RESEARCH ARTICLE

Dynamics of the stream–lake transitional zone affect littoral lake metabolism

Nicole K. Ward[1] · Jennifer A. Brentrup[2] · David C. Richardson[3] · Kathleen C. Weathers[4] · Paul C. Hanson[5] · Russell J. Hewett[6] · Cayelan C. Carey[1]

Abstract

Lake ecosystems, as integrators of watershed and climate stressors, are sentinels of change. However, there is an inherent time-lag between stressors and whole-lake response. Aquatic metabolism, including gross primary production (GPP) and respiration (R), of stream–lake transitional zones may bridge the time-lag of lake response to allochthonous inputs. In this study, we used high-frequency dissolved oxygen data and inverse modeling to estimate daily rates of summer epilimnetic GPP and R in a nutrient-limited oligotrophic lake at two littoral sites located near different major inflows and at a pelagic site. We examined the relative importance of stream variables in comparison to meteorological and in-lake predictors of GPP and R. One of the inflow streams was substantially warmer than the other and primarily entered the lake's epilimnion, whereas the colder stream primarily mixed into the metalimnion or hypolimnion. Maximum GPP and R rates were 0.2–2.5 mg O_2 L^{-1} day^{-1} (9–670%) higher at littoral sites than the pelagic site. Ensemble machine learning analyses revealed that > 30% of variability in daily littoral zone GPP and R was attributable to stream depth and stream–lake transitional zone mixing metrics. The warm-stream inflow likely stimulated littoral GPP and R, while the cold-stream inflow only stimulated littoral zone GPP and R when mixing with the epilimnion. The higher GPP and R observed near inflows in our study may provide a sentinel-of-the-sentinel signal, bridging the time-lag between stream inputs and in-lake processing, enabling an earlier indication of whole-lake response to upstream stressors.

Keywords Ecosystem function · Machine learning · Littoral · Pelagic · GPP · R

Introduction

Lake ecosystems are sentinels of change in the landscape, since they integrate watershed and climate stressors (Adrian et al. 2009; Williamson et al. 2009). However, teasing out the complex ways in which lakes integrate landscape and meteorological drivers is often difficult due to long time lags between upland drivers and downstream responses and high variability within response variables (Wilkinson et al. 2020). Volumetric lake metabolism, including gross primary production (GPP) and respiration (R), is an integrated metric of lake ecosystem functioning and state (Solomon et al. 2013). Lake metabolism patterns are dominated by trophic state, a variable that generally remains constant from year to year and corresponds to the long water residence times in lakes, from years to decades. Whereas stream metabolism patterns are dominated by nutrients loads, a variable that may change more quickly, corresponding to the shorter water residence times in streams, from minutes to days (Hotchkiss

✉ Nicole K. Ward
nkward@vt.edu; nicole.ward@state.mn.us

1 Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA

2 Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

3 Biology Department, SUNY New Paltz, New Paltz, NY, USA

4 Cary Institute of Ecosystem Studies, Millbrook, NY, USA

5 Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA

6 Department of Mathematics, Virginia Tech, Blacksburg, VA, USA

et al. 2018). Since lake metabolism is an integrated metric of ecosystem state, as increasing upstream stressors drive a nutrient-limited lake toward a trophic state change, increases in-lake metabolism may be detected before more commonly measured variables such as phosphorus (Richardson et al. 2017). As such, lake metabolism near stream–lake transitional zones may provide a sentinel-of-the-sentinel signal, serving as an intermediary indication of the time-lagged whole-lake response to shorter-term stream inflow stressors. Specifically, lake metabolism near stream–lake transitional zones may help identify how different inflow streams contribute to in-lake response and provide key insights into freshwater cross-ecosystem connections (Hotchkiss et al. 2018; Hanson et al. 2015).

Previous lake metabolism work has focused on pelagic metabolism to represent whole-lake conditions or on spatial variability in metabolism to improve accuracy of whole-lake estimates. Pelagic lake metabolism is generally considered to be representative of whole-lake processing that depends on lake water residence times, typically on the scale of years to decades (Hotchkiss et al. 2018). In addition, pelagic lake carbon is often autochthonous (Hotchkiss et al. 2018), and thus, patterns in GPP and R are generally driven by water-column conditions (Hoellein et al. 2013). As a result, the magnitude of pelagic GPP and R can be an indicator of overall lake trophic state (Solomon et al. 2013) and the initiation of eutrophication in oligotrophic lakes (Richardson et al. 2017).

Littoral zone GPP and R is generally higher in comparison to pelagic sites (Sadro et al. 2011; Van de Bogert et al. 2012; Cavalcanti et al. 2016; Tonetta et al. 2016). Higher rates of GPP and R in the littoral zone are often attributed to greater substrate-surface water interactions, resulting in the sediment release of nutrients and carbon, and other littoral conditions, such as macrophytes and greater light availability (Vadeboncoeur et al. 2006; Simčič & Germ 2009; Cavalcanti et al. 2016; Tonetta et al. 2016). However, littoral sites near the mouths of inflow streams may have localized water residence times much shorter than the lake epilimnetic pelagic zone and are likely more indicative of lake connections with the upstream landscape than pelagic sites (Chmiel et al. 2020). Thus, another mechanism for higher GPP and R in littoral areas may be due to river water stimulation (Johengen et al. 2008).

While stream inflows may provide key nutrient subsidies in nutrient-limited oligotrophic lake epilimnia (MacIntyre et al. 2006), the capacity for inflow streams to increase littoral GPP and R will also be dependent on the characteristics of the stream–lake transitional zone. We define the stream–lake transitional zone to be a hypothesized "activated ecosystem control point" (following Bernhardt et al. 2017), requiring the combination of appropriate abiotic conditions and delivery of limiting nutrients to stimulate disproportionately high rates of biogeochemical processing. First, density-based physical mixing in the stream–lake transitional zone will determine if inflowing stream water mixes with the lake surface water or plunges into deeper layers. In stream–lake transitional zones where water density differences are primarily temperature-driven, mixing is determined by the inflowing stream water temperature and the temperature profile in the littoral zones, with warmer stream water mixing with the epilimnion and colder stream water sinking to deeper layers as underflow or entering the metalimnion as interflow (Imberger and Hamblin 1982; Vincent et al. 1991; MacIntyre et al. 2006; Rueda and MacIntyre 2010; Cortés et al. 2014). Second, the nutrient (e.g., nitrogen and phosphorus) and carbon concentrations and stoichiometry in the inflowing streams relative to the lake will determine if the inflowing water provides a subsidy to GPP and R in the littoral zone (MacIntyre et al. 2006). The inflowing stream could serve as a subsidy if it has higher nutrient and carbon concentrations than the littoral epilimnion, and if the epilimnion is nutrient or carbon limited. Conversely, metabolism may decrease via a dilution effect if the inflowing streams have lower nutrient concentrations or via decreased light penetration if the stream increases turbidity in the lake. Accordingly, the spatial and temporal extent of the stream–lake transitional zone will vary depending on the stream–lake mixing conditions as will the magnitude of biogeochemical response.

In this study, we examined if the magnitude of littoral GPP and R response was related to stream–lake mixing conditions at two littoral sites located near major inflows to a nutrient-limited oligotrophic lake. In particular, we were interested in comparing the relative importance of stream-related predictor variables to meteorological and in-lake predictor variables of GPP and R. We compared volumetric metabolism in the epilimnion at two littoral sites to estimates at a pelagic site to test if the commonly observed pattern of higher littoral zone GPP and R was also observed in our study lake. The overall aim of our study was to determine if GPP and R near stream–lake transitional zones was impacted by the water temperature and biogeochemistry of the inflow streams. Altered GPP and R at transitional zones may indicate lake integration of stream inflows in ways that bridge the time-lag between upstream stressors and whole-lake response. New insights on stream–lake transitional zones may inform monitoring protocols for identifying sources of nutrient pollution in catchments and advance our overall understanding of in-lake responses to upstream stressors.

Methods

Study area description

Lake Sunapee (New Hampshire, USA; 43° 24′ N, 72° 3′ W) is an oligotrophic, dimictic lake. Lake Sunapee has a

surface area of 16.55 km^2, a mean water residence time of 3.1 years, volume of 1.88×10^8 m^3, mean depth of 10 m, maximum depth of 33 m, mean June–August thermocline depth of 7 m (Carey et al. 2014a; Richardson et al. 2017), and ice cover generally from December–January to March–April (Bruesewitz et al. 2015). The glacially formed lake is irregularly shaped with high shoreline complexity (shoreline development index of 3.6 due to the multiple coves or bays adjacent to the main basin). Each of the two focal coves in our study have two inflowing streams, though the primary inflow to each cove (subwatersheds highlighted in Fig. 1) provide the large majority of surface flow to the cove. The watershed is experiencing increasing conversion of forest to cleared land for housing and urban development (Ward et al. 2020).

Lake Sunapee is classified as oligotrophic using trophic state indices based on pelagic chlorophyll-*a*, Secchi depth, and total phosphorus concentrations, but daily rates of pelagic R and GPP have been increasing since 2007, indicating potential trophic state change (Carey et al. 2014a; Richardson et al. 2017). Although there have been no detectable changes in pelagic total phosphorus concentrations over the past 3 decades (Steele et al. 2021), littoral zone total phosphorus concentrations are increasing (Richardson et al. 2017). Lake Sunapee is nutrient limited with co-limitation of nitrogen and phosphorus (Ward et al. 2020; Carey et al. 2014b).
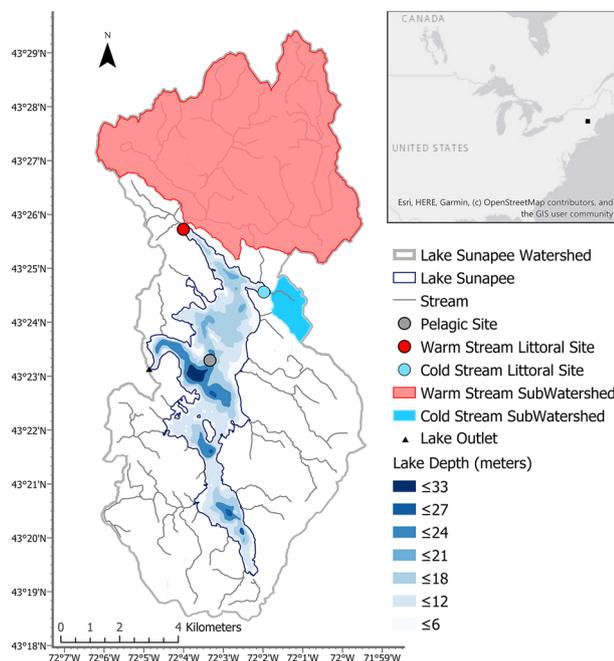

Fig. 1 Study area map: Lake Sunapee watershed and metabolism sample sites, New Hampshire, USA; upper bound of lake depth is inclusive

Study design overview

We deployed dissolved oxygen (DO) and water temperature sensors in three different oligotrophic lake sites—one site in the lake pelagic zone and two littoral sites near different inflow streams. From these data, we quantified site-specific metabolism and compared daily metabolism estimates among sites. We used an ensemble machine learning approach to quantify the relative contribution of meteorological, lake, and stream-related predictors to GPP and R estimates at the two littoral sites. In addition, we conducted the same machine learning predictor analysis on the pelagic site GPP and R to identify if relationships between the streams and in-lake metabolism were unique to the littoral sites. Ensemble machine learning approaches are well-designed for datasets with high noise and predictor variable co-linearities (Crisci et al. 2012; Boehmke and Greenwell 2020), common features of ecological data which often complicate ecosystem metabolism analyses (Coloso et al. 2011; Rose et al. 2014; Dormann et al. 2013). We specifically focused on predictor variables that are commonly measured by freshwater monitoring programs.

Littoral site characterization and field data collection

From three in-lake buoy sites (Table 1; Fig. 1), we collected epilimnetic dissolved oxygen (1–1.75 m below the surface) and water profile temperature measurements during the summer stratified period in 2018 (June through September). The pelagic site buoy was established in 2007 and is monitored by the Lake Sunapee Protective Association. It is located near the deepest part of the lake, at a site which is 12.5 m deep and 1150 m from the closest shore. The two littoral sites were established for this study and strategically located near major inflow streams that flow into coves of the lake (Fig. 1), with the expectation that those coves are likely largely influenced by the inflow streams entering them and less influenced by dynamics at the deepest part of the lake. Both of the coves included in this study have similar morphometry, sediment substrate, very low macrophyte cover, and maximum depths close to the average thermocline depth. Moreover, the depths at the opening of both coves to the main basin were generally shallower than the thermocline.

The sub-catchments of the two inflow sites differ substantially in size and hydrology, but both are primarily forested and have no agricultural land (Ward et al. 2020). The northern sub-catchment has increasing housing development pressure and is the largest sub-catchment (32.6 km^2) to the lake, flows through two lentic ecosystems (Little Lake Sunapee and Otter Pond) before entering Lake Sunapee, and provides ~50% of the lake's stream inflow volume (Schloss 1990). The north inflow

Table 1 Sampling site characteristics in Lake Sunapee, NH, USA

Site	Description	Buoy deployment duration (2018)	Buoy site depth (m)	Depth of DO sensor (m)	Distance to nearest inflow (m)	DO sensor type	Thermistor depths (m)
Pelagic	Near deepest location of lake	24 May–21 Sept	12.5	1.0	1150	Onset HOBO U26	0.5–9.5 by 1 m increments
Warm-stream littoral	Northwest cove of lake, near warm-inflow stream	3 Jun–21 Sept	7.5	1.75	160	PME miniDOT	0.1, 0.5–3.0 by 0.5 m, 3.0–7.0 by 1 m
Cold-stream littoral	Northeast cove of lake, near cold inflow stream	1 Jun–21 Sept	7.0	1.75	220	PME miniDOT	0.1, 0.5–6.5 by 0.5 m

DO dissolved oxygen

is warmer than the other inflows (Ewing et al. 2021) because of its long residence time in the two upstream lakes and flow over a dam spillway before entering Lake Sunapee. Consequently, we refer to the littoral site adjacent to the north inflow hereafter as the "warm-stream littoral site." The warm-stream littoral site was in the northwest cove of the lake in 7.5 m depth and 160 m from the warm-inflow stream (Fig. 1). In contrast, the sub-catchment on the northeast side of the lake is much smaller (1.7 km²), contributing ~3% of the lake's stream inflow volume and is characterized by cooler water temperatures and more variable stream discharge (Ewing et al. 2021). Consequently, we refer to the littoral site adjacent to the northeast inflow hereafter as the "cold-stream littoral site." The cold-stream littoral site was in the northeast cove of the lake at 7 m depth and 220 m from the cold inflow stream (Fig. 1). Due to the large water temperature differences between the two inflow streams, we assumed the density differences were primarily due to temperature. However, future study to fully resolve differences between the inflow streams should include density differences due to suspended sediments.

We monitored stream water temperature, depth, total nitrogen (TN), total phosphorus (TP), and dissolved organic carbon (DOC) in the warm stream 200 m upstream of where it entered the lake and the cold stream 50 m upstream of where it entered the lake during the study period. The locations of stream sampling were determined by where we were logistically able to access the stream. Water temperature and depth were measured with in-stream HOBO Water Level Loggers (Onset Corporation, Bourne, Massachusetts) and recorded at 15-min intervals (Ewing et al. 2021). For the water chemistry data, grab samples were collected weekly June through July and once every two weeks August through September. When flows were elevated (i.e., reaching approximately halfway between baseflow stage and bank-full stage) due to precipitation, automated ISCO samplers were triggered to collect hourly samples. We analyzed the water chemistry samples following the methods of Murphy and Riley (1962), Brenton and Arnett (1993), EPA (1993), Patton and Kryskalla (2003), and APHA (2005) (Supplemental Text 1).

All three lake buoys were equipped with a surface (1–1.75 m) DO sensor, underwater light sensor (HOBO pendant loggers; Onset Corporation, Bourne, Massachusetts) and a thermistor chain with temperature sensors (HOBO pendant loggers; Onset Corporation, Bourne, Massachusetts) at 0.5–1 m intervals from the surface to the bottom (Table 1; littoral buoy data available at Ward et al. 2021; pelagic buoy data available at LSPA 2020b). The DO sensors at the littoral sites were PME miniDOTs collecting data every 10 min (Precision Measurement Engineering, Vista, California) and the pelagic site DO sensor was a HOBO U26 collecting data every 15 min (Onset Corporation, Bourne, Massachusetts). To facilitate comparison across sites, we standardized DO sensor values using calibrated manual YSI ProDSS ODO/CT (YSI corp., Yellow Springs, Ohio) measurements collected weekly to monthly throughout the sampling period (Figs. S1, S2). Manual YSI versus high-frequency sensor temperature measurements were very close to 1:1, indicating similarly estimated DO concentrations at saturation between the YSI and high-frequency sensor (Fig. S2d–f), but comparisons of DO concentrations were offset from the 1:1 line (Fig. S2a–c). To enable comparisons across sites, we corrected the high-frequency sensor DO concentrations using the corresponding YSI comparison DO measurements (Fig. S2a, b, c). At each buoy, we also collected weekly grab samples during June and July, and monthly samples during August and September using a Van Dorn water sampler (Wildco, Yulee, Florida) for TN, TP, and DOC at the depth of the DO sensor, following the same methods of analysis as the stream grab samples described in Supplemental Text 1.

Data analysis

Littoral site characterization

We used the field data to calculate a variety of water physical indices to compare the two littoral sites. For both littoral sites, we calculated three metrics of stream–lake mixing:

nominal intrusion depth, inflow Froude number, and transition Richardson number. Each metric required water density, which we calculated using observed water temperature, an assumed salinity of zero, and the "water.density()" function in the rLakeAnalyzer R package (Winslow et al. 2019). The nominal intrusion depth was calculated as the lake depth of minimum density difference between stream inflow water and epilimnetic lake water based on littoral buoy water temperature profiles (sensu Cortés et al. 2014). On days when inflow stream water was denser than epilimnion water at each littoral site, we also characterized the plunging inflows, i.e., inflows with greater density due to cooler temperatures than the surface layer of the lake, to determine potential for interflow incorporation into the epilimnion using the dimensionless Froude number and transition Richardson number (Ri_{12}) (sensu Cortés et al. 2014). All three of these metrics (nominal intrusion depth, inflow Froude number, and transition Richardson number) were calculated from lake water temperature profiles and stream water temperature measurements, and the inflow Froude number also used estimated stream velocity (Table S1).

We used non-parametric Mann–Whitney U tests to compare nominal intrusion depth at the two littoral sites and to examine stream–lake differences in water density, TN, TP, and DOC concentrations at each littoral site to determine the potential for stream inflows to supplement lake nutrient concentrations. To supplement the site comparison of water density differences, a comparison of inflow Froude number and transition Richardson number is presented in the supplement, though patterns follow established site differences due to water density. All analyses were conducted in R (R Core Team, 2021; Version 4.0.4).

GPP and R estimates

We estimated metabolism metrics using inverse modeling with maximum likelihood to fit gross primary production (GPP) and respiration (R) as model parameters to predict observed DO concentrations ("metab.mle()" function from the LakeMetabolizer R package in Winslow et al. 2016; sensu Van de Bogert et al. 2007; Hanson et al. 2008; Solomon et al. 2013; Richardson et al. 2017). We estimated daily GPP and R rates in the summer stratified period, from when sensors were deployed in late May–early June (Table 1) until 22 September, as fall turnover occurred on 23 September (defined by a density difference between 2 m from the surface and 2 m above the sediments less than 0.1 kg m^{-3}, sensu Andersen et al. 2017).

We used the methods of Richardson et al. (2017) and Brentrup et al. (2021) to estimate metabolism at the three sites (Supplemental Text 2). All metabolism estimates are published in the EDI repository (Ward et al. 2022) and we did not quantify model process errors. Daily rates of R and

GPP at all sites were generally low (often 0.2–0.5 mg O$_2$ L^{-1} Day^{-1}) and did not exhibit temporal autocorrelation (Figs. S4, S5), as determined by the autocorrelation "acf()" and partial autocorrelation "pacf()" function in the stats R package, enabling us to conduct metabolism site comparisons without needing to take autocorrelation into account.

We examined two aspects of site-specific GPP and R (detailed below). First, we determined if differences among the pelagic and littoral sites in our study followed previously established patterns; specifically, higher littoral zone GPP and R in comparison to pelagic epilimnetic GPP and R (e.g., Sadro et al. 2011; Cavalcanti et al. 2016; Tonetta et al. 2016; Van de Bogert et al. 2012). Second, we used an ensemble machine learning approach to identify potential predictors of GPP and R, and how those predictors might differ between sites. We included water temperature as a potential predictor variable, so did not apply temperature-scaling corrections to GPP and R estimates.

Site-to-site differences in GPP and R

To test for differences in daily GPP and R among the three sample locations in the lake, we used non-parametric paired Mann–Whitney U tests due to non-normal distributions of GPP and R. The paired Mann–Whitney U tests with Bonferroni-corrected α for multiple comparisons identified pairwise differences among littoral and pelagic site epilimnetic GPP and R.

Littoral site GPP and R predictor analysis

To determine if littoral GPP and R were related to stream-related drivers, we first grouped all potential predictor variables ($n = 17$) into three categories (Table S1, Figs. S6–S9). The predictor variables included: (1) meteorological predictors ($n = 7$): wind speed and direction, air temperature, degree day, cumulative degree day, surface photosynthetically active radiation (PAR) (LSPA et al. 2020a), and precipitation from the North American Land Data Assimilation System (Xia et al. 2012); (2) lake predictors ($n = 5$) were pelagic site Schmidt stability and littoral site epilimnion water temperature, seiche period, underwater light, and littoral site Schmidt stability. Schmidt stability and seiche period were calculated using MATLAB Lake Analyzer (Read et al. 2011), derived from water temperature profiles at the Pelagic Site (LSPA et al. 2020b) and the two littoral sites. Water temperature and underwater light data are available in Ward et al. (2021); and (3) stream-related predictors ($n = 5$): stream water temperature difference from littoral site epilimnion, stream depth as a proxy for discharge, nominal intrusion depth, inflow Froude number, and transition Richardson number (derived from Ewing et al. 2021; Ward et al. 2021). As is common in analyzing the outputs of simulation

models (Snortheim et al. 2017), some predictor variables used in our machine learning analysis were also used as input variables to the metabolism model. For example, wind speed is included in the metabolism model to estimate changes in DO through atmospheric gas exchange, so that those changes in DO are not attributed to GPP or R rates. However, wind speed also drives mixing in the lake, which can directly affect GPP and R, and thus warrants investigation as a predictor, thereby meeting criteria for inclusion (Prairie and Bird 1989). We were careful in our interpretation to not weight too heavily the influence of these potentially confounded predictors.

We aggregated all variables to the daily scale and added a lag-1 predictor variable for precipitation and stream predictors to account for a one-day lag between stream conditions and delivery to the littoral zone of the lake (Table S1). We did not include grab samples and stream depth-based ISCO samples of TN, TP, and DOC due to the coarse timescale of those observations relative to the high-frequency data.

There are several challenging aspects of identifying and quantifying relationships between a suite of potential predictors and daily metabolism estimates. Metabolism signals tend to have high noise (Coloso et al. 2011; Rose et al. 2014), and ecological predictors tend to have co-linearities that make it difficult to discern their effects (Dormann et al. 2013). Issues of high noise and predictor co-linearities have been confronted by ensemble machine learning approaches (Crisci et al. 2012; Boehmke and Greenwell 2020), which give careful consideration to the model fitting process and validation methods (Vabalas et al. 2019).

Ensemble machine learning models are composed of multiple machine learning models, which are each referred to as base learners. The base learners of our ensemble machine learning model were: (1) random forest, due to its ability to work well with outliers, noisy data, and when one or two predictors may overwhelm the model prediction (Boehmke and Greenwell 2020); (2) regularized regression, due to its ability to address multicollinearity (Dormann et al. 2013) and smaller datasets with many predictor variables; and (3) eXtreme gradient boosting (XGB), due to its ability to overcome overfitting issues (Boehmke and Greenwell 2020). Base learners need to be tuned, trained, validated, and summarized before their predictions are used to evaluate the potential importance of predictors of lake metabolism. To assess the tuning process itself, we conducted a nested cross-validation model fitting assessment to compensate for biased fit assessments common with smaller sample sizes ($n = \sim 50$–100; Vabalas et al. 2019). A full description of our base-learner tuning method and nested cross-validation process is in Supplemental Text 3. Final ensemble and base-learner model hyperparameters and model performance are in Tables S2, S3, and Fig. S10.

To quantify and visualize the predictor–response variable relationships in the final ensemble model for GPP and for R at each littoral site, we examined variable importance for the top 10 predictor variables in each final ensemble model using a permutation method. Since variables that contribute to the overall model prediction in variable importance plots may not correspond to changes in magnitude of the response variable, we also examined specific predictor–response relationships for top predictors. For this step, we generated partial prediction and individual conditional expectation (ICE) plots for the top stream-related driver and top non-stream-related driver identified in the variable importance analysis, following Boehmke and Greenwell (2020) (additional information in Supplemental Text 3). Importantly, the magnitudes of metabolism metric response to predictor variables were discerned through the partial prediction and ICE plots to identify if variables identified in the variable importance plots contributed to ecologically meaningful relationships. We specifically focused the machine learning analysis on predictors of littoral metabolism to identify if littoral site GPP and R was associated with stream-related variables. In addition, we conducted the same machine learning analysis for the pelagic site to confirm that relationships between the streams and in-lake metabolism were unique to the littoral sites. All analyses were run in R (R Core Team, 2021; version 4.0.4) and the code is available in Zenodo (Ward 2021).

Results

Littoral sites

The littoral sites had contrasting stream–lake transitional zone characteristics. The water density of the warm stream was equal to or less than the density of the surface layer (0–2 m) at the warm-stream littoral site on 95 out of 112 (85%) days (Fig. 2b; Mann–Whitney $U = 5100$, $p = 0.9$), indicating direct mixing of the inflow stream with the surface mixed layer in the receiving littoral zone on those days. Conversely, the water density of the cold stream was greater than the surface layer at the cold-stream littoral site on 111 of 112 (99%) days (Fig. 2b; $U = 12,058$, $p < 0.001$). As a result, the nominal intrusion depth, or the lake depth of minimum density difference between inflow and lake water (sensu Cortés et al. 2014), at the warm-stream littoral site was significantly shallower than at the cold-stream littoral site (Fig. 2a; $U = 925$, $p < 0.001$). Throughout the study, the thermocline was generally at or near the bottom of the thermistor chain at the littoral sites (Fig. S3). Surface water (0–2.5 m) temperatures were very similar at the warm-stream littoral site, cold-stream littoral site, and the pelagic site (mean = 22.8, 23.2, and 22.7 °C, respectively; maximum = 26.6, 27.2, and 26.2 °C, respectively; minimum = 16.8, 17.2, and 16.9 °C, respectively).
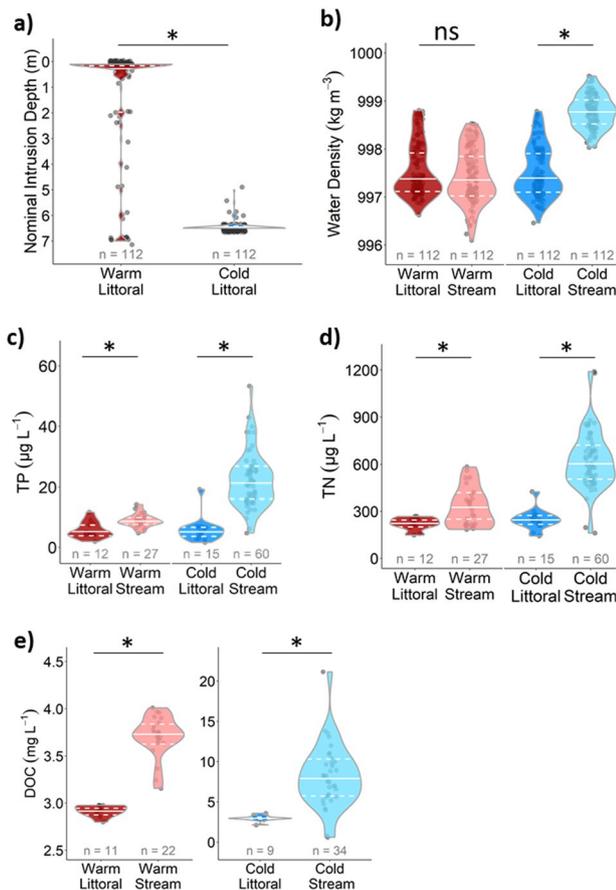
Fig. 2 Comparison of **a** the cold-stream littoral site vs. warm-stream littoral site **a** nominal intrusion depth, **b** the water density in the warm-inflow stream to warm-stream littoral site and comparison of the water density in the cold inflow stream to cold-stream littoral site, **c** comparison of total phosphorus (TP) in both streams and littoral sites, **d** total nitrogen (TN) in both streams and littoral sites, and **e** dissolved organic carbon (DOC) in both streams and littoral sites. *Indicates p value < 0.05; ns indicates not significant Mann–Whitney U test. Note that the y-axis in panel a is reversed to show the water's surface at the top of the plot, and sediments at the bottom; solid white line indicates median, dashed white line indicates interquartile range

Altogether, the stream–lake transitional zone physical metrics indicated the warm stream was more often mixing with the surface water of the littoral zone, while the cold stream was more often entering the lake as interflow or underflow. Since the inflow Froude number and transition Richardson number primarily quantify mixing potential for colder inflow streams entering warmer water, we present the metrics here for days when each inflow stream was more dense than the receiving littoral surface water. For the days when the water density in the warm stream was greater than the littoral surface layer, the mean inflow Froude number was 11.9 (SD = 4.84, maximum = 29.0, minimum = 6.6, $n = 20$) and the mean transition Richardson number (Ri_{12}) was 5.0 (SD = 7.2, maximum = 31.4, minimum = − 0.04, $n = 20$).

The mean inflow Froude number for the cold stream was 1.3 (SD = 0.20, maximum = 1.75, minimum = 0.84, $n = 83$), indicating a quickly plunging flow (~ 4 m of stream–lake interface, Johnson and Stefan 1988). The mean Ri_{12} for the cold stream was 0.54 (SD = 0.34, maximum = 1.83, minimum = 0, $n = 83$), suggesting potential for distinct underflow on most days (when $Ri_{12} << 1$) and incorporation of the interflow into the epilimnion on other days ($Ri_{12} \geq 1$; Cortés et al. 2014).

Nutrient and DOC concentrations in the inflow streams in comparison to the littoral sites in the lake indicate potential for delivery of nutrient and carbon subsidies (Fig. 2c–e). The mean TP, TN, and DOC concentrations in the inflow streams were all significantly greater than the surface water at the corresponding warm-stream littoral site and cold-stream littoral site (Mann–Whitney U TP: $U = 269.5$ and 863, both $p < 0.001$, TN: $U = 237.5$ and 872, $p = 0.01$ and < 0.001, DOC: $U = 242$ and 288, both $p < 0.001$; Fig. 2c–e). Surface water TP, TN, and DOC was not significantly different between the warm-stream littoral site, cold-stream littoral site, and pelagic site (Mann–Whitney U all $p > 0.05$). A higher n value is likely needed to differentiate any existing site-to-site differences in this oligotrophic lake with relatively low nutrient and carbon concentrations in pelagic and littoral zones ($n < 15$ at each lake site).

Site-to-site differences in GPP and R

Metabolism estimates of GPP and R at the littoral sites were greater than at the pelagic site (Figs. 3, 4). GPP estimates at the warm-stream littoral site (median = 0.22, maximum = 2.89 mg O_2 L^{-1} day^{-1}, coefficient of variation (CV) = 132%) and at the cold-stream littoral site (median = 0.22, maximum = 0.63 mg O_2 L^{-1} day^{-1}, CV = 49%) were significantly greater than at the pelagic site (median = 0.18, maximum = 0.43 mg O_2 L^{-1} day^{-1}, CV = 39%; both Mann–Whitney $U \geq 3494$, both $p \leq 0.004$; Figs. 3, 4, Table 2). The highest daily rates of GPP were estimated at the warm-stream littoral site (Figs. 3, 4, Table 2), but overall, GPP at the warm-stream littoral site was not significantly greater than the cold-stream littoral site ($U = 2884$, $p = 0.22$; Figs. 3, 4, Table 2). The R estimates at the warm-stream littoral site (median = 0.22, maximum = 2.65 mg O_2 L^{-1} day^{-1}, CV = 126%) were significantly greater than at the pelagic site (median = 0.17, maximum = 0.56 mg O_2 L^{-1} day^{-1}, CV = 57%; $U = 4185$, $p < 0.001$; Figs. 3, 4, Table 2), but not significantly greater than at the cold-stream littoral site (median = 0.19, maximum = 0.61 mg O_2 L^{-1} day^{-1}, CV = 58%; $U = 3159$, $p = 0.03$, greater than Bonferroni-corrected α of 0.02 for three comparisons). The cold-stream littoral site R estimates were not significantly greater than the pelagic site R estimates ($U = 3165$, $p = 0.07$).

Fig. 3 Daily estimates of GPP (triangles) and R (circles) at **a** the pelagic site, **b** the warm-stream littoral site (note the different *y*-axis scale), **c** the cold-stream littoral site, and **d** all three sites, shown together for direct comparison using Loess curves
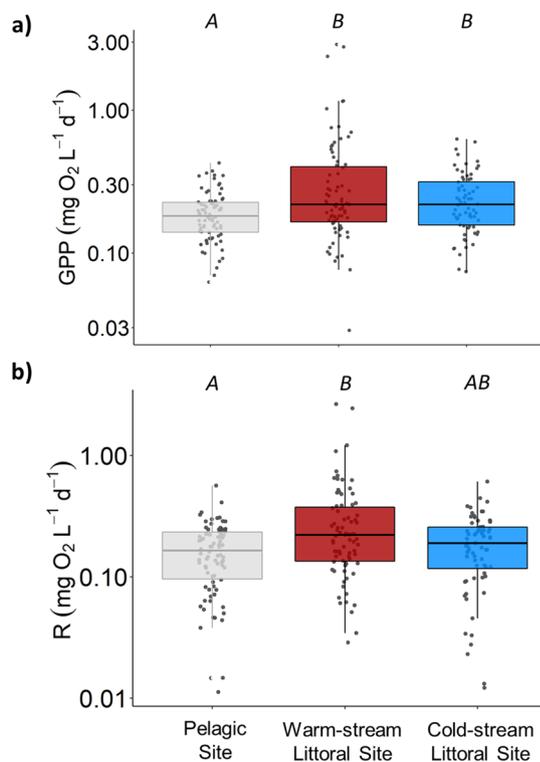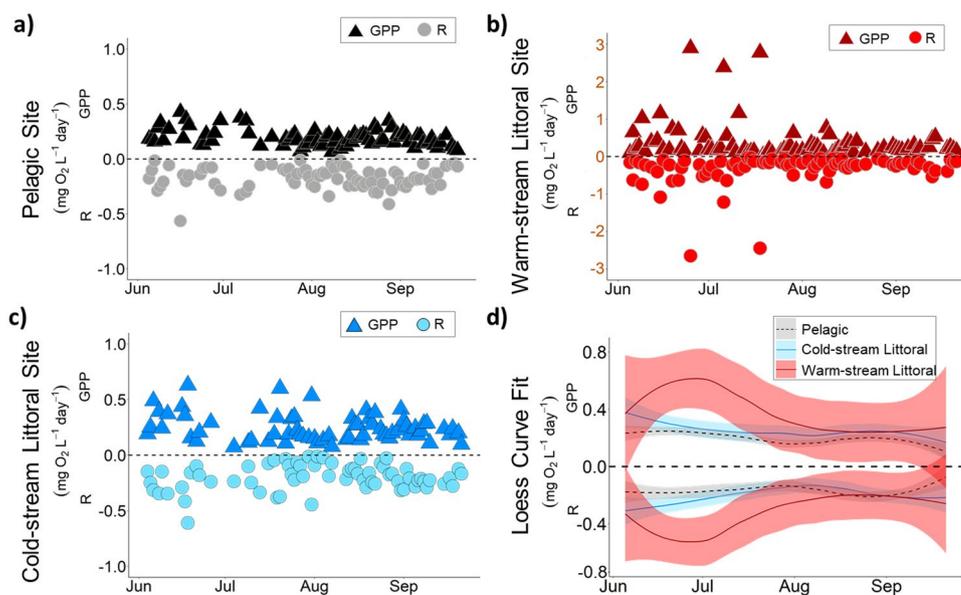
Fig. 4 Comparison of metabolism estimates among the pelagic site, cold-stream littoral site, and warm-stream littoral site in Lake Sunapee, NH, USA. Estimates of **a** gross primary production (GPP) and **b** respiration (R) are displayed on a $\log_{10}$ *y*-axis for easier visual comparison with boxplots denoting median and quartile ranges. Unique letters above boxplots indicate significant Mann–Whitney *U* comparisons ($p < 0.02$, Bonferroni-corrected α for 3 comparisons; see Table 2)

Predictors of littoral site GPP and R

Overall, the machine learning predictor analysis identified that both meteorological and stream variables affected GPP and R at the warm-stream littoral site (Fig. 5) and GPP at the cold-stream littoral site. Whereas primarily stream variables affected R at the cold-stream littoral site (Fig. 6). In contrast, at the pelagic site, primarily meteorological variables affected GPP and R (Fig. 7).

Though meteorological variables contributed most to GPP at the warm-stream littoral site ($62\% \pm 16\%$, 1 S.D. of the ensemble model), stream-related variables also affected GPP response ($34\% \pm 7\%$; Fig. 5). The variable importance analysis identified that the most important predictor of warm-stream littoral site GPP was wind direction ($37\% \pm 10\%$ of the model). When wind was out of the northwest, GPP was 0.41 mg O_2 L^{-1} day^{-1} higher (122% increase) than when wind was from any other direction (Fig. 5c). The most important stream-related driver was the temperature difference between the stream and littoral zone epilimnion water ($19\% \pm 3\%$ of the model; Fig. 5a). Similar to the effect of wind direction, warmer stream water in comparison to littoral zone epilimnion water resulted in a 0.43 mg O_2 L^{-1} day^{-1} (126%) increase in GPP (Fig. 5e).

Similar to warm-stream littoral site GPP, the warm-stream littoral site R was modulated by both meteorological variables ($66\% \pm 11\%$ of the model) and stream-related variables ($26\% \pm 7\%$; Fig. 5). The variable importance analysis identified that the most important predictor of warm-stream littoral R was the previous day's precipitation (precipitation

Table 2 Mann–Whitney *U* site-to-site comparisons of metabolic rates in Lake Sunapee, NH, USA

Metabolism metric	Site (*n*)	Site (*n*)	*U*	*p* value
GPP	Pelagic (81)	Warm-stream littoral (79)	2273	**0.0008**
	Pelagic (81)	Cold-stream littoral (67)	2024	**0.004**
	Cold-stream littoral (67)	Warm-stream littoral (79)	2447	0.22
R	Pelagic (81)	Warm-stream littoral (79)	2253	**0.0006**
	Pelagic (81)	Cold-stream littoral (67)	2324	0.07
	Cold-stream littoral (67)	Warm-stream littoral (79)	2192	0.03

P values < 0.02 (Bonferroni-corrected *α* for 3 comparisons) are bolded

GPP gross primary production, *R* respiration, *n* sample size, *U* Mann–Whitney *U* test statistic

lag-1; 33% ± 5% of the model; Fig. 5b). When total daily precipitation was greater than 2 cm, R on subsequent days was 0.30 mg O_2 L^{-1} day^{-1} higher (104% increase) than after days with no precipitation (Fig. 5d). The most important stream-related predictor for R at this site was the water temperature difference between the stream and littoral zone epilimnion (stream minus epilimnion T; 13% ± 3% of the model; Fig. 5b). The partial prediction analysis for R showed a smaller response to stream minus epilimnion temperature than GPP: the warmer stream water in comparison to littoral zone epilimnion water resulted in a 0.17 mg O_2 L^{-1} day^{-1} (56%) increase in R (Fig. 5f).

At the cold-stream littoral site, stream variables contributed 49% (± 8%) for GPP and 69% (± 21%) for R to the overall ensemble machine learning predictor–response model for each metabolism metric (Fig. 6a, b). Similar to the warm-stream littoral site, the variable importance analysis identified that the top predictor for cold-stream littoral GPP was wind direction, which contributed 23% (± 5%) to the overall predictor variable–response relationship in the ensemble machine learning model (Fig. 6a). In contrast to the wind direction patterns at the warm-stream littoral site (Fig. 5c), the highest magnitude GPP at the cold-stream littoral site was associated with wind out of the south (Fig. 6c), though the increase of the GPP in response to wind direction was small (0.02 mg 0_2 L^{-1} day^{-1}, 7% increase; Fig. 6c). The most important stream-related predictor for GPP was the dimensionless inflow Froude number, which contributed 18% (± 2%) to the overall predictor variable–response relationship (Fig. 6a). The inflow Froude number resulted in an increase in GPP of 0.03 mg O_2 L^{-1} day^{-1} (11% increase) when the inflow Froude number was greater than 1.75, indicating greater stream mixing with the surface water (Fig. 6e).

In contrast to all other GPP and R responses, the cold-stream littoral R was driven by both stream variables (69% ± 21% of the model) and lake variables (21% ± 5%; Fig. 6b). The variable importance analysis identified that the top predictor for R at this site was a stream-related variable: the inflow Froude number, which contributed 47% (± 14%)

to the overall predictor variable–response relationship in the ensemble machine learning model (Fig. 6b). As the inflow Froude number increased from 0.8 to 1.8, indicating more mixing potential with littoral zone surface water, R increased by 0.07 mg O_2 L^{-1} day^{-1} (43% increase; Fig. 6d). The top non-stream-related driver was pelagic Schmidt stability, though the associated change in R was quite small (< 0.01 mg O_2 L^{-1} day^{-1}, 3% change; Fig. 6f).

The warm-stream and cold-stream littoral sites exhibited similar top predictor variables for GPP, but the strength of the predictors varied, as indicated by the partial prediction and ICE plots. Wind direction was the top predictor variable for GPP at both littoral sites (Figs. 5a, 6a); however, there was a much larger predictive range in the warm-stream littoral GPP response (0.41 mg O_2 L^{-1} day^{-1}, Fig. 5c) compared to the cold-stream littoral GPP response (0.02 mg O_2 L^{-1} day^{-1}, Fig. 6c). Similarly, the temperature difference between the stream and littoral surface water was the second-most important predictor of GPP at both sites (Figs. 5a, 6a); however, the predictive ranges and patterns were very different. At the warm-stream littoral site, increases in the stream-epilimnion temperature difference caused an increase in the GPP of 0.43 mg O_2 L^{-1} day^{-1} (126% increase), especially when stream temperatures were 3 °C warmer than the littoral lake temperature (Fig. 5e). Conversely, the temperature difference was negatively related to GPP (0.02 mg O_2 L^{-1} day^{-1}, 7% decrease) at the cold-stream littoral site, though the stream water temperature was always cooler than the cold-stream littoral site (Fig. 6e).

In contrast to GPP, the top predictor variables for R were different between the two littoral sites (Figs. 5b, 6b), in which the top two predictors for warm-stream littoral R were meteorological variables, but the top two predictors for cold-stream littoral R were stream variables. Similar to cross-site patterns in GPP, the predictive range in the warm-stream littoral R response was larger than the cold-stream littoral site. For example, the top stream-related predictor for warm-stream littoral R (but third-most important predictor overall in variable importance analysis, Fig. 5b), stream minus epilimnion temperature,
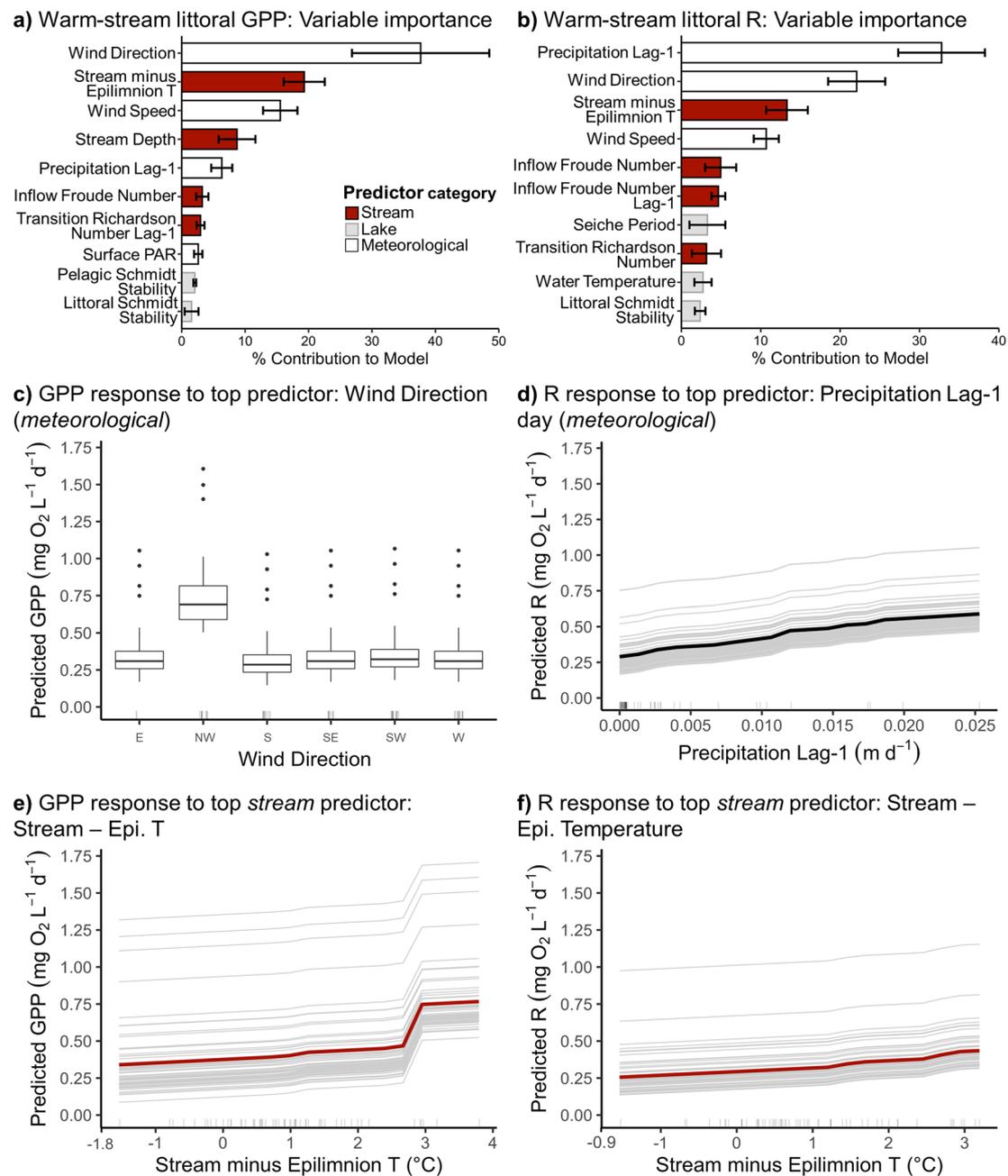
a) Warm-stream littoral GPP: Variable importance

b) Warm-stream littoral R: Variable importance

c) GPP response to top predictor: Wind Direction (*meteorological*)

d) R response to top predictor: Precipitation Lag-1 day (*meteorological*)

e) GPP response to top *stream* predictor: Stream – Epi. T

f) R response to top *stream* predictor: Stream – Epi. Temperature

Fig. 5 Predictor analysis results for the warm-stream littoral site in Lake Sunapee, NH, USA. Predictor variable importance plots, shown as percent contribution to ensemble machine learning model for GPP **a** and R **b**; partial prediction (thick line) and individual conditional expectation (ICE; thin lines) plots indicate presence of predictor–response relationships for the top non-stream-related predictor for GPP (**c**; categorical variable shown with box plot) and R (**d**) and the top stream-related predictor for GPP (**e**) and R (**f**). Each ICE line represents the focal predictor variable varying along the *x*-axis for each combination of other predictor variables observed. Partial dependence line is the average of ICE lines. Grey vertical dashes on *x*-axis indicate predictor variable observations

corresponded to an increase of 0.18 mg O_2 L^{-1} day^{-1} (70% increase; Fig. 5f). The top predictor for cold-stream littoral R, which was the stream-related variable of inflow Froude number, corresponded to a 0.07 mg O_2 L^{-1} day^{-1} (43%) increase (Fig. 6d).

In sum, stream variables affected daily rates of littoral lake metabolism, but were largely disconnected from pelagic metabolism response. Though both meteorological and stream variables were identified as important predictors of GPP and R from the variable importance analysis at
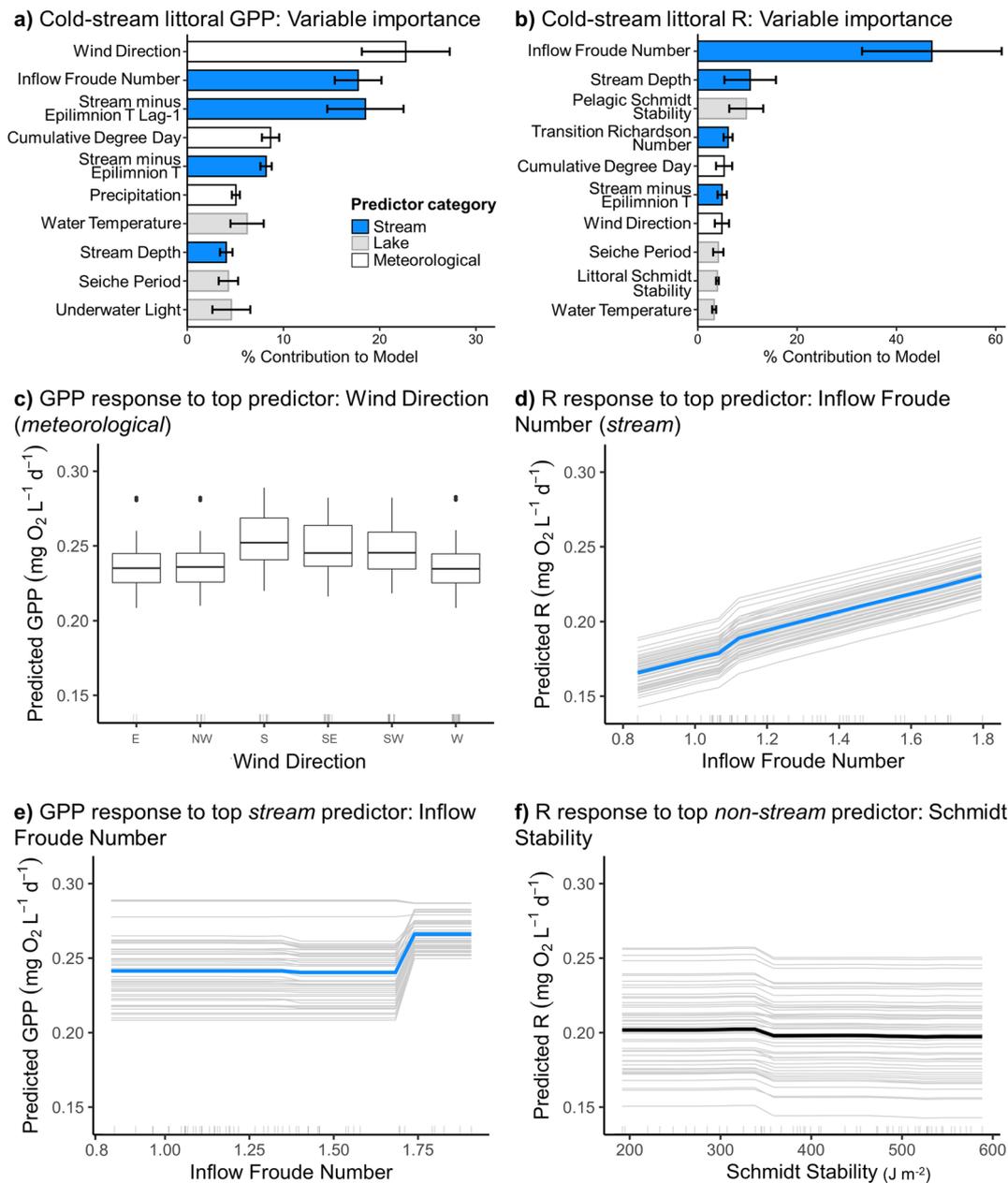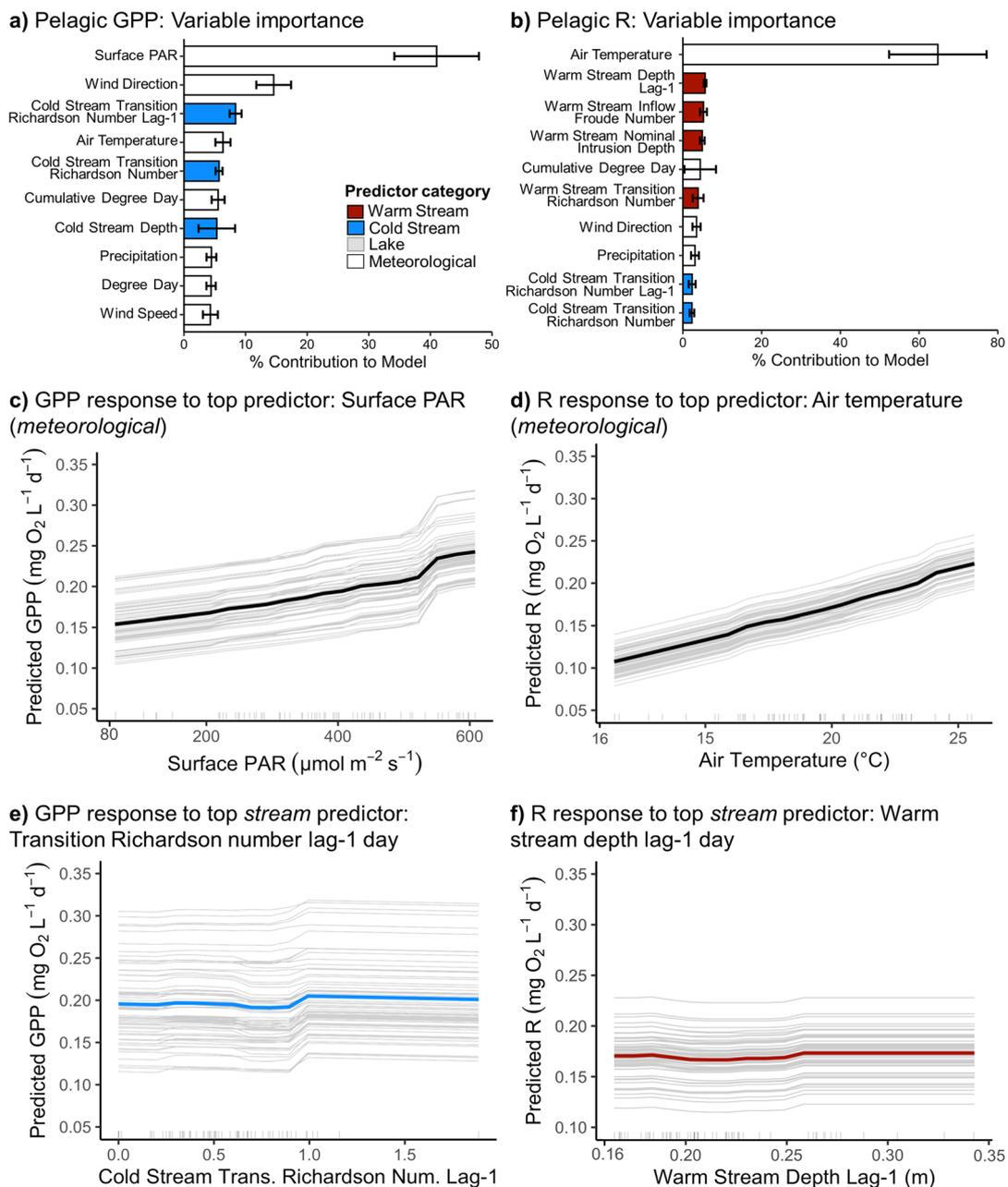
Fig. 6 Predictor analysis results for the cold-stream littoral site in Lake Sunapee, NH. Predictor variable importance plots, shown as percent contribution to ensemble machine learning model for GPP (**a**) and R (**b**); partial prediction (thick line) and individual conditional expectation (ICE; thin lines) plots indicate presence of predictor–response relationships for the top non-stream-related predictor for GPP (**c**); categorical variable shown with box plot and top predictor for R (**d**) and the top stream-related predictor for GPP (**e**) and top non-stream-related predictor for R (**f**). Each ICE line represents the focal predictor variable varying along the *x*-axis for each combination of other predictor variables observed. Partial dependence line is the average of ICE lines. Grey vertical dashes on *x*-axis indicate predictor variable observations

all three sites, the partial prediction and ICE analysis shows that stream variables contributed to ecologically meaningful responses in littoral GPP and R (Figs. 5, 6). In contrast, only meteorological variables contributed to ecologically meaningful responses in pelagic GPP and R (Fig. 7).

Discussion

The higher rates of littoral epilimnetic GPP and R observed with increased epilimnetic mixing at the stream–lake interface provides evidence that the stream–lake transitional zone

Fig. 7 Predictor analysis results for the pelagic site in Lake Sunapee, NH. Predictor variable importance plots, shown as percent contribution to ensemble machine learning model for GPP (**a**) and R (**b**); partial prediction (thick line) and individual conditional expectation (ICE; thin lines) plots are indicate presence of predictor–response relationships for the top non-stream-related predictor for GPP (**c**) and R (**d**) and the top stream-related predictor for GPP (**e**) and R (**f**). Each ICE line represents the focal predictor variable varying along the x-axis for each combination of other predictor variables observed. Partial dependence line is the average of ICE lines. Grey vertical dashes on x-axis indicate predictor variable observations

may function as an activated ecosystem control point (sensu Bernhardt et al. 2017). Accordingly, omitting these zones in whole-lake estimates may underestimate overall lake metabolism and leave an important aspect of lake metabolism spatial variability unaccounted for. The connection of higher near-stream littoral rates of GPP and R to inflow stream conditions bridges the inherent time-lag between upstream stressors and whole-lake response. Thus, the stream–lake transitional zone may provide a sentinel-of-the-sentinel signal in this nutrient-limited lake (Ward et al. 2020; Carey et al. 2014b) showing increasing productivity and the potential for future eutrophication (Richardson et al. 2017).

Littoral GPP and R indicate activated ecosystem control points

Activated ecosystem control points require the combination of favorable abiotic conditions and delivery of limiting nutrients to stimulate disproportionately high rates of biogeochemical processing (Bernhardt et al. 2017). The disproportionately high littoral GPP and R rates we observed in Lake Sunapee were dependent on greater stream–epilimnion mixing (favorable abiotic conditions) and stream nutrient subsidies (delivery of limiting nutrients), suggesting the stream–lake transitional zone functions as an activated ecosystem control point. Important context for interpretation of the stream–lake transitional zone as an activated ecosystem control point is that the lake is nutrient limited (Ward et al. 2020; Carey et al. 2014b), and the inflow streams have significantly higher nutrient concentrations than the receiving lake waters (Fig. 2).

While inputs from both warm and cold streams were shown to increase littoral metabolism, the mechanisms appear to differ. The physical mixing metrics indicated the warm-stream inputs entered into the surface mixed layer of the lake (sensu Vincent et al. 1991) and cold-stream inputs may have partially been entrained into the surface waters, creating conditions in which both streams likely provided nutrient subsidies (following MacIntyre et al. 2006). Though we were unable to measure daily stream and littoral zone nutrients and did not observe consistent patterns between our intermittent stream nutrient observations and estimated GPP and R (Figs. S11, S12), the GPP and R increases with greater stream mixing potentially support a nutrient-subsidy effect, following previous studies (Imberger and Hamblin 1982; Rueda and MacIntyre 2010; Cortés et al. 2014). Since the warm-stream inflow had lower nutrient concentrations than the cold-stream inflow, it is likely that a stronger signal of GPP and R stimulation would be found in a warm-inflow stream with higher nutrient concentrations. In addition, the different stream inflows may contribute different types of organic matter, resulting in different metabolic responses (Marcarelli et al. 2011).

Common pelagic epilimnetic metabolism predictors, such as water temperature and PAR, did not significantly contribute to the littoral GPP and R responses, highlighting the unique dependence of the littoral sites on the stream–lake mixing conditions. If the study were extended to include seasons of lake thermal mixing and ice cover, the more commonly considered in-lake and meteorological predictors may become more important for littoral GPP and R response. The significance of wind as a predictor of GPP response at both littoral sites (Figs. 5c, e, 6c, e) is likely a function of the cove and subbasin morphometry of the lake (Fig. 1). The warm-stream littoral site is most protected from wind exposure when wind is out of the northwest and the cold-stream littoral site is most protected from wind out of the south or southeast, corresponding to the wind direction associated with the highest rates of GPP at each littoral site. With minimal wind-induced mixing and surface disturbance, the water column may be more favorable for or decrease the horizontal variability of phytoplankton (Stockwell et al. 2020; Cyr 2017). Lower wind-induced mixing may enable the stream to affect littoral conditions more directly through nutrient subsidy or transport of DO from the stream to the littoral site. Similarly, precipitation effects could increase DO transport from the stream to the littoral zone, and depending on the timescale of influence relative to 24-h metabolism rates derived from diel DO curves, the estimated GPP and R rates may be affected. A more thorough consideration of hydrodynamic interactions would require detailed lake physics modeling of the spatial and temporal extent of the stream–lake transitional zone under different wind exposures and stream discharge.

Further research is needed to resolve the extent to which stream–lake transitional zones function as activated ecosystem control points. For example, a more detailed horizontal and vertical spatial analysis of metabolism in the lake would inform if the GPP and R rates observed in this study contribute "disproportionately" to whole-lake metabolism (Sadro et al. 2011). Both stream–lake transitional zone littoral sites were selected based on their similar depth, substrate, and lack of macrophytes to control for littoral zone influences on metabolism, however, including non-stream littoral sites could help clarify understanding of the unique influence of the streams on littoral GPP and R. In addition, a detailed physical tracking of the stream plumes (Rueda and MacIntyre 2009; Vincent et al. 1991) and GPP and R response in the lake would be required to identify the spatial and temporal dynamics of the stream–lake transitional zone as an ecosystem control point (Krause et al. 2017).

Linking lake metabolism to upstream stressors through the stream–lake transitional zone

The partial dependence of littoral GPP and R on stream-related predictors supports previous studies finding water residence time as a controlling variable (e.g., Catalán et al. 2016; Casas-Ruiz et al. 2017; Hotchkiss et al. 2018). Littoral metabolism near inflowing streams may provide a key intermediary measure linking small stream and large lake residence times. Pelagic metabolism estimates, which are more representative of water-column conditions in lakes with multi-year residence times (Hoellein et al. 2013; Hotchkiss et al. 2018), may not detect near-real time tributary effects on water quality in lakes like Lake Sunapee. Following the scaling of dominant metabolism behavior across water residence times (Hotchkiss et al. 2018), summer seasonal average chlorophyll-*a* in Lake Sunapee is more dependent on

climate-related variables (long-term response to climate) whereas summer maximum chlorophyll-*a* is more dependent on external nutrient load (short-term response to streams) (Ward et al. 2020).

Though it is well documented that metabolism can have high spatial variability in lakes (Sadro et al. 2011; Van de Bogert et al. 2012; Cavalcanti et al. 2016; Tonetta et al. 2016), some of the variability may be due to upstream stressors. Assessing variability in GPP and R at littoral sites near inflows may more accurately reflect land use changes happening in different sub-catchments. The contribution of individual streams to pelagic GPP and R is also likely dependent on water residence time near the inflow. Major storm events increase stream–lake connectivity, lowering water residence time in sub-basins, and delivering nutrients and DOC to the pelagic zone (Gallardo et al. 2012; Vachon and del Giorgio 2014; Zwart et al. 2017).

The littoral zone response to stream stressors is likely highly variable and context dependent. Though stream nutrient subsidies can stimulate littoral GPP and R, increased DOC or suspended solids may limit GPP by reducing water clarity (Kelly et al. 2018; Olson et al. 2020). High inflow DOC can cause light limitation and decrease GPP, but the median cold-stream inflow DOC concentration of 8.6 mg L^{-1} is at the lower end of concentrations where DOC-induced light limitation could be expected (Finstad et al. 2014; Thrane et al. 2014; Seekell et al. 2015). Therefore, there is continued potential for GPP stimulation with greater inflow DOC concentrations (Kelly et al. 2018) at the cold-stream littoral site. Elevated DOC and TP concentrations at our littoral sites may be particularly acute following precipitation events. The disproportionately high amount of dissolved organic and particulate matter transport that can occur during high stream flow events (Newbold et al. 1997) may create optimum conditions for the stream–lake transitional zone to function as an activated ecosystem control point (Bernhardt et al. 2017). Given the likely short water residence times in the littoral sites we selected, further work characterizing the dynamic nutrient, DOC, suspended sediment, and light conditions at these sites may be key in linking landscape drivers to lake responses in space and time.

Machine learning provides new insights on ecosystem metabolism

The identification of covariate relationships between littoral zone metabolism metrics and stream-related predictor variables was uniquely enabled by our ensemble machine learning approach. Metabolism datasets are typically noisy (Coloso et al. 2011), especially in oligotrophic lakes, since they capture multiple processes including photosynthesis, autotrophic and heterotrophic respiration, and integrate from a volume of water surrounding the sensor (Staehr et al.

2010). Further, many predictor variables for metabolism are collinear (Giling et al. 2017), for example, wind and water-column stability. Our ensemble machine learning approach used regularized regression as a base-learner model, which includes a penalty function to reduce conflated signals of collinear predictor variables (Boehmke and Greenwell 2020), with improvements over commonly used statistical methods, such as multiple regression in differentiating the effects of multiple collinear predictor variables (Lucas 2020).

The ensemble machine learning model exhibited systematic bias, resulting in a smaller range of GPP and R predictions than the range of metabolism model estimated GPP and R (Fig. S10). If the machine learning model were used in a predictive capacity, bias correction techniques should be applied to improve estimation of the tails of the distribution (Belitz and Stackelberg, 2021). In addition, unmeasured predictor variables may contribute to the observed variability in metabolism metrics. For example, both the concentrations and forms of nitrogen and phosphorus entering the stream–lake transitional zone are likely important, especially in oligotrophic lakes where co-limitation is common (Lewis et al. 2020). Since water temperature and season are known to affect metabolism, the lower range in machine learning predicted GPP and R may also be due to the comparatively narrow range of predictor and response variable magnitudes observed during the summer stratified period only (e.g., less than 10 °C range in epilimnetic water temperature, in comparison to the 25 °C range across the entire year; see Brentrup et al. 2021). Lake Sunapee's epilimnion is autotrophic in the summer and heterotrophic in the winter, where the dominant environmental predictors of GPP and R may vary with season (Brentrup et al. 2021). Similarly, the relative importance of stream–related drivers likely varies with season. The loss of thermal stratification in the lake would increase surface mixing with the cold stream; however, this effect may be dampened by the greater volume of receiving lake water and may be less relevant to nutrient availability than the vertical mixing within the lake. Further, food web processes, ranging from microbial community dynamics (Warnecke et al. 2005) and associated internal recycling of nutrients (Fenchel 2008) to fish consumer effects on nutrient processing (Vanni et al. 2013), likely also affect Lake Sunapee's GPP and R (Stewart et al. 2018).

Overall, the machine learning approach enabled us to conduct our exploratory analysis with novel predictor variables for highly noisy littoral metabolism, including collinear variables with no prior assumptions about the shape of the predictor–response relationship and existence of interactive effects. The machine learning approach would only be strengthened with the addition of predictor variables discussed above and extension of the study through space and time.

Implications for Lake Sunapee

The results of this study provide insight into how stream–lake connections may affect Lake Sunapee in the future. As of 2020, the Lake Sunapee region has already experienced a 1.4 ℃ increase in mean annual air temperature over pre-1970 annual averages (Ward et al. 2020) and is facing increasing development pressure (LSPA et al. 2020c). Cold-stream inflows to Lake Sunapee may warm through both a decreased canopy cover resulting in more direct warming of the stream from sunlight (Nelson et al. 2009; Kaushal et al. 2010) and an increase in the relative contribution of surface versus groundwater to streamflow (LeBlanc et al. 1997). As cold-stream inflows warm, they will mix more directly with the lake epilimnion, stimulating higher GPP and R near the stream–lake transitional zone. Though groundwater-based streams may be less sensitive to climate and land use effects than surface water streams (Luce et al. 2014), storm events have the potential to create short periods when cold, groundwater-based streams are dominated by surface flow (Shanley and Peters 1988). If cold streams warm faster than the lake epilimnion and enter the lake more as surface flow, the stream water and associated solutes will have a higher potential to stimulate surface littoral GPP and R. In particular, if the cold stream in our study warms faster than the lake epilimnion, the significantly higher stream nutrient and DOC concentrations (Fig. 2c–e) will further stimulate GPP and R. The increased productivity will potentially contribute to declines in localized water quality in the cove.

While changes in the depth of stream–lake mixing will change the location of peak GPP and R due to the nutrient-subsidy effect, the implications of different depths of mixing on whole-lake metabolism are complex. For example, stream water entering a lake at different depths results in different mixing of organic matter quality and may initiate a "priming effect," increasing the total availability of organic matter at the depth of mixing (Bouffard and Perga 2016). In addition, stratified lakes may have an autotrophic epilimnion and a heterotrophic hypolimnion during the summer (Staehr et al. 2012) and, thus, changes in the depth of stream–lake mixing during the summer may alter the relative contributions of each zone to overall lake metabolism.

Conclusions

Our study, examining epilimnetic GPP and R in two littoral and one pelagic location of a lake, demonstrates that metabolism near stream–lake transitional zones are uniquely related to stream variables. Therefore, GPP and R signals from the stream–lake transitional zones, within the context of this nutrient-limited lake showing early indications of eutrophication (Ward et al. 2020; Richardson et al. 2017), may provide a sentinel-of-the-sentinel signal. While littoral metabolism is generally higher than pelagic metabolism, directly linking these disproportionately high GPP and R rates in stream–lake transitional zones to stream-related drivers through time may bridge the inherent time-lag between upstream stressors and whole-lake responses. Estimates of whole-lake metabolism, particularly in nutrient-limited lakes should specifically include littoral locations near stream–lake transitional zones where GPP and R may be disproportionately high. Use of high-frequency dissolved oxygen sensors for examining sensitive ecosystem function metrics such as metabolism is developing quickly for management applications (Jankowski et al. 2021). As such, for lakes approaching a trophic state change, where mitigation of nutrient pollution is a key management goal (Jeppesen et al. 2010), monitoring lake water quality with more sensitive metrics of change (i.e., GPP, R) will be particularly helpful. Strategically estimating these metabolism metrics near inflows may help reveal the connection between upstream inputs and downstream lake processing, providing key insights into how lakes act as sentinels to changes in the surrounding catchment and informing strategic monitoring to prevent future declines in water quality.

resolution stream temperature, pressure, and estimated depth from transducers in streams in the Lake Sunapee watershed, New Hampshire, USA 2010–2018 ver 4. *Environmental Data Initiative,* https://doi.org/10.6073/pasta/9921cdeb6291b1e43251a3d776942c71. LSPA, K.C. Weathers, & B.G. Steele. 2020a. High-frequency Weather Data at Lake Sunapee, New Hampshire, USA 2007–2019 *Environmental Data Initiative,* https://doi.org/10.6073/pasta/698e9ffb0cdcda81ecf7188bff54445e. LSPA, K.C. Weathers, and B.G. Steele. 2020b. Lake Sunapee Instrumented Buoy: High Frequency Water Temperature and Dissolved Oxygen Data–2007–2019 ver 1. Environmental Data Initiative. https://doi.org/10.6073/pasta/70c41711d6199ac2758764ecfcb9815e. Ward NK, Brentrup JA, Johnson AE, Carey CC, Weathers KC, Fichter JR (2021) Underwater temperature, light, and dissolved oxygen data from 3 mini-buoys in Lake Sunapee, NH, USA from June–October 2018 ver 1. Environmental Data Initiative. https://doi.org/10.6073/pasta/b73f0b4a97ccef80249eb09b619979d5. Ward, N.K., J.A. Brentrup, D.C. Richardson, and C.C. Carey. 2022. Lake ecosystem metabolism estimates from 3 locations in Lake Sunapee, NH, USA during the summer stratified period from June to September 2018 ver 1. Environmental Data Initiative. https://doi.org/10.6073/pasta/c54eeb67934e3c2576fcec4402767747. Ward, N.K. 2021. nkward/LakeMetabolismML: Lake Metabolism Predictor Analysis (Version v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.5021337.

Declarations

Conflict of interest The authors declare no conflicts of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication All the co-authors consent to this manuscript submission.

References

Adrian R, O'Reilly CM, Zagarese H, Baines SB, Hessen DO, Keller W, Livingstone DM, Sommaruga R, Straile D, Van Donk E, Weyhenmeyer GA, Winder M (2009) Lakes as sentinels of climate change. Limnol Oceanogr 54(6 part 2):2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

Andersen MR, Sand-Jensen K, Woolway RI, Jones ID (2017) Profound daily vertical stratification and mixing in a small, shallow, wind-exposed lake with submerged macrophytes. Aquat Sci 79:395–406. https://doi.org/10.1007/s00027-016-0505-0

APHA (2005) Standard methods for the examination of water and wastewater, 21st edn. American Public Health Association, American Water Works Association, Water Environment Federation, Washington DC

Belitz K, Stackelberg PE (2021) Evaluation of six methods for correcting bias in estimates from ensemble tree machine learning regression models. Environ Model Softw 139:105006. https://doi.org/10.1016/j.envsoft.2021.105006

Bernhardt ES, Blaszczak JR, Ficken CD, Fork ML, Kaiser KE, Seybold EC (2017) Control points in ecosystems: moving beyond the hot spot hot moment concept. Ecosystems 20:665–682. https://doi.org/10.1007/s10021-016-0103-y

Boehmke B, Greenwell B (2020) Hands-on machine learning with R. CRC Press, Taylor & Francis Group, Boca Raton

Bouffard D, Perga ME (2016) Are flood-driven turbidity currents hot spots for priming effects in lakes? Biogeosciences 13:3573–3584. https://doi.org/10.5194/bg-13-3573-2016

Brenton RW, Arnett TL (1993) Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory; determination of dissolved organic carbon by UV-promoted persulfate oxidation and infrared spectrometry. U.S. Geological Survey, Denver

Brentrup JA, Richardson DC, Carey CC, Ward NK, Bruesewitz DA, Weathers KC (2021) Under-ice respiration rates shift the annual carbon cycle in the mixed layer of an oligotrophic lake from autotrophy to heterotrophy. Inland Waters 11(1):114–123. https://doi.org/10.1080/20442041.2020.1805261

Bruesewitz DA, Carey CC, Richardson DC, Weathers KC (2015) Under-ice thermal stratification dynamics of a large, deep lake revealed by high-frequency data. Limnol Oceanogr 60:347–359. https://doi.org/10.1002/lno.10014

Carey CC, Weathers KC, Ewing HA, Greer ML, Cottingham KL (2014a) Spatial and temporal variability in recruitment of the cyanobacterium *Gloeotrichia echinulata* in an oligotrophic lake. Freshw Sci 33:577–592. https://doi.org/10.1086/675734

Carey CC, Cottingham KL, Weathers KC, Brentrup JA, Ruppertsberger NM, Ewing HA, Hairston NG (2014b) Experimental blooms of the cyanobacterium *Gloeotrichia echinulata* increase phytoplankton biomass, richness and diversity in an oligotrophic lake. J Plankton Res 36(2):364–377. https://doi.org/10.1093/plankt/fbt105

Casas-Ruiz JP, Catalán N, Gómez-Gener L, von Skiller D, Obrador B, Kothawala DN, López P, Sabater S, Marcé R (2017) A tale of pipes and reactors: controls on the in-stream dynamics of dissolved organic matter in rivers. Limnol Oceanogr 62(S1):S85–S94. https://doi.org/10.1002/lno.10471

Catalán N, Marcé R, Kothawala DN, Tranvik LJ (2016) Organic carbon decomposition rates controlled by water retention time across inland waters. Nat Geosci 9:501–504. https://doi.org/10.1038/NGEO2720

Cavalcanti JR, da Motta-Marques D, Fragoso CR Jr (2016) Process-based modeling of shallow lake metabolism: Spatio-temporal variability and relative importance of individual processes. Ecol Model 323:28–40. https://doi.org/10.1016/j.ecolmodel.2015.11.010

Chmiel HE, Hofmann H, Sobek S, Efremova T, Pasche N (2020) Where does the river end? Drivers of spatiotemporal variability in CO_2 concentration and flux in the inflow area of a large boreal lake. Limnol Oceanogr 65:1161–1174. https://doi.org/10.1002/lno.11378

Coloso JJ, Cole JJ, Pace ML (2011) Difficulty in discerning drivers of lake ecosystem metabolism with high-frequency data. Ecosystems 14:935–948. https://doi.org/10.1007/s10021-011-9455-5

Cortés A, Fleenor WE, Wells MG, de Vicente I, Rueda FJ (2014) Pathways of river water to the surface layers of stratified reservoirs. Limnol Oceanogr 59(1):233–250. https://doi.org/10.4319/lo.2014.59.1.0233

Crisci C, Ghattas B, Perera G (2012) A review of supervised machine learning algorithms and their applications to ecological data. Ecol Model 240:113–122. https://doi.org/10.1016/j.ecolmodel.2012.03.001

Cyr H (2017) Winds and the distribution of nearshore phytoplankton in a stratified lake. Water Res 122:114–127. https://doi.org/10.1016/j.watres.2017.05.066

Dormann CF, Elith J, Bacher S, Buchman C, Carl G, Carré G, Garciá Marquéz JR, Gruber B, Lafourcade B, Leitão PJ, Münkemüller MC, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lauthenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

EPA (US Environmental Protection Agency) (1993) Method 353.2, Revision 2.0: Determination of Nitrate-Nitrite Nitrogen by automated colorimetry. https://www.epa.gov/sites/production/files/2015-08/documents/method_353-2_1993.pdf. Accessed 3 April 2021

Ewing HA, Weathers KC, Steele BG (2021) High resolution stream temperature, pressure, and estimated depth from transducers in six streams in the Lake Sunapee watershed, New Hampshire, USA 2010–2018. Environ Data Initiat. https://doi.org/10.6073/pasta/12a18d9d178b17d2ac333930f3efca9f(Accessed2021-03-16)

Fenchel T (2008) The microbial loop—25 years later. J Exp Mar Biol Ecol 366(1–2):99–103. https://doi.org/10.1016/j.jembe.2008.07.013

Finstad AG, Helland IP, Ugedal O, Hesthagen T, Hessen DO (2014) Unimodal response of fish yield to dissolved organic carbon. Ecol Lett 17:36–43. https://doi.org/10.1111/ele.12201

Gallardo B, Español C, Comin FA (2012) Aquatic metabolism short-term response to the flood pulse in a Mediterranean floodplain. Hydrobiologia 693:251–264. https://doi.org/10.1007/s10750-012-1126-9

Giling DP, Staehr PA, Grossart HP, Andersen MR, Boehrer B, Escot C, Evrendilek F, Gómez-Gener L, Honti M, Jones ID, Karakaya N, Laas A, Moreno-Ostos E, Rinke K, Scharfenberger U, Schmidt SR, Weber M, Woolway RI, Zwart JA, Obrador B (2017) Delving deeper: metabolic processes in the metalimnion of stratified lakes. Limnol Oceanogr 62:1288–1306. https://doi.org/10.1002/lno.10504

Hanson PC, Carpenter SR, Kimura N, Wu C, Cornelius SP, Kratz TK (2008) Evaluation of metabolism models for free-water dissolved oxygen methods in lakes. Limnol Oceanogr Methods 6(9):454–465. https://doi.org/10.4319/lom.2008.6.454

Hanson PC, Pace ML, Carpenter SR, Cole JJ, Stanley EH (2015) Integrating landscape carbon cycling: research needs for resolving organic carbon budgets of lakes. Ecosystems 18:363–375. https://doi.org/10.1007/s10021-014-9826-9

Hoellein TJ, Bruesewitz DA, Richardson DC (2013) Revisiting Odum (1956): a synthesis of aquatic ecosystem metabolism. Limnol Oceanogr 58(6):2089–2100. https://doi.org/10.4319/lo.2013.58.6.2089

Hotchkiss ER, Sadro S, Hanson PC (2018) Toward a more integrative perspective on carbon metabolism across lentic and lotic inland waters. Limnol Oceanogr Lett 3(3):57–63. https://doi.org/10.1002/lol2.10081

Imberger J, Hamblin PF (1982) Dynamics of lakes, reservoirs, and cooling ponds. Annu Rev Fluid Mech 14:153–187. https://doi.org/10.1146/annurev.fl.14.010182.001101

Jankowski KJ, Mejia FH, Blaszcak JR, Holtgrieve GW (2021) Aquatic ecosystem metabolism as a tool in environmental management. Wires Water (early View). https://doi.org/10.1002/wat2.1521

Jeppesen E, Moss B, Bennion H, Carvalho L, DeMeester L, Friberg N, Gessner MO, Lauridsen TL, May L, Meerhoff M, Olafsson JS, Soons MB, Verhoeven JTA (2010) Interaction of climate change and eutrophication. In: Kernan M, Battarbee R, Moss B (eds) Climate change impacts on freshwater ecosystems. Blackwell Publishing Ltd, Oxford, pp 119–151

Johengen TH, Biddanda BA, Cotner JB (2008) Stimulation of Lake Michigan plankton metabolism by sediment resuspension and river runoff. J Gt Lakes Res 34(2):213–227. https://doi.org/10.3394/0380-1330(2008)34[213:SOLMPM]2.0.CO;2

Johnson TR, Stefan HG (1988) Experimental study of density induced plunging into reservoirs and coastal regions. Report 254, St. Anthony Falls Hydraulics Laboratory, Minneapolis

Kaushal SS, Likens GE, Jaworski NA, Pace ML, Sides AM, Seekell D, Belt KT, Secor DH, Windgate RL (2010) Rising stream and river temperatures in the United States. Front Ecol Environ 8(9):461–466. https://doi.org/10.1890/090037

Kelly PT, Solomon CT, Zwart JA, Jones SE (2018) A framework for understanding variation in pelagic gross primary production of lake ecosystems. Ecosystems 21:1364–1376. https://doi.org/10.1007/s10021-018-0226-4

Krause S, Lewandowski J, Grimm NB, Hannah DM, Pinay G, McDonald K, Martí E, Argerich A, Pfister L, Klaus J, Battin T, Larned ST, Schelker J, Fleckenstein J, Schmidt C, Rivett MO, Watts G, Sabater F, Sorolla A, Turk V (2017) Ecohydrological interfaces as hotspots of ecosystem processes. Water Resour Res 53:6359–6376. https://doi.org/10.1002/2016WR019516

LeBlanc RT, Brown RD, FitzGibbon JE (1997) Modeling the effects of land use change on the water temperature in unregulated urban streams. J Environ Manage 49(4):445–469. https://doi.org/10.1006/jema.1996.0106

Lewis ASL, Kim BS, Edwards HL, Wander HL, Garfield CM, Murphy HE, Poulin ND, Princiotta SD, Rose KC, Taylor AE, Weathers KC, Wigdahl-Perry CR, Yokota K, Richardson DC, Bruesewitz DA (2020) Prevalence of phytoplankton limitation by both nitrogen and phosphorus related to nutrient stoichiometry, land use, and primary producer biomass across the northeastern United States. Inland Waters 10(1):42–50. https://doi.org/10.1080/20442041.2019.1664233

LSPA, Weathers KC, Steele BG (2020a) High-frequency Weather Data at Lake Sunapee, New Hampshire, USA 2007–2019. Environ Data Initiat. https://doi.org/10.6073/pasta/698e9ffb0cdcda81ecf7188bff54445e

LSPA, Weathers KC, Steele BG (2020b) Lake Sunapee instrumented buoy: high frequency water temperature and dissolved oxygen data—2007–2019 ver 1. Environ Data Initiat. https://doi.org/10.6073/pasta/70c41711d6199ac2758764ecfcb9815e

Lake Sunapee Protective Association, DK Water Resource Consulting, & Stone Environmental (2020c) Lake Sunapee watershed management plan. Sunapee (NH): Lake Sunapee Protective Association. https://static1.squarespace.com/static/54bfbf13e4b08839cc010a41/t/5ef3476843817b4f70ea3ac2/1593001836774/LS_2020c_WMP_Complete.pdf. Accessed 10 Mar 2021

Lucas TCD (2020) A translucent box: interpretable machine learning in ecology. Ecol Monogr 90(4):e01422. https://doi.org/10.1002/ecm.1422

Luce C, Staab B, Kramer M, Wenger S, Isaak D, McConnell C (2014) Sensitivity of summer stream temperatures to climate variability in the Pacific Northwest. Water Resour 50:3428–3443. https://doi.org/10.1002/2013WR014329

MacIntyre S, Sickman JO, Goldthwait SA, Kling GW (2006) Physical pathways of nutrient supply in a small, ultraoligotrophic arctic lake during summer stratification. Limnol Oceanogr 51(2):1107–1124. https://doi.org/10.4319/lo.2006.51.2.1107

Marcarelli AM, Baxter CV, Mineau MM, Hall RO (2011) Quantity and quality: unifying food web and ecosystem perspectives on the role of resource subsidies in freshwaters. Ecology 92(6):1215–1225. https://doi.org/10.1890/10-2240.1

Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. Anal Chim Acta 27:31–36. https://doi.org/10.1016/S0003-2670(00)88444-5

Nelson KC, Palmer MA, Pizzuto JE, Moglen GE, Angermeier PL, Hilderbrand RH, Dettinger M, Hayhoe K (2009) Forecasting the combined effects of urbanization and climate change on stream ecosystems: from impacts to management options. J Appl Ecol 46:154–163. https://doi.org/10.1111/j.1365-2664.2008.01599.x

Newbold JD, Bott TL, Kaplan LA, Sweeney BW, Vannote RL (1997) Organic Matter Dynamics in White Clay Creek, Pennsylvania, USA. J N Amer Benthol Soc 16(1):46–50. https://doi.org/10.2307/1468231

Olson CR, Solomon CT, Jones SE (2020) Shifting limitation of primary production: experimental support for a new model in lake ecosystems. Ecol Lett 23(12):1800–1808. https://doi.org/10.1111/ele.13606

Patton CJ, Kryskalla JR (2003) Methods of analysis by the U.S. Geological survey national water quality laboratory—evaluation of alkaline persulfate digestion as an alternative to Kjeldahl digestion for determination of total and dissolved nitrogen and phosphorus in water. U.S. Geological Survey, Denver

Prairie YT, Bird DF (1989) Some misconceptions about the spurious correlation problem in the ecological literature. Oecologia 81:285–288. https://doi.org/10.1007/BF00379817

R Core Team (2021) R: a language and environment for statistical computing: R Foundation for Statistical Computing, Vienna. www.R-project.org

Read JS, Hamilton DP, Jones ID, Muraoka K, Winslow LA, Kroiss R, Wu CH, Gaiser E (2011) Derivation of lake mixing and stratification indices from high-resolution lake buoy data. Environ Model Softw 26(11):1325–1336. https://doi.org/10.1016/j.envsoft.2011.05.006

Richardson DC, Carey CC, Bruesewitz DA, Weathers KC (2017) Intra- and inter-annual variability in metabolism in an oligotrophic lake. Aquat Sci 79:319–333. https://doi.org/10.1007/s00027-016-0499-7

Rose KC, Winslow LA, Read JS, Read EK, Solomon CT, Adrian R, Hanson PC (2014) Improving the precision of lake ecosystem metabolism estimates by identifying predictors of model uncertainty. Limnol Oceanogr-Meth 12(5):303–312. https://doi.org/10.4319/lom.2014.12.303

Rueda FJ, MacIntyre S (2009) Flow paths and spatial heterogeneity of stream inflows in a small multibasin lake. Limnol Oceanogr 54(6):2041–2057. https://doi.org/10.4319/lo.2009.54.6.2041

Rueda FJ, MacIntyre S (2010) Modelling the fate and transport of negatively buoyant storm-river water in small multi-basin lakes. Environ Model Softw 25(1):146–157. https://doi.org/10.1016/j.envsoft.2009.07.002

Sadro S, Melack JM, MacIntyre S (2011) Spatial and temporal variability in the ecosystem metabolism of a high-elevation lake: integrating benthic and pelagic habitats. Ecosystems 14:1123–1140. https://doi.org/10.1007/s10021-011-9471-5

Schloss JA (1990) Lake Sunapee nutrient budget study. University of New Hampshire, Durham. https://extension.unh.edu/resources/files/Resource005326_Rep7480.pdf. Accessed 23 Mar 2021.

Seekell DA, Lapierre J, Karlsson J (2015) Trade-offs between light and nutrient availability across gradients of dissolved organic carbon concentration in Swedish lakes: implications for patterns in primary production. Can J Fish Aquat Sci 72(11):1663–1671. https://doi.org/10.1139/cjfas-2015-0187

Shanley JB, Peters NE (1988) Preliminary observations of streamflow generation during storms in a forested Piedmont watershed using temperature as a tracer. J Contam Hydrol 3(2–4):349–365. https://doi.org/10.1016/0169-7722(88)90040-X

Simčič T, Germ M (2009) Organic matter degradation through respiration in littoral and pelagial including profundal zones of an oligotrophic lake assessed by electron transport system activity. Hydrobiologia 635:137–146. https://doi.org/10.1007/s10750-009-9906-6

Snortheim CA, Hanson PC, McMahon KD, Read JS, Carey CC, Dugan HA (2017) Meteorological drivers of hypolimnetic anoxia in a eutrophic, north temperature lake. Ecol Model 343:39–53. https://doi.org/10.1016/j.ecolmodel.2016.10.014

Solomon CT, Bruesewitz DA, Richardson DC, Rose KC, Van de Bogert MC, Hanson PC, Kratz TK, Larget B, Adrian R, Babin BL, Chiu CY, Hamilton DP, Gaiser EE, Hendricks S, Istvánovics V, Laas A, O'Donnell DM, Pace ML, Ryder E, Staehr PA, Torgersen T, Vanni MJ, Weathers KC, Zhu G (2013) Ecosystem respiration: Drivers of daily variability and background respiration in lakes around the globe. Limnol Oceanogr 58(3):849–866. https://doi.org/10.4319/lo.2013.58.3.0849

Staehr PA, Bade D, Van de Bogert MC, Koch GR, Willliamson C, Hanson P, Cole JJ, Kratz T (2010) Lake metabolism and the diel oxygen technique: state of the science. Limnol Oceanogr Methods 8(11):628–644. https://doi.org/10.4319/lom.2010.8.0628

Staehr AP, Christensen JPA, Batt RD, Read JS (2012) Ecosystem metabolism in a stratified lake. Limnol Oceanogr 57(5):1317–1330

Steele BG, Weathers KC, Lake Sunapee Protective Association (2021) Quality controlled in situ data from multiple locations in Lake Sunapee, NH, USA from the Lake Sunapee Protective Association's Long-term Monitoring Program, 1986–2020 (Version v2020.1). Zenodo. https://doi.org/10.5281/zenodo.4652076

Stewart SD, Hamilton DP, Baisden WT, Verburg P, Duggan IC (2018) The role of mobile consumers in lake nutrient cycles: a brief review. Hydobiologia 818:11–29. https://doi.org/10.1007/s10750-018-3603-2

Stockwell JD, Doubek JP, Adrian R, Anneville O, Carey CC, Carvalho L, De Senerpont Domis LN, Dur G, Frassl MA, Grossart HP, Ibelings BW, Lajeunesse MJ, Lewandowska AM, Llames ME, Matsuzaki SIS, Nodine ER, Nõges P, Patil VP, Pomati F, Rinke K, Rudstam LG, Rusak JA, Salmaso N, Seltmann CT, Straile D, Thackeray SJ, Thiery W, Urrutia-Cordero P, Venail P, Verburg P, Woolway RI, Zohary T, Andersen MR, Bhattacharya R, Hejzlar J, Janatian N, Kpodonu ATNK, Williamsom TJ, Wilson HL (2020) Storm impacts on phytoplankton community dynamics in lakes. Glob Change Biol 26(5):2756–2784. https://doi.org/10.1111/gcb.15033

Thrane J, Hessen DO, Andersen T (2014) The absorption of light in lakes: negative impact of dissolved organic carbon on primary productivity. Ecosystems 17:1040–1052. https://doi.org/10.1007/s10021-014-9776-2

Tonetta D, Staehr PA, Schmitt R, Petrucio MM (2016) Physical conditions driving the spatial and temporal variability in aquatic metabolism of a subtropical coastal lake. Limnologica 58:30–40. https://doi.org/10.1016/j.limno.2016.01.006

Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019) Machine learning algorithm validation with a limited sample size. PLoS ONE 14(11):e0224365. https://doi.org/10.1371/journal.pone.0224365

Vachon D, del Giorgio PA (2014) Whole-lake CO2 dynamics in response to storm events in two morphologically different lakes. Ecosystems 17:1338–1353. https://doi.org/10.1007/s10021-014-9799-8

Vadeboncoeur Y, Kalff J, Christoffersen K, Jeppesen E (2006) Substratum as a driver of variation in periphyton chlorophyll and productivity in lakes. J N Am Benthol Soc 25(2):379–392. https://doi.org/10.1899/0887-3593(2006)25[379:SAADOV]2.0.CO;2

Van de Bogert MC, Carpenter SR, Cole JJ, Pace ML (2007) Assessing pelagic and benthic metabolism using free water measurements. Limnol Oceanogr Methods 5(5):145–155. https://doi.org/10.4319/lom.2007.5.145

Van de Bogert MC, Bade DL, Carpenter SR, Cole JJ, Pace ML, Hanson PC, Langman OC (2012) Spatial heterogeneity strongly affects estimates of ecosystem metabolism in two north temperate lakes. Limnol Oceanogr 57(6):1689–1700. https://doi.org/10.4319/lo.2012.57.6.1689

Vanni MJ, Boros G, McIntyre PB (2013) When are fish sources vs. sinks of nutrients in lake ecosystems? Ecology 94(10):2195–2206. https://doi.org/10.1890/12-1559.1

Vincent WF, Gibbs MM, Spigel RH (1991) Eutrophication processes regulated by a plunging river inflow. Hydrobiologia 226:51–63. https://doi.org/10.1007/BF00007779

Ward NK, Steele BG, Weathers KC, Cottingham KL, Ewing HA, Hanson PC, Carey CC (2020) Differential responses of maximum versus median chlorophyll-*a* to air temperature and nutrient loads in an oligotrophic lake over 31 years. Water Resour Res. https://doi.org/10.1029/2020WR027296

Ward NK, Brentrup JA, Johnson AE, Carey CC, Weathers KC, Fichter JR (2021) Underwater temperature, light, and dissolved oxygen data from 3 mini-buoys in Lake Sunapee, NH, USA from June–October 2018 ver 1. Environ Data Initiat. https://doi.org/10.6073/pasta/b73f0b4a97ccef80249eb09b619979d5

Ward NK, Brentrup JA, Richardson DC, Carey CC (2022) Lake ecosystem metabolism estimates from 3 locations in Lake Sunapee, NH, USA during the summer stratified period from June to September 2018 ver 1. Environ Data Initiat. https://doi.org/10.6073/pasta/c54eeb67934e3c2576fcec4402767747

Ward NK (2021) nkward/LakeMetabolismML: Lake Metabolism Predictor Analysis (Version v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.5021337

Warnecke F, Sommaruga R, Sekar R, Hofer RS, Pernthaler J (2005) Abundances, identity, and growth state of Actinobacteria in mountain lakes of different UV transparency. Appl Environ Microbiol 71(9):5551–5559. https://doi.org/10.1128/AEM.71.9.5551-5559.2005

Wilkinson GM, Walter J, Fleck R, Pace ML (2020) Beyond the trends: the need to understand multiannual dynamics in aquatic ecosystems. Limnol Oceanogr Lett 5(4):281–286. https://doi.org/10.1002/lol2.10153

Williamson CE, Saros JE, Schindler DW (2009) Sentinels of change. Science 323:887–888. https://doi.org/10.1126/science.1169443

Winslow LA, Zwart JA, Batt RD, Dugan HA, Woolway RI, Corman JR, Hanson PC, Read JR (2016) LakeMetabolizer: and R package for estimating lake metabolism from free-water oxygen using diverse statistical models. Inland Waters 6(4):622–636. https://doi.org/10.1080/IW-6.4.883

Winslow LA, Read JR, Woolway RI, Brentrup JA, Leach T, Zwart J, Albers S, Collinge D (2019) rLakeAnalyzer: Lake Physics Tools. R package version 1.11.4.1. https://CRAN.R-project.org/package=rLakeAnalyzer

Xia Y, Mitchell K, Ek M, Sheffield J, Cosgrove B, Wood E, Luo L, Alonge C, Wei H, Meng J, Livneh B, Lettenmaier D, Koren V, Duan Q, Mo K, Fan Y, Mocko D (2012) Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J Geophys Res Atmos 117(D3):D03109. https://doi.org/10.1029/2011jd016048

Zwart JA, Sebestyen SD, Solomon CT, Jones SE (2017) The influence of hydrologic residence time on lake carbon cycling dynamics following extreme precipitation events. Ecosystems 20:1000–1014. https://doi.org/10.1007/s10021-016-0088-6