Supporting Information for

Anoxia decreases the magnitude of the carbon, nitrogen, and phosphorus sink in freshwaters in *Global Change Biology*

Cayelan C. Carey*, Paul C. Hanson, R. Quinn Thomas, Alexandra B. Gerling, Alexandria G. Hounshell, Abigail S. L. Lewis, Mary E. Lofton, Ryan P. McClure, Heather L. Wander, Whitney M. Woelmer, B. R. Niederlehner, Madeline E. Schreiber

*Corresponding author. Email: cayelan@vt.edu

This PDF file includes:

Supporting Information Text 1. Description of field sampling methods at Falling Creek Reservoir.

Overview

Throughout the REDOX experiment in 2013-2019, Falling Creek Reservoir (FCR) was intensively monitored for water temperature, dissolved oxygen, chemistry, and phytoplankton (chlorophyll-a) concentrations. The two major inflows into FCR were also monitored for discharge, temperature, dissolved oxygen, and chemistry. The frequency of monitoring varied by time of year: during the spring (March-May), field sampling occurred weekly to fortnightly; from May-October, sampling occurred 1-2 times per week; and from November-March, sampling occurred approximately once every month or every other month. Monitoring occurred at the deepest lacustrine site in FCR next to the dam (see Fig. 2 in main text) as well as at the primary inflow stream to FCR (Tunnel Branch) from 2013-2019. In 2019, a second smaller inflow was also monitored (Falling Creek). The reservoir's biogeochemistry was not routinely sampled prior to the onset of REDOX in 2013.

All monitoring data are available with associated metadata in the Environmental Data Initiative repository (Carey et al. 2019, Carey et al. 2020, Carey et al. 2021a, Carey et al. 2021b, Carey et al. 2021d, Carey et al. 2021e), and described below.

Lacustrine monitoring

At the lacustrine sampling site, high-frequency (4 Hz) depth profiles of water temperature, dissolved oxygen, and chlorophyll-a were collected with a SeaBird Conductivity, Temperature, and Depth (CTD) profiler (Sea-Bird Scientific, Bellevue, WA, USA) on each sampling day from 2013-2019 (Carey et al. 2021b). On the few sampling days when the CTD

was not available, we substituted a YSI handheld probe (ProPlus with Quattro cable or ProODO optical dissolved oxygen meter; YSI Inc., Yellow Springs, OH, USA) to measure depth profiles of temperature and dissolved oxygen (Carey et al. 2021e). We collected paired CTD and YSI depth profiles on 60 days during 2015-2019, yielding n=674 observations throughout the water column for which we had replicate temperature and dissolved oxygen measurements. A comparison of the CTD and YSI data indicate that the methods were quantitatively similar (temperature Spearman's rho = 0.98, mean bias = 0.48^{o}C; dissolved oxygen rho = 0.82, mean bias = 0.44 mg L^{-1}).

We also collected water samples for total and dissolved nutrient (nitrogen and phosphorus) and organic carbon analysis from the reservoir's water treatment extraction depths (0.1, 1.6, 2.8, 3.8, 5.0, 6.2, 8.0, and 9.0 m) using a 4-L Van Dorn sampler (Wildco, Yulee, FL, USA). Water was filtered through GF/F (0.7 um) filters into acid-washed 125 mL HDPE bottles and immediately frozen for nitrate, ammonium, dissolved reactive phosphorus, and dissolved organic carbon analysis (Carey et al. 2021d). Unfiltered water was also frozen in separate acid-washed 125 mL HDPE bottles for total nitrogen and total phosphorus analysis (Carey et al. 2021d), as well as for total organic carbon analysis on a subset of sampling days in 2014 (Carey et al. 2018). In 2014, water samples were collected from 0.1, 5, and 9 m for dissolved reactive silica analysis (Carey et al. 2020). We focused our analysis on organic C, rather than inorganic C, because of the important role of reservoirs in burying this pool in the global C cycle (Mendonça et al. 2017), and because previous work indicates that most terrestrial dissolved inorganic C loads are rapidly emitted to the atmosphere (McDonald et al. 2013). All laboratory analysis methods are described in Supporting Information Text 2.

Inflow monitoring

At Tunnel Branch, the primary inflow to FCR (Fig. 2 in main text), we measured discharge and water temperature every 15 minutes using an INW Aquistar PT2X pressure sensor (INW, Kirkland, Washington, USA) throughout the study period. From 15 May 2013 to 6 June 2019, the weir was rectangular before it was converted to a V-notch weir on 7 June 2019. Rating curves and equations for calculating discharge, as well as data quality checks on the weir conversion, can be found in the metadata for the discharge dataset, available in the Environmental Data Initiative repository (Carey et al. 2021a). On each monitoring day from 2013-2019, we also measured water temperature and dissolved oxygen using a YSI handheld probe and collected water samples for chemical analyses, as described above.

Monitoring at Falling Creek, the second largest inflow to FCR, began in February 2019. Falling Creek did not have a weir, so we measured discharge at this inflow using a flowmeter (either a model FP 111 propeller flowmeter from Forestry Suppliers, Inc., Jackson, MS, USA or a model Flo-Mate 2000656 instantaneous velocity meter from Marsh-McBirney, Frederick, MD, USA). We used the flowmeter to measure velocity at 0.1-m increments along a transect across the inflow stream. These measurements were subsequently used to calculate discharge following Gordon (2004). We periodically verified flowmeter measurements using both salt injection and velocity floats. Details regarding these ancillary measurements as well as all inflow discharge data are in the Environmental Data Initiative repository (Carey et al. 2019). We also measured water temperature and dissolved oxygen using a YSI probe and collected water samples for chemical analyses on each monitoring day at this inflow.

Reservoir water budget

We used our inflow monitoring data to estimate a water budget for FCR. We based our analysis of the water budget of the reservoir upon Munger et al. (2019), who calculated that the major stream inflow (Tunnel Branch) provided 56 ± 23% and 66 ± 31% (1 S.D.) of the total water input into the reservoir in 2014 and 2015, respectively. Precipitation contributed a relatively small proportion of the budget (3 ± 7% and 5 ± 13%, respectively), with all other water sources (the smaller second stream inflow, groundwater, and surficial runoff) summed together as unmeasured inflows and calculated from subtraction, given that Munger et al. (2019) had accurate daily measurements of the reservoir's volume from the dam operators. We used our 2019 measurements of the contribution of the second smaller stream (Falling Creek) to FCR (Carey et al. 2019) to calculate that it contributed on average 69% of the inflow of Tunnel Branch during that monitoring period.

Thus, given that Tunnel Branch contributed 56-66% of the reservoir's water inputs (Munger et al. 2019) and Falling Creek's inflow represented ~69% of Tunnel Branch (Carey et al. 2019), we estimate that Falling Creek contributed ~38-46% of the reservoir's water inputs. Altogether, assuming that precipitation contributed 3-5% of the water budget, that leaves ~0-3% of the reservoir's budget that could be due to groundwater and runoff. While approximate, these calculations help demonstrate that the vast majority of the reservoir's inputs (>97%) are likely attributable to the two stream inflows.

Sediment flux chambers

Benthic flux chambers were deployed in triplicate over the sediments in the hypolimnion of FCR during two 10-day experiments in summer 2018 (21 June – 2 July and 13 – 23 August) to measure diffusive solute fluxes across the sediment-water interface (Krueger et al. 2020). The

flux chambers isolated 64.86 L of hypolimnetic water and 0.27 m^2 of the hypolimnetic sediment surface. Each chamber contained an optical dissolved oxygen sensor (InsiteIG Model 31, Slidell, LA, USA) to measure temperature and dissolved oxygen. The sensors were connected to a data logger (Gantzer Water, Livingston, TX, USA), which collected data on 2-minute intervals.

During deployment, the chambers were slowly lowered from a boat through the water column to the bottom and were then flushed with hypolimnetic water for 90 – 120 minutes using a circulation pump until dissolved oxygen and temperature within the chamber stabilized. After deployment, the chambers were left undisturbed for 24 hours before water samples were collected from the chamber every 3 days for 10 days (Krueger et al. 2020). The chambers typically went anoxic within 2-3 days. During sampling, water within the chamber and tubing was slowly circulated for one minute before water samples were gently removed via tubing for dissolved organic carbon, ammonium, dissolved reactive phosphorus, dissolved iron, and dissolved manganese without disturbing the sediments (Krueger et al. 2020). Iron and manganese analytical methods and data are published in Krueger et al. (2020). Dissolved organic carbon, ammonium, and dissolved reactive phosphorus analytical methods are described in Supporting Information Text 2 and published in Carey et al. (2022b).

The water chemistry samples were used to calculate sediment release rates for each solute into the hypolimnion. Fluxes were calculated using the equation:

$$J = b \times (V/A) \qquad \text{(eqn. S1)}$$

Where J is the flux of the solute (mmol m^{-2} d^{-1}), b is the slope of the best fit line of the solute concentrations plotted over time (mmol m^{-2} d^{-1}), V is the volume of the flux chamber (64.86 L), and A is the surface area of the flux chamber (0.27 m^2).

Supporting Information Text 2. Laboratory water chemistry analysis detailed methods.

Overview and quality assurance/quality control procedures

Below, we describe the laboratory methods and quality control procedures we used to quantify carbon (dissolved organic carbon, total organic carbon), nitrogen (ammonia, nitrate, total nitrogen), phosphorus (dissolved reactive phosphorus and total phosphorus), and dissolved reactive silica in reservoir water samples.

For quality control and assurance of data, method detection limits, limits of quantitation, and long-term averages were calculated. Generally, method detection limits (MDL) were determined as a one-sided 99% confidence interval from repeated measurements of a low concentration standard (USEPA 1997). Limits of quantitation (LOQ) were calculated as 10 times the standard deviation (Currie 1968). These measures were calculated at least once each season. Long-term averages were calculated over all determinations within the study period and reported with a 95% confidence interval.

Values below the MDL were kept as reported by the instrument but flagged in the database as being below detection. Values below the LOQ were kept and not flagged. Negative values were set to zero and flagged in the database. If a sample was analyzed repeatedly and multiple valid measurements for an analyte were obtained, as might happen if a signal peak was improperly integrated for only one of several analytes reported by a particular instrument, the mean of the multiple valid measurements was calculated and reported. All chemical measurements used in the calibration and validation of the FCR GLM-AED model are published in the Environmental Data Initiative (EDI) repository (Carey et al. 2021d).

Carbon

Total and dissolved organic carbon concentrations were both determined by infrared absorbance after organic carbon was released from the samples as carbon dioxide by either heated persulfate digestion (APHA 2017e) or high temperature combustion (APHA 2017d). Long-term mean MDL and LOQ were 0.34 (0.11 to 0.56) and 1.13 (0.37 to 1.89) mg L^{-1}, respectively (calculated from n = 12 MDL determinations).

Nitrogen

Ammonia nitrogen (hereafter, ammonium) concentrations were determined colorimetrically using the phenate method on a Lachat QuikChem 8500 Flow Injection Analyzer (Lachat Instruments, Loveland, CO, USA; Lachat 2007a, APHA 2017a). We adopted a common modification and used sodium dichloroisocyanuric acid as the source of hypochlorite ion because of its longer shelf life (following Grasshoff and Johannsen 1972, Zhang et al. 1997). Long-term mean MDL and LOQ were 2.9 (2.2-3.7) and 9.2 (7.0-11.3) µg L^{-1}, respectively (calculated from n = 32 MDL determinations).

Nitrate plus nitrite nitrogen (hereafter, nitrate) concentrations were determined colorimetrically using cadmium reduction followed by the Griess reaction on a Lachat QuikChem 8500 Flow Injection Analyzer (Lynch 2007, APHA 2017b). Long-term mean MDL and LOQ were 1.8 (1.4 to 2.2) and 5.7 (4.4 to 7.1) µg L^{-1}, respectively (calculated from n = 41 MDL determinations).

Total nitrogen concentrations were determined after an alkaline persulfate digestion (Patton and Kryskalla 2003) at sub-boiling temperatures (Doyle et al. 2004, Huang and Zhang 2009), followed by a colorimetric determination of the resulting nitrate using cadmium reduction

and the Griess reaction. The method was modified by the addition of 0.25 M sodium hydroxide to the sample before color development as described by Egan (2013). Due to known contribution of reagents to nutrient burdens, concentrations were corrected by the mean of process blanks. Long-term mean MDL and LOQ were 14.2 (7.5 to 20.1) and 45.1 (23.8 to 66.5) µg L^{-1}, respectively (calculated from n = 53 MDL determinations).

Phosphorus

Dissolved reactive phosphorus concentrations were determined colorimetrically using the molybdenum blue method on a Lachat QuikChem 8500 Flow Injection Analyzer (Lachat 2007b, APHA 2017c). Long-term mean MDL and LOQ were 2.5 (2.0 to 3.0) and 7.6 (6.0 to 9.2) µg L^{-1}, respectively (calculated from n = 31 MDL determinations). For this analyte in particular, caution is warranted when interpreting potential differences in concentrations below the LOQ.

Total phosphorus concentrations were determined after an alkaline persulfate digestion (Patton and Kryskalla 2003) at sub-boiling temperatures (Zhang et al. 1997, Doyle et al. 2004), followed by a colorimetric determination of the resulting dissolved reactive phosphorus concentration using the molybdenum blue method. Due to known contribution of reagents to nutrient burdens, concentrations were corrected by the mean of process blanks. Long-term mean MDL and LOQ were 3.7 (2.9 to 4.5) and 11.7 (9.2 to 14.2) µg L^{-1}, respectively (calculated from n = 54 MDL determinations). As for total phosphorus, caution is warranted when interpreting potential differences in concentrations below the LOQ.

Silica

Dissolved reactive silica was determined colorimetrically using ammonium molybdate

followed by sodium sulfite reduction (Wetzel and Likens 2000) using a Shimadzu UV 1601 spectrophotometer (Shimadzu Corp., Kyoto, Japan). Dissolved reactive silica was only analyzed for a subset of depths in 2014, and a limit of detection (LOD) and limit of quantitation (LOQ) were estimated from the calibration curve (IUPAC 1997, Ranke 2018, Team 2020). The estimated LOD was 0.39 mg L^{-1} and the estimated LOQ was 0.70 mg L^{-1}.

Instrument overlap

Some analytical instruments were replaced or updated during the 2013-2019 study period. When transitioning between instruments, we conducted overlap studies (Table S1). Correspondences in concentrations determined from the old vs. new instruments were assessed with method comparison regression. We used Passing-Bablock regression (Passing and Bablok 1983), a non-parametric form of error-in-variables regression. Estimates of slope and intercept were determined and confidence intervals for the regression coefficients were quantified by bootstrapping (Manuilova and Schuetzenmeister 2014). Good correspondence yields a slope near 1 and an intercept near 0, and a 95% confidence interval that does not include these best-case values indicates a statistically significant difference. If an intercept was significantly different from 0, we compared the magnitude of this difference to the MDL. If the difference was within 2× of the MDL, the correspondence was judged to be acceptable for inclusion of both old and new instrument concentrations in subsequent data analyses. If a slope was significantly different from 1, we evaluated both the magnitude of the difference and plots of residuals vs. concentrations. If there was no clear pattern over increasing concentration, the correspondence was judged acceptable.

Supporting Information Text 3. Model description, driver data, configuration, calibration, scenarios, and goodness-of-fit statistics.

Model description

The General Lake Model coupled to Aquatic EcoDynamics modules (GLM-AED) simulates the dominant biogeochemical processes controlling freshwater oxygen and carbon, nitrogen, and phosphorus cycling (Farrell et al. 2020, Ward et al. 2020, Hipsey 2022). Water column oxygen dynamics are modeled as a function of atmospheric oxygen exchange, sediment oxygen demand, organic matter mineralization, chemical oxidation (e.g., nitrification), and phytoplankton photosynthesis and respiration. Carbon state variables in GLM-AED include methane (CH_4), dissolved inorganic carbon (DIC), two dissolved organic carbon (DOC) pools (recalcitrant and labile), particulate organic carbon (POC), and coarse particulate organic matter (CPOM). Carbon processes include sediment fluxes of CH_4, DOC, and DIC; CH_4 oxidation; mineralization of DOC; decomposition of POC and CPOM; and phytoplankton C fixation, respiration, excretion, and death. Nitrogen state variables in GLM-AED include ammonium (NH_4^+), nitrate (NO_3^-), recalcitrant and labile dissolved organic nitrogen (DON), and particulate organic nitrogen (PON). Nitrogen processes include sediment fluxes of NH_4^+, NO_3^-, and DON; mineralization of DON; decomposition of PON; nitrification; denitrification; anaerobic ammonium oxidation (anammox); dissimilatory nitrate reduction to ammonium (DNRA); phytoplankton uptake of NH_4^+ and NO_3^-; phytoplankton excretion of DON; and phytoplankton mortality, which affects PON. Phosphorus state variables include DRP, recalcitrant and labile dissolved organic phosphorus (DOP), and particulate organic phosphorus (POP). Phosphorus processes include sediment fluxes of DRP and DOP; mineralization of DOP; decomposition of POP; phytoplankton uptake of DRP; phytoplankton excretion of DOP; and phytoplankton

mortality, which affects POP. Total pools - i.e., TOC, TN, and TP - are calculated from summing all of their respective fractions, including phytoplankton C, N, and P pools.

Driver data

GLM-AED simulates the dominant biogeochemical processes controlling freshwater oxygen and C, N, and P cycling (Hipsey 2014, Farrell et al. 2020, Ward et al. 2020, Hipsey 2022) and requires three driver datasets: meteorological data; inflow stream data (which consists of discharge, water temperature, and chemistry); and outflow water discharge. We developed each of these driver datasets for Falling Creek Reservoir (FCR) for 15 May 2013 to 31 December 2019, as detailed below. All model configuration files and driver data are available in the Environmental Data Initiative repository (Carey et al. 2022a), and the R code (v.3.6.3) to generate the driver files are available in the Zenodo repository (Carey et al. 2022b).

Meteorological driver data: GLM-AED was forced with hourly meteorological data (air temperature, relative humidity, shortwave and longwave radiation, wind speed, and precipitation) from NASA's North American Land Data Assimilation System (NLDAS-2; Xia et al. 2012) for FCR.

Inflow stream driver data: We developed driver datasets of daily discharge and water temperature for the two primary surface streams entering into FCR, Falling Creek and Tunnel Branch (Fig. 2 in main text), using manually collected data and sensor observations from 2013-2019. The total discharge and temperature of Tunnel Branch, the largest stream that enters into FCR, were measured continuously throughout the study period at a weir with pressure and temperature sensors measuring every 15 minutes (Carey et al. 2021a).

Falling Creek, the smaller stream entering into FCR, was monitored less regularly than the larger inflow and did not have a weir (Supporting Information Text 1), so its daily total discharge was summed from separate models that calculated precipitation-driven flow and baseflow. First, we estimated daily precipitation-driven flow for both Falling Creek and Tunnel Branch using a simple hydrological model based on their delineated watershed areas and NLDAS-2 precipitation and air temperature following Ward et al. (2020); code available in Carey et al. (2022b). Second, we subtracted the modeled daily precipitation-driven flow from total discharge measured at the Tunnel Branch weir to estimate its daily baseflow. Third, we compared baseflow data collected manually in both streams using flowmeters and salt injection on n=19 days without any recent precipitation in 2019 (Carey et al. 2021a), and calculated the mean ratio of baseflow in Falling Creek to Tunnel Branch from those 19 days of data. Fourth, we then multiplied that ratio by the daily baseflow in Tunnel Branch to calculate daily baseflow in Falling Creek. Finally, we summed Falling Creek's daily baseflow and its daily precipitation-driven flow to calculate daily total discharge for that stream.

Water temperature was measured at an upstream littoral site in FCR near where Falling Creek entered the reservoir at approximately the same frequency as the lacustrine monitoring (Supporting Information Text 1). This littoral temperature dataset (Carey et al. 2021b) was linearly interpolated and used to calculate Falling Creek's daily inflow temperature (Carey et al. 2022b).

The inflow water chemistry driver data for Tunnel Branch and Falling Creek were determined from manual grab sample observations (Carey et al. 2021d). Both streams are well-mixed, so their dissolved oxygen concentrations were estimated assuming 100% saturation using the *rMR* package (Moulton 2018). Surface grab samples at Tunnel Branch for NO_3^-, NH_4^+, DRP,

DOC, TN, and TP were collected on approximately the same frequency as the reservoir water column chemistry samples (Supporting Information Text 1), and linearly interpolated to a daily time step.

GLM-AED requires additional solute driver data for inflows, including dissolved silica (Si); particulate organic fractions of C, N, and P; two dissolved organic fractions of C, N and P (labile and recalcitrant); dissolved CH_4; and DIC; which were estimated from intermittent grab samples or literature ratios to develop the daily inflow dataset. Specifically, following Wetzel (2001)'s synthesis of the distribution of dissolved, organic, and total fractions of C, N, and P in north temperate lakes similar to FCR, we assumed that: 1) 10% of the DOC, DON, and DOP pools was labile and 90% was recalcitrant; 2) POC concentrations were 10% of DOC concentrations; 3) DON concentrations were five times greater than PON concentrations; and 4) 30% of total organic P is in the DOP fraction. PON and POP concentrations were determined via subtraction, and DIC and dissolved Si concentrations were set at the median of long-term measurements (Munger et al. 2016, Carey et al. 2020).

Stream water chemistry was measured intermittently at Falling Creek, so its daily solute concentrations were calculated from multiplying Tunnel Branch's daily solute concentrations by ratios of the two stream's solutes measured on n=24 days in 2019 (Carey et al. 2021d).

We represented FCR's hypolimnetic oxygenation system (HOx) by a submerged inflow in the model. The mass of dissolved oxygen in the submerged inflow matched the mass of dissolved oxygen added in the reservoir as part of HOx operations during the field manipulation (Carey et al. 2022a, Carey et al. 2022b), but with a reduced inflow volume (compared to the volume of water pumped through the HOx system) and no other solutes. We simulated the HOx system in this way to recreate how it functions in the reservoir: the HOx system deployed in FCR

does not affect water balance or thermal structure because it returns the same volume of water it extracts from 8 m in the hypolimnion for oxygenation onshore back to 8 m in the reservoir without altering its temperature (Gerling et al. 2014). Consequently, because GLM-AED can simulate submerged inflows but not outflows, we reduced the inflow volume of the submerged stream to avoid altering the model's water budget. We did not add any solutes (other than dissolved oxygen) in this inflow to ensure that any effects of the HOx system on hypolimnetic chemistry were due to *in situ* reservoir processes, not inflow chemistry dynamics. The water temperature of this submerged inflow was set to observations of water temperature at 8 m, the depth at which the HOx injects the oxygenated water into the hypolimnion (Gerling et al. 2014).

Outflow water balance: The daily outflow discharge was estimated as the sum of the two surface streams' daily inflow discharge, as the reservoir did not exhibit large changes in water level during the 2013-2019 study and was managed to maintain full pond conditions (Munger et al. 2019). GLM-AED determines the physical and chemical properties of the outflow using the state of the modeled reservoir.

Model configuration

We simulated two sediment zones in the model, based on the bathymetry of FCR (Gerling et al. 2014), to represent separate epilimnetic and hypolimnetic sediment dynamics. The two sediment zones had different temperature and heat dynamics, as well as different sediment flux rates for oxygen, NH_4^+, NO_3^-, and DRP, following previous work in FCR (Gerling et al. 2016, Krueger et al. 2020, McClure et al. 2020). Phytoplankton were divided into three groups that represented the dominant taxa in FCR: cyanobacteria, diatoms, and green algae (Carey et al. 2021c).

The total model simulation period was 15 May 2013 to 31 December 2019, which was chosen because the weir on the largest stream inflow was deployed on 14 May 2013. GLM-AED was run on an hourly time step throughout the simulation period, which encompassed a wide range of meteorological conditions, inflow volumes, HOx operation levels, and resulting hypolimnetic oxygen concentrations (Carey et al. 2022a). Initial water column profiles of water temperature and water chemistry were chosen to represent typical conditions observed in the reservoir in May (Carey et al. 2022a).

Model calibration

We divided the total simulation period into calibration (15 May 2013 - 31 December 2018) and validation (1 January 2019-31 December 2019) periods for model verification. As GLM-AED is a 1-D model, we focused model calibration on the full depth profile at the deepest site at the reservoir, where the long-term monitoring data were collected (Fig. 2 in main text). Earlier studies indicate that oxygen dynamics at the deep site are representative of upstream sites in the reservoir (Gerling et al. 2014, McClure et al. 2018). The monitoring site is immediately adjacent to the reservoir outflow, thereby ensuring that the model accurately captured downstream export of C, N, and P.

We calibrated GLM-AED to observed conditions in a three-step approach, following (Ladwig et al. 2021). First, we conducted a global sensitivity analysis to identify the most important parameters for simulating water temperature, dissolved oxygen, NH_4^+, NO_3^-, DRP, and DOC following Morris (1991). Second, we calibrated the identified sensitive parameters (Table S2) using the covariance matrix adaptation evolution strategy (CMA-ES) for automated numerical optimization to minimize root mean square error (RMSE) between observations and

model output (Hansen 2016), using all sampling depths in the water column (0.1, 1.6, 2.8, 3.8, 5.0, 6.2, 8.0, and 9.0 m). Initial parameter values were set at model defaults (Hipsey et al. 2019, Hipsey 2022) and run for 1000 iterations. Third, since we found strong trade-offs in performance for NH_4^+ and NO_3^- automated calibrations, whereby the best parameters for one solute degraded the performance of the other, we manually chose parameter values that exhibited their best combined performance, following Ladwig et al. (2021). All model parameter values are available in the Environmental Data Initiative repository (Carey et al. 2022a).

Goodness-of-fit metrics

We calculated multiple goodness-of-fit metrics to assess the model's performance during the calibration period, the validation period, and the total simulation period for hypolimnetic water temperature, thermocline depth, dissolved oxygen, total and dissolved fractions of N and P, DOC, and water chemistry ratios. These metrics included RMSE, the coefficient of determination (R^2), percent bias, and normalized mean absolute error (NMAE), which were chosen because they were the four most common goodness-of-fit metrics used in 328 freshwater modeling studies reviewed by Soares and Calijuri (2021). We followed the approach described by Soares and Calijuri (2021) as most commonly used for calculating the goodness-of-fit metrics to ensure that we were appropriately benchmarking our model fit; specifically, we compared model output from the whole water column to monthly field observations.

Model scenarios

We examined the effects of two different oxygen scenarios on the calibrated GLM-AED model: one in which the model was forced with a high level of experimental oxygenation to keep

the hypolimnion oxic throughout the summer thermally stratified period (15 May - 15 October) over the 2013-2019 simulation period, and one in which zero oxygen was added to the hypolimnion during 2013-2019, so hypolimnetic anoxia quickly set up after the onset of summer thermal stratification each year. The high level of experimental oxygenation in the oxic scenario matched the level of maximum HOx operation in FCR (Carey et al. 2022a). These scenarios were instantiated in the model by modifying the concentration of oxygen in the submerged inflow oxygen driver file. All other driver data (meteorology, surface stream inflows, outflow) were held constant.

Parameter sensitivity analysis

To examine how the uncertainty of our model outputs was affected by the model parameterization, we conducted an additional sensitivity analysis in which we doubled and halved the calibrated values of highly sensitive parameters for DOC, NH_4^+, NO_3^-, and DRP using a one-step-at-a-time (OAT) approach (following Brett et al. 2016). We then re-calculated the summer hypolimnetic concentrations of DOC, NH_4^+, NO_3^-, and DRP in the anoxic and oxic model scenarios for each variable and compared anoxic and oxic concentrations with paired t-tests.

We found that while modeled concentrations of DOC, NH_4^+, NO_3^-, and DRP changed in response to parameter value changes, the direction and magnitude of the results fundamentally stayed the same (Fig. S1). Regardless of the parameter values, anoxia significantly increased summer DOC, NH_4^+, and DRP, and decreased NO_3^-, with statistically significant differences between anoxic and oxic scenarios.

Supporting Information Text 4. Empirical data, comparison of model output and observations, and biogeochemical rates.

Aggregated among the July 15 – October 1 summer days with the least oxygenation (2018, 2019), the median observed hypolimnetic DOC, NH_4^+, and DRP concentrations were 2.0, 6.9, and 1.3× higher than in summers with the highest oxygenation (2016, 2017; Fig. S2), respectively. Following the patterns exhibited by the dissolved fractions, median observed hypolimnetic TN and TP concentrations were both 2.4× higher in the low vs. high oxygenation years (Fig. S2). Conversely, median observed hypolimnetic NO_3^- concentrations were 5× lower in summers with low oxygenation than summers with high oxygenation (Fig. S2).

GLM-AED was able to reproduce observed physical and chemical dynamics when comparing model output and empirical data for the full water column (Table 1), the hypolimnion (the focus of this study; Fig. 3), as well as for the epilimnion (Fig. S3) during the seven-year REDOX field manipulation. In addition to the state variables reported in the main text, the time series comparison of the modeled and observed focal hypolimnetic (9 m) stoichiometric ratios are presented in Fig. S4.

The biogeochemical rates presented in Fig. 7 in the text show that the sediment fluxes of DOC, NH_4^+, and DRP dominate the responses of C, N, and P cycling to changes in oxygen availability in the reservoir. To better visualize the role of the other fluxes, we present a modified version of Fig. 7 without the sediment fluxes included (Fig. S5).

Supporting Information Text 5. Statistical analyses to compare anoxic vs. oxic model scenario output.

In the oxic model scenario, oxygen was added into the hypolimnion of Falling Creek Reservoir throughout the thermally stratified period during each summer of the seven-year study. In the anoxic scenario, no oxygen was added to the hypolimnion, resulting in prolonged hypolimnetic anoxia each summer. We compared the median summer hypolimnetic (9 m) concentrations (Table S3), hypolimnetic molar ratios (Table S4), and reservoir retention of carbon, nitrogen, and phosphorus (Table S5) between the two scenarios with paired t-tests. Daily values were aggregated to summer medians, which were not temporally autocorrelated. We also compared annual rates of sediment burial of particulate organic carbon, nitrogen, and phosphorus between the two scenarios (Table S6).

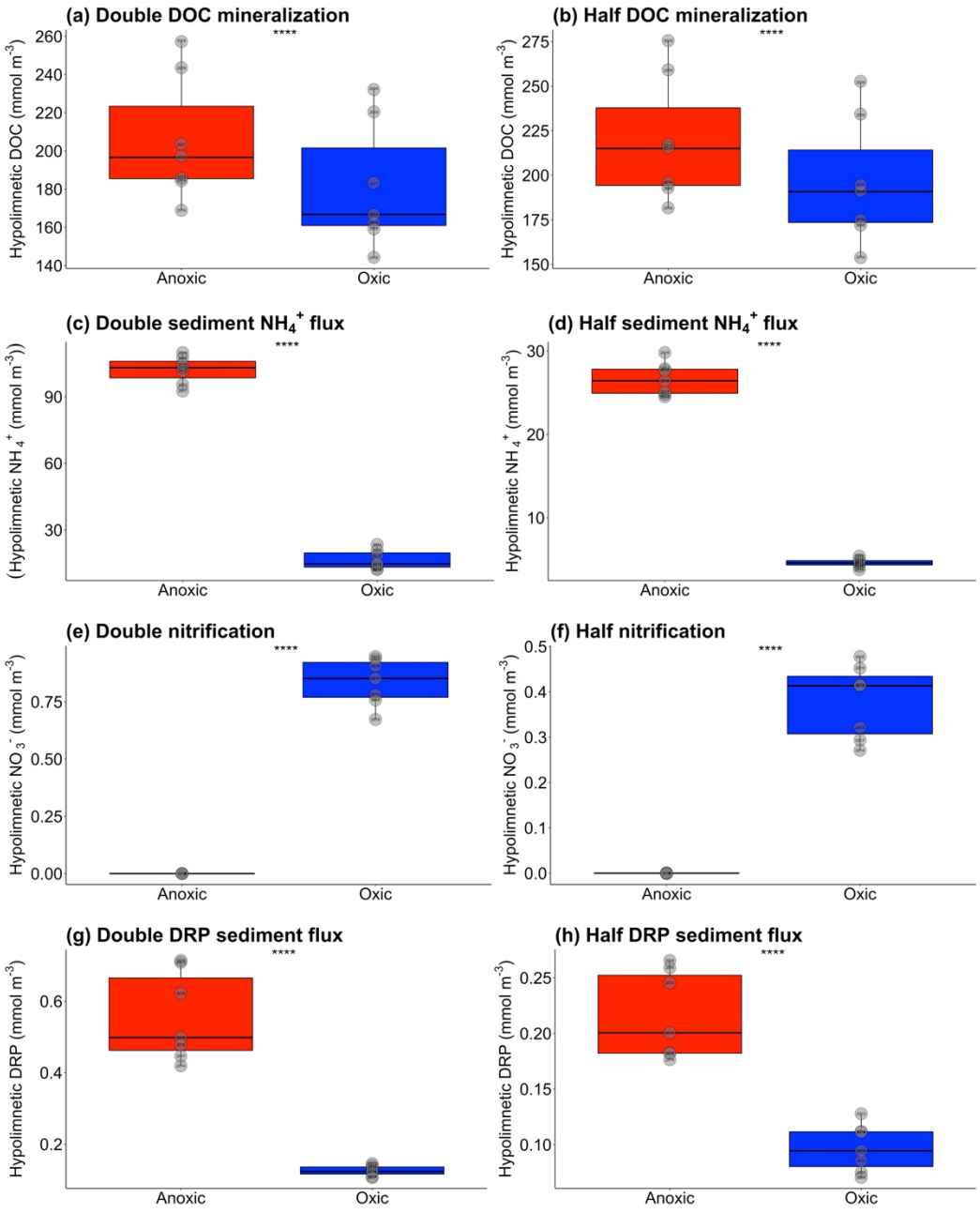

Fig. S1. Anoxic vs. oxic model scenario results were robust to model parameterization. Doubling and halving sensitive parameters for the mineralization rate of dissolved organic carbon (DOC; a,b), sediment flux rate of ammonium (NH_4^+; c,d), nitrification rate of nitrate (NO_3^-; e,f), and sediment flux rate of dissolved reactive phosphorus (DRP; g,h) did not alter the overall patterns observed in the anoxic (red) and oxic (blue) scenarios for median hypolimnetic (9 m) concentrations calculated with the calibrated parameters in Fig. 5 in the main text. Concentrations represent the medians from each summer stratified period (July 15 - October 1) during all years of this study. The **** denotes that the difference between the median summer anoxic and oxic scenario concentrations was highly statistically significant (all paired t-tests $p \leq 0.0001$). Note varying y-axes among panels.

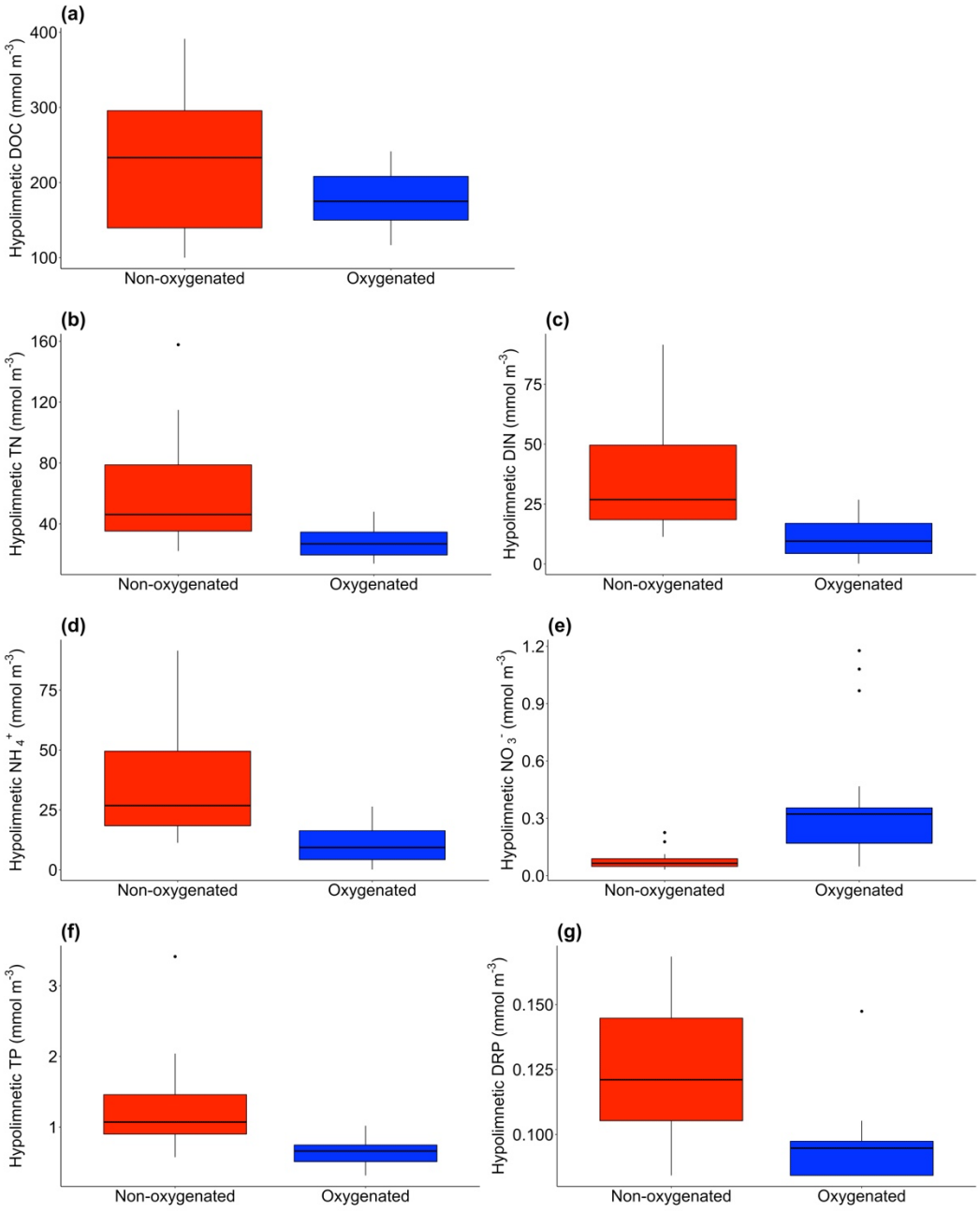

Fig. S2. Empirical data from non-oxygenated vs. oxygenated summer days mirror the anoxic vs. oxic model scenario results. We pooled all observed hypolimnetic (9 m) C, N, and P samples from when the HOx was deactivated during the two summers with the least oxygenation (2018, 2019) ("Non-oxygenated") and compared them with concentrations measured from when the HOx was activated during the two summers with the most continuous oxygenation ("Oxygenated"; 2016, 2017). Similar to Fig. 5 model scenario results, oxygenation altered bottom-water concentrations of dissolved organic carbon (DOC; a), total nitrogen (TN; b), dissolved inorganic nitrogen (DIN; c), ammonium (NH_4^+; d), nitrate (NO_3^-; e), total phosphorus (TP; f), and dissolved reactive phosphorus (DRP; g) concentrations during Falling Creek Reservoir's stratified period (July 15 - October 1). Note varying y-axes among panels.

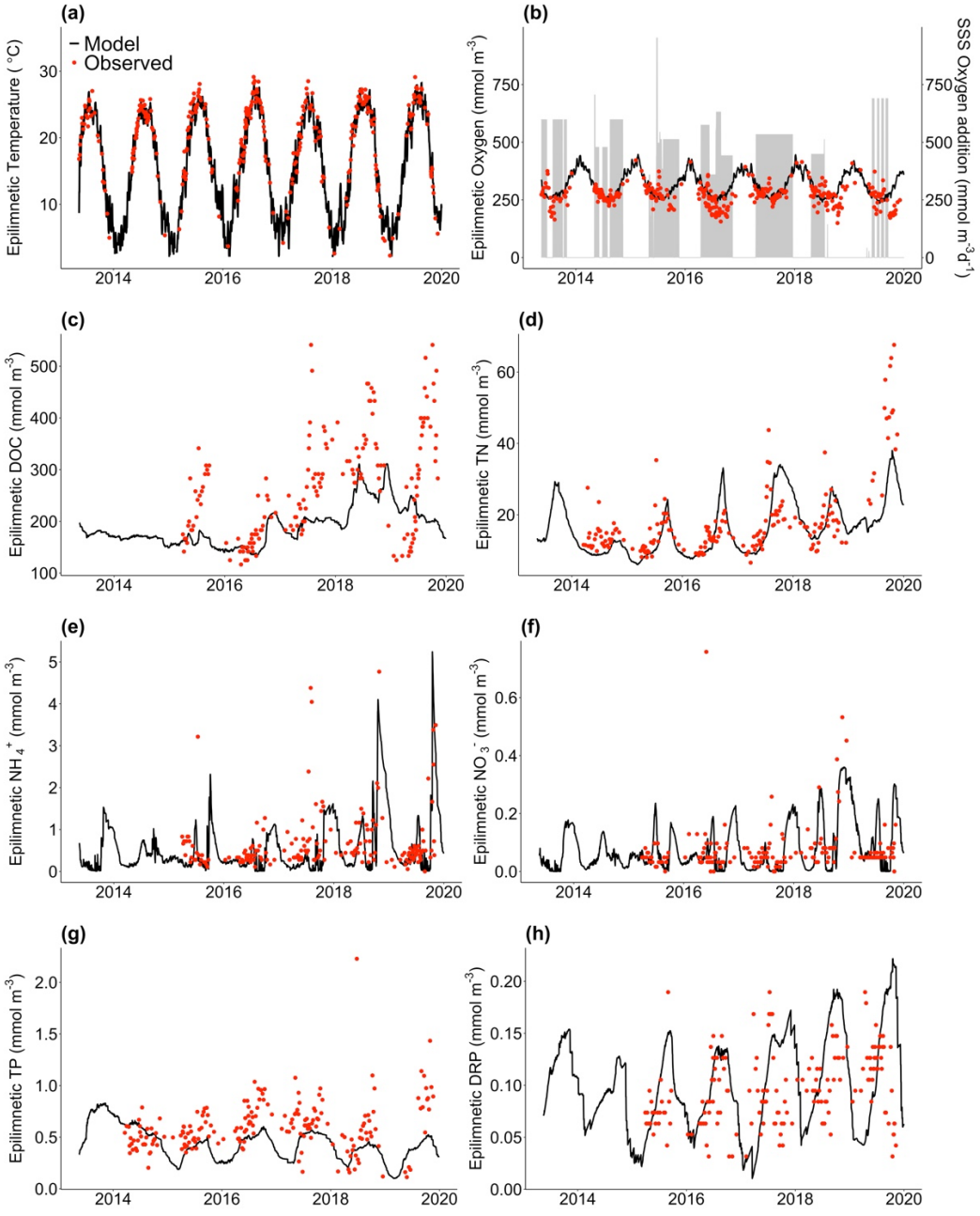

Fig. S3. Modeled vs. observed epilimnetic dynamics in Falling Creek Reservoir. Comparison of modeled (black line) vs. observed (red points) epilimnetic water temperature (a), dissolved oxygen (b), dissolved organic carbon (DOC; c), total nitrogen (TN; d), ammonium (NH_4^+; e), nitrate (NO_3^-; f), total phosphorus (TP; g), and dissolved reactive phosphorus (DRP; h). Epilimnetic water temperature and dissolved oxygen data are from 1 m depth while the water chemistry concentrations are from 1.6 m depth. The grey shaded areas in panel (b) represent the periods and addition rates of oxygen injection into the hypolimnion from the hypolimnetic oxygenation system (HOx) during the seven-year field manipulation. Note varying y-axes among panels.

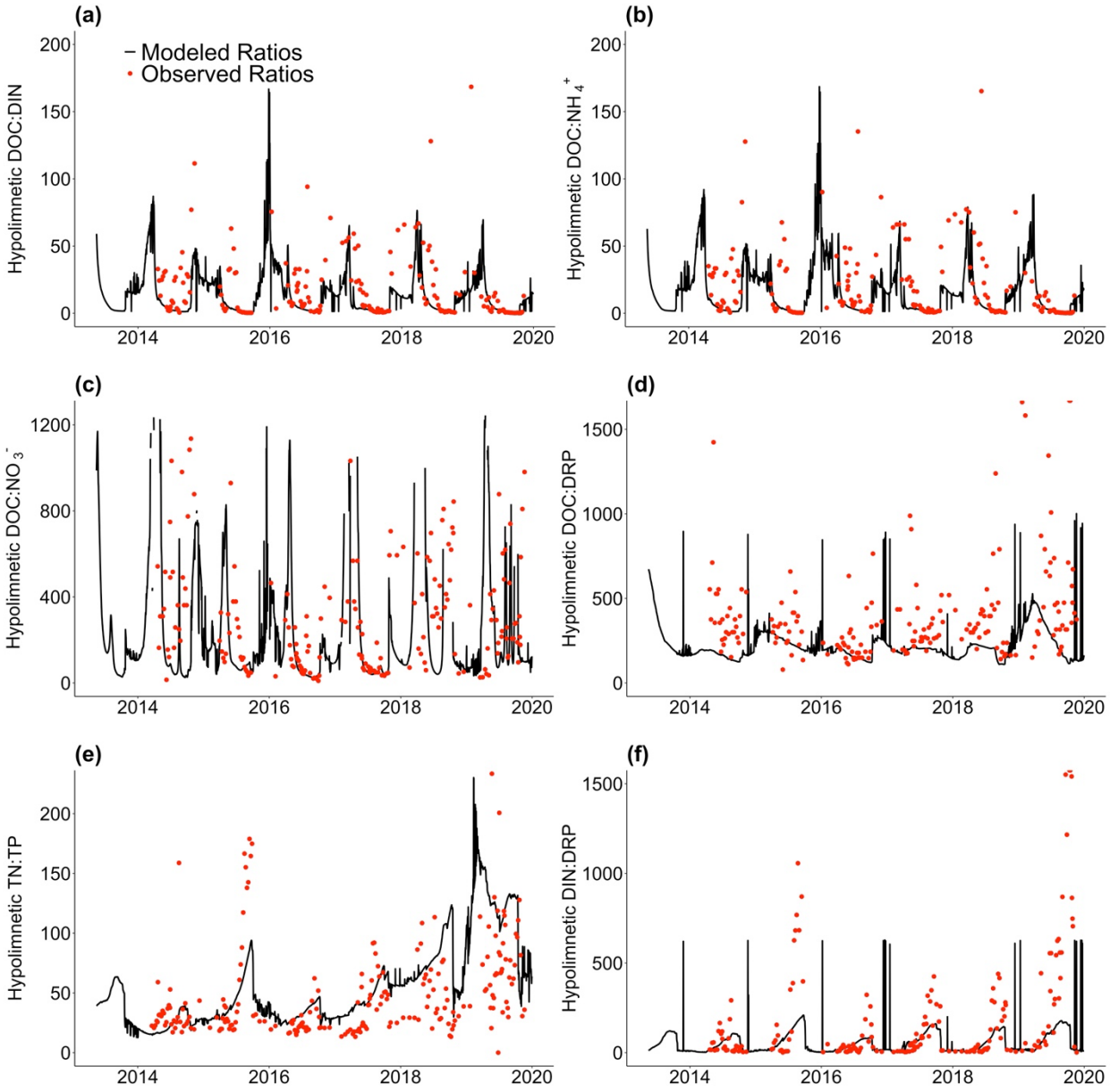

Fig. S4. Modeled vs. observed hypolimnetic stoichiometry in Falling Creek Reservoir. Comparison of modeled (black line) vs. observed (red points) hypolimnetic (9 m) water chemistry molar ratios of dissolved organic carbon to dissolved inorganic nitrogen (DOC:DIN; a), DOC to ammonium (DOC:NH_4^+; b), DOC to nitrate (DOC: NO_3^-; c), DOC to dissolved reactive phosphorus (DOC:DRP; d), total nitrogen to total phosphorus (TN:TP; e), and DIN to DRP (DIN:DRP; f). Note the varying y-axes among panels. Because nitrate concentrations were often below method detection limits, resulting in very large erroneous ratios, only ratios <1250 are presented in panel (c).

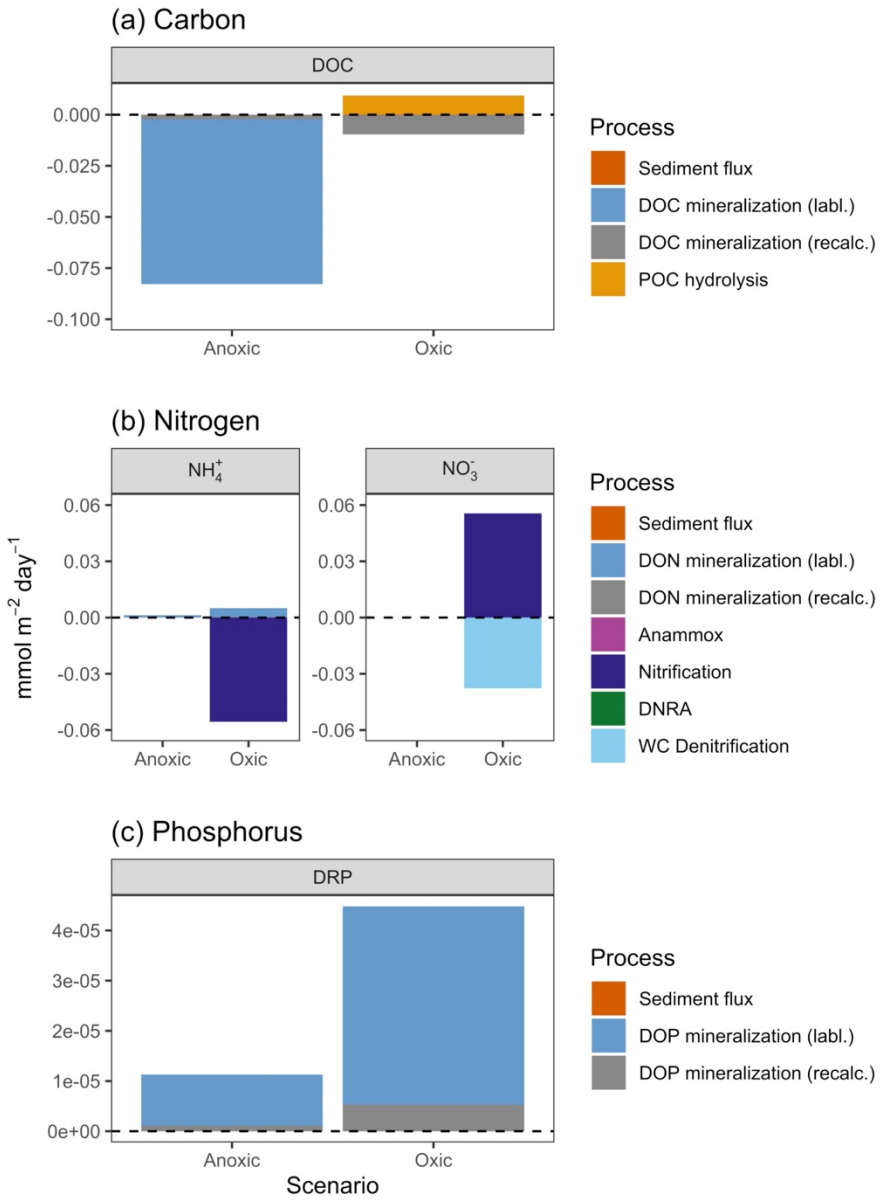

Fig. S5. The rates of processes altering the biogeochemical cycling of dissolved carbon, nitrogen, and phosphorus in response to anoxia, with sediment fluxes excluded. Comparison of the other biogeochemical processes altering dissolved pools of carbon (dissolved organic carbon, DOC; a), nitrogen (ammonium, NH_4^+, and nitrate, NO_3^-; b), and phosphorus (dissolved reactive phosphorus, DRP; c) in the hypolimnion of Falling Creek Reservoir under anoxic vs. oxic model scenarios. Rates shown represent the median contribution of each process to hypolimnetic concentrations of DOC, NH_4^+, NO_3^-, and DRP during Falling Creek Reservoir's summer stratified period (July 15 - October 1) for all years of this study, with the sediment fluxes of all four analytes removed. Positive rates indicate that the process increased hypolimnetic concentrations; negative rates indicate that the process decreased hypolimnetic concentrations. Mineralization is shown separately for both labile (labl.) and recalcitrant (recalc.) dissolved organic pools. Note the varying y-axes among panels and that some rates are so small – even when the sediment fluxes are removed from the figure – that they are not visible.

Table S1. Overview of results of overlap studies between old and new analytical chemistry instruments for each biogeochemical analyte. LOQ denotes limit of quantitation and MDL denotes method detection limit.

Analyte	Old instrument	New instrument	Timing of transition	Number of samples in overlap study	Results of overlap study
Total and dissolved organic carbon	OI Analytical Model 1010 Total Organic Carbon Analyzer (a heated persulfate digestion instrument)	Elementar vario TOC select (a high temperature combustion instrument)	January 2017	302	Only 4% of the observations were below the LOQ. When all samples were included, the Passing-Babcock regression slope was 1.22 (1.17-1.28), the intercept was 0.06 (-0.04 to 0.16), and the Pearson correlation coefficient was 0.967. The slope above 1 suggests that the new method using high temperature combustion may have greater recoveries than the old method using heated persulfate and that this difference may increase with concentration. Plots of residuals over concentration showed no pattern. We chose not to correct concentrations determined before the instrument transition to a modeled equivalence to concentrations determined by the new instrument (e.g., Newell and Morrison 1993) and instead evaluated potential discontinuities over time during that period with particular caution.
Ammonium	Lachat QuikChem 8500 Series 1 Flow Injection Analyzer	Lachat QuikChem 8500 Series 2 Flow Injection Analyzer	Fall 2018	359	We found 24.5% of observations were below the LOQ. When all data were included, the Passing-Babcock regression slope was 1.02 (1.01-1.02), the intercept was -0.48 (-0.80 to 0.06), and the Pearson correlation coefficient was 0.997. While the slope was significantly different from 1, this difference was small. Plots of residuals showed no patterns. When restricted to concentrations above the LOQ, the analysis was largely unchanged. Passing-Babcock regression slope was 1.02 (1.01-1.02), the intercept was -0.54 (-1.00 to 0.25), and the Pearson correlation coefficient was 0.997.
Nitrate	Lachat QuikChem 8500 Series 1 Flow Injection Analyzer	Lachat QuikChem 8500 Series 2 Flow Injection Analyzer	Fall 2018	354	More than half the observations (58.2%) were below the LOQ. When all samples were included the Passing-Babcock regression slope was 1.01 (0.99-1.04), the intercept was -1.05 (-1.26 to -0.84) and the Pearson correlation coefficient was 0.981. While the intercept was significantly different from zero,

					differences between instruments were smaller than the long-term MDL (1.05 and 1.7, respectively). When restricted to concentrations above the LOQ, the analysis was largely unchanged. Passing-Babcock regression slope was 1.02 (1.00-1.06), the intercept was -1.29 (-2.00 to -0.58), and the Pearson correlation coefficient was 0.978.
Total nitrogen	Lachat QuikChem 8500 Series 1 Flow Injection Analyzer	Lachat QuikChem 8500 Series 2 Flow Injection Analyzer	Fall 2018	326	Only one observation was below the LOQ (0.3%). The Passing-Babcock regression slope was 0.98 (0.94-1.02), the intercept was 18.03 (9.01-28.42), and the Pearson correlation coefficient was 0.982. While the intercept was significantly different from zero, the magnitude of the difference between instruments was close to the long-term MDL (18.03 and 14.2 $\mu g\ L^{-1}$ respectively).
Dissolved reactive phosphorus	Lachat QuikChem 8500 Series 1 Flow Injection Analyzer	Lachat QuikChem 8500 Series 2 Flow Injection Analyzer	Fall 2018	356	More than half the observations (64.9%) were below the LOQ. When all samples were included, the Passing-Babcock regression slope was 0.76 (0.71-0.80), the intercept was 0.27 (-0.22 to 0.63), and the Pearson correlation coefficient was 0.958. The method comparison plot shows a flattening of the slope at concentration below the LOQ. Plots of residuals suggest a U-shape with an increase at the highest and lowest concentrations. When restricted to concentrations above the LOQ, Passing-Babcock regression slope was 0.99 (0.92-1.06), the intercept was -3.39 (-4.79 to -2.12), and the Pearson correlation coefficient was 0.966. Plots of residuals over concentration were randomly distributed around 0. While the intercept was significantly different from zero, the magnitude of the difference between instruments was close to the long-term MDL (3.39 and 2.5, respectively). For this analyte in particular, caution is warranted when interpreting potential differences in concentration below the LOQ.
Total phosphorus	Lachat QuikChem 8500 Series 1 Flow Injection Analyzer	Lachat QuikChem 8500 Series 2 Flow Injection Analyzer	Fall 2018	336	More than half the observations were below the LOQ (56.0%). The Passing-Babcock regression slope was 0.95 (0.92-0.98), the intercept was 3.24 (2.78-3.72), and the Pearson correlation coefficient was 0.954. When restricted to concentrations above the LOQ, Passing-Babcock regression slope was 1.05 (1.01-1.09), the intercept was 0.78 (-0.39 to

					1.82), and the Pearson correlation coefficient was 0.960. While the slope was significantly different from 1, this difference was small. Plots of residuals showed no patterns. For this analyte in particular, caution is warranted when interpreting potential differences in concentration below the LOQ.

Table S2. GLM-AED parameters that were identified as highly sensitive for the focal state variables in the Falling Creek Reservoir model. The focal state variables were water temperature, dissolved oxygen, dissolved organic carbon (DOC), ammonium (NH_4^+), nitrate (NO_3^-), and dissolved reactive phosphorus (DRP). The focal parameters that were doubled and halved in the model scenario sensitivity analysis (Supporting Information Text 3) are shown in bold, which were chosen because of their importance in modeling DOC, NH_4^+, NO_3^- and DRP.

State variables	**Calibrated parameters**
Temperature	Coef_mix_hyp, sw_factor, lw_factor, ch, sed_temp_mean, sed_temp_amplitude, sed_temp_peak_doy
Dissolved oxygen	Fsed_oxy, Ksed_oxy
DOC	**Kdom_minerl**, Kpom_hydrol
NH_4^+	**Fsed_amm**, Ksed_amm, Rnitrif, Knitrif, theta_nitrif, Kdnra_oxy, Kanmx_amm
NO_3^-	**Rnitrif**, Fsed_nit
DRP	**Fsed_frp**, Ksed_frp

Table S3. Paired t-test results comparing median summer stratified period (July 15-October 1) hypolimnetic (9 m) concentrations in anoxic vs. oxic model scenarios for Falling Creek Reservoir during the seven years of the study. The variables compared were total organic carbon (TOC), dissolved organic carbon (DOC), total nitrogen (TN), dissolved inorganic nitrogen (DIN), ammonium (NH_4^+), nitrate (NO_3^-), total phosphorus (TP), and dissolved reactive phosphorus (DRP). df denotes degrees of freedom.

Variable	**t-value**	**df**	**p-value**
TOC	18.35	6	p<0.0001
DOC	16.08	6	p<0.0001
TN	24.61	6	p<0.0001
DIN	23.93	6	p<0.0001
NH_4^+	12.99	6	p<0.0001
NO_3^-	-9.34	6	p<0.0001
TP	9.79	6	p<0.0001
DRP	18.35	6	p<0.0001

Table S4. Paired t-test results comparing median summer stratified period (July 15 - October 1) total and dissolved molar ratios of hypolimnetic (9 m) carbon, nitrogen, and phosphorus concentrations for Falling Creek Reservoir during the seven years of the study. The variables compared included the ratios of total organic carbon to total nitrogen (TOC:TN), TOC to total phosphorus (TOC:TP), dissolved organic carbon to dissolved inorganic nitrogen (DOC:DIN), DOC to ammonium (DOC:NH_4^+), DOC to nitrate (DOC:NO_3^-), DOC to dissolved reactive phosphorus (DOC:DRP), TN:TP, and DIN:DRP. df denotes degrees of freedom. Because NO_3^- concentrations functionally went to zero during anoxia, DOC:NO_3^- could not be calculated and hence there are no t-test results for this ratio.

Variable	**t-value**	**df**	**p-value**
TOC:TN	-16.14	6	p<0.0001
TOC:TP	-3.08	6	0.02
DOC:DIN	-12.36	6	p<0.0001
DOC:NH_4^+	-11.76	6	p<0.0001
DOC:NO_3^-	.	.	.
DOC:DRP	-6.69	6	0.0005
TN:TP	6.45	6	0.0007
DIN:DRP	4.36	6	0.005

Table S5. Paired t-test results comparing median summer stratified period (July 15-October 1) retention of carbon, nitrogen, and phosphorus entering into Falling Creek Reservoir in anoxic vs. oxic model scenarios during the seven years of the study. The variables compared were total organic carbon (TOC), dissolved organic carbon (DOC), total nitrogen (TN), dissolved inorganic nitrogen (DIN), ammonium (NH_4^+), nitrate (NO_3^-), total phosphorus (TP), and dissolved reactive phosphorus (DRP). df denotes degrees of freedom.

Variable	**t-value**	**df**	**p-value**
TOC	12.80	6	p<0.0001
DOC	12.92	6	p<0.0001
TN	16.44	6	p<0.0001
DIN	9.40	6	p<0.0001
NH_4^+	6.19	6	0.0008
NO_3^-	-4.51	6	0.004
TP	4.72	6	0.003
DRP	8.83	6	0.0001

Table S6. Mean (±1 S.D.) rates of annual sediment burial (g m^{-2} yr^{-1}) of particulate organic carbon (POC), particulate organic nitrogen (PON), and particulate organic phosphorus (POP) in Falling Creek Reservoir in anoxic vs. oxic model scenarios of the seven years of the study, with paired t-test results comparing the anoxic vs. oxic rates among years. df denotes degrees of freedom. Data from 2013 were excluded from this analysis because the study started in May of that year.

Variable	**Anoxic scenario annual burial rate**	**Oxic scenario annual burial rate**	**t-value**	**df**	**p-value**
POC	0.95 (±0.08)	0.90 (±0.08)	-4.70	5	0.005
PON	0.27 (±0.05)	0.24 (±0.04)	-4.33	5	0.007
POP	0.023 (±0.004)	0.022 (±0.004)	-2.85	5	0.035

References

APHA. 2017a. Method 4500-NH3 H. Nitrogen (ammonia) flow injection analysis. Standard Methods for the Examination of Water and Wastewater, 23rd Ed. American Public Health Association, American Water Works Association, Washington, DC.

APHA. 2017b. Method 4500-NO3 I. Nitrogen (nitrate) cadmium reduction flow injection method. Standard Methods for the Examination of Water and Wastewater, 23rd Ed. American Public Health Association, American Water Works Association, Washington, DC.

APHA. 2017c. Method 4500-P G. Flow injection analysis for orthophosphate. Standard Methods for the Examination of Water and Wastewater, 23rd Ed. American Public Health Association, American Water Works Association, Washington, DC.

APHA. 2017d. Method 5310B. Total organic carbon (TOC): high-temperature combustion method. Standard Methods for the Examination of Water and Wastewater, 23rd Ed. American Public Health Association, American Water Works Association, Washington, DC.

APHA. 2017e. Method 5310C. Total organic carbon (TOC): persulfate-ultraviolet or heated-persulfate oxidation method. Standard Methods for the Examination of Water and Wastewater, 23rd Ed. American Public Health Association, American Water Works Association, Washington, DC.

Brett, M. T., S. K. Ahopelto, H. K. Brown, B. E. Brynestad, T. W. Butcher, E. E. Coba, C. A. Curtis, J. T. Dara, K. B. Doeden, K. R. Evans, L. Fan, J. D. Finley, N. J. Garguilo, S. M. Gebreeyesus, M. K. Goodman, K. W. Gray, C. Grinnell, K. L. Gross, B. R. E. Hite, A. J. Jones, P. T. Kenyon, A. M. Klock, R. E. Koshy, A. M. Lawler, M. Lu, L. Martinkosky, J.

R. Miller-Schulze, Q. T. N. Nguyen, E. R. Runde, J. M. Stultz, S. Wang, F. P. White, C. H. Wilson, A. S. Wong, S. Y. Wu, P. G. Wurden, T. R. Young, and G. B. Arhonditsis 2016. The modeled and observed response of Lake Spokane hypolimnetic dissolved oxygen concentrations to phosphorus inputs. Lake and Reservoir Management **32**:246-258.

Carey, C. C., J. P. Doubek, R. P. McClure, and P. C. Hanson. 2018. Oxygen dynamics control the burial of organic carbon in a eutrophic reservoir. Limnology and Oceanography Letters **3**:293-301.

Carey, C. C., J. P. Doubek, J. H. Wynne, and B. R. Niederlehner. 2020. Dissolved silica time series for Beaverdam Reservoir, Carvins Cove Reservoir, Claytor Lake, Falling Creek Reservoir, Gatewood Reservoir, Smith Mountain Lake, and Spring Hollow Reservoir in southwestern Virginia, USA during 2014 ver 1. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/353aa34bedfd68800c12275633811341

Carey, C. C., A. G. Hounshell, M. E. Lofton, B. Birgand, B. J. Bookout, R. S. Corrigan, A. B. Gerling, R. P. McClure, and W. M. Woelmer. 2021a. Discharge time series for the primary inflow tributary entering Falling Creek Reservoir, Vinton, Virginia, USA 2013-2021 ver 7. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/8d22a432aac5560b0f45aa1b21ae4746

Carey, C. C., A. S. Lewis, R. P. McClure, A. B. Gerling, S. Chen, A. Das, J. P. Doubek, D. W. Howard, M. E. Lofton, K. D. Hamre, and H. L. Wander. 2021b. Time series of high-frequency profiles of depth, temperature, dissolved oxygen, conductivity, specific conductivity, chlorophyll a, turbidity, pH, oxidation-reduction potential, photosynthetic active radiation, and descent rate for Beaverdam Reservoir, Carvins Cove Reservoir,

Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in Southwestern Virginia, USA 2013-2020 ver 11. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/5448f9d415fd09e0090a46b9d4020ccc

Carey, C. C., M. E. Lofton, W. M. Woelmer, K. D. Hamre, A. Breef-Pilz, J. P. Doubek, and R. P. McClure. 2021c. Time-series of high-frequency profiles of fluorescence-based phytoplankton spectral groups in Beaverdam Reservoir, Carvins Cove Reservoir, Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in southwestern Virginia, USA 2014-2020 ver 5. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/54d4bd2fee1e52e36e2b0f230912d2da

Carey, C. C., R. Q. Thomas, and P. C. Hanson. 2022a. General Lake Model-Aquatic EcoDynamics model parameter set for Falling Creek Reservoir, Vinton, Virginia, USA 2013-2019 ver 1. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/9f7d037d9a133076a0a0d123941c6396

Carey, C. C., R. Q. Thomas, R. P. McClure, A. G. Hounshell, W. M. Woelmer, H. L. Wander, and A. S. L. Lewis. 2022b. CareyLabVT/FCR-GLM: FCR GLM-AED model, data, and code for the paper "Anoxia decreases the magnitude of the carbon, nitrogen, and phosphorus sink in freshwaters" in Global Change Biology by Carey et al., v1.2. Zenodo repository. https://doi.org/10.5281/zenodo.6520742

Carey, C. C., H. L. Wander, W. M. Woelmer, M. E. Lofton, A. Breef-Pilz, J. P. Doubek, A. B. Gerling, A. G. Hounshell, R. P. McClure, and B. R. Niederlehner. 2021d. Water chemistry time series for Beaverdam Reservoir, Carvins Cove Reservoir, Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in southwestern Virginia,

USA 2013-2020 ver 8. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/8d83ef7ec202eca9192e3da6dd34a4e0

Carey, C. C., W. M. Woelmer, J. T. Maze, and A. G. Hounshell. 2019. Manually-collected discharge data for multiple inflow tributaries entering Falling Creek Reservoir and Beaverdam Reservoir, Vinton, Virginia, USA in 2019 ver 4. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/4d8e7b7bedbc6507b307ba2d5f2cf9a2

Carey, C. C., J. H. Wynne, H. L. Wander, R. P. McClure, K. J. Farrell, A. Breef-Pilz, J. P. Doubek, A. B. Gerling, K. D. Hamre, A. G. Hounshell, A. S. Lewis, M. E. Lofton, and W. M. Woelmer. 2021e. Secchi depth data and discrete depth profiles of photosynthetically active radiation, temperature, dissolved oxygen, and pH for Beaverdam Reservoir, Carvins Cove Reservoir, Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in southwestern Virginia, USA 2013-2020, ver 8. Environmental Data Initiative repository. https://doi.org/10.6073/pasta/3e9f27971e353c8a80840b5e99a67d0c

Currie, L. A. 1968. Limits for qualitative detection and quantitative determination. Application to radiochemistry. Analytical Chemistry **40**:586-593.

Doyle, A., M. N. Weintraub, and J. P. Schimel. 2004. Persulfate digestion and simultaneous colorimetric analysis of carbon and nitrogen in soil extracts. Soil Science Society of America Journal **68**:669-676.

Egan, L. 2013. Determination of total nitrogen in manual persulfate digests. Low flow method. QuikChem Method 10-107-04-4-C. Lachat Instruments Loveland, CO.

Farrell, K. J., N. K. Ward, A. I. Krinos, P. C. Hanson, V. Daneshmand, R. J. Figueiredo, and C. C. Carey. 2020. Ecosystem-scale nutrient cycling responses to increasing air temperatures vary with lake trophic state. Ecological Modelling **430**:109134.

Gerling, A., R. Browne, P. Gantzer, M. Mobley, J. Little, and C. Carey. 2014. First report of the successful operation of a side stream supersaturation hypolimnetic oxygenation system in a eutrophic, shallow reservoir. Water Research **67**:129-143.

Gerling, A. B., Z. W. Munger, J. P. Doubek, K. D. Hamre, P. A. Gantzer, J. C. Little, and C. C. Carey. 2016. Whole-catchment manipulations of internal and external loading reveal the sensitivity of a century-old reservoir to hypoxia. Ecosystems **19**:555-571.

Gordon, N. D., T.A. McMahon, B.L. Finlayson. 2004. Stream Hydrology: an Introduction for Ecologists, 2nd edition. John Wiley & Sons, Inc., New York.

Grasshoff, K., and H. Johannsen. 1972. A new sensitive and direct method for the automatic determination of ammonia in sea water. ICES Journal of Marine Science **34**:516-521.

Hansen, N. 2016. The CMA evolution strategy: A tutorial. ArXiv **160400772**.

Hipsey, M. R. 2022. Modelling Aquatic Eco-Dynamics: Overview of the AED modular simulation platform (v0.9.0). Zenodo. https://doi.org/10.5281/zenodo.6516222

Hipsey, M. R., L. C. Bruce, C. Boon, B. Busch, C. C. Carey, D. P. Hamilton, P. C. Hanson, J. S. Read, E. de Sousa, M. Weber, and L. A. Winslow. 2019. A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). Geoscience Model Development **12**:473-523.

Hipsey, M. R., Bruce, L.C., Hamilton, D.P. 2014. GLM - General Lake Model: Model overview and user information. The University of Western Australia, Perth, Australia.

Huang, X.-L., and J.-Z. Zhang. 2009. Neutral persulfate digestion at sub-boiling temperature in an oven for total dissolved phosphorus determination in natural waters. Talanta **78**:1129-1135.

IUPAC. 1997. International Union of Pure and Applied Chemistry 1997 IUPAC Compendium of Analytical Nomenclature, Definitive Rules, 3rd edition. Blackwell Science, London.

Kara, E. L., P. Hanson, D. Hamilton, M. R. Hipsey, K. D. McMahon, J. S. Read, L. Winslow, J. Dedrick, K. Rose, C. C. Carey, S. Bertilsson, D. da Motta Marques, L. Beversdorf, T. Miller, C. Wu, Y.-F. Hsieh, E. Gaiser, and T. Kratz. 2012. Time-scale dependence in numerical simulations: Assessment of physical, chemical, and biological predictions in a stratified lake at temporal scales of hours to months. Environmental Modelling & Software **35**:104-121.

Krueger, K. M., C. E. Vavrus, M. E. Lofton, R. P. McClure, P. Gantzer, C. C. Carey, and M. E. Schreiber. 2020. Iron and manganese fluxes across the sediment-water interface in a drinking water reservoir. Water Research **182**:116003.

Lachat. 2007a. Determination of ammonia by flow injection analysis. Low flow method. QuikChem Method 10-107-06-1-J. Lachat Instruments Loveland, CO.

Lachat. 2007b. Determination of orthophosphate by flow injection analysis. Low flow method. QuikChem Method 10-115-10-1-Q. Lachat Instruments Loveland, CO.

Ladwig, R., P. C. Hanson, H. A. Dugan, C. C. Carey, Y. Zhang, L. Shu, C. J. Duffy, and K. M. Cobourn. 2021. Lake thermal structure drives interannual variability in summer anoxia dynamics in a eutrophic lake over 37 years. Hydrology and Earth System Sciences **25**:1009-1032.

Lynch, D. 2007. Determination of nitrate/nitrite by flow injection analysis. Low flow method. QuikChem Method 10-107-04-1-L. Lachat Instruments, Loveland, CO.

Manuilova, E., and A. Schuetzenmeister. 2014. mcr: Method Comparison. Regression. R package version 1.2.1.

McClure, R., K. Hamre, B. Niederlehner, Z. Munger, S. Chen, M. Lofton, M. Schreiber, and C. Carey. 2018. Metalimnetic oxygen minima alter the vertical profiles of carbon dioxide and methane in a managed freshwater reservoir. The Science of the Total Environment **636**:610-620.

McClure, R. P., M. E. Lofton, S. Chen, K. M. Krueger, J. C. Little, and C. C. Carey. 2020. The magnitude and drivers of methane ebullition and diffusion vary on a longitudinal gradient in a small freshwater reservoir. Journal of Geophysical Research: Biogeosciences **125**:e2019JG005205.

McDonald, C. P., E. G. Stets, R. G. Striegl, and D. Butman. 2013. Inorganic carbon loading as a primary driver of dissolved carbon dioxide concentrations in the lakes and reservoirs of the contiguous United States. Global Biogeochemical Cycles **27**:285-295.

Mendonça, R., R. A. Müller, D. Clow, C. Verpoorter, P. Raymond, L. J. Tranvik, and S. Sobek. 2017. Organic carbon burial in global lakes and reservoirs. Nature Communications **8**:1694.

Morris, M. D. 1991. Factorial sampling plans for preliminary computational experiments. Technometrics **33**:161-174.

Moulton, T. L. 2018. rMR: Calculating Metabolic Rates and Critical Tensions. R package version 1.1.0.

Munger, Z., C. Carey, A. Gerling, K. Hamre, J. Doubek, S. Klepatzki, R. McClure, and M. Schreiber. 2016. Effectiveness of hypolimnetic oxygenation for preventing accumulation of Fe and Mn in a drinking water reservoir. Water Research **106**:1-14.

Munger, Z. W., C. C. Carey, A. B. Gerling, J. P. Doubek, K. D. Hamre, R. P. McClure, and M. E. Schreiber. 2019. Oxygenation and hydrologic controls on iron and manganese mass budgets in a drinking-water reservoir. Lake and Reservoir Management **35**:277-291.

Newell, A. D., and M. L. Morrison. 1993. Use of overlap studies to evaluate method changes in water chemistry protocols. Water, Air, and Soil Pollution **67**:433-456.

Passing, H., and Bablok. 1983. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. Journal of Clinical Chemistry and Clinical Biochemistry **21**:709-720.

Patton, C. J., and J. R. Kryskalla. 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. USGS Water-Resources Investigations Report 2003-4174.

R Core Team. 2020. R: A language and environment for statistical computing. Version 3.6.3. R Foundation for Statistical Computing, Vienna, Austria.

Ranke, J. 2018. chemCal: Calibration functions for analytical chemistry. R package version 0.2.1.

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

Soares, L. M. V., and M. d. C. Calijuri. 2021. Deterministic modelling of freshwater lakes and reservoirs: Current trends and recent progress. Environmental Modelling & Software **144**:105143.

USEPA. 1997. Appendix B to Part 136 - Definition and Procedure for the Determination of the Method Detection Limit - Revision 1.11. U.S. Environmental Protection Agency.

Ward, N. K., B. G. Steele, K. C. Weathers, K. L. Cottingham, H. A. Ewing, P. C. Hanson, and C. C. Carey. 2020. Differential responses of maximum versus median chlorophyll-a to air temperature and nutrient loads in an oligotrophic lake over 31 years. Water Resources Research **56**:e2020WR027296.

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems. 3rd edition. Academic Press, New York.

Wetzel, R. G., and G. E. Likens. 2000. Exercise 7. Inorganic nutrients: Nitrogen, phosphorus, and other nutrients. Pages 85-111 *in* Limnological Analyses. 3rd ed. Springer, New York.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, D. Lettenmaier, V. Koren, Q. Duan, K. Mo, Y. Fan, and D. Mocko. 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. Journal of Geophysical Research: Atmospheres **117**:D03109.

Zhang, J. Z., P. B. Ortner, C. J. Fischer, and L. D. Moore. 1997. Method 349.0 Determination of ammonia in estuarine and coastal waters by gas segmented flow continous flow colorimetric analysis. Methods for determination of chemical substances in marine and

estuarine environmental matrices, 2nd ed. Environmental Protection Agency, Washington, D.C.