LIMNOLOGY and OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 67, 2022, 1313–1327

doi: 10.1002/lno.12079

High-resolution water-quality and ecosystem-metabolism modeling in lowland rivers

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

Abstract

High-resolution monitoring of water quality and ecosystem functioning over large spatial scales in expansive lowland river catchments is challenging. Therefore, we need modeling tools to predict these processes at locations where observations are absent. Here, we present a new approach to estimate ecosystem metabolism underpinned by a high-resolution, process-based model of in-stream flows and water quality. The model overcomes the current challenges in metabolism modeling by accounting for oxygen transport under varying flows and oxygen transformations due to biogeochemical processes. We implement the model in a 62-km-long stretch of the River Thames, England, using observations spanning 2 yr. Model outputs suggest that the river is primarily autotrophic from mid-spring to mid-summer due to high biomass during low-flow periods, and is heterotrophic during the rest of the year. Ecosystem respiration in upstream reaches is driven mainly by biochemical oxygen demand, autotrophic respiration, and nitrification processes, whereas downstream sites also show a control of benthic oxygen demand in addition to the aforementioned processes. Using empirical modeling, we analyze the sensitivity of our estimated metabolism rates to multiple environmental stressors. Results demonstrate that empirical models could be useful for rapid river health assessments, but need improvements to reproduce peak metabolism rates. The process-based model, although more complex than existing in situ approaches to metabolism quantification, allows inference when gaps in continuous observations are present. The model offers additional benefits for predicting metabolism rates under future scenarios of environmental change incorporating multiple stressor effects.

Assessments of river ecosystem health have traditionally relied on structural indicators such as channel morphology, water quality, or the composition of biological communities (Von Schiller et al. 2017). However, with the advances in high-resolution monitoring techniques (Rode et al. 2016), sensor networks and linked modeling tools are gaining traction for prediction of functional indicators such as ecosystem metabolism. Conventional methods of metabolism modeling, based on Odum's (1956) open-channel approach, estimate metabolism rates at a river-reach level using continuous dissolved oxygen (DO) measurements at a single site (e.g., single station method, Izagirre et al. 2007) or two sites over a reach (e.g., two station method, Hall Jr and Tank 2005; Halbedel and Büttner 2014). These models do not account for the influence of upstream changes on the downstream DO advection and transformations within the river network. Moreover, these models do not specifically account for changes in river hydrology and biogeochemistry (exceptions include Payn et al. (2017); Segatto et al. 2020), which could have a critical impact on DO dynamics in the river. Therefore, such models may provide biased interpretations of metabolism estimates (Payn et al. 2017) if the metabolic regime is sensitive to changes in these environmental stressors at the time-step of calculation.

*Correspondence: devpat@ceh.ac.uk

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: DP conceived the presented work with scientific guidance from MH. DP and MH developed the model. LEB and FE contributed to statistical analysis. DP analysed the data. ML, MB and GO provided management and scientific direction of the contributory monitoring programmes. PS contributed to field data collection. LA and DN contributed to laboratory analysis of the water quality data. DP wrote the manuscript with contributions and review from MH and LEB.

Ecosystem metabolism characterizes carbon fixation and mineralization through gross primary production (GPP) and ecosystem respiration (ER). GPP and ER are sensitive to multiple stressors, which act independently or in combination with other stressors (Heathwaite 2010; Von Schiller et al. 2017) often presenting a complex interplay of controls. ER is regulated by water temperature (Demars et al. 2011; Perkins et al. 2012) and organic matter supply (Young et al. 2008). Often, light (Mulholland et al. 2001) and in some cases, nutrient availability (Guasch et al. 1995) control GPP. Flow is also an important regulator of GPP. Flooding disrupts GPP seasonality through scouring of benthic producers and organic matter (Uehlinger 2006) as well as by reducing light availability in sediment-mobilized turbid waters (Aspray et al. 2017). Slow-flowing rivers with clear waters are typically autotrophic due to high light availability and stable flow regimes (Acuña et al. 2011) whilst faster-flowing rivers are typically heterotrophic. Measuring these spatial and temporal dynamics can be difficult, since the spatial resolution of sensor networks is largely limited by logistics of multi-site set up, maintenance, and data collection/validation.

GPP and ER estimates from conventional DO mass-balance models are usually empirically related to different environmental stressors to evaluate their sensitivity to these stressors (Izagirre et al. 2008; Beaulieu et al. 2013; Aspray et al. 2017). However, it is also important to associate these stressors mechanistically to GPP and ER to better understand the underlying controls of metabolic regimes in rivers as well as to predict changes in river metabolism outside the range of the available observations. Whereas several mechanistic water quality models (e.g., QUAL2E (Brown and Barnwell 1987), RWQM (Reichert et al. 2001)) include biochemical processes that affect DO transformations in the water column, they are generally only tested at daily to weekly time-steps. The coarse testing limits their use for metabolism estimation, which is susceptible to sub-daily changes in hydrology and biochemical water quality (Roberts et al. 2007; Izagirre et al. 2008). Therefore, here we combine DO and metabolism modeling using an existing, hourly scale, mechanistic water quality model, the hourly Quality Evaluation and Simulation Tool for River-systems (QUESTOR) model (Pathak et al. 2021). We implement this approach in a lowland river, the River Thames, in southern England.

QUESTOR is a process-based, in-stream water quality model, which simulates hourly scale variation and transport of river flows, water temperature, DO, nutrients, and phytoplankton biomass in a river network. The model has been previously tested to predict diel variation in physico-chemical water quality and phytoplankton biomass in the lower River Thames (Pathak et al. 2021). We advance this study to estimate metabolism rates from the DO mass-balance module. The hourly QUESTOR model simulates diel changes in the environmental stressors (e.g., light, temperature, flow, nutrients) and their resulting impact on ecosystem productivity and respiration. The model, however, has a relatively more complex structure with many model parameters, thus carrying a risk of attached uncertainties and parameter equifinality during the calibration. Nonetheless, we reduce parameter uncertainties during the calibration process by making use of the abundant literature on water quality modeling that exists for the River Thames catchment (Whitehead and Hornberger 1984; Waylett et al. 2013; Whitehead et al. 2015; Hutchins et al. 2018).

Whilst process-based models realistically represent the sensitivities of the system to key drivers (Hrachowitz et al. 2014) and capture the short term dynamics (Jankowski et al. 2021), empirical models have long-standing pedigree in providing insight into ecosystems' response to multiple stressors (Izagirre et al. 2008; Beaulieu et al. 2013). Specifically, with the development of machine learning techniques, empirical models can utilize data to learn and increasingly improve model performance (Elith et al. 2008; Feld et al. 2016). Therefore, in addition to the process-based model, we also use an empirical approach to assess the sensitivity of modeled metabolism rates to multiple stressors. A comparison of both models enables us to test if the empirical approach can provide accurate predictions and substitute process-based modeling for rapid assessments of river ecosystem health.

Our main aims here are as follows.

1. To develop a process-based approach for coupled modeling of in-stream hydrology, biochemical water quality, and ecosystem metabolism in lowland rivers.
2. To analyze spatio-temporal variation in the metabolic regime within the modeled river network (Thames, England).
3. To perform a sensitivity analysis of GPP and ER to physico-chemical determinands using random forest machine learning technique and generalized least squares (GLS) regression modeling.

Methods

Modeling approach

We use the process-based, hourly QUESTOR model tested in the lower Thames by Pathak et al. (2021). The model is a pseudo 1-D model (strictly speaking 0-D) and assumes perfect mixing within a reach (river network description in Table S1). The model assumes a fixed channel width with rectangular cross section and represents river reaches as a set of non-linear reservoirs or well-mixed tanks in series (Fig. 1). Solute dynamics and transport in the model are described with a mass-balance approach using ordinary differential equations, which are numerically solved using an explicit fourth-order Runge–Kutta–Merson differential equation solver (DASCRU; Fox 1962). The key variables represented by equations are flow, water temperature, photosynthetically active radiation (PAR), chlorophyll *a* (Chl *a*), BOD (biochemical oxygen demand), DO, nitrate (NO^-_3), ammonium (NH^+_4), particulate organic nitrogen (PON), organic phosphorus (P_{org}), and inorganic phosphorus (SRP). A detailed set of equations for all the key variables is provided

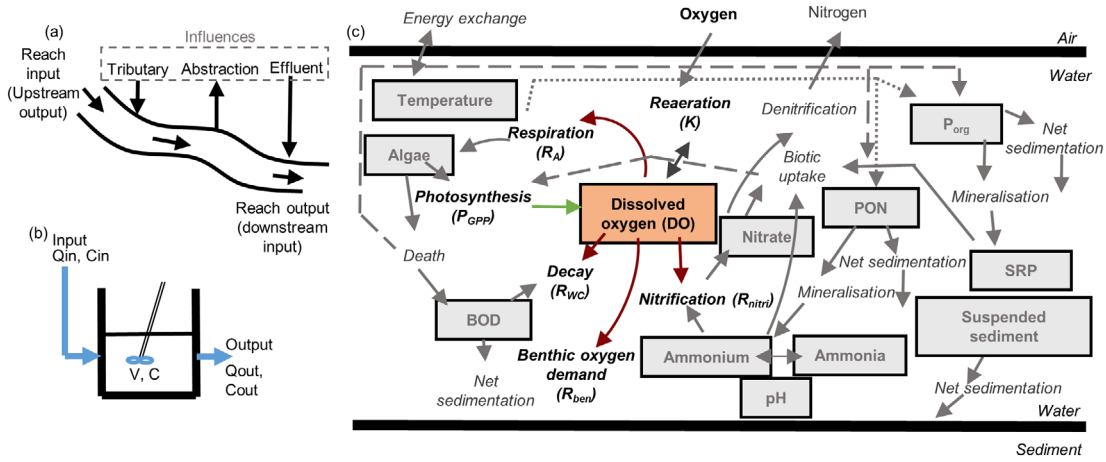

Fig. 1. Model structure and processes (modified after Pathak et al. 2021). (**a**) The schematic of a typical reach, (**b**) the conceptualization of reaches (C, solute concentration; Q, flow; V, volume), and (**c**) describes the water quality determinands and processes in the model. BOD, biochemical oxygen demand; PON, particulate organic nitrogen; P_{org}, organic phosphorus; SRP, inorganic phosphorus; SS, suspended sediment.

in Text S1. Here, we summarize the flow and DO modules of the hourly model, and describe the equations for ecosystem metabolism estimation.

Flow routing module

A simple mass-balance of incoming and outgoing flows is used. The incoming flow in the reach is the balance of the upstream flows plus point-source discharges minus abstractions. Flow inputs from the tributaries at the main channel confluence are scaled upwards based on the location (often some distance upstream) of the gauging station and the contributing catchment area (Hutchins et al. 2020). Therefore, this indirectly includes groundwater contribution at the tributary confluences with the main Thames. Groundwater contribution to and from the main Thames is assumed to be in balance. Outflow of water from a reach is calculated as

$$\frac{dQ_{\text{out}}}{dt} = \frac{Q_{\text{in}} - Q_{\text{out}}}{\tau(1-c)} \tag{1}$$

where Q_{in} is the total flow into the reach ($m^3\ s^{-1}$), Q_{out} is the flow out of the reach ($m^3\ s^{-1}$), t is the time-step (h), τ is the residence time (h) derived by $l/bQ_{\text{out}}{}^c$, l is the length of the reach (m), b and c are reach-specific constants. Constants b and c are calibrated from flow-velocity relationships ($v = bQ_{\text{out}}{}^c$), which characterize the hydromorphology and lock operations in the river (Whitehead and Hornberger 1984; Waylett et al. 2013). The flow routing model facilitates modeling of river residence time, which allows us to account for the influence of hydrological variation on DO dynamics as discussed in the next section.

Dissolved oxygen module

The processes controlling DO concentrations within a reach include (1) DO advection; (2) production of oxygen from GPP; (3) loss of oxygen from ER; and (4) oxygen change from reaeration

$$\frac{dC_{\text{DO},o}}{dt} = \frac{1}{\tau}\left(C_{\text{DO},i} - C_{\text{DO},o} + W\right) + P_{\text{GPP}} - R_{\text{ER}} + F \tag{2}$$

where $C_{\text{DO},i}$ is the input DO concentration (mg L^{-1}), $C_{\text{DO},o}$ is the output DO concentration (mg L^{-1}), W is the aeration at weirs (mg O_2 L^{-1}), P_{GPP} is the gross primary production (mg O_2 L^{-1} h^{-1}), R_{ER} is the ecosystem respiration (mg O_2 L^{-1} h^{-1}), and F is the aeration at the air-water surface (mg O_2 L^{-1} h^{-1}).

Ecosystem metabolism

Oxygen production. Rate of oxygen production (P_{GPP}, mg O_2 L^{-1} h^{-1}) in the river (Eq. 3) is given by

$$P_{\text{GPP}} = P_P + P_N \tag{3}$$

where P_P is the photosynthetic production (mg O_2 L^{-1} h^{-1}) and P_N is the oxygen produced during nitrate assimilation by phytoplankton (mg O_2 L^{-1} h^{-1}). Although the hourly model supports modeling of macrophytes and benthic algae, this study only includes modeling of phytoplankton biomass, since it is the dominant driver of metabolism in the study stretch (Whitehead and Hornberger 1984; Lázár et al. 2012).

The hourly model not only includes the influence of phytoplankton biomass on DO variations, but also accounts for the influence of light, temperature, and nutrient availability on photosynthesis. These details are explained in the description of the phytoplankton model in Text S1. Here, we only summarize the equations directly relevant to the DO model. P_P is a function of photosynthetic rate and phytoplankton concentration in the water column

$$P_P = k^{\text{pho}}\left(\frac{32}{12}\Delta\right) \tag{4}$$

where k^{pho} is the gross photosynthetic rate of autotrophs representing increase in Chl *a* during photosynthesis (mg L^{-1} h^{-1}) and Δ is the ratio of carbon to Chl *a* in autotrophs (50) (Bowie et al. 1985). The ratio of 32/12 represents the mass of oxygen produced in photosynthesis or consumed in respiration per unit mass of carbon fixed.

P_N is estimated as

$$P_N = k^{\text{pho}}(1 - n_{\text{pref}})\alpha \tag{5}$$

where n_{pref} is the autotroph preference for ammonia (Eq. 6) and α is the ratio of nitrogen to Chl *a* in autotrophs (10) (Bowie et al. 1985)

$$n_{\text{pref}} = \frac{k_{\text{pref}} C_{\text{NH4,o}}}{k_{\text{pref}} C_{\text{NH4,o}} + (1 - k_{\text{pref}}) C_{\text{NO3,o}}} \tag{6}$$

Here, $C_{\text{NH4,o}}$ is the NH^{+}_{4} concentration (mg L^{-1}), $C_{\text{NO3,o}}$ is the NO^{-}_{3} concentration (mg L^{-1}), and k_{pref} is the preference factor for ammonia over nitrate (Table S2).

Oxygen depletion. Rate of oxygen depletion (R_{ER}, mg O_2 L^{-1} h^{-1}) in the river includes four pathways

$$R_{\text{ER}} = R_A + R_{\text{nitri}} + R_{\text{ben}} + R_{\text{wc}} \tag{7}$$

where R_A is the autotrophic respiration (mg O_2 L^{-1} h^{-1}), R_{nitri} is the assimilation of oxygen in the process of nitrification of ammonium to nitrate (mg O_2 L^{-1} h^{-1}), R_{ben} is the sediment oxygen demand (benthic respiration) (mg O_2 L^{-1} h^{-1}), and R_{wc} is the BOD in the water column (mg O_2 L^{-1} h^{-1}).

R_A is a function of phytoplankton respiration rate (k^{res}, mg L^{-1} h^{-1}) and temperature (Eq. 8). k^{res} is modeled in the phytoplankton model as described in Text S1.

$$R_A = k^{\text{res}}\left(\frac{32}{12}\Delta\right) \tag{8}$$

where R_{nitri} involves conversion of ammonium to nitrate and requires oxygen for this conversion

$$R_{\text{nitri}} = 4.57 k_{\text{nit}} C_{\text{NH4,o}}\left(\frac{C_{\text{DO},o}}{C_{\text{DO},o} + S_{\text{nitri}}}\right) \tag{9}$$

where k_{nit} is the nitrification rate (h^{-1}) and S_{nitri} is the DO half-saturation concentration for nitrification (mg L^{-1}). The coefficient, 4.57 is derived from the stoichiometry of the reactions and represents the oxygen required to convert ammonia to nitrate.

R_{ben} (Eq. 10) represents the transfer of oxygen between the overlying water and the sediments (Cox 2003). The benthic respiration rate (k_{ben}, h^{-1}) (Eq. 11) is simulated as a function of stream depth and temperature, where the stream depth and water column oxygen concentration represents the availability of oxygen to the bed.

$$R_{\text{ben}} = k_{\text{ben}} C_{\text{DO},o} \tag{10}$$

$$k_{\text{ben}} = \frac{k_{\text{ben20}}}{z}\theta^{(T - T_{\text{ref}})} \tag{11}$$

where k_{ben20} is a unitless coefficient provided by the user during the calibration, z is the mean water depth of the reach (m), T is the water temperature (°C), T_{ref} is the reference temperature (20°C), and θ is a temperature correction factor (1.08).

R_{wc} represents the carbonaceous deoxygenation where oxygen in the water column is consumed by heterotrophic bacteria

$$R_{\text{wc}} = k_{\text{bod}} C_{\text{BOD},o}\left(\frac{C_{\text{DO},o}}{C_{\text{DO},o} + S_{\text{bod}}}\right) \tag{12}$$

where k_{bod} is the rate of loss of DO as BOD decays (h^{-1}) and S_{bod} is the half-saturation concentration for the use of DO to satisfy BOD (mg L^{-1}). Note that sedimentation and phytoplankton death also influence BOD, as described in Text S1.

Process rate coefficients for nitrification and BOD in the model are temperature-dependent

$$k_T = k_{T\text{ref}}\theta^{(T - T_{\text{ref}})} \tag{13}$$

where k_T is the process rate at T°C (h^{-1}) and $k_{T\text{ref}}$ is the process rate (h^{-1}) at a reference temperature (20°C).

Reaeration. Reaeration estimation (K) in the model accounts for reaeration at the water surface and at weirs. Reaeration at the water surface represents the rate of change in DO concentration (F, mg O_2 L^{-1} h^{-1}) in the water column via the exchange at the air-water interface. F is represented by a transfer coefficient (k_{rea}, h^{-1}, see Text S1) and a DO deficit term, which is the difference between the saturated DO concentration (O_{sat}, mg L^{-1}) and the actual DO concentrations ($C_{\text{DO},o}$, mg L^{-1}) in the water column

$$F = k_{\text{rea}}(O_{\text{sat}} - C_{\text{DO},o}) \tag{14}$$

Weirs in the river create a head loss, which can aerate or deaerate water depending upon the upstream DO concentrations, creating an instantaneous change in the DO concentrations. Hence, it is important to consider weirs, especially in a heavily regulated river like the Thames. The aeration effect of weirs (W, mg L^{-1}) is calculated as

$$W = O_{\text{sat}} - \left[\frac{O_{\text{sat}} - C_{\text{DO},o}}{R_{\text{ODR}}}\right] \tag{15}$$

where R_{ODR} is the oxygen deficit ratio (Text S1).

Study area and model application

The Thames is a lowland river in southern England with a length of 354 km and a catchment area of around 10,000 km^2 (Marsh and Hannaford 2008). Much of the catchment is underlain by Chalk and Oolitic Limestone aquifers that provide around 40% of the public water supply in the catchment (Bloomfield et al. 2011). In this study, we focused on a 62-km stretch in the lower part of the catchment from Caversham to Runnymede (see Fig. S1). This river stretch comprises 14 locks and weirs. Effluents from six sewage treatment works along the river stretch affect the water quality. During the monitoring period (2013–2014), flow varied from 5 to 360 $m^3\ s^{-1}$. Residence time varied from 9 to 112 h at 90^{th} and 10^{th} percentile flows, respectively. A favorable combination of light availability, high temperatures, and long residence times during mid-spring to mid-summer causes excessive phytoplankton production, which results in large fluctuations in river DO levels (Bowes et al. 2016; Pathak et al. 2021).

River hydrology and water quality monitoring is routinely carried out in the River Thames. Detailed description of data sources, model development, and application is given by Pathak et al. (2021) and in Text S1. The model estimates GPP and ER at the end of each reach in the network (Table S1), but for brevity, we focus our discussion on outputs at Sonning (upstream site) and Runnymede (downstream end).

Empirical analysis

We performed site-wise (Sonning and Runnymede) GLS regression at weekly resolution to examine the sensitivity of GPP and ER to observations of multiple physico-chemical determinands. The observations of physico-chemical determinands were available on a weekly basis and comprised flow, PAR, T, dissolved inorganic nitrogen (DIN), SRP, dissolved organic carbon (DOC), and suspended sediment (SS) concentration. Temperature observations were transformed to $1/(k_bT)$ as per the Metabolic Theory of Ecology, where T is the temperature in Kelvin and k_b is the Boltzmann constant (8.62×10^{-5} eV K^{-1}). GLS was used to account for the residual autocorrelation using the *nlme* package (Pinheiro et al. 2017). Selection of relevant predictors for GLS models was carried out following Feld et al. (2016). Specifically, we performed an exploratory analysis using the random forest (Breiman 2001) machine learning technique (*randomForestSRC* package, Ishwaran and Kogalur 2017) to derive the hierarchy of the most influential stressors and interactions that explain GPP and ER dynamics. Please *see* Text S1 for details.

We also used river water fluorescence observations (collected at weekly resolution from January to July in 2013 and less frequently at other times during 2013–2014 at Sonning and Runnymede sites: for details, see Old et al. 2019) to explore whether ER prediction could be improved by adding water fluorescence information in the GLS models. Fluorescence signals contain information about organic matter composition. Specifically, tryptophan-like fluorescence represents degradable organic matter from farm wastes and sewage discharges and hence, can be related to river BOD (here, R_{wc}) (Hudson et al. 2008). Therefore, our goal was to check if the modeled R_{wc} could be explained in terms of tryptophan-like fluorescence.

Results

Model performance

The model satisfactorily reproduced flow, physico-chemical water quality, and biomass variation along the river stretch (Table S4, Fig. S2) (details in Pathak et al. 2021). High diel fluctuations in DO coincided with high phytoplankton blooms and low flows (Fig. S2). DO levels were slightly over-estimated in spring and under-estimated during the rest of year. Seasonality and timings of high diel fluctuations were well-captured by the model with Nash–Sutcliffe Efficiency > 0.47 and percentage error in mean up to 11% (Table S4). Sonning showed overall over-estimation (up to 14%) of DO concentrations as opposed to Taplow and Windsor, which showed slightly under-estimated (up to 8%) DO concentrations (Table S4). Overall, the model satisfactorily captured the seasonality of DO concentrations along the river stretch.

Spatio-temporal variation in ecosystem metabolism

GPP followed phytoplankton seasonality showing maximum productivity during the biomass-growing season and low productivity during the rest of the year (Fig. 2). Peak GPP was higher in 2013 (up to 21 mg $O_2\ L^{-1}\ d^{-1}$) compared to 2014 (> 10 mg $O_2\ L^{-1}\ d^{-1}$) due to relative inter-annual magnitudes of phytoplankton blooms (Fig. S2). Increase in nutrient concentrations did not result in increase in GPP. In contrast, primary production during the growing season reduced nutrient concentrations through uptake (Fig. S8). During the growing season, ER was dominated by autotrophic respiration and more or less mirrored the GPP trend, although with a lesser magnitude (< 10 mg $O_2\ L^{-1}\ d^{-1}$). In comparison to Runnymede (downstream end), Sonning (upstream site) was characterized by higher nitrification loss throughout the year and higher R_A during the growing season, resulting in generally higher ER upstream.

Values of k_{rea} varied from 0.2–1.1 d^{-1} to 0.3–2.6 d^{-1} at the upstream (Sonning) and downstream (Runnymede) site, respectively. Higher k_{rea} values at the downstream site were due to shallower depths (mean depth = 1.7 m) and faster velocities (mean velocity = 0.73 m s^{-1}) compared to the upstream site, which showed mean depth and velocity of 2.3 m and 0.49 m s^{-1} as calculated in the model. Total estimated reaeration ranged from −1.6 to +1.1 mg $O_2\ L^{-1}\ d^{-1}$ and −4.6 to +2.2 mg $O_2\ L^{-1}\ d^{-1}$ at Sonning and Runnymede, respectively. Runnymede showed higher reaeration during the biomass-growing season when DO saturation went up to 130–150% (Fig. S3). On average over a day during the growing season, O_{sat} at Sonning and Runnymede varied by 14% and 18%,

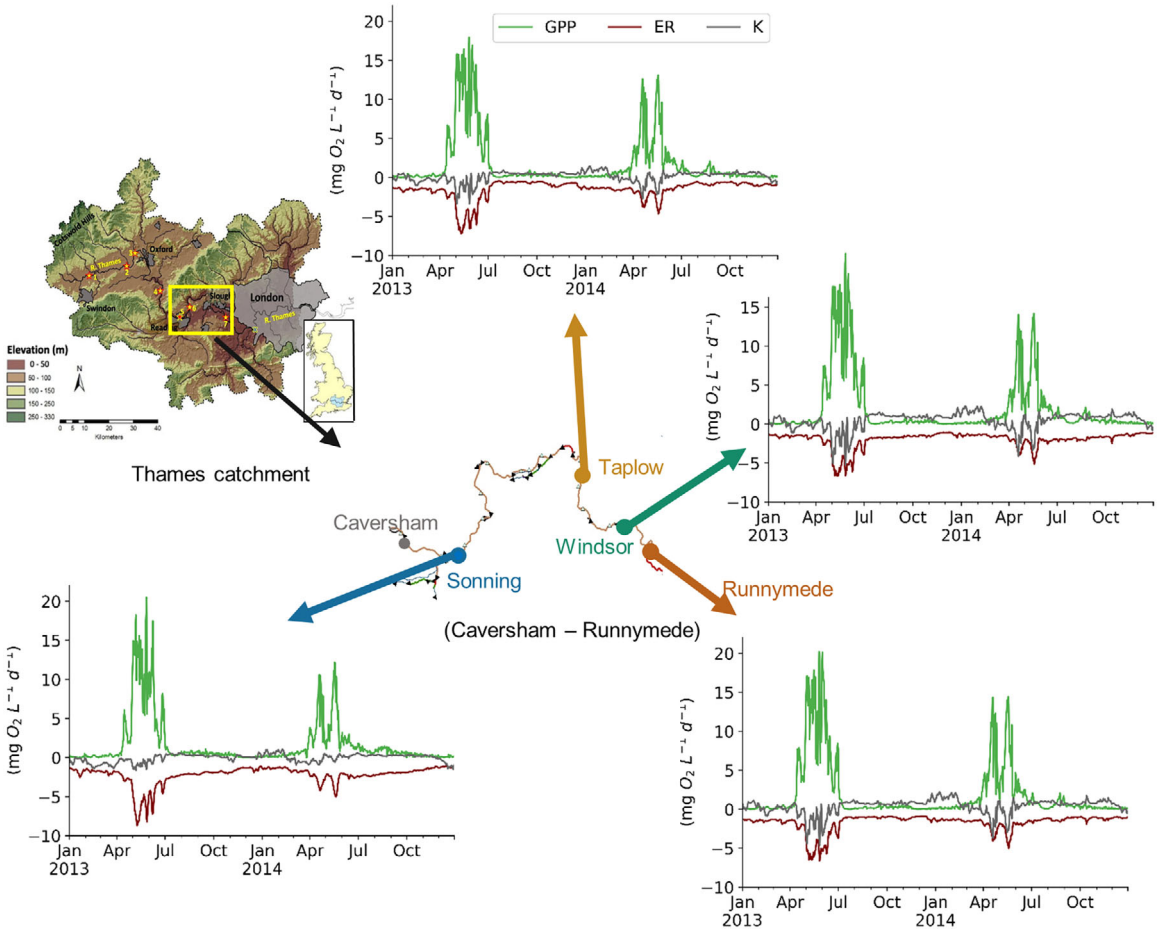

Fig. 2. River Thames catchment and time-series of modeled gross primary production (GPP), ecosystem respiration (ER), and reaeration (K) aggregated at daily scale at Sonning, Taplow, Windsor, and Runnymede sites during 2013–2014. The catchment map at the top left corner is adapted from Bowes et al. (2016).

respectively. During the 2-yr modeling period, DO at Sonning was super-saturated for around 40% of the time (more frequently than at Runnymede [20%]), as reflected by the negative reaeration at Sonning during most of the modeling period (Fig. 2).

The lower River Thames, particularly in the downstream reaches, was dominantly autotrophic during April–June with GPP/ER > 1 and mainly heterotrophic (GPP/ER < 1) during the rest of the year (Fig. 3a). GPP increased up to four times as high as ER during April–June due to large algal blooms. Annual net ecosystem productivity (= GPP − ER) estimates were −192 mg O_2 L^{-1} yr^{-1} and 87 mg O_2 L^{-1} yr^{-1}, and annual GPP/ER ratios were 0.8 and 1.1 at Sonning and Runnymede, respectively. Mid-reaches showed both overall autotrophy at Taplow (net ecosystem productivity = 102 mg O_2 L^{-1} yr^{-1}, GPP/ER = 1.2) and heterotrophy (net ecosystem productivity = −14 mg O_2 L^{-1} yr^{-1}, GPP/ER = 1) at Windsor. Excluding the upstream reaches, annual GPP/ER ratio along the channel was close to 1 in the lower Thames.

Downstream reaches showed a lag of up to 4 h for GPP to peak from the upstream site, Sonning. Hourly GPP increased downstream, but ER variation showed no such trend. Average annual hourly ER increased in response to temperature increase during the day and subsided during the night with temperature decrease, showing a hysteresis effect (Fig. S4). Average annual hourly GPP also showed a hysteresis effect with PAR (Fig. S4). Mean daily GPP and ER at upstream (Sonning) and downstream (Runnymede) sites during the 2-yr period varied from 1.7 ± 3.2 to 2.0 ± 3.8 mg O_2 L^{-1} d^{-1}, and 2.2 ± 1.2 to 1.5 ± 1.2 mg O_2 L^{-1} d^{-1}, respectively.

Relative contribution of autotrophic primary production and respiration was maximum during April–June (Fig. 4). The rest of the year was characterized by low oxygen production and R_A throughout the river stretch. These months, however, showed another oxygen source through diffusion from air at all sites except Sonning, where diffusion was mainly into the atmosphere (DO sink). During autumn and winter at Windsor and Runnymede, oxygen addition from reaeration exceeded oxygen production from GPP. R_{wc} and R_{nitri} at Sonning (total 79% of ER) and Taplow (total 70% of ER) mainly governed ER. At Windsor and Runnymede, ER also

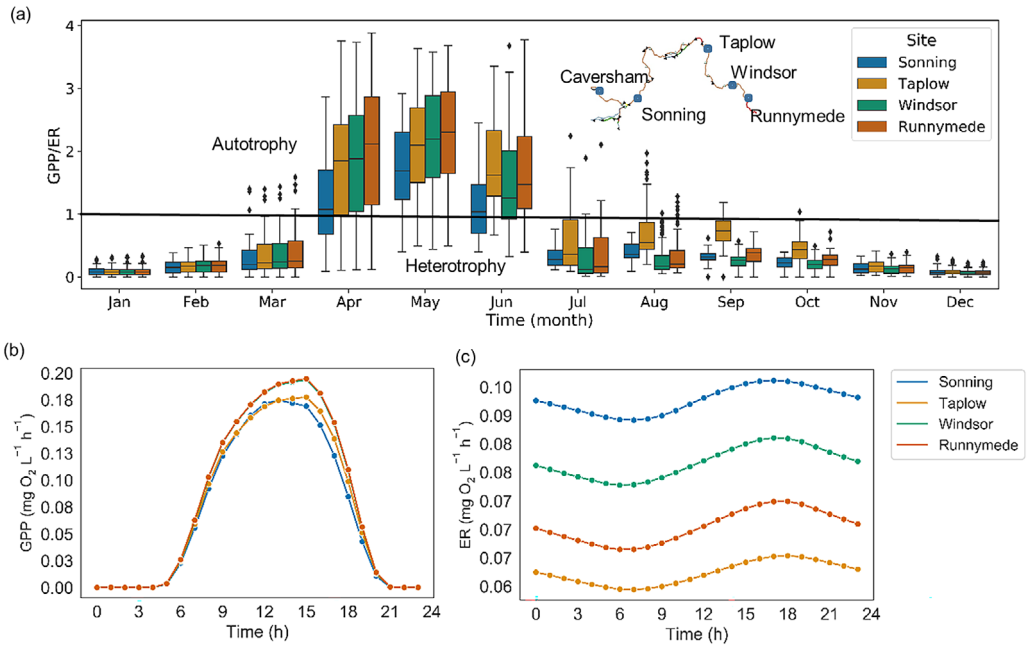

Fig. 3. Box plot of monthly variation in GPP/ER ratio (**a**) and time-series of average hourly variation in GPP (**b**) and ER **(c)** at all calibration sites during 2013–2014. ER, ecosystem respiration; GPP, gross primary production.

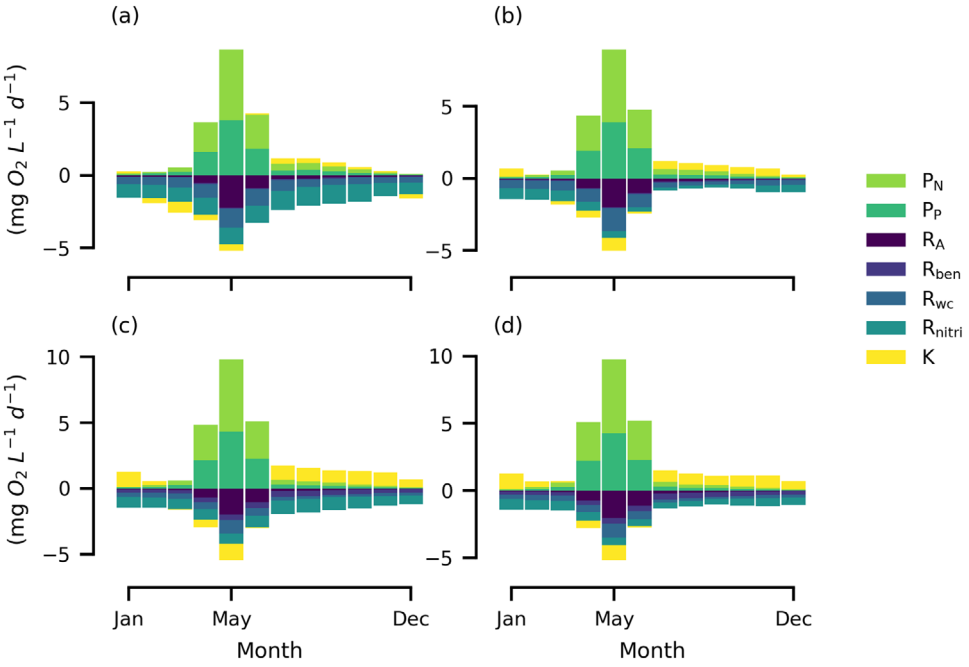

Fig. 4. Average monthly metabolic source and sink pathways at calibration sites (**a**) Sonning, (**b**) Taplow, (**c**) Windsor, and (**d**) Runnymede during 2013–2014. K, reaeration; P_N, oxygen produced during nitrate assimilation by phytoplankton; P_P, autotrophic production; R_A, autotrophic respiration; R_{ben}, benthic oxygen demand; R_{wc}, DO loss due to BOD decay; R_{nitri}, DO loss from nitrification.

included significant contribution from benthic communities (17–19%) in addition to the aforementioned processes (~ 60%).

Sensitivity of river metabolism to multiple stressors

The variables retained in the best approximating GLS models of GPP included PAR, $1/k_bT$, SRP, flow, and seasonality

components. For ER, the best approximating GLS models included SRP, flow, SS, and seasonality components. Inclusion of seasonality components (sine/cosine terms) improved GLS model performance in all cases. For GPP at Sonning for example, a model with seasonality plus physico-chemical predictors (Akaike information criterion = 75.58) performed better than the model with either only seasonality predictors (Akaike information criterion = 136.77) or only physico-chemical predictors (Akaike information criterion = 82.61). Final GLS models at both sites showed good agreement between observed and fitted values with $r > 0.8$ (Table 1). However, high values were under-estimated by these models, especially for ER (Fig. 5). Under-estimation of ER, when investigated using a comparison of R_{wc} and tryptophan-like fluorescence component (Fig. S10), revealed a positive relationship at Sonning ($r = 0.45$, $p < 0.05$). Runnymede, however, did not show a significant relationship between R_{wc} and the tryptophan-like fluorescence component.

Significant interaction effects between $1/k_bT \times$ SRP and PAR $\times$ $1/k_bT$ were found at Sonning for GPP variation (Table 1). The interaction between $1/k_bT$ and SRP at Sonning (Fig. 6a) is an opposing interaction, i.e., the effect of one variable is reversed above a certain limit of another variable. Another significant interaction between PAR and $1/k_bT$ is observed to be antagonistic, i.e., one variable attenuates the effect of the other variable (Fig. 6b). Similar to Sonning, Runnymede also showed an opposing interaction between $1/k_bT$ and SRP (Fig. 6c). The final GLS models of ER did not include any interaction effects.

Table 1. Summaries for best approximating generalized least squares models for gross primary production (GPP) and ecosystem respiration (ER) with autoregressive structure of order 1 including standardized effect size (SES), standard error of the estimate (SE), *t*-test value of the coefficient and its associated *p* value and the Pearson's product moment correlation coefficient (*r*) for the model fits.

Variable	SES	SE	*t* value	*p*	*r*
GPP-Sonning					
(intercept)	0.132	0.118	1.124	0.264	0.89
PAR	0.369	0.052	7.088	0.000	
$1/k_bT$	−0.223	0.140	−1.599	0.114	
SRP	−0.048	0.074	−0.647	0.520	
Flow	0.370	0.107	3.445	0.001	
Sine component	−0.063	0.173	−0.363	0.718	
Cosine component	−0.540	0.159	−3.395	0.001	
$1/k_bT \times$ SRP	0.273	0.093	2.950	0.004	
PAR $\times$ $1/k_bT$	0.071	0.044	1.615	0.110	
GPP-Runnymede					
(intercept)	0.078	0.171	0.453	0.651	0.81
PAR	0.416	0.058	7.175	0.000	
$1/k_bT$	0.136	0.180	0.752	0.454	
SRP	−0.102	0.085	−1.195	0.235	
Flow	0.497	0.126	3.950	0.000	
Sine component	−0.304	0.230	−1.321	0.190	
Cosine component	−0.667	0.209	−3.194	0.002	
$1/k_bT \times$ SRP	0.259	0.107	2.428	0.017	
ER-Sonning					
(intercept)	−0.005	0.109	−0.047	0.963	0.87
Cosine component	−0.640	0.130	−4.932	0.000	
Sine component	0.067	0.172	0.389	0.698	
SRP	−0.385	0.069	−5.540	0.000	
SS	0.147	0.068	2.163	0.033	
Flow	−0.255	0.163	−1.570	0.120	
ER-Runnymede					
(intercept)	−0.003	0.107	−0.032	0.975	0.88
SRP	−0.544	0.060	−9.023	0.000	
SS	0.127	0.067	1.902	0.061	
Flow	0.312	0.152	2.052	0.043	
Sine component	−0.176	0.160	−1.104	0.273	
Cosine component	−0.612	0.120	−5.092	0.000	

Discussion

Estimating metabolism with process-based modeling

The model presented here has several advantages over conventional open-channel methods. The hourly model supports network-scale prediction of metabolism rates unlike open-channel methods that are generally applied at a river-reach scale. Network-scale modeling allows us to translate the influence of what is happening upstream in the river (e.g., flow management, sewage discharges, etc.) to the downstream DO dynamics. The model presented here is also particularly advantageous over open-channel methods when DO observations are not continuously available. In such scenarios, the estimation of metabolism rates outside of the available data periods relies on further assumptions (see Bernhardt et al. 2018). The hourly model overcomes this challenge as it simulates DO using observations of environmental variables and their process-linkages with DO dynamics, thus eliminating the dependence of metabolism estimation strictly on continuous DO measurements. The model can derive the relative contribution of autotrophic and heterotrophic respiration (here $R_A/P_{GPP} = 0.32$ and 0.35 for Sonning and Runnymede, respectively). Estimates of R_A/P_{GPP} are useful to estimate the autotrophic base of food webs (Hall Jr and Beaulieu 2013) and to calculate carbon spiraling in rivers (Newbold et al. 1982). Furthermore, the model can be adapted to study the impact of land use changes by translating diffuse nutrient fluxes into tributary inputs. The model is also useful to predict changes in metabolic regime of rivers under different climate and management scenarios (Hutchins et al. 2018).

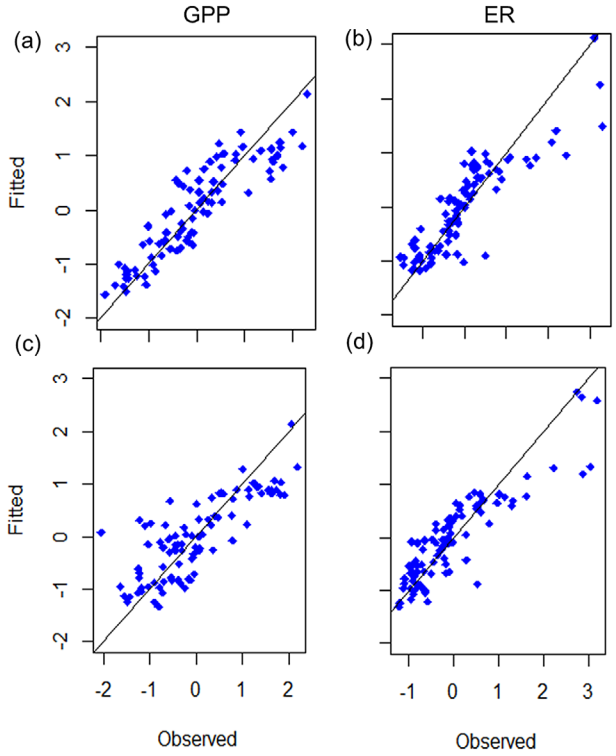

Fig. 5. Fitted vs. observed values of standardized gross primary production (GPP) and ecosystem respiration (ER) in generalized least squares models at Sonning (**a**, **b**) and Runnymede (**c**, **d**) sites. The black line in the plots is the $y = x$ line.

In spite of the satisfactory reproduction of river water quality dynamics, process-based models invariably include some uncertainties linked to input data quality, process simplifications in the model structure, and/or from process knowledge gaps (Hrachowitz et al. 2014). For example, the model in this study assumes that *Stephanodiscus hantzschii* diatoms dominate phytoplankton biomass in the lower Thames throughout the year as found by Read et al. (2014). In reality, multiple algal groups may thrive together and result in within-year compositional change. However, testing of the model using parameters reflecting different phytoplankton groups has reinforced the assumption of diatom dominance (Pathak et al. 2021).

The model also takes a simple approach to calculate reaeration flux. The model uses a previously developed

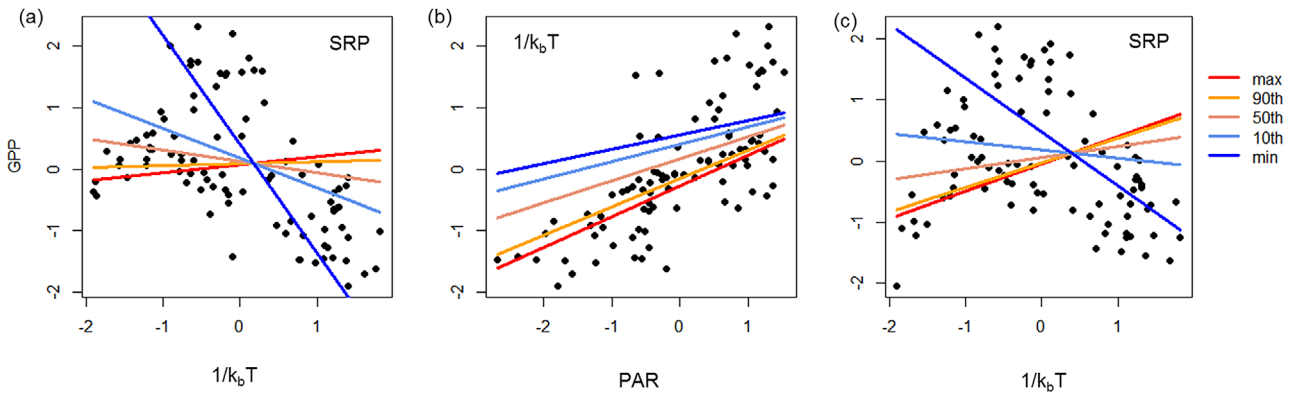

Fig. 6. Pairwise interactions in generalized least squares models for gross primary production (GPP) at (**a**) Sonning ($1/k_bT$, SRP), (**b**) Sonning (PAR, $1/k_bT$), and (**c**) Runnymede ($1/k_bT$, SRP). Lines represent fitted response to one variable while keeping the second variable values fixed at minimum, maximum, 10th, 50th, and 90th percentiles. $1/k_bT$, transformed water temperature; PAR, photosynthetically active radiation; SRP, inorganic phosphorus.

empirical equation (Owens et al. 1964) to calculate the reaeration coefficient. Out of several equations reported in the literature, it is recommended (Young et al. 2004; Aristegi et al. 2009) to use the Owens et al. (1964) equation for rivers with low velocities such as the River Thames. Additionally, by comparing the commonly used empirical equations (O'Connor and Dobbins 1958; Churchill et al. 1962; Owens et al. 1964) in water quality models (Chapra 2008), we found that the choice of equation did not affect metabolism rates largely because the reaeration coefficients were relatively low (average k_{rea} 0.5–1.3 d^{-1}) in the River Thames (Fig. S6). These approaches may be unreliable to use with open-channel methods as (1) it is possible to get a good model fit to the observations because of equifinality in the model and (2) the errors in reaeration estimates will directly translate to GPP and ER estimates (Holtgrieve et al. 2010). However, GPP and ER in our model are simulated from biomass variation and underlying biochemical processes. Moreover, we found that metabolism rates were insensitive to the reaeration values derived for the flow regime of the River Thames (Fig. S6), thus reducing the uncertainties related to reaeration fluxes.

Using the process-based model, we made a detailed quantification of metabolism fluxes, and although the application presented here is data intensive, it can be implemented with lower resolution inputs as shown by Pathak et al. (2021). The model application here uses hourly scale inputs of light, water temperature, DO, and Chl *a*, and daily scale input of flow. However, model sensitivity analysis by Pathak et al. (2021) suggests that the model outputs are not sensitive to the timescale of water quality inputs, but are highly sensitive to that of radiation inputs. If the model is driven by weekly DO, temperature, and Chl *a* observations, instead of hourly as presented, there is little loss of performance at the downstream sites (e.g., at Windsor, NSE values for DO and Chl *a* change from 0.59 to 0.57 and from 0.80 to 0.73, respectively). The outcome from the model sensitivity analysis is reassuring for model applications elsewhere since water quality determinands are irregularly monitored at high-resolution in rivers. Unlike water quality, flow is often routinely (e.g., daily) monitored in rivers, and high-resolution (e.g., hourly) radiation information is easier to obtain either directly or indirectly based on catchment location and sunshine hours (Pathak et al. 2021).

It is still difficult to gather process-rate information in rivers, which is also the case in the lower Thames. For this study, BOD information was only available at a monthly scale and data on benthic oxygen demand were absent. Despite scarce data, our estimates of GPP and ER rates agree well with the findings of Hutchins et al. (2020), who used the Delta method (Chapra and Di Toro 1991) to estimate metabolism rates in the lower Thames. For prediction of the specific metabolic pathways, we still have more confidence at a larger temporal scale due to the lack of high-resolution process-rate information. Though these limitations may introduce uncertainties during model calibration, other extensive applications of the QUESTOR model in the River Thames (Waylett et al. 2013; Hutchins et al. 2018, 2020) provide confidence in the parameter calibration in this study as the calibrated values (Pathak et al. 2021) lie within similar ranges.

Multiple stressor controls on metabolism dynamics

Faster decline of GPP than ER (Fig. 3a) indicates a decrease in primary productivity compared to biological activity and a shift in river energetics from autotrophic to heterotrophic. The shift in river energetics implies reliance of the metabolic regime on stored algae from summer and/or allochthonous carbon sources during autumn and winter. We find that PAR, water temperature, SRP, and flow variation mainly control the GPP dynamics in the lower Thames (Table 1). Nitrate concentrations in the river are present in excess throughout the year and do not limit primary production. Phosphorus concentrations, on the other hand, decrease with high biomass growth and become limiting in summer (Pathak et al. 2021). Light availability, as commonly observed (Bott et al. 1985; Mulholland et al. 2001), increases the GPP with increase in photosynthetic production. High GPP occurred at mid-temperatures (Fig. 6), which is similar to the findings of Bowes et al. (2016) and Pathak et al. (2021), who reported optimum temperature ranges (~ 11–18°C) for high phytoplankton growth in the lower River Thames. An opposing interaction between $1/k_bT$ and SRP (Fig. 6) shows that GPP increases with water temperature, but only at low SRP levels. SRP depletion with increased GPP (Fig. S8) indicates biomass uptake (Bowes et al. 2016). Hence, we believe that the opposing interaction between $1/k_bT$ and SRP is more of a causal effect that occurs during the growing season when phytoplankton utilizes SRP and peaks with increase in temperature.

The GLS model derived a positive slope to represent the overall response of GPP to flow variation. However, a closer look at the partial dependence plot of the GPP-flow relationship (Fig. S8) showed that GPP increased only up to a certain flow threshold and began decreasing with further increase in flow (Pathak et al. 2021). Maximum GPP occurred during mid-spring to mid-summer due to the presence of large phytoplankton blooms during periods of low flows (Bowes et al. 2016; Pathak et al. 2021). The rest of the year represented extremely low GPP due to low phytoplankton biomass (Figs. S2 and 2). Such a seasonal variation in GPP is commonly observed in temperate rivers, where GPP peaks during periods of high light availability and low flows (Roberts et al. 2007) and significantly reduces during high flows that flush away primary producers (Wang et al. 2019).

High ER occurred during mid-spring to mid-summer in response to high GPP, reflecting high autotrophic respiration of phytoplankton biomass (Fig. 2). A strong coupling between GPP and ER is common in rivers as a major part of the organic matter produced during photosynthesis is immediately respired by autotrophs and their closely associated heterotrophs (Hall Jr and Beaulieu 2013). During the biomass-growing season, Sonning represented higher overall ER

compared to Runnymede because of added contributions from nitrification and BOD decay processes. Lower velocity at Sonning may have promoted higher respiration from organisms suspended in the water column with more residence time to utilize the DO in the reach. Through empirical modeling, we derived SRP, flow, and SS to be the most important controls of ER in the river (Table 1). Addition of nutrients in the river did not result in increased biomass growth. On the contrary, we observed high biomass growth coinciding with low SRP levels, suggesting algal uptake. Higher biomass should have resulted in increased decomposer activity (Pascoal et al. 2003) and high resource availability for feeders (Niyogi et al. 2003), causing higher microbial respiration in addition to the high R_A and both contributing to increase in ER. Thus, the co-occurrence of high ER (from algal growth) and SRP depletion may have put SRP in the list of important predictors for ER variation.

On the other hand, high ER in response to high suspended sediments probably indicates organic matter delivery attached to sediments (Roberts et al. 2007; Aspray et al. 2017). Runnymede showed high ER in response to increase in flow, which can again be related to flushing of upstream biomass and organic matter supply along with sediment delivery. However, Sonning showed a negative correlation between flow and ER. Similar to the relationship between flow and GPP, the partial dependence plot (Fig. S9) of ER in response to flow at Sonning showed high ER at mid-flows that decreased at very high flows, as opposed to Runnymede that showed constant ER after a certain flow threshold was reached. GPP at Runnymede still decreased after a certain flow threshold was reached. The relationship between flow-ER at Runnymede indicates that in spite of the biomass flushing, there is still an allochthonous organic matter supply that supports high respiration.

Comparing modeling approaches

In spite of the overall good model performance ($r > 0.8$), peak metabolism rates were under-estimated in the GLS models. Some information about rapidly changing dynamics could have been lost in the empirical modeling as this approach uses weekly time-scale information about the environmental stressors to predict GPP and ER. The under-estimation of GPP can be attributed to the under-estimation of Chl *a* concentrations in the process-based model (Table S4). ER under-estimation suggests that some important metabolism controls might be missing in the empirical analysis. For example, only a limited number of physico-chemical controls were directly included. Land use pressures can also be important as these can contribute large amounts of nutrients and fine particles in the river, and influence primary producers and heterotrophs (dos Reis Oliveira et al. 2019). However, we have accounted for these influences through proxy variables such as suspended sediment and nutrient concentrations. Although we included flow as a control variable, specific matrices of hydrology (e.g., low-flow events and duration) may improve the model performance. Grazers may also influence ER through phytoplankton predation (Welker and Walz 1998) and oxygen consumption through respiration (Garnier et al. 1999) with strong seasonal patterns (Schöl et al. 2002). Hence, further research on these controls will be useful to improve model predictions and explain the role of seasonality components in the GLS models. The process-based model, on the other hand, includes the influences of these controls directly (e.g., simulation of hydrology, nutrients, and sediment concentrations) or indirectly (e.g., grazing through calibration of death constant).

Additionally, information about organic matter composition may help explain ER under-estimation and improve model performance. R_{wc} at Sonning increased with the tryptophan-like fluorescence component (Fig. S10), which is expected because it represents the presence of organic matter that can be easily degraded by microbes, resulting in high R_{wc}. Runnymede, on the other hand, did not show a significant relationship, probably because of the limited ability of the process-based model to accurately represent BOD fluxes. As discussed earlier, there is a paucity of BOD data, making it difficult to estimate the BOD decay rate parameter precisely in the model. The process-based model cannot incorporate any additional site-specific sources/sinks of BOD (e.g., internal BOD sources from higher trophic levels). Additionally, the poorer fits and the under-estimation of Chl *a* concentrations at Runnymede (Table S4) suggest that R_{wc} from phytoplankton death is not represented accurately at this site, which may have resulted in a weak relationship between R_{wc} and organic matter availability. Nevertheless, a strong relationship at Sonning still suggests that empirical model performance at the upstream end can be improved with detailed information about organic matter composition as it can be directly linked to BOD (Hudson et al. 2008). Use of water fluorescence indicators (such as full spectra fluorescence excitation–emission matrix or sensors designed to identify tryptophan at specific wavelengths) as an alternative in the absence of BOD information in the river is an important area of future research as it can potentially improve ER prediction in both process-based and empirical approaches.

Although the empirical approach under-estimates the peak values, it is mostly able to mimic the process-based predictions. Combining a physics-based approach with empirical analysis provides powerful possibilities. For example, the empirical models derived in this study can be used for rapid river health assessments across large areas when setting up a complex, process-based model is not feasible. Empirical approaches also provide information about important environmental stressors and their interactions for GPP and ER variation. These established relationships between metabolism rates and environmental stressors can be useful to infer the degradation or recovery of river health following management actions (Jankowski et al. 2021), although the variable

importance and effect sizes of environmental stressors should be considered (Feld et al. 2016). A process-based approach, on the other hand, presents a readily available tool to study river ecosystem functioning in response to changing multiple environmental stressors (Heathwaite 2010). The process-based model in this study can be improved further to create a management tool by linking it with multiple stressor effects such as flow, temperature, nutrients and sediment modifications derived from the empirical approach. Overall, the comparison of empirical and process-based approach provides useful insights into modeling limitations and directions for future work.

Summary

An approach to estimate ecosystem metabolism rates (GPP, ER) in lowland rivers with a network-scale, process-based water quality model overcomes the current challenges in metabolism modeling by accounting for oxygen advection under varying flows and oxygen transformations due to biogeochemical processes. Only a few river modeling studies (Payn et al. 2017; Segatto et al. 2020) have attempted to overcome these challenges, but at a much smaller spatial scale (e.g., reach level). The model can easily be extended to an entire catchment if more observations in the catchment are available (Hutchins et al. 2020). The approach presented here uses a previously tested high-resolution river model for water quality prediction in the lower Thames (Pathak et al. 2021). Instead of continuous DO measurements, the process-based approach relies on biomass variation, and the physics of the underlying hydrological- and biochemical-process dynamics to estimate GPP and ER. Therefore, the model has a potential to predict metabolism rates (1) for periods when gaps in continuous DO observations are present; (2) at sites within the modeled river network where continuous monitoring is not carried out; and (3) under future environmental and anthropogenic changes. The model presented here is a step forward in high-resolution modeling of long-term, network-scale predictions of river ecosystem functioning, which in turn, can support ecosystem health assessments using functional indicators (Von Schiller et al. 2017).

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 (Pathak et al. 2020). 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/).

References

Acuña, V., C. Vilches, and A. Giorgi. 2011. As productive and slow as a stream can be—the metabolism of a Pampean stream. J. North Am. Benthol. Society **30**: 71–83. doi:10.1899/09-082.1

Aristegi, L., O. Izagirre, and A. Elosegi. 2009. Comparison of several methods to calculate reaeration in streams, and their effects on estimation of metabolism. Hydrobiologia **635**: 113–124. doi:10.1007/s10750-009-9904-8

Aspray, K. L., J. Holden, M. E. Ledger, C. P. Mainstone, and L. E. Brown. 2017. Organic sediment pulses impact rivers across multiple levels of ecological organization. Ecohydrology **10**: e1855. doi:10.1002/eco.1855

Beaulieu, J. J., C. P. Arango, D. A. Balz, and W. D. Shuster. 2013. Continuous monitoring reveals multiple controls on ecosystem metabolism in a suburban stream. Freshwater Biol. **58**: 918–937. doi:10.1111/fwb.12097

Bernhardt, E. S., and others. 2018. The metabolic regimes of flowing waters. Limnol. Oceanogr. **63**: S99–S118. doi:10.1002/lno.10726

Bloomfield, J., S. Bricker, and A. Newell. 2011. Some relationships between lithology, basin form and hydrology: A case study from the Thames basin, UK. Hydrol. Processes **25**: 2518–2530. doi:10.1002/hyp.8024

Bott, T., J. Brock, C. Dunn, R. Naiman, R. Ovink, and R. Petersen. 1985. Benthic community metabolism in four temperate stream systems: An inter-biome comparison and evaluation of the river continuum concept. Hydrobiologia **123**: 3–45. doi:10.1007/BF00006613

Bowes, M., and others. 2016. Identifying multiple stressor controls on phytoplankton dynamics in the river Thames (UK) using high-frequency water quality data. Sci. Total Environ. **569**: 1489–1499. doi:10.1016/j.scitotenv.2016.06.239

Bowie, G.L., and others. (1985) Rates, constants, and kinetics formulations in surface water quality modeling. U.S. Environmental Protection Agency, Washington, EPA/600/3-85/040.

Breiman, L. 2001. Random forests. Machine Learning **45**: 5–32. doi:10.1023/A:1010933404324

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

Chapra, S. C. 2008. Surface water-quality modeling. Waveland Press. ISBN: 1478608307, 9781478608301.

Chapra, S. C., and D. M. Di Toro. 1991. Delta method for estimating primary production, respiration, and reaeration in

streams. J. Environ. Eng. **117**: 640–655. doi:10.1061/(ASCE)0733-9372(1991)117:5(640)

Churchill, M. A., H. L. Elmore, and E. A. Buckingham. 1962. The prediction of stream reaeration rates. J. Sanitary Eng. **88**: 1–46. doi:10.1061/JSEDAI.0000390

Cox, B. 2003. A review of dissolved oxygen modelling techniques for lowland rivers. Sci. Total Environ. **314**: 303–334. doi:10.1016/S0048-9697(03)00062-7

Demars, B. O., and others. 2011. Temperature and the metabolic balance of streams. Freshwater Biol. **56**: 1106–1121. doi:10.1111/j.1365-2427.2010.02554.x

dos Reis Oliveira, P. C., H. G. van der Geest, M. H. Kraak, and P. F. Verdonschot. 2019. Land use affects lowland stream ecosystems through dissolved oxygen regimes. Sci. Rep. **9**: 19685. doi:10.1038/s41598-019-56046-1

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. J. Animal Ecol. **77**: 802–813. doi:10.1111/j.1365-2656.2008.01390.x

Feld, C. K., P. Segurado, and C. Gutiérrez-Cánovas. 2016. Analysing the impact of multiple stressors in aquatic biomonitoring data: A 'cookbook' with applications in R. Sci. Total Environ. **573**: 1320–1339. doi:10.1016/j.scitotenv.2016.06.243

Fox, L. 1962. Chebyshev methods for ordinary differential equations. *In* The Computer Journal. doi:10.1093/comjnl/4.4.318

Garnier, J., G. Billen, and L. Palfner. 1999. Understanding the oxygen budget and related ecological processes in the river Mosel: The RIVERSTRAHLER approach, p. 151–166. *In* Hydrobiologia. Springer. doi:10.1023/A:1003894200796

Guasch, H., E. Martí, and S. Sabater. 1995. Nutrient enrichment effects on biofilm metabolism in a Mediterranean stream. Freshwater Biol. **33**: 373–383. doi:10.1111/j.1365-2427.1995.tb00399.x

Halbedel, S., and O. Büttner. 2014. MeCa, a toolbox for the calculation of metabolism in heterogeneous streams. Methods Ecol. Evol. **5**: 971–975. doi:10.1111/2041-210X.12207

Hall, R. O., Jr., and J. L. Tank. 2005. Correcting whole-stream estimates of metabolism for groundwater input. Limnol. Oceanogr.: Methods **3**: 222–229. doi:10.4319/lom.2005.3.222

Hall, R. O., Jr., and J. J. Beaulieu. 2013. Estimating autotrophic respiration in streams using daily metabolism data. Freshwater Sci. **32**: 507–516. doi:10.1899/12-147.1

Heathwaite, A. 2010. Multiple stressors on water availability at global to catchment scales: Understanding human impact on nutrient cycles to protect water quality and water availability in the long term. Freshwater Biol. **55**: 241–257. doi:10.1111/j.1365-2427.2009.02368.x

Holtgrieve, G. W., D. E. Schindler, T. A. Branch, and A'mar, Z.T. 2010. Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. Limnol. Oceanogr. **55**: 1047–1063. doi:10.4319/lo.2010.55.3.1047

Hrachowitz, M., and others. 2014. Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. Water Resour. Res. **50**: 7445–7469. doi:10.1002/2014WR015484

Hudson, N., A. Baker, D. Ward, D. M. Reynolds, C. Brunsdon, C. Carliell-Marquet, and S. Browning. 2008. Can fluorescence spectrometry be used as a surrogate for the biochemical oxygen demand (BOD) test in water quality assessment? An example from south West England. Sci. Total Environ. **391**: 149–158. doi:10.1016/j.scitotenv.2007.10.054

Hutchins, M., C. Abesser, C. Prudhomme, J. Elliott, J. Bloomfield, M. Mansour, and O. Hitt. 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. Sci. Total Environ. **631**: 962–986. doi:10.1016/j.scitotenv.2018.03.052

Hutchins, M., G. Harding, H. Jarvie, T. Marsh, M. Bowes, and M. Loewenthal. 2020. Intense summer floods may induce prolonged increases in benthic respiration rates of more than one year leading to low river dissolved oxygen. J. Hydrol. X **8**: 100056. doi:10.1016/j.hydroa.2020.100056

Ishwaran, H. and Kogalur, U. (2017) Random forests for survival, regression, and classification (RF-SRC), R. package version 2.5.0. https://cran.r-project.org/package=randomForestSRC

Izagirre, O., M. Bermejo, J. Pozo, and A. Elosegi. 2007. RIVERMET©: An excel-based tool to calculate river metabolism from diel oxygen–concentration curves. Environ. Model. Software **22**: 24–32. doi:10.1016/j.envsoft.2005.10.001

Izagirre, O., U. Agirre, M. Bermejo, J. Pozo, and A. Elosegi. 2008. Environmental controls of whole-stream metabolism identified from continuous monitoring of Basque streams. J. North Am. Benthol. Society **27**: 252–268. doi:10.1899/07-022.1

Jankowski, K. J., F. H. Mejia, J. R. Blaszczak, and G. W. Holtgrieve. 2021. Aquatic ecosystem metabolism as a tool in environmental management. Wiley Interdiscip. Rev.: Water **8**: e1521. doi:10.1002/wat2.1521

Lázár, A. N., A. Wade, P. Whitehead, C. Neal, and M. Loewenthal. 2012. Reconciling observed and modelled phytoplankton dynamics in a major lowland UK river, the Thames. Hydrol. Res. **43**: 576–588. doi:10.2166/nh.2012.029

Marsh, T. and Hannaford, J. (2008) UK Hydrometric Register. A catalogue of river flow gauging stations and observation wells and boreholes in the United Kingdom together with summary hydrometric and spatial statistics, Bailrigg, England: Centre for Ecology & Hydrology. ISBN: 9780955767227. http://nora.nerc.ac.uk/id/eprint/3093

Mulholland, P., and others. 2001. Inter-biome comparison of factors controlling stream metabolism. Freshwater Biol. **46**: 1503–1517. doi:10.1046/j.1365-2427.2001.00773.x

Newbold, J., P. Mulholland, J. Elwood, and O'neill, R. 1982. Organic carbon spiralling in stream ecosystems. Oikos **38**: 266–272. doi:10.2307/3544663

Niyogi, D. K., K. S. Simon, and C. R. Townsend. 2003. Breakdown of tussock grass in streams along a gradient of agricultural development in New Zealand. Freshwater Biol. **48**: 1698–1708. doi:10.1046/j.1365-2427.2003.01104.x

O'Connor, D. J., and W. E. Dobbins. 1958. Mechanism of reaeration in natural streams. Trans. Am. Society Civil Eng. **123**: 641–666. doi:10.1061/TACEAT.0007609

Odum, H. T. 1956. Primary production in flowing waters 1. Limnol. Oceanogr. **1**: 102–117. doi:10.4319/lo.1956.1.2.0102

Old, G., and others. 2019. Using dissolved organic matter fluorescence to identify the provenance of nutrients in a lowland catchment; the river Thames, England. Sci. Total Environ. **653**: 1240–1252. doi:10.1016/j.scitotenv.2018.10.421

Owens, M., R. Edwards, and J. Gibbs. 1964. Some reaeration studies in streams. Air Water Pollut. **8**: 469–486.

Pascoal, C., M. Pinho, F. Cássio, and P. Gomes. 2003. Assessing structural and functional ecosystem condition using leaf breakdown: Studies on a polluted river. Freshwater Biology **48**: 2033–2044. doi:10.1046/j.1365-2427.2003.01130.x

Pathak D., and others. (2020). Supporting data for 'Hourly prediction of phytoplankton biomass and its environmental controls in lowland rivers' [Data set]. In Water Resources Research (1.0.0). Zenodo. 10.5281/zenodo.4288254

Pathak, D., and others. 2021. Hourly prediction of phytoplankton biomass and its environmental controls in lowland rivers. Water Resour. Res. **57**: e2020WR028773. doi:10.1029/2020WR028773

Payn, R. A., R. Hall Jr., T. A. Kennedy, G. C. Poole, and L. A. Marshall. 2017. A coupled metabolic-hydraulic model and calibration scheme for estimating whole-river metabolism during dynamic flow conditions. Limnol. Oceanogr.: Methods **15**: 847–866. doi:10.1002/lom3.10204

Perkins, D. M., G. Yvon-Durocher, B. O. Demars, J. Reiss, D. E. Pichler, N. Friberg, M. Trimmer, and G. Woodward. 2012. Consistent temperature dependence of respiration across ecosystems contrasting in thermal history. Glob. Chang. Biol. **18**: 1300–1311. doi:10.1111/j.1365-2486.2011.02597.x

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B. and Maintainer, R. (2017) Package 'nlme'. Linear and nonlinear mixed effects models, version 3. https://CRAN.R-project.org/package=nlme

Read, D. S., M. J. Bowes, L. K. Newbold, and A. S. Whiteley. 2014. Weekly flow cytometric analysis of riverine phytoplankton to determine seasonal bloom dynamics. Environ. Sci.: Processes Impacts **16**: 594–603. doi:10.1039/C3EM00657C

Reichert, P., D. Borchardt, M. Henze, W. Rauch, P. Shanahan, L. Somlyódy, and P. Vanrolleghem. 2001. River water quality model no. 1 (RWQM1): II. Biochemical process equations. Water Sci. Technol. **43**: 11–30. doi:10.2166/wst.2001.0241

Roberts, B. J., P. J. Mulholland, and W. R. Hill. 2007. Multiple scales of temporal variability in ecosystem metabolism rates: Results from 2 years of continuous monitoring in a forested headwater stream. Ecosystems **10**: 588–606. doi:10.1007/s10021-007-9059-2

Rode, M., and others. 2016. Sensors in the stream: The high-frequency wave of the present. Environ. Sci. Technol. **50**: 10297–10307. doi:10.1021/acs.est.6b02155

Schöl, A., V. Kirchesch, T. Bergfeld, F. Schöll, J. Borcherding, and D. Müller. 2002. Modelling the chlorophyll a content of the River Rhine–Interrelation between riverine algal production and population biomass of grazers, rotifers and the zebra mussel, *Dreissena polymorpha*. Int. Rev. Hydrobiol. **87**: 295–317. doi:10.1002/1522-2632(200205)87:2/3<295::AID-IROH295>3.0.CO;2-B

Segatto, P. L., T. J. Battin, and E. Bertuzzo. 2020. Modeling the coupled dynamics of stream metabolism and microbial biomass. Limnol. Oceangr. **65**: 1573–1593. doi:10.1002/lno.11407

Uehlinger, U. 2006. Annual cycle and inter-annual variability of gross primary production and ecosystem respiration in a floodprone river during a 15-year period. Freshwater Biol. **51**: 938–950. doi:10.1111/j.1365-2427.2006.01551.x

Von Schiller, D., and others. 2017. River ecosystem processes: A synthesis of approaches, criteria of use and sensitivity to environmental stressors. Sci. Total Environ. **596**: 465–480. doi:10.1016/j.scitotenv.2017.04.081

Wang, J., Z. Zhang, and B. Johnson. 2019. Low flows and downstream decline in phytoplankton contribute to impaired water quality in the lower Minnesota River. Water Res. **161**: 262–273. doi:10.1016/j.watres.2019.05.090

Waylett, A., M. Hutchins, A. Johnson, M. Bowes, and M. Loewenthal. 2013. Physico-chemical factors alone cannot simulate phytoplankton behaviour in a lowland river. J. Hydrol. **497**: 223–233. doi:10.1016/j.jhydrol.2013.05.027

Welker, M., and N. Walz. 1998. Can mussels control the plankton in rivers?—A planktological approach applying a Lagrangian sampling strategy. Limnol. Oceanogr. **43**: 753–762. doi:10.4319/lo.1998.43.5.0753

Whitehead, P., and G. Hornberger. 1984. Modelling algal behaviour in the river Thames. Water Res. **18**: 945–953. doi:10.1016/0043-1354(84)90244-6

Whitehead, P. G., G. Bussi, M. J. Bowes, D. S. Read, M. G. Hutchins, J. A. Elliott, and S. J. Dadson. 2015. Dynamic modelling of multiple phytoplankton groups in rivers with an application to the Thames river system in the UK.

Environ. Model. Software **74**: 75–91. doi:10.1016/j.envsoft.2015.09.010

Young, R., C. Townsend, and C. Matthaei. 2004. Functional indicators of river ecosystem health–an interim guide for use in New Zealand. Cawthron Rep. **870**: 495–523.

Young, R. G., C. D. Matthaei, and C. R. Townsend. 2008. Organic matter breakdown and ecosystem metabolism: Functional indicators for assessing river ecosystem health. J. North Am. Benthol. Society **27**: 605–625. doi:10.1899/07-121.1

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 dataset 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 the Environment Agency for the high-frequency water quality data.

Conflict of Interest

None declared.

Submitted 19 May 2021
Revised 10 March 2022
Accepted 27 March 2022

Associate editor: Robert O. Hall Jr.