LIMNOLOGY and OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 68, 2023, 2461–2475

doi: 10.1002/lno.12435

Coupled primary production and respiration in a large river contrasts with smaller rivers and streams

Sarah S. Roley [1,2*] **Robert O. Hall Jr.** [3] **William Perkins** [2] **Vanessa A. Garayburu-Caruso** [1,2] **James C. Stegen** [1,2]

[1]School of the Environment, Washington State University, Richland, Washington, USA
[2]Pacific Northwest National Laboratory, Richland, Washington, USA
[3]Flathead Lake Biological Station, University of Montana, Polson, Montana, USA

Abstract

Although time series in ecosystem metabolism are well characterized in small and medium rivers, patterns in the world's largest rivers are almost unknown. Large rivers present technical difficulties, including depth measurements, gas exchange (K, d^{-1}) estimates, and the presence of large dams, which can supersaturate gases. We estimated reach-scale metabolism for the Hanford Reach of the Columbia River (Washington state, USA), a free-flowing stretch with an average discharge of 3173 m^3 s^{-1}. We calculated K from semi-empirical models and directly estimated it from tracer measurements. We fixed K at the median value from these calculations (0.5 d^{-1}), and used maximum likelihood to estimate reach-scale, open-channel metabolism. Both gross primary production (GPP) and ecosystem respiration (ER) were high (GPP range: 0.3–30.8 g O_2 m^2 d^{-1}, ER range: 0.8–30.6 g O_2 m^2 d^{-1}), with peak GPP and ER occurring in the late summer or early fall. GPP increased exponentially with temperature, consistent with metabolic theory, while light was seasonally saturating. Annual average GPP, estimated at 1500 g carbon m^{-2} yr^{-1}, was in the top 2% of estimates for other rivers. GPP and ER were tightly coupled and 90% of GPP was immediately respired, resulting in net ecosystem production near 0. Patterns in the Hanford Reach contrast with those in small-medium rivers, suggesting that metabolism magnitudes and patterns in large rivers may not be simply scaled from knowledge of smaller rivers.

Gross primary production (GPP) and ecosystem respiration (ER), the production and consumption of organic carbon (C), reflect aggregate metabolic activity of an ecosystem. The magnitude of GPP indicates the amount of new organic C entering the river, while net ecosystem production (NEP) expresses the degree of heterotrophy and the relative importance of allochthonous vs. autochthonous inputs (Meyer and Edwards 1990; Hoellein et al. 2013). Seasonal patterns of GPP, ER, and NEP define the productivity regime of the river (Savoy et al. 2019). As a result, characterizing reach-scale metabolism is a key method to understanding the fundamental drivers of river food webs and biogeochemistry.

The metabolism of small-medium streams and rivers is reasonably well known (Hoellein et al. 2013; Appling et al. 2018*b*), but estimates for large rivers are much rarer. Single-day estimates show that GPP and ER increase with mean discharge (Q), to a peak of 10 m^3 s^{-1}, and then decline (Hall et al. 2016) and annual GPP and ER increase with annual solar energy (Bernhardt et al. 2022). Time series metabolism data show light appears to drive seasonal and day-to-day variation in GPP, with turbidity controlling light attenuation (Dodds et al. 2013; Hall et al. 2015). Across streams of all sizes, autotrophy (i.e., GPP : ER > 1) occurs only occasionally (Quay et al. 1995; Hoellein et al. 2013; Bernhardt et al. 2022), but can occur below big dams (Genzoli and Hall 2016). Dams and their management also influence metabolism via their effects on turbidity, Q, and algal blooms (Genzoli and Hall 2016; Deemer et al. 2022). With the exception of Dodds et al. (2013) and Quay et al. (1995), however, all metabolism estimates are on channels with $Q < 1000\ m^{-3}\ s^{-1}$. As a result, it is unclear if

*Correspondence: sarah.roley@wsu.edu

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

Author Contribution Statement: S.S.R. conceptualized the study, collected data, analyzed data, and drafted the manuscript. S.S.R. is responsible for the integrity of the data, analysis, and presentation of the findings as a whole. R.O.H. developed methods, conceptualized the study, collected data, and edited the manuscript. W.P. analyzed hydrodynamic data and edited the manuscript. V.G.C. collected data and edited the manuscript. J.C.S. collected data and edited the manuscript.

and how metabolic patterns from small rivers scale to truly large rivers, which we define as the very largest rivers, often with continental-scale drainages. The metabolic regimes of such rivers are virtually unknown.

Large lowland rivers (sensu Keith et al. 2022; e.g., the Columbia, Mississippi, and Yukon in North America) integrate across huge, sometimes continental-scale, watersheds, with their activity representing the net effect of watershed land use, climate, and soils, along with nutrient and C processing in tributaries. Most water flows through such large rivers before reaching the ocean; the large river is the last opportunity for nutrient transformations before the water reaches the ocean. River primary production can also be a large flux of organic matter to marine environments (Behnke et al. 2023). Moreover, large rivers host complex food webs, including iconic fishes (e.g., sturgeon, salmon) and mammals (Keith et al. 2022). Humans demand many ecosystem services from large rivers: irrigation, hydroelectric power, food, shipping, industrial intake, cooling, recreation. Metabolism metrics offer fundamental information on overall biological activity, allowing us to better understand these ecosystems and to better manage them for a multitude of sometimes competing ecosystem services.

The lack of estimates from large lowland rivers is likely a result of several difficulties associated with larger channels. Depth measurements and sensor deployment are more difficult, although not insurmountable. Gas exchange (K) estimates are more challenging: unlike in small-medium rivers, measuring K via tracer gas injections is difficult (though not impossible, Ho et al. 2011). The assumptions behind the one-station methods (i.e., using continuous dissolved oxygen [DO] and temperature data from a single station, Odum 1956; Hall and Hotchkiss 2017) used in small-medium streams may not be met in large rivers. First, transverse mixing is lower, and a sensor at a single location may not reflect DO and temperature concentrations of the full channel, bank to bank. Large rivers often have discontinuities (usually dams) that break the assumption of a long reach with uniform conditions. In addition, water passing over large dams is typically super-saturated with gases, including oxygen (O_2), because the pressure of water dropping from a tall height pushes atmospheric gases into the water beyond saturation. As a result, large dams can interfere with metabolism estimates that rely on changes in O_2 saturation. As a result, researchers who wish to estimate metabolism in large rivers must first assess whether one-station assumptions are met and, if not, either employ two-station methods or adapt existing one-station methods.

Here, we apply existing reach-scale, open-channel, one-station metabolism tools to a continuous DO and temperature dataset in the Columbia River, to estimate GPP and ER and controls on their seasonal variation. To achieve this overall objective, we first assess metabolism methods best suited for large rivers, specifically by exploring methods for estimating K in large rivers and accounting for super-saturation effects of an upstream dam.

Methods

We estimated continuous open-channel metabolism for > 2 yr in the Hanford Reach of the Columbia River (May 2018 through December 2020), using the one-station method. This method requires continuous DO and temperature data, along with reliable estimates of gas exchange (K, d^{-1}). To get these estimates, we had to account for the effects of an upstream dam on DO and consider how to estimate K. We then examined physical attributes: light, temperature, and Q, and how they covaried with GPP and ER.

Site description

The Columbia River is the fifth-largest river in North America. It originates in the Canadian Rockies and flows south, into Washington state (USA), before forming the border between the states of Washington and Oregon and discharging into the Pacific Ocean. Flow rates are influenced by mountain snowmelt and dam operations. The Hanford Reach is a free-flowing section of the Columbia River, stretching 168 km from the Priest Rapids Dam near Mattawa, Washington, USA to the McNary Dam near Pasco, Washington, USA. In the Hanford Reach, the Columbia is an 8^{th}-order river until its confluence with the Snake River, which occurs just upstream of the McNary Dam.

We measured metabolism at the Hanford site, 85 km downstream of the Priest Rapids Dam and just north of the city of Richland, Washington (coordinates are 46.3729 N, −119.272 W). Average annual discharge for this part of the Hanford Reach is 3173 $m^{-3}\,s^{-1}$ and average width is 683 m. The surrounding land use is primarily sagebrush steppe, most of which is part of the Hanford Reach National Monument or the Saddle Mountain National Wildlife Refuge (Fritz et al. 2007; Fritz and Arntzen 2007). The land immediately adjacent to our sensors is part of the Hanford Nuclear Site (Fig. 1). In-stream habitat is somewhat homogeneous across the Hanford Reach, with the bottom composed mostly of gravel and cobble. The river is generally either constrained to a single channel or splits into two channels as it moves around islands. Channel width varies tremendously; for example, at average annual Q, transect widths range from 246 to 1060 m.

The phytoplankton population in this reach has not been assessed since 1979, but at that time, most of the photosynthesizing organisms were planktonic. Both planktonic and benthic populations were composed primarily of diatoms (Neitzel et al. 1982). There are few emergent plants in the reach and submergent vegetation is not well characterized. The water is generally clear, with annual median turbidity < 3 nephelometric turbidity units (Poston et al. 2009). The river has low to medium nutrient availability, with nitrate concentrations < 0.9 mgL^{-1} and some locations < 0.5 mgL^{-1} (DOE 2022).

Dam operations and DO

We deployed a HOBO U26 DO logger to measure DO and temperature every 15 min. The logger was deployed inside a

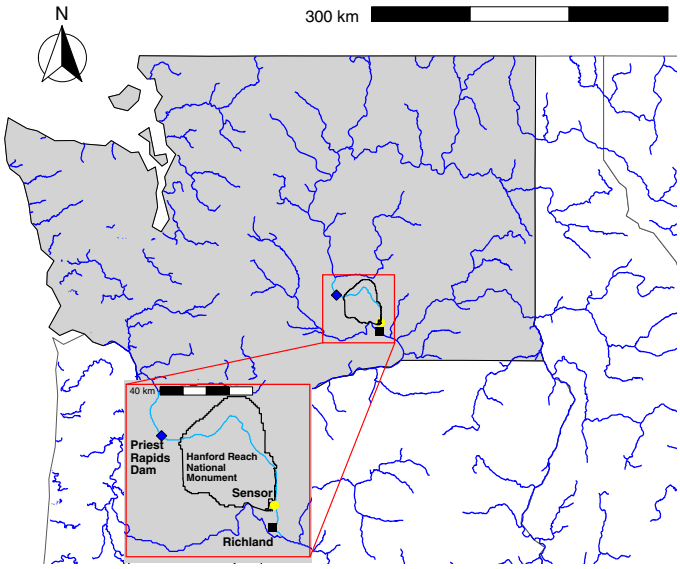

Fig. 1. Location of Hanford Reach within Washington state, USA, showing the location of the DO sensor and the upstream dam. The dam is 85 km upstream of the sensor.

completely submerged, perforated polyvinyl chloride (PVC) pipe that allowed water to flow through while protecting the sensor. We calibrated the sensor every month and replaced the sensor cap every 6 months, per manufacturer's recommendations. We checked and cleaned the data at every download. We did not check for heterogeneity in DO, because hydrodynamic analyses of this reach showed full transverse mixing (*see* "Hydrodynamics" section).

The Priest Rapids dam releases water into its spillway to aid in fish migration and for other dam operations, with releases occurring continuously during the fall and spring. The spillway pushes dissolved atmospheric gas into the river water, thus supersaturating gases when the spillway is operating. The effects of those supersaturated gases were evident downstream at our sensor, where DO remained supersaturated for days at a time, even as it exhibited typical diel patterns of increased DO during the day.

To check how much the spillway explained the supersaturation of DO, we accessed a continuous record of total dissolved gas (TDG), which is collected hourly by Grant County Public Utilities District (http://www.cbr.washington.edu/dart/) at Vernitas Bridge, 15 km downstream of the dam. We plotted the relationship between the spillway flow rate and TDG saturation and fitted a spline curve (Supporting Information Fig. S1). The relationship suggested that TDG saturation covaried with spillway flow rate. Next, we measured DO and argon (Ar) saturation via Membrane Inlet Mass Spectrometry (MIMS). Ar saturation is influenced by physical processes only, while DO saturation can be influenced by physical, chemical, and biological processes. DO and Ar were oversaturated by nearly the same amount (Supporting Information Fig. S2), demonstrating that any biological or chemical processes between the dam and the TDG sensor were insufficient to influence DO saturation. As a result, the TDG sensor is a reasonable indicator of upstream DO concentrations and further evidence that all-day DO supersaturation was caused by dam operations.

ER cannot be reliably estimated when equilibrium concentration of gas exceeds that of saturation concentration. Because some of the DO present at our sensor was a result of dam operations, as opposed to a reflection of biological activity, we subtracted DO present because of dam operations. This correction is based on dissolved gas concentrations near the dam, travel time (τ), and K. The correction had the effect of lowering O_t at the downstream station to saturation concentration (O_{sat}) in the absence of any biological change in the reach. We assumed an exponential loss of supersaturated DO as if it were a tracer gas (Hall and Ulseth 2019):

$$O_{down,(t+\tau)} = O_{sensor,(t+\tau)} - \left(O_{up,t} - O_{sat,up,t}\right)e^{-K\tau} \quad (1)$$

where $O_{down,(t+\tau)}$ is the corrected DO value used in our models, $O_{sensor,(t+\tau)}$ is the DO concentration reported by the sensor at the Hanford site (in mg L^{-1}), $O_{up,t}$ is the DO reported by the TDG sensor at Vernitas Bridge, $O_{sat,up,t}$ is the DO concentration at 100% saturation at the Vernitas Bridge, K is temperature-corrected K_{600} for O_2 and τ is the travel time (d) between the TDG sensor and the DO sensor at the 300 area, calculated as reach length divided by average velocity (v). v varied with Q as $v = 0.553 \times \log(Q) - 3.15$, with Q in $m^3\ s^{-1}$, as determined from the Modular Aquatic Simulation System 1D (MASS1) model (*see* "Hydrodynamics" section).

We did this correction only for times when DO at the TDG sensor was >100% and the spillway was operating (i.e., spillway $Q > 0$). When spillway daily mean $Q > 4245\ m^3\ s^{-1}$, we did not attempt to estimate metabolism at all, because diel DO curves at the sensor did not follow modelable patterns, even with the oxygen correction. Given daily average Q ranging from 1198 to 7493 when the spillway was operating, τ ranged from 0.6 to 1.3 d (Supporting Information Fig. S3). At this range, water was constrained within the river channel. With a K_{600} of 0.5 d^{-1} (*see* Gas exchange section below), the correction to DO at the bottom of the reach ranged from < 0.01 mg L^{-1} when the spillway Q was low, to a maximum of 3.5 mg L^{-1}, when the spillway Q was high (Fig. 2). We corrected the DO at each time point, with an average daily DO correction ranging from 0 to 2 mg L^{-1}. Although the amount of extra DO at the sensor generally increased with spillway Q, there were distinct seasonal patterns, likely because spillway releases can occur during times of both low and high flows; spillway Q is not necessarily correlated with total Q. As a result, sometimes high spillway Q corresponded with a long travel time, allowing more time for extra O_2 to escape.

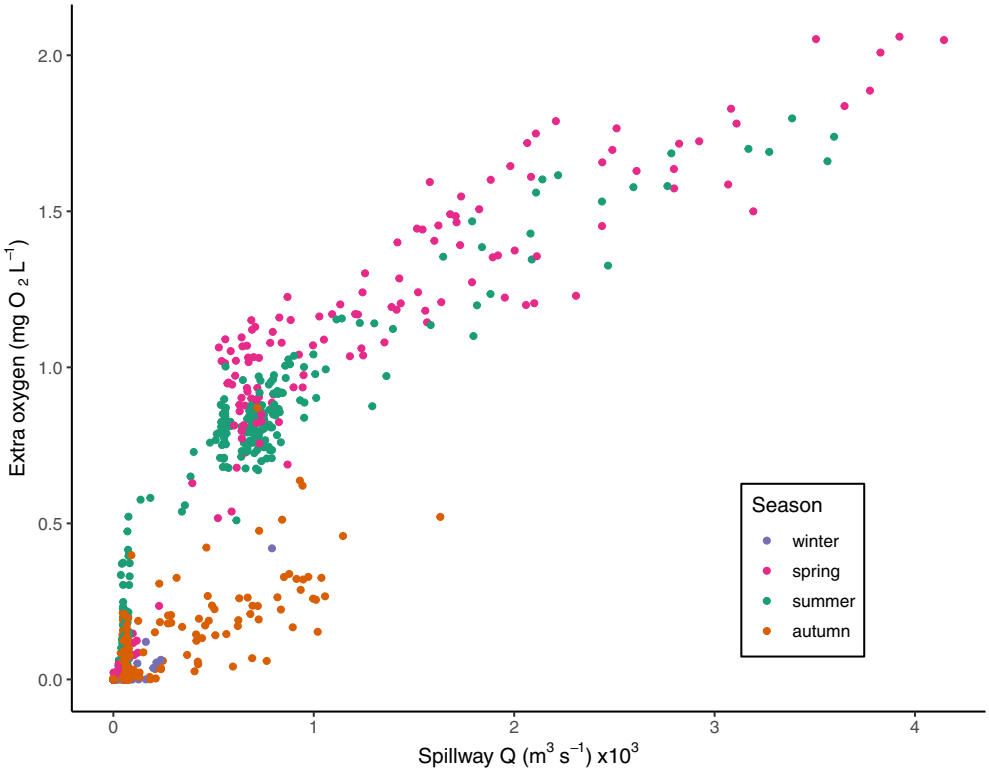

Fig. 2. The amount of extra oxygen, that is, amount of oxygen subtracted from sensor values, based on Eq. (1), generally increased with *Q* at the spillway. Each point is the daily average.

Gas exchange

We had three options for estimating K_{600} (K normalized to a Schmidt number of 600): modeled based on the DO time series, empirically measured based on downstream changes in gas concentrations, and calculated using semi-empirical equations (Hall and Ulseth 2019). We employed all three methods after observing apparently implausible estimates in K_{600} derived from O_2 models. Each method resulted in a different K_{600} value, and so we examined the sensitivity of metabolism estimates by fixing K_{600} at a range of values and examining resultant changes in the metabolic parameters that we actually care about: GPP and ER.

K as a free parameter in oxygen models

We modeled K_{600} as a free parameter, estimated via Bayesian inference using streamMetabolizer (Appling et al. 2017). This method estimated K_{600} as a piecewise function of discharge and was partially pooled across days, a method that can simultaneously estimate metabolism and gas exchange (Appling et al. 2018*a*). We tested the effect of varying priors across two orders of magnitude, but the model returned nearly identical ranges of K_{600}, regardless of priors. With the K_{600} values returned by streamMetabolizer, many of the ER estimates were positive, which is impossible. Furthermore, if the modeled estimates of K_{600} were correct, the extra DO from the dam would have escaped before reaching the DO sensor, but our measured DO data clearly indicates that was not the case. Lastly, estimates were too high; the gas exchange velocity (k_{600}) was around 14 m d^{-1}, which approaches that of the Colorado River in the Grand Canyon, a whitewater river. Because the modeled K_{600} did not match expectations, we estimated K_{600} with empirical and predictive methods.

Direct measures of argon to estimate K

Our empirical K_{600} estimates came from measurements of downstream Ar decline. Typically, K_{600} is estimated via propane or sulfur hexafluoride injections, but large rivers are not conducive to such techniques (Hall and Ulseth 2019). Instead, we took advantage of the upstream dam, which effectively injects atmospheric gases (dinitrogen [N_2], O_2, and Ar) into the river continuously when the spillway is running. We sampled dissolved gas concentrations 4 km downstream of the dam. We then took samples of the same "packet" of water when it reached the downstream sensor, usually 18–24 h later (travel time plotted in Supporting Information Fig. S3). The sampling time was based on a relationship between velocity, estimated with the MASS1 model (*see* "Hydrodynamics" section), and *Q* at the Priest Rapids USGS gauge (12472800). We sampled dissolved gas at *Q* ranging from 1984 to 2926 $m^3 s^{-1}$, and most of our metabolism estimates were made with *Q* from 2000 to 4000 $m^3 s^{-1}$. All of our measurements were thus within the typical range of *Q*, albeit not encompassing the full

range. Our estimates of K did not vary with Q, however (*see* below), despite measurements made over a range of nearly 1000 $m^3 s^{-1}$, so we are confident this sampling was adequate.

To sample dissolved gases, we filled 12-mL Exetainers (Labco, Lampeter, Wales, UK), using a dissolved gas sampler. The sampler was a 3.8-cm diameter PVC pipe, with a 1-cm flexible tube affixed close to the bottom. The bottom of the PVC was closed with a rubber stopper. We filled the pipe with water and then filled the Exetainer from the bottom by placing the flexible tubing on the bottom of the Exetainer and filling to an inverted meniscus. To prevent biological activity from changing gas concentrations, we preserved the sample with 0.1 mL of 50% zinc chloride. For each sample, we recorded barometric pressure (BP, mm Hg) and water temperature with a YSI ProODO (YSI, Yellow Springs). We measured dissolved Ar (A_{meas}) on the samples using MIMS following Kana et al. (1994) and Hall and Madinger (2018). We calculated fractional excess Ar (A) in samples as $A = (A_{meas}/A_{sat}) - 1$. We estimated the downstream decline for each of 4 separate sampling events using a log-transformed exponential decay model with variable intercepts (one for each day) and a completely pooled slope (K). Because all sample events were close in time, we had no reason to believe that gas exchange (i.e., K) would vary from day to day.

$$\log(A_{t,j}) = \log(A_{0,j}) - Kt + \varepsilon_{t,j} \quad (2)$$

where $A_{t,j}$ is A at travel time t downstream and sampling event j, and $A_{0,j}$ is A at the upstream station at sampling event j. We fit this model using Bayesian inference because we have prior information on K_{600} from other means. Prior probability for K was $K \sim \text{gamma}(3.7, 6.7)$; based on moment matching from a distribution that has a mean = 0.55 and SD = 0.29 and derives from empirical estimates from large, flat rivers (Alin et al. 2011). We used gamma distribution for this prior because there is no way for K to be negative. We also tested a minimally informative prior, $K \sim \text{gamma}(1,0.1)$. We used package *brms* in R (Bürkner 2021).

K predicted from semi-empirical equations

We predicted K_{600} from three semi-empirical equations, which are based on a combination of hydraulic parameters and empirically-derived coefficients. There are a vast array of these equations; we chose three that were reliable over a wide range of Q and had been successfully used in a reach-scale metabolism context (Cox 2003; Dodds et al. 2013). First, we used an equation from Isaacs and Gaudy Jr. (1968), the same used by Dodds et al. (2013). Of the models that use simple hydraulic parameters, this one performs the best (Cox 2003).

$$K_{600} = 4.75\left(\frac{v}{z^{1.5}}\right)1.0241^{(T-20)} \quad (3)$$

where z is average depth and T is temperature. We used 17°C for temperature, which gives us K_{600}, because K_{600} is simply K_{O_2} at 17°C.

Next, we applied a model from Parkhurst and Pomeroy (1972). This model is generally reliable over a wide range of Q and had the lowest standard error of prediction, compared to a suite of similar models (Cox 2003).

$$K_{600} = 23.04\frac{(1+0.17F^2)(Sv)^{0.375}}{z}1.0212^{(T-20)} \quad (4)$$

where S is the slope and F is the Froude number, calculated as $F = v/(\sqrt{gz})$, where g = acceleration due to gravity. Both S and F are unitless. As in (3), we used 17°C for T so that it would return K_{600}.

Finally, we applied a model from Dobbins (1964, 1965). Like (4), (5) was reliable over a wide range of Q (Cox 2003).

$$K_{600} = 1.7535\frac{C_A A_d E^{0.375}}{C_4^{1.5} z}\coth\frac{(2.751BE^{0.125})}{C_4^{0.5}} \quad (5)$$

where $C_A = 1.0 + F^2$, $A_d = 9.68 + 0.054(T - 20)$, $E = Svg$, $B = 0.9760 + 0.0137(30 - T)^{1.5}$, and $C_4 = 0.9 + F$. As above, we substituted 17°C for T.

As a final check on gas exchange, we collated estimates from large rivers and estuaries from Alin et al. (2011). We compiled all estimates that reported some measure of central tendency from their Table 5. We estimated the mean and uncertainty of this estimate by using a bootstrap. The mean and the median of all the empirical and calculated estimates was K_{600} = 0.5 d^{-1} (*see* "Results" section), and so we fixed K_{600} at 0.5. One problem with applying a single K_{600} value is that in reality, K probably varies day-to-day. However, our sensitivity analysis demonstrated that as long as K_{600} is between 0.2 and 1 d^{-1}, it has little effect on GPP and ER estimates, suggesting that even if K_{600} was not always exactly 0.5 d^{-1}, our GPP and ER estimates are nonetheless robust.

Metabolism estimates

We estimated metabolism with the one-station method, using inverse modeling. This method uses the following equation from Hall and Hotchkiss (2017):

$$\frac{\Delta DO}{\Delta t} = GPP + ER + (K \times (DO_{sat} - DO)) \quad (6)$$

where DO_{sat} is the concentration of DO at saturation, given the temperature and BP.

The estimates presented here were estimated via maximum likelihood, using streamMetabolizer (Appling et al. 2017), with K_{600} fixed at 0.5 d^{-1}. In addition to DO and temperature, maximum likelihood requires light data, which we estimated based on our site location, using the *calc light* function in the streamMetabolizer package in R version 3.4.4 (Appling et al. 2017). This function bases light values on the site location and time of year. We chose to estimate light in this way,

because point estimates, while effective at detecting cloud cover, may not apply to the full 85-km reach.

We also estimated metabolism via Bayesian inference, where we estimated K from the data. As discussed above, Bayesian inference returned unexpectedly high estimates of K_{600} and ER, and so we chose a fixed K based on other methods. Because we were only solving for 2 unknowns (GPP and ER), we did not risk equifinality with maximum likelihood inference (Appling et al. 2018*a*). Furthermore, models based on maximum likelihood run far faster than Bayesian methods. The Bayesian method does produce a full probability distribution for each day's GPP and ER, with inferences for within-day uncertainty, but we did not need those daily uncertainty estimates to address our research goals. For our site and data set, maximum likelihood with fixed K was the most reliable method.

Readers may wonder why we did not use a two-station model (Hall et al. 2016) to estimate metabolism given that the dam was close enough to the downstream site (85 km away) to influence DO concentrations. We lacked a suitable upstream DO sensor for which to conduct a two-station analysis. The TDG sensor may have approximated O_2 concentrations, but we doubt that this sensor is accurate for fine-scale O_2 dynamics. However, our one-station approach contains elements of a two station by correcting for upstream supersaturation.

Hydrodynamics

We used output from the MASS1 model for hydrodynamic information. MASS1 was calibrated and validated for the Hanford Reach (Niehus et al. 2014). In a one-station model, the reach length (i.e., upstream extent over which the DO data integrate) varies with velocity (v, m s^{-1}) and thus Q, and we defined that reach length, D_u, as equivalent to 80% turnover distance, $D_u = 1.6v/K$ (Chapra and Di Toro 1991), where 80% represents turnover of most O_2, but not as much as the long turnover distance for 95%. Using reach-scale average v from MASS1 (Richmond and Perkins 2009) and a mean K of 0.5 d^{-1} (*see* "Gas Exchange" section), we found that the reach length always exceeded the distance to the dam (> 85 km). Because the dam effectively resets gas concentrations and because we corrected our DO data for dam O_2 inputs, the reach length was the entire distance to the dam.

Although the Columbia River is wide in the Hanford Reach (average of 683 m), it has sufficient transverse mixing to reasonably assume that the DO data apply to the full river width. A modeling study of this reach found consistent lateral distribution of TDG (Richmond et al. 2000) and bottle incubations showed consistent photosynthesis rates across lateral transects (Neitzel et al. 1982). Similarly, field studies at other Columbia and lower Snake River dams found that TDG is generally well mixed across the river within a few kilometers of the dam (e.g., USACE 2002; Juul 2003), and those results are likely applicable here. As a result, our estimates of average depth, $\overline{z}$, include the full width of the river, extending upstream to the dam.

If our assumptions about transverse mixing are incorrect, we expect only a minor impact on metabolism estimates. Our metabolism estimates would apply only to the section of river with the DO sensor, and the DO correction for the dam would potentially be incorrect, if that section had a different velocity than the river average. However, metabolism occurs across the entire width of the river and the Columbia is rather homogeneous in its habitats—the river is constrained within a gorge and its entire width is composed of gravel and cobble. As a result, we do not expect much heterogeneity in metabolism, either, and if the river is not well mixed, effects on metabolism estimates are likely small.

To calculate average depth, we established transects every 0.5 km along the entire Hanford Reach. At each transect, we used MASS1 to estimate cross-sectional area, width, hydraulic depth, and hydraulic radius for a typical range of Q. MASS1 uses measured bathymetry at each transect, measured inflow at the upstream boundary, and water surface elevation at the downstream boundary to make its estimates at each transect (Niehus et al. 2014). For each value of Q, we calculated reach-average depth ($\overline{z}$) by taking the average hydraulic radius across all transects in the reach. We used those values to develop a $Q-\overline{z}$ relationship. Finally, we used MASS1 to estimate reach-scale Q every 15 min, based on inflow and water surface elevation data, and used the $Q-\overline{z}$ relationship to calculate $\overline{z}$.

Statistical inference of metabolism

To understand the fate of GPP and the relationship between GPP and ER, we estimated the fraction of GPP that is immediately respired (AR_f), which includes autotrophic respiration and some small fraction of heterotrophic respiration (Hall and Beaulieu 2013). We estimated AR_f with quantile regression at the 0.9 quantile, using the *quantreg* package in R.

Explanatory variables for GPP included temperature, light, and Q. We expressed temperature as $1/(K_B \times T)$, where K_B is the Boltzmann constant (8.61×10^{-5} eV K^{-1}) and T is temperature in Kelvin. By using $1/(K_B \times T)$, we can interpret the absolute value of the slope of log(GPP) and $1/(K_B \times T)$ (Welter et al. 2015) in the context of metabolic theory (Brown et al. 2004). We used daily total light, from the National Solar Radiation Data Base (https://nsrdb.nrel.gov/), in W m^{-2}, for the light predictor. We converted those values to total daily light, in kJ m^{-2} d^{-1}. We also used the light data to calculate light use efficiency (LUE) by converting GPP to kJ with a conversion of 1 g O_2 = 14 kJ (Kirk et al. 2021).

To estimate the covariates of GPP, we fit Generalized Least Squares (GLS) models to GPP, with $1/(K_B \times T)$, light, and Q as potential explanatory variables, using the using first-order autoregressive AR(1) correlation structure in the *nlme* package in R. GLS allows for autocorrelation in residuals, a common property in time series. All our explanatory variables were temporally autocorrelated and so we included an auto-correlation

parameter in each model. We fit each potential driver individually and in combination and compared AIC scores to rank the models. We also calculated the R^2_{lik}, a pseudo-r^2 developed for GLS models, using the *rr2* package in R (Ives 2019; Ives et al. 2019). We examined all residuals for homogeneity of variance, and the residuals from all models reported here met that assumption. All statistical tests and figures were completed in R version 4.2.1.

Results

Gas exchange

Measures of gas exchange based on Ar decline, predictions based upon models, and literature values all converged on K_{600} between 0.2 and 1 d^{-1} with a central tendency of 0.5 d^{-1}. Given a mean depth of 4.8 m, the gas exchange velocity k_{600} (m d^{-1}) therefore ranged from 1.0 to 4.8 m d^{-1}, with a mean near 1.9 m d^{-1} (Fig. 3). streamMetabolizer provided K_{600} values that averaged around 3 d^{-1} ($k_{600} = 14$ m d^{-1}), which we found unlikely and provided untenable estimates of metabolism.

Estimates of metabolism were insensitive to *K*

Each method of estimating *K* gave a distinct range of estimates, but many overlapped, and all were <1 (Fig. 3), except for those from the inverse modeling approach (not shown in figure). Parkhurst and Pomeroy (1972) estimates were lower than our empirical estimates, while the other two were higher. The mean and median value was 0.5 d^{-1}.

Estimates of GPP were insensitive to our range in uncertainty in gas exchange; at $K_{600} = 0.2$, average GPP was 11.4 g O_2 m^{-2} d^{-1} and at $K_{600} = 3$, average GPP was 11.8 g O_2 m^{-2} d^{-1}, a difference of 3.8% (Table 1). The median daily difference of GPP between $K_{600} = 0.2$ and $K_{600} = 3$ was 0.8 g O_2 m^{-2} d^{-1} and the median daily difference between $K_{600} = 0.2$ and $K_{600} = 1.0$ was 0.2 g O_2 m^{-2} d^{-1} (Fig. 4). As a result, our GPP estimates are highly robust. ER estimates were more sensitive to K_{600}, but especially when $K_{600} > 1$. From $K_{600} = 0.2$ to $K_{600} = 3$, the difference in ER was 59%, but from $K_{600} = 0.2$ to $K_{600} = 1$, the difference in ER was only 9%. The median daily difference between $K_{600} = 0.2$ and $K_{600} = 3$ was 5.2 g O_2 m^{-2} d^{-1}, but the median daily difference between $K_{600} = 0.2$ and $K_{600} = 1.0$ was 1.7 g O_2 m^{-2} d^{-1} (Fig. 4).

The selection of K_{600} also influenced the number of nonsensical ER and GPP estimates. We defined nonsensical estimates as ER > 0 g O_2 m^{-2} d^{-1} and GPP < 0 g O_2 m^{-2} d^{-1}. For ER, there were 5 or fewer nonsensical (i.e., positive) estimates when $0.2 \le K_{600} \le 1.0$. At $K_{600} = 3.0$, 25% of ER estimates were nonsensical. For GPP, $<2\%$ of estimates were nonsensical (i.e., negative) for any value of K_{600}, although there were more nonsensical values at the lowest values of K_{600} (Table 1).

Metabolism patterns

GPP ranged from 0.3 to 30.8 g O_2 m^{-2} d^{-1}, with a median value of 10.5 g O_2 m^{-2} d^{-1}. In general, the lowest GPP occurred

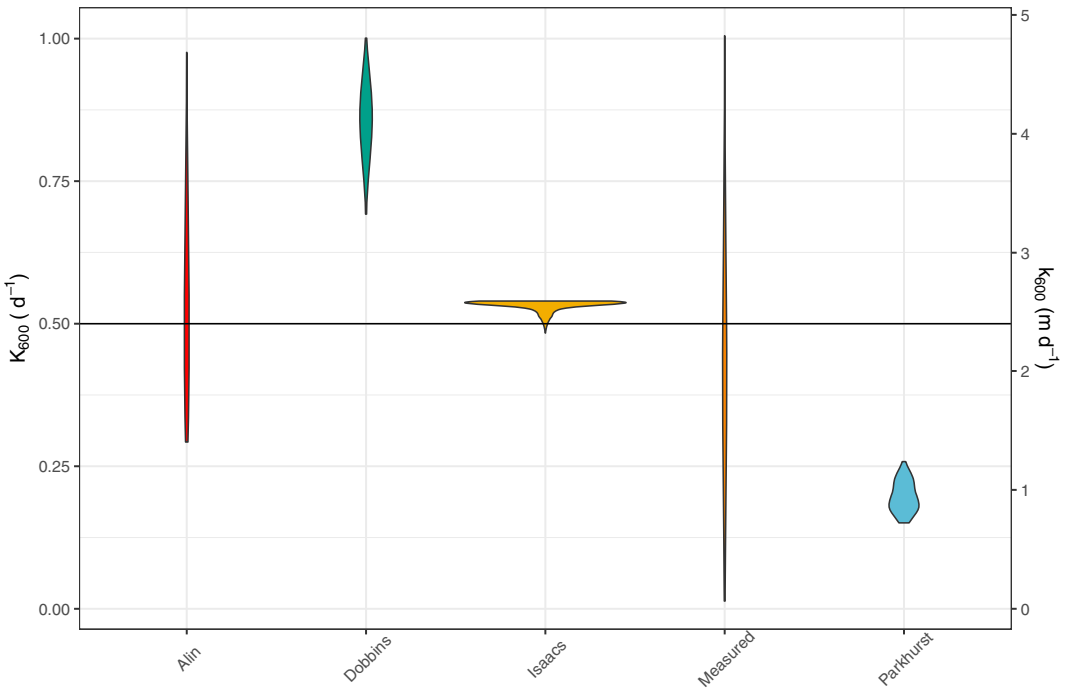

Fig. 3. K_{600} values produced with three semi-empirical models (Dobbins, Isaacs, and Parkhurst), along with a bootstrapped estimate from our measured values, and a bootstrapped range of estimates from large rivers and estuaries (Alin et al. 2011). The solid line is the value of K_{600} used in our metabolism estimates.

Table 1. Sensitivity of GPP and ER to a range of K_{600}. Mean GPP and mean ER refer to the mean of all successfully modeled days. GPP Removed and ER Removed refer to the number of estimates removed because they yielded non-sensical estimates (i.e., negative values of GPP and positive values of ER).

K_{600} (d^{-1})	Mean GPP (g O_2 m^{-2} d^{-1})	Mean ER (g O_2 m^{-2} d^{-1})	GPP removed	ER removed
0.2	11.38	−11.85	15	11
0.4	11.39	−11.71	6	5
0.5	11.40	−11.60	5	5
0.6	11.40	−11.47	4	5
0.8	11.42	−11.14	2	3
1.0	11.44	−10.74	2	3
1.5	11.50	−9.51	2	5
2.0	11.58	−8.05	2	24
2.5	11.69	−6.47	2	102
3.0	11.84	−4.84	2	213

in winter, with a gradual increase in spring. GPP rose through summer, peaking in August or September (Fig. 5). ER ranged from −0.8 to −30.6 g O_2 m^{-2} d^{-1}, with a median value of −10.3 g O_2 m^{-2} d^{-1} (Fig. 5). The seasonal patterns in ER mirrored those of GPP (Fig. 5). Indeed, GPP almost perfectly predicted ER ($|\mathrm{ER}| = 0.87\,\mathrm{GPP} + 0.94$, $r^2 = 0.92$, SE of slope = 0.01).

GPP covaried with $1/(K_B \times T)$ and light, but not with Q. Peak GPP occurred in late summer and early fall, when temperatures were highest (Fig. 6A). GPP increased exponentially with $1/(K_B \times T)$ and the absolute value of the slope of log (GPP) and $1/(K_B \times T)$ was 0.42. GPP and light fit a power law model with a slope of 0.25, which indicates a loss in efficiency as light increases. In other words, at low light, GPP increased rapidly, but at moderate to high light, the rate of increase attenuated (Fig. 6B). As a result, maximum GPP did not occur at maximum light; it occurred at maximum temperature (Fig. 6B). Consistent with this, the power law relationship between GPP and light suggested that LUE declined as light increased. Indeed, mean LUE was highest in autumn (2.4 ± 2.5 SD), followed by winter (2.1 ± 2.1), then summer (1.0 ± 0.5 SD), and lowest in spring (0.77 ± 0.3 SD). The best model, according to AIC and R^2 values, included both light and $1/(K_B \times T)$ (Table 2). Q was not an effective predictor of GPP (Fig. 6C).

NEP was nearly 0, reflecting the close relationship between ER and GPP. It was negative on about half the dates (58%) and averaged −0.22 g O_2 m^{-2} d^{-1} (Fig. 6D). On an annual scale, the Hanford Reach was slightly heterotrophic, but autotrophy occurred about half of the time, primarily in spring and summer, occasionally in autumn, and almost never in the winter. Quantile regression (Hall and Beaulieu 2013) revealed that 90% of GPP was respired immediately (within 1 d, Fig. 6D), meaning that even though the reach may be semi-

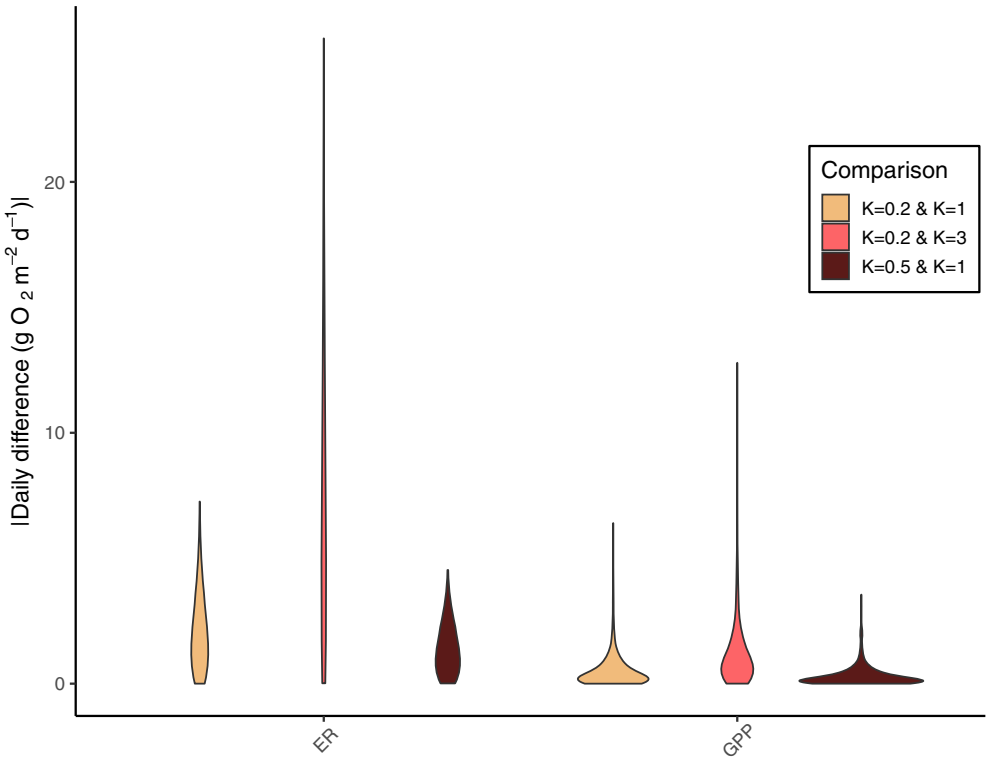

Fig. 4. Absolute value of the daily differences in GPP and ER, when estimated with various values of K_{600}.

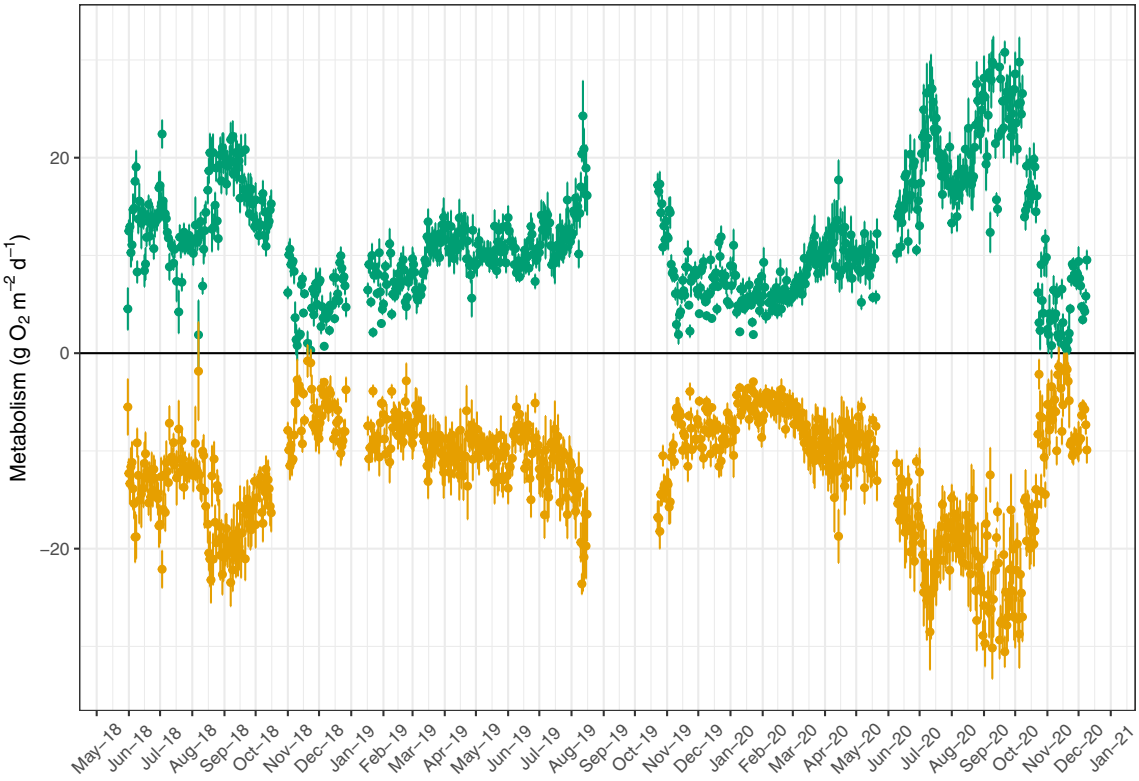

Fig. 5. GPP and ER in the Hanford Reach. GPP is in teal and ER is in orange.

frequently autotrophic, it is not likely to accumulate or export much organic C derived from photosynthesis.

Discussion

Estimating metabolism in large rivers

A key challenge with estimating metabolism in large rivers is attaining reasonable estimates of K_{600}. In our case, metabolism estimates were relatively insensitive to the range of estimated K. This relative insensitivity is likely a function of having low K, relative to other rivers (Ulseth et al. 2019). Why was K so small in the Columbia River and why did K not change with Q, as typically expected (Appling et al. 2018*b*)? The Columbia River is deep, and so a relatively small proportion of its water interacts with the atmosphere, limiting exchange. As Q increases, turbulence may increase (increasing K), but so does depth (decreasing K). These two changes likely cancel each other out, similar to the Klamath River where increased Q caused idiosyncratic changes to K (Genzoli and Hall 2016). We note, however, that despite our K estimates being on the low end, they did vary fivefold (from 0.2 to 1.0 d^{-1}). If we were estimating K for the purpose of estimating gas fluxes, such as CO_2 or CH_4, this error would be problematic, because an uncertainty on gas exchange scales 1:1 with fluxes, resulting in a five-fold uncertainty.

Estimates of K_{600} from the DO data themselves, using streamMetabolizer, were likely far too high and provided nonsensical estimates of ER. The use of modeling approaches such as streamMetabolizer to estimate metabolism and gas exchange concomitantly has been hugely helpful in estimating riverine metabolism where it is not possible to measure gas exchange directly, for example, in hundreds of rivers where the USGS monitors DO (Appling et al. 2018*b*). Such statistical tools have allowed large-scale syntheses of annual metabolism (Bernhardt et al. 2022) that were heretofore not possible. But our findings in the Columbia River demonstrate that deriving K_{600} from O_2 models may not always be correct, even when K_{600} is low and GPP is high, two scenarios that should, in theory, promote the use of modeling approaches for K_{600} (Hall and Ulseth 2019). We suggest that when investigators wish to generate high-confidence estimates of metabolism time series or gas emission for any particular river they should question the use of any single method for gas exchange and instead try multiple approaches to increase confidence in the final estimates of gas exchange and, therefore, metabolism.

Discontinuities, like dams, are prevalent in large rivers, presenting both benefits and problems. Dams function as a continuous tracer gas injection when their spillways result in supersaturation, allowing researchers to estimate K_{600} empirically, as we have done here. They also reset gas concentrations

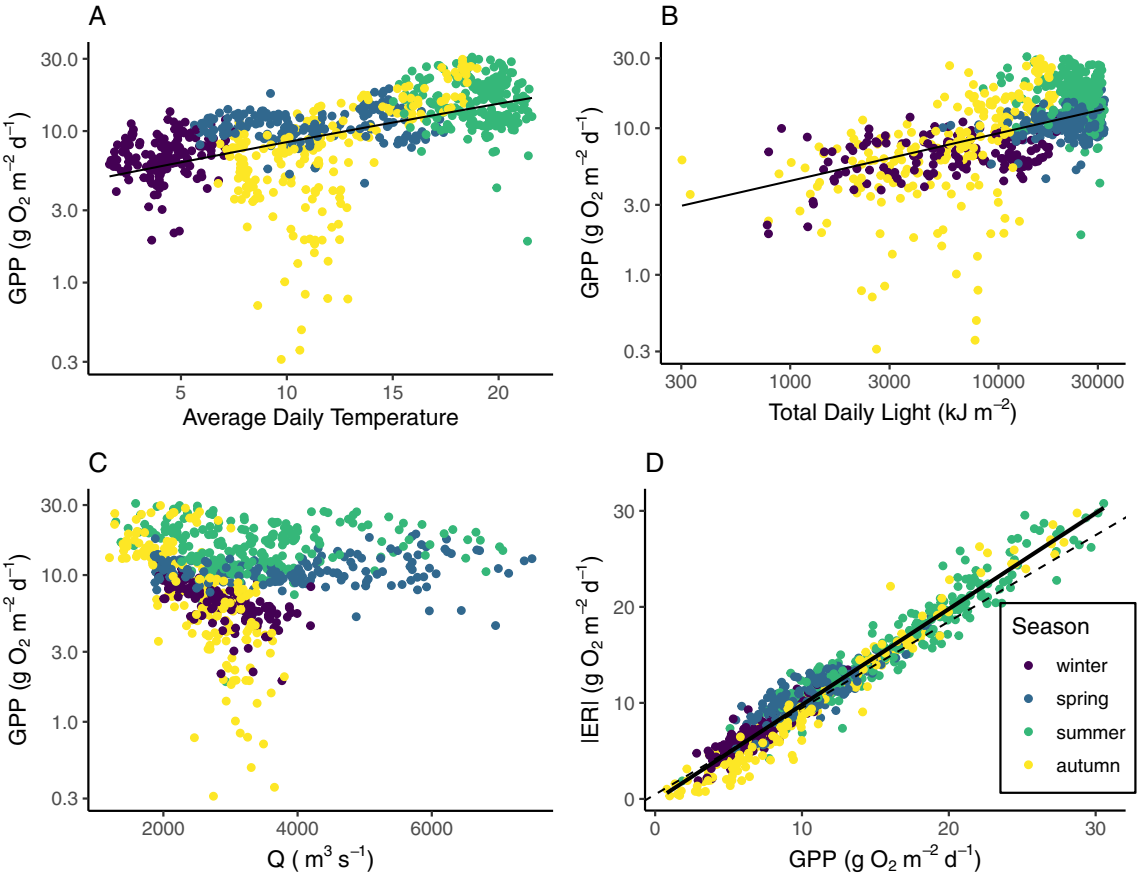

Fig. 6. Drivers of metabolism in the Hanford Reach: (**A**) average daily temperature, expressed in °C; (**B**) total daily light, expressed in kJ m^{-2}; (**C**) discharge (Q); and (**D**) the relationship between ER and GPP. Note that (**A**) and (**B**) are plotted on a log scale. The fitted lines are those reported in Table 2, and were fit via gls with a corAR1 autocorrelation structure, except (**D**), which was fit with a linear model. The dashed line in (**D**) shows the fraction of GPP that is immediately respired, fit with quantile regression at the 0.9 quantile.

and define the reach, by providing an upstream boundary. At the same time, supersaturation of river gases can obscure DO patterns. We found that applying a DO correction alleviated this problem. The potential risk with DO correction is misattribution of the source of supersaturation. We are confident that the extra DO was a result of spillway operations, because of the tight relationship between supersaturation and spillway flow. An alternative explanation is that the reach is highly autotrophic, with GPP exceeding ER to such a high degree that the combination of nighttime respiration and gas exchange is insufficient to return the river to equilibrium. There are physiological constraints to such an explanation, however; autotrophs must respire. As a result, a completely biological explanation for continuous oversaturation is

Table 2. Coefficients, standard error, AIC, and pseudo-r^2 for models that predict log (GPP). K_B is the Boltzmann constant, in eV, and T is temperature in Kelvin.

Model	Variable	Coefficient	SE	AIC	R^2
$\log(\text{GPP}) = 1/(K_B \times T)$	$1/(K_B \times T)$	−0.42	0.06	572	0.68
	Intercept	19.24	2.30		
$\log(\text{GPP}) = \log(\text{light})$	Light	0.33	0.02	450	0.73
	Intercept	−0.77	0.23		
$\log(\text{GPP}) = 1/(K_B \times T) + \log(\text{light})$	Light	0.30	0.02	441	0.73
	$1/(K_B \times T)$	−0.22	0.05		
	Intercept	8.55	2.26		

possible, but would be limited to places and times where gas exchange was low and the biota were respiring a relatively small proportion of GPP. This scenario may have occurred on a few dates in the Hanford Reach: even after DO correction, there were 107 d in which DO remained supersaturated all day. In most cases, these days were very close to equilibrium, with daily minimum DO saturation between 100% and 101%. Nonetheless, this finding suggests that metabolism may contribute in a small way to supersaturated DO concentrations if NEP is positive and K_{600} is low.

Dams also create unsteady flow conditions, which violate the assumptions of one-station metabolism. Others have dealt with unsteady flow by coupling a two-station metabolism model with a flow routing model (Pathak and Demars 2023) or with a Eulerian model (Payn et al. 2017). Although we used one-station metabolism, our DO corrections accounted for the lag time between the upstream boundary and the DO sensor. The lag time was calculated for each time point in our data (every 15 min). In that sense, we account for variable flow. Our approach is less computationally intense, but required an existing, detailed reach-scale hydrologic model that allowed us to obtain high-confidence reach-scale velocity. In contrast to our site, Bayesian inference produced reasonable K_{600} estimates in the Colorado River, when using informed priors (Payn et al. 2017). Similar to our site, constant K_{600}, based on empirical data, worked well for the River Otra (Pathak and Demars 2023).

Finally, reach-scale average depth and velocity are difficult to obtain in large rivers. Errors in depth scale 1 : 1 with errors in metabolism, so good depth estimates are critical to accurate metabolism estimates. Hanford Reach hydrodynamics and channel dimensions have been well characterized (Perkins et al. 2004; Richmond and Perkins 2009; Niehus et al. 2014), allowing us to apply existing models. Such efforts also provided appropriate data for the semi-empirical K_{600} calculations. Other large river reaches may not have similar existing hydrological information. In summary, estimating metabolism in large rivers requires additional hydrological and geomorphological considerations, along with a healthy dose of skepticism: simply placing a DO and temperature sensor close to a USGS gauge and hoping that streamMetabolizer works is insufficient. Prior to embarking on metabolism estimates in large rivers, we suggest researchers consider the feasibility of depth measurements, options for estimating K_{600}, discontinuities with the turnover distance of DO, transverse mixing, variable flow, and other unknown unknowns.

Metabolism patterns

Annual GPP and ER exceeded most estimates (Bernhardt et al. 2022) showing that the Columbia River is one of the most metabolically active rivers ever measured. Assuming a 1:1 M ratio of O_2 : C, average annual metabolism on the 2-yr time series was GPP = 1530 g C m^{-2} yr^{-1} and ER = 1500 g C m^{-2} yr^{-1}. These values are in the upper 2% of GPP and 3% for ER of the 222 rivers in Bernhardt et al. (2022). The median values of GPP and ER exceeded all GPP and ER values estimated in the lower Mississippi River (Dodds et al. 2013) and the Colorado River (Hall et al. 2015). The generally high rates of GPP and ER are likely due to the low turbidity of the Columbia River, along with its deep depth: there is more water and resultant metabolic activity per m^2 in the Columbia compared to most rivers. There is little basis to compare our estimates here with other findings in the Columbia River. Weekly measurements of primary production via ^{14}C from 1977 to 1980 were lower than our measurements, with a maximum value of 6 g O_2 m^{-2} d^{-1} (Neitzel et al. 1982), compared to our maximum GPP of 30 g O_2 m^{-2} d^{-1} and median GPP of 10.5 g O_2 m^{-2} d^{-1}. The lower values from the ^{14}C methods could reflect changes in production through time, bottle effects, the exclusion of benthic GPP, or the fact that ^{14}C always measures something less than gross production (Wetzel and Likens 2010).

The seasonal metabolism patterns in the Hanford Reach were distinct from those in most rivers, which typically exhibit either a spring peak or a summer peak in GPP (Savoy et al. 2019), although there are exceptions (Bernhardt et al. 2022). In the Hanford Reach, GPP and ER peaked during August and September of 2018 and during September of 2020. Data are missing from August to October 2019, but the highest rates occurred just before data loss, in early August. These patterns are consistent with seasonal measurements of ^{14}C uptake in the Hanford Reach, in which C uptake peaked in summer and fall (Neitzel et al. 1982). This early fall peak corresponds with peak temperature, but after peak light availability.

There appears to be a hysteretic pattern with light and GPP. Spring and fall have the same range of total daily light, but fall has far higher GPP and higher LUE. We hypothesize that this delay in response to light occurs because the upstream reservoir drives metabolism patterns in the Hanford Reach. As temperatures increase, so do phytoplankton populations. A larger algal population can fix more C, in aggregate, than will a smaller population, even if the photosynthesis rates of individual cells are not as high as during times of peak light. As a result, GPP increased with temperature. The travel time in the Hanford Reach is generally < 1 d, so if most of the photosynthesis is planktonic, it's unlikely that the population is increasing appreciably within the Hanford Reach unless the plankton have a doubling time < 1 d. Most likely, the planktonic population behind the Priest Rapids dam increases throughout the summer, peaking in early fall when temperatures are maximal. The phytoplankton population in the Hanford Reach likely reflects the dynamics behind the dam; indeed, phytoplankton density increased in the Hanford Reach after construction of the Priest Rapids Dam (Neitzel et al. 1982). In addition, it is possible that submerged plants and/or benthic algae reach peak biomass and metabolism in the fall, at the end of the growing season. Given the 4.8 m average water depth, we expect that most metabolism, expressed areally, is planktonic.

When LUE is calculated for streams and rivers from open-sky irradiance, as we did here, the mean value is 0.5%, with a range up to 3% (Kirk et al. 2021). We observed an overall mean LUE of 1.5%, with the autumn mean of 2.4% and some individual dates > 10%. This high LUE is somewhat surprising, given the depth of the Columbia River and the resultant light attenuation. It further suggests that much of the GPP is planktonic, rather than benthic. The seasonal variation in LUE potentially reflects both biological and physical drivers. In winter, light levels are quite low and primary producers thus must use more of the available light and they may use various physiological adaptations to do so (e.g., Felip and Catalan 2000). In autumn, light is similar to spring, yet LUE was much higher. We suggest this higher LUE is a result of higher algal biomass; i.e., the interaction between light and temperature.

Light and $1/(K_B \times T)$ explained a large fraction of variation in GPP. The power law model relating light to GPP had a slope < 1 (0.3), meaning there was a loss in efficiency as light increased, consistent with the seasonal changes in LUE. Metabolic theory predicts an exponential relationship between GPP and temperature, with the absolute value of the slope of log(GPP) and $1/(K_B \times T)$ representing the activation energy. Activation energy for GPP at the cellular level is assumed to be 0.32 eV (Allen et al. 2005) and we calculated a similar slope (0.4 when $1/(K_B \times T)$ was modeled alone and 0.2 when modeled with light). Although remarkably similar to theory, this alignment may be coincidental. Ecosystem-level activation energy can be amplified by biomass accrual and dampened by other limiting factors, such as nutrients or light (Welter et al. 2015). The interacting effects of light, temperature, and primary producer population size may have resulted in model coefficients that aligned with theory. Overall, the seasonal patterns and the model coefficients suggest that light has a stronger effect on GPP during winter, early spring, and late fall, while temperature was the primary driver during spring, summer, and early fall, when light is saturating.

Metabolism in large lowland rivers

Other large rivers with continuous open-channel metabolism estimates include the Mississippi, Loire, Thames, and Colorado Rivers (Dodds et al. 2013; Hall et al. 2015; Appling et al. 2018*b*; Diamond et al. 2022; Pathak et al. 2022). Of these, the Mississippi River is most similar in size to the Columbia and it has two attempts at estimating metabolism. One is Dodds et al. (2013), who estimated metabolism at weekly intervals in the Mississippi River near Baton Rouge, Louisiana. They fit models to data, but with a fixed K calculated from Isaacs and Gaudy Jr. (1968). This site had a benefit of no flow obstructions for > 1000 km upstream, meaning it likely met the assumptions of one-station, open-channel metabolism. The second estimate was derived from data in Appling et al. (2018*b*) at the Mississippi River farther upstream, at Clinton, Iowa. This site had periods of time with very high GPP (> 15 g O_2 m^{-2} d^{-1}) and low ER (0–5 g O_2 m^{-2} d^{-1}), which is thermodynamically impossible because primary producers must respire (Hall and Beaulieu 2013). This result shows that at best some of the assumptions of one-station metabolism were unmet, perhaps because of the extensive lock and dam infrastructure along this reach of the Mississippi. The Clinton site modeled gas exchange as a free parameter, but values of K_{600} were low and tightly constrained at 0.5–1 d^{-1}, close to the value for the Columbia River. In combination with our observations in the Columbia, these two examples show that in large rivers, K_{600} can be low. The example from Iowa also shows that assumptions of one-station metabolism can be easily violated.

Columbia River metabolism magnitude and patterns contrast with those in other large rivers. The Mississippi and Colorado River had substantially lower GPP and ER (Mississippi River GPP < 4 g O_2 m^{-2} d^{-1}, |ER| < 7 g O_2 m^{-2} d^{-1}; Colorado River GPP ranged from 0 to 3 g O_2 m^{-2} d^{-1}, with a mean of 0.8, Hall et al. 2015; Dodds et al. 2013; Appling et al. 2018*b*). In the Mississippi River, |ER| increased with Q, likely because higher Q resulted in more allochthonous inputs (Dodds et al. 2013). Turbidity controlled GPP via its effects on light transmission in the Colorado River (Hall et al. 2015) and turbidity may have controlled Mississippi River GPP, as well (Dodds et al. 2013). In both cases, temperature was of secondary importance or had minimal predictive power. Both rivers are more turbid than the Columbia River. In the Loire River, temperature had a larger effect on GPP than light, but the abiotic drivers of metabolism depended somewhat on ecosystem state (i.e., phytoplankton- or macrophyte-driven, Diamond et al. 2022). The Thames had a summer peak in productivity when light and temperature were high (Pathak et al. 2022). A recent synthesis predicts that a non-turbid river without seasonal canopy cover (such as the Columbia) will have maximum GPP during the time of maximum incident light (Bernhardt et al. 2022). The Columbia River deviates from this broad pattern, but so does the Mississippi River, albeit in a different way. We do not yet understand the governing mechanisms for metabolism in truly large rivers and indeed, those governing mechanisms may be site-specific.

Autotrophy and coupled respiration

As a whole, streams and rivers are rarely autotrophic (Hoellein et al. 2013; Bernhardt et al. 2022), because they receive terrestrial inputs along their length, and those inputs fuel respiration. The Hanford Reach deviates from this general pattern: because it is just barely heterotrophic on an annual scale, GPP exceeds ER on nearly half of the days, and most of this GPP is immediately respired. Autochthonous production thus fuels this ecosystem. This deviation from typical streams likely occurs because the Hanford Reach is effectively disconnected from its watershed. The upstream boundary is a dam, which accumulates organic matter and keeps it out of the Hanford Reach. In addition, the Hanford Reach has no major

tributary inputs, because the surrounding landscape receives just 12–25 cm of rain per year. As a result, allochthonous inputs in this reach are restricted to those from irrigation return flows and from vegetation immediately adjacent to the river. Finally, substrate within the Hanford Reach comprises cobble, gravel, and boulders, and so likely has minimal C stored within the channel to cause temporal lags in respiration.

The lower Mississippi River was strongly heterotrophic on an annual scale, with a few dates of autotrophy and a weaker relationship between GPP and ER (Dodds et al. 2013). Similarly, the Amazon River appears to be strongly heterotrophic (Quay et al. 1995), but the Loire River exhibited heterotrophy only after shifting from a planktonic-dominated metabolism to a benthic-dominated ecosystem (Diamond et al. 2022). The geological context of these rivers is distinct from the Columbia, however. The Klamath River exhibited similar patterns to the Columbia, with AR_f as high as 0.92. Autochthony in the Klamath was attributed, in part, to downstream export of organic C (Genzoli and Hall 2016). The Klamath, although much smaller than the Columbia, shares some key characteristics: it has large dams and flows through a bedrock canyon, resulting in minimal allochthonous inputs during summer baseflow.

In addition to geology, the food web may influence the carbon balance of rivers; specifically, plankton-dominated rivers appear more likely to demonstrate autotrophy than benthic- or plant-dominated food webs. The Murray River, where most GPP is planktonic, has a tight GPP : ER relationship, with NEP close to 0 (Oliver and Merrick 2006), similar to the Columbia. The Loire River was autotrophic on an annual scale when it was dominated by phytoplankton, but heterotrophic when the river shifted to plant dominance (Diamond et al. 2022). The Thames and Seine, also plankton-dominated, are seasonally autotrophic, but are heterotrophic on an annual scale, likely because they receive plentiful anthropogenic allochthonous carbon inputs (Escoffier et al. 2018; Pathak et al. 2022).

Conclusion

So far, truly large rivers appear to be idiosyncratic, with their metabolism patterns influenced by site-specific conditions, such as geology, food web, dams, climate, and surrounding land use. As more large river reaches are studied, perhaps distinct regimes will appear. Given that only 36 rivers are larger than the Columbia, it is also possible that each has its own metabolic pattern and generalization will elude us. Based on our experience with the Hanford Reach of the Columbia River, metabolism in large rivers may require correcting DO data for dam O_2 inputs and creative approaches to estimating K. Our metabolism estimates revealed notably high rates of GPP and ER, with GPP driven primarily by temperature and secondarily by light, while ER was almost solely a function of daily GPP. These patterns contrast with those in small–medium streams and suggest that the world's largest rivers require specific attention when studying the carbon cycle of river networks.

Data availability statement

All data and R code associated with this paper are published on ESS-DIVE (Roley et al. 2023), available at https://data.ess-dive.lbl.gov/datasets/doi:10.15485/1985922.

References

Alin, S., F. M. d. Fátima, C. Salimon, J. E. Richey, G. W. Holtgrieve, A. V. Krusche, and A. Snidvongs. 2011. Physical controls on carbon dioxide transfer velocity and flux in low-gradient river systems and implications for regional carbon budgets. J. Geophys. Res. **116**: G01009.

Allen, A., J. Gillooly, and J. Brown. 2005. Linking the global carbon cycle to individual metabolism. Funct. Ecol. **19**: 202–213.

Appling, A. P., R. O. Hall, M. Arroita, and C. B. Yackulic. 2017. streamMetabolizer: Models for estimating aquatic photosynthesis and respiration.

Appling, A. P., R. O. Hall, C. B. Yackulic, and M. Arroita. 2018*a*. Overcoming equifinality: Leveraging long time series for stream metabolism estimation. Eur. J. Vasc. Endovasc. Surg. **123**: 624–645.

Appling, A. P., and others. 2018*b*. The metabolic regimes of 356 rivers in the United States. Sci. Data **5**: 180292.

Behnke, M. I., and others. 2023. Aquatic biomass is a major source to particulate organic matter export in large arctic rivers. Proc. Natl. Acad. Sci. USA **120**: e2209883120.

Bernhardt, E. S., and others. 2022. Light and flow regimes regulate the metabolism of rivers. Proc. Natl. Acad. Sci. USA **119**: e2121976119.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology **85**: 1771–1789.

Bürkner, P.-C. 2021. Bayesian item response modeling in R with brms and Stan. J. Stat. Softw. **100**: 1–54.

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.

Cox, B. 2003. A review of dissolved oxygen modelling techniques for lowland rivers. Sci. Total Environ. **314**: 303–334.

Deemer, B. R., C. B. Yackulic, R. O. Hall Jr., M. J. Dodrill, T. A. Kennedy, J. D. Muehlbauer, D. J. Topping, N. Voichick, and M. D. Yard. 2022. Experimental reductions in sub-daily flow fluctuations increased gross primary productivity for 425 river kilometers downstream. PNAS Nexus **1**: pgac094.

Diamond, J. S., F. Moatar, M. J. Cohen, A. Poirel, C. Martinet, A. Maire, and G. Pinay. 2022. Metabolic regime shifts and ecosystem state changes are decoupled in a large river. Limnol. Oceangr. **67**: S54–S70.

Dobbins, W. E. 1964. BOD and oxygen relationships in streams. J. Sanit. Eng. Div. **90**: 53–78.

Dobbins, W. E. 1965. Closure to "BOD and oxygen relationship in streams". J. Sanit. Eng. Div. **91**: 49–55.

Dodds, W., A. M. Veach, C. M. Ruffing, D. M. Larson, J. L. Fischer, and K. H. Costigan. 2013. Abiotic controls and temporal variability of river metabolism: Multiyear analyses of Mississippi and Chattahoochee River data. Freshw. Sci. **32**: 1073–1087.

DOE. 2022. Hanford annual site environmental report for calendar year 2021. US Department of Energy, Richland Operations Office.

Escoffier, N., N. Bensoussan, L. Vilmin, N. Flipo, V. Rocher, A. David, F. Métivier, and A. Groleau. 2018. Estimating ecosystem metabolism from continuous multi-sensor measurements in the Seine River. Environ. Sci. Pollut. Res. **25**: 23451–23467.

Felip, M., and J. Catalan. 2000. The relationship between phytoplankton biovolume and chlorophyll in a deep oligotrophic lake: Decoupling in their spatial and temporal maxima. J. Plankton Res. **22**: 91–106.

Fritz, B., N. Kohn, T. Gilmore, D. McFarland, E. Arntzen, R. Mackely, G. Patton, D. Mendoza, and A. Bunn. 2007. Investigation of the hyporheic zone at the 300 area, Hanford site. Technical report, Pacific Northwest National Lab. (PNNL).

Fritz, B. G., and E. V. Arntzen. 2007. Effect of rapidly changing river stage on uranium flux through the hyporheic zone. Groundwater **45**: 753–760.

Genzoli, L., and R. O. Hall. 2016. Shifts in Klamath River metabolism following a reservoir cyanobacterial bloom. Freshw. Sci. **35**: 795–809.

Hall, R. O., and J. J. Beaulieu. 2013. Estimating autotrophic respiration in streams using daily metabolism data. Freshw. Sci. **32**: 507–516.

Hall, R. O., C. B. Yackulic, T. A. Kennedy, M. D. Yard, E. J. Rosi-Marshall, N. Voichick, and K. E. Behn. 2015. Turbidity, light, temperature, and hydropeaking control primary productivity in the Colorado River, Grand Canyon. Limnol. Oceangr. **60**: 512–526.

Hall, R. O., J. L. Tank, M. A. Baker, E. J. Rosi-Marshall, and E. R. Hotchkiss. 2016. Metabolism, gas exchange, and carbon spiraling in rivers. Ecosystems **19**: 73–86.

Hall, R. O., and R. Hotchkiss. 2017. Stream metabolism, p. 219–234. *In* G. A. Lamberti and F. R. Hauer [eds.], Methods in stream ecology. Academic Press.

Hall, R. O., and H. L. Madinger. 2018. Use of argon to measure gas exchange in turbulent mountain streams. Biogeosciences **15**: 3085–3092.

Hall, R. O., and A. J. Ulseth. 2019. Gas exchange in streams and rivers. Wiley Interdiscip. Rev. Water **7**: e1391.

Ho, D. T., P. Schlosser, and P. M. Orton. 2011. On factors controlling air–water gas exchange in a large tidal river. Estuar. Coast. **34**: 1103–1116.

Hoellein, T. J., D. A. Bruesewitz, and D. C. Richardson. 2013. Revisiting Odum (1956): A synthesis of aquatic ecosystem metabolism. Limnol. Oceangr. **58**: 2089–2100.

Isaacs, W. P., and A. F. Gaudy Jr. 1968. Atmospheric oxygenation in a simulated stream. J. Sanit. Eng. Div. **94**: 319–344.

Ives, A., D. Li, and M. A. Ives. 2019. Package "rr2": R^2s for regression models. doi:10.1093/sysbio/syy060

Ives, A. R. 2019. R^2s for correlated data: Phylogenetic models, lmms, and glmms. Syst. Biol. **68**: 234–251.

Juul, S. 2003. An assessment of selected water quality parameters for the priest rapids hydroelectric project. Technical report, Public Utility District No. 2 of Grant County.

Kana, T., C. Darkangelo, M. Hunt, J. Oldham, G. Bennett, and J. Cornwell. 1994. Membrane inlet mass spectrometer for rapid high-precision determination of N_2, O_2, and Ar in environmental water samples. Anal. Chem. **66**: 4166–4170.

Keith, D. A., and others. 2022. A function-based typology for Earth's ecosystems. Nature **610**: 513–518.

Kirk, L., R. T. Hensley, P. Savoy, J. B. Heffernan, and M. J. Cohen. 2021. Estimating benthic light regimes improves predictions of primary production and constrains light-use efficiency in streams and rivers. Ecosystems **24**: 825–839.

Meyer, J. L., and R. T. Edwards. 1990. Ecosystem metabolism and turnover of organic carbon along a blackwater river continuum. Ecology **71**: 668–677.

Neitzel, D. A., T. L. Page, and R. W. H. Jr. 1982. Mid-Columbia River microflora. J. Freshwater Ecol. **1**: 495–505.

Niehus, S., W. Perkins, and M. Richmond. 2014. Simulation of Columbia River hydrodynamics and water temperature from 1917 through 2011 in the Hanford Reach. Final Report PNWD-3278, Battelle-Pacific Northwest Division, Richland, Washington, 99352.

Odum, H. T. 1956. Primary production in flowing waters. Limnol. Oceangr. **1**: 102–117.

Oliver, R. L., and C. J. Merrick. 2006. Partitioning of river metabolism identifies phytoplankton as a major contributor in the regulated Murray River (Australia). Freshwater Biol. **51**: 1131–1148.

Parkhurst, J. D., and R. D. Pomeroy. 1972. Oxygen absorption in streams. J. Sanit. Eng. Div. **98**: 101–124.

Pathak, D., and others. 2022. High-resolution water-quality and ecosystem-metabolism modeling in lowland rivers. Limnol. Oceangr. **67**: 1313–1327.

Pathak, D., and B. O. L. Demars. 2023. Metabolism modeling in rivers with unsteady flow conditions and transient storage zones. Eur. J. Vasc. Endovasc. Surg. **128**: e2022JG007245.

Payn, R. A., R. O. Hall, T. A. Kennedy, G. 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.

Perkins, W., M. Richmond, and G. McMichael. 2004. Two-dimensional modeling of time-varying hydrodynamics and juvenile Chinook salmon habitat in the Hanford reach of

the Columbia River. *In* G. Sehlke, D. F. Hayes, and D. K. Stevens [eds.], Critical transitions in water and environmental resources management. Proceedings of the 2004 World Water and Environmental Resources Congress, June 27–July 1, ASCE (American Society of Civil Engineers), Salt Lake City, Utah.

Poston, T., J. Duncan, and R. Dirkes. 2009. Hanford annual site environmental report for calendar year 2008. Pacific Northwest National Laboratory.

Quay, P. D., D. Wilbur, J. E. Richey, A. H. Devol, R. Benner, and B. R. Forsberg. 1995. The ^{18}O:^{16}O of dissolved oxygen in rivers and lakes in the Amazon Basin: Determining the ratio of respiration to photosynthesis rates in freshwaters. Limnol. Oceangr. **40**: 718–729.

Richmond, M. C., W. A. Perkins, and Y. Chien. 2000. Numerical model analysis of system-wide dissolved gas abatement alternatives. Technical report, Pacific Northwest National Lab (PNNL).

Richmond, M. C., and W. A. Perkins. 2009. Efficient calculation of dewatered and entrapped areas using hydrodynamic modeling and GIS. Environ. Model. Software **24**: 1447–1456.

Roley, S. S., R. O. Hall, V. A. Garayburu-Caruso, W. A. Perkins, and J. C. Stegen. 2023. Data and scripts associated with "Coupled primary production and respiration in a large river contrasts with smaller rivers and streams". River Corridor and Watershed Biogeochemistry SFA, ESS-DIVE repository [Dataset]. doi:10.15485/1985922

Savoy, P., A. P. Appling, J. B. Heffernan, E. G. Stets, J. S. Read, J. W. Harvey, and E. S. Bernhardt. 2019. Metabolic rhythms in flowing waters: An approach for classifying river productivity regimes. Limnol. Oceangr. **64**: 1835–1851.

Ulseth, A. J., R. O. Hall Jr., M. Boix Canadell, H. L. Madinger, A. Niayifar, and T. J. Battin. 2019. Distinct air–water gas exchange regimes in low-and high-energy streams. Nat. Geosci. **12**: 259–263.

USACE. 2002. Dissolved gas abatement study, final report. Technical report, US Army Corps of Engineers.

Welter, J. R., J. P. Benstead, W. F. Cross, J. M. Hood, A. D. Huryn, P. W. Johnson, and T. J. Williamson. 2015. Does N_2 fixation amplify the temperature dependence of ecosystem metabolism? Ecology **96**: 603–610.

Wetzel, R. G., and G. E. Likens. 2010. Limnological analyses, 3rd ed. Springer.

Acknowledgments

Thanks to Lupita Renteria and Flora Yeager for field assistance. Two anonymous reviewers provided helpful comments that improved this paper. R.O.H. was supported by the Modelscape project, NSF 2019528. This study was partially funded by the River Corridor Science Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). PNNL is operated by Battelle Memorial Institute for the US DOE under Contract No. DE-AC05-76RL01830. The SFA is supported by the US DOE, Office of Biological and Environmental Research (BER), Environmental System Science (ESS) Program.

Conflict of Interest

None declared.

Submitted 13 March 2023
Revised 30 August 2023
Accepted 10 September 2023

Associate editor: Ryan A. Sponseller