Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR028773

Hourly Prediction of Phytoplankton Biomass and Its Environmental Controls in Lowland Rivers

Devanshi Pathak[1,2], **Michael Hutchins**[1], **Lee Brown**[2], **Matthew Loewenthal**[3], **Peter Scarlett**[1], **Linda Armstrong**[1], **David Nicholls**[1], **Michael Bowes**[1], **and François Edwards**[1]

[1]UK Centre for Ecology and Hydrology, Wallingford, UK, [2]School of Geography and water@leeds, University of Leeds, Leeds, UK, [3]Environment Agency, National Water Quality Instrumentation Service, Reading, UK

Key Points:

- Hourly model is developed for lowland river phytoplankton prediction and is tested in the River Thames using high-frequency observations
- Blooms are mainly influenced by physical controls and develop at low-mid flows (21–63 $m^3 s^{-1}$) and temperatures (11–18°C)
- High-resolution modeling improves phytoplankton prediction in the river

Supporting Information:

- Supporting Information S1
- Data Set S1

Correspondence to:
D. Pathak,
devpat@ceh.ac.uk

Citation:
Pathak, D., Hutchins, M., Brown, L., Loewenthal, M., Scarlett, P., Armstrong, L., et al. (2021). Hourly prediction of phytoplankton biomass and its environmental controls in lowland rivers. *Water Resources Research*, *57*, e2020WR028773. https://doi.org/10.1029/2020WR028773

Received 8 SEP 2020
Accepted 6 JAN 2021

Abstract High-resolution river modeling is valuable to study diurnal scale phytoplankton dynamics and understand biomass response to short-term, rapid changes in its environmental controls. Based on theory contained in the Quality Evaluation and Simulation Tool for River-systems model, a new river model is developed to simulate hourly scale phytoplankton growth and its environmental controls, thus allowing to study diurnal changes thereof. The model is implemented along a 62 km stretch in a lowland river, River Thames (England), using high-frequency water quality measurements to simulate flow, water temperature, dissolved oxygen, nutrients, and phytoplankton concentrations for 2 years (2013–2014). The model satisfactorily simulates diurnal variability and transport of phytoplankton with Nash and Sutcliffe Efficiency (NSE) > 0.7 at all calibration sites. Even without high-frequency data inputs, the model performs satisfactorily with NSE > 0.6. The model therefore can serve as a powerful tool both for predictive purposes and for hindcasting past conditions when hourly resolution water quality monitoring is unavailable. Model sensitivity analysis shows that the model with cool water diatoms as dominant species with an optimum growth temperature of 14°C performs the best for phytoplankton prediction. Phytoplankton blooms are mainly controlled by residence time, light and water temperature. Moreover, phytoplankton blooms develop within an optimum range of flow (21–63 $m^3 s^{-1}$). Thus, lowering river residence time with short-term high flow releases could help prevent major bloom developments. The hourly model improves biomass prediction and represents a step forward in high-resolution phytoplankton modeling and consequently, bloom management in lowland river systems.

1. Introduction

The biomass and composition of phytoplankton are important indicators of water quality and the biological health of rivers (Villegas & de Giner, 1973). Phytoplankton communities are a major source of food for primary consumers through organic carbon production, and act as the primary source of oxygen in many rivers (Köhler, 1995). However, rivers also suffer if there is excessive phytoplankton growth since it may cause oxygen depletion (Hilton et al., 2006), produce harmful toxins (e.g., cyanobacteria), increase water treatment costs (P. Whitehead & Hornberger, 1984), and interfere with fishing and other recreational activities (Paerl & Huisman, 2009). Hence, it is vital to build an understanding of the controls on phytoplankton, to predict and prevent harmful growth in rivers, especially under the changing climate and environmental conditions (Read et al., 2014).

Phytoplankton growth in rivers is influenced by different environmental controls, and their influence may vary depending upon the river characteristics and local conditions (Reynolds, 2000). River phytoplankton development is often linked to increase in nutrient concentrations (Dodds, 2006; Minaudo et al., 2018). However, a number of recent studies have also highlighted the role of physical factors such as residence time (Reynolds, 2000), light availability (Domingues et al., 2011), and temperature (Canale & Vogel, 1974) in shaping river phytoplankton populations. These environmental controls generally act in combination to control phytoplankton blooms in rivers such as the combination of flow and light (Hardenbicker et al., 2014), flow, temperature, and nutrients (Larroudé et al., 2013), or flow, temperature, and light (Balbi, 2000; M. Bowes et al., 2016).

Several studies have addressed environmental controls of phytoplankton growth in lowland rivers around the world. Lowland rivers are heavily impacted by water abstractions, artificial flow regulation, physical modifications of the channel as well as substantial pollution load discharges from sewage and agricultural runoff (M. Hutchins et al., 2018). These slow flowing rivers typically have long residence time, which provides sufficient time for water quality to be sensitive to in-stream biogeochemical processes and for phytoplankton to utilize nutrients and grow (Reynolds, 2000). For example, River Thames (United Kingdom) exhibited high phytoplankton biomass only at low flows (Lack, 1971). River Meuse (Belgium) showed a combination of physical factors such as flow, temperature, and light as important biomass controls (Everbecq et al., 2001). Total phosphorus in the Rideau (Canada) (Basu & Pick, 1997) and grazing in the Moselle (France and Germany) (Descy et al., 2003) have also been found to be important biomass controls. However, most previous studies examined changes in these multiple control relationships at daily to monthly time scales, with higher resolution diurnal dynamics often being overlooked. Phytoplankton biomass in rivers have shown to respond to rapid changes in its environmental controls (M. Bowes et al., 2016), and thus, it is crucial to understand these shorter scale dynamics to predict phytoplankton growth and bloom timings more accurately.

In the River Thames, a regulated lowland river in southern England, substantial efforts have been made to understand phytoplankton response through process-based modeling (e.g., Lázár et al., 2016; P. Whitehead & Hornberger, 1984; P. G. Whitehead et al., 2015). These studies, however, have shown limited predictive ability in modeling large and rapidly developing phytoplankton biomass, suggesting that there is still a need to better understand process interactions (M. Bowes et al., 2016). For example, there has been uncertainty as to which phytoplankton groups dominate the response. Some studies found green algae to be dominant during peaks in the summer (Lack, 1971; Ruse & Love, 1997) yet subsequent studies have found cool water diatoms to predominate with peaks instead during spring and autumn seasons (M. Bowes et al., 2016; Read et al., 2014). The daily time-step Quality Evaluation and Simulation Tool for River-systems (QUESTOR) model (Boorman, 2003a) has been developed to simulate phytoplankton (M. Hutchins et al., 2010) and also been extensively applied in the River Thames (M. Hutchins et al., 2016, 2018; Waylett et al., 2013). However, consistent with other models, QUESTOR applications have also had only limited success in simulating phytoplankton biomass with overestimation in mid-summer (M. Hutchins et al., 2016).

Various process-based river models (e.g., Brown & Barnwell, 1987; Everbecq et al., 2001; Reichert et al., 2001) have been applied worldwide to understand phytoplankton dynamics, but these models are rarely tested with high-frequency observations to explore shorter scale dynamics. Phytoplankton modeling applications are generally limited to weekly to daily time-steps. High-resolution modeling has been challenging because of high computational requirements and a lack of high-frequency monitoring data. Even the high-resolution hourly modeling studies done so far (Martin et al., 2013; Minaudo et al., 2018; Suarez et al., 2019; Van Griensven & Bauwens, 2005) do not test the simulations with high-frequency observations of all key variables (temperature, chlorophyll, and dissolved oxygen [DO]) and only report model testing against daily (or coarser scale) observations. However, it is now easier to monitor water quality at higher temporal resolutions with the development of low-cost, robust water quality sensors (Rode et al., 2016). Models can utilize high-frequency datasets to understand phytoplankton growth and its environmental controls at finer resolutions, and thus, ensure early warnings of blooms in river systems.

The present study was undertaken by modifying an existing water quality model, QUESTOR, to run at shorter time steps and testing against high-frequency (hourly) water quality measurements (chlorophyll-a [Chl-a], DO, and water temperature) at two locations in the River Thames. The model testing was reinforced with testing against daily flow observations and weekly water quality observations at other locations and for other determinands. The overall aim of this study was to test the hypothesis that hourly time-step modeling can improve prediction of phytoplankton biomass and to demonstrate the utility of the model to study phytoplankton dynamics and its controls in lowland rivers. Specific objectives were to:

(1) Develop a model to predict hourly variation and transport of in-stream flow, temperature, nutrients, DO, and phytoplankton biomass in the lower Thames
(2) Identify an accurate model structure that represents dominant phytoplankton groups in the lower Thames using a model comparison and sensitivity analysis
(3) Illustrate the extent to which low-frequency water quality observations, used as inputs to the model in the absence of high-frequency observations, can still provide a basis for satisfactory explanation of phytoplankton dynamics in the catchment

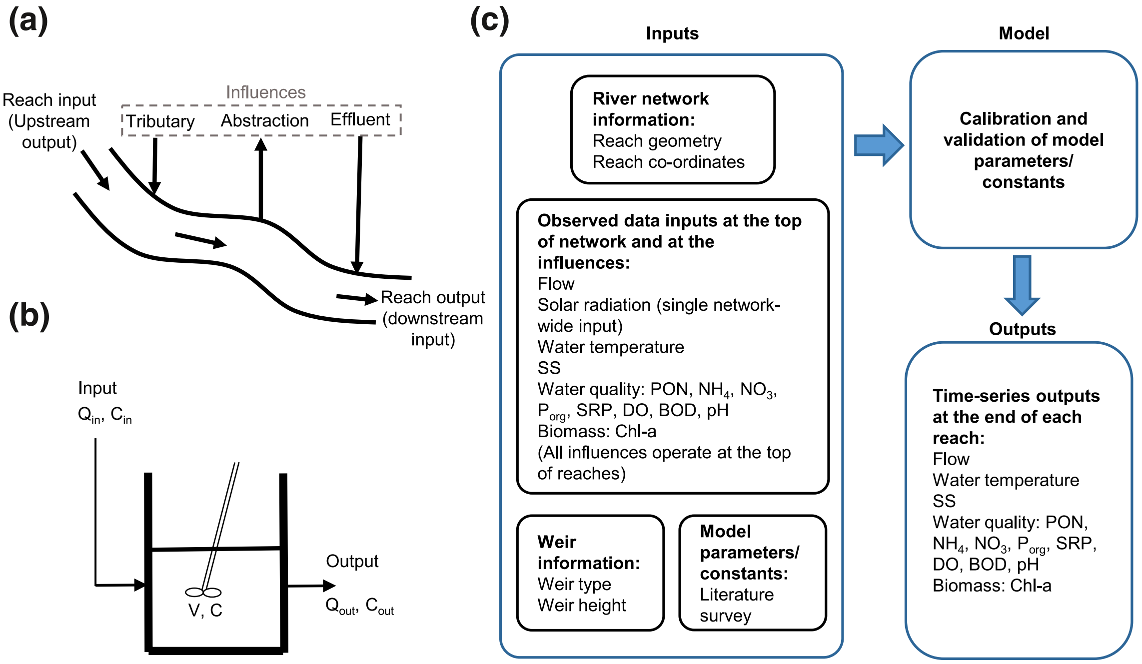

Figure 1. Model structure. (a) represents the schematic of a typical reach in the model, (b) represents the conceptualization of reaches in the model and (c) shows the input and output information in the hourly model. (a) and (b) are modified after P. Whitehead et al. (1997). In (b), V represents volume of water in a reach, C represents concentration of water quality determinand, Q_{in} and Q_{out} are input and output flows, C_{in} and C_{out} are input and output concentrations in a reach.

(4) Identify favorable environmental conditions for photosynthetic production using the hourly model outputs and evaluate association of phytoplankton biomass with multiple environmental controls using boosted regression trees (BRT) technique

2. Methods

2.1. Model Theory

QUESTOR is an in-stream, process-based water quality model that allows users to represent rivers as a network of reaches. River reaches are modeled as a set of nonlinear reservoirs or well-mixed tanks in series (Figure 1). The hourly model is a pseudo 1-D (strictly speaking 0-D) model and assumes fixed channel width with rectangular cross section. The model simulates dynamic solute transport within the river network using ordinary differential equations (ODEs) with a mass-balance approach. The numerical solution of the ODEs is implemented using an explicit fourth-order Runge-Kutta-Merson differential equation solver (DASCRU). This operates at variable temporal resolution much finer than hourly reporting of results. The equations characterize major processes affecting model determinands, but include empirical coefficients which need to be calibrated (Boorman, 2003b). We modify the *Stephanodiscus hantzschii* version (SH module, Waylett et al., 2013) of daily time-step QUESTOR model to account for hourly variations in physico-chemical and biological water quality. The key determinands in the model include flow, water temperature, photosynthetically active radiation (PAR), Chl-a, biochemical oxygen demand (BOD), DO, nitrate (NO_3), ammonium (NH_{4+}), particulate organic nitrogen (PON), as well as organic and inorganic phosphorus in the water column (Figure 2). The equations for the new hourly model version of phytoplankton are explained below and the remaining set of equations for other determinands are provided in Text S1. The model parameters are listed in Table S3.

The growth of phytoplankton is estimated using a fixed stoichiometry model where the ratios (by weight) Chl-a:C:N:P are 1:50:10:1 (M. Hutchins et al., 2010). The new version required modifications in the phytoplankton growth rate parameters as described in Equation 1,

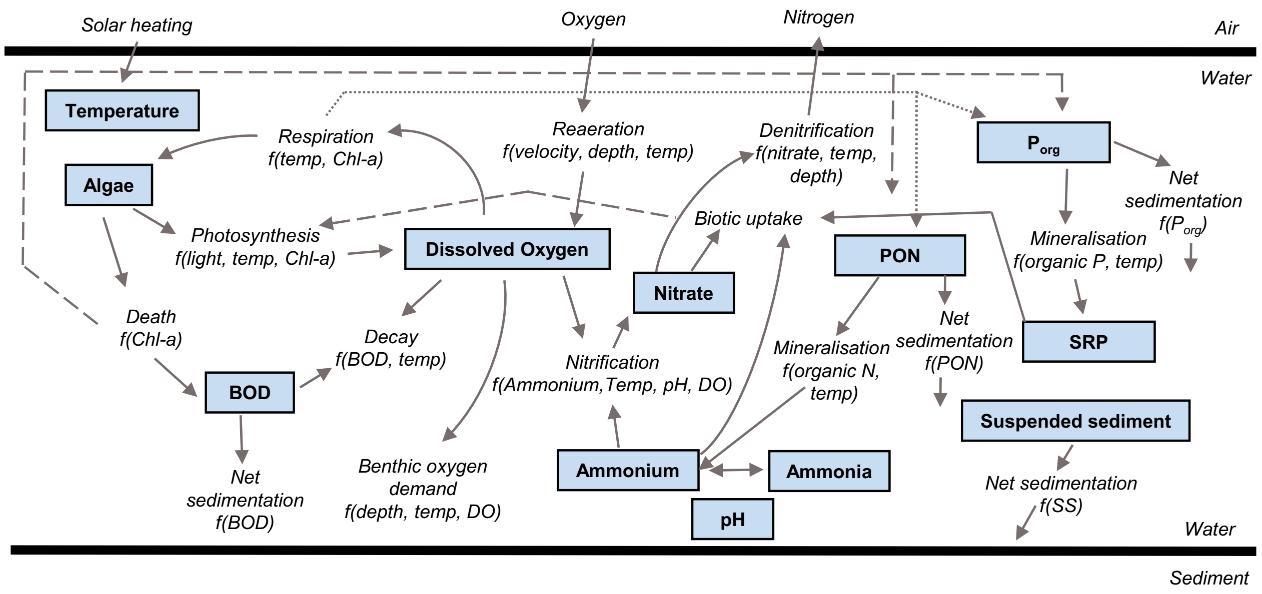

Figure 2. In-stream water quality determinands and processes represented in the model (modified after Eatherall et al., 1998).

$$k^{\text{pho}} = \text{Phy}.k_{\text{ref}}^{\text{pho}}.e\left(-\frac{\left(\text{Acutal Temp} - \text{Optimum Temp}\right)^2}{a^2}\right).f\left(N\right).f\left(L\right) \tag{1}$$

k^{pho} = photosynthetic rate (mg L^{-1} h^{-1}); Phy = concentration of Chl-a (mg L^{-1}); $f(N)$ and $f(L)$ = nutrients and light limitation factors, each holding values between 0 and 1; Optimum Temp = 14°C and a = 8°C (Descy et al., 2003); $k_{\text{ref}}^{\text{pho}}$ = maximum phytoplankton growth rate (h^{-1}) at T_{ref}; T_{ref} = 20°C.

Temperature limitation factor ($f(T)$) is estimated using Equation 2,

$$f\left(T\right) = e\left(-\frac{\left(\text{Acutal Temp} - \text{Optimum Temp}\right)^2}{a^2}\right) \tag{2}$$

The calculation of nutrient limitation uses Michaelis Menten kinetics (Equation 3),

$$f\left(N\right) = \min\left(\frac{N}{N + k_N}, \frac{P}{P + k_P}\right) \tag{3}$$

N = nitrate-N plus Ammonium-N (mg L^{-1}); P = inorganic (soluble reactive phosphorus [SRP]) plus organic phosphorus (mg L^{-1}); k_N (mg N L^{-1}) and k_P (mg P L^{-1}) are the half-saturation constants for N and P in phytoplankton, respectively.

To account for light limitation, attenuation with depth (γ, m^{-1}) is described by the Beer-Lambert Law (Equation 4),

$$\gamma = \gamma_{\text{base}} + L_{\text{SS}}.\text{SS} + L_{\text{Phy}}.\text{Phy} \tag{4}$$

γ_{base} = light extinction coefficient in clean water (m^{-1}); SS = concentration of suspended sediment (mg L^{-1}); L_{ss} = light attenuation with depth due to suspended sediment (m^{-1} mg^{-1} L); L_{phy} = light attenuation with depth due to phytoplankton (m^{-1} mg^{-1} L).

Estimation of photolimitation with respect to phytoplankton-specific optimum intensities (Steele, 1962) in the model requires hourly inputs of incoming radiation and a constant value of optimum light intensity (Equation 5),

$$f(L) = \frac{2.718}{\gamma d}.\left[e^{-\frac{R_s L_1 L_2}{L_{\text{opt}}} e^{-\gamma d}} - e^{-\frac{R_s L_1 L_2}{L_{\text{opt}}}} \right] \tag{5}$$

d = water column depth (m); R_s = radiation at the surface not reflected (W m^{-2}) (i.e., raw data × L_3); L_1 = fraction of incoming radiation that is visible light; L_2 = fraction of visible light used for phytoplankton; L_3 = fraction of light reaching water surface that is not reflected; L_{opt} = optimum light intensity for phytoplankton (W m^{-2}).

Respiration calculation requires estimates of respiration fraction and maximum phytoplankton growth rate (Equation 6),

$$k^{\text{res}} = \text{Phy}.k_{\text{ref}}^{\text{res}}.k_{\text{ref}}^{\text{pho}}.\theta^{(T-T\text{ref})} \tag{6}$$

k^{res} = phytoplankton respiration rate (mg L^{-1} h^{-1}); $k_{\text{ref}}^{\text{res}}$ = reference respiration fraction for phytoplankton (as fraction of $k_{\text{ref}}^{\text{pho}}$); θ = Arrhenius factor for temperature dependencies (θ=1.08); T = temperature (°C); T_{ref} = 20°C.

The death of the autotrophs is a combination of grazing and nonpredatory mortality. In QUESTOR, death rate is estimated from algal growth limitation due to unsuitable nutrient and light conditions and the calibration of the death rate constant compensates for the death from grazing (Equation 7),

$$k^{\text{death}} = \text{Phy}.k_{\text{ref}}^{\text{death}}.k_{\text{ref}}^{\text{pho}}.\left[1-\left(f(N).f(L)\right)\right].\theta^{(T-T\text{ref})} \tag{7}$$

k^{death} = phytoplankton death rate (mg L^{-1} h^{-1}); $k_{\text{ref}}^{\text{death}}$ = reference death fraction for phytoplankton (as fraction of $k_{\text{ref}}^{\text{pho}}$).

2.2. Study Area

The Thames catchment is situated in southern England with an area of 9,948 km^2 at the tidal limit (Waylett et al., 2013). The catchment has a population of around 15 million people with its uplands characterized by arable and pasture, and the lowland areas covered mainly by urban land uses (M. Hutchins et al., 2018). Mean annual precipitation and mean daily temperature in the catchment are 700 mm and 11°C, respectively (Crossman et al., 2013). The catchment receives around 40% of its water supply from groundwater sources, which are mainly characterized by Oolitic Limestone and Cretaceous Chalk aquifers (Crossman et al., 2013). The River Thames is a heavily regulated river with 45 locks and weirs along its course. Catchment rivers that have extensive lock systems or are connected to canals are characterized by higher phytoplankton biomass compared to the unconnected rivers because of longer residence times (M. Bowes et al., 2012). In general, high phytoplankton biomass is observed in the middle and lower reaches of the catchment and phytoplankton blooms mainly occur during March to July (M. Bowes et al., 2012). This study focuses on a 62 km long stretch in the catchment from Caversham to Runnymede (Figure 3). The river stretch receives inputs from major tributaries such as Kennet, Lodden, and Wye rivers, and Sewage Treatment Works (STW) effluents. In-stream flows are regulated by 14 locks and weirs along the stretch, sometimes resulting in long residence times from reduced flow velocities.

2.3. Data Source and Model Application

The River Thames is one of the most intensively studied and monitored rivers in the United Kingdom (M. J. Bowes et al., 2018). The model development here makes use of a combination of weekly to hourly scale existing flow and water quality data of 2 years 2013–2014. Gauged daily flow data were obtained from

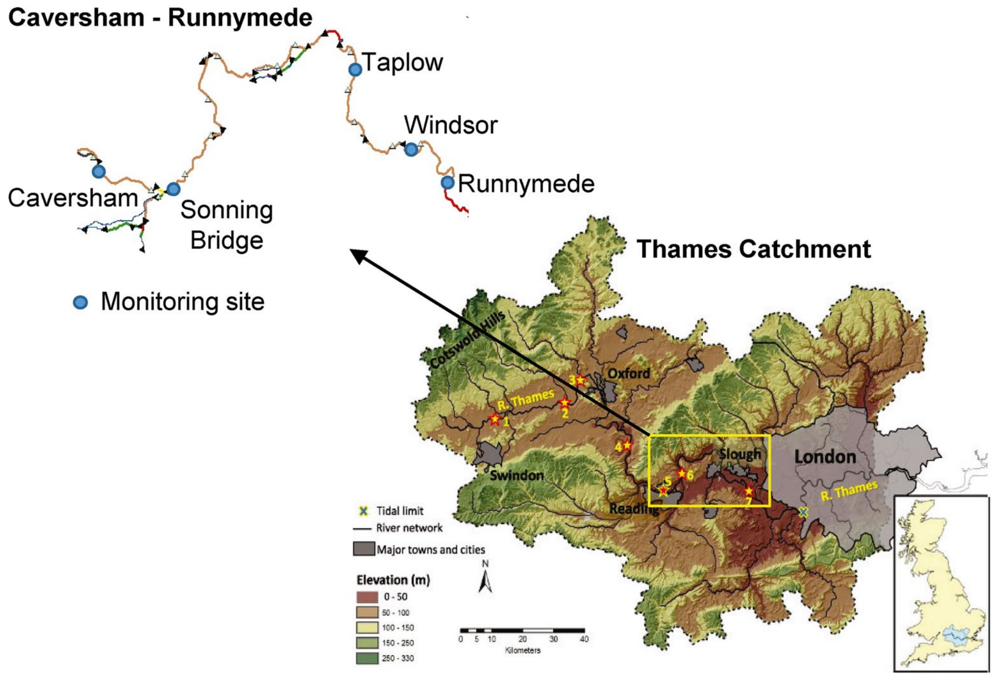

Figure 3. River Thames catchment with monitoring locations. Site 5 = Caversham; Site 7 = Runnymede. (modified after M. Bowes et al., 2016).

the NRFA (NERC, National River Flow Archive, http://www.ceh.ac.uk/data/nrfa/) and were interpolated to hourly time-steps for this application. Nutrient data at multiple sites along the Thames and its major tributaries, sampled at weekly intervals, were obtained from the UK Centre for Ecology & Hydrology's (UKCEH) Thames Initiative research platform (M. J. Bowes et al., 2018). Hourly water temperature, DO and Chl-a concentrations at Caversham, Taplow and Windsor sites (Figure 3) were generated by the Environment Agency's (EA) National Water Quality Instrumentation Service, using YSI6600 sensors, calibrated every 3 weeks. The quality control procedure for the data collection is provided by Waylett et al. (2013). Chl-a concentration provides a proxy for phytoplankton concentration (M. Bowes et al., 2012). For chlorophyll observations, aside from the hourly frequency sensor data, the model uses UKCEH's standard laboratory methanol extraction-based weekly chlorophyll observations at Sonning Bridge and Runnymede sites (Figure 3).

The model uses single hourly radiation time-series for the whole catchment obtained from the British Atmospheric Data Centre (MIDAS Landsat data) (http://archive.ceda.ac.uk/) for Little Rissington near the River Windrush in Gloucestershire (NGR 4299 2107). The hourly radiation time-series was modified to account for canopy shading from riparian trees using a fraction of potential solar radiation reaching the river surface as recommended by Waylett et al. (2013). Data to define tributary and sewage inputs were derived from the online UKCEH (doi: https://doi.org/10.5285/e4c300b1-8bc3-4df2-b23a-e72e67eef2fd) and EA (http://environment.data.gov.uk/water-quality/view/landing) datasets. Weir height and type within the river stretch were adapted from the previous model application in the River Thames (P. Whitehead & Hornberger, 1984). Flow-velocity relationships were derived using a set of three linearized velocity equations that reflect the river hydromorphology and the lock operations in the river (Waylett et al., 2013; P. Whitehead & Hornberger, 1984).

To establish confidence in the model calibration, testing of phytoplankton response during bloom periods is crucial. Often only one bloom period occurs each year. The hourly model was calibrated using the observed data of the year 2013, which encompassed one large prolonged bloom and one medium-sized bloom, and the model setup was validated using the observed data of the year 2014, which covered two distinct medium-sized blooms. The whole study stretch was divided into 23 reaches (Table S1), accounting for the influence of tributaries, weir locations, abstractions and sewage treatment works. The hourly model used data at the top of the modeled river stretch (here Caversham) and at the influences as inputs (Figures 1 and 3). Model calibration was carried out using observations at four sites (Sonning Bridge, Taplow, Windsor, and Runnymede) in the stretch. The calibration was performed using a sequential procedure determinand by determinand, by modifying process-rate parameters, working downstream from site to site. The order of this one by one parameter calibration and the list of sources and sinks of the variables affected by each processes are explained in detail elsewhere (Waylett et al., 2013). The model performance is judged using a combination of Nash and Sutcliffe Efficiency (NSE) and percentage error in mean (PBIAS) statistics (Text S1).

2.4. Sensitivity Analysis

Previous studies have reported dominance of diatom populations in spring and autumn seasons and green algal groups in summer periods (Lack, 1971; Read et al., 2014; Ruse & Love, 1997). The difference in the timings of dominance can be attributed to their different temperature requirements for growth. Green algae have a higher optimum growth temperature that results in their abundance in summer due to higher temperatures compared to spring/autumn. We tested the hourly model with three algal group representations, namely, *Stephanodiscus hantzschii* (SH), Green algae (GA) (including chlorophytes and cryptophytes) and small centric diatoms (SCD) (up to 15 μm diameter), differentiated based on their temperature preferences. The three models, SH, GA, and SCD, use optimum temperatures of 14°C, 24.5°C, and 21°C and a values of 8°C, 14°C, and 12°C, respectively in Equation 1 that have been optimized for another lowland river (Descy et al., 2003). All three models were calibrated for growth, death and respiration rate parameters and were compared with observations to derive the best model that represent phytoplankton populations in the River Thames. We also investigate the temperature constraints on Thames phytoplankton by assessing the goodness of fit of model outputs under a range of assumptions of optimum temperature for growth ranging from 10°C to 26°C with an increment of 2°C.

To evaluate the model sensitivity to change in the temporal resolution of the input data, we test the hourly model setup with low-frequency water quality inputs. Our model has a requirement of hourly time-step input data, which may not be easier to fulfil at times and require recourse to extensive interpolation of sparse data values. Data regarding radiation and hydrology are often available from routine monitoring, but high-frequency water quality information is still difficult to gather. The model setup here uses daily flow data, and hourly water temperature, DO, and chlorophyll data as inputs at Caversham. We filtered weekly data points (Monday, 11 a.m.) from these high-frequency datasets, and interpolated the weekly spot samples' time-series to hourly time-steps to use as inputs in the hourly model. This way we assess model's applicability both to generate past conditions pre-2000 in the Thames before high-frequency monitoring was established and in other catchments where only low-frequency flow and water quality monitoring is practiced.

2.5. Statistical Analysis

We use BRT, a machine learning technique, to evaluate the associations between chlorophyll and multiple environmental variables. BRT can handle continuous, collinear variables, support nonlinear variables with missing data and help identify interactions between explanatory variables (Elith et al., 2008). Recently, it has been widely used to link biological water quality with multiple environmental variables simulated from process-based models (Feld et al., 2016; Rankinen et al., 2019). Here, we use the simulated/calculated environmental variables (flow, water temperature, nitrate, SRP, and PAR) from the hourly model and link them with continuous chlorophyll observations during 2013–2014 at Windsor. We use R packages gbm (Ridgeway, 2020) and dismo (Hijmans et al., 2017) to run BRT analysis. We also use pair-wise boxplots at Windsor site to identify favorable environmental conditions within which phytoplankton blooms develop. Hourly model outputs of controlling variables were divided in 10 equal quantile groups, and were plotted against chlorophyll simulations to derive environmental bounds of phytoplankton growth.

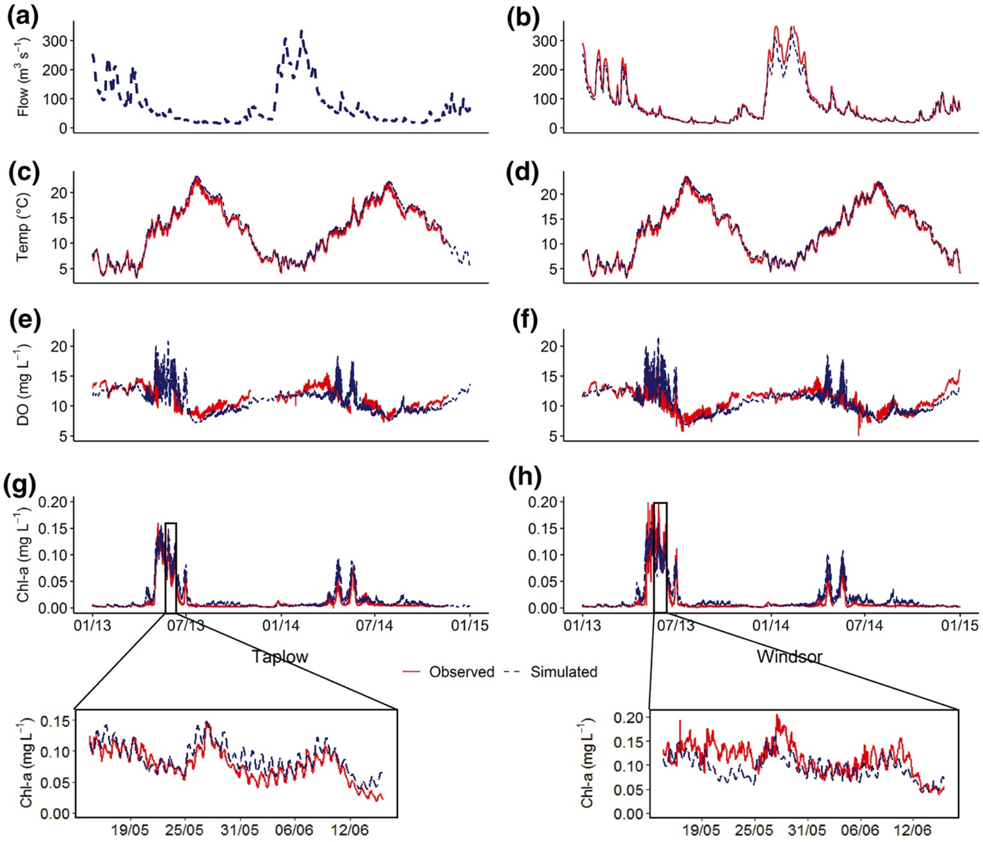

Figure 4. Modeled and observed hourly flow, water temperature (Temp), DO, and chlorophyll concentrations (Chl-a) at Taplow and Windsor for calibration and validation runs. Note that observed flow data were not available at Taplow. DO, dissolved oxygen.

3. Results

We present the model results in four main parts: (1) the hourly model performance in simulating environmental controls and phytoplankton biomass, (2) temperature preferences of the dominant phytoplankton species in the model, (3) model testing with lower temporal resolution of input data, and (4) identification of environmental controls and their influence on phytoplankton biomass.

3.1. Model Calibration and Validation

3.1.1. Flow and Water Temperature Prediction

The Thames catchment is characterized by high winter flows that decrease in early spring and remain very low during summer and autumn. The hourly model successfully captures this seasonality in flow at Windsor (Figure 4). The overall flow simulation indicates an underestimation of flow volume at Windsor with a percentage mean error (PBIAS) of −9.88 and −11.58 for the calibration and validation periods, respectively (Table 1). In spite of the underestimation, the model satisfactorily simulates flow variation at Windsor with very good NSE values of 0.96 and 0.95 for calibration and validation periods, respectively.

Water temperature in the model is controlled mainly by heat gain or loss from radiation, canopy shading, and the temperature of flow volumes entering the main channel. The average hourly temperature variation shows a clear cycle throughout the year, with daily minimum temperature observed in the early morning hours (6:00–8:00) and daily maximum temperature observed in the late afternoon (16:00–17:00). We also compared the temperature simulations with hourly observations at Taplow and Windsor (Figure 4), and with weekly observations at Sonning Bridge and Runnymede sites (Table 1). This comparison shows that

Table 1
Model Performance Statistics for Calibration (2013) and Validation Periods (2014)

		Sonning bridge		Taplow		Windsor		Runnymede	
Period	Determinand	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)
Calib (2013)	Flow					0.96	−9.88		
	Temp	0.99	−2.14	0.98	3.80	0.98	3.52	0.99	1.58
	DO	0.52	12.22	0.49	−4.47	0.46	−0.93		
	Chl-a	0.81	−19.9	0.87	26.16	0.80	12.71	0.73	−34.64
	Nitrate	0.21	−4.17					0.31	1.74
	SRP	0.77	2.92					0.75	16.24
Valid (2014)	Flow					0.95	−11.58		
	Temp	0.98	−3.87	0.98	3.61	0.99	2.35	0.99	−0.9
	DO	0.25	14.99	0.43	−5.11	0.58	−4.29		
	Chl-a	0.78	−20.06	−0.19	60.93	0.20	76.73	0.77	−16.58
	Nitrate	−0.07	−5.97					0.46	1.55
	SRP	0.80	5.47					0.71	16.11

Abbreviations: DO, dissolved oxygen; NSE, Nash and Sutcliffe Efficiency; PBIAS, percentage error in mean; SRP, soluble reactive phosphorus.

Notes. Calib and valid represents calibration and validation periods, respectively. Chl-a and Temp represents chlorophyll and water temperature, respectively. NSE stands for Nash-Sutcliffe Efficiency and PBIAS represents percentage error in mean.

the temperature model performs in strong agreement with the observations throughout the study stretch, with NSE ≥ 0.98 and mean errors within ±4% at all calibration sites (Table 1).

3.1.2. Water Chemistry and DO Prediction

The phosphorus model successfully captures the seasonal trend and magnitude of SRP concentrations with NSE > 0.7 at Sonning Bridge and Runnymede (Figure 5). Mean error for SRP modeling is around 16% at Runnymede, whereas Sonning Bridge shows lower errors (3%–6%). The nitrogen model shows relatively poorer fits for nitrate with NSE values ranging from −0.07 to 0.46. As opposed to the SRP model, the nitrate model performs better at Runnymede (NSE = 0.31, 0.46) compared to Sonning Bridge (NSE = 0.21, −0.07)

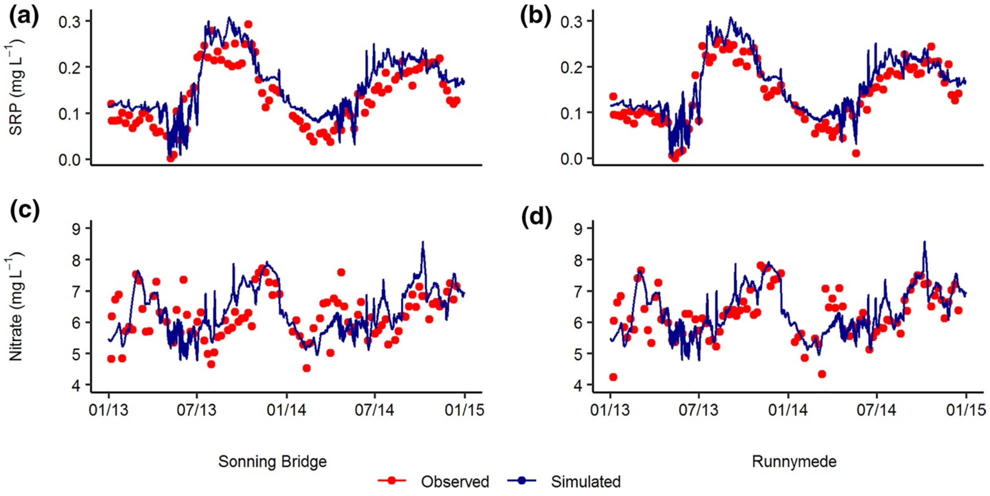

Figure 5. Hourly modeled and observed SRP and nitrate concentrations at Sonning Bridge and Runnymede for calibration and validation runs. SRP, soluble reactive phosphorus.

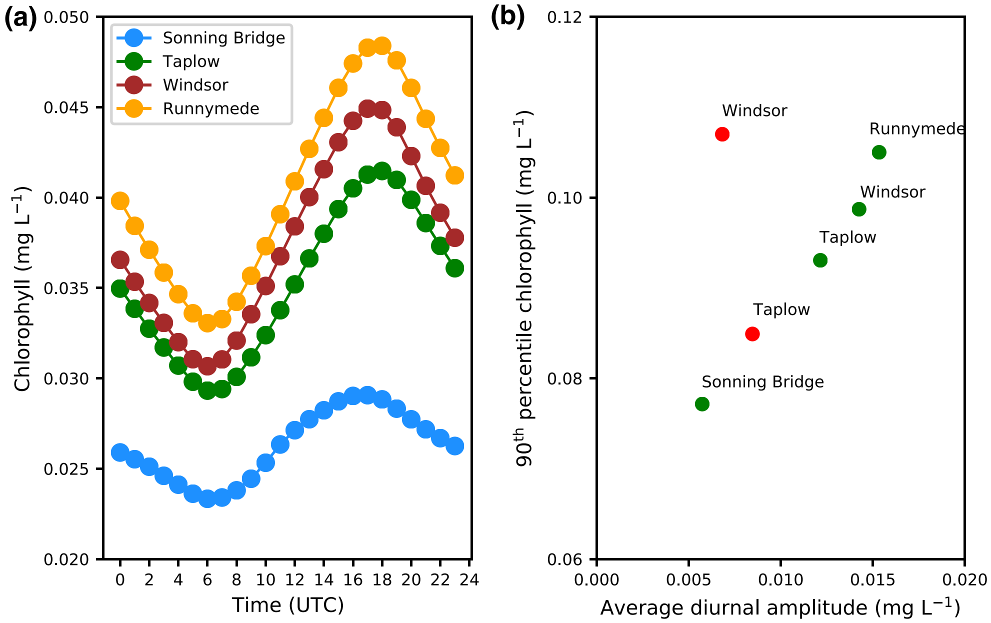

Figure 6. Spatial and temporal variation of phytoplankton biomass. (a) represents average diurnal phytoplankton variation in the model during growing period (April–July) at all calibration sites and (b) shows modeled (green markers) and observed (red markers) spatial variation in 90th percentile chlorophyll concentration with its average diurnal amplitude during growing period.

for calibration and validation periods. Sonning Bridge shows an overall underestimation (up to 6%), and Runnymede shows an overall overestimation (up to 2%) of nitrate concentrations.

DO concentrations in the catchment begin increasing in spring, but drops to minimum levels in mid-summer. We observe a high diurnal variability in DO, and this coincides with increased photosynthesis and respiration during phytoplankton blooms, both successfully captured by the model (Figure 4). The magnitude of DO concentrations, during the bloom period, is overestimated. During the rest of the year, DO concentrations are generally underestimated. The calibration run shows satisfactory DO fits with NSE > 0.45 at all calibration sites (Table 1). The validation run also shows satisfactory fits at Taplow and Windsor (NSE > 0.43). Model performance at Sonning Bridge, however, drops slightly in the validation run (NSE = 0.25). The model satisfactorily captures the seasonal behavior of DO concentrations at all calibration sites (Figure 4, Table 1).

3.1.3. Phytoplankton Prediction

Phytoplankton observations during the calibration period (2013) shows much higher levels of peak blooms (up to 0.2 mg L^{-1}) compared to the validation period (2014) with less than half the magnitude (up to 0.1 mg L^{-1}) of the 2013 blooms. Peak levels in the calibration period are underestimated, but are overestimated for the validation period (Figure 4). Phytoplankton model shows a good performance for the calibration run with NSE > 0.7 at all sites (Table 1). Mean errors, however, are relatively high (−35%–26%) for the calibration run. The calibration run indicates an overall underestimation at Sonning Bridge and Runnymede, and an overestimation at Taplow and Windsor. For the validation run, the model performs well (NSE > 0.7, PBIAS up to 20%) at Sonning Bridge and Runnymede, but with relatively poorer fits at Windsor (NSE = 0.20, PBIAS = 77%) and Taplow (NSE = −0.19, PBIAS = 61%). On average during the growing season, daily minimum and maximum phytoplankton levels are modeled around 6:00 h and 17:00–18:00 h, respectively (Figure 6a). Modeled diurnal variability of phytoplankton agrees well with the observations, but the model underestimates biomass magnitude. The model predicts an increase in the bloom size and diurnal amplitude from upstream to downstream (Figure 6b). Observations also show an increase in the bloom size from Taplow to Windsor, but show reduction in the diurnal amplitude. Overall, the model identifies the timing of multiple blooms and collapses during the growing season and successfully models phytoplankton dynamics along the river stretch.

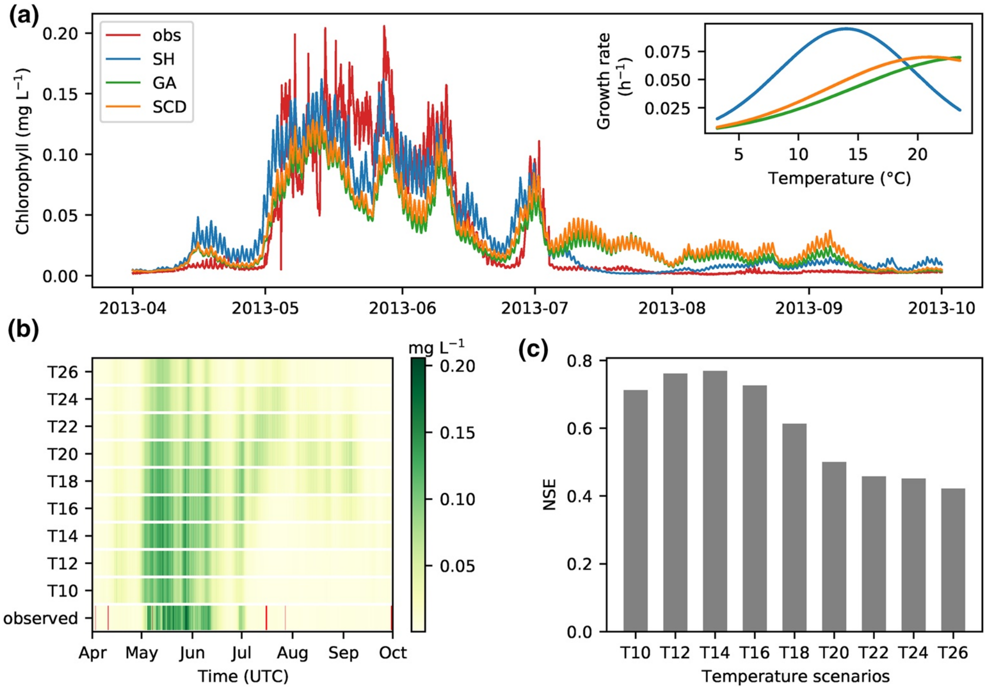

Figure 7. Temperature preferences of phytoplankton populations. (a) shows a comparison of modeled and observed chlorophyll concentrations for the April–September, 2013 period for SH, GA, and SCD models at Windsor. (b) represents a comparison of model performance for a range of optimum temperatures from 10°C to 26°C with an increment of 2°C at Windsor. T10 to T26 represent different optimum temperature scenarios for phytoplankton growth. Red lines represent missing data periods. (c) shows goodness of fit statistics (NSE) for all temperature scenarios from T10 to T26 at Windsor. GA, Green algae; SCD, small centric diatoms; SH, *Stephanodiscus hantzschii*.

3.2. Optimum Temperature Representation for Phytoplankton Growth

All three models, SH, GA, and SCD, simulate high biomass during April–September as shown in Figure 7a, which is when we mainly see the differences in performance among the models. Although all three models perform well (NSE > 0.6) at Windsor during this period, SH captures the high concentrations most accurately, which is crucial for water quality management. Moreover, GA and SCD models overestimate biomass from mid-July to mid-September as opposed to the SH model that performs the best to capture low concentrations as well. High optimum growth temperatures in GA (24.5°C, a = 14°C) and SCD (21°C, a = 12°C) models prompt the algal growth after July with increasing water temperature (Figure 7a). Growth rate in the SH model, on the other hand, starts decreasing after an optimum temperature of 14°C (a = 8°C), which agrees well with the observations. Sensitivity analysis with the SH model shows better performance (Figure 7c) at low optimum temperatures (10°C–16°C) to simulate phytoplankton blooms during May–June (Figure 7b). At higher temperatures (>16°C), there is an overestimation of biomass after July. Lower temperatures (<12°C), on the other hand, underestimate the blooms at the start of July. The best model fit (NSE = 0.77) is obtained for T14 scenario at an optimum temperature of 14°C (Figure 7c).

3.3. Model Sensitivity

The distribution of the simulated chlorophyll concentrations in the weekly input run do not change significantly from that of the hourly input run (Figure 8). Model performance declines only marginally with weekly inputs (Table 2) and the model captures the phytoplankton blooms. However, the weekly input run still shows a bigger lag in simulating timings of the development and collapse of blooms compared to the hourly input run (Figure 8). The weekly input run also simulates higher phytoplankton growth during September and October months, than the hourly input run, when no blooms are seen in the observed data. In spite of this, the NSE statistics for the weekly input run remain above 0.6 at all sites for the calibration period (Table 2). Even for the validation period, the weekly input run shows NSE > 0.7 at Sonning Bridge and Runnymede. Statistically,

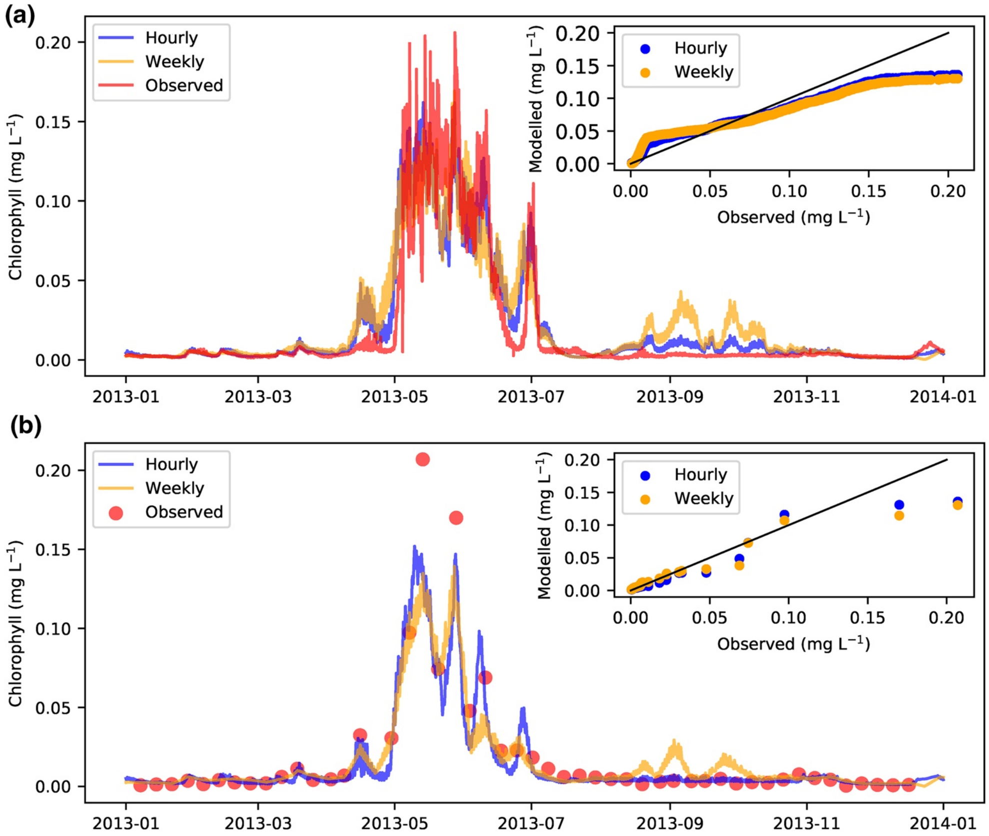

Figure 8. Comparison of modeled and observed chlorophyll concentrations for hourly and weekly input runs at Sonning Bridge (a) and Windsor (b). Inset figures represent quantile-quantile plots of model performances at Sonning Bridge and Windsor.

model performance at Taplow and Windsor is poor in both the runs for the validation period, although timings of blooms are represented well.

Importantly the weekly input run still uses hourly radiation time-series; both runs use single radiation time-series for the entire river stretch. We also tested the model with low-frequency radiation inputs, but it led to a significant and much larger drop in the model performance with NSE values changing from 0.87 (Taplow), 0.8 (Windsor), and 0.73 (Runnymede) to −2.57 (Taplow), −1.51 (Windsor), and 0.24 (Runnymede) with weekly radiation inputs. Even the daily scale radiation inputs affected the model performance heavily with NSE values of −2.46 (Taplow), −1.60 (Windsor), and 0.10 (Runnymede). This explains the key role of radiation inputs

Table 2
Comparison of Phytoplankton Model Performance for Hourly and Weekly Input Runs for Calibration and Validation Periods

		Sonning bridge		Taplow		Windsor		Runnymede	
Period	Model	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)	NSE (−)	PBIAS (%)
Calib (2013)	Hourly	0.81	−19.9	0.87	26.16	0.8	12.71	0.73	−34.64
	Weekly	0.83	−12.78	0.76	38.32	0.73	23.46	0.64	−29.63
Valid (2014)	Hourly	0.78	−20.06	−0.19	60.93	0.20	76.73	0.77	−16.58
	Weekly	0.83	−5.85	−0.82	90.62	−0.19	104.81	0.70	−3.72

Abbreviations: NSE, Nash and Sutcliffe Efficiency; PBIAS, percentage error in mean.

Notes. Calib and valid represents calibration and validation periods, respectively. NSE stands for Nash-Sutcliffe Efficiency and PBIAS represents percentage error in mean.

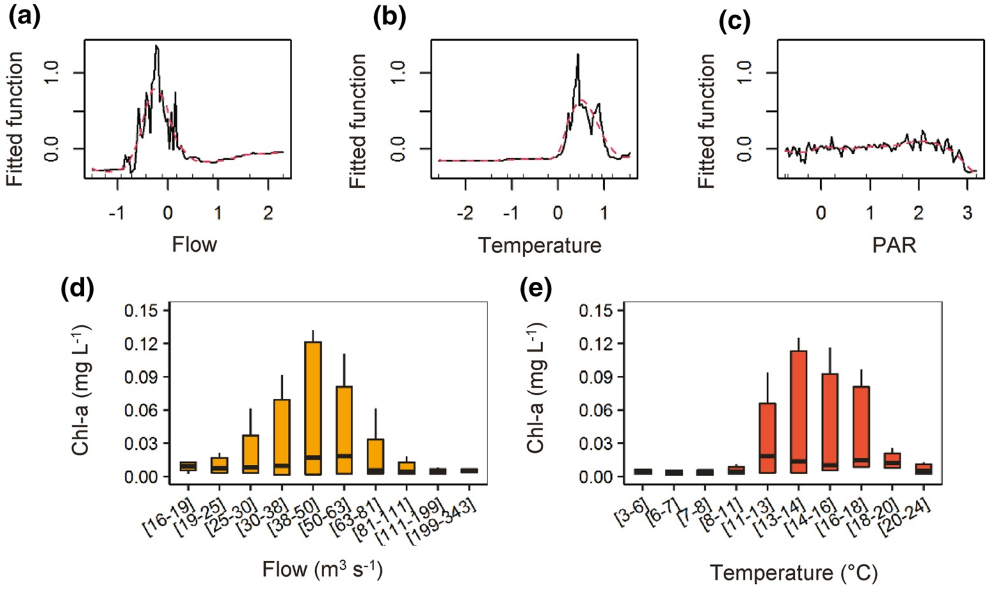

Figure 9. Multiple environmental controls of phytoplankton growth. (a–c) represents partial dependence plots for the modeled flow, water temperature and PAR with normalized fitted values of observed chlorophyll concentrations at Windsor in the BRT model. (d–e) represents boxplots of modeled hourly flow and water temperature against modeled hourly chlorophyll concentrations at Windsor. The hinges represent 10th and 90th percentile, and the line within the box represents median chlorophyll concentration. BRT, boosted regression trees; PAR, photosynthetically active radiation.

in modeling phytoplankton dynamics. Therefore, we conclude that the hourly model can successfully reproduce phytoplankton dynamics with low-frequency flow and water quality input data in the lower Thames, and the only input necessary at high-resolution in the model is radiation.

3.4. Physico-Chemical Controls on Phytoplankton Growth

A multiple stressor analysis using BRT technique shows maximum association of SRP (58%) with chlorophyll followed by water temperature (21.6%), flow (11.6%), nitrate (7.8%), and PAR (1.1%) (Figure S4). The reduction in nutrients (SRP and nitrate) with high chlorophyll levels suggests that nutrient levels during the growing season are influenced by phytoplankton growth rather than the other way around. It is not always clear which the dependent variable is since nutrients and chlorophyll relationships are intertwined, and hence, SRP and nitrate were removed from the list of predictors. A BRT model excluding the nutrient predictors highlights flow (55%) as the most influencing control followed by water temperature (38%) and PAR (7%). PAR does not represent a strong relationship with biomass variation in the BRT model (Figure 9c). Phytoplankton biomass increases with increase in flow and temperature until a certain limit of these controls is reached, and then shows a reversal of response with continued increase in the controls (Figure 9a and 9b). These relationships are also supported by pair-wise boxplots (Figures 9d–9e), which we use to identify the environmental bounds of flow and temperature within which phytoplankton blooms develop. High phytoplankton populations >0.03 mg L^{-1} develop with increase in temperature (11–18°C), but are not sustained at higher temperatures. Similarly, blooms only develop at low to mid flows between 30 and 63 m^3 s^{-1} at Windsor. This suggests that there is an optimum window of these controls where phytoplankton can bloom, and that outside this window, growth is not as strong.

Table 3
Seasonal Average of Limiting Factors for Phytoplankton Growth at Windsor Site as Calculated in the Hourly Model

Season	Light	Nitrogen	Phosphorus	Temperature
Spring	0.258	0.984	0.903	0.786
Summer	0.308	0.984	0.946	0.639
Autumn	0.130	0.986	0.954	0.850
Winter	0.097	0.984	0.924	0.431

Seasonal variation in growth-limiting controls of phytoplankton was assessed using the hourly model outputs. Table 3 represents how light, temperature and nutrient limitations co-vary seasonally in the model. Note that higher values in Table 3 indicate fewer limitations on phytoplankton growth. The hourly model shows light and temperature as key controls that limit phytoplankton growth throughout the year. Seasonal light factor varies from 0.1 to 0.31, and temperature factor varies from 0.43 to 0.85. Nitrogen does not limit phytoplankton growth with its limiting coefficient always being >0.98. Average seasonal phosphorus coefficient is also higher than 0.9 in all seasons, and remain >0.8 during the entire monitoring period (except for a few days in May, 2013). Chemical controls do not show limitations on phytoplankton growth, but physical controls show a significant influence.

4. Discussion

The hourly model presented here successfully simulates phytoplankton biomass along with other key determinands including flow, water temperature, nutrients and DO along the lower Thames. High-frequency observations within the catchment are utilized as model inputs as well as to support calibration of in-stream process-rate parameters that influence river water quality. Moreover, a model with low-frequency flow and water quality inputs is also able to characterize phytoplankton dynamics in the catchment, but still requires high-resolution light information. The model representation with dominant species as *Stephanodiscus hantzschii* best explains the phytoplankton variability in the catchment, which is also applicable for many lowland rivers worldwide. Lower River Thames shows significant association of chlorophyll concentrations with residence time, water temperature and light intensity. In the following sections, we discuss the model results from the lower Thames application and review model applicability as a phytoplankton prediction and management tool for lowland rivers.

4.1. Hourly Model Performance

4.1.1. Modeling Environmental Controls

The hourly model successfully simulates the physico-chemical and biological water quality variation across the 62 km length of the River Thames. The model performance is evaluated according to the guidelines provided by Moriasi et al. (2007). Flow simulations in the model make use of data of tributary inputs, abstraction volumes and sewage releases, giving very good fits (NSE > 0.9) for flows. Minor differences between the model and observations might be due to several reasons. Differences in the flow volume estimation (PBIAS = −9.88, −11.58) could be because QUESTOR does not include a hydrological component to account for rainfall-runoff processes. Flow simulations highly rely on the calibration of flow routing parameters and a correct representation of water inputs and outputs in the river. The flow routing parameters, adapted from previous studies in the River Thames (Waylett et al., 2013; P. Whitehead & Hornberger, 1984), are well calibrated for this river. Errors might be attributable to additional influences not currently represented in the model, but these cannot be identified with the available information. Groundwater discharge into the floodplain/river might be important, but the overestimation of low flows suggests it is less likely. Additionally, sewer overflows from STWs in urban areas close to the river are not specifically represented, which if included, could boost simulated high flows. Nevertheless, the overall model performance for flow is very good, and the model is able to capture the seasonal variability and flow magnitude.

For nutrients, the model performs better in simulating SRP (NSE > 0.7) compared to nitrate (NSE < 0.5). Despite the good NSE fits for SRP concentrations, there is a slight overestimation (Figure S2). Due to this, phosphorus mineralization process was switched-off for the entire stretch (Table S4) as there was already a sufficient pool of inorganic phosphorus simulated from upstream transport, tributary and sewage works inputs. For nitrate simulations, the model simulates a rapid increase in nitrification rate downstream, with the process rate becoming more than twice the upstream rate after merging of a heavily nutrient-enriched tributary, the River Wye (M. Bowes et al., 2012). The overestimation of the extreme low (<5 percentile) and high (>99 percentile) nitrate concentrations (Figure S2) might be because the nitrogen sources from incoming tributaries are not well characterized. Currently, the model uses low-resolution time-series of tributary nutrient inputs, and the nitrate fits may improve with higher resolution data inputs.

Diurnal variability in DO concentrations increases with biomass during growing season because of the increase in the autotrophic production and respiration. However, during this period, the magnitude of DO is overestimated in spite of the good simulation of phytoplankton blooms. This may happen if one or more of the other processes influencing DO variation such as reaeration, benthic oxygen demand, BOD decay, and so on are not accurately represented in the model. However, measurements of these processes are often scarce or absent in rivers and it is difficult to pin-point the processes influencing DO fits in the lower Thames due to lack of data availability. Nevertheless, the model does not show large errors in DO estimates and represent only a minor underestimation at Taplow (PBIAS = −4.47, −5.11) and Windsor (PBIAS = −0.93, −4.29) during both years. The overall seasonality of DO concentrations is also satisfactorily (NSE > 0.45) captured by the hourly model.

4.1.2. Modeling Phytoplankton Biomass

To reproduce phytoplankton dynamics, the model uses high-frequency sensor measurements, which also support the calibration of phytoplankton growth, death and respiration rate constants. Phytoplankton process rates in rivers around the world have been observed within ranges of 0.06–3 d^{-1} (growth) and 0.06–0.17 d^{-1} (death) (Bowie et al., 1985; Everbecq et al., 2001; Reichert et al., 2001), whereas for the River Thames, studies have reported 0.2–1.35 d^{-1} for growth rate and 0.05–0.23 d^{-1} for death rate (M. Hutchins et al., 2016; Waylett et al., 2013; P. G. Whitehead et al., 2015). The maximum growth rate calibrated for our model for the lower Thames is higher (2.28 day^{-1}), and the death and respiration fractions are 0.1 (Table S4). The death constant in the model accounts for grazing and nonpredatory mortality. Although phytoplankton mortality is a complex process and grazing rates in the river may vary spatially and temporally, the lack of grazing abundance data requires the model to be simple in its representation of mortality to avoid an overparameterized model. The growth, respiration and death constants mainly control the timing and magnitude of phytoplankton blooms, which the model is able to capture broadly.

We observe high photosynthetic production and respiration during April–July, when the environmental conditions accommodate high phytoplankton growth. The model simulates a clear diurnal cycle during the growing season, when autotrophic production is maximum. The model successfully captures the timing and magnitude of major peaks at all calibration sites. Previous phytoplankton modeling studies at daily to annual scale in the River Thames have reported NSE values of −5.350–0.228 (Waylett et al., 2013), 0.34–0.75 (P. G. Whitehead et al., 2015) and mean error values up to ±30% (M. Hutchins and Bowes 2018; M. Hutchins et al. 2016). A recent study (M. Hutchins et al., 2020) with daily time-step QUESTOR model in the Thames for 2013–2014 reported NSE statistics for chlorophyll between −0.17 and 0.22 at lower Thames reaches. Our study simulates hourly phytoplankton variation for the same period, and reports NSE > 0.73 at all calibration sites and mean errors ranging from −35% to 26%. Better performance of the hourly model, compared to the previous modeling studies in the Thames, confirms our hypothesis that high-resolution modeling can improve the predictions of timing and magnitude of phytoplankton blooms.

Model sensitivity testing with different algal groups derives SH model as the best representing model of phytoplankton growth in the lower Thames. Although the assumptions of green algae and small centric diatoms dominance also provide satisfactory fits, these models do not capture the peak blooms as well as the SH model does and overestimate low concentrations. Previously, mixed-phytoplankton populations have been reported to best represent phytoplankton biomass in the River Thames using the daily time-step QUESTOR (Waylett et al., 2013), where phytoplankton groups were allowed to thrive regardless of temperature. However, the hourly model, with better agreement with observations, suggests that the phytoplankton biomass in the river do not survive at higher temperatures. Thus, modeling studies at coarser resolution can sometimes result in misleading interpretations about the dominant algal communities and river processes.

The model performs best with an optimum temperature of 14°C in the lower Thames, which is also observed in other lowland rivers (Descy et al., 2003). *Stephanodiscus hantzschii* is found to be dominant in many lowland, temperate rivers (Everbecq et al., 2001) offering a wider applicability of the hourly model. Observational studies (M. Bowes et al., 2012; Lack, 1971) including the flow cytometry analysis (Read et al., 2014) in the River Thames showed dominance of diatoms in spring. However, during summer, a lower biomass is observed in the river and smaller pico-chlorophytes dominate the community (Read et al., 2014). Hence, our assumption about the dominant species works well for modeling phytoplankton seasonality. However, it is important to note that the dominance of phytoplankton groups in rivers may change over a year. In reality, different phytoplankton groups compete for resources and their dominance depends on multiple environmental factors that are not just limited to temperature. The current model structure does not incorporate these processes, but future research on model development should focus on including the interplay between different phytoplankton groups.

4.1.3. Model Uncertainties

Process-based water quality models include uncertainties introduced from several sources such as input data quality, conceptual simplifications causing structural uncertainty, and limitations in process understanding of the modeler because of lack of sufficient data (Abbaspour et al., 2015). This is especially a problem in large lowland catchments with inputs from a considerable number of sources. The hourly model in this study is based on the QUESTOR model, which has been widely applied in rivers across Europe. QUESTOR has been tested and subjected to comprehensive sensitivity analysis elsewhere (Deflandre et al., 2006; M. G. Hutchins & Hitt, 2019). Moreover, extensive QUESTOR applications in the River Thames provide confidence in calibration of the hourly model parameters and optimized values lie within similar ranges. In this study, we address the importance of inputs relative to that of in-stream processes for model predictions by testing the model with low-frequency input data. Through this exercise, we find that the model outputs are not sensitive to the time-scale of flow and water quality inputs, but are highly sensitive to the radiation inputs. The hourly model requires high-resolution radiation information to estimate phytoplankton growth. Obtaining high-resolution radiation data is feasible either directly or indirectly based on catchment location and sunshine hours, unlike high-resolution river water quality data that are often difficult to obtain. High-frequency inputs should reduce uncertainties in the model, but sparse data from tributaries may still introduce some uncertainties. However, assessment of model uncertainty requires a much more comprehensive analysis, and is out of the scope of this study. The hourly model application in this study is a step forward in high-resolution phytoplankton modeling, and demonstrates an efficient and skilful modeling tool for simulating hourly to annual scale variation in phytoplankton biomass and its controls.

4.2. Environmental Controls in the Lower Thames

BRT analysis provides an insight in to the nature and importance of associations of phytoplankton with multiple environmental variables, and the box-plots provide information about environmental bounds of these variables that promote harmful algal blooms. River Thames exhibits high nutrient availability throughout the year, and nutrients are consumed by phytoplankton during high growth (M. Bowes et al., 2016). Instead of nutrient concentrations influencing phytoplankton growth, we find that in fact, the stronger control is the reverse, that of biomass on nutrient availability. Nitrate is present in excess throughout both years due primarily to diffuse input of nitrate-rich groundwaters. The delivery of nutrients to the Thames from diffuse agricultural sources is primarily during winter and autumn high flows. In contrast, phosphorus addition from point sources is constant throughout the year, resulting in high SRP levels during low summer flows (Jarvie et al., 2002). Low flows and elevated SRP levels, in theory, should promote algal growth in the river (Hilton et al., 2006). However, high chlorophyll concentrations coinciding with low SRP levels and low flows (Figure S4) in this study suggest that (i) the residence time in the river is long enough for phosphorus uptake by phytoplankton biomass for autotrophic production, and (ii) as blooms develop, phosphorus levels start depleting in the river and become limiting.

After excluding the nutrients from the list of controls, the BRT analysis shows highest relative influence of flow (55%) followed by water temperature (38%) and PAR (7%). PAR is sufficiently available throughout the year accounting for the pattern in the relationship between chlorophyll and PAR (Figure 9c). This contrasts with findings from the process-based modeling exercise, where high-resolution light information is a crucial predictor to model accurate timing and magnitude of phytoplankton blooms. Light is a complex parameter to consider compared to other environmental variables as the influence of light on phytoplankton can be a function of past light information in terms of its timing, periodicity, and intensity over time (M. Bowes et al., 2016). Moreover, phytoplankton growth is also influenced by seasonal changes in riparian shading (M. Hutchins et al., 2010), which are hard to capture in the BRT model using only a simple measurement of PAR. Hourly process-based modeling, on the other hand, accounts for these details, albeit with specific assumptions about incoming light information and riparian shading patterns.

Flow and temperature showed an important control on phytoplankton growth. The interactions between flow versus temperature and PAR versus flow also showed significant influences. High PAR promotes phytoplankton growth only at low-mid flows. Moreover, large blooms are observed when temperature and

Table 4
Comparison of Environmental Bounds (For Chl-a>0.03 mg L^{1}) Along the River Stretch

	Observed		Modeled		
Variable	Caversham[a]	Windsor	Sonning bridge	Windsor	Runnymede
Flow ($m^3 s^{-1}$)	<30	32–68	28–51	30–63	21–54
Temp (°C)	9–19	10–17	11–18	11–18	11–18

[a] Thresholds at Caversham were reported by M. Bowes et al. (2016).
Note. Temp represents water temperature.

flow interacts within specific ranges. Phytoplankton growth at Caversham has been reported to respond to certain flow and temperature thresholds (M. Bowes et al., 2016). Downstream of Caversham, lower and upper flow bounds exist for phytoplankton growth (Table 4). Majority of high simulated chlorophyll (>0.03 mg L^{-1}) concentrations only occur when flows are between 21 and 63 $m^3 s^{-1}$ in contrast to M. Bowes et al. (2016) who only found high concentrations below 30 $m^3 s^{-1}$ from analyzing 5 years (2009–2013) of high-frequency measurements at Caversham. Flow bounds identified from the observations at Windsor (32–68 $m^3 s^{-1}$) are in a similar range to the modeled bounds (Table 4). Phytoplankton blooms only develop at low flows (<51–63 $m^3 s^{-1}$ here), when the residence time is long enough for phytoplankton growth. However, the presence of a lower flow bound (>21–30 $m^3 s^{-1}$ here), below which high concentrations do not occur, also suggests that the phytoplankton biomass cannot remain suspended when the flow becomes too low (<20 $m^3 s^{-1}$), and settle in the river (Balbi, 2000).

Due to SH model assumption of an optimum temperature for growth of 14°C, a consistent temperature range, within which high concentrations are simulated (11°C–18°C [Table 4]), is apparent throughout the stretch. Observed data show a temperature range of 10°C–17°C at Windsor, whereas M. Bowes et al. (2016) reported a temperature range of 9°C–19°C at Caversham. The slightly different bounds for temperature are likely indicative of a more-complex system than that represented in the model. These differences are likely to reflect largely dominant cool-water centric diatoms, but with secondary influences from other groups, such as groups thriving in warmer conditions and attached algae mobilized under turbulent conditions at lower temperatures. Important environmental controls found in this study (flow, light, and temperature) are consistent with findings from other lowland rivers around the world such as the Murray (Bormans & Webster, 1999), the Meuse (Everbecq et al., 2001), and the Severn (Reynolds & Glaister, 1993). Some lowland studies have also highlighted the importance of chemical (nutrients) and biological (grazing) controls, but these are shown to become significant only when physical constraints are reduced (Billen et al., 1994; Gosselain et al., 1998; Reynolds & Descy, 1996).

4.3. Flow Regulation and Water Quality

Physical factors, such as river residence time and flow variability, play an important role in controlling phytoplankton growth in lowland rivers (Reynolds, 2000). High flows prevent growth by rapid flushing of phytoplankton biomass. The lower Thames reaches are deeper than the upstream reaches with slow-moving water enriched by high floods in the winter and consistent low flows during the rest of the year. Moreover, locks and weirs for navigation throughout the river stretch slow the flow, resulting in increased residence times (M. Hutchins & Bowes, 2018). Median residence time in the river stretch in our study during 2013–2014 was 40 h, which varied from 9 to 112 h at very high (90th percentile) and very low (10th percentile) flows, respectively. One possible solution to avoid high phytoplankton growth in rivers could be via the maintenance of river flow above critical thresholds (Wang et al., 2019), in line with environmental flow concepts (Poff & Zimmerman, 2010). Experimental flow releases have proved to facilitate maintenance of chemical and biological water quality in regulated rivers elsewhere (Gillespie et al., 2020; Lind et al., 2007). As discussed in the previous section, high phytoplankton concentrations in the lower Thames are only encountered at low flows below 60 $m^3 s^{-1}$. Although more evidence is required to use this threshold as a minimum environmental flow in the river, short pulses of high flow release could act as a measure to prevent large algal bloom developments in regulated, lowland rivers like Thames.

5. Conclusion

We develop a new hourly river model for a 62 km stretch in the lower River Thames. By including an hourly mass balance, the model successfully simulates flow, water temperature, DO, nutrients, and phytoplankton biomass in the river. The model satisfactorily captures diurnal variation of phytoplankton dynamics as well as the magnitude and timing of bloom events. The hourly model in this study represents better goodness of fits compared to the previous daily time-step modeling studies in the Thames, and thus, confirms our hypothesis that high temporal-resolution modeling improves phytoplankton growth prediction. The model can predict phytoplankton dynamics from low-resolution water flow and quality with hourly resolution data only needed for solar radiation. This offers a possibility of model application in catchments where high-frequency measurements are not available.

From a range of algal groups tested under sensitivity analysis, a model assuming *Stephanodiscus hantzschii* with optimum growth at 14°C to predominate best represented biomass variation. *Stephanodiscus hantzschii* is also predominant in many lowland rivers worldwide. The model can easily be applied elsewhere and also be adapted in its parameterization to reflect dominance by different species if needed. Phytoplankton growth in lower Thames is mainly influenced by hydrological (residence time) and physical controls (water temperature and light intensity), which is typically found in lowland rivers worldwide. We observe that phytoplankton blooms only develop within specific flow bounds (21–63 $m^3 s^{-1}$). Identification of flow bounds is useful to prevent major bloom developments and to maintain river water quality. Hence, short-term high-flow release (here, above 60 $m^3 s^{-1}$), as experimented in other regulated rivers, could form a potential management strategy in critical situations.

This is the first study where high-resolution hourly model is validated against similarly high-frequency biomass observations. To our knowledge, river modeling studies with all environmental controls at such a temporal and spatial extent have not previously been undertaken. It offers the following powerful possibilities:

(1) Feasibility for hourly model application in any river with a single continuous water quality monitoring site in the lower reaches
(2) Reconstruction of past long-term changes in hourly water quality dynamics before continuous monitoring with sensors was widely available
(3) Application to provide early warnings of phytoplankton blooms as well as to evaluate management strategies using scenario analysis
(4) Hourly scale DO curves and biomass information can be further interpreted to evaluate ecosystem metabolism and to identify low night-time oxygen levels that may threaten ecological health
(5) The costs for high-frequency monitoring over multiple sites within the river network can be reduced if a reliable modeling tool such as the one described in this study is available

Data Availability Statement

The hourly data for water temperature, chlorophyll and dissolved oxygen in the lower Thames were made available from the Environment Agency and can be downloaded from Zenodo data repository (https://doi.org/10.5281/zenodo.4288254). Daily flow data are available at the NRFA (NERC, National River Flow Archive, http://www.ceh.ac.uk/data/nrfa/). Weekly water quality data can be found at (1) the UK Centre for Ecology & Hydrology's Thames Initiative research platform (https://doi.org/10.5285/e4c300b1-8bc3-4df2-b23a-e72e67eef2fd) hosted by the UK NERC Environmental Information Data Centre and (2) Environment Agency's water quality data archive (http://environment.data.gov.uk/water-quality/view/landing). Radiation information is available at British Atmospheric Data Centre (MIDAS Landsat data) (http://archive.ceda.ac.uk/).

Acknowledgments
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 765553. The weekly water quality monitoring data set used in this study was gathered under the UKCEH Thames Initiative, which was funded by the Natural Environment Research Council (NEC04877). The authors would like to thank the National River Flow Archive for providing the river flow data, and Environment Agency for the high-frequency water quality data.

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., & Kløve, B. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of hydrology*, *524*, 733–752.

Balbi, D. M. (2000). Suspended chlorophyll in the River Nene, a small nutrient-rich river in eastern England: long-term and spatial trends. *The Science of the Total Environment*, *251*, 401–421.

Basu, B., & Pick, F. (1997). Phytoplankton and zooplankton development in a lowland, temperate river. *Journal of Plankton Research*, *19*(2), 237–253.

Billen, G., Garnier, J., & Hanset, P. (1994). Modelling phytoplankton development in whole drainage networks: The RIVERSTRAHLER model applied to the Seine river system. In J. P. Descy, C. S. Reynolds, & J. Padisák (Eds.), *Phytoplankton in turbid environments: Rivers and shallow lakes. Developments in Hydrobiology* (Vol. 100). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-2670-2_11

Boorman, D. B. (2003). LOIS in-stream water quality modeling. Part 1. Catchments and methods. *The Science of the Total Environment*, *314*, 379–395.

Boorman, D. B. (2003). LOIS in-stream water quality modeling. Part 2. Results and scenarios. *The Science of the Total Environment*, *314*, 397–409.

Bormans, M., & Webster, I. T. (1999). Modeling the spatial and temporal variability of diatoms in the River Murray. *Journal of Plankton Research*, *21*(3), 581–598.

Bowes, M. J., Armstrong, L. K., Harman, S. A., Wickham, H. D., Nicholls, D. J., Scarlett, P. M., et al. (2018). Weekly water quality monitoring data for the River Thames (UK) and its major tributaries (2009–2013): the Thames Initiative research platform. *Earth System Science Data*, *10*(3), 1637–1653.

Bowes, M., Gozzard, E., Johnson, A., Scarlett, P., Roberts, C., Read, D., et al. (2012). Spatial and temporal changes in chlorophyll-a concentrations in the River Thames basin, UK: Are phosphorus concentrations beginning to limit phytoplankton biomass? *The Science of the Total Environment*, *426*, 45–55.

Bowes, M., Loewenthal, M., Read, D., Hutchins, M., Prudhomme, C., Armstrong, L., et al. (2016). Identifying multiple stressor controls on phytoplankton dynamics in the River Thames (UK) using high-frequency water quality data. *The Science of the Total Environment*, *569*, 1489–1499.

Bowie, G. L., Mills, W. B., Porcella, D. B., Campbell, C. L., Pagenkopf, J. R., Rupp, G. L., et al. (1985). Rates, constants, and kinetics formulations in surface water quality modeling. *Environmental Protection Agency*, *600*, 3–85.

Brown, L. C., & Barnwell, T. O. (1987). *The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual EPA/600/3-87/007 (NTIS PB87202156)*. Washington, DC: US Environmental Protection Agency, Office of Research and Development.

Canale, R. P., & Vogel, A. H. (1974). Effects of temperature on phytoplankton growth. *Journal of the Environmental Engineering Division*, *100*(1), 231–241.

Crossman, J., Whitehead, P., Futter, M., Jin, L., Shahgedanova, M., Castellazzi, M., & Wade, A. (2013). The interactive responses of water quality and hydrology to changes in multiple stressors, and implications for the long-term effective management of phosphorus. *The Science of the Total Environment*, *454*, 230–244.

Deflandre, A., Williams, R., Elorza, F., Mira, J., & Boorman, D. (2006). Analysis of the QUESTOR water quality model using a Fourier amplitude sensitivity test (FAST) for two UK rivers. *The Science of the Total Environment*, *360*(1–3), 290–304.

Descy, J. P., Everbecq, E., Gosselain, V., Viroux, L., & Smitz, J. (2003). Modeling the impact of benthic filter-feeders on the composition and biomass of river plankton. *Freshwater Biology*, *48*(3), 404–417.

Dodds, W. K. (2006). Eutrophication and trophic state in rivers and streams. *Limnology & Oceanography*, *51*(1), 671–680.

Domingues, R. B., Anselmo, T. P., Barbosa, A. B., Sommer, U., & Galvão, H. M. (2011). Light as a driver of phytoplankton growth and production in the freshwater tidal zone of a turbid estuary. *Estuarine, Coastal and Shelf Science*, *91*(4), 526–535.

Eatherall, A., Boorman, D., Williams, R., & Kowe, R. (1998). Modeling in-stream water quality in LOIS. *The Science of the Total Environment*, *210*, 499–517.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813.

Everbecq, E., Gosselain, V., Viroux, L., & Descy, J.-P. (2001). Potamon: A dynamic model for predicting phytoplankton composition and biomass in lowland rivers. *Water Research*, *35*(4), 901–912.

Feld, C. K., Segurado, P., & Gutiérrez-Cánovas, C. (2016). Analysing the impact of multiple stressors in aquatic biomonitoring data: A 'cookbook'with applications in R. *The Science of the Total Environment*, *573*, 1320–1339.

Gillespie, B. R., Kay, P., & Brown, L. E. (2020). Limited impacts of experimental flow releases on water quality and macroinvertebrate community composition in an upland regulated river. *Ecohydrology*, *13*(2), e2174.

Gosselain, V., Viroux, L., & Descy, J. P. (1998). Can a community of small-bodied grazers control phytoplankton in rivers?. *Freshwater Biology*, *39*(1), 9–24.

Hardenbicker, P., Rolinski, S., Weitere, M., & Fischer, H. (2014). Contrasting long-term trends and shifts in phytoplankton dynamics in two large rivers. *International Review of Hydrobiology*, *99*(4), 287–299.

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. *Circles*, *9*(1), 1–68.

Hilton, J., O'Hare, M., Bowes, M. J., & Jones, J. I. (2006). How green is my river? A new paradigm of eutrophication in rivers. *The Science of the Total Environment*, *365*(1–3), 66–83.

Hutchins, M., Abesser, C., Prudhomme, C., Elliott, J., Bloomfield, J., Mansour, M., & Hitt, O. (2018). Combined impacts of future land-use and climate stressors on water resources and quality in groundwater and surface waterbodies of the upper Thames river basin, UK. *The Science of the Total Environment*, *631*, 962–986.

Hutchins, M., & Bowes, M. (2018). Balancing water demand needs with protection of river water quality by minimising stream residence time: an example from the Thames, UK. *Water Resources Management*, *32*(7), 2561–2568.

Hutchins, M., Harding, G., Jarvie, H., Marsh, T., Bowes, M., & Loewenthal, M. (2020). Intense summer floods may induce prolonged increases in benthic respiration rates of more than one year leading to low river dissolved oxygen. *Journal of Hydrology X*, *8*, 100056.

Hutchins, M. G., & Hitt, O. E. (2019). Sensitivity of river eutrophication to multiple stressors illustrated using graphical summaries of physics-based river water quality model simulations. *Journal of hydrology*, *577*, 123917.

Hutchins, M., Johnson, A., Deflandre-Vlandas, A., Comber, S., Posen, P., & Boorman, D. (2010). Which offers more scope to suppress river phytoplankton blooms: reducing nutrient pollution or riparian shading? *The Science of the Total Environment*, *408*(21), 5065–5077.

Hutchins, M., Williams, R., Prudhomme, C., Bowes, M., Brown, H., Waylett, A., & Loewenthal, M. (2016). Projections of future deterioration in UK river quality are hampered by climatic uncertainty under extreme conditions. *Hydrological Sciences Journal*, *61*(16), 2818–2833.

Jarvie, H. P., Neal, C., Williams, R. J., Neal, M., Wickham, H. D., Hill, L. K., et al. (2002). Phosphorus sources, speciation and dynamics in the lowland eutrophic River Kennet, UK. *The Science of the Total Environment*, *282*, 175–203.

Köhler, J. (1995). Growth, production and losses of phytoplankton in the lowland River Spree: carbon balance. *Freshwater Biology*, *34*(3), 501–512.

Lack, T. (1971). Quantitative studies on the phytoplankton of the Rivers Thames and Kennet at Reading. *Freshwater Biology*, *1*(2), 213–224.

Larroudé, S., Massei, N., Reyes-Marchant, P., Delattre, C., & Humbert, J. (2013). Dramatic changes in a phytoplankton community in response to local and global pressures: A 24-year survey of the river L oire (France). *Global Change Biology*, *19*(5), 1620–1631.

Lázár, A. N., Wade, A. J., & Moss, B. (2016). Modeling primary producer interaction and composition: an example of a UK Lowland River. *Environmental Modeling & Assessment*, *21*(1), 125–148.

Lind, P., Robson, B., & Mitchell, B. (2007). Multiple lines of evidence for the beneficial effects of environmental flows in two lowland rivers in Victoria, Australia. *River Research and Applications*, *23*(9), 933–946.

Martin, N., McEachern, P., Yu, T., & Zhu, D. Z. (2013). Model development for prediction and mitigation of dissolved oxygen sags in the Athabasca River, Canada. *The Science of the Total Environment*, *443*, 403–412.

Minaudo, C., Curie, F., Jullian, Y., Gassama, N., & Moatar, F. (2018). QUAL-NET, a high temporal-resolution eutrophication model for large hydrographic networks. *Biogeosciences*, *15*(7), 2251–2269.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Paerl, H. W., & Huisman, J. (2009). Climate change: A catalyst for global expansion of harmful cyanobacterial blooms. *Environmental Microbiology Reports*, *1*(1), 27–37.

Poff, N. L., & Zimmerman, J. K. (2010). Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. *Freshwater Biology*, *55*(1), 194–205.

Rankinen, K., Bernal, J. E. C., Holmberg, M., Vuorio, K., & Granlund, K. (2019). Identifying multiple stressors that influence eutrophication in a Finnish agricultural river. *The Science of the Total Environment*, *658*, 1278–1292.

Read, D. S., Bowes, M. J., Newbold, L. K., & Whiteley, A. S. (2014). Weekly flow cytometric analysis of riverine phytoplankton to determine seasonal bloom dynamics. *Environmental Science: Processes & Impacts*, *16*(3), 594–603.

Reichert, P., Borchardt, D., Henze, M., Rauch, W., Shanahan, P., Somlyódy, L., & Vanrolleghem, P. (2001). River water quality model no. 1 (RWQM1): II. Biochemical process equations. *Water Science and Technology*, *43*(5), 11–30.

Reynolds, C. (2000). Hydroecology of river plankton: The role of variability in channel flow. *Hydrological Processes*, *14*(16–17), 3119–3132.

Reynolds, C., & Descy, J.-P. (1996). The production, biomass and structure of phytoplankton in large rivers. *Large Rivers*, 161–187.

Reynolds, C., & Glaister, M. (1993). Spatial and temporal changes in phytoplankton abundance in the upper and middle reaches of the River Severn. *River Systems*, *9*, 1–22.

Ridgeway, G. (2020). Generalized boosted models: A guide to the gbm package. *R Package Vignette*, *1*(1), 2007.

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., et al. (2016). Sensors in the stream: The high-frequency wave of the present. *Environmental Science & Technology*, *50*(19), 10297–10307. https://doi.org/10.1021/acs.est.6b02155

Ruse, L., & Love, A. (1997). Predicting phytoplankton composition in the River Thames, England. *Regulated Rivers: Research & Management*, *13*(2), 171–183.

Steele, J. H. (1962). Environmental control of photosynthesis in the sea. *Limnology & Oceanography*, *7*(2), 137–150.

Suarez, V. V. C., Brederveld, R. J., Fennema, M., Moreno-Rodenas, A., Langeveld, J., Korving, H., et al. (2019). Evaluation of a coupled hydrodynamic-closed ecological cycle approach for modeling dissolved oxygen in surface waters. *Environmental Modelling & Software*, *119*, 242–257.

Van Griensven, A., & Bauwens, W. (2005). Application and evaluation of ESWAT on the Dender basin and the Wister Lake basin. *Hydrological Processes: An International Journal*, *19*(3), 827–838.

Villegas, I., & de Giner, G. (1973). Phytoplankton as a biological indicator of water quality. *Water Research*, *7*(3), 479–487.

Wang, J., Zhang, Z., & Johnson, B. (2019). Low flows and downstream decline in phytoplankton contribute to impaired water quality in the lower Minnesota River. *Water Research*, *161*, 262–273.

Waylett, A., Hutchins, M., Johnson, A., Bowes, M., & Loewenthal, M. (2013). Physico-chemical factors alone cannot simulate phytoplankton behavior in a lowland river. *Journal of hydrology*, *497*, 223–233.

Whitehead, P. G., Bussi, G., Bowes, M. J., Read, D. S., Hutchins, M. G., Elliott, J. A., & Dadson, S. J. (2015). Dynamic modeling of multiple phytoplankton groups in rivers with an application to the Thames river system in the UK. *Environmental Modelling & Software*, *74*, 75–91.

Whitehead, P., & Hornberger, G. (1984). Modeling algal behavior in the River Thames. *Water Research*, *18*(8), 945–953.

Whitehead, P., Howard, A., & Arulmani, C. (1997). Modeling algal growth and transport in rivers: a comparison of time series analysis, dynamic mass balance and neural network techniques. *Hydrobiologia*, *349*(1–3), 39–46.

References From the Supporting Information

Boorman, D. B. (2003). LOIS in-stream water quality modelling. Part 1. Catchments and methods. *Science of the Total Environment*, *314*, 379–395.

DOE. (1973). *Aeration at weirs*. London, UK: Department of the Environment.

Waylett, A., Hutchins, M., Johnson, A., Bowes, M., & Loewenthal, M. (2013). Physico-chemical factors alone cannot simulate phytoplankton behaviour in a lowland river. *Journal of hydrology*, *497*, 223–233.

Whitehead, P., & Hornberger, G. (1984). Modelling algal behaviour in the River Thames. *Water Research*, *18*(8), 945–953.