LIMNOLOGY and OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 9999, 2024, 1–15

doi: 10.1002/lno.12650

Does polymixis complicate prediction of high-frequency dissolved oxygen in lakes and reservoirs?

Caleb J. Robbins,[1,2*,a] **Jeffrey M. Sadler**,[3] **Dennis Trolle**,[4] **Anders Nielsen**,[4] **Nicole D. Wagner**,[5] **J. Thad Scott**[1]

[1]Department of Biology, Center for Reservoir and Aquatic Systems Research, Baylor University, Waco, Texas, USA
[2]Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA
[3]Department of Biosystems and Agricultural Engineering, Oklahoma State University, Stillwater, Oklahoma, USA
[4]WaterITech, Skanderborg, Denmark
[5]Department of Biological Sciences, Oakland University, Rochester, Michigan, USA

Abstract

As lake and reservoir ecosystems transition across major environmental regimes (e.g., mixing regime) resulting from anthropogenic change, setting predictive expectations is imperative. We tested the hypothesis that (dissolved) oxygen is more predictable in monomictic reservoirs that thermally stratify throughout the summer (warm) season compared to polymictic reservoirs that stratify intermittently. Using two-hourly vertical profiles of oxygen, we compared daily-aggregated errors of oxygen predictions from random forests across and within two monomictic and two polymictic reservoirs in the south-central (subtropical) USA. Although one monomictic reservoir was typically more predictable than the polymictic reservoirs, the hypereutrophic, small monomictic reservoir had less predictable oxygen patterns potentially related to rapid oxygen cycling and intrusions of oxygenated waters in the hypolimnion without mixing. Daily mixing did not relate strongly to model errors. Water temperature, depth, and wind were the most important predictors, but were not clearly related to season or mixing. Lastly, we compared multiple model types (regression, neural network, and process-based) in one polymictic reservoir to test how our interpretations of oxygen predictability were sensitive to model type, finding that the models generally agreed; however, the process-based model poorly predicted oxygen in the middle of the vertical profiles (5 m) where most models performed poorly due to a temporally unstable, vacillating metalimnion. Our results suggest predicting reservoir oxygen dynamics may be easier in stratified reservoirs, but eutrophication and complex hydrodynamics may cause forecasting surprises especially for those who use or manage water resources in mono- or dimictic reservoirs.

Lake and reservoir mixing regimes are predicted to change due to local and global environmental change (Woolway and Merchant 2019). For example, some lakes will transition from monomictic to polymictic, either by climate change or by shallowing from sedimentation fill, and it is unknown how this will impact the ability to make informed management decisions (Taranu et al. 2010; Woolway and Merchant 2019; Kornijów 2023). With increased availability of high temporal resolution monitoring data and more accessible software, data-driven machine learning methods are particularly primed to rapidly catalog and understand predictability of diverse environmental variables. In addition, space-for-time studies of predictability across environmental gradients now might set expectations for future predictability when major drivers of a system's ecological or biogeochemical dynamics (e.g., mixing regime) are nonstationary (Thomas et al. 2018; Rissman and Wardropper 2021).

*Correspondence: caleb_robbins@baylor.edu

[a]Present address: Department of Biology, Center for Reservoir and Aquatic Systems Research, Baylor University, Waco, Texas, USA

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

Author Contribution Statement: AN: Conceptualization (supporting); Methodology (equal), Writing – review and editing (equal). CR: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Methodology (equal); Writing – original draft (lead). JS: Methodology (equal); Writing – review and editing (equal). JTS: Conceptualization (supporting); Data curation (supporting); Project administration (lead); Resources (lead); Writing – review and editing (equal). DT: Conceptualization (supporting); Methodology (equal); Writing – review and editing (equal). NW: Data curation (supporting); Writing – review and editing (equal).

Dissolved oxygen is a principal biogeochemical variable in aquatic ecosystems, so predicting lake and reservoir oxygen dynamics is imperative to managing aquatic ecosystem function and services, especially as oxygen patterns worldwide become increasingly variable (Carey 2023). Causes of oxygen patterns are relatively well understood. For example, weather dynamics such as wind and air temperature, interacting with lake morphology, are key drivers of oxygen dynamics such as hypoxia duration in lakes and reservoirs via effects on lake mixing and metabolic rates (Woolway et al. 2017; Cortés et al. 2021; Ishikawa et al. 2021). In contrast, the predictability (rather than causes) of oxygen dynamics is poorly quantified, creating a problem for anticipating and managing effects of high or, especially, low oxygen, such as metal desorption, nutrient release, and spatial distributions of oxygen-sensitive organisms (Müller et al. 2012; Lofton et al. 2022; Carey 2023). Water temperature is a key control on oxygen and is relatively easy to predict, so a common assumption may be that oxygen patterns follow water temperature; studies predicting water temperature dynamics regularly point to impacts on oxygen (e.g., Butcher et al. 2015; Mi et al. 2019; Thomas et al. 2020). However, oxygen concentrations are additionally controlled by biological (e.g., phytoplankton biomass) and chemical (e.g., redox) factors that decouple patterns of oxygen from those of temperature, adding complexity, variability, and potentially hindering prediction over near-term timescales (Carey 2023). Studies are needed to understand patterns and drivers of oxygen predictability, but especially as related to environmental gradients that are expected to be nonstationary due to anthropogenic pressures.

Lake mixing patterns control the relative importance across space and time of oxygenating (e.g., photosynthesis) and deoxygenating (e.g., respiration) processes in water bodies and may be critical for forecasting oxygen and related biogeochemical or water quality patterns (Jane et al. 2021). In lakes and reservoirs with stable (seasonal) stratification, the persistence of presumably distinct hypo- and epilimnetic layers may facilitate predictability of oxygen vertically and through time. Therefore, epilimnetic oxygen can be expected to largely follow diel rhythms of photosynthesis and hypolimnetic oxygen generally decreases throughout the stratified period, possibly even becoming anoxic due to the dominance of microbial respiration. In polymictic systems, however, mixing of oxygen-rich surface waters into deeper depths frequently occurs, rapidly equilibrating chemically distinct epi- and hypolimnia to oxygen concentrations largely defined by the volumes of mixing water and the biological and chemical oxygen demand of the hypolimnion (Hammond et al. 2023; Wagner et al. 2023). Periods of stratification and mixing in the summer can quickly alternate, sometimes on daily timescales (MacIntyre et al. 2002; Wilhelm and Adrian 2008; Wagner et al. 2023), potentially worsening predictions of lake physiochemistry in the summer compared to cooler seasons characterized by constant holomixis (Durell et al. 2023). The temporality of stratification may therefore influence the predictability of oxygen, particularly on hourly to daily timescales that are critical for adaptively managing activities such as raw drinking water intake.

We asked how reservoir mixing conditions influence model predictions of daily oxygen dynamics using vertical profiles of high-frequency ($\sim$ every 2 h) oxygen concentrations across two monomictic and two polymictic reservoirs. Specifically, we hypothesized that frequent breakdown of stratification induces rapid variability in oxygen that diminishes oxygen predictability. We analyzed oxygen predictability both across and within lakes, predicting that (1) oxygen would be more predictable in monomictic reservoirs compared to polymictic reservoirs, (2) oxygen in polymictic reservoirs would be less predictable during the warm season (when stratification occurs intermittently), and (3) mixing conditions, that is, days with highly variable thermal stratification strength, would negatively correlate with oxygen predictability. We targeted daily prediction errors because this is the temporal scale at which there can be substantial variability, particularly from mixing dynamics, relevant to management action for activities like drinking water intake and treatment (e.g., from near-term, 1–10 d forecasts; Carey et al. 2022, Wagner et al. 2023). We used the machine learning algorithm random forest to generate predictions with commonly available weather variables and water temperature profiles as predictor variables. In this sense, we define predictability as a realized predictability (Pennekamp et al. 2019) in the context of driver variables that (1) have known causal relationships (either directly or indirectly) with stratification and oxygen patterns, and (2) are generally accessible, especially in a true forecasting context via, for example, the National Oceanic and Atmospheric Administration's Global Ensemble Forecasting System (GEFS), at similar temporal scale as the high-frequency water monitoring data. We used root mean square error (RMSE; *see* Methods) as a quantitative measure. We additionally explored how predictor variables causally associated with mixing (wind, water temperature, depth) contributed to prediction skill across reservoirs, seasons, and mixing conditions using Shapley Additive exPlanations (SHAP). Finally, we compared linear regression, machine learning, and process-based modeling approaches in a single polymictic reservoir to test how different modeling approaches were influenced by mixing conditions and to better evaluate our choice of the random forest algorithm for the focal across- and within-reservoir comparisons. We highlight how increasingly available high-frequency monitoring technology can be combined with accessible, highly predictive machine learning methods to analyze drivers of ecological predictability across environmental gradients (here, lake and reservoir mixing patterns). Our results uniquely suggest that the short-term, management relevant predictability of biogeochemical patterns and water quality varies across multiple environmental and even vertical gradients. Making accurate predictions of oxygen on these short timescales (e.g., near-term forecasts), particularly in metalimnia that are spatially highly variable, will often be difficult, requiring multiple years of high-frequency data or complex modeling to capture dynamic, complex oxygen patterns.

Methods

Study sites and high-frequency data collection

We collected high-frequency water temperature and oxygen profiles at four south-central US reservoirs varying in size, trophic status, and mixing regime (Table 1). All reservoirs are located in a humid subtropical climate and do not form any significant ice cover during the winter. Coordinates georeference the spatial location of the profiles and are shown in Fig. 1. Richland-Chambers Reservoir (32.04129°N, −96.20726°E) is a large (170 km^2), shallow (mean depth = 8 m) reservoir in east-central Texas, USA, near the city of Corsicana, ~ 120 km southeast of the Dallas-Fort Worth Metroplex (Fig. 1). Eagle Mountain Lake (32.87513°N, −97.46076°E) abuts northwest Fort Worth, Texas and has a mean depth of 6 m and surface area 36.5 km^2. These reservoirs are eutrophic (CJR, unpublished data; Chrzanowski and Grover 2005). Tarrant Regional Water District owns and operates the two Texas reservoirs to serve drinking water needs to the city and surrounding areas of Fort Worth, Texas. Lake Maumelle (34.86095°N, −92.50311°E) is a 36 km^2, mesotrophic reservoir with a mean depth of 7.5 m, operated by Central Arkansas Water to provide drinking water to the city of Little Rock (Shaver 2015). Lake Fayetteville (36.13475°N, −94.13813°E) is a small (0.8 km^2), hypereutrophic reservoir with a mean depth of ~ 3 m which serves as a secondary water supply source for the City of Fayetteville, Arkansas and is otherwise used for recreation (Grantz et al. 2012; Haggard et al. 2023).

Oxygen and water temperature profiles at Richland-Chambers, Eagle Mountain, and Fayetteville were collected using an autonomous vertical profiler (Instrumental Solutions, Inc.; as described in Wagner et al. 2023). At Eagle Mountain and Fayetteville, the profiler was mounted on an anchored pontoon with solar power, while at Richland-Chambers, the profiler was mounted to the wall of a water intake building, continuously powered by an electrical outlet, on the Chambers creek arm of the reservoir (Fig. 1). The profiler comprised a Linux computer as a central controller, cellular data card, and a Hydrolab HL7 water quality sonde (Ott Hydromet) hoisted by a motorized reel. The HL7 sondes were switched out for maintenance and calibration at least every 2 weeks to prevent probe biofouling and check for instrument drift. Measurement accuracy for temperature was 0.1°C and for oxygen was 0.2 mg L^{-1}. Profiles at Richland-Chambers and Eagle Mountain were measured every 2 h at 0.5 m depth intervals from 0 to 10 m. At Fayetteville, electrical malfunction of the solar panel prevented profiles from being collected during daytime hours, therefore we profiled every 2 h at 0.5 m depth intervals from 0 to 8.5 m each day from 7:00 p.m. to 5:00 a.m. Each profile takes 30 (Lake Fayetteville) to 40 min to complete.

Table 1. Information on each reservoir in this study, including the number (*N*) of dissolved oxygen data points used for model training and testing, overall R^2, and distributions of model error (root mean square error, RMSE, for the testing set) for each reservoir. All model RMSE in the manuscript are based on data from the testing set.

Reservoir	Mixing regime	Trophic state	Mean depth (m)	Surface area (km^2)	*N* (training)	*N* (testing)	Overall R^2	Overall RMSE	RMSE min	RMSE 5%	RMSE 50%	RMSE 95%	RMSE max
Eagle Mountain	Polymictic	Eutrophic	6	37	23,226	22,596	0.79	1.36	0.04	0.18	1.06	2.37	3.73
Richland-Chambers	Polymictic	Eutrophic	8	170	26,654	25,939	0.74	1.61	0.02	0.1	1.06	3.02	4.58
Fayetteville	Monomictic	Eutrophic	3	0.8	7873	8125	0.78	2.12	0.02	0.06	0.8	4.21	11.5
Maumelle	Monomictic	Mesotrophic	7.5	36	26,913	27,157	0.89	0.89	0	0	0.35	2.08	4.8

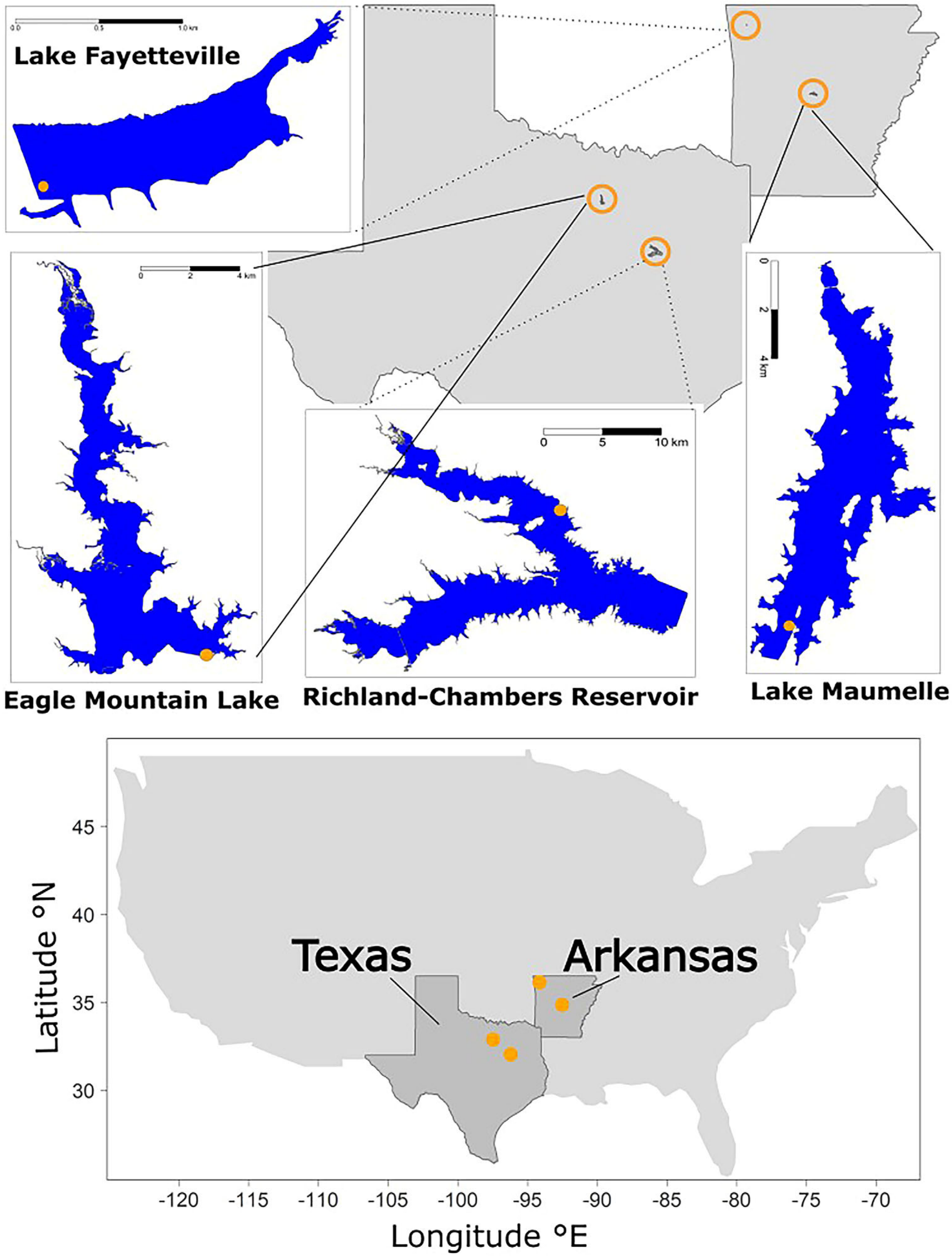

Fig. 1. Map of study reservoirs located in Texas and Arkansas, USA. North is oriented at the top of each reservoir map except for Maumelle, where north is toward the right of the figure. Orange circles on each reservoir mark the location of high-frequency water quality profilers.

Time series spanned April–December 2021 for Richland-Chambers, April–October 2019 for Eagle Mountain, and April–September 2021 for Fayetteville. At Maumelle, data were collected with a buoyed chain (PME) of 10 temperature and dissolved oxygen sensors (In-Situ Roxygen Pro) spaced ~ 1 m apart to 10 m depth (J. Fleming USGS pers. comm.). These data are collected and managed by the US Geological Survey (USGS) as monitoring location 072632995. We selected a subset of the available Maumelle time series that spanned dates similar to the other time series, from April to December 2021.

Hourly local weather data were obtained from the National Oceanographic and Atmospheric Administration (NOAA) Integrated Surface Database (ISD) using the R package worldmet (Carslaw 2023). ISD data collected at airports ranged from 10 to 30 km away from the profile locations. Weather conditions can vary substantially across such distances, so we stress that all models are learning correlative associations between weather variables and oxygen. Indeed, for the purpose of prediction, this is a value of highly flexible, data-driven machine learning methods. Even calibration of process-based models can tune parameters that control the sensitivities of hydrodynamics to weather conditions, optimizing process rate responses to given weather conditions rather than assuming perfectly known relationships between weather conditions and lake behavior (*see* Modeling section). We acquired air temperature, wind speed and wind direction at 10 m (transformed to north and west wind velocities), barometric pressure, dew point temperature, and cloud cover fraction. Reservoir inflow, only used for the process-based model GOTM-WET at Richland-Chambers (*see* Multi-model comparison section below) was obtained using measured discharge at USGS gages on the two major inflows, Richland Creek and Chambers Creek. Because these gages were upstream from the reservoir (but below any major inflows or impoundments), discharge was scaled to the most upstream location of the reservoir using the drainage area method, and then summed into a single inflow.

Modeling

Oxygen data were evenly split into training and testing sets to create 10-d periods ("validation periods") for model validation (Fig. 2). This splitting captured the general variability in oxygen dynamics across the available data in each reservoir while holding out ample data for testing across different reservoir conditions and avoiding a random split of the data that could significantly undersample cool or warm water periods (Supporting Information Header S1; Supporting Information Figs. S1–S4 show data sensitivity analyses, including a downsampling approach to validate inter-reservoir comparisons given Fayetteville profiling limitations in Study sites and high-frequency data collection). Random samples of the data are also improper tests of skill for random forests predicting time series (Regier et al. 2023).

We compared reservoir oxygen predictability using the machine learning algorithm random forest, which is based on decision trees that partition the predictor space using splitting rules. We fit, tuned, and compared random forest models using the ranger package within the tidymodels framework and collection of packages (Wright and Ziegler 2017; Kuhn and Wickham 2020) in R. Models were always fit with 500 trees (James et al. 2021), but we tuned hyperparameters m_try, which controls the number of randomly selected predictors for each tree, and min_n, which controls the minimum node size for each tree. Model selection and tuning used depth-stratified 10-fold cross-validation with the 10-fold partitioning repeated five times, using *vfold_cv*() on the training data. Final models to predict testing set oxygen used the combination of hyperparameters with the lowest root mean square error (RMSE) during the cross-validation routine on the training data.

We evaluated several other models to predict oxygen in Richland-Chambers Reservoir only to test how different modeling approaches might influence the efficacy of our oxygen predictions: (1) Linear regression, that is, ordinary least squares regression, (2) lasso regression, which is regularized linear regression that effectively adjusts regression coefficients to improve out-of-sample prediction, (3) Long short-term memory recurrent neural network (LSTM; Hochreiter and Schmidhuber 1997), which learns how to "retain" and "forget" potentially predictive information through time series, and (4) GOTM-WET, a process-based model built on differential equations representing the coupled (1-dimensional, vertical) hydrodynamics and ecological dynamics within a water body. The train-test split for each of these models was the same as used for the Richland-Chambers random forest. The lasso model was trained as the random forest models, where the hyperparameter lambda, the regularization parameter, was tuned via an identical 10-fold cross-validation procedure. The LSTM hyperparameters were also tuned using 10-fold cross-validation in Python. Tuning resulted in using a single LSTM layer with 10 hidden nodes. GOTM-WET was calibrated to the training data by continuously narrowing model parameters using the differential evolution algorithm in the Parallel Sensitivity and Auto-Calibration (parsac) tool in Python (Bruggeman and Bolding 2020). GOTM-WET was run continuously over the entire time series, including the test data periods, and was therefore not restarted with updated state parameters at the beginning of each 10-d period. *See* Supporting Information Header S2 for more details on hyperparameter tuning, and LSTM and GOTM-WET model training.

We used the standard deviation of daily Schmidt stability to indicate reservoir mixing. Briefly, Schmidt stability measures the resistance of a lake to mixing, or rather the energy needed to overcome the potential energy inherent to the lake's vertical thermal gradient (Idso 1973). Higher values of Schmidt stability therefore indicate strong thermal stratification, while values closer to zero indicate mixed conditions. We calculated stability for each depth profile with the function *schmidt_stability* in the R package rLakeAnalyzer (Winslow et al. 2019), which takes the vertical profile of water temperature and bathymetry as its main arguments. Higher standard deviation of the daily Schmidt stability suggests a mix between stratified and mixed conditions, while lower standard deviation of Schmidt stability suggests stable thermal conditions.

We also explored the importance of different predictors across and within reservoirs using Shapley Additive exPlanations (SHAP). Because we assumed that our predictors (features) were correlated, which confounds SHAP interpretation, we used the shapr package in R that implements SHAP that take feature correlation into account (Aas et al. 2021; Sellereite et al. 2023). SHAP values have a theoretical basis in game theory and represent the contribution of a feature to individual model predictions. More specifically, a SHAP value is the contribution of a feature to the difference between a model's prediction and the mean of the response variable across the dataset. We scaled each feature contribution as a percentage of

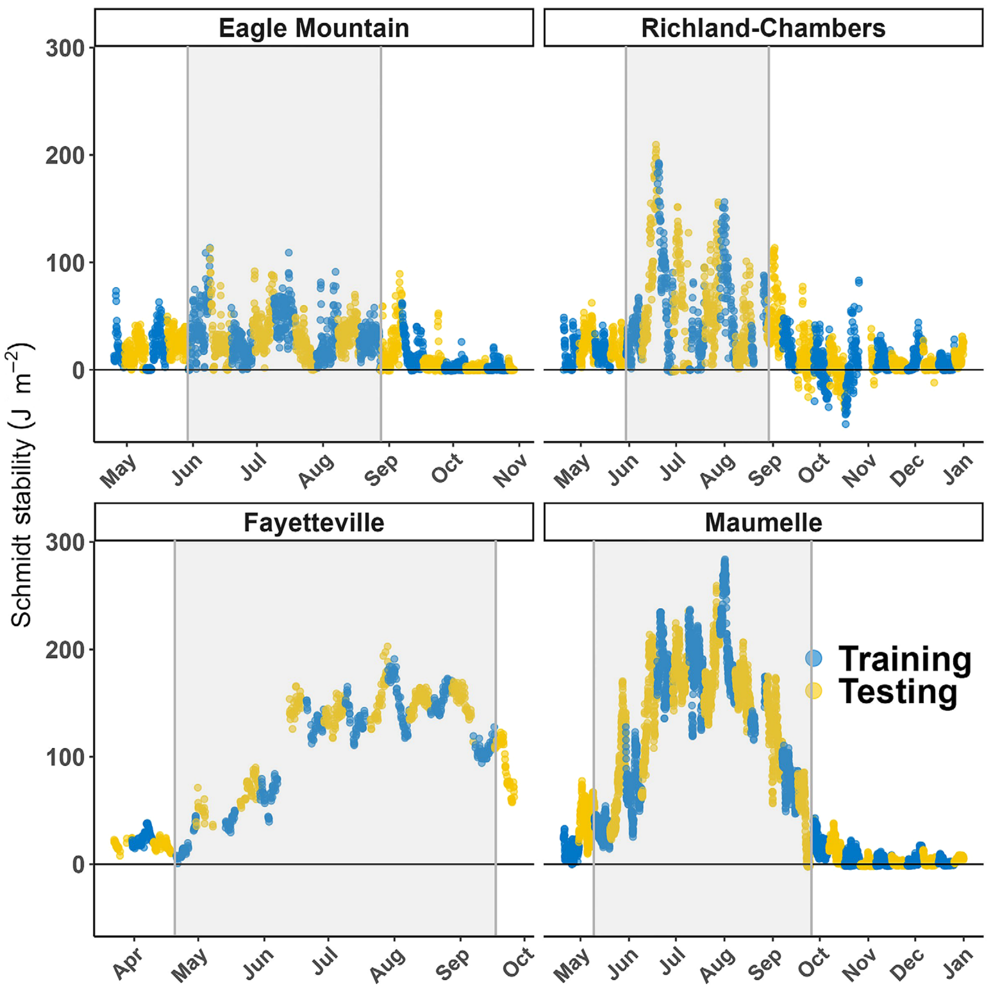

Fig. 2. Time series of Schmidt stability for each reservoir. Blue data points coincide with time periods of model training (though Schmidt stability was not a predictor variable), and yellow coincide with the model testing set. Schmidt stability estimates the reservoir's resistance to mixing, so higher values indicate more strongly stratified conditions and 0 indicates completely mixed conditions. The gray areas denote the "warm" season when intermittent mixing and stratification occur in polymictic reservoirs (Eagle Mountain, Richland-Chambers) and when stable stratification occurs in monomictic reservoirs (Fayetteville, Maumelle).

each prediction to facilitate interpreting predictor importance across reservoirs.

Statistical analysis of model prediction errors

We used linear mixed effects models to assess how RMSE (daily prediction error of oxygen) and predictor importance (in the random forest models) differ across (1) reservoirs, depths, and seasons, and (2) reservoirs, depths, and daily mixing. We tested models including random intercepts for each combination of reservoir, depth, and validation period, and included an AR(1) autoregressive model in the residuals. We also used variance weighting with the varIdent function in the R package nlme (Pinheiro and Bates 2000; Pinheiro et al. 2023). Final model selection was based on visualization of residuals for heteroscedasticity across the fitted values and by predictor variables, normality, and reduction in autocorrelation (Zuur et al. 2009). For inference of effect sizes between levels of predictor variables, we present the 95% confidence intervals (CIs) of the contrasts between levels, but note that in all cases where these CIs did not contain zero the corresponding p values (corrected for multiple comparisons via Tukey's method) were less than 0.05. Model estimated marginal means and contrasts were calculated using the package emmeans (Lenth 2023).

All R modeling and analysis was done in v4.2.3 (R Core Team 2023) and Python modeling in v3.9.0. All data, model training, and analyses are archived in Zenodo at DOI: 10.5281/zenodo.10403565.

Results

Mixing, temperature, and oxygen dynamics

Temporal patterns of Schmidt stability showed Eagle Mountain Lake and Richland-Chambers Reservoir were polymictic, while Lake Fayetteville and Lake Maumelle were monomictic (Fig. 2). The warm water season where polymictic reservoirs could occasionally stratify and monomictic reservoirs were stratified generally stretched from May through September or October. Fayetteville at the end of data availability was beginning to destratify, but the warm season at other reservoirs was completely shouldered by cool seasons of fully mixed conditions.

The two polymictic reservoirs, Eagle Mountain and Richland-Chambers, had similar temperature regimes to the two monomictic reservoirs, Fayetteville and Maumelle (Fig. 3, top panel). The main difference was that bottom temperatures in the polymictic reservoirs were generally warmer and more variable owing to the periodic mixing of warm surface waters. Hypolimnetic temperatures in the bottom waters of the monomictic reservoirs were relatively stable, suggesting

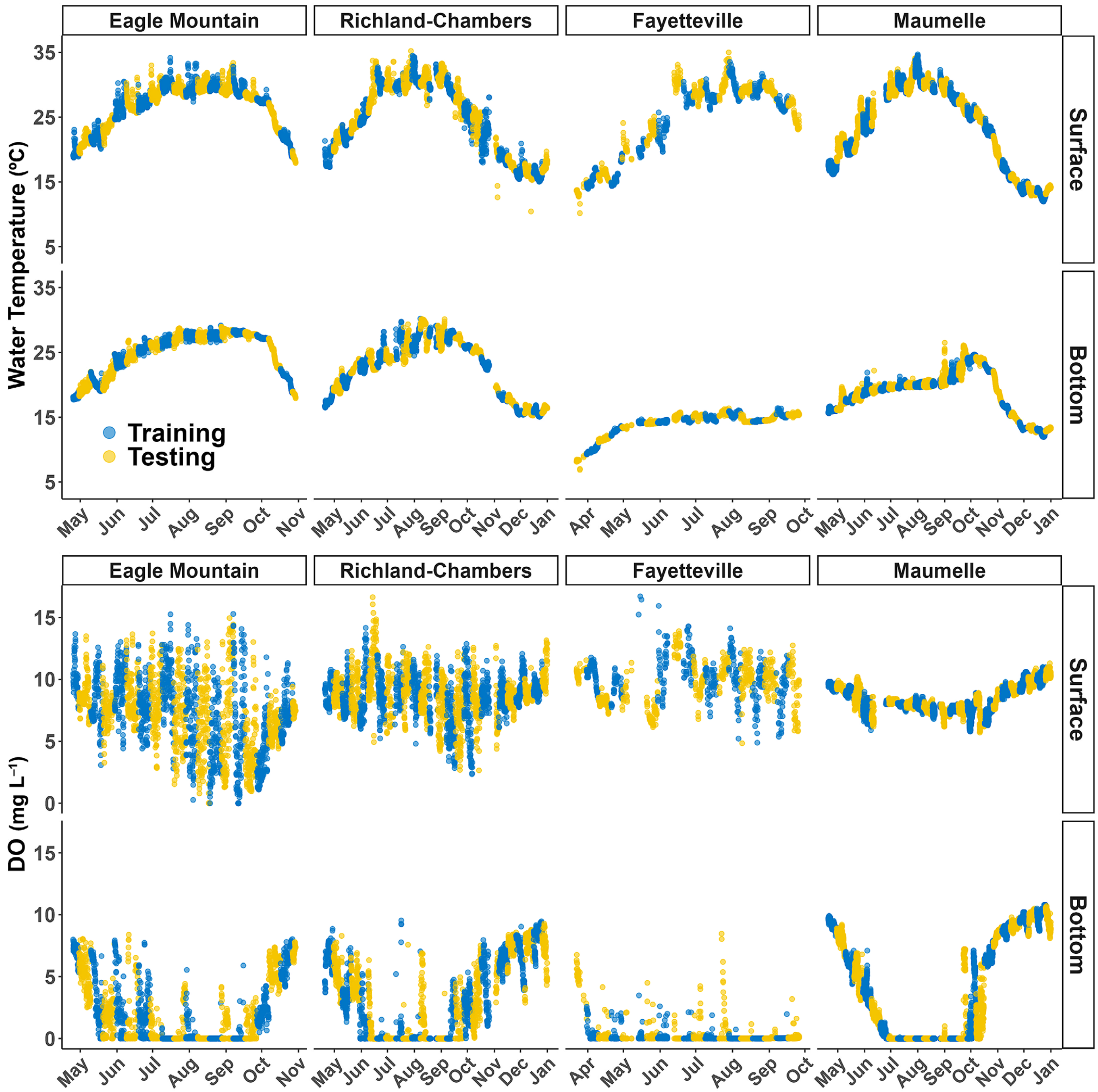

Fig. 3. Time series of temperature (top panel) and dissolved oxygen (DO; bottom panel) at the surface and bottom (8.5–10 m; *see* Methods) measurement stations for each reservoir. Data were measured at approximately two-hourly intervals. Blue data points were used for model training, while yellow data points were held as the testing set. Note that the *x*-axis has slightly different resolution depending on the reservoir due to differences in length of available time series.

limited mixing and heat exchange with surface waters that were ~10–15°C warmer throughout the stratified warm season.

Oxygen (dissolved oxygen) profiles between the two reservoir mixing regimes were markedly different (Fig. 3, bottom panel). Eagle Mountain and Richland-Chambers had strong diel swings at the surface (as did the monomictic reservoirs), but occasional mixing of anoxic bottom water induced hypoxia (< 2 mg L^{-1}) in the surface waters, inducing temporally highly variable surface water oxygen. Similarly, mixing events could oxygenate anoxic bottom waters to hypoxia or greater (e.g., > 5 mg L^{-1}). Anoxia was much more stable in the hypolimnia of the monomictic reservoirs. At Maumelle, oxygen decreased steadily in the hypolimnion after the reservoir stratified during the warm season, until reaching a stable anoxic state until the fall mixing event in late September. Fayetteville hypolimnetic oxygen was generally anoxic but occasionally oxygenated above hypoxic levels despite elevated Schmidt stability indicating a lack of a convective mixing event.

Reservoir model error comparison

Supporting Information Figs. S5–S17 show raw predicted oxygen vs. observed oxygen in each reservoir throughout the test set period and across the top, middle, and bottom depths.

Dissolved oxygen was fairly well-predicted in each reservoir (Table 1), with R^2 on the entire testing data varying from a minimum 0.73 for Richland-Chambers to a maximum of 0.89 for Maumelle (Table 1). RMSE on the entire testing data ranged from 0.89 mg L^{-1} for Maumelle to 2.1 mg L^{-1} for Fayetteville, suggesting Maumelle oxygen was generally more predictable than the other reservoirs. Mean RMSE differed by reservoir ($p < 0.0001$), as well as an interaction of season and depth ($p = 0.0012$). Depth-averaged oxygen concentrations in Richland-Chambers and Eagle Mountain, the two polymictic reservoirs, were less predictable than Maumelle oxygen in both seasons with the estimated mean differences in RMSE between reservoirs ranging from 0.4 (CI: 0.01–0.79) to 0.7 (CI: 0.30–1.2). Yet, in contrast to our prediction that the polymictic reservoirs would be less predictable than the monomictic reservoirs, Fayetteville oxygen RMSE was 0.85 (CI: 0.15–1.6) mg L^{-1} greater than Maumelle oxygen during the warm season.

Seasonal (cool or warm) factors generally had weak and uncertain effects on daily-aggregated RMSE within reservoirs, which did not support our prediction that the polymictic reservoirs would become less predictable during the warm stratifying season (Fig. 4). In fact, bottom RMSE decreased by 0.61 (contrast CI: 0.14–1.1) and 0.66 (contrast CI: 0.3–1.1) mg L^{-1} in the warm season compared to the cool season at Richland-Chambers and Maumelle, respectively, and increased by 1.4 mg L^{-1} (contrast CI: 0.35–2.4) in warm season surface predictions at Fayetteville, where we expected RMSE to stay the same or decrease. Generally, bottom oxygen was better predicted than other depths during the warm season. Maumelle bottom mean RMSE during the warm season was only 0.32 (CI: 0.0–0.7 mg L^{-1}) likely due

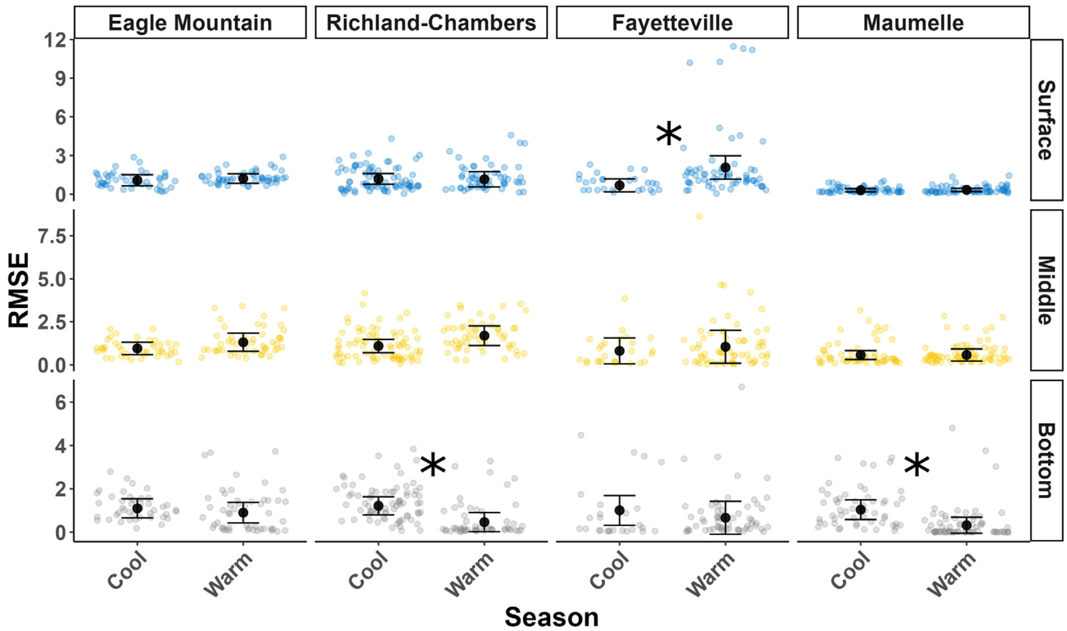

Fig. 4. Comparison of daily root mean square error (RMSE (mg L^{-1}); colored small points) of dissolved oxygen (DO) predictions between seasons at two polymictic (Eagle Mountain, Richland-Chambers) and two monomictic (Fayetteville, Maumelle) reservoirs. The warm season was the period during which the polymictic reservoirs would intermittently mix and stratify. Large black points and error bars are estimated marginal means ± 95% confidence interval. Model predictions were made at the surface (0 m), middle (5 m) and bottom (8.5–10 m; *see* Methods). Asterisks indicate significant ($p < 0.05$) differences between seasonal means within reservoirs and depths.

to temporally stable anoxia from approximately June 14 to September 22 eliminating variability in oxygen. Although anoxic conditions were present in almost all the testing data at the bottom of Fayetteville, as well, Fayetteville anoxic conditions were more tenuous than those observed at Maumelle despite no evidence of strong convective mixing (Fig. 2). Consequently, Fayetteville bottom oxygen predictions (RMSE: 0.83 mg L^{-1}, CI: 0.19–1.5) were less precise than Maumelle bottom oxygen during the warm season, and more similar to warm season bottom RMSE in the polymictic reservoirs Eagle Mountain (0.77 mg L^{-1}, CI: 0.36–1.2) and Richland-Chambers (0.55 mg L^{-1}, CI: 0.16–0.94).

Mixing (daily Schmidt stability standard deviation) had only minimal linear influence on the mean predicted daily RMSE at any reservoir or depth (Fig. 5). This effect was limited to the middle and surface depths of Eagle Mountain where RMSE increased by 0.030 (CI: 0.009–0.052) and 0.026 (CI: 0.007–0.046) mg L^{-1} per unit change in mixing.

Predictor importance

Across reservoirs, the most important predictors were depth, water temperature, and day of year, which on average respectively contributed 28%, 20%, and 19% of predictive force (as normalized SHAP) to oxygen predictions (Supporting Information Fig. S18). The other predictors (dew point, cloud cover, air temperature, air pressure, south + west wind), contributed 4–10% on average. Note that quantitative comparisons of SHAP in the following text are absolute differences of these percentages (i.e., treating "%" like a measurement unit).

We initially expected that variables strongly associated with mixing (wind, water temperature), as well as the importance of depth, would vary between polymictic vs. monomictic reservoirs, and between cool and warm seasons within

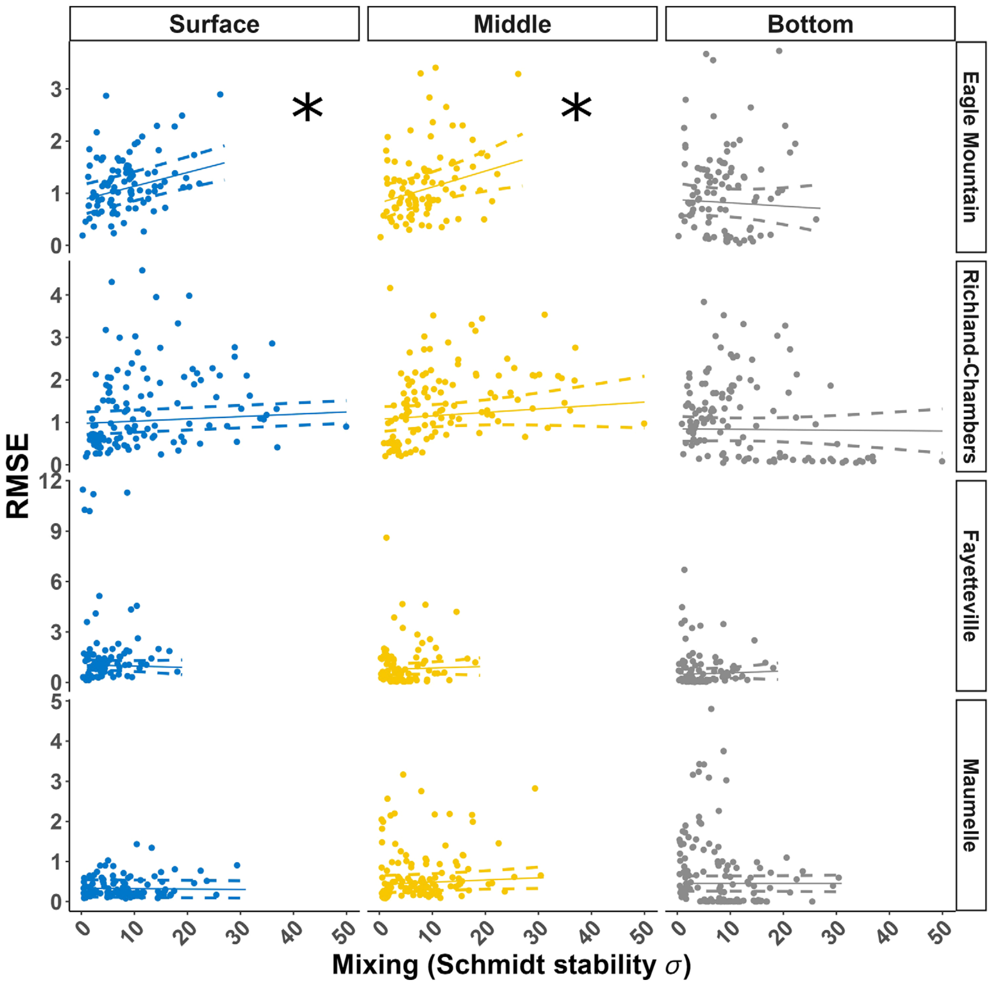

Fig. 5. Daily dissolved oxygen (DO) prediction errors (root mean square error [RMSE]) regressed on daily reservoir mixing at two polymictic (Eagle Mountain, Richland-Chambers) and two monomictic (Fayetteville, Maumelle) reservoirs. RMSE is given for predictions at the surface, middle (5 m), and bottom (8.5–10 m; *see* Methods). The standard deviation of Schmidt stability for each day was used to estimate the amount of mixing. Estimated regression line (solid) and its 95% confidence intervals (dotted) are derived from linear mixed model with variance weighting, residual autocorrelation structure, and random intercept. Significant slopes (*$p < 0.05$) were only detected at the middle and surface depths of Eagle Mountain.

polymictic reservoirs (Fig. 6). On average, wind was a more important predictor of oxygen concentrations in the polymictic reservoirs but this was mostly driven by depth-dependent differences, and qualitatively, wind was much more variably important to oxygen predictions in polymictic reservoirs. Wind was a more important predictor in the polymictic than monomictic reservoirs only at middle depths, where wind was ~ 9–10% (range of low to high contrast CIs: 1.4–19.2) more important to average oxygen predictions in each of the polymictic reservoirs contrasted with each monomictic reservoir. However, this difference in importance of wind between mixing regimes was constant across the warm and cool seasons.

During the cool season, water temperature contributed on average 7–14% (range of CIs: 0.5–21) less to oxygen predictions in the two polymictic reservoirs compared to Maumelle (but not Fayetteville), regardless of depth. During the warm season, the importance of water temperature across reservoirs was depth dependent. Water temperature was 8.5–12% more important in the polymictic reservoirs and Fayetteville than Maumelle at the surface (range of CIs: 1.5–19), and 7–10% less important to bottom predictions in the polymictic reservoirs compared to Maumelle (range of CIs: 0.08–18).

Depth was generally an important predictor of oxygen for surface and bottom oxygen predictions, but was 16–30% less important (regardless of season) for the middle depth oxygen predictions in Eagle Mountain, Richland-Chambers, and Fayetteville (Fig. 6). Only at Maumelle did depth tend to be more important to predictions during the warm season, where the importance of depth increased 12–16% Maumelle (range of lowest to highest contrast CIs: 7–22%) in warm compared to cool seasons at every depth.

Mixing was occasionally linearly related to variable importance (Supporting Information Fig. S19). Water temperature SHAP (%) increased with mixing at the surface (slope = 0.23, CI: 0.13–0.34) and middle (slope = 0.12, CI: 0.005–0.23) depths of Richland-Chambers, but decreased with mixing at the surface of Maumelle (slope = −0.25, CI: −0.46 to −0.05). Wind SHAP was negatively related to mixing at the surface (slope = −0.14, CI: −0.25 to −0.04) and middle (slope = −0.25, CI: −0.37 to −0.14) depths of Richland-Chambers.

Multi-model comparison at Richland-Chambers Reservoir

No one model always outperformed the others, but differences in model performance depended on the season and depth of the predictions (three-way interaction of season, model type, and depth category $p < 0.0001$; Fig. 7). At the surface, mean daily RMSE was not discernibly different across any model regardless of season. The linear regression-based methods performed comparatively to the machine learning methods, excepting lasso RMSE was ~ 1.3 (both CI: 0.3–2.4) mg L^{-1} greater than the same predictions made with random

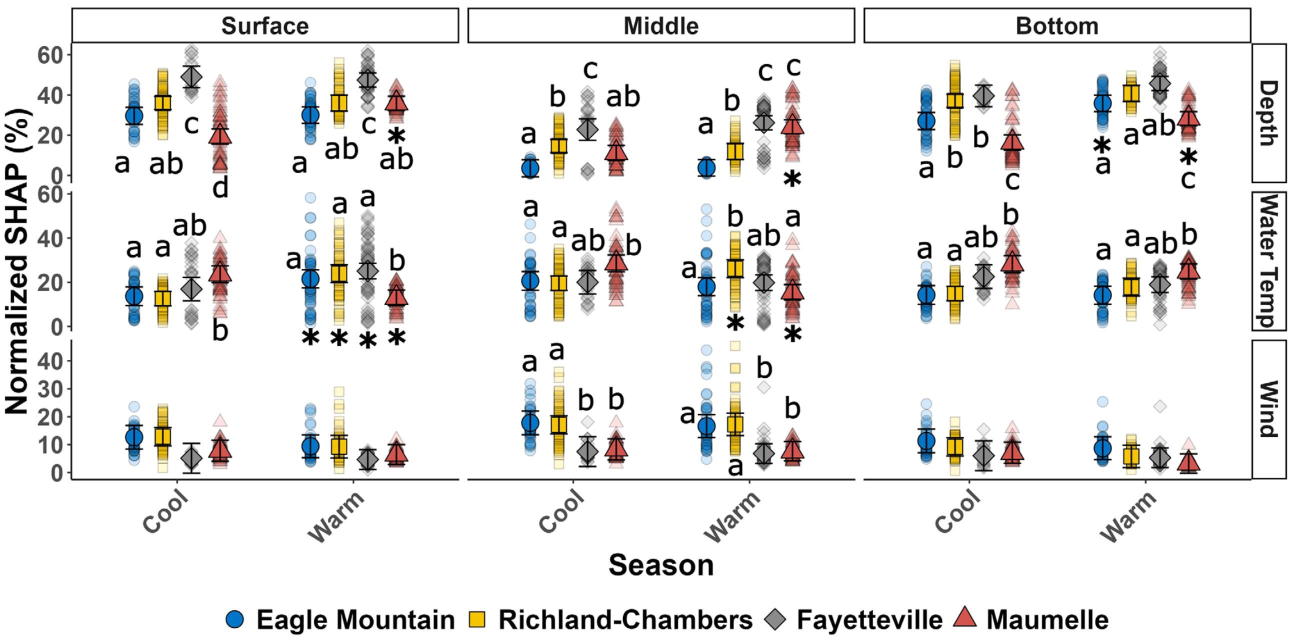

Fig. 6. Shapley Additive exPlanations (SHAP), normalized as a percent contribution to the model dissolved oxygen (DO) prediction, for three primary variables (depth, top; water temperature, middle; wind speed, bottom) we expected to be important across monomictic (Maumelle, Fayetteville) and polymictic (Eagle Mountain, Richland-Chambers) reservoirs. SHAP was calculated for model predictions made at the surface (0 m), middle (5 m), and bottom (8.5–10 m; *see* Methods). The warm season was the period during which the polymictic reservoirs would intermittently mix and stratify. Lettering indicates differences across reservoirs but within season. Asterisks represent a significant seasonal change within depth and within a reservoir from the corresponding cool season SHAP value.

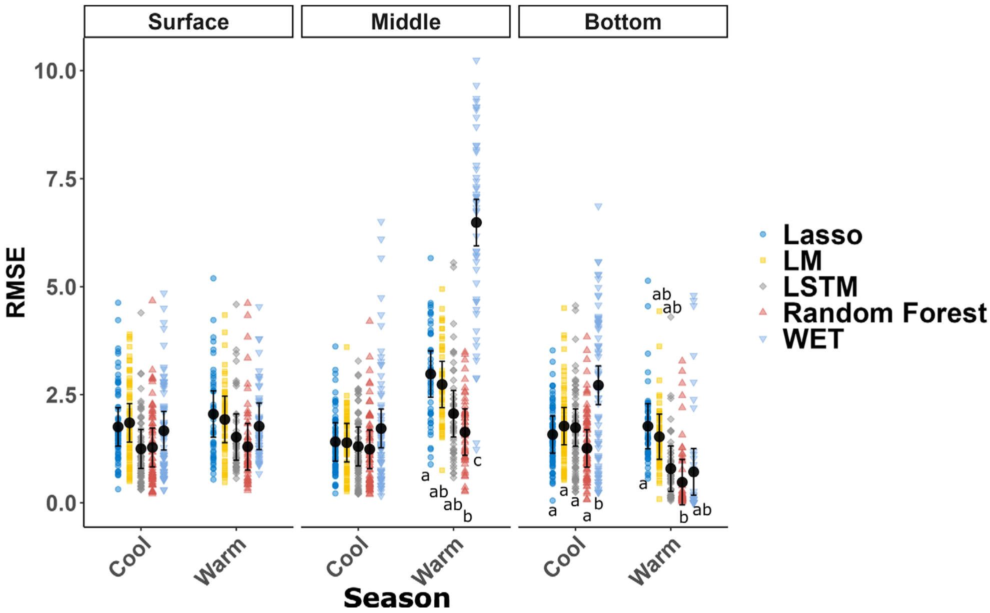

Fig. 7. Daily prediction errors (root mean square error [RMSE]) for Richland-Chambers dissolved oxygen (DO) profiles at the surface, middle (5 m) and bottom (10 m). Models are lasso regression, ordinary least squares regression (LM), long short-term memory neural network (LSTM), random forest, and GOTM-WET (process-based model). The warm season was the period during which the reservoir would intermittently mix and stratify. Different letters indicate significant ($p < 0.05$) differences across models within seasons.

forest. While differences in mean daily RMSE were often indiscernible across models, the bottom and middle predictions from GOTM-WET were poor during the warm season. Daily RMSE was 0.9–1.5 mg L^{-1} higher for GOTM-WET than all other models at bottom predictions during the cool season. Moreover, while most model predictions at the middle depth were worse during the warm season, RMSE from the process-based model GOTM-WET was 3.5–4.9 (range of CI: 2.4–5.5) mg L^{-1} higher than the other models.

There was little consistent pattern suggesting any particular model errors were directly influenced by mixing conditions. Only middle depth RMSE for GOTM-WET and surface RMSE for LSTM were influenced by daily mixing (Supporting Information Fig. S20).

Discussion

Thermal stratification of lakes and reservoirs can largely isolate photosynthetically active surface waters (epilimnion) from bottom water (hypolimnion) that gradually loses oxygen. We therefore hypothesized that this stratification creates relatively predictable patterns of oxygen; conversely, that the frequent breakdown of this stratification yields less predictable oxygen. Specifically, we predicted that oxygen would be less predictable (1) in polymictic reservoirs compared to monomictic, (2) within polymictic reservoirs during the warm season, and (3) during polymictic (intermittently stratified) conditions. None of these predictions were strictly true because lake categories based on thermal stratification did not clearly delineate predictability and instability rarely influenced daily RMSE. In fact, models were generally skilled at using weather variables to predict oxygen across a broad variety of conditions.

Our results suggest that reservoir mixing regime (specifically, poly vs. monomictic) in a strict sense poorly determines oxygen predictability, and other factors such as eutrophication and lake morphology should be more broadly explored to understand oxygen predictability. However, our results do indicate that thermal stratification increases the predictability of oxygen depending on the specific environmental context, congruent with our main hypothesis. Maumelle, a monomictic reservoir, was generally more predictable than the other three reservoirs. During the warm season when the reservoir was fully stratified, predictability increased only at the bottom. Maumelle bottom mean RMSE during the warm season was only 0.32 due to temporally stable anoxia from approximately June 14 to September 22 eliminating variability in oxygen. Although Fayetteville was consistently thermally stratified (Fig. 2), hypolimnetic oxygen was not consistently anoxic like at Maumelle. Fayetteville is by a wide margin the smallest reservoir in this study, and rapid hydrodynamic changes over these small spatial scales might bring oxygenated waters to the profiler. Lake morphology can influence the predictability of water temperature (Thomas et al. 2023). Morphology may similarly relate to complex hydrodynamic patterns that could drive oxygen predictability on different time scales than the random forest models here could identify, for example, as lagged effects of river inflows or cumulative effects of wind-driven water circulation. Hydrodynamically driven oxygen dynamics that are rarer in the data and poorly correlated with some function of the predictors will be less predictable.

We hypothesize that eutrophication may also decrease oxygen predictability by increasing rates of change of oxygen over short time scales. Similar to hypolimnetic model predictions, Maumelle surface oxygen was highly predictable compared to oxygen at the other three reservoirs, likely due to Maumelle surface oxygen varying by only a few mg L^{-1} across most training and testing splits. In contrast, oxygen at other reservoirs was considerably more variable. In the most extreme case at Eagle Mountain, surface oxygen ranged from 14 mg L^{-1} to nearly anoxic within days, driven by both stratification dynamics and eutrophication (Wagner et al. 2023). Future studies should elucidate mechanisms of biogeochemical predictability across eutrophic reservoirs.

Vertically, middle depths (5 m in this study) were the worst predicted, in line with observations in other predictive studies (Saber et al. 2020; Lin et al. 2023). This is not particularly unexpected as a metalimnion is often located there when lakes and reservoirs are stratified, representing the interface where downward convective mixing and upward biological oxygen demand interact. Although many predictions in the literature are concerned with surface and bottom oxygen dynamics, many raw water intakes are located at intermediate depths (e.g., 3–7 m), such as at Richland-Chambers. Water quality can vary tremendously depending on depth. For example, managers may be sensitive to intake of anoxic water because it contains desorbed metal contaminants that require expensive chemical pretreatment. Forecasting oxygen at depths around which a metalimnion temporally vacillates above and below (e.g., "thermocline deepening"; Lofton et al. 2022), exacerbated by eutrophic conditions that rapidly change oxygen concentrations, may be particularly difficult but be of particular importance to the utility of lake forecasting for management-focused end-users.

Mixing regime (across-reservoir comparisons) and mixing (within-reservoir) inconsistently related to the importance (SHAP) of wind or water temperature, indicating complex relationships among the predictor variables drove oxygen predictions during mixing. Although temporally variable, wind was qualitatively more important in predicting oxygen in polymictic reservoirs, suggesting the importance of capturing wind patterns to predict frequent convective mixing. In particular, wind was most important at the middle depths of the polymictic reservoirs, suggesting the importance of wind as driving a highly variable mixed layer depth where epilimnetic and hypolimnetic waters meet when the reservoirs are not mixed. In contrast to the importance of wind between mixing regimes, mixing (within-reservoir) was inconsistently correlated with SHAP values; for instance, mixing was positively correlated with the importance of water temperature at the surface and middle of polymictic Richland-Chambers and not for any depth at the other polymictic reservoir, Eagle Mountain. This illustrates that data-driven ML methods can identify different correlations between oxygen and driver variables depending on the data, even in similar systems. The different correlations (and SHAP variable importance) across reservoirs could be attributed, at least partially, to differences in (1) weather data accuracy and (2) differential influences of weather on lake hydrodynamics and oxygen production/consumption at the exact spatial locations of the sensors. Flexible methods such as random forest are potentially less sensitive to these data biases than simple linear models or complex process-based models precisely because they are parameterizing a high number of correlations found in the data, providing predictive accuracy at the expense of interpretability. Nevertheless, tools like SHAP (and tools not used here such as partial dependence plots) can help scientists and managers explore, understand, and communicate patterns in their specific systems, being careful to separate correlation from causation.

Despite being one of the most important variables to our models according to SHAP, knowing water temperature did not always lead to accurate daily oxygen predictions. Lake and reservoir water quality predictions often very reasonably focus on water temperature (e.g., Thomas et al. 2020). Water temperature directly controls oxygen solubility, directly controls biogeochemical rates, and vertical profiles of temperature indicate mixing. Oxygen is often less predictable than temperature and near-term forecasts of oxygen are less skilled than those of temperature (Arhonditsis and Brett 2004; Saber et al. 2020). Oxygen may become even less predictable as the climate continues to change, driving oxygen patterns in lakes and reservoirs further away from historical trends (Pilla et al. 2020; Jane et al. 2021; Carey 2023). Water temperature profiles indicate stratification, which should reliably suggest, for example, that anoxic hypolimnia will remain hypoxic if temperature profiles are predicted to remain stratified. However, as seen for oxygen dynamics at Fayetteville, oxic conditions can occur in the hypolimnion without evidence of mixing. This underscores the possibility that oxygen forecasts may be complicated not only by temporally complex interactions between physical mixing and environmental conditions, but also variables that are less easily forecastable like horizontal hydrodynamics (Carey 2023). Our results suggest that future studies should increasingly study the predictability of oxygen in addition to water temperature, for example, to better understand when and where investments in oxygen forecasting can be valuable, or where temperature will be an acceptably accurate indicator of oxygen patterns for specific management applications.

The inter-model comparison at Richland-Chambers showed that random forest was the best performing model alongside the LSTM. We expected the complex mixing and oxygen dynamics at Richland-Chambers to be better predicted by the LSTM as a complex deep learning approach incorporating explicit time series dynamics. No model is perfectly predictive, but we argue random forest was a robust modeling choice to provide estimates of realized predictability that reflect usable information from the predictors at each reservoir. Random forests are well known as highly flexible and skillful predictive algorithms, with the added benefit that they are computationally light to train compared to deep learning (Tyralis et al. 2019). In addition, predictability as conceptualized and estimated in this study also incorporates an aspect of ease—finding a significantly better predicting model at each reservoir would likely require obtaining much more data volume (likely multiyear) or exploring more difficult-to-obtain predictor variables that additionally may be less applicable in a true forecasting context. In support of this argument, first, our sensitivity analysis (Supporting Information Header S1) showed that random forest performance on the test set was not data limited at any reservoir. Second, as predictors we used (1) weather variables that are often available via weather agencies such as NOAA or by local weather monitoring stations, and (2) water temperature which is usually measured simultaneously with oxygen and could be well-predicted such that the temperature predictions could be incorporated as driver variables in oxygen forecasting models. Other forecastable predictors (e.g., reservoir inflow, if inflow is gaged) are more difficult to incorporate at relevant temporal scales, but should be explored in future water quality prediction studies.

Although random forests and LSTM outperformed the other model types, we were somewhat surprised by the consistency of model errors across all model types in the inter-model comparison more generally. We especially might have expected that models incorporating temporally lagged effects of predictors would perform better by using temporally autocorrelated information naturally found in many time series. As noted in the previous paragraph, LSTM performed essentially as well as the random forest model. The process-based model GOTM-WET, based on time-explicit differential equations, generally also performed well, but at certain times produced some of the largest model errors in the Richland-Chambers inter-model comparison. This does not necessarily imply that GOTM-WET poorly predicts oxygen dynamics in a general sense. In a time series context, GOTM-WET and other process models can strongly benefit from data assimilation involving reinitializing the model with updated states of temperature and oxygen (Moore 2020; Wander et al. 2024). Incorporation of temporal information implicitly occurred to varying extents for the other models, for example with a day-of-year predictor included in the regression, lasso, and random forests (even though lagged oxygen observations were not explicit predictors in these models). Furthermore, we trained random forest configurations at all reservoirs with lagged predictors; cross-validation at each lake consistently selected the predictor set with non-time lagged predictors. It was also somewhat surprising that, despite their relatively inflexible functional form, linear regression and lasso methods only performed significantly worse than any other methods during the warm season for middle and bottom predictions. Increasingly complex correlations that simpler linear methods cannot capture might exist between weather drivers and oxygen dynamics when they are spatially and temporally disconnected from processes at the air–water interface (e.g., during polymictic periods). In these cases, more complex ML modeling such as random forests or deep learning can be powerful.

Predictive science builds an important basis for the management of the environment and our resources, including reservoirs (Houlahan et al. 2017; Carey et al. 2022). Our results suggest that differences in predictability of biogeochemical dynamics and water quality exist across reservoirs. Specifically, our results imply that predicting oxygen can be easier in less nutrient rich, monomictic (or dimictic) reservoirs that have summer-long stable periods of thermal stratification. Our study furthermore shows that even stratified reservoirs can exhibit complex oxygen dynamics. We hypothesize that eutrophication, reservoir morphology, and horizontal hydrodynamics strongly determine the ease of predictability of oxygen time series. Accurately predicting oxygen in such water bodies may require multiple years of data and/or more complex modeling approaches that incorporate machine learning with 3D hydrodynamics or other process-based models (Read et al. 2019; Lin et al. 2023). Accuracy is also relative, defined by values and management objectives that should drive what level of predictability is worth improving through more complex modeling or data collection (Elliott-Graves 2020). As lake and reservoir processes in the Anthropocene become increasingly nonstationary and transition across fundamental regimes such as mixing, trophic status, and climate, water quality may become more predictable or less predictable on short timescales (e.g., near-term forecasts of 1–10 d). In a space-for-time approach, benchmarking predictive models of biogeochemical and ecological time series across these regimes may develop management expectations and forecasting priorities (Brookes et al. 2014).

Data availability statement

Data and analyses are archived in Zenodo at https://zenodo.org/records/11412820 (DOI: 10.5281/zenodo.10403565).

References

Aas, K., M. Jullum, and A. Løland. 2021. Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. Artif Intell **298**: 103502. doi:10.1016/j.artint.2021.103502

Arhonditsis, G. B., and M. T. Brett. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. Mar. Ecol. Prog. Ser. **271**: 13–26. doi:10.3354/meps271013

Brookes, J. D., C. C. Carey, D. P. Hamilton, L. Ho, L. van der Linden, R. Renner, and A. Rigosi. 2014. Emerging challenges for the drinking water industry. Environ. Sci. Technol. **48**: 2099–2101. doi:10.1021/es405606t

Bruggeman, J., and K. Bolding, 2020. Parsac: Parallel sensitivity analysis and calibration. doi:10.5281/ZENODO.4280520.

Butcher, J. B., D. Nover, T. E. Johnson, and C. M. Clark. 2015. Sensitivity of lake thermal and mixing dynamics to climate change. Clim. Change **129**: 295–305. doi:10.1007/s10584-015-1326-1

Carey, C. C. 2023. Causes and consequences of changing oxygen availability in lakes. Inland Waters **13**: 316–326. doi:10.1080/20442041.2023.2239110

Carey, C. C., and others. 2022. Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting. Inland Waters **12**: 107–120. doi:10.1080/20442041.2020.1816421

Carslaw, D. 2023. worldmet: Import surface meteorological Data from NOAA integrated surface database (ISD).

Chrzanowski, T. H., and J. P. Grover. 2005. Temporal coherence in Limnological features of two southwestern reservoirs. Lake Reserv. Manag. **21**: 39–48. doi:10.1080/07438140509354411

Cortés, A., and others. 2021. Prediction of hypoxia in eutrophic polymictic lakes. Water Resour. Res. **57**: e2020WR028693. doi:10.1029/2020WR028693

Durell, L., J. T. Scott, and A. S. Hering. 2023. Hybrid forecasting for functional time series of dissolved oxygen profiles. Data Science in Science **2**: 2152401. doi:10.1080/26941899.2022.2152401

Elliott-Graves, A. 2020. The value of imprecise prediction. Philos. Theory Pract. Biol. **12**.

Grantz, E. M., A. Kogo, and J. T. Scott. 2012. Partitioning whole-lake denitrification using in situ dinitrogen gas accumulation and intact sediment core experiments. Limnol. Oceanogr. **57**: 925–935. doi:10.4319/lo.2012.57.4.0925

Haggard, B. E., E. Grantz, B. J. Austin, A. L. Lasater, L. L. Haddock, A. Ferri, N. D. Wagner, and J. T. Scott. 2023. Microcystin shows thresholds and hierarchical structure with physiochemical properties at Lake Fayetteville, Arkansas, May through September 2020. J. ASABE **66**: 307–317. doi:10.13031/ja.15273

Hammond, N. W., F. Birgand, C. C. Carey, B. Bookout, A. Breef-Pilz, and M. E. Schreiber. 2023. High-frequency sensor data capture short-term variability in Fe and Mn concentrations due to hypolimnetic oxygenation and seasonal dynamics in a drinking water reservoir. Water Res. **240**: 120084. doi:10.1016/j.watres.2023.120084

Hochreiter, S., and J. Schmidhuber. 1997. Long short-term memory. Neural Comput. **9**: 1735–1780. doi:10.1162/neco.1997.9.8.1735

Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2017. The priority of prediction in ecological understanding. Oikos **126**: 1–7. doi:10.1111/oik.03726

Idso, S. B. 1973. On the concept of lake stability. Limnol. Oceanogr. **18**: 681–683. doi:10.4319/lo.1973.18.4.0681

Ishikawa, M., T. Bleninger, and A. Lorke. 2021. Hydrodynamics and mixing mechanisms in a subtropical reservoir. Inland Waters **11**: 286–301. doi:10.1080/20442041.2021.1932391

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2021. Tree-based methods, p. 327–365. *In* An introduction to statistical learning. Springer. doi:10.1007/978-1-0716-1418-1_8

Jane, S. F., and others. 2021. Widespread deoxygenation of temperate lakes. Nature **594**: 66–70. doi:10.1038/s41586-021-03550-y

Kornijów, R. 2023. Lake aging concept revisited: Alternative comprehensive approach to lake ontogeny. Can. J. Fish. Aquat. Sci. **80**: 1385–1393. doi:10.1139/cjfas-2023-0027

Kuhn, M., and H. Wickham. 2020. Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles. https://www.tidymodels.org

Lenth, R. 2023. emmeans: Estimated marginal means, aka least-squares means.

Lin, S., D. C. Pierson, R. Ladwig, B. M. Kraemer, and F. R. S. Hu. 2023. Multi-model machine learning approach accurately predicts lake dissolved oxygen with meteorological and hydrological input. SSRN Scholarly Paper.

Lofton, M. E., D. W. Howard, R. P. McClure, H. L. Wander, W. M. Woelmer, A. G. Hounshell, A. S. L. Lewis, and C. C. Carey. 2022. Experimental thermocline deepening alters vertical distribution and community structure of phytoplankton in a 4-year whole-reservoir manipulation. Freshw. Biol. **67**: 1903–1924. doi:10.1111/fwb.13983

MacIntyre, S., J. R. Romero, and G. W. Kling. 2002. Spatial-temporal variability in surface layer deepening and lateral advection in an embayment of Lake Victoria, East Africa. Limnol. Oceanogr. **47**: 656–671. doi:10.4319/lo.2002.47.3.0656

Mi, C., A. Sadeghian, K.-E. Lindenschmidt, and K. Rinke. 2019. Variable withdrawal elevations as a management tool to counter the effects of climate warming in Germany's largest drinking water reservoir. Environ. Sci. Eur. **31**: 19. doi:10.1186/s12302-019-0202-4

Moore, T. N. 2020. Predicting in-lake responses to short and long-term changes using lake physical models. PhD thesis. Dundalk Institute of Technology.

Müller, B., L. D. Bryant, A. Matzinger, and A. Wüest. 2012. Hypolimnetic oxygen depletion in eutrophic lakes. Environ. Sci. Technol. **46**: 9964–9971. doi:10.1021/es301422r

Pennekamp, F., and others. 2019. The intrinsic predictability of ecological time series and its potential to guide forecasting. Ecol. Monogr. **89**: e01359. doi:10.1002/ecm.1359

Pilla, R. M., and others. 2020. Deeper waters are changing less consistently than surface waters in a global analysis of 102 lakes. Sci. Rep. **10**: 20514. doi:10.1038/s41598-020-76873-x

Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-plus. Springer. doi:10.1007/b98882

Pinheiro, J., and others. 2023. nlme: Linear and nonlinear mixed effects models. doi:10.7759/cureus.50549

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Read, J. S., and others. 2019. Process-guided deep learning predictions of Lake water temperature. Water Resour. Res. **55**: 9173–9190. doi:10.1029/2019WR024922

Regier, P., M. Duggan, A. Myers-Pigg, and N. Ward. 2023. Effects of random forest modeling decisions on biogeochemical time series predictions. Limnol. Oceanogr.: Methods **21**: 40–52. doi:10.1002/lom3.10523

Rissman, A. R., and C. B. Wardropper. 2021. Adapting conservation policy and administration to nonstationary conditions. Soc. Nat. Resour. **34**: 524–537. doi:10.1080/08941920.2020.1799127

Saber, A., D. E. James, and D. F. Hayes. 2020. Long-term forecast of water temperature and dissolved oxygen profiles in deep lakes using artificial neural networks conjugated with wavelet transform. Limnol. Oceanogr. **65**: 1297–1317. doi:10.1002/lno.11390

Sellereite, N., M. Jullum, and A. Redelmeier. 2023. shapr: Prediction explanation with dependence-aware Shapley values.

Shaver, R. R. 2015. Overview of the phytoplankton and water quality of Lake Maumelle for 2009. MS thesis. Univ. of Arkansas.

Taranu, Z. E., D. Köster, R. I. Hall, T. Charette, F. Forrest, L. C. Cwynar, and I. Gregory-Eaves. 2010. Contrasting responses of dimictic and polymictic lakes to environmental change: A spatial and temporal study. Aquat. Sci. **72**: 97–115. doi:10.1007/s00027-009-0120-4

Thomas, M. K., S. Fontana, M. Reyes, M. Kehoe, and F. Pomati. 2018. The predictability of a lake phytoplankton community, over time-scales of hours to years. Ecol. Lett. **21**: 619–628. doi:10.1111/ele.12927

Thomas, R. Q., R. J. Figueiredo, V. Daneshmand, B. J. Bookout, L. K. Puckett, and C. C. Carey. 2020. A near-term iterative forecasting system successfully predicts reservoir hydrodynamics and partitions uncertainty in real time. Water Resour. Res. **56**: e2019WR026138. doi:10.1029/2019WR026138

Thomas, R. Q., R. P. McClure, T. N. Moore, W. M. Woelmer, C. Boettiger, R. J. Figueiredo, R. T. Hensley, and C. C. Carey. 2023. Near-term forecasts of NEON lakes reveal gradients of environmental predictability across the US. Front. Ecol. Environ. **21**: 220–226. doi:10.1002/fee.2623

Tyralis, H., G. Papacharalampous, and A. Langousis. 2019. A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water **11**: 910. doi:10.3390/w11050910

Wagner, N. D., F. S. Osburn, C. J. Robbins, M. R. Ernst, J. Owens, S. M. Powers, and J. T. Scott. 2023. Lake stability and anoxia dynamics revealed from high frequency vertical profiling in a eutrophic polymictic reservoir. Inland Waters **13**: 167–181. doi:10.1080/20442041.2022.2161264

Wander, H. L., R. Q. Thomas, T. N. Moore, M. E. Lofton, A. Breef-Pilz, and C. Carey. 2024. Data assimilation experiments inform monitoring needs for near-term ecological forecasts in a eutrophic reservoir. Ecosphere **15**: e4752. doi:10.1002/ecs2.4752

Wilhelm, S., and R. Adrian. 2008. Impact of summer warming on the thermal characteristics of a polymictic lake and consequences for oxygen, nutrients and phytoplankton. Freshw. Biol. **53**: 226–237. doi:10.1111/j.1365-2427.2007.01887.x

Winslow, L., J. Read, R. Woolway, T. Leach, J. Zwart, S. Albers, and D. Collinge. 2019. rLakeAnalyzer: Lake physics tools.

Woolway, R. I., and C. J. Merchant. 2019. Worldwide alteration of lake mixing regimes in response to climate change. Nat. Geosci. **12**: 271–276. doi:10.1038/s41561-019-0322-x

Woolway, R. I., P. Meinson, P. Nõges, I. D. Jones, and A. Laas. 2017. Atmospheric stilling leads to prolonged thermal stratification in a large shallow polymictic lake. Clim. Change **141**: 759–773. doi:10.1007/s10584-017-1909-0

Wright, M. N., and A. Ziegler. 2017. ranger: A fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. **77**: 1–17. doi:10.18637/jss.v077.i01

Zuur, A., E. N. Ieno, N. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer Science & Business Media. doi:10.1007/978-0-387-87458-6

Acknowledgments

Tarrant Regional Water District provided funding to install and maintain profilers at Richland-Chambers and Eagle Mountain, with specific support from Mark Ernst, Jennifer Owens, and Donna Stephens. Support for Caleb Robbins was provided by the Baylor University Office of the Vice Provost for Research. Brad Austin helped install and maintain the profiler at Fayetteville. Isabelle Anderson, Ashlynn Boedecker, Brittany Carnathan, Felicia Osburn, Jasmine Stovall, and Jingyu Wang variously aided with profiler maintenance. Thanks to Reed Green and Joey Fleming for specific information on the Maumelle buoy. The computing for the LSTM portion of the project was performed at the High Performance Computing Center at Oklahoma State University supported in part through the National Science Foundation grant OAC-1531128.

Conflict of Interest

The authors declare no conflicts of interest.

Submitted 22 December 2023
Revised 01 June 2024
Accepted 12 July 2024

Associate editor: Werner Eckert