Article (refereed) - postprint

Page, Trevor; Smith, Paul J.; Beven, Keith J.; Jones, Ian D.; Elliott, J. Alex; Maberly, Stephen C.; Mackay, Eleanor B.; De Ville, Mitzi; Feuchtmayr, Heidrun. 2017. **Constraining uncertainty and process-representation in an algal community lake model using high frequency in-lake observations**.

# 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] and 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*

[**]*Corresponding author: t.page@lancaster.ac.uk*

**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 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 & Irish (2000); Thackeray *et al*. (2006); Maberly *et al*. (2011) 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.

**Table 1.** Primary characteristics of the study lakes.

| Lake | Area ($km^2$) | Volume ($Mm^3$) | Mean depth (m) | Max depth (m) | Catchment area ($km^2$) | Mean residence time (days) | Trophic state |
|---|---|---|---|---|---|---|---|
| Windermere South Basin | 6.7 | 113 | 16.8 | 42 | 231 | 100 | Mesotrophic |
| Esthwaite Water | 1.0 | 6.7 | 6.9 | 15.5 | 17 | 100 | Eu-mesotrophic |
| Bassenthwaite Lake | 5.3 | 28 | 5.3 | 19 | 360 | 30 | Meso-eutrophic |

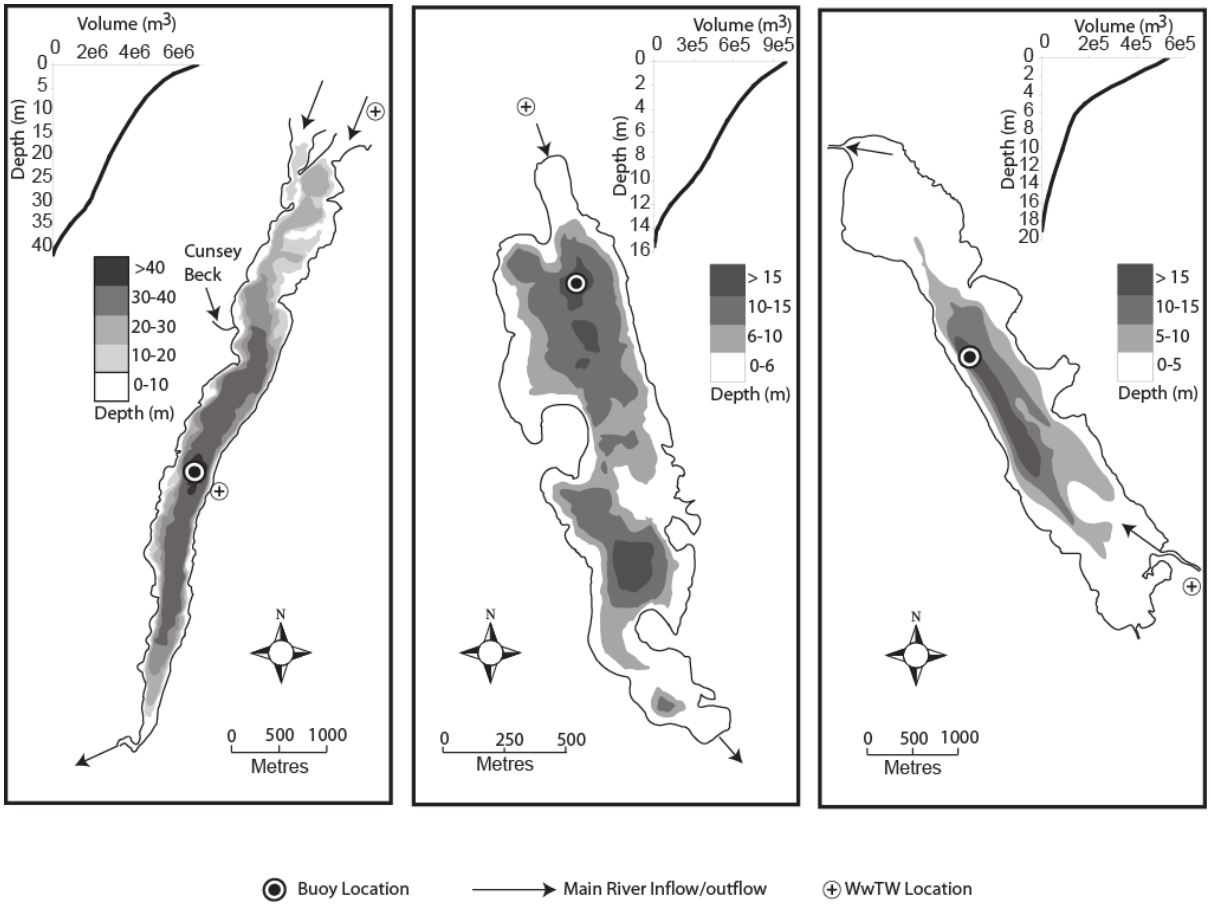


***Figure 1.*** *Bathymetric map and inset hyposographic curve for (a) Windermere South Basin*, (b) Esthwaite Water** and (c) Bassenthwaite Lake*. * Redrawn from Ramsbottom, 1976; ** Redrawn from Mackay et al., 2012.*

## 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 (Eqn. 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:

$$F_{n,j} = \frac{K_z}{z_n - z_{n-1}} . \frac{C_n - C_{n-1}}{A} \quad (1)$$

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 $i$, $l_i$, is calculated using:

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

Where: $I_{surf}$ is the daily surface light flux (see Reynolds et al., 2001), $d_i$ is the depth from the lake surface, $\varepsilon$ is the light extinction coefficient resulting from the sum of lake-specific abiotic extinction ($\varepsilon_b$*; 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 parameter $\varepsilon_a$.

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):

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

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:

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

where $r'_{(\theta,l)}$ is the growth rate at a given temperature *(θ)* and daily photoperiod (*l*) and $r'_P$, $r'_N$, $r'_{Si}$ are the growth rates determined by phosphorous, nitrogen and silica concentrations. The final growth rate ($r'_{cor(\theta,l)}$) 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.

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

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

### 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 species chosen. The C-S-R functional phytoplankton classification of Reynolds (1988) is used to classify phytoplankton into morphologically defined groups relating to broad ecological strategies. The primary groups are: C-types, which are invasive, ecological pioneers that are small with high surface-to-volume ratios (e.g. *Chlorella*, and *Plagioselmis*); S-types which are 'stress tolerators' that tolerate relatively low nutrient availability and strong stratification (e.g. *Woronichinia, Microcystis* and *Oocystis)*; and R-types which can harvest sufficient light at low levels to be able to maintain growth and are hence tolerant of well-mixed, intermittently insolated environments (e.g. *Asterionella*, *Aulacoseira and Oscillatoria)*. Also important for the lakes studied here, are CS-types, whose characteristics are intermediate between those of C and S species (e.g. *Dolichospermum, Aphanizomenon* and *Ceratium*) and CSR-types (e.g*. Cryptomonas*) that are intermediate between C-, S- and R-types.

## 2.3 Modelling Hypotheses and scenarios

Two hypotheses were tested to improve the model: 1) a modification of the method for estimating mixed depth, which affects the light climate for algae and hence population dynamics and 2) A modified relationship between phosphorus concentrations and river inflow magnitude aimed at improving the timing of phosphorus fluxes and subsequently algal community dynamics.

### 2.3.1 Hourly estimation of mixed depth

In model space, the degree to which algae are exposed to light is necessarily a simplification and controlled primarily by the mixed depth, the way light is "averaged" throughout the mixed depth and the light extinction coefficient ($\varepsilon$). Under the standard model formulation used here, the mixed depth is estimated using daily averaged (of hourly) temperature depth profiles using a density gradient method (Read *et al*., 2011). However, analysis of the hourly temperature profile data showed that there were periods of *temporary stratification* that were not captured by daily averages, especially during the onset of stratification, (Fig. 2). We therefore postulated that ***the daily mixed depth estimate is represented better by the distribution (or a moment of the distribution) of hourly mixed depth estimates for a given day rather than an estimate using the daily averaged temperature profile***.

### 2.3.2 Inflow-dependent phosphorus inputs

The ***standard model*** representation of diffuse soluble reactive phosphorus (SRP)-river inflow relationships identified here (see section 2.4.2.1 for a description) were developed using all the available nutrient data at monthly resolution. It is well-known, however, that low resolution routine monitoring tends to underestimate P concentrations at high flows for river-catchments where diffuse sources dominate and where there tends to be an increase in concentration with flow (e.g. see Johnes, 2007; Cassidy and Jordan, 2011). In the case of Windermere South Basin, approximately 85% of the flow- dependent P inputs are delivered via its North Basin, any modification to the diffuse SRP- inflow relationship will implicitly include effects from misrepresentation of upstream lake P inputs. We hypothesised that ***diffuse SRP concentrations are linearly related to inflow magnitude***; a description of the implementation of this hypothesis is provided in section 2.4.2.1 below.

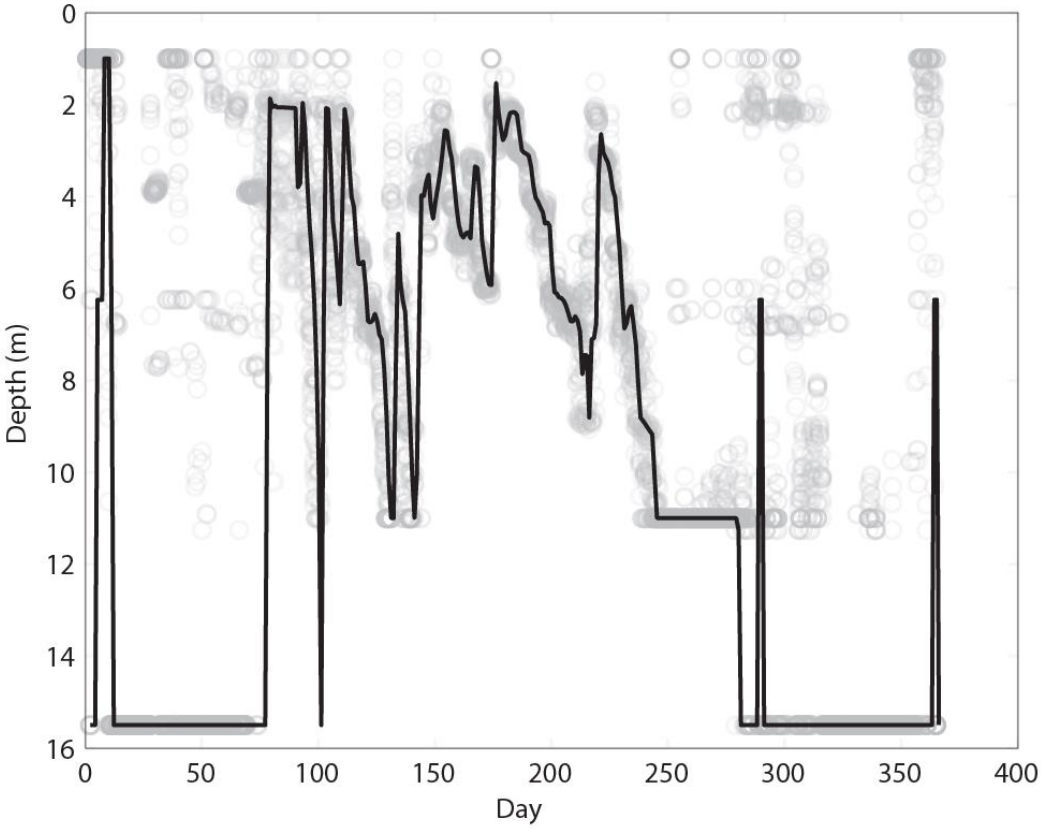


***Figure 2.*** *Comparison of the standard model formulation mixed depth estimates based upon daily averaged temperature profiles (black line) with individual hourly mixed depth estimates for the same day (grey circles) for Esthwaite Water 2009. The distribution of hourly estimates for each day was sampled to provide a modified representation of the daily depth for the modelling scenarios (Table 2).*

### 2.3.3 Modelling scenarios

The factorial combination of the two time resolutions for mixed depth and treatments of SRP input led to four scenarios (Table 2).

**Table 2.** Modelling scenarios.

| Scenario | Mixed depth | P-inflow |
|---|---|---|
| S1 | Daily average | Standard model representation |

| | | |
|---|---|---|
| S2 | Daily average | Hypothesised representation |
| S3 | Sampled from hourly dist. | Standard model representation |
| S4 | Sampled from hourly dist. | Hypothesised representation |

## 2.4 Modelling methodology

Here we assess PROTECH under the GLUE-LoA methodology. The philosophy underlying GLUE recognises that given the significant uncertainties associated with modelling environmental systems there will be multiple model structures and parameter set combinations that provide 'acceptable' simulations (the *equifinality* thesis; Beven, 2006). As parameter sets (rather than individual parameter values) and different model structures are evaluated, interaction between parameters and structures that lead to acceptable simulations is implicitly taken into account. The use of GLUE with explicit LoA takes into account uncertainties associated with input and evaluation data, as well as incommensurability (e.g. the mismatch between variables in model space and those observed in the real system) such that models that might be useful in prediction are not falsely deemed unacceptable (Beven, 2006, 2012; Blazkova and Beven, 2009; Liu *et al*., 2009). LoAs are absolute ranges, associated with specified criteria, within which simulation outputs are required to fall to be deemed acceptable and which should ideally be defined *a priori.* The rationale used in deriving the LoA for each lake-year considered here is described in detail below.

Monte Carlo sampling was employed to explore the model parameter space from *a priori* defined ranges for each parameter (Table Supp. 1). Where no information is available regarding the prior probability distributions of parameters, a uniform distribution was sampled. Where prior knowledge about parameter distributions and covariation of parameters is known it can be incorporated within the sampling strategy. For each simulation, model performance was assessed by LoA (as discrete acceptance criteria) as well as a *likelihood measure* or *weighting* which expresses the degree of fit to the evaluation data. The likelihood measures used for this study are specified below (Eqns. 8-14). Models deemed unacceptable based on the LoA were rejected and played no further part in the

analysis. All acceptable simulations were used in the generation of likelihood-weighted uncertainty bounds using:

$$P(\hat{Z}_t < z_t) = \sum_{j=1}^{j=N} L\left[ M(\Theta_j) \mid \hat{Z}_{t,j} < z_t \right] \qquad (6)$$

where $P$ is the prediction quantile for $\hat{Z}$ (the value of variable *Z* at time *t* simulated by model $M(\Theta_j)$) being less than *z*, *L* is the likelihood weighting (a scaled form of eqn. 14 such that all weightings sum to unity) associated with model $M(\Theta_j)$, $\Theta_j$ is the $j^{th}$ parameter set and *N* is the number of acceptable models.

### 2.4.1 Sampling model parameters

The model parameters for each lake and scenarios and their ranges for the uncertainty analysis where these were varied are show in Table Supp. 1. For parameters that were varied, Monte Carlo sampling from uniform distributions was employed. The parameters varied were those shown to be the most sensitive from previous unpublished work, past analyses (e.g. see Elliott *et al*., 1999) and initial simulations undertaken for the present study. These include those which determine the source, magnitude and dynamics of nutrient inputs, the representation of underwater light and the magnitude of eddy diffusion between metalayers as described above. For each of the scenarios and for each lake-year considered 100,000 simulations were carried out.

### 2.4.2 Nutrient inputs

All three lakes are impacted by diffuse nutrient sources as well as significant point sources of P from WwTW. Additionally, Esthwaite Water is known to be affected by significant internal sources of P (Mackay *et al*., 2014) but it was assumed, for the purposes of this study, that as Windermere and Bassenthwaite Lake were unlikely to be anoxic during the study period they were not subject to significant internal P releases.

#### 2.4.2.1 Diffuse nutrient inputs

Measured lake outflows for each lake were available from the United Kingdom Environment Agency (National River Flow Archive: http://www.ceh.ac.uk/data/nrfa/) at a daily resolution. Inflows were assumed to equal outflows and were treated in a deterministic manner. The ***standard model*** treatment of diffuse nutrient inputs for all scenarios, for Si and $NO_3$-N and S1 and S3 for SRP is as follows. All available nutrient concentrations (for all rivers where nutrient data were available) were associated with the lake outflow magnitude the

observation day. A “regionalised” relationship was the developed using discrete outflow magnitude classes, to which a Gamma distribution was fitted to all concentrations associated within that outflow class. At each simulation time step, each nutrient concentration was sampled from its respective Gamma distribution of the flow class associated with the observed daily flow. The magnitude of the inputs was also modified using a multiplier which was constant for the duration of each simulation (parameter $P_{fact}$: Table Supp. 1).

Where diffuse P inputs were treated differently under the hypothesized scenarios S2 and S4, input concentrations for each time step (*Pi*) were estimated using the flow-proportional relationship:

$$P_i = P_{min} + \frac{Q_i}{Q'} P_{max} \qquad (7)$$

where $P_{min}$ and $P_{max}$ are parameters which define the minimum and maximum P concentration, $Q_i$ is the inflow at timestep *i* and $Q'$ is a normalising flow value (set to the mean of the years of interest as a first approximation).

#### 2.4.2.2 Upstream lake inputs

Windermere South Basin is subject to significant upstream lake inputs (estimated to be approximately 85% of the catchment area-weighted inflow), primarily from Windermere North Basin. Data were only available for Windermere and upstream lake inputs for other lakes were represented by the inflow-dependent nutrient relationships. For Windermere, upstream lake inputs of SRP, $NO_3$-N, Si and Chlorophyll *a* were sampled from a distribution for each day of the year. The day-specific distribution was developed using data from the fortnightly long-term monitoring record (2006-2012). As multiple observations were not available for each day of the year, concentrations for that day were represented by observations within a ‘moving window’ of 20 days and a Gamma distribution was fitted to all points within the window.

#### 2.4.2.3 Wastewater treatment works P inputs

Inputs of SRP from wastewater treatment works were treated in the same way as upstream lake inputs but as a mass per day and were modified using a multiplier (parameter $WwTW_{fact}$: Table Supp. 1). The only data available were for P inputs to Windermere for the years 2002 to 2007 (Maberly and Elliott, 2009). Distributions for Esthwaite Water and Bassenthwaite Lake were scaled (using approximate population statistics) versions of those developed for Windermere so that the seasonality of inputs associated with tourist populations was retained.

#### 2.4.2.4 Internal lake P fluxes

Hypolimnetic and epilimnetic SRP fluxes were considered only for Esthwaite Water. As a way of constraining the hypolimnetic P fluxes we used year-specific estimates from observations and calculations reported by Mackay *et al*. (2014). These observations included the temporal dynamics of oxygen depletion and SRP concentrations at depths of 0.5 m, 11 m, and 14 m for SRP and every 1 m for oxygen concentration. Given an estimated depth to deoxygenated waters (assumed to be below 1 g $m^{-3}$ of dissolved oxygen) the sediment area in contact with deoxygenated water was calculated at each time step. It was assumed that SRP was released when waters were deoxygenated and was treated as a threshold without varying degrees of release. The mass of SRP released into the hypolimnion was estimated using the bed area associated with deoxygenated waters and a parameter specifying the mass of SRP released per $m^2$ ($\boldsymbol{P_{Hypo}}$; Table Supp.1) which was fixed for the duration of any given simulation. Epilimnetic P inputs were included using the method of Mackay *et al*. (2014) who employed SRP release estimates, of 0.46 mg $m^{-2}$ $d^{-1}$ based on Steinman *et al*. (2009). This value was modified by the parameter ($\boldsymbol{P_{Epi}}$; Table Supp.1) used in conjunction with the epilimnetic bed area (calculated each day within the model) to provide a daily mass input to the mixed layer.

### 2.4.3 *Lake temperature and mixed depth estimates*

Under the United Kingdom Lake Ecological Observatory Network (UKLEON) project (http://www.ceh.ac.uk/our-science/projects/uk-lake-ecological-observatory-network-ukleon) high frequency (4 minute) observations are being collected using a network of automatic lake monitoring systems including those associated with the buoy located at the three study lakes (Fig. 1). These included, among other variables, a meteorological station on the station and a thermistor chain.

The temperature in each vertical layer of the model and the mixed depth were estimated using the high frequency buoy observations. At each time step, either hourly or daily depending upon the scenario, the mixed depth was estimated using thermistor chain data and a critical density gradient method (Read *et al*., 2011) which identifies the thermocline based upon a critical water density gradient ($\Delta\rho$).

Mixed depth sampling for S1 and S2 utilised mixed depth estimated from average daily lake temperature data. The uncertainty associated with the mixed depth estimates was taken into account by varying $\Delta\rho$. The minimum, best estimate and maximum $\Delta\rho$ (0.05, 0.1 and 0.15 respectively) were used to estimate 3 mixed depths for each simulation day. Additional uncertainties (associated with representation of the whole lake by the buoy location and

temperature interpolation error) were estimated to be +/- 0.5 m of the estimated mixed depth as a first approximation. For each day, an estimate of the mixed depth was sampled randomly from the range described above. The mixed depth estimate was correlated to the previous day's sample with the correlated random number ($Rmd_i$) using:

$$Rmd_i = Rmd_i \times C + Rmd_{i-1} \times (1 - C) \quad (8)$$

where, *i* denotes timestep and *C* is a correlation coefficient (nominally set to 0.75). Each mixed depth estimate ($M_{di}$) at each timestep was weighted using a triangular fuzzy membership function (Eqn. 9) constructed from the range of likely mixed depths described above.

$$W_i = 1 - \left(\frac{(E_i - Md_i)}{E_{\max} - E_i}\right) : where\ E_i > M_{di}$$

$$W_i = 1 - \left(\frac{(Md_i - E_i)}{E_i - E_{min}}\right) : where\ E_i < M_{di} \quad (9)$$

$$W_i = 1 : where\ M_{di} = E_i$$

where: $W_i$ is the individual weighting for timestep *i*, $E_i$ is the expected value of mixed depth and $E_{min}$ and $E_{max}$ are the minimum and maximum of the fuzzy range. The overall weight ($W_s$) is the mean of all *N* weights:

$$W_s = \frac{\sum_i^N (W_i)}{N} \quad (10)$$

and is an *a priori* weighting that represents the confidence in the sequence of mixed depth estimates for a given simulation. This weighting was combined with simulation performance and propagated to the results of the uncertainty analysis using Eqn. 14 below.

For the scenarios utilising hourly temperature data (S3 and S4**)**, mixed depth was estimated for each hour of each day using the density gradient method to provide a distribution of hourly mixed depths for each day. This distribution was sampled using an additional parameter (*Mp*; Table Supp.1) specifying the percentile of the distribution to be used for the duration of each simulation.

### *2.4.4 Choosing the simulated phytoplankton*

The taxa chosen to represent the algal community for each lake-year considered were the top 8 species observed (ranked by biovolume magnitude; see Table Supp.2). In each case, the sum of the biovolumes of the species chosen was greater than 90% of the total annual biomass.

### 2.4.5 Evaluation of simulations and defining Limits of Acceptability

The *initial* LoA were defined *a priori* using the available data and literature sources together with uncertainty estimates elicited from experts associated with the UKLEON project. The *initial* LoA did not explicitly include allowance for the uncertainty associated with nutrient inputs as these inputs were modified by parameters to be constrained under GLUE-LoA.

#### 2.4.5.1 Initial Limits of Acceptability

Chlorophyll *a* observations were the primary modelling constraint. Each observation is derived from a water sample integrated over 0-5 m depth (Esthwaite Water and Bassenthwaite Lake) or 0-7 m depth (Windermere) (see Maberly *et al*., 2010) collected at the buoy location (see Fig. 1). There are three primary sources of uncertainty associated with the chlorophyll *a* observations: sampling error associated with the integrated water samples themselves, analytical error associated with the laboratory-based chlorophyll measurement and the error associated with in-lake spatial and temporal variability. We estimated the **sampling/analytical error** to be approximately +/- 8% using data from replicate samples taken under UKLEON combined with published estimates (Knowlton *et al*., 1984 and Mackay *et al*., 2011). The uncertainty associated with spatial heterogeneity is more difficult to estimate and varies over time (Elliott and Defew, 2012) and between species (e.g. wind-blown cyanobacteria species can be particularly heterogeneous: George and Heaney, 1978); we estimated the **overall error** to be in the order of +/-25%.

The model was also constrained using algal community structure (also collected at the buoy location: Fig. 1). To avoid over constraint, both observations and simulations of algal species were represented as functional algal types (R-types and CS-types), rather than individual species; the use of individual species has the potential to spuriously reject simulation because…….. Constraining simulations on functional type does retain our ability to reject simulations that may achieve acceptable chlorophyll *a* concentrations, but which do not simulate well the dynamics of the algal community. The algal species “counts” themselves are robust, in terms of relative abundance, but will have unquantified errors associated with sample heterogeneity, counter fatigue and between-counter variation (Thackeray *et al*., 2012). Given the higher level of uncertainty associated with these data and the uncertainty

associated with conversion to biovolume and subsequently chlorophyll *a,* we estimated the **sampling/analytical error** to be +/- 25% and the **overall error** to be +/- 50%.

#### 2.4.5.2 Relaxed Limits of Acceptability

Nutrient input uncertainties including the interaction between the different sources (particularly for P), can have significant knock-on effects as the year-long simulations progress. Relaxed LoAs were developed to allow for uncertainties associated with nutrient inputs, during periods of the year when nutrients are believed to be limiting. The consequences of relaxation, however, mean that knock-on effects on model state variables (such as the P concentration in the mixed layer) are not well-constrained, making definition of the LoA later in the year, when nutrients are no longer limiting growth, problematic. Limits of Acceptability for functional types suffer from similar problems. For example, CS-types tend to be present throughout the stratified period and will hence be more affected by misrepresentation of P inputs. This reasoning provides significant scope for relaxing the LoA such that we do not reject an appropriate model falsely; however, it is worth reiterating that the aim of constraining the sources and timing of nutrient inputs meant that relaxation was minimised for each lake-year. The LoA were relaxed differently for the periods deemed to be predominantly *nutrient limited* or *light limited*. These year-specific periods were estimated using observations of chlorophyll *a* and residual nutrient concentrations and are shown in Table Supp. 3 together with the associated percentage deviations representing the LoA for both chlorophyll *a* and functional types.

#### 2.4.5.3 Timing errors and minimum error magnitude

For both, *initial* and *relaxed* LoA, a minimum absolute error was set to avoid over-constraint by very low observed concentrations: this was set at 5 mg $m^{-3}$ (2 mg $m^{-3}$ for Windermere 2008) for chlorophyll *a* and 10 mg $m^{-3}$ for R and CS functional types. To allow for unquantified uncertainties associated with model forcing, the LoA were expanded temporally to allow for timing errors in simulations. A first-approximation estimate of +/- 10 days ($e_t$ in Eqn. 12) was used as the “window” for an acceptable simulation (Eqn. 12 and Fig. Supp. 1).

#### 2.4.5.4 Weighting acceptable simulations

Simulations which fall within the LoA are assigned a likelihood weighting (*L*) based upon their goodness-of-fit to the observations. The uncertainty embodied in the LoAs described above, was defined by a trapezoidal fuzzy weighting measure (Eqn. 11) for each observation timestep (*i*). This formulation gives an equal weighting (of 1) to all simulations that fall within the *sampling/analytical* error bounds; simulations that fall between the *sampling/analytical*

error and the overall error were given a lower weighting as they approach the LoA (the *overall error*) outside of which they were given a zero weighting as defined by:

$$L_i = 1 - \left(\frac{(E_i - e_{a,i}) - S_i}{(E_i - e_{a,i}) - E_{\mathrm{min,i}}}\right), where: (E_i - e_{a,i}) > S_i > E_{min,i}$$

$$L_i = 1 - \left(\frac{S_i - (E_i + e_{a,i})}{E_{\mathrm{max,i}} - (E_i + e_a)}\right), where: (E_i + e_{a,i}) > S_i > E_{max,i}$$

$$L_i = 1, \qquad where: (E_i - e_{a,i}) < S_i < (E_i + e_{a,i}) \qquad (11)$$

$$L_i = 0, \qquad where: E_{min,i} > S_i > E_{max,i}$$

and where, $S_i$ is the simulated estimate, $E_i$ is the expected or observed value, $e_{a,i}$ is the analytical error and $E_{min,i}$ and $E_{max,i}$ are the are the overall error. The individual likelihood weights were modified further to allow for timing errors using:

$$L_i = \max\left(L_{i,\Delta t} \times \left|\frac{\Delta t}{e_t}\right|\right) \qquad (12)$$

where, $\Delta t$ is the timing error associated with the simulated variable and $e_t$, is the acceptable timing error. The overall weighting for any given criterion for the simulation period $L_c$ is given by:

$$L_c = \frac{\sum_i^N (L_i)}{N} \qquad (13)$$

where, $N$ is the number of time steps where observed data are available. $L_c$ is common to all observed criteria i.e. Chlorophyll ($L_{Chl}$), R-types ($L_R$), CS-types ($L_{CS}$) and the overall weighting for the simulation $L_s$ is given by:

$$L_s = [(L_{Chl} + L_R + L_{CS}) \times W_s] \qquad (14)$$

and where $W_s$ is unity for S3 and S4 owing to the different sampling strategy and unity for the absolute comparison of fit presented in Table 3. The weighting $W_s$ is however used in determining the final uncertainty estimates (Eqn. 6) for S1 and S2.

## 3 Results and Discussion

Simulation results for the scenarios are presented in this section and are discussed in terms of goodness-of-fit to the available observations and LoA. As a way of comparing the overall performance of modelling scenario, each was assigned an *integrated score* (Table 3). The integrated score was calculated using trapezoidal numerical integration of *all* acceptable *overall* likelihood weightings using Eqn. 14 where $W_s$ was set to unity to enable comparison based solely on goodness-of-fit.

### 3.1 Simulation results: S1

For all lake-years considered, no simulations were acceptable in terms of falling within the stringent **initial LoA** defined above. In fact, no model simulations fell within the specified ranges for chlorophyll *a* alone: i.e. without any additional constraint associated with the LoA based on functional algal types. This is not unusual in environmental modelling applications given the complexity of the uncertainties involved (e.g. Beven *et al*., 2007; Liu *et al*. 2009; Van Straten and Keesman, 1991), particularly when using multi-criteria LoA (e.g. Blazkova and Beven, 2009; Brazier *et al*., 2000). Using the **relaxed LoA**, acceptable simulations were obtained for all lake-years apart from Esthwaite Water 2009 where no simulations were acceptable based upon chlorophyll *a* or community structure and only the chlorophyll *a* LoA could be met for Bassenthwaite Lake (Table 3). Subsequently, in this section results for Esthwaite Lake 2009 relate to the dynamics of simulations which achieved the highest overall weightings and for Bassenthwaite Lake relate to the LoA for chlorophyll *a* only.

Simulations for three of the six lake-years (Windermere 2009, 2010 and Esthwaite Water 2009) showed a general tendency for under-prediction of biomass at the beginning of the year (predominantly in the pre-stratification period) if the biomass towards the end of the year was well-simulated; where simulations provided adequate fits to the early part of the year, there was a systematic overestimation of biomass during and after destratification. This apparent hysteresis was the most distinctive feature of the S1 simulations and is highlighted in Figs. 3a-d by the comparison of two *sets* of simulations which fit either the early or late part of each year but which yield similar goodness-of-fit weightings (calculated using Eqn. 14). The two *sets* of highly-weighted simulations were separated using different ranges of the parameter $\boldsymbol{\varepsilon}_b$ (simulations were most sensitive to $\boldsymbol{\varepsilon}_b$ during the periods of interest). Simulation hysteresis was present but lower for Windermere 2008 and apparently absent for Esthwaite Water 2008, apart from the under prediction of the observation on day 78.

There was also a tendency for there to be too much biomass during the period where the observed chlorophyll *a* concentration "*crashes"* after the spring diatom bloom (e.g. Figs. 3 b and d) which can be as a result many different phenomena including nutrient limitation, zooplankton grazing and sometimes the effects of deep mixing events. Determination of which of these phenomena drive the observed pattern (in both the real system and in model space) is not straightforward as misrepresentation of the mixed depth and consequent light regime or incorrect representation of nutrient inputs could contribute to a similar pattern. However, analysis of concurrent residual nutrient concentrations suggests that a lack of P limitation (possibly together with Si in some cases) exacerbated by a poor simulation of early growth was the most likely cause.

### 3.2 Simulation results: S2, S3 and S4

Implementing the modelling hypotheses had various effects with some simulation improvements in chlorophyll *a* dynamics for some periods and others where simulations were poorer. Where there were improvements these were not enough to allow any simulations to fall within the stringent **initial LoA**. Using the **relaxed LoA**, acceptable simulations were obtained for all lake-years considered except Bassenthwaite Lake where, similarly to S1, the algal community structure was not simulated well. Evaluation of the goodness-of-fit discussed in this section considers both the entire time series and different

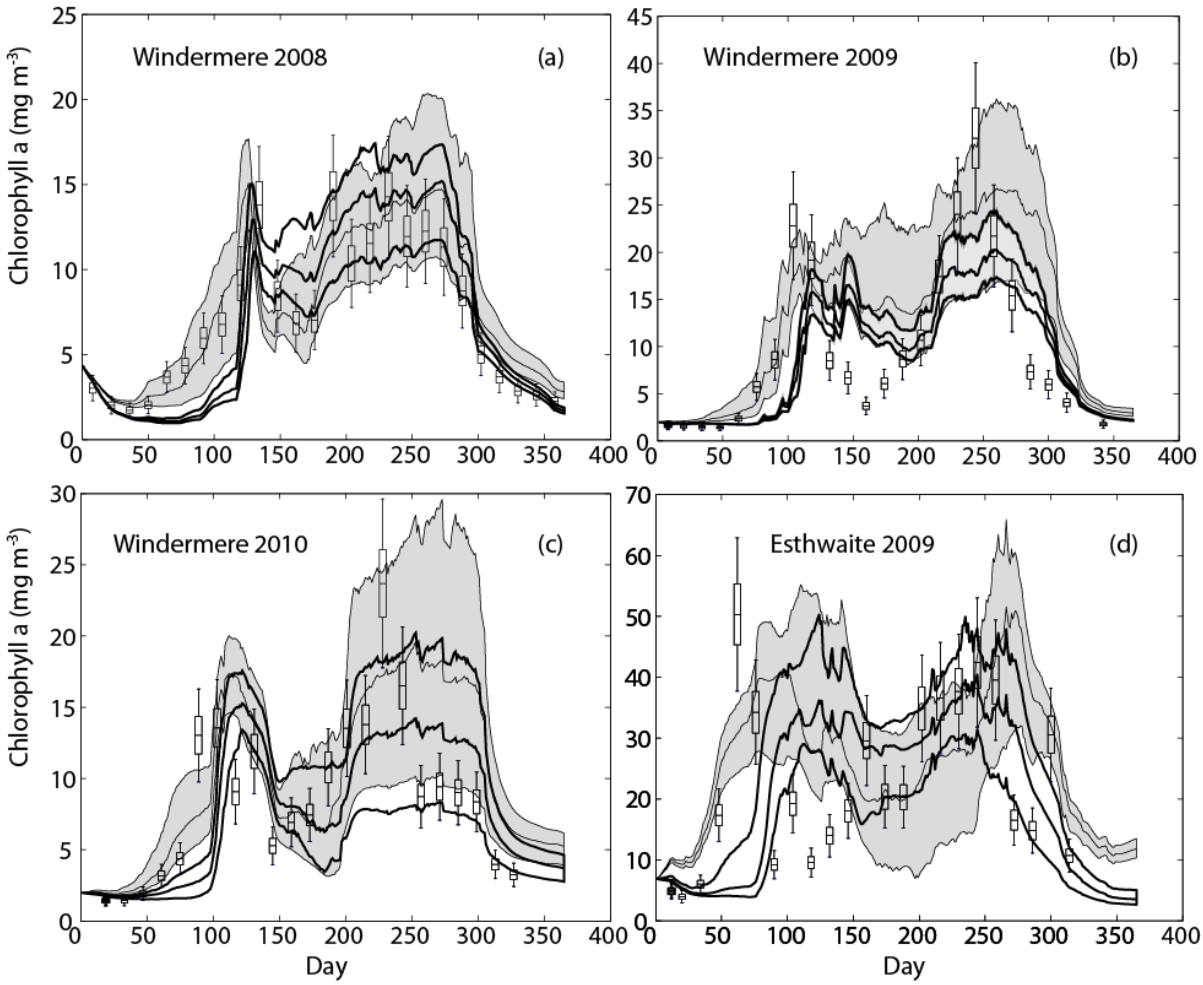


***Figure 3.*** *High-weighted sets of simulations which fit either the early (grey shaded area) or the late (thick black lines) part of each year; the sets were isolated using different ranges of the parameter* $\boldsymbol{\varepsilon}_b$ *and are represented by 3 lines showing the 5th, 50th and 95th percentiles of the likelihood-weighted distributions for: Windermere (a) 2008, (b) 2009, (c) 2010; (d) Esthwaite Water 2009; the box and whisker plots indicate the initial LoA without allowance for timing errors for clarity and where the boxes denote the sampling/analytical error and the whiskers the overall error.*

periods of interest in comparison to the simulations associated with the S1 results: in particular in context with periods where either light or nutrients were deemed to be the most limiting for algal growth. The discussion of simulation dynamics during these periods is qualitative and is based upon simulations that fell within the **relaxed LoA** for chlorophyll *a*, R-type species and CS-type species unless specified. In general, although the *integrated score* showed that some improvements were achieved using the new representation of P inputs (S2 and S4, Table 3), the differences were small and were more apparent in combination with the alternative treatment of mixed depth: for these reasons the majority of the discussion below focusses on the effects of changing the representation of mixed depth

estimates and subsequent algal exposure to light. Of the six lake-years considered, four showed an improved *integrated score* using the model structural changes implemented with Esthwaite Water 2009 showing the most significant differences resulting from improved algal dynamics. Two of the lake years had poorer overall fits to the observed data, one of those significantly.

Simulations for Windermere 2008 were slightly worse using hourly mixed depths (Fig. 4a) primarily as a result of an unobserved “spike” of biomass simulated at approximately day 310 which was simulated as a result of an occurrence of *temporary stratification* within the model.

The improved simulations for Windermere 2009 (Fig. 4b) were achieved using S3 and S4 and resulted in an overall reduction in hysteresis in the predicted biomass relative to the observations. The significant deviation between simulated and observed chlorophyll *a* at around day 125 to 175, where too much biomass was simulated, was apparent under all scenarios (Figs. 3b and 4b). Evidence from the observed data for this period indicates that the loss of biomass in the real system is associated with P and Si limitation which could not be simulated using the sampled nutrient inputs and which was compounded by the knock-on effects of the under estimation of biomass (and associated lack of nutrient uptake) around days 100 to 120.

The simulation of R-type species growth was improved for Windermere 2010 in both the spring bloom and the resurgent population after approximately day 200 (Fig. 4c). Similarly to 2009, the higher biomass between days 220 and 250 was not simulated well with any of the model implementations, primarily owing to hysteresis effects but also because of an *apparent* misrepresentation of P inputs during a specific inflow event. For model runs which achieved high concentrations for this period, simulation of the low concentrations observed (of primarily R-type species) in the subsequent days (approximately days 250-300) was not possible. Given that observations of residual concentrations of P and Si were observed to be relatively high during this period, too much available light, because of the use of hourly mixed depths, is a possible cause.

A consistent pattern for all 3 years of simulations for Windermere showed a lack of sufficient loss (or too much growth) of algal biomass, particularly towards the end of the year. In the “real” system net-losses are observed to be more rapid when the mixed depth is *estimated* to be greater than approximately 15 to 20 m, whereas in model space rapid loss occurs at a greater depth (approximately between 20 and 25 m) indicating a systematic difference in the model representation.

For Esthwaite Water 2008, and for periods where R-type species dominated (approx. days 0-140; Fig. 4d), improved simulation dynamics were achieved using S3 and S4, where S1 resulted in an overestimate around days 100 to 120. After day 250, the hourly mixed depth representation produced more dynamic responses than the “smoothed” response associated with the daily mixed depth; the more dynamic responses are, however, difficult to associate with improved simulations given the frequency of observations available. Using the hourly mixed depths gave a poorer representation of the chlorophyll *a* dynamics between days 140 and 250 (Fig. 4d) where CS-type species were observed to be dominant; the over-estimation of R-type species during this period led to a lower integrated score.

An improvement in simulation dynamics and overall fit was achieved using S3 and S4 for Esthwaite Water 2009 (Table 3 and Fig. 4e). In particular, the simulation of rapid growth from around day 40 to 90 was made possible, although the model was still not able to simulate the peak observed chlorophyll *a* concentration on day 62. The hourly mixed depth estimates of S3 and S4 provide good simulations of the observed chlorophyll *a* dynamics around days 280-300 which were not simulated well under S1 or S2 (Fig. 4e) and which subsequently led to rejection of all simulations for these scenarios (Table 3). The algal population dynamics from approximately day 90 to day 160 were not simulated well by any of the implemented model structures and was apparently a result of the misrepresentation of P inputs on the limitation of growth, although this was likely to be compounded by the underestimation of growth between timesteps 40 to 90.

In the case of Bassenthwaite Lake where algal community structure was not simulated well for any of the scenarios, only the **relaxed** LoA for chlorophyll *a* were used for model rejection. Under S3 and S4 a marginal overall improvement in integrated score was achieved (Table 3), but simulations also gave periods of poor fit which appear to be associated with too much available light: these periods were at the extremes of the year coincident with periods of reverse stratification (Fig. 4f).

### 3.3 Parameter sensitivities

The importance of available light in simulated algal dynamics is supported by the fact that $\varepsilon_b$ was consistently the most sensitive parameter for all lake-years and all scenarios and that acceptable parameter values were constrained significantly from the initial range sampled. This is shown in the examples of (Figs. 5 a and b) which are one-dimensional representations of the multidimensional parameter space, presented as scatter plots of parameter value versus likelihood-weighting; it can be seen that the acceptable simulations are located in a smaller range than sampled. Simulations were also sensitive to the various

parameters which control the dynamics of P inputs but to a lesser extent (e.g. Figs. 5 c and d). The apparent insensitivity for some lake-years is likely to be associated with interaction between the different P sources, particularly in Esthwaite Water where internal P sources were included.

***Table 3. Integrated scores for each scenario***

| Lake | Year | S1 | S2 | S3 | S4 |
|---|---|---|---|---|---|
| Windermere | 2008 (43$^{\Psi}$) | 19.4 (21.9$^{\Psi}$) | 19.3 (21.82$^{\Psi}$) | 18.3 (20.85$^{\Psi}$) | 18.4 (20.85$^{\Psi}$) |
| Windermere | 2009 (35$^{\Psi}$) | 9.32 (12.85$^{\Psi}$) | 9.64 (12.4$^{\Psi}$) | 10.06 (14.23$^{\Psi}$) | 11.32 (14.95$^{\Psi}$) |
| Windermere | 2010 (32$^{\Psi}$) | 14.48 (18.99$^{\Psi}$) | 14.21 (18.83$^{\Psi}$) | 15.38 (19.76$^{\Psi}$) | 15.64 (19.99$^{\Psi}$) |
| Esthwaite Water | 2008 (51$^{\Psi}$) | 17.04 (24.79$^{\Psi}$) | 17.97 (25.41$^{\Psi}$) | 14.1 (21.2$^{\Psi}$) | 14.4 (21.6$^{\Psi}$) |
| Esthwaite Water | 2009 (45$^{\Psi}$) | 0 | 0 | 15.17 (19.95$^{\Psi}$) | 18.70 (22.49$^{\Psi}$) |
| Bassenthwaite Lake | 2010 (38$^{\Psi}$) | 11.70 (15.25$^{\Psi}$)* | 11.97 (16.62$^{\Psi}$)* | 12.3 (15.17$^{\Psi}$)* | 12.6 (16.37$^{\Psi}$)* |

** Values given are for lake-years where all simulations were rejected based upon LoA for functional algal types and are presented for comparison; the values presented are calculated using all simulations using the chlorophyll a LoA alone but include the weightings for goodness of fit to functional species types; $^{\Psi}$maximum value attainable if simulations fell within the sampling/analytical error range for all LoA criteria at all observation timesteps.*

### 3.4 Implications for modelling and future research

In interpreting the simulation results from the previous sections, resolution of the causes of poor model fits to observations is difficult given the complex interactions between phenomena that control growth and loss in both model space and real lake systems. This is made more difficult by the potential for significant knock-on effects of simulation errors from previous timesteps. There are, however, a few salient results of which we can be more confident that provide a better representation of the lake systems studied here and some which remain hypotheses to be tested.

In terms of appropriate representation of algal exposure to underwater light, representing temporary (sub-daily) stratification events significantly improved simulation dynamics for some periods and gave a smaller but systematic improvement to the balance of growth during stratifying and destratifying periods. However, some periods were not simulated as well as the standard model formulation and simulations still exhibit residual hysteresis for some lake-years. It is possible that the disparity between the degree of epilimnetic mixing in real systems and the modelling representation (i.e. particularly the assumption of complete epilimnetic mixing) is important in this respect. Although representing temporary stratification gave some improvements, because the density gradient estimate of mixed depth (calculated on the basis of isothermal conditions) does not describe the degree of epilimnetic mixing, the representation of algal exposure to light may still need to be improved. For example, when stratification is indicated by isothermal conditions, but there is little mixing in the real system,

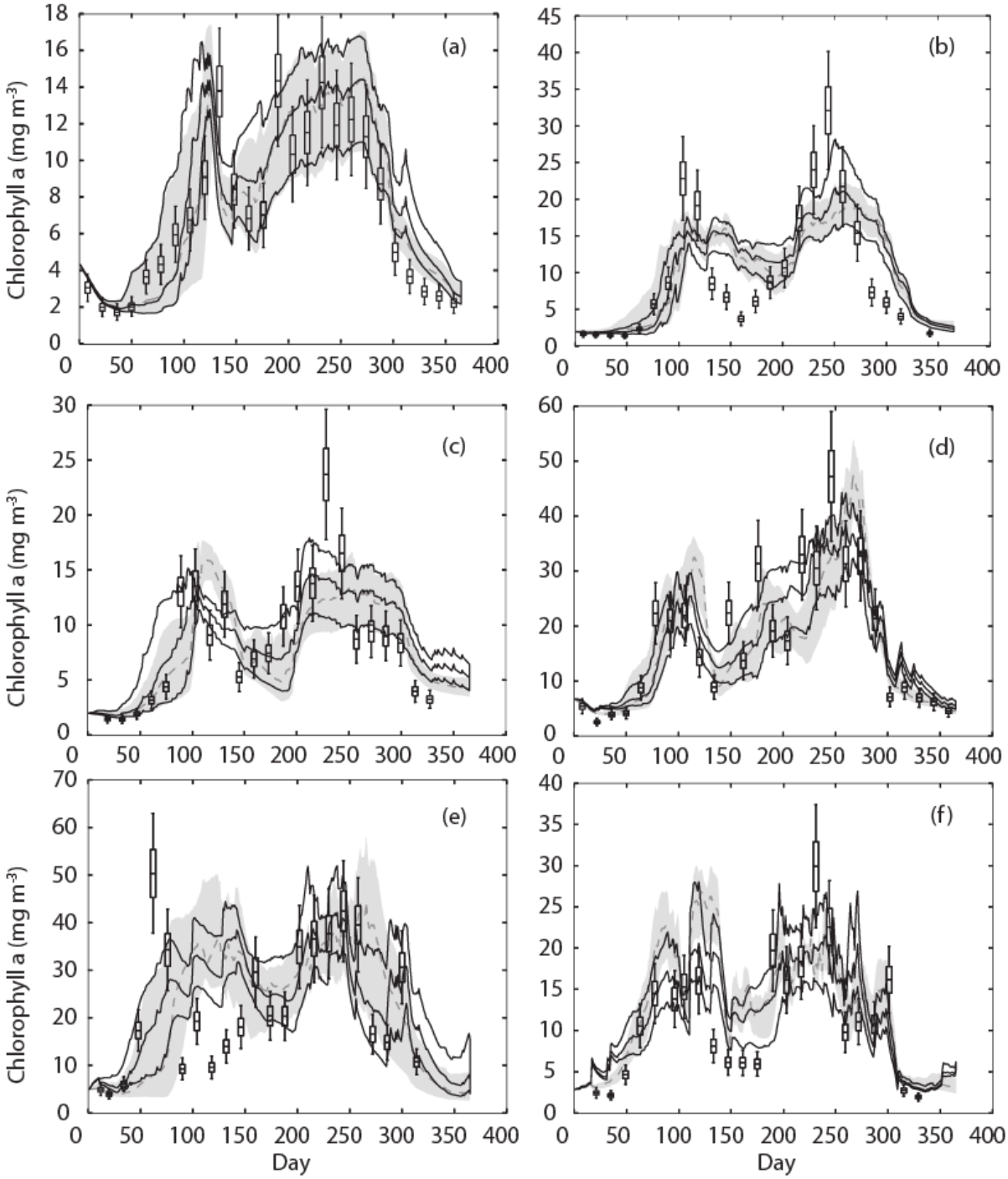


***Figure 4.*** *Comparison of S1 acceptable simulations (grey shaded uncertainty envelope and dashed line) and S4 (solid black lines) for chlorophyll a; the uncertainty estimates represent the 5th, 50th and 95th percentiles of the likelihood-weighted distributions for: Windermere (a) 2008** (b) 2009** (c) 2010 Esthwaite Water (d) 2008** (e) 2009 and Bassenthwaite Lake 2010** (f); the box and whisker plots indicate the analytical error and the overall error as defined in section (3.4.4) respectively; The LoA including timing errors are not shown for*

*clarity. ** where all simulations were rejected the highest likelihood-weighted simulations were used for comparison.*

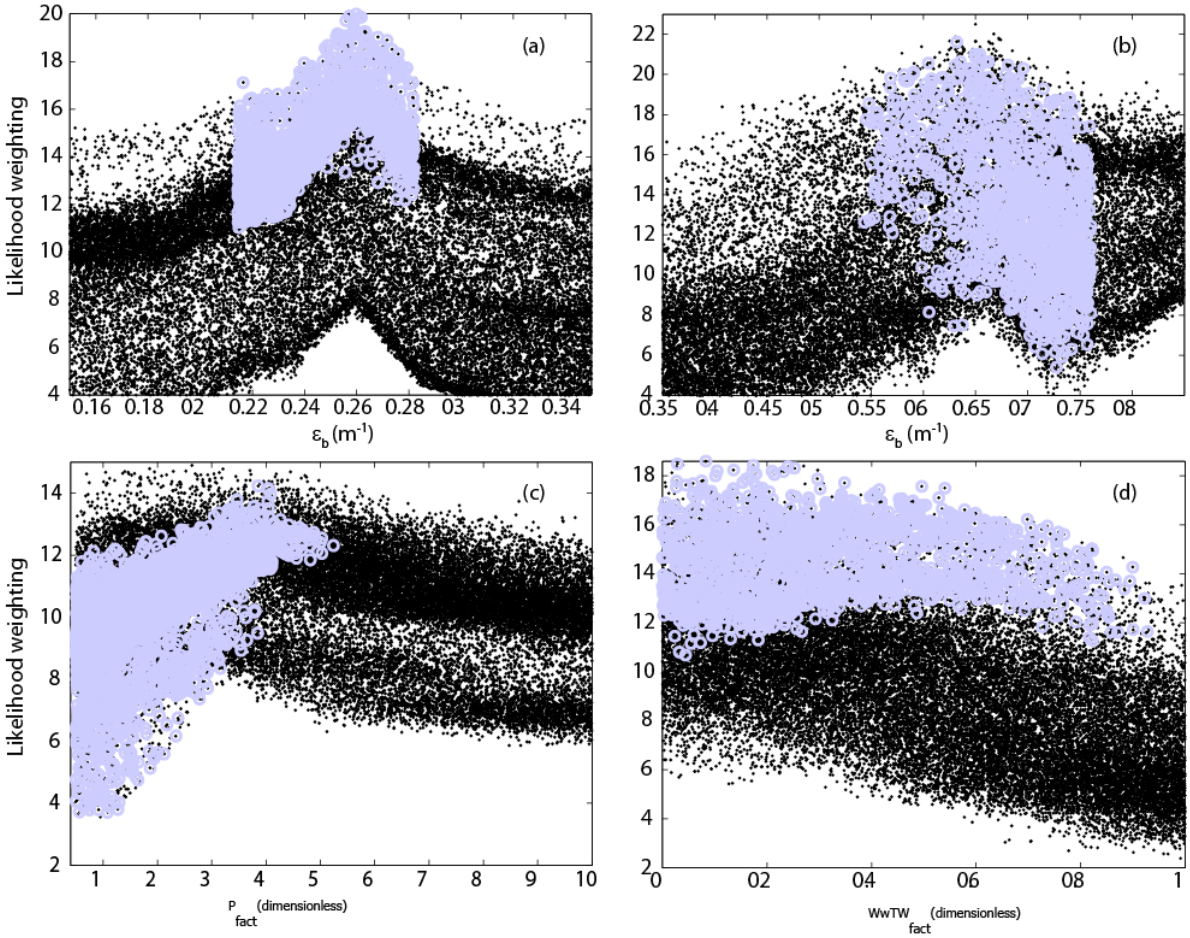


***Figure 5.*** *Scatter plots of likelihood weighting Vs. parameter value for: (a)* $\boldsymbol{\varepsilon}_b$ *(Windermere 2010: S3) (b)* $\boldsymbol{\varepsilon}_b$ *(Esthwaite 2008: S1) (c)* $P_{fact}$ *(Windermere 2009: S3) and (d)* $WwTW_{fact}$ *(Windermere 2010: S3). Grey circles are acceptable parameter sets and black dots all samples.*

algae will grow at different light dependent rates at different depths such that the modelling assumption of complete mixing will not be appropriate (i.e. the average of the growth at different depths is not equal to the growth under average light conditions). Conversely, when a strongly stratified lake begins to overturn, a significant amount of energy is required to deepen the mixed layer by a small amount so that the change in mixed depth is strongly linked to mixing: averaged light conditions are hence more likely to be appropriate in this case. Although an accurate representation of mixing in a 1-D model is not realistic, it may be possible with further analyses of high resolution data to derive improved state-dependant indices of mixing so that exposure to light is improved.

An alternative hypothesis for the simulation hysteresis is associated with the representation of biomass loss via flushing. Flushing of algae may also be biased because of the assumption of instantaneous mixing. In real, three dimensional lake systems, different fractions of a lake are flushed more efficiently than others and therefore there will be a spectrum of residence times that vary temporally (e.g. with different mixed depths, lake mixing regimes and inflow magnitudes). It may be the case that some horizontal disaggregation of the mixed layer could improve simulations, in a similar manner to the Aggregated Dead Zone approach for river systems where multiple stores (often two in parallel) are used to simulate well-mixed and poorly mixed fractions of a river reach. This approach has primarily been used for conservative chemical tracers (e.g. Beer and Young 1984; Wallis *et al*., 1989 and Barraclough *et al.,* 1994) but has also been used to explain the relatively high (given their relatively short mean residence times) plankton concentrations observed in some rivers (Reynolds *et al.,* 1991; Reynolds 2000; Istvanovics and Honti, 2011). It is possible that using different modelling configurations of the well-flushed and poorly-flushed fractions of lakes could provide both higher and lower concentrations of biomass under different conditions, compared to those simulated using a single well-mixed store.

The timing of nutrient inputs is crucial, in particular, to avoid severe knock-on effects from input errors during subsequent timesteps. Data relating to the sources and timing of nutrient fluxes tend to be lacking, owing to the significant costs associated with the high frequency sampling required, but are critical for reducing the uncertainties associated with algal modelling (Saloranta and Anderson, 2007; Missaghi *et al*., 2013). For some of the lake-years studied here modified diffuse P-river flow relationships were identified; the relationships constrained were far from clear and were affected significantly by knock-on effects from simulation errors. Improvement of nutrient input dynamics is a priority but will be hampered while other systematic simulation errors, that have significant effects on algal growth, remain.

In terms of forecasting algal blooms, mitigation of uncertainties associated with nutrient inputs can be achieved to some degree by data assimilation and will be dependent on the frequency of the observations. Mitigating systematic errors associated with algal exposure to light may be more challenging and will require state-dependent functional relationships to be identified. These two priorities are currently being explored.

## 4 Conclusions

The process-representation of the algal community model PROTECH was tested using the extended Generalised Likelihood Uncertainty Estimation technique which employs pre-defined Limits of Acceptability for determination of model adequacy. Testing was a precursor to modification of the model for real-time forecasting of algal communities which places different demands on the model in terms of the accuracy required for simulation estimates. For consistency with the data available to develop the forecasting system, high resolution observations were used to force the model, minimising simulation uncertainties associated with some elements of the abiotic nature of the lakes.

Two modelling hypotheses were tested, under four scenarios, which considered the representation of algal exposure to light and the timing and magnitude of diffuse SRP inputs. It was found that when using the initial (stringent) Limits of Acceptability all simulations were unacceptable. Relaxed Limits of Acceptability which provided allowance for errors associated with model forcing inputs were developed and acceptable simulations were identified. Modifying the way the mixed depth (strictly depth of epilimnion) was represented provided some simulation improvements for periods when the systems were light limited and an overall improvement for some of the lake-years considered. However some residual systematic errors, which manifest themselves as a hysteretic effect on biomass, remain. Although simulations for some of the lake-years were improved by modification of the diffuse P input-inflow relationship, they were limited by other simulation errors which have significant knock-on effects on residual nutrient concentrations. Nutrient inputs are likely to be a significant limiting factor for simulating algal community dynamics and particularly for the accuracy required for real-time forecasting, but they are difficult to constrain using modelling approaches where other simulation errors exist. Never the less, improved observations of the timing and magnitude of nutrient fluxes would greatly enhance our ability to reduce modelling input uncertainties and focus on model process representation.

By taking a hypothesis-driven approach within the Generalised Likelihood Uncertainty Estimation framework, which employs pre-defined Limits of Acceptability, has helped improve the model's representation of epilimnetic depth and identify new modelling hypotheses which may further improve simulations. These relate to the disparity between the degree of epilimnetic mixing in real systems and the modelling assumption of instantaneous epilimnetic mixing. Although an accurate description of mixing in a 1-D model may not be possible, further analyses of high resolution data may allow the identification of state-dependant indices of mixing and subsequent algal exposure to light that will further improve model representation. It is also possible that the representation of the flushing of algae is misrepresented and that it may be improved by considering what fraction of the epilimnion is actively mixed during different periods.

## Figure Captions

Figure 1. Plan view and inset of bathymetric curve for (a) Windermere South Basin*, (b) Esthwaite Water** and (c) Bassenthwaite Lake*. * Redrawn from Ramsbottom, 1976; ** Redrawn from Mackay *et al*., 2012.

Figure 2. Comparison of the standard model mixed depth estimates based upon daily averaged temperature profiles (black line) with individual hourly mixed depth estimates for the same day (grey circles) for Esthwaite Water 2009. The distribution of hourly estimates for each day was sampled to provide a modified representation of the daily depth for the modelling scenarios (Table 2).

Figure 3. High-weighted *sets* of simulations which fit either the early (grey shaded area) or the late (thick black lines) part of each year; the *sets* were isolated using different ranges of the parameter $\boldsymbol{\varepsilon}_b$ and are represented by 3 lines showing the 5$^{th}$, 50$^{th}$ and 95$^{th}$ percentiles of the likelihood-weighted distributions for: Windermere (a) 2008, (b) 2009, (c) 2010; (d) Esthwaite Water 2009; the box and whisker plots indicate the initial LoA without allowance for timing errors for clarity and where the boxes denote the sampling/analytical error and the whiskers the overall error.

Figure 4. Comparison of S1 acceptable simulations (grey shaded uncertainty envelope and dashed line) and S4 (solid black lines) for chlorophyll *a*; the uncertainty estimates represent the 5$^{th}$, 50$^{th}$ and 95$^{th}$ percentiles of the likelihood-weighted distributions for: Windermere (a) 2008** (b) 2009** (c) 2010 Esthwaite Water (d) 2008 (e) 2009 ** and Bassenthwaite Lake 2010** (f); the box and whisker plots indicate the analytical error and the overall error as defined in section (3.4.4) respectively; The LoA including timing errors are not shown for clarity. ** where all simulations were rejected the highest likelihood-weighted simulations were used for comparison.

Figure 5. Scatter plots of likelihood weighting Vs. parameter value for: (a) $\boldsymbol{\varepsilon}_b$ (Windermere 2010: S3) (b) $\boldsymbol{\varepsilon}_b$ (Esthwaite 2008: S1) (c) $P_{fact}$ (Windermere 2009: S3) and (d) $WwTW_{fact}$ (Windermere 2010: S3). Grey circles are acceptable parameter sets and black dots all samples.

Figure Supp.1. Example Limits of Acceptability; two-dimensional representation of weightings base upon observed chlorophyll *a* concentrations (*Initial* LoA for Windermere 2008); inset shows a three dimensional example of the shape of the weighting function at each observation timestep.

**Acknowledgements**

This work was supported by the Natural Environment Research Council projects: the United Kingdom Lake Ecological Observatory Network (UKLEON; grant number NE/I007407/1) and the Consortium on Risk in the Environment: Diagnostic, Integration, Benchmarking, Learning and Elicitation (CREDIBLE; grant number NE/J017299/1).

**References**

Barraclough, A., Freestone, R., Guymer, I. and O'Brien, R.T. (1994). Evaluation of the Aggregated Dead Zone (ADZ) method as a River Catchment Management Tool applied to the rivers Aire and Derwent in Yorkshire. Proceedings of 2nd Int. Conf. on Hyd. Modelling, Stratford-upon-Avon, U.K., 14-16 June, 439-449.

Beer, T. and Young, P.C. (1984). Longitudinal Dispersion in Natural Streams. Proc. A.S.C.E., J. Env. Eng. Div., 109, 1049-1067.

Bennion, H., Monteith, D. and Appleby, P. (2000). Temporal and geographical variation in lake trophic status in the English Lake District: evidence from (sub)fossil diatoms and aquatic macrophytes. Freshwater Biology, 45(4), 1365-2427, doi: 10.1046/j.1365-2427.2000.00626.x

Beven, K.J. (2006). A manifesto for the equifinality thesis. Journal of Hydrology 320

(1–2), 18–36.

Beven, K. J, (2012). Causal models as multiple working hypotheses about environmental processes, Comptes Rendus Geoscience, Académie de Sciences, Paris, 344, 77–88, doi:10.1016/j.crte.2012.01.005.

Beven, K.J. and Binley, A.M. (1992). The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298.

Beven, K. J. and Binley, A. M. (2014). GLUE, 20 years on. Hydrol. Process. 28(24):5897-5918, DOI: 10.1002/hyp.10082.

Beven, K.J., Page, T., McGechan, M. (2007). Uncertainty estimation in phosphorus models. In: Modeling phosphorus in the environment. Boca Ranton : CRC Press p. 131-160. 30 p. ISBN: 0849337771.

Blazkova, S. and Beven, K.J. (2009). A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation : Skalka catchment, Czech Republic. Water Resources Research. 45, W00B16.

Brazier, R.E., Beven, K.J., Freer, J. and Rowan, J.S. (2000). Equifinality and uncertainty in physically based soil erosion models: application of the GLUE methodology to WEPP-the water erosion prediction project-for sites in the UK and USA. Earth Surface Processes and Landforms. 25, 8, 825-845.

Brookes, J.D. and Carey, C.C. (2011).Resilience to blooms. Science 334 (6052), 46-47; DOI: 10.1126/science.1207349.

Carmichael, W.W. (1992). A status report on planktonic cyanobacteria (blue-green algae) and their toxins. EPA/600/R-92-079, Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH. 141 pp.

Dodds W.K. Bouska , W.W., Eitzmann , J. L., Pilger , T. J., Pitts, K. L., Riley, A.J. Schloesser, J.T. and Thornbrugh, D.J. (2009). Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. Environ. Sci. Technol, 43, 12–19.

Dong X., Bennion H., Maberly S.C., Sayer C.D., Simpson G.L. and Battarbee R.W. (2012). Nutrients provide a stronger control than climate on diatom communities in Esthwaite Water: Evidence from monitoring and palaeolimnological records over the past 60 years. Freshwater Biology 57, 2044-2056.

Smith, V.H. (2003). Eutrophication of Freshwater and Coastal Marine Ecosystems: A Global Problem. Environ Sci & Pollut Res. 10 (2), 126 - 139.

Cassidy, R. and Jordan, P. (2011). Limitations of instantaneous water quality sampling in surface-water catchments: Comparison with near-continuous phosphorus time-series data. Journal of Hydrology, 405, (1–2), 182-193.

Elliott, J.A. (2010). The seasonal sensitivity of Cyanobacteria and other phytoplankton to changes in flushing rate and water temperature. Global Change Biology, 16, 864-876.

Elliott, J. A., Irish, A. E. and Reynolds, C, S., (2010). Modelling phytoplankton dynamics in fresh waters: affirmation of the PROTECH approach to simulation. Freshwater Reviews, 3, 75-96.

Elliott, J.A. (2012). Predicting the impact of changing nutrient load and temperature on the phytoplankton of England's largest lake, Windermere. Freshwater Biology, 57, 400-413.

Elliott, J.A., Jones, I.D. and Page, T. (2009). The importance of nutrient source in determining the influence of retention time on phytoplankton: an explorative modelling study of a naturally well-flushed lake. Hydrobiologia, 627, 129-142.

Elliott, J.A., Irish, A.E. and Reynolds, C.S. (1999). Sensitivity analysis of PROTECH, a new approach in phytoplankton modelling. J Hydrobiologia, 414, 45-51.

Elliott, J. A. and L. Defew (2012). Modelling the response of phytoplankton in a shallow lake (Loch Leven, UK) to changes in lake retention time and water temperature. Hydrobiologia, 681, 105-116.

Elliott, J.A. and Thackeray, S.J. (2004). The simulation of phytoplankton in shallow and deep lakes using PROTECH. Ecological Modelling 178, 357–369.

George, D. G. and Heaney, S. I. (1978). Factors influencing the spatial distribution of phytoplankton in a small productive lake. Journal of Ecology, 66(1),133-155.

Hall, G.H., Maberly, S.C., Reynolds, C.S., Winfield, I.J., James, B.J.,Parker, J.E., Dent, M.M., Fletcher, J.M., Simon, B.M. & Smith, E.(2000). Feasibility study on the restoration of three Cumbrian lakes. Centre for Ecology and Hydrology Windermere, Ambleside, UK. 82 pp.

Hamilton, D.P., Schladow, S.G. (1997). Prediction of water quality in lakes and reservoirs. Part I. Model description. Ecol. Model. 96, 91–110.

Heany, S.I., Corry, J. E. and Lishman, J. P. (1992). Changes of water quality and sediment phosphorus of a small productive lake following decreased phosphorus loading. Centre for Ecology and Hydrology Windermere, Ambleside, UK. 14 pp.

Hipsey, M.R., and Hamilton, D.P. 2008. Computational Aquatic Ecosystem Dynamic Model: CAEDYM v3 Science Manual. Centre for Water Research Report.

Ho, J. C. and Michalak, A. M. (2015). Challenges in tracking harmful algal blooms: A synthesis of evidence from Lake Erie, Journal of Great Lakes Research, 41(2), 317-325: doi.org/10.1016/j.jglr.2015.01.001.

Istvanovics, V. and Honti, M. (2011). Phytoplankton growth in three rivers: The role of meroplankton and the benthic retention hypothesis. Limnol. Oceanogr., 56(4), 2011, 1439–1452.

Johnes, P.J. (2007). Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. Journal of Hydrology, 332(1–2), 241–258.

Knowlton, M.F., Hoyer, M.V. and Jones, J.R. (1984). Sources of variability in phosphorus and chlorophyll a and their effects on use of lake survey data. Journal of the American Water Resources Association 20(3), 397–408.

Liu, Y., Freer, J., Beven, K.J., Matgen (2009). Towards a limits of acceptability approach to the calibration of hydrological models: extending observation error. Journal of Hydrology. 367(1-2), 93-103.

Maberly, S.C. and Elliott, J.A. (2009). Options for the remediation of Windermere: PROTECH modelling of the effects of different management scenarios. NERC/Centre for Ecology and Hydrology, 28pp. (CEH Report Ref: LA/C03623/2).

Maberly, S.C. and Elliott, J.A. (2012). Insights from long-term studies in the Windermere catchment: external stressors, internal interactions and the structure and function of lake ecosystems. Freshwater Biology, 57, 233,243: doi:10.1111/j.1365-2427.2011.02718.x

Maberly, S.C., De Ville, M.M., Thackeray, S.J., Feuchtmayr, H., Fletcher, J.M.,James, J.B., Kelly, J.L., Vincent, C.D., Winfield, I.J., Newton, A., Atkinson, D., Croft,A., Drew, H., Saag, M., Taylor, S., Titterington, H. (2011). A survey of the lakes of the English Lake District: The Lakes Tour 2010. NERC/Centre for Ecology and Hydrology,137pp. (CEH Project Number. Report to: Environment Agency, North West Region and Lake District National Park Authority: downloaded Jan 2015 from: http://nora.nerc.ac.uk/14563/2/N014563CR.pdf

Mackay, E. B., Jones, I. D., Folkard, A. M. and Barker, P. (2012). Contribution of sediment focussing to heterogeneity of organic carbon and phosphorus burial in small lakes. Freshwater Biology, 57, 290-304.

Mackay, E.B., Jones, I.D., Thackeray, S.J. and Folkard, A.M. (2011). Spatial heterogeneity in a small, temperate lake during archetypal weak forcing conditions. Fundamental and Applied Limnology, 179, 27-40.

Mackay, E.B. Folkard, A.M. and Jones, I.D. (2014). Interannual variations in atmospheric forcing determine trajectories of hypolimnetic soluble reactive phosphorus supply in a eutrophic lake. Freshwater Biology, 59, 1646 – 1658.

Madgwick G., Jones I.D., Thackeray S.J., Elliott J.A. & Miller H.J. (2006). Phytoplankton communities and antecedent conditions: high resolution sampling in Esthwaite Water. Freshwater Biology, 51, 1798–1810.

Metcalf, J.S. and Codd, G.A. (2009). Cyanobacteria, neurotoxins and water resources: are there implications for human neurodegenerative disease? Amyotrophic Lateral Sclerosis 10, suppl. 2, 74-78 (2009).

Michalak, A., M. (2016). Study role of climate change in extreme threats to water quality. Nature 535, 349-350.

Missaghi, S., Hondzo, M. and Melching, C. (2013). Three-dimensional lake water quality modeling: sensitivity and uncertainty analyses. Journal of Environmental Quality, 42(6), 1684-1698.

Omlin, ., Brun, R. and Reitchert, P. (2000). Biogeochemical model of Lake Zürich model: sensitivity, identifiability and uncertainty analysis. Ecol. Modell. 141,105-123.

Paerl, H.W. and Huisman, J.  (2008). Blooms like it hot. Science, 4, 320(5872):57-8. doi: 10.1126/science.1155398. DOI: 10.1126/science.1155398

Pretty, J. N., Mason, C. F., Nedwell, D. B., Hine, R. E., Leaf, S., and Dils, R. (2003). Environmental Costs of Freshwater Eutrophication in England and Wales. Environ. Sci. Technol., 37(2), 201-208.

Ramsbottom A.E. (1976). Depth Charts of the Cumbrian Lakes. Freshwater Biological Association Scientific Publication No. 33, Ambleside, UK.

Read J.S., Hamilton, D.P., Jones, I.D., Muraoka, K., Winslow, L.A. , Kroiss, R. , Wu, C.H. and Gaiser. E. (2011). Derivation of lake mixing and stratification indices from high-resolution lake buoy data. Environmental Modelling and Software. 26, 1325-1336.

Reynolds C.S. (1988). Functional morphology and the adaptive strategies of freshwater phytoplankton. In: Growth and Reproductive strategies of Freshwater Phytoplankton (Ed. C.D. Sandgren), pp. 388–433. Cambridge, University Press, New York.

Reynolds CS, Carling PA and Beven KJ. (1991). Flow in river channels; new insights into hydraulic retention. Archiv fuer Hydrobiologie. 121, 171-179.

Reynolds C.S. (2000). Hydroecology of river plankton: the role of variability in channel flow. Hydrol. Process. 14, 3119-3132.

Reynolds C.S., Irish A.E. and Elliott J.A. (2001). The ecological basis for simulating phytoplankton responses to environmental change (PROTECH). Ecological Modelling, 140, 271–291.

Rigosi, A., Carey, C.C., Ibelings, B. W. and Brookes, J. D. (2014). The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. Limnol. Oceanogr. 59(1), 2014, 99–114; doi:10.4319/lo.2014.59.01.0099.

Saloranta, T., M. and Andersen, T. (2007). MyLake - A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. Ecological Modelling, 207 (1), 45-60.

Smith, V.H., (2003). Eutrophication of Freshwater and Coastal Marine Ecosystems: A Global Problem. Environ Sci & Pollut Res. 10 (2) 126-39.

Steinman, A., Chu, X. and Ogdahl, M. (2009). Spatial and temporal variability of internal and external phosphorus loads in Mona Lake, Michigan. Aquat. Ecol. 43, 1 -18.

Thackeray, S.J. Noges, P., Dunbar, M. J., Dudley, B. J., Skjelbred, B., Morabito, G., Carvalho, L., Phillips, G. Mischke, U., Catalan, J., de Hoyos, C., Laplace, C., Austoni, M., Padedda, B. M., Maileht, K., Pasztaleniec, A., Järvinen, M., Lyche Solheim, A. and Clarke, R.T. (2013). Quantifying uncertainties in biologically-based water quality assessment: a pan-European analysis of lake phytoplankton community metrics. Ecological Indicators. 29. 34-47. doi:10.1016/j.ecolind.2012.12.010

Van Straten, G and Keesman, K J, (1991). Uncertainty propagation and speculation in projective forecasts of environmental change, J. Forecasting, 10, 163-190.

Wallis, S.G., Young, P.C. and Beven. K.J. (1989). Experimental Investigations of the Aggregated Dead Zone Model for Longitudinal Solute Transport in Stream Channels. Proc. Inst. Civil Eng., Part 2. 87, 1-22.

World Health Organization (1999). Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management. I. Chorus and J. Bartram (Eds.). E & FN Spon, London, UK (1999).

## Supplementary information

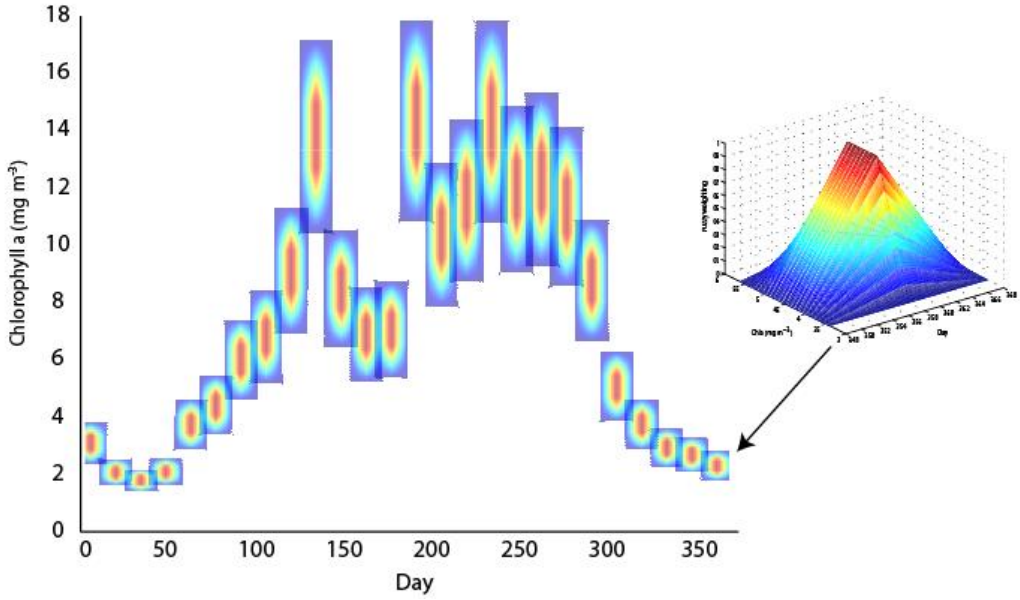


***Figure Supp 1.*** *Example Limits of Acceptability; two-dimensional representation of weightings base upon observed chlorophyll a concentrations (Initial LoA for Windermere 2008); inset shows a three dimensional example of the shape of the weighting function at each observation timestep.*

***Table Supp. 1. Model parameters varied and ranges sampled for each lake-year and each of the modelling scenarios (S1-S4; Table 2). See text for explanation of the parameters.***

| Parameter | Lake | year | S1 | S2 | S3 | S4 |
|---|---|---|---|---|---|---|
| Background light extinction coef. | Windermere | 08/09/10 | 0.1-0.5 | 0.1-0.5 | 0.1-0.5 | 0.1-0.5 |
| | Bassenthwaite | 10 | 0.35-0.85 | 0.35-0.85 | 0.35-0.85 | 0.35-0.85 |
| $\varepsilon_b$ ($m^{-1}$) | Esthwaite | 08/09 | 0.35-0.85 | 0.35-0.85 | 0.35-0.85 | 0.35-0.85 |
| Mixed depth percentile *Mp* (%) | All lakes | 08/9/10 | - | - | 10-90 | 10-90 |
| Diffuse P input multiplier | Windermere | 08/9/10 | 0.2-10 | - | 0.2-10 | - |
| $P_{fact}$ (dimensionless) | Bassenthwaite | 10 | 0.5-1.5 | - | 0.5-1.5 | - |
| | Esthwaite | 08/09 | 0.2-2 | - | 0.2-2 | - |
| Diffuse Si input multiplier | Windermere | 08/9/10 | 0.2-2.5 | 0.2-2.5 | 0.2-2.5 | 0.2-2.5 |
| $Si_{fact}$ (dimensionless) | Bassenthwaite | 2010 | 0.5-1.5 | 0.5-1.5 | 0.5-1.5 | 0.5-1.5 |
| | Esthwaite | 2008/09 | 0.4-2.5 | 0.4-2.5 | 0.4-2.5 | 0.4-2.5 |
| Diffuse N input multiplier | Windermere | 08/9/10 | 0.4-1.5 | - | 0.4-1.5 | - |
| $N_{fact}$ (dimensionless) | Bassenthwaite | 10 | 0.5-2.5 | - | 0.5-2.5 | - |
| | Esthwaite | 08/09 | 0.4-1.5 | - | 0.4-1.5 | - |
| Inlow-P relationship | Windermere | 08/9/10 | - | 2-12 / 2-300 | - | 2-12 / 2-300 |
| | Bassenthwaite | 10 | - | 0.05-2 / 0.05-15 | - | 0.05-2 / 0.05-15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Pmin / Pmax (mg m-3) | Esthwaite | 08/09 | - | 2-50 / 2-700 | - | 2-50 / 2-700 |
| WwTW P input multiplier $WwTW_{fact}$ (dimensionless) | Windermere | 08/09/10 | 0.01-0.9 | 0.01-0.9 | 0.01-0.9 | 0.01-0.9 |
| | Bassenthwaite | 10 | 0-1 | 0-1 | 0-1 | 0-1 |
| | Esthwaite | 08/09 | 0.01-1.2 | 0.01-1.2 | 0.01-1.2 | 0.01-1.2 |
| Hypolimnetic P modifier $P_{Hypo}$ (mg $m^{-2}$) | Esthwaite | 08/09 | 2-8 | 2-8 | 2-8 | 2-8 |
| Epilimnetic P modifier $P_{Epi}$ (dimensionless) | Esthwaite | 08/09 | 0.5-1.5 | 0.5-1.5 | 0.5-1.5 | 0.5-1.5 |
| Vertical eddy diffusivity $K_z$ ($m^2$ $d^{-1}$) | All lakes | 08/9/10 | 0.05-0.4 | 0.05-0.4 | 0.05-0.4 | 0.05-0.4 |
| Metalayer depth $ML_d$ (m) | All lakes | 08/9/10 | 1.1 | 1.1 | 1.1 | 1.1 |
| Light extinction (algae) $\varepsilon_a$ ($m^2mg^{-1}$) | All lakes | 08/9/10 | 0.01 | 0.01 | 0.01 | 0.01 |

***Table Supp. 2. Species used to represent algal communities. Functional types follow* Reynolds (1988).**

| **Windermere** | **Functional type** | **Bassenthwaite Lake** | **Functional type** | **Esthwaite Water** | **Functional type** |
|---|---|---|---|---|---|
| *Aphanizomenon flos-aquae* | CS | *Aulacoseira* | R | *Asterionella* | R |
| *Aulacoseira* | R | *Asterionella* | R | *Aulacoseira* (2008); *Fragilaria crotonensis (2009)* | R |
| *Asterionella* | R | *Cryptomonas* | CSR | *Aphanizomenon flos-aquae* | CS |
| *Cryptomonas* | CSR | *Dolichospermum* | CS | *Aphanothece clathrata* | CS |
| *Dolichospermum* | CS | *Monoraphidium* | CR | *Cryptomonas* | CSR |
| *Monoraphidium* | CR | *Paulschulzia tenera* | S | *Dictyosphaerium pulchellum* | R |
| *Oscillatoria* | R | *PseudDolichospermum* | R | *Dolichospermum* | CS |
| *Paulschulzia tenera* | S | *Pseudosphaerocystis lacustris* | S | *Eudorina* | S |

***Table Supp. 3 Estimated periods of nutrient limitation for each lake-year and percentage error for chlorophyll a, R-type and CS-type species for nutrient limited and light limited periods***

| Lake | Year | Start (day) | End (day) | LoA (% error) nutrient limited period (Chl*a* /R and CS) | LoA% light limited period (Chl*a* /R and CS) |
|---|---|---|---|---|---|
| Windermere | 2008 | 135 | 250 | 50/75 | 35/50 |
| Windermere | 2009 | 100 | 280 | 75/95 | 50/75 |
| Windermere | 2010 | 110 | 260 | 70/95 | 50/75 |
| Esthwaite Water | 2008 | 90 | 250 | 75/95 | 50/75 |
| Esthwaite Water | 2009 | 60 | 270 | 75/95 | 50/75 |
| Bassenthwaite Lake | 2010 | 50 | 315 | 50/60 | 35/50 |