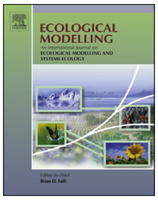
Ecological Modelling 357 (2017) 1–13

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Constraining uncertainty and process-representation in an algal community lake model using high frequency in-lake observations

Trevor Page[a,*], Paul J. Smith[a,c], Keith J. Beven[a], Ian D. Jones[b], J. Alex Elliott[b], Stephen C. Maberly[b], Eleanor B. Mackay[b], Mitzi De Ville[b], Heidrun Feuchtmayr[b]

[a] Lancaster Environment Centre,Library Avenue, Lancaster University, Lancaster, LA1 4YQ, UK
[b] Lake Ecosystems Group, Centre for Ecology & Hydrology, Lancaster Environment Centre,Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK
[c] ECMWF, Shinfield Park, Reading, RG2 9AX, UK

ARTICLE INFO

Article history:
Received 8 February 2017
Received in revised form 14 April 2017
Accepted 14 April 2017
Available online 12 May 2017

Keywords:
Algal bloom
Forecasting
GLUE
PROTECH model
Uncertainty

ABSTRACT

Excessive algal blooms, some of which can be toxic, are the most obvious symptoms of nutrient enrichment and can be exacerbated by climate change. They cause numerous ecological problems and also economic costs to water companies. The process-representation of the algal community model PROTECH was tested within the extended Generalised Likelihood Uncertainty Estimation framework which includes pre-defined Limits of Acceptability for simulations. Testing was a precursor to modification of the model for real-time forecasting of algal communities that will place different demands on the model in terms of a) the simulation accuracy required, b) the computational burden associated with the inclusion of forecast uncertainties and c) data assimilation. We found that the systematic differences between the model's representation of underwater light compared to the real lake systems studied and the uncertainties associated with nutrient fluxes will be the greatest challenges when forecasting algal blooms.

1. Introduction

Algal blooms are a globally significant problem affecting water resources, recreation and ecosystems (Carmichael, 1992; Smith, 2003; World Health Organization, 1999). These problems are particularly acute when blooms include significant cyanobacteria populations as some species can produce toxins that cause adverse health effects to humans and affect wildlife (Metcalf and Codd, 2009). Water companies face associated problems such as blocked filters, poor taste and odour and, in more extreme cases, high levels of algal-derived toxins. Managing these effects costs greater than £50 million per year in the UK (Pretty et al., 2003) and billions of dollars annually in the US (Dodds et al., 2009; Michalak, 2016). Implementation of mitigation strategies is becoming more expensive owing to increases in the frequency of blooms (Ho and Michalak, 2015) as a result of nutrient enrichment and climate change (Brookes and Carey, 2011; Paerl and Huisman, 2008; Rigosi et al., 2014) and the effectiveness of interventions is, in some cases, being compromised. It is therefore beneficial to be able to forecast algal blooms to allow the most cost-effective management strategies to be implemented.

One algal model that has been used in lakes and reservoirs around the world is PROTECH (Elliott et al., 2009; Elliott, 2010, 2012; Reynolds et al., 2001). PROTECH was used here because it explicitly simulates the dynamics of lake algal community structure and hence algal types of particular interest including cyanobacteria. As real-time forecasting of algal blooms is becoming a priority for the management of lakes and reservoirs used for water supply and recreation, one of the aims of this study is to test the model as a precursor to modification for forecasting purposes. Real-time forecasting places different demands on the model in terms of the accuracy and resolution required for simulation estimates, the computational burden associated with the inclusion of forecast uncertainties and in the way that data assimilation of observations is structured. Access to high-frequency data does, however, provide opportunities to improve model process-representation consistent with these requirements. The sensitivity of the PROTECH phytoplankton growth equations has been assessed and was shown to be robust (Elliott et al., 1999); consequently, in this study, we primarily consider the model's abiotic environment, including water temperature, underwater light, mixing processes and nutrient input dynamics. Sensitivity and uncertainty analyses were carried out within a hypothesis testing framework where different

---

[*] Corresponding author.
E-mail address: t.page@lancaster.ac.uk (T. Page).

http://dx.doi.org/10.1016/j.ecolmodel.2017.04.011
0304-3800/© 2017 Elsevier B.V. All rights reserved.

model representations were considered as competing hypotheses and accepted or rejected based upon specific criteria. This was achieved using the extended Generalised Likelihood Uncertainty Estimation Framework (GLUE; Beven and Binley, 1992) where the criteria for acceptance are formalised *Limits of Acceptability* (LoA) for model simulations (GLUE-LoA; Beven, 2006, 2012; Beven and Binley, 2014; Blazkova and Beven, 2009; Liu et al., 2009). Hypotheses are tested under this approach where interactions between the uncertainties arising from model structural components, parameters, model inputs and observations used for model constraint are taken into account. Using LoA has the advantages that explicit representation can be made for the variability of errors (e.g. non-stationary/state-dependent errors and correlation of errors) at individual observation times and/or locations and is a natural way to combine different types of observation. This approach is critically important for focussing on how different sources of uncertainty determine model acceptability, affect the assessment of modelling hypotheses and inform strategies used when implementing the model to make predictions.

2. Methods

2.1. Study lakes

The study area is located in the English Lake District of North West England which is a hilly region with a landscape and lakes shaped by glaciation. The land use is predominantly upland unimproved grassland, grazed by sheep and the region is extremely popular with tourists throughout the year, particularly during summer. The three study lakes, Windermere, Bassenthwaite Lake and Esthwaite Water, are among the best studied lakes in the world (Maberly and Elliott, 2012) and differ in area, depth, extent of summer stratification, hydraulic residence times and trophic state (Fig. 1; Table 1). For more information see Talling (1999); Reynolds and Irish (2000); Maberly et al. (2010) Mackay et al. (2014). In this study for Windermere we simulate only the South Basin of Windermere rather than the whole lake. It receives inputs directly from the larger North Basin and indirectly from Esthwaite Water via Cunsey Beck. For this study, simulations were made for six lake-years where high resolution and high quality data were available: 2008–2010 for Windermere, 2008 and 2009 for Esthwaite and 2010 for Bassenthwaite.

2.2. The PROTECH model

2.2.1. General description

PROTECH (Reynolds et al., 2001) is an algal community lake model that runs on a daily time-step. It is a 1-D model where the lake is represented by 0.1 m horizontal layers each with a volume calculated by interpolation of lake bathymetric data. The model has routines which calculate stratification and destratification and determine the depth to the top of the thermocline for each time step. In the model representation, the top of the thermocline is considered the depth at which all layers above are fully mixed: referred to as the *mixed depth* for the purposes of this study. The layers from the surface to the mixed depth are treated as homogeneous and are instantaneously mixed at each time step. The model also has the ability to represent vertical eddy diffusion fluxes (of energy and nutrients; see Elliott and Thackeray, 2004) which is particularly important for simulating the behaviour of lakes with significant sediment-derived internal P fluxes. Eddy diffusion is represented using a simplified function where groups of model layers (*metalayers of depth* ML_d) are homogenized and mixing occurs across the boundary between them (Eq. (1)). The degree of mixing is specified by an eddy diffusivity parameter (K_z) that is assigned a fixed value for the duration of a simulation and is used to calculate the flux (F) of a given substance (j) for metalayer n using:

$$F_{n,j} = \frac{K_z}{z_n - z_{n-1}} \cdot \frac{C_n - C_{n-1}}{A} \tag{1}$$

Where: A is the area of the plane of contact between metalayers, z is the depth at the centre of each metalayer and C is the mean concentration of the metalayer in question.

River inputs drive fluxes of diffuse nutrients as well as the flushing of algae. Riverine inputs include algal *inocula* which are set to a ‘background’ chlorophyll *a* concentration for the time of year; for each day this inocula is distributed equally across the species simulated. Upstream lake inputs are added proportionally (using proportion of overall catchment area drained) to river inputs but are given the algal concentrations associated with the upstream lake, where it is possible to represent them.

Underwater light for model layer i, li, is calculated using:

$$l_i = I_{surf} . e^{(-\varepsilon . d_i)} \tag{2}$$

Where: Isurf is the daily surface light flux (see Reynolds et al., 2001), di is the depth from the lake surface, ε is the light extinction coefficient resulting from the sum of lake-specific abiotic extinction (εb; *a model parameter which is fixed for the duration of a simulation*) and the extinction of light associated with the concentration of algae at each time-step multiplied by the εa parameter

In the layers from the surface to the mixed depth, the light is *averaged* (using the geometric mean) to represent the amount of light to which algae are exposed. This averaging is based on the assumption that the algae spend an equal time in each layer down to the mixed depth for the duration of the time step.

Once the environment for algal growth of each layer is determined, algal population dynamics are simulated using the following state variable equation which describes the change in chlorophyll *a* concentration (X) of each algal species considered (Reynolds 1988):

$$\frac{\Delta X}{\Delta t} = \left(r' - S - G - D\right) . X \tag{3}$$

where r′ is the growth rate, S is the settling loss, G is the grazing loss and D is the loss caused by flushing. The growth rate (r') is defined for each layer using:

$$r' = \min\left\{r'_{(\theta)}, r'_{(P)}, r'_{(N)}, r'_{(Si)}\right\} \tag{4}$$

where $r'_{(\theta,I)}$ is the growth rate at a given temperature (θ) and daily photoperiod (I) and r'_{P}, r'_{N}, r'_{Si} are the growth rates determined by phosphorous, nitrogen and silica concentrations. The final growth rate ($r'_{cor(\theta,I)}$) is a corrected rate allowing for dark respiration using equation 5. This is required as the model growth equations are net of basal metabolism but not dark respiration burden.

$$r'_{corr(\theta,I)} = R_{d(\theta)} \cdot r'_{(\theta,I)} - \left(1 - R_{d(\theta)}\right) . r'_{(\theta,I)} \tag{5}$$

Where $R_{d(\theta)}$ is the dark respiration rate at temperature θ.

2.2.2. Simulating the dynamics of algal species

PROTECH simulates the dynamics of the species chosen to represent the algal community of a given lake. Species are represented by their morphology, nutrient requirements (i.e. silica requirement and nitrogen fixing ability) and their vertical movement strategies. The number of species simulated is nominally eight (although unlimited) and are chosen to represent the dominant functional types of the system of interest (see Table Supp. 2). Modelling results are thus primarily interpreted on the basis of the behaviour of the functional algal community rather than the dynamics of specific species simulated, to avoid overconstraint on the specific