasets to derive general models for such site-dependent relationships, it is possible to improve the model reliability by acknowledging these so-called random effects. Yet with LMM it is not necessary to split the data in any way, but all gathered information can be efficiently used (so called 'borrowing strength theory'). Despite this clear advantage over simple regression, for the review only seven LMM papers for phytoplankton was found, of which two were actually using a non-linear, so called general additive method (Carvalho *et al*. 2011, Salmaso *et al*. 2012).

The hierarchical model for chlorophyll *a* introduced by Malve and Qian (2006) was used as a basis for the LLR modelling tool, which has been developed to ease the use of models in WFD-related management of lakes. LLR gives estimates of the loading reduction that is required to have the concentrations of total phosphorus, total nitrogen and chlorophyll *a* under the lake type specific class limits for good water quality (see Table 1 in IV). The assumption of a continuously stirred tank reactor (CSTR) is applied, hence all the values are given as averages for the lake retention time (in years). To improve the model in Malve and Qian (2006), more recent data from Finnish lakes were used, with almost double the number of observations. In addition, the lake types in the model were updated. The dataset consists of 36942 in-lake observations of chlorophyll *a*, total nitrogen and total phosphorus concentrations collected in July–August in 1990–2007 from the upper 2 m. All Finnish lake types are rather well represented in the data. In LLR, the hierarchical chlorophyll *a* model is linked to the simple steady-state mass-balance model of Vollenweider (1968) modified by Chapra (1975). Hence, there is a chain where external phosphorus and nitrogen loadings are first converted into in-lake concentrations of phosphorus and nitrogen. The median concentration estimates are then used in the chlorophyll *a* model to better estimate the target loading that would eventually lead to good ecological condition for the lake. The Marcov chain Monte Carlo methods were used for solving the posterior distributions for settling velocity in the mass-balanse model and for estimating all unknown parameters for the hierarchical chlorophyll *a* model. The final outputs of LLR are given as probability distributions.

With the application to Lake Kuortaneenjärvi, the actual management facilitating properties of the LLR tool were demonstrated and the benefits of this kind of modelling approach were introduced. The measured total phosphorus and nitrogen concentrations from the deepest part of Lake

Kuortaneenjärvi were used as input for LLR (HERTTA database of Finnish Environmental Administration). The data were from 1991–2013 and, because of the CSTR assumption, yearly averages were calculated. The amount of external loading to Lake Kuortaneenjärvi has not been measured, hence the external loading and water outflow estimates for years 1991–2013 were derived from the nutrient loading estimation tool WSFS-Vemala. The WSFS-Vemala tool is based on a modelling system which includes the simulation of hydrology, nutrient leaching from fields and forests as well as nutrient transport in rivers and lakes (Markus Huttunen, Finnish Environment Institute, written communication).

3 RESULTS AND DISCUSSION

3.1 Effect of higher temperature on the phytoplankton community in Lake Pyhäjärvi

The PROTECH output was in good congruence with the observed values, while manipulating the input data showed that rising temperatures pose a risk for more frequent HABs in the future, as the Cyanobacteria are able to out-compete the other phytoplankton groups in warmer conditions (I). The total chlorophyll *a* values did not change markedly, because the nutrient levels were left intact. Hence, the percentage of cyanobacteria from the total phytoplankton biomass is a valuable measure when assessing the ecological quality of lakes and monitoring its development. The change in windiness did not have any additional effect on the mixing of the water column, that already at present mixes very easily. The results with higher temperatures indicate that, although the present water quality of Lake Pyhäjärvi is good, the in-lake nutrient levels may still need to be decreased to prevent future cyanobacterial blooms. There are parallel results from Lake Pyhäjärvi showing that more zooplankton is needed to control the cyanobacteria biomasses at higher temperatures (Malve *et al*. 2007). The ongoing management actions are thus crucial for the good condition of the lake and they may have to be further enhanced, especially when studies indicate that internal loading has a significant role in the phosphorus cycle of Lake Pyhäjärvi, and should be taken into consideration as a significant source of phosphorus.

A shift in the timing of the phytoplankton spring bloom was expected, but that did not happen even though the ice thaw took place a month earlier with warmer temperatures. Similar problems have been reported with applications for other ice-covered lakes (Markensten and Pierson 2007), and the problem may derive from the cold water temperatures that followed the early ice break. PROTECH does not have a routine for ice and snow, so it was linked with the MyLake model to enhance the physical description of the ice-cover period. Obviously the under-ice phytoplankton growth (Vehmaa and Salonen 2009, Kiili *et al*. 2009) is also an unfamiliar phenomenon for PROTECH, which raises a question about the validity of the PROTECH outputs near these extreme periods.

The benefit of higher temperature for cyanobacteria is evident (Jöhnk *et al*. 2008, Paerl and Huisman 2008), and has been shown in many previous modelling studies (reviewed in Elliott 2012). Our warmest study year 2008 (warmest in SW Finland since 1961, Ventelä *et al*. 2011) also served as a good comparison point for the model-implied phytoplankton development in other years, when temperatures were increased. The observations were in good congruence to those predicted by the model. Nevertheless the problems with the spring peak, which to some extent determines the community development later in the year, are unfortunate. Although Lake Pyhäjärvi has relatively comprehensive and frequent observational data available by Finnish standards, phytoplankton data are lacking from late autumn to early spring and a lot of interpolation, averaging and assumptions had to be made when collating all the necessary input data for the model. In addition, PROTECH does not include a function for internal loading, so an estimate to represent its existence had to be added. This all makes the question about missing uncertainty analysis very relevant, because already the input data are prone to error and it cannot be guaranteed the model is able to describe correctly the function of a system with internal loading, not to mention one that is ice-covered for about one-third of the year. Although application with PROTECH does not include any tuning of parameters, the absence of uncertainty analysis considering all the other issues means the results are essentially a qualitative description, and the modelling effort is disproportionate compared to the information gained. For instance, concrete figures of permissible loading levels in the future, to compensate for the effect of warming, cannot be derived from the model.

Because of these reservations and the lack of data to resolve them, or to develop the model to be more suitable for boreal conditions, the feasibility of deterministic models like PROTECH for studies with Finnish lakes is poor at this stage. However, if this kind of approach is chosen, commitment to monitor some lake(s) extensively enough for an indefinite period is needed (c.f. the research and model development done on Lake Constance) to gain a robust scientific basis for complicated modelling studies of boreal aquatic ecosystems. Monitoring the factors most connected to the features of boreal lakes will be challenging per se; thus it would be advisable to test the existing lake models as a basis for phytoplankton modelling in Finland and, if possible, add the 'boreal compartments' into those rather than develop completely new models (this approach in general is discussed by Trolle *et al*. 2012).

The review of alternative modelling techniques (II) originated from the experiences with Lake Pyhäjärvi PROTECH application and also from some other attempts (unpublished) to apply first PROTECH and then the phytoplankton part under development of the simpler model MyLake for other Finnish lakes. The problem with the lack of data became even more emphasised with other lakes, as Lake Pyhäjärvi has perhaps the most thorough observations from the lake and its surroundings. The development of MyLake on the other hand would have needed even better datasets than that from Lake Pyhäjärvi. Because of the uncertainties the review focused to introduce statistical methods, which have gained less attention within aquatic modelling. They would be more suitable for limited data sets and better serve the lake management work as well.

Based on the reviewed studies, Structural Equation Modelling (SEM) and Linear Mixed Modelling (LMM) seem to be suitable alternatives for modelling phytoplankton and water quality related issues. These kinds of alternative modelling approaches should be encouraged to obtain more information about their performance in various aquatic systems and to enhance the possibility of utilizing them more efficiently. The methods have not been widely used in aquatic studies, but have many favourable features for analysing ecological observation data, including assumptions that are easier to meet and compatibility with Bayesian inference. Applications in other fields of science are numerous, but for our review only handful of phytoplankton related applications came up in the search (II).

3.2 Factors affecting the development of phytoplankton in Lake Valkea-Kotinen

A realistic model was created for Lake Valkea-Kotinen using SEM that indicates the interactions between phytoplankton (expressed as chlorophyll *a*) and some basic variables (III). However, some variables like phosphate and different fractions of nitrogen were omitted from the model because the available data did not fully support their inclusion. For phosphate that may have been because its values were under the detection limit on many sampling occasions and this has been problematic in other SEM studies as well (Arhonditsis *et al*. 2006). The relationship between epilimnetic nutrients and phytoplankton is also affected by the presence of motile algae that can access additional nutrient resources in the deeper water (Salonen and Rosenberg 2000). Nevertheless, nutrients, which now included only total phosphorus, together with temperature had a clear positive impact on phytoplankton development. The negative effect of water colour and grazing was also included in the model, although their effect was not as clear.

In addition, it was possible to evaluate potential reasons for unclear zooplankton–phytoplankton interactions in Lake Valkea-Kotinen (Lehtovaara *et al*. 2014, Arvola *et al*. 2014). The negative effect of grazing on phytoplankton was also weak in SEM, probably because the phytoplankton community in 1990–1995 was dominated by the motile Raphidophyte *Gonyostomum semen* which is inedible for many zooplankton species (Jones *et al*. 1999, Lebret *et al*. 2012, Peltomaa *et al*. 2013). The effect of zooplankton on *G. semen* may indeed be negligible or they may even further boost its growth by eliminating the competing species and recycling nutrients. From the posterior distributions of

Bayesian analysis it was possible to see which parameter estimates were most uncertain, and the inconsistencies with zooplankton were also elucidated in those. On the other hand, the zooplankton groups in SEM did not cover all the possible grazers. There is also a clear seasonal pattern in the highest densities of each group (Lehtovaara *et al*. 2014); thus some information about the true grazing pressure and of the reasons it may vary was missing.

The negative effect of higher water colour on phytoplankton growth was also shown in the model. Phytoplankton photosynthesis suffers from the declining light levels, but the link between the two variables could also partly be explained by the fact that colour values tend to be lower in the summer months when phytoplankton peak. This should be considered especially because the dominant *G. semen* is shown to favour brown waters (Trigal *et al*. 2013). These kinds of species-specific questions require species-level modelling, and although a promising start the present Lake Valkea-Kotinen SEM could be improved in several ways, mostly with more rigorous data allowing more specific questions and also with higher number of observations. However, even in this form the SEM can be used as adequate prior information for studies of other lakes and it was already possible to gain a more detailed overview of the system function than some other studies applying stepwise regression (Arvola *et al*. 2014) for the Lake Valkea-Kotinen data, or examining temporal trends in the data (Peltomaa *et al*. 2013).

More applications with SEM would be needed to better evaluate its performance. Data sets with n~200 are not too challenging to gather, especially when there is some freedom to decide what to measure. SEM can be used for dynamic estimations of the system, coupled with uncertainty analysis (e.g. Arhonditsis *et al*. 2007b), which makes it an interesting alternative to deterministic models. Yet its strength, especially with boreal lakes at the moment, is in the possibility to describe interactions within these systems as a basis for further model development. However, as seen with the effect of water colour on phytoplankton in the Lake Valkea-Kotinen SEM, caution should be exercised when interpreting the results and considering the possible influence of factors not included in the model.

3.3 Estimates of target nutrient loading for Lake Kuortaneenjärvi

Acknowledging the effect of lake type on nutrient–chlorophyll *a* relationship improved the fit in a hierarchical chlorophyll *a* model for Finnish lakes, which now provides a better basis for estimates than one with 'global parameters' (IV). In addition to global intercept and slopes of total phosphorus, total nitrogen and that of their interaction lake-type-specific intercepts and slopes of phosphorus and nitrogen were gained. Those clearly show that the effect of nutrients varies between lakes (Fig. 4 in IV).The posterior box plots indicate that some groups could be combined at least inside the model to gain more group-specific data, and on the other hand the lakes in northern Lapland are a quite heterogeneous group. The LLR output showed that the fit for the nutrient models was reasonable. The probability of reaching the 'good' class limit for phosphorus in Lake Kuortaneenjärvi with the current phosphorus loading level (1.45 g m^2 a^{-1}) is only 4 %. The most probable phosphorus status is poor (52 %). The yearly average of observed total phosphorus concentrations never reached the good class limit during the study period, and poor water quality is apparent. The LLR estimate for target nutrient loading is derived from the median outflow and median external loading level, so the relationship between external loading and in-lake nutrient concentration is linear. In order to achieve good status on average, the external phosphorus loading should be not more than 1.0 g m^2 a^{-1}. Therefore, the phosphorus loading should be reduced by 30 %. For total nitrogen the most probable status with average nitrogen loading (27 g m^2 a^{-1}) is moderate (46 %). The probability of achieving good status with current loading level is also 46 %. As the critical nitrogen loading is 23 g m^2 a^{-1} the required loading reduction is about 13 %. It should be noted that these values are the minimum requirements, because as a default LLR gives target load estimates that lead to good water quality with 50 % probability. This represents a more political view of lake management.

With estimated median total phosphorus and total nitrogen concentrations the median chlorophyll *a* concentration in Lake Kuortaneenjärvi is 25 µg l^{-1}, which corresponds quite well to the observed average 28 µg l^{-1}. The probability distribution of chlorophyll *a* shows that the lake is in moderate condition (100 %). The uncertainty in the chlorophyll *a* model is small and thus the probability distribution is quite narrow. However, because LLR uses median nutrient values for chlorophyll *a* estimates the overall uncertainty in the model chain is larger, because with nutrient values diverging from the median the chlorophyll *a* estimate changes (illustrated in Fig. 8 of IV). Yet, even when using 25th and 75th percentiles for nutrients the density peak is still located at the good quality limit of 20 µg chlorophyll *a* l^{-1}. Lake Kuortaneenjärvi chlorophyll *a* contours as a function of nutrient loadings show the median chlorophyll *a* estimate with different phosphorus and nitrogen loading combinations (Fig. 7 in IV). In order to achieve good chlorophyll *a* levels in the lake, it would be reasonable to reduce the loading of both nutrients. It is unlikely that reducing nitrogen loading alone near its high water quality levels would lead to chlorophyll *a* concentration under the good quality level, as the contours imply. There may be some problems in the phosphorus model because Lake Kuortaneenjärvi suffers from internal loading, which may also partly explain the weak response of observed concentration values to changes in external loading (see Fig. 6 A of IV). Including internal loading into LLR is one future aim, since it quite commonly affects the nutrient balance in eutrophic lakes.

The highest uncertainty in LLR arises from the missing loading values, and hence it is not possible to give detailed estimates of the needed loading reduction. Then again, it is not assumed the loading values from VEMALA differ drastically from the real ones because yearly averages are used. This does not reduce the importance of proper loading measurement, as the management actions need to be targeted somehow. Hopefully, as automatic monitoring techniques develop and adaptive monitoring and management strategies evolve, there will be more data available. Most importantly, with more comprehensive nutrient balance data, the whole nutrient retention model could be improved, because utilizing similar hierarchical model structure as with chlorophyll *a* would then be possible (Cheng *et al*. 2010).

Despite the loading, the phosphorus and chlorophyll *a* concentrations in the lake are almost invariably too high (also in the observations). The outputs of LLR thus give an estimate of the severity of the problem and indicate how extensive management actions should be prepared. Reducing the phosphorus loading by 30 %, even if having some degree of uncertainty, is challenging. If the aim is to increase the probability of success in lake management (from the 50:50 situation), the loading reduction has to be even more. It has been estimated, that 65 % of the present phosphorus loading to Lake Kuortaneenjärvi comes from agriculture, 22 % from forestry and 10 % from scattered settlements (Väisänen 2013), and the needed reduction of ~30 % puts high pressure on agriculture.

LLR has been applied already for several lakes during WFD-related work. There has been an urgent need for fast and simple assessment tools when implementing the WFD for which LLR is especially tailored. Hopefully when the management projects start there will be funding directed to monitoring some of the lakes during and after. The collected data will then enable better evaluation of whether the predictions of LLR hold and also improve the model further. Although the approach in LLR sounds modest, it is justifiable and represents a significant improvement in situations where the only other option would be using the mass-balance equation of Vollenweider (1968) alone or estimating the development of chlorophyll *a* concentration from a regression line. It serves a good starting point for lake management, but with yearly averages little can be said about the system function.

4 CONCLUSIONS

There is a real risk that phytoplankton modelling for scientific purposes, and all the lake ecosystem modelling that could follow, will remain at a modest level in Finland. If there is not a stronger commitment to collecting high quality time-series of environmental data with sufficient temporal and spatial resolution for phytoplankton studies, detailed models cannot be set up or properly validated and nor can the existing models be developed further to be applied to boreal lakes. This would be most unfortunate, as the changing environment and climate add new pressures on boreal lakes, and it should be possible to study the possible impacts more efficiently, combining data collection, modelling and critical evaluation into an iterative process. The studies and data collection should also be extended outside the growing season. The change in temperatures and rainfall during autumn–spring introduces most drastic changes to the transport of nutrients and organic matter into lakes and also to the thermal properties of lakes. These changes may affect the phytoplankton development in spring, but also in the following growing season. In addition, the model development cannot be based solely on the routinely collected monitoring data, because it is often too restricting. There are on-going discussions about potential intensive monitoring sites that could enhance the overall lake ecosystem research, but as yet these schemes are still open.

All the studies made for this thesis suffered from the issue of shortage of available data, although they represented very different approaches and varying levels of complexity. The results from the PROTECH application to Lake Pyhäjärvi, although improving the understanding of the lake function, raised more questions than the model was able to answer. This was mostly because many processes particularly important for boreal lakes are not currently included in the model. The statistical methods SEM and LMM would also require better data if applied for purely ecological questions. These two approaches, however, seem more appropriate at the moment for further studies, because they have more flexibility around the data issue and, on the other hand, they could also indicate some guidelines for the development of deterministic models. They (especially LMM) also offer a way to utilize more efficiently existing environmental monitoring data, as the inevitable uncertainty and error in the modelling can be quantified when Bayesian analysis is included in the modelling process. Besides being an important feature for assessing model performance, examining the sources of uncertainty allows the monitoring effort to be best targeted in order to reduce the uncertainty. The false assumption that useful model predictions can be produced without uncertainty increases the risk of poor decisions and unfortunately may weaken the justification for the importance of adaptive management, which would serve both model development and evaluation of the efficiency of management actions.

Despite the data issue, there is now a better overall picture of how to continue the utilization of models in phytoplankton studies. The focus should be more in explorative modelling studies and data analysis, because these present studies already indicated that there are many unanswered questions, which should not in the end be seen as drawbacks. Instead, they should be seen as an inspiration for further studies which will benefit the implementation of some sophisticated modelling techniques. On some occasions even approximate model outputs can be valuable, although the shortcomings of the approach need to be indicated properly. Yet, in a strict funding situation it should also be evaluated what is the sufficient role of models when addressing eutrophication problems in practice. If the modelling is costly, and mainly confirms what is already expected, and if the modeller cannot present any quantitative estimation of the model uncertainly, then it would be better to target the resources towards management actions and monitoring. In the end it will benefit both lake recovery and modelling.

Acknowledgements

First I would like to express my gratitude to my supervisors Professor Roger Jones, Research Manager and Adjunct Professor Timo Huttula and Dr Olli Malve. Without Timo and his persistent work to establish a lake modelling unit at Jyväskylä, taking this path would have never been possible. Although sometimes a rocky path, it is definitely the one I would choose again. The advice and support from Roger during these years have been invaluable. Despite you being the only non-modeller supervisor, I always left your room with my head full of ideas. First and foremost I have you to thank for being able to complete my PhD work and this thesis. From Olli I found someone who shares my ideas and visions and the enthusiasm for water management issues. You have always shown interest and support towards my work, and I am grateful for all the advice you provided when I decided to shift towards the more statistical side.

I am obliged to the Maj and Tor Nessling foundation, the VALUE doctoral school and Maa- ja vesitekniikan tuki ry. for the funding provided for my work. I also want to thank the Finnish Environment Institute (SYKE) for being able to perform my work at the Jyväskylä office and attend the many activities and work done at SYKE. The Centre for Ecology and Hydrology, Lancaster kindly enabled my one month stay at the Lake Ecosystem Group while I was learning to use PROTECH with Dr Alex Elliott. I want to thank Alex, Professor Emeritus Jouko Sarvala and Professor Lauri Arvola for their valuable help with my work and many good ideas concerning the studies. The help from Petri Kiuru with MyLake was also indispensable. Adjunct Professor Anne-Mari Ventelä, Dr Tiina Tulonen and Antti Taskinen have contributed this work as co-authors of the research papers, and the first two are much to thank for the phytoplankton data available for modelling.

I feel almost privileged that I got to share an office with a statistician, Niina Kotamäki for most of this work. With you I had the most interesting discussions and created the best ideas. You have been a great help and hopefully our biologist-statistician symbiosis will continue. There are numerous people at SYKE and at the WETsection at the University of Jyväskylä who have helped me in many different ways or just been great co-workers and friends to have lunch or coffee with. Also the meetings with the members of the VALUE doctoral school have been memorable and offered a possibility to share 'the agony of PhD studies'. Finally, I want to thank the pre-examiners Research Professor Tom Frisk and Professor Frank Peeters for their comments on the thesis.

To be blessed with supporting and ever encouraging family and friends is the biggest fortune to have. No drawback or problem could ever become larger than the tranquillity and optimism I found from your company. Especially my dearests Ville, Leevi and Viljo; you fill my life with love, laughter and adventures. More than that Ville, you have eased my life tremendously during these years by being a superb father and an amazing husband.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Mallien hyödyntäminen boreaalisten järvien kasviplanktonin runsauteen ja kasviplanktonyhteisön dynamiikkaan vaikuttavien tekijöiden tarkastelussa

Kasviplanktonilla tarkoitetaan mikroskooppisen pieniä leviä, jotka yleisimmin ajelehtivat vapaina veden mukana. Kasviplanktonlajeja on tuhansia, ja niitä esiintyy niin suolaisessa kuin makeassakin vedessä. Muiden kasvien tavoin kasviplankton kykenee yhteyttämään ja pienestä koostaan huolimatta sitomaan yhteyttämisprosessissa merkittäviä määriä ilmakehän hiilidioksidia. Samalla vapautuu happea, ja yhteyttämisessä varastoitunut hiili ja energia siirtyvät edelleen muiden eliöiden hyödynnettäväksi. Kasviplanktonilla on siis äärimmäisen tärkeä, vaikka jokseenkin näkymätön rooli ilmakehän hiilidioksidi- ja happipitoisuuden säätelijänä sekä vesiekosysteemien perustuottajana.

Tietyissä tilanteissa—erityisesti silloin, kun kasvuun tarvittavia ravinteita, fosforia ja typpeä, on runsaasti saatavilla—kasviplankton voi kuitenkin runsastua niin, että levien muodostamat massaesiintymät ovat selkeästi havaittavissa. Tällöin kasviplankton koetaan vähintäänkin esteettiseksi haitaksi ja massaesiintymien syntyä pyritään ehkäisemään. Erityisen haitallisiksi massaesiintymät muodostuvat ihmisille ja muille eliöille silloin, kun niiden taustalla on myrkyllisiä yhdisteitä tuottavia lajeja. Makeissa vesissä näitä ovat eräät sinilevät eli syanobakteerit, jotka luetaan kasviplanktoniin, vaikka ne rakenteeltaan muistuttavatkin enemmän bakteereita. Myrkylliset sinilevät tuottavat ihoa ärsyttäviä yhdisteitä sekä maksa- ja hermomyrkkyjä, joihin liittyviä eläinten ja ihmistenkin kuolemia on raportoitu. Kaikki sinilevälajit eivät ole myrkyllisiä, mutta koska lajimääritykset ja yhdisteiden toteaminen vedestä vaativat laboratoriotutkimuksia, yleisenä suosituksena on välttää veden kaikenlaista käyttöä sinilevien massaesiintymän eli sinileväkukinnan aikana ja pian sen jälkeen. Sinileväkukintoja esiintyy erityisesti tyynien ja lämpimien kesäpäivien aikana. Kaasurakkuloiden avulla sinilevät nousevat tyynen veden pintaan, ja mitä lämpimämpää vesi on, sitä nopeammin ne kasvavat ja lisääntyvät verrattuna muihin lajeihin.

Erilaisten mallien avulla voidaan tarkastella lähemmin kasviplanktonin runsastumiseen liittyviä tekijöitä ja numeerisiin ennusteisiin tukeutuen tehdä päätöksiä tarvittavista toimenpiteistä massaesiintymien ehkäisemiseksi. Tämän väitöskirjatyön ensimmäisessä osassa (I) tutkittiinkin determinististä PROTECH-mallia käyttäen, miten ilmastonmuutoksesta johtuva lämpötilojen nousu vaikuttaa haitallisten sinilevien runsauteen. Mallin antaman ennusteen mukaan sinilevät hyötyvät selvästi korkeammista lämpötiloista ja voivat runsastua muiden lajien kustannuksella. Tämä voi tapahtua siitä huolimatta, että kasvuun tarvittavien ravinteiden pitoisuudet pysyvät nykyisellä, kohtuulliseksi arvioidulla tasolla. Ilmastonmuutos voi siis luoda lisäpaineita järviin päätyvän fosfori- ja typpikuorman vähentämiseen.

PROTECHin kaltaisissa deterministisissä malleissa kasviplanktonin kasvuun ja (eri lajien) runsauteen vaikuttavat prosessit on määritetty yhtälöiden avulla. Näin ollen voidaan tehdä yksiselitteisiä ennusteita siitä, miten esimerkiksi lämpötilan tai ravinnepitoisuuden muutos vaikuttaa kasviplanktonin kasvunopeuteen tai miten kasviplanktonyhteisö kehittyy vuoden mittaan. Deterministinen lähestymistapa ei kuitenkaan ole ongelmaton eikä mallin antama yksiselitteinen ennuste virheetön. Erityisesti mallien suuri havaintoaineiston tarve järvestä ja sen ympäristöstä vaikeuttaa ennusteiden tekoa vähän tutkituille järville. Koska kasviplanktonin kasvuun vaikuttavia tekijöitä on lukuisia, malleista tulee helposti myös erittäin monimutkaisia kokonaisuuksia, joiden käyttö vaatii runsaasti laskentatehoa ja asiantuntemusta. Kaikkia kasviplanktonin kannalta oleellisia prosesseja ja vuorovaikutuksia ei edes tunneta, joten mallit väistämättä ovat yksinkertaistuksia todellisesta järviekosysteemistä ja niiden antamissa tuloksissa on epävarmuutta ja virhettä. Esimerkiksi boreaalisille järville tyypillisiä piirteitä, kuten selvää vuodenaikaisuutta, ei ole huomioitu muualla kehitetyissä malleissa. Niiden avulla ei siis voida luotettavasti tarkastella, miten esimerkiksi järvien jääpeitteisen ajan lyhentyminen vaikuttaa kasviplanktoniin (I). Prosessien lisääminen malleihin ja mallien toiminnan asianmukainen testaus vaatii sekin kattavan havaintoaineiston, jota on toistaiseksi heikosti saatavilla suomalaisista järvistä.

Koska havaintoaineiston puute on merkittävä ongelma, työn seuraavassa vaiheessa tarkasteltiin tilastollisia mallinnusmenetelmiä, sillä niiden avulla hajanaisen, mahdollisesti useista eri järvistä kerätyn ja laajuudeltaan rajatun aineiston hyödyntämiseen on paremmat edellytykset. Katsaus hierarkkisen mallinnuksen ja rakenneyhtälömallien kasviplanktoniin liittyviin sovelluksiin osoitti, että deterministisille malleille on olemassa varteenotettavia vaihtoehtoja, vaikka ne ovatkin saaneet vähemmän huomiota osakseen (II). Determinististiin malleihin nähden menetelmät ovat myös yksinkertaisempia sekä usein helpompia ja halvempia käyttää. Lisäksi mallituloksiin aina liittyvä epävarmuus voidaan ottaa paremmin huomioon ja ilmaista numeerisesti (II, III, IV). Nämä kaikki ovat tärkeitä ominaisuuksia erityisesti silloin, kun malleilla halutaan tukea käytännön vesienhoitotyötä.

EU:n vesipuitedirektiivi velvoittaa kaikkien jäsenalueiden järvien olevan hyvässä ekologisessa tilassa jo lähivuosina. Tämä tarkoittaa kasviplanktonin osalta ennen kaikkea sitä, että sen kasvua edistävien ravinteiden, erityisesti fosforin ja typen, pitoisuudet järvissä saadaan riittävän alhaisiksi. Kasviplanktonin määrää kuvastavan klorofylli *a*:n riippuvuutta kokonaisfosfori- ja kokonaistyppipitoisuuksista suomalaisissa järvissä tarkasteltiin hierarkkisen mallin avulla yhdistämällä yli 2000 järven havainnot (IV). Näin ollen yksittäisille, vähemmän tutkituille järville voidaan tehdä tarkempia ennusteita hyödyntäen muistakin järvistä kerättyä aineistoa. Menetelmä vastaa tavallista lineaarista regressiota, mutta hierarkkisen mallirakenteen avulla voidaan huomioida, miten järven ominaisuudet, kuten koko, syvyys ja veden väri, vaikuttavat siihen, kuinka voimakkaasti kasviplanktonin määrä kasvaa ravinnepitoisuuksien kohotessa. Liittämällä tämä hierarkkinen klorofyllimalli yksinkertaiseen ravinnemalliin saatiin malliketju ja työkalu, jonka avulla fosfori- ja typpikuormitus voidaan ravinnepitoisuuksien kautta johtaa klorofylli *a* pitoisuudeksi—voidaan siis tarkastella, miten kuormituksen pieneminen vaikuttaa järven kasviplanktonin määrään. Mallityökalu on jo sen sisältämien muuttujien suhteen erittäin yksinkertainen, mutta käytännössä klorofylli *a* sekä kokonaisravinteet ovat helposti mitattavia perusmuuttujia, joista löytyy havaintoja vesipuitedirektiivin toimeenpanoa varten luokitelluista järvistä. Hoitotoimenpiteitä vaativia järviä on arvioiden mukaan noin 700, joten yksinkertaisuus myös mahdollistaa työkalun käytön mahdollisimman monen järven hoidon ja kunnostuksen tueksi. Mallitulosten epävarmuutta ja virhettä voidaan tarkastella tehokkaasti Bayes-lähestymistapaa hyödyntäen. Tuloksina ei anneta vain yksittäisiä lukuja, vaan ennuste ilmaistaan eri tulosvaihtoehtojen todennäköisyysjakaumana. Samalla voidaan tarkastella, millaisia kuormitusvähennyksiä vähintään tarvitaan, jotta järven tila olisi todennäköisemmin hyvä kuin huono. Arvio on hyödyllinen, sillä kuormituksen vähentäminen on hankalaa ja kallista.

Rakenneyhtälömallien avulla tutkittiin tarkemmin kasviplanktonin kehitykseen vaikuttavia tekijöitä pienessä humusjärvessä (III). Rakenneyhtälömallinnuksessa havaintoaineiston tarve tutkimusjärvestä on suurempi kuin hierarkkisessa mallinnuksessa. Mitä yksityiskohtaisempia kysymyksiä tarkastellaan, sitä suuremmaksi tarve kasvaa. Rakenneyhtälömallien kiistattomana etuna kuitenkin on, että niillä voidaan determinististen mallien tavoin tarkastella syy–seuraus-suhteita useiden muuttujien välillä. Malliin saatiin oletusten mukaisesti lisättyä ravinteiden ja lämpötilan positiivinen sekä veden värin ja eläinplanktonin negatiivinen vaikutus kasviplanktoniin. Humuksen ja raudan aiheuttama veden väri (tummuus) on boreaalisten järvien kohdalla mielenkiintoinen muuttuja, sillä se vaikuttaa yhteyttämiseen käytettävän valon määrään ja veden lämpötilaan. Lauhtuvien talvien myötä humuksen kulkeutumisen valuma-alueelta järviin on arvioitu lisääntyvän. Bayes-analyysin yhdistäminen rakenneyhtälömalleihin mahdollisti tässäkin tapauksessa epävarmuuksien tarkastelun. Kun lisäksi havaittiin, mihin muuttujiin suurin epävarmuus kohdistuu, ne voidaan jatkotutkimuksissa pyrkiä ottamaan lähempään tarkasteluun.

Mallien avulla voidaan siis etsiä vastauksia monentyyppisiin kysymyksiin. Mallien käyttöä kasviplanktonin tutkimuksessa tulisikin edistää, sillä niiden avulla voidaan tukea sekä perustutkimusta että käytännön vesienhoitotyötä. Kaikkien mallien ja menetelmien kehityksen perusedellytyksenä on kuitenkin nykyistä kattavamman havaintoaineiston kerääminen. Koska mallinnusmenetelmiä on useita ja eri menetelmillä on omat vahvuutensa, niitä tulisi käyttää monipuolisemmin, jopa yhdistellen, erilaisiin kysymyksiin. Erilaisten mallien testaaminen samoilla aineistoilla ja tulosten keskinäinen vertailu on myös tarpeen mahdollisten ongelmakohtien selvittämiseksi.

REFERENCES

Anderson T.R. 2005. Plankton functional type modelling: Running before we can walk? *J. Plankton Res*. 27: 1073–1081.

Antikainen R., Haapanen R., Lemola R., Nousiainen J.I. & Rekolainen S. 2008. Nitrogen and phosphorus flows in the Finnish agricultural and Forest Sectors, 1910–2000. *Water Air Soil Poll*. 194: 163–177.

Arhonditsis G.B. & Brett M.T. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Mar. Ecol. Prog. Ser*. 271: 13–26.

Arhonditsis G.B., Paerl H.W., Valdes-Weaver L.M., Stow C.A., Steinberg L.J. & Reckhow K.H. 2007a. Application of Bayesian structural equation modeling for examining phytoplankton dynamics in the Neuse River Estuary (North Carolina, USA). *Estuar. Coast Shelf S*. 72: 63–80.

Arhonditsis G.B., Stow C.A., Paerl H.W., Valdes-Weaver L.M., Steinberg L.J. & Reckhow K.H. 2007b. Delineation of the role of nutrient dynamics and hydrologic forcing on phytoplankton patterns along a freshwater-marine continuum. *Ecol. Model.* 208: 230–246.

Arhonditsis G.B., Stow C.A., Steinberg L.J., Kenney M.A., Lathrop R.C., McBride S.J. & Reckhow K.H. 2006. Exploring ecological patterns with structural equation modeling and Bayesian analysis. *Ecol. Model.* 192: 385–409.

Arvola L., Salonen K., Keskitalo J., Tulonen T., Järvinen M. & Huotari J. 2014. Plankton metabolism and sedimentation in a small boreal lake—a long-term perspective. *Boreal Env. Res*. 19: 83–96.

Benincà E., Huisman J., Heerkloss R., Jöhnk K., Branco P., Van Nes E.H., Scheffer M. & Ellner S.P. 2008. Chaos in a long-term experiment with a plankton community. *Nature* 451: 822–825.

Bolker B.M., Brooks M.E., Clark C.J., Geange S.W., Poulsen J.R., Stevens M.H.H. & White J.S.S. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol*. 24: 127–135.

Caminiti J.E. 2004. Catchment modelling—a resource manager’s perspective. *Environ. Modell. Softw*. 19: 991–997.

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

Carvalho L., Miller (nee Ferguson) C.A., Scott E.M., Codd G.A. Davies P.S. & Tyler, A.N. 2011. Cyanobacterial blooms: statistical models describing risk factors for national-scale lake assessment and lake management. *Sci. Total Environ*. 409: 5353–5358.

Chapra S. 1975. Comment on "An empirical method of estimating the retention of phosphorus in lakes" by W. B. Kirchner and P. J. Dillon. *Water Resour. Res*. 11: 1033–1034.

Chen Q. & Mynett A. 2006. Modelling algal blooms in the Dutch coastal waters by integrated numerical and fuzzy cellural automata approaches. *Ecol. Model.* 199: 73–81.

Cheng V., Arhonditsis G.B. & Brett M.T. 2010. A revaluation of lake-phosphorus loading models using a Bayesian hierarchical framework. *Ecol. Res.* 25: 59–76.

Chou W.-R., Fang L.-S., Wang W.-H. & Tew K.S. 2012. Environmental influence on coastal phytoplankton and zooplankton diversity: a multivariate statistical model analysis. *Environ. Monit. Assess*. 184: 5679–5688.

Clark J.S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters.* 8: 2–14.

Clark J.S., Carpenter S.R., Barber M., Collins S., Dobson A., Foley J.A., Lodge D.M., Pascual M., Pielke R. Jr., Pizer W., Pringle C., Reid W.V., Rose K.A., Sala O., Schlesinger W.H., Wall D.H. & Wear D. 2001. Ecological Forecasts: An Emerging Imperative. *Science* 293: 657–660.

Codd G.A., Morrison L.F. & Metcalf J.S. 2005. Cyanobacterial toxins: risk management for health protection. *Toxicol. Appl. Pharmacol*. 203: 264–272.

Doherty J. & Christensen S. 2011. Use of paired simple and complex models to reduce predictive bias and quantify uncertainty. *Water Resour. Res*. 47: W12534, doi:10.1029/2011WR010763.

Ekholm P., Malve O. & Kirkkala T. 1997. Internal and external loading as regulators of nutrient concentrations in the agriculturally loaded Lake Pyhäjärvi (southwest Finland). *Hydrobiologia* 345: 3–14.

Elliott J.A. 2012. Is the future blue-green? A review of the current model predictions of how climate change could affect pelagic freshwater cyanobacteria. *Water Res*. 46: 1364–1371.

Elser J.J. & Urabe J. 1999. The stoichiometry of consumer-driven nutrient recycling: theory, observations, and consequences. *Ecology* 80: 735–751.

European Parliament & Council 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities* L 327: 1–73.

Flynn K.J. 2003. Modelling multi-nutrient interactions in phytoplankton; balancing simplicity and realism. *Prog. Oceanog.* 56: 249–279.

Gudimov A., O'Connor E., Dittrich M., Jarjanazi H., Palmer M.E., Stainsby E., Winter J.G., Young J.D. & Arhonditsis G.B. 2012. Continuous Bayesian Network for Studying the Causal Links between Phosphorus Loading and Plankton Patterns in Lake Simcoe, Ontario, Canada. *Environ. Sci. Technol*. 46: 7283–7292.

Hershberger S.L., Marcoulides G.A. & Parramore M.M. 2003. Structural equation modeling: an introduction. In: Pugesek B.H., Tomer A. & von Eye A. (eds), *Structural Equation Modeling: Applications in Ecological and Evolutionary Biology*, Cambridge University Press, Cambridge, pp. 3–41.

Holling C.S. (ed.) 1978. *Adaptive Environmental Assessment and Management*. Wiley, Chichester.

Holmes J. & Clark R. 2008. Enhancing the use of science in environmental policy-making and regulation. *Environ. Sci. Policy* 11: 702–711.

IPCC 2007. *Synthesis report. Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change*. IPCC, Geneva.

Jackson L.J., Lauridsen T.L., Søndergaard M. & Jeppesen E. 2007. A comparison of shallow Danish and Canadian lakes and implications of climate change. *Freshwat. Biol*. 52: 1782–1792.

Jakeman A.J., Letcher R.A. & Norton J.P. 2006 Ten iterative steps in development and evaluation of environmental models. *Environ. Modell. Softw*. 21: 602–614.

Jöhnk K., Huisman J., Sharples J., Sommeijer B., Visser P.M. & Stroom J.M. 2008. Summer heatwaves promote blooms of harmful cyanobacteria. *Glob. Change Biol*. 14: 495–512.

Johnson J.B. & Omland K.S. 2004. Model selection in ecology and evolution. *Trends Ecol. Evol*. 19: 101–108.

Jones R.I., Grey J., Sleeo D. & Arvola L. 1999. Stable isotope analysis of zooplankton carbon nutrition in humic lakes. *Oikos* 86: 97–104.

Jørgensen S.E. 2008. Overview of the model types available for development of ecological models. *Ecol. Model.* 215: 3–9.

Jørgensen S.E. 2010. A review of recent developments in lake modelling. *Ecol. Model.* 221: 689–692.

Keller W. 2007. Implications of climate warming for Boreal Shield lakes: a review and synthesis. *Environ. Rev.* 15: 99–112.

Kettunen J., Leonov A.V. & Varis O. 1987. Examination of model adequacy and analysis of phosphorus dynamics in lake Kuortaneenjärvi—A case study with two lake models. *Publications of the Water and Environmental Research Institute* 3: 48–54.

Kettunen J., Sirviö H. & Varis O. 1989. An analysis of a sampling design—A case study of lake eutrophication. *Comput. Statist. Data Anal.* 8: 81–91.

Kiili M., Pulkkanen M. & Salonen K. 2009. Distribution and development of under-ice phytoplankton in 90-m deep water column of Lake Päijänne (Finland) during spring convection. *Aquat. Ecol*. 43: 693–705.

Korhonen J.J., Wang J. & Soininen J. 2011. Productivitydiversity relationships in lake plankton communities. *PLoS ONE* 6: e22041, doi:10.1371/journal.pone.0022041.

Kruk C. & Segura A.M. 2012. The habitat template of phytoplankton morphology-based functional groups. *Hydrobiologia* 698: 191–202.

Lamon E.C. III & Qian S.S. 2008. Regional scale stressor-response models in aquatic ecosystems. *J. Am. Water Resour. As.* 44: 771–781.

Landsberg J.H. 2010. The effects of Harmful Algal Blooms on Aquatic Organisms. *Rev. Fish. Sci*. 10: 113–390.

Lawson L.M., Hofmann E.E. & Spitz Y.H. 1995. Time series sampling and data assimilation in a simple marine ecosystem model. *Deep-Sea Research II* 43: 625–651.

Lebret K., Fernandez M.F., Hagman C.H.C., Rengefors K. & Hansson L.A. 2012. Grazing resistance allows bloom formation and may explain invasion success of Gonyostomum semen. *Limnol. Oceanogr*. 57: 727–734.

Lehtoranta J. & Gran V. 2002. Sedimentistä veteen vapautuvat ravinteet Säkylän Pyhäjärvellä [Nutrient flux from sediment into water in Lake Säkylän Pyhäjärvi]. *Finnish Environment Institute Mimeographs* 247: 1–31. [In Finnish with English summary]

Lehtovaara A., Arvola L., Keskitalo J., Olin M., Rask M., Salonen K., Sarvala J., Tulonen T. & Vuorenmaa J. 2014. Responses of zooplankton to long-term environmental changes in a small boreal lake. *Boreal Env. Res*. 19: 97–111.

Lepistö L., Antikainen S. & Kivinen J. 1994. The occurrence of *Gonyostomum semen* (Her.) Diesing in Finnish lakes. *Hydrobiologia* 273: 1–8.

Liu Y., Guo H. & Yang P. 2010. Exploring the influence of lake water chemistry on chlorophyll *a*: a multivariate statistical model analysis. *Ecol. Model*. 221: 681–688.

Magnuson J.J., Robertson D.M., Benson B.J., Wynne R.H., Livingstone D.M., Arai T., Assel R.A., Barry R.G., Card V., Kuusisto E., Granin N.G., Prowse T.D., Stewart K.M. & Vuglinski V.S. 2000. Historical trends in lake and river ice cover in the Northern Hemisphere. *Science* 289: 1743–1746.

Malve O. 2007. *Water quality predictions for river basin management*. Doctoral Dissertation. Helsinki University of Technology, Helsinki.

Malve O. & Qian S.S. 2006. Estimating nutrients and chlorophyll *a* relationships in Finnish lakes. *Environ. Sci. Technol*. 40: 7848–7853.

Malve O., Laine M., Haario H., Kirkkala T. & Sarvala J. 2007. Bayesian modelling of algal mass occurrences—using adaptive MCMC methods with a lake water quality model. *Environ. Modell. Softw*. 22: 966–977.

Markensten H. & Pierson D.C. 2007. Weather driven influences on phytoplankton succession in a shallow lake during contrasting years: Application of PROTBAS. *Ecol. Model*. 207: 128–136.

Mason C. 2002. *Biology of freshwater pollution*. Pearson Education Limited, Essex.

Matthews B. & Pomati F. 2012. Reversal in the relationship between species richness and turnover in a phytoplankton community. *Ecology* 93: 2435–2447.

Mooij W.M., Trolle D., Jeppesen E., Arhonditsis G., Belolipetsky P.V., Chitamwebwa D.B.R., Degermendzhy A.G., DeAngelis D.L., De Senerpont Domis L.N., Downing A.S., Elliott J.A., Fragoso C.R. Jr., Gaedke U., Genova S.N., Gulati R.D., Håkanson L., Hamilton D.P., Hipsey M.R., Hoen J., Hülsmann S., Los F.H., Makler-Pick V., Petzoldt T., Prokopkin I.G., Rinke K., Schep S.A., Tominaga K., Van Dam A.A., Van Nes E.H., Wells S.A. & Janse J.H. 2010. Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquat. Ecol*. 44: 633–667.

Mowrer H.T. 2000. Uncertainty in natural resources decision support systems: sources, interpretation, and importance. *Comput. Electron. Agr*. 27: 139–154.

Naden P.S., Allott N., Arvola L., Järvinen M., Jennings E., Moore K., Nic Aonghusa C., Pierson D. & Schneiderman E. 2010. Modelling the impacts of climate change on dissolved organic carbon. In: George D.G. (ed.), *The Impact of Climate Change on European Lakes*, Springer, pp. 221–252.

Niemi J. 1979. Application of an ecological simulation model to Lake Päijänne. *Publications of the water research institute* 28: 1-39.

Nõges T. 2009. Relationships between morphometry, geographic location and water quality parameters of European lakes. *Hydrobiologia* 633: 33–43.

Nürnberg G.K., Tarvainen M., Ventelä A.-M. & Sarvala J. 2012. Internal phosphorus load estimation during biomanipulation in a large polymictic and mesotrophic lake. *Inland Waters* 2: 147–162.

Oreskes N., Shrader-Frechette K. & Belitz K. 1994. Verification, validation and confirmation of numerical models in the earth sciences. *Science*. 263: 641–646.

Otto S.P. & Day T. 2007. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, New Jersey.

Paerl H.W. & Huisman J. 2008. Blooms like it hot. *Science* 320: 57–58.

Paterson A.M., Cumming B.F., Smol J.P. & Hall R.I. 2004. Marked recent increases of colonial scaled chrysophytes in boreal lakes: implications for the management of taste and odour events. *Freshwat. Biol*. 49: 199–207.

Pearl J. 1995. Causal diagrams for empirical research. *Biometrika* 82: 669–688.

Peck S.L. 2004. Simulation as experiment: a philosophical reassessment for biological modeling. *Trends Ecol. Evol*. 19: 530-534.

Peltomaa E. & Ojala A. 2010. Size-related photosynthesis of algae in a strongly stratified humic lake. *J. Plankton Res*. 32: 341–355.

Peltomaa E., Ojala A., Holopainen A.-L. & Salonen K. 2013. Changes in phytoplankton in a boreal lake during a 14-year period. *Boreal Env. Res*. 18: 387–400.

Ptacnic R., Lepisto L., Willen E., Brettum P., Andersen T., Rekolainen S., Solheim A.L. & Carvalho L. 2008. Quantitative responses of lake phytoplankton to eutrophication in Northern Europe. *Aquat. Ecol.* 42: 227–236.

Ramin M., Labencki T., Boyd D., Trolle D. & Arhonditsis G.B. 2012. A Bayesian synthesis of predictions from different models for setting water quality criteria. *Ecol. Model*. 242: 127–145.

Rautio L.M. & Aaltonen E.-K. 2006. *Kuortaneenjärvi Lapuanjoen helmi*. Länsi-Suomen ympäristökeskus, Vaasa.

Refsgaard J.C., van der Sluijs J.P., Højberg A.L. & Vanrolleghem P.A. 2007. Uncertainty in the environmental modelling process—A framework and guidance. *Environ. Modell. Softw*. 22: 1543–1556.

Reynolds C.S. 2006. *Ecology of Phytoplankton*. Cambridge University Press, Cambridge.

Reynolds C.S., Irish A.E. & Elliott J.A. 2001. The ecological basis for simulating phytoplankton responses to environmental change (PROTECH). *Ecol. Model*. 140: 271–291.

Robson B.J. 2014. State of the art in modelling of phosphorus in aquatic systems: Review, criticisms and commentary. *Environ. Modell. Softw*. http://dx.doi.org/10.1016/j.envsoft.2014.01.012.

Robson B.J., Hamilton D.P., Webster I.T. & Chan T. 2008. Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries. *Environ. Modell. Softw*. 23: 369–384.

Rode M., Arhonditsis G., Balin D., Kebede T., Krysanova V., van Griensven A. & van der Zee S.E.A.T.M. 2010. New challenges in integrated water quality modelling. *Hydrol. Process*. 24: 3447–3461.

Romero J.R., Antenucci J.P. & Imberger J. 2004.One- and three-dimensional biogeochemical simulations of two different reservoirs. *Ecol. Model*. 174: 143–160.

Rykiel E.J. Jr. 1996. Testing ecological models: the meaning of validation. *Ecol. Model*. 90: 229–244.

Salmaso N. 2011. Interactions between nutrient availability and climatic fluctuations as determinants of the long-term phytoplankton community changes in Lake Garda, Northern Italy. *Hydrobiologia* : 660: 59–68.

Salmaso N., Buzzi F., Garibaldi L., Morabito G. & Simona M. 2012. Effects of nutrient availability and temperature on phytoplankton development: a case study from large lakes south of the Alps. *Aquat. Sci.* 74: 555–570.

Salonen K. & Rosenberg M. 2000. Advantages from diel vertical migration can explain the dominance of Gonyostomum semen (Raphidophyceae) in a small, steeply-stratified humic lake. *J. Plankton Res*. 22: 1841–1853.

Saloranta T.M., & Andersen T. 2007. MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecol. Model*. 207: 45–60.

Scheffer M. & Beets J. 1994. Ecological models and the pitfalls of causality. *Hydrobiologia*. 275/276: 115–124.

Scheffer M. & Carpenter S.R. 2003. Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends Ecol. Evol*. 18: 648–656.

Shatwell T., Köhler J. & Nicklisch A. 2008 Warming promotes cold-adapted phytoplankton in temperate lakes and opens a loophole for Oscillatoriales in spring. *Glob. Change Biol*. 14: 2194–2200.

Shipley B. 2002. *Cause and correlation in biology: A user's guide to path analysis, structural equations and causal inference*. Cambridge University Press, Cambridge.

Soininen J., Tallberg P. & Horppila J. 2005. Phytoplankton community assembly in a large boreal lake—deterministic pathways or chaotic fluctuations? *Freshwat. Biol.* 50: 2076–2086.

Stomp M., Huisman J., Mittelbach G.G., Litchman E. & Klausmeier C.A. 2011. Large-scale biodiversity patterns in freshwater phytoplankton. *Ecology* 92: 2096–2107.

Sutherland W.J. 2006. Predicting the ecological consequences of environmental change: a review of the methods. *J. Appl. Ecol*. 43: 599–616.

Trigal C., Hallstan S., Johansson K.S.L. & Johnson R.K. 2013. Factors affecting occurrence and bloom formation of the nuisance flagellate Gonyostomum semen in boreal lakes. *Harmful Algae* 27: 60–67.

Trolle D., Hamilton D.P., Hipsey M.R., Bolding K., Bruggeman J., Mooij W.M., Janse J.H., Nielsen A., Jeppesen E., Elliott J.A., Makler-Pick V., Petzoldt T., Rinke K., Flindt M.R., Arhonditsis G.B., Gal G., Bjerring R., Tominaga K., Hoen J., Downing A.S., Marques D.M., Fragoso C.R. Jr., Søndergaard M. & Hanson P.C. 2012. A community-based framework for aquatic ecosystem models. *Hydrobiologia* 683: 25–34.

Väisänen S. (ed.) 2013. Mallit avuksi vesienhoidonsuunnitteluun GisBloom-hankkeen pilottialueilla [Models in support of river basin management planning in GisBloom pilot areas, Guidance document of applying the demonstrated methods in river basin management planning]. *Reports of the Finnish Environment Institute* 29 :1-188. [In Finnish with English abstract].

Vehmaa A. & Salonen K. 2009 Development of phytoplankton in Lake Paajarvi (Finland) during under-ice convective mixing period. *Aquat. Ecol*. 43: 707–713.

Ventelä A.-M., Tarvainen M., Helminen H. & Sarvala J. 2007. Long-term management of Pyhäjärvi (southwest Finland): eutrophication, restoration—recovery? *Lake Reservoir Manage.* 23: 428–438.

Ventelä A.-M., Kirkkala T., Lendasse A., Tarvainen M., Helminen H. & Sarvala J. 2011. Climate-related challenges in long term management of Säkylän Pyhäjärvi (SW Finland). *Hydrobiologia* 660: 49–58.

Virtanen M. 2009. *Mathematical modelling of flow and transport as link to impacts in multidiscipline environments*. Doctoral dissertation, University of Oulu, Oulu.

Vollenweider R.A. 1968. *The scientific basis of lake and stream eutrophication, with particular reference to phosphorus and nitrogen as eutrophication factors*. Paris: Organization for Economic Co-operation and Development. Technical report: DAS/CSI/68.27.

Vuorenmaa J., Arvola L. & Rask M. (eds.) 2011. Hämeen ympäristö muutoksessa, Kaksikymmentä vuotta ympäristön huippututkimusta Valkea-Kotisen alueella [Long-term environmental change in Häme, Twenty years of environmental research and monitoring at Valkea-Kotinen supersite]. *The Finnish Environment* 34: 1-104. [In Finnish with English abstract]

Vuorio K., Järvinen M., Moilanen S. & Bilaletdin Ä. 2013. Piilevien aiheuttamat kalanpyydysten likaantumisongelmat—Pyhäjärven eteläosa, Kulovesi ja Roine [Diatom induced fouling of fishing nets—Lakes Pyhäjärvi, Kulovesi and Roine]. *Reports of the Finnish Environment Institute* 22: 1-37. [In Finnish with English abstract]

Wagner C. & Adrian R. 2009. Cyanobacteria dominance: Quantifying the effects of climate change. *Limnol. Oceanog*. 54: 2460–2468.

Wagner T., Soranno P.A., Webster K.E. & Spence Cheruvelil K. 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes. *Freshwat.Biol*. 56: 1811–1824.

White S.H., Duivenvoorden L.J. & Fabbro L.D. 2005. Impacts of a toxic Microcystis bloom on the macroinvertebrate fauna in Lake Elphinstone, Central Queensland, Australia. *Hydrobiologia* 548: 117–126.

Whittingham M.J., Stephens P.A., Bradbury R.B. & Freckleton R.P. 2006. Why do we still use stepwise modelling in ecology and behaviour? *J. Anim. Ecol*. 75: 1182–1189.

Zuur A.F., Ieno E.N. & Elphick C.S. 2010. A Protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*. 1: 3–14.

Zuur A.F., Ieno E.N., Walker N.J., Saveliev A.A. & Smith G. 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.

ORIGINAL PAPERS

I

MODELLING THE IMPACT OF HIGHER TEMPERATURE ON THE PHYTOPLANKTON OF A BOREAL LAKE

by

Anita Pätynen, J. Alex Elliott, Petri Kiuru, Jouko Sarvala, Anne-Mari Ventelä & Roger I. Jones 2014

Boreal Environment Research 19: 66–78.

II

ALTERNATIVE APPROACHES TO MODELLING LAKE ECOSYSTEMS

by

Anita Pätynen, Niina Kotamäki & Olli Malve 2013

Freshwater Reviews 6: 63-74.

III

CAUSAL ANALYSIS OF PHYTOPLANKTON DEVELOPMENT IN A SMALL HUMIC LAKE USING STRUCTURAL EQUATION MODELLING

by

Anita Pätynen, Niina Kotamäki, Lauri Arvola, Tiina Tulonen & Olli Malve 2014

Submitted manuscript

IV

ENHANCING LAKE MANAGEMENT WITH MODEL BASED ESTIMATES OF TARGET NUTRIENT LOAD REDUCTION

by

Niina Kotamäki, Anita Pätynen, Antti Taskinen, Timo Huttula & Olli Malve
2014

Manuscript