A near-term iterative forecasting system successfully predicts reservoir hydrodynamics and partitions uncertainty

R. Quinn Thomas[1], Renato J. Figueiredo[2], Vahid Daneshmand[2], Bethany J. Bookout[3], Laura K. Puckett[1], and Cayelan C. Carey[3]

[1]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, 24061, USA

[2]Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, 32611, USA

[3]Department of Biological Sciences, Virginia Tech, Blacksburg, VA, 24061, USA

Corresponding author: R. Quinn Thomas (rqthomas@vt.edu)

Key Points:

- We created a near-term iterative lake water temperature forecasting system that uses sensors, data assimilation, and hydrodynamic modeling
- FLARE quantifies the uncertainty in each daily forecast and provides an open-source, generalizable system for water quality forecasting
- 16-day forecasted temperatures were within 0.91℃ over 100 days in a reservoir case study

Abstract

Freshwater ecosystems are experiencing greater variability due to human activities, necessitating new tools to anticipate future water quality. In response, we developed and operationalized a near-term iterative water temperature forecasting system (FLARE – Forecasting Lake And Reservoir Ecosystems) that is generalizable for lakes and reservoirs. FLARE is composed of: water quality and meteorology sensors that wirelessly stream data, a data assimilation algorithm that uses sensor observations to update predictions from a hydrodynamic model and calibrate model parameters, and an ensemble-based forecasting algorithm to generate forecasts that include uncertainty. Importantly, FLARE quantifies the contribution of different sources of uncertainty (parameters, driver data, initial conditions, and process) to each daily forecast of water temperature at multiple depths. We applied FLARE to a temperate reservoir during a 100-day period that encompassed stratified and mixed thermal conditions and found that daily forecasted water temperatures were on average within 0.91℃ at all depths of the reservoir over a 16-day forecast horizon. FLARE successfully predicted the onset of fall turnover eight days in advance, and identified meteorology driver data and downscaling as the dominant sources of forecast uncertainty. Overall, FLARE provides an open-source and easily-generalizable system for water quality forecasting for lakes and reservoirs to improve management.

Keywords: Data assimilation, Ecological forecasting, Ensemble Kalman filter, FLARE, General Lake Model, Water temperature

1 Introduction

As a result of human activities, ecosystems around the globe are increasingly changing [*Stocker et al. 2013, Ummenhofer and Meehl 2017*], making it challenging for resource managers to consistently provision vital ecosystem services [*West et al. 2009*]. In particular, managers of freshwater ecosystems, which have been more degraded than any other ecosystem on the planet [*Millennium Ecosystem Assessment 2005*], are seeking new tools to anticipate future change and ensure clean water for drinking, fisheries, irrigation, industry, and recreation [*Brookes et al. 2014*].

In response to this need, near-term iterative ecological forecasting has emerged as a solution to provide stakeholders, managers, and policy-makers crucial information about future ecosystem conditions [*Clark et al. 2001, Dietze et al. 2018, Luo et al. 2011*]. Here, we define a near-term iterative forecast as a projection of future ecosystem states with fully-specified uncertainties, generated from predictive models that can be constantly updated with new data as they become available [*Clark et al. 2001*]. Importantly, a near-term iterative forecast is not created from merely one ecosystem simulation, but an ensemble of simulations that enable quantification of the uncertainty in the forecast contributed by different sources [i.e., parameters, driver data, initial conditions, and process; *Dietze 2017a, Dietze 2017b*]. Because quantifying the sources of forecast uncertainty is a major goal of ecological forecasting research, multiple approaches have been developed to estimate and reduce forecast uncertainty [e.g., Bayesian state-space modeling, particle filters, and ensemble filters; *Dietze 2017a*]. Fully-specified uncertainty provides both an assessment of confidence in a forecast for managers as they interpret the forecasts for decision-making and valuable information for researchers about how to improve forecasts.

Forecasts of water temperature are particularly valuable for managers that oversee drinking water supply lakes and reservoirs, as waterbody temperatures can be very dynamic due to meteorological forcing, management, and seasonality [e.g., *Klug et al. 2012, Mi et al. 2019, Schmidt et al. 2018, Sharma et al. 2015*]. Because water temperature is closely related to many water quality metrics, including microbial and algal growth, dissolved oxygen saturation, the release of chemical constituents from sediments into the water column, and habitat suitability for organisms [e.g., fish; *Butcher et al. 2015, Carey et al. 2012, Delpla et al. 2009, Jöhnk et al. 2008*], water temperature profile data are used to determine withdrawal depths for water treatment, extraction schedules for hydropower generation, and *in situ* water quality management [*Çalışkan and Elçi 2008, Casamitjana et al. 2003, Weber et al. 2017*]. Water temperature depth profiles also determine the strength of thermal stratification, i.e., if there are discrete epilimnetic (surface) and hypolimnetic (bottom) layers or isothermal (fully-mixed) conditions [*Read et al. 2011*]. When waterbodies transition from stratified to mixed conditions during the onset of fall turnover, reduced nutrients and metals that accumulated in the hypolimnion during the summer are mixed throughout the water column, decreasing water quality [*Cooke et al. 2005, Effler and Matthews 2008*]. Consequently, near-term iterative forecasts of water temperature profiles would allow managers to preemptively respond to impending poor water quality during fall turnover and other episodic events (e.g., storms) that alter water temperature and thermal stratification.

Here, we introduce a forecasting system (FLARE, Forecasting Lake And Reservoir Ecosystems) that generates automated 16-day water temperature forecasts and is generalizable to many lakes and reservoirs (Figure 1). FLARE is composed of: 1) water quality and meteorology sensors deployed in a lake or reservoir that wirelessly stream data, 2) a data assimilation algorithm that uses sensor observations to update water temperature predictions from a

hydrodynamic model and to calibrate model parameters, and 3) an ensemble-based forecasting algorithm to generate forecasts that quantify the sources of forecast uncertainty. FLARE quantifies uncertainty from model process (i.e., the capacity of a calibrated model to reproduce past observations), model parameters, initial conditions (i.e., the uncertainty observed in water temperatures on the first day of the forecast), and driver data (i.e., the uncertainty in future weather forecasts that are needed to run the hydrodynamic model). The forecasting system samples from these sources of uncertainty to generate probability distributions for water temperature at multiple depths and can generate probability distributions of hydrodynamic events such as the occurrence of fall turnover.

We set up the forecasting system to automatically generate probabilistic water temperature forecasts for a drinking water reservoir over 100 days to address the following questions: 1) How does forecasting performance differ among three key stages of lake thermal dynamics (summer stratification, fall turnover, and fall mixing)?, 2) How well does the forecasting system predict the onset of fall turnover?, and 3) What are the contributions of different sources of uncertainty to the forecasts?

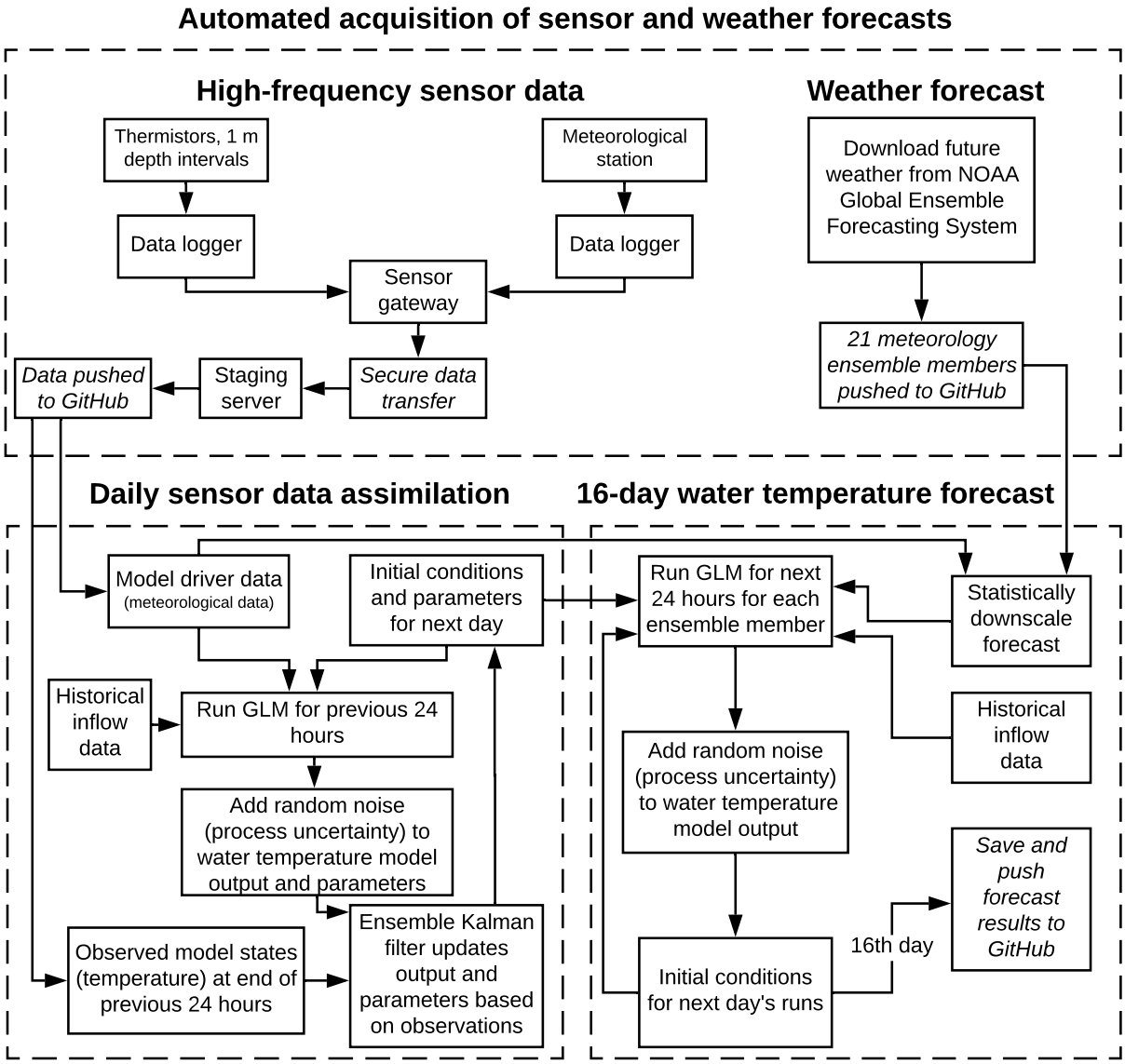

Figure 1. A conceptual diagram describing the data assimilation and forecasting workflow for the FLARE (Forecasting Lake And Reservoir Ecosystems) system. The high-frequency sensor data and NOAA 16-day weather forecasts are transferred to a GitHub repository where the data assimilation algorithm uses the output of the previous day's data assimilation as initial conditions for the current day, assimilates the new sensor observations to update water temperature predictions and model parameters, and passes the updated temperatures and parameters to the forecasting algorithm as initial conditions for generating predictions of future ecosystem states with their uncertainty. The forecasting system spatially and temporally downscales the 16-day NOAA Global Ensemble Forecasting System forecasts to local site conditions at the hourly scale and propagates the process uncertainty through the 16-day forecast horizon.

2 Methods

We developed a forecasting system (FLARE; Forecasting Lake And Reservoir Ecosystems) that predicts water temperature at any set of specified depths in a lake or reservoir for a 16-day time horizon using a physics-based hydrodynamic model. The system uses observed data and data assimilation techniques to generate the initial conditions, parameters, and uncertainty estimates for a forecast into a 16-day future. We describe the forecasting methods below and the data assimilation methods in Supporting Information A.

2.1 *Hydrodynamic model*

FLARE simulated reservoir hydrodynamics with the General Lake Model (GLM), a one-dimensional (1-D) vertical stratification model [*Hipsey et al. 2019*]. We used GLM because: 1) the model has successfully reproduced observed water temperature profiles in lakes around the world with varying mixing regime, climate, and morphology [*Bruce et al. 2018*]; 2) GLM is an open-source, community-developed model and thus scalable to other waterbodies for future forecasting applications [*Snortheim et al. 2017*]; and 3) GLM has low computational needs, enabling many model ensemble members to be run quickly and efficiently, a requirement for near-term iterative forecasting.

To enable generalization to other lakes or reservoirs, we set all but three highly sensitive parameter values equal to the values reported in the default GLM version 3.0 model [*Hipsey et al. 2019*]. The three parameters, selected using the methods described in Supplemental Information A), were: a scalar for incoming shortwave radiation (sw_factor) and two parameters defining the sediment temperature in the deep and shallow reservoir zones (zone1temp: 5 – 9.3 m deep; zone2temp: 0 – 5 m deep). For driver data, GLM requires hourly meteorological data on

138 downwelling shortwave radiation (W m^{-2}), downwelling longwave radiation (W m^{-2}), air
139 temperature (℃), wind speed (m s^{-1}), relative humidity (%), and precipitation (m day^{-1}) as well
140 as daily rates of water inflow (m^3 day^{-1}), inflow temperature (℃), and daily rates of outflow (m^3
141 day^{-1}) [*Hipsey et al. 2019*].
142
143 2.2 *Ensemble forecasting approach*
144 The FLARE system uses an ensemble approach to numerically simulate and propagate forecast
145 uncertainty into the future. The ensemble forecasting is based on eqn. 1:

146 $$x_{f+1}^{i} = G\left(x_{f}^{i}, \alpha_{t}^{i}, D_{f}^{i}\right) + MVN(0, \Sigma_{t}) \quad \text{(eqn. 1)}$$

147 where G() is the GLM model that requires a vector of water temperatures at the modeled depths
148 as initial conditions (x_f^i), a vector of the four calibrated model parameters (α_t^i), and a set of
149 model drivers (i.e., weather, inflows, and outflows). The index *i* in eqn. 1 represents the *i*th
150 ensemble member and the subscript *f* is the day in the future (e.g., if *f*=1, this would be referring
151 to tomorrow). The *t* subscript denotes the time step in the daily in the data assimilation. MVN is
152 the multivariate normal distribution with a mean of zero and a covariance matrix of Σ_t.
153 The contribution of process uncertainty to the total forecast uncertainty is represented by
154 adding a multivariate normal random variable to the GLM predictions (MVN(0, Σ_t) in eqn 1).
155 This process uncertainty is centered on zero with a covariance matrix (Σ_t) that represents the
156 uncertainty at each depth (the diagonals of Σ_t) and the covariance of uncertainty between depths
157 (the off-diagonals of Σ_t). Σ_t is calculated from the residuals in the data assimilation before the
158 states were updated using the observations. Eqn. 1 is applied at the daily time step and the
159 random process uncertainty is added each day of the forecast. In eqn. 1, the subscript *t* on
160 Σ_t signifies that the covariance matrix does not change through the forecast because *t* does not

incrementally increase when forecasting into the future (*t* only increments in the daily data assimilation that sets the initial conditions, parameters, and process uncertainty for the forecast).

The 16-day forecasts require initial conditions of the water temperature at each modeled depth ($x^{i}_{f=0} = x^{i}_{t}$, where *t* equals the time step of the data assimilation when the forecast into the future was initialized) and initial model parameters (α^{i}_{t}) on day 0 (f = 0) of the forecast. The variance in $x_t$and α_t across ensemble members when the forecast was initialized represents the contribution of initial conditions and parameters, respectively, to the total forecast uncertainty. To generate these initial conditions while also calibrating the three focal parameters, FLARE assimilates temperature sensor observations from the previous day into the GLM using an Ensemble Kalman Filter [EnKF; *Evensen 2003, Evensen 2009*]. We used data assimilation rather than simply specifying the initial conditions from the observations because the EnKF: 1) enabled the generation of initial conditions when sensor data were not available (e.g., during maintenance), 2) mechanistically-interpolated water temperature between sensor depths, 3) enabled the calibration of model parameters, and 4) generated a historical data product of water temperature with spatial and temporal gap-filling. The EnKF method of data assimilation is well-suited for non-linear mechanistic models like the GLM and enables ensemble-based forecasts of future states [*Dietze 2017a*]. Our implementation of the EnKF with state augmentation to calibrate parameters (Supporting Information A) follows Zhang et al. (2017).

To quantify the contribution of future meteorological conditions to the total forecast uncertainty, each water temperature ensemble member was assigned one of the ensemble members from the National Oceanic and Atmospheric Administration Global Ensemble Forecasting System (NOAA GEFS). NOAA GEFS provides 21 different modeled representations of future weather conditions at a 16-day time horizon. Each day that a forecast

was generated, we downloaded the meteorological driver data required by the GLM from NOAA GEFS for the 0:00:00 UTC forecast using the rNOMADS package in R [*Bowman 2019*].

Additionally, our forecasts included the uncertainty in temporally and spatially-downscaling the gridded NOAA GEFS meteorological forecasts. NOAA GEFS forecasts were available for 6-hour periods over a 16-day horizon at a 1×1° spatial resolution, which needed to be translated into hourly meteorology driver data for GLM that was specific to local site conditions. Detailed information on the meteorology downscaling are available in Supporting Information B and summarized here. First, we estimated the linear relationship between the daily mean observed value for a variable (e.g., daily mean air temperature) and the mean of the NOAA GEFS ensemble on the 1st day of the NOAA GEFS forecast. We used observations between April 4, 2018 and December 6, 2018 in the linear regression. Second, we calculated the residuals in the linear regression for each day and meteorological variable. Third, we calculated the covariance matrix in the residuals among the meteorological variables. Fourth, to propagate the uncertainty in the downscaling process for the meteorology variables, we sampled from a multivariate distribution with the mean equal to a vector of daily means for the meteorology variables from a NOAA GEFS forecast ensemble member and the covariance equal to the covariance of residuals described above. We repeated the sampling from the multivariate distribution for each day within each NOAA GEFS ensemble member, resulting in an ensemble of meteorological drivers that represented both NOAA GEFS forecast and downscaling uncertainty. For example, 21 draws from this multivariate distribution for each of the 21 NOAA ensemble members resulted in 441 unique ensemble members describing uncertainty in the weather drivers for a particular day in the future. Finally, we converted the daily downscaled value for each meteorological variable to the hourly time resolution needed as model driver data.

We used the sub-daily variation inherent in the original 6-hour NOAA forecast to first linearly convert the daily downscaled data to the 6-hour time scale and then second, used a linear spline function to convert the 6-hour data to 1-hour time resolution. In the case of shortwave radiation, we used solar geometry to convert the daily downscaled shortwave radiation to a 1-hour time resolution.

2.3 *Application of forecasting system*

2.3.1 Site description

We applied and evaluated FLARE at Falling Creek Reservoir (FCR), a dimictic, eutrophic reservoir located in Vinton, Virginia, USA (37.30°N,79.84°W). FCR is a shallow (maximum depth=9.3 m, mean depth=4 m), small (surface area=0.119 km^2) reservoir [*Chen et al. 2018*]. The lake exhibits summer thermal stratification from May to October and is ice-covered from January to February or March [*Carey et al. 2019c*]. FCR is primarily fed by one upstream tributary and was maintained at full pond throughout this study by the Western Virginia Water Authority (WVWA), who own and manage the reservoir as a drinking water supply [*Gerling et al. 2014, Gerling et al. 2016*].

We monitored FCR water temperature [*Carey et al. 2019b*] and meteorology [*Carey et al. 2019a*] with high-frequency sensors during 11 July to 19 December 2018 and measured the inflow discharge rate of the primary tributary entering FCR through a weir [*Carey et al. 2018*]. Descriptions of the sensor array and methods for real-time wireless transfer of data to cloud storage are in Supporting Information C. Because we were unable to wirelessly connect the weir sensor to the cloud to transmit the inflow data in real-time, we averaged the previous five years' data measured on a given day to serve as inflow driver data for forecasting.

230

231 2.3.2 Description of forecasting analysis

232 Our application of the forecasting system was divided into two periods: the spin-up
233 period and the forecasting period. The spin-up period was from 11 July to 25 August 2018 and
234 was used to develop the Σ_t process uncertainty covariance matrix and to constrain the three
235 parameters that were calibrated using the EnKF. In this period, the EnKF was used to update the
236 states and parameters using observed meteorology as the drivers. We used N=441 ensemble
237 members in the EnKF so that each ensemble member can be associated with one of the 441
238 weather ensemble members when forecasting (e.g., 21 NOAA GEFS ensemble members each
239 with 21 downscaling ensemble members). We modeled water temperature on 0.33-m depth
240 intervals starting at the surface at 0.1 m through 9.3 m at the sediments. This resulted in 29
241 model states of water temperature depths in the x matrix (eqn. 1). We had sensor observations for
242 10 of the 29 model depths (0.1, 1, 2, 3, 4, 5, 6, 7, 8, and 9 m; Supporting Information C).
243 The second period was the forecasting period when 16-day forecasts were generated each
244 day between 26 August and 3 December 2018. The 100 daily forecasts developed over this
245 period included summer stratified, fall turnover, and fall mixed conditions in the reservoir.
246 During this period, the model states and parameters were advanced one day using observed
247 meteorology and updated using the EnKF. These updated states and parameters were used as the
248 initial conditions for each 16-day forecast, which started at midnight of the current day.

249

250 2.3.3 Evaluation of forecasts

251 We evaluated the forecasts using three metrics that assessed the skill, bias, and quality of
252 confidence intervals. Skill was assessed using the root mean square error (RMSE) of the mean

253 forecasted water temperature and the observed water temperature at each day in the 16-day
254 forecast horizon. Bias was assessed using the absolute difference between the mean forecasted
255 water temperature and the observed water temperature at each day in the 16-day forecast. We
256 averaged the RMSE and bias for each day within the 16-day forecast horizon over the 100
257 forecasts generated between 26 August and 3 December. We compared the forecast RMSE to the
258 RMSE of a null persistence model that assumed water temperature did not change over the 16-
259 day horizon. The quality of the confidence intervals was assessed by calculating the total number
260 of observations contained in a specific confidence interval. We considered a forecast to be well-
261 calibrated if its 90% confidence intervals contained 90% of the observations over the 100 days of
262 forecasts. The confidence interval would be over-confident if fewer than 90% of observations
263 were contained in this interval and would be under-confident if more than 90% observations
264 were contained. We calculated the proportion of observations in the 10, 20, 30, 40, 50, 60, 70,
265 80, and 90th confidence intervals for 1-day, 7-day, and 16-day time horizons.
266 We evaluated the ability of the forecasting system to predict the day that fall turnover
267 was first observed at the reservoir. Following McClure et al. [*2018*], we defined fall turnover as
268 the first day in autumn when water temperatures at 1 m and 8 m depths are within 1°C. In the
269 forecasts prior to turnover, we calculated the proportion of ensemble members that predicted the
270 1 m and 8 m water temperatures to be ≤1°C different on each day in the forecast. We expected
271 this proportion to increase for the day of turnover relative to other days as the onset of turnover
272 approached.
273
274

2.3.4 Partitioning of uncertainty

To compare the contributions of the different sources of forecast uncertainty, we simulated each uncertainty source individually and compared the variance in each source to the variance in the total forecast uncertainty. Initial condition uncertainty was isolated by only including uncertainty in the x_{t-1}^{i} on the first day of the forecast – i.e., no process error was added, the mean parameter values were used, and only one NOAA GEFS ensemble member was used without downscaling uncertainty. Additionally, we examined the influence of gaps in high-frequency water temperature sensor data availability on forecast uncertainty by again isolating initial condition uncertainty but only allowing weekly observations to update model states in the EnKF (simulating if only weekly sampling data were available for EnKF updating instead of 10-minute resolution data). In this case, we set initial condition uncertainty equal to the uncertainty in the EnKF after six days without observations.

Model process uncertainty was isolated by sampling from the process uncertainty while initializing all ensemble members with the ensemble mean (removing initial condition uncertainty), using the mean parameter values, and using one member of the NOAA GEFS without downscaling. Parameter uncertainty was isolated in a similar way to process uncertainty.

The meteorological driver data uncertainty was separated into two components. First, the uncertainty in the NOAA forecast was isolated by using the 21 ensemble members without downscaling, process, parameters, and initial conditions uncertainty included. The downscaling uncertainty was isolated by adding downscaling uncertainty to a single NOAA ensemble member while not adding process, parameters, and initial conditions uncertainty.

We repeated the forecasting uncertainty partitioning for three different forecasts (1 September, 18 October, and 1 December) to represent the three different stages of lake thermal dynamics that occurred during the forecasting period.

3. Results

3.1 *Observational water temperature data*

Falling Creek Reservoir exhibited summer thermal stratification from the beginning of the monitoring period in July until the onset of fall turnover on 21 October, and then remained mixed until the end of December 2018 (Figure 2A). During the summer thermally-stratified period, observed water temperatures at the surface reached up to 29.6^{o}C on 11 July before cooling in mid-September. Fall turnover was preceded by cooling surface water temperatures,

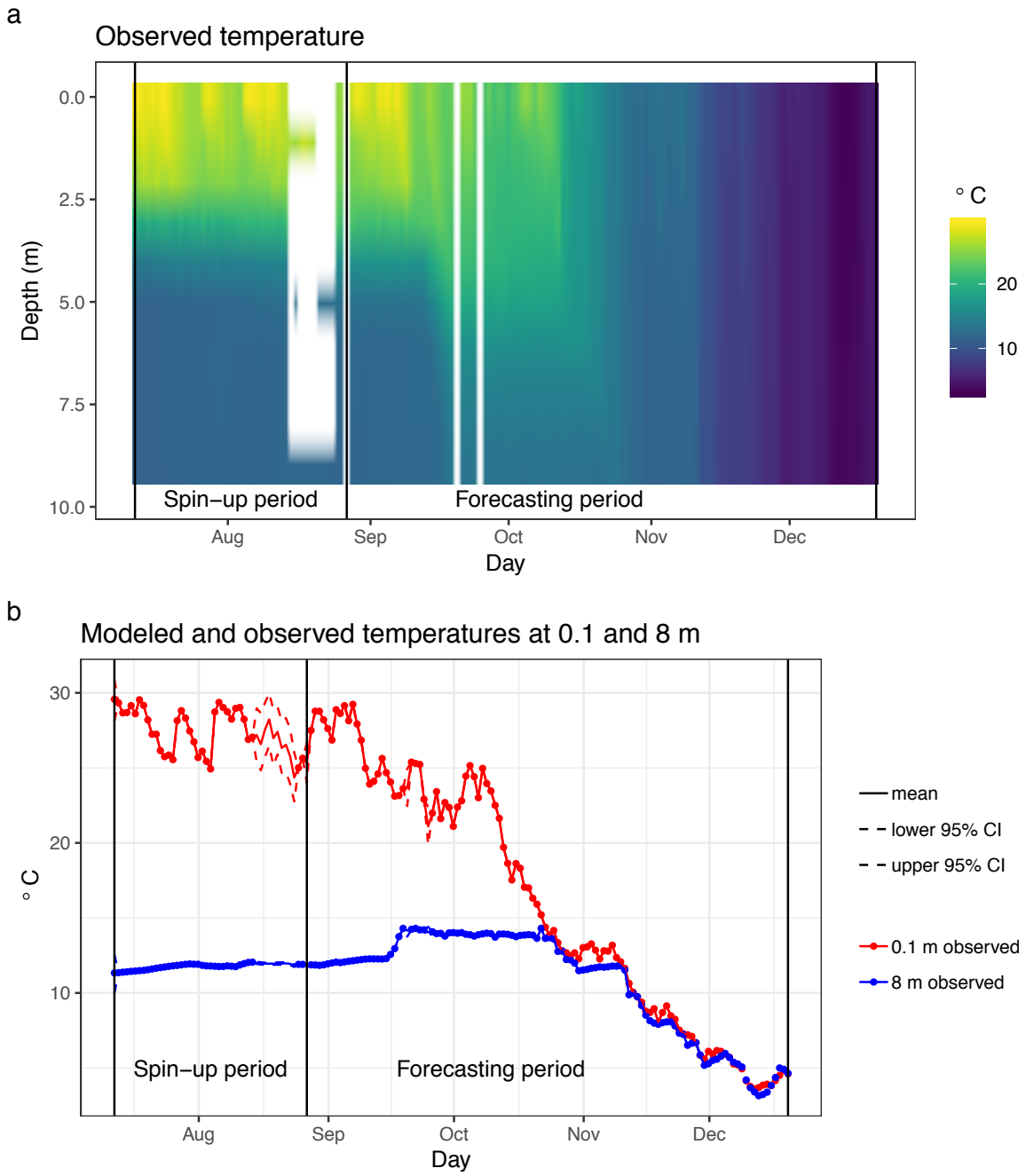

Figure 2. Water temperature in Falling Creek Reservoir during the spin-up period (11 July – 25 August 2018) and forecasting period (26 August to 3 December 2018). (a) A heat map showing the observed temperatures through the water column for the spin-up period when only data assimilation was used to update model temperatures and parameters and the forecasting period when data assimilation was used and a 16-day forecast was generated each day. White values designate days with missing water temperature sensor data. (b) A comparison of the observed water temperature for two depths (0.1 m: red, 8 m: blue) with the modeled temperature (0.1 m: black, 8 m: gray) following updating with data assimilation. The 95% confidence intervals are shown as dashed lines.

318 which decreased from 25.0°C to 15.2°C in the 14 days prior to 21 October. After turnover, the 1
319 and 8 m depth thermistors recorded water temperatures that had a mean difference of 0.51°C
320 (±0.4°C, 1 S.D) throughout the fall mixing period.
321
322 3.2 *GLM and ensemble Kalman filter (EnKF) performance*
323 The GLM simulation of water temperature with the daily EnKF parameter and initial
324 condition updating was able to successfully simulate reservoir water temperatures throughout the
325 spin-up period (11 July – 25 August) and forecasting period (26 August – 3 December; Figure
326 2B). Following daily assimilation of observed water temperature data, the RMSE for the water
327 temperature predicted by GLM at the surface (0.1 m) and 8.0 m depth were both 0.01 °C during
328 the spin-up and forecasting periods. During the forecasting period, these post-assimilation water
329 temperature predictions were used as initial conditions for the 16-day forecasts. There were two
330 1-day thermistor data gaps due to sensor maintenance in September (Figure 2B) that highlight
331 the value of the EnKF for updating initial conditions when observational data are unavailable.
332 The three GLM parameters were well-constrained by data assimilation (Supporting Information
333 Figure 1).
334
335 3.3 *Forecast performance*
336 Every day between 26 August and 3 December, the forecasting system generated 16-day
337 forecasts of water temperature for the entire water column, successfully capturing summer
338 stratified, fall turnover, and fall mixed conditions (Figure 3). In general, the forecast accuracy
339 was high throughout the 16-day forecast horizon, with forecasted water temperatures within
340 0.14°C for 0.1 m and 0.41°C for 8.0 m of observed temperatures throughout August-December,

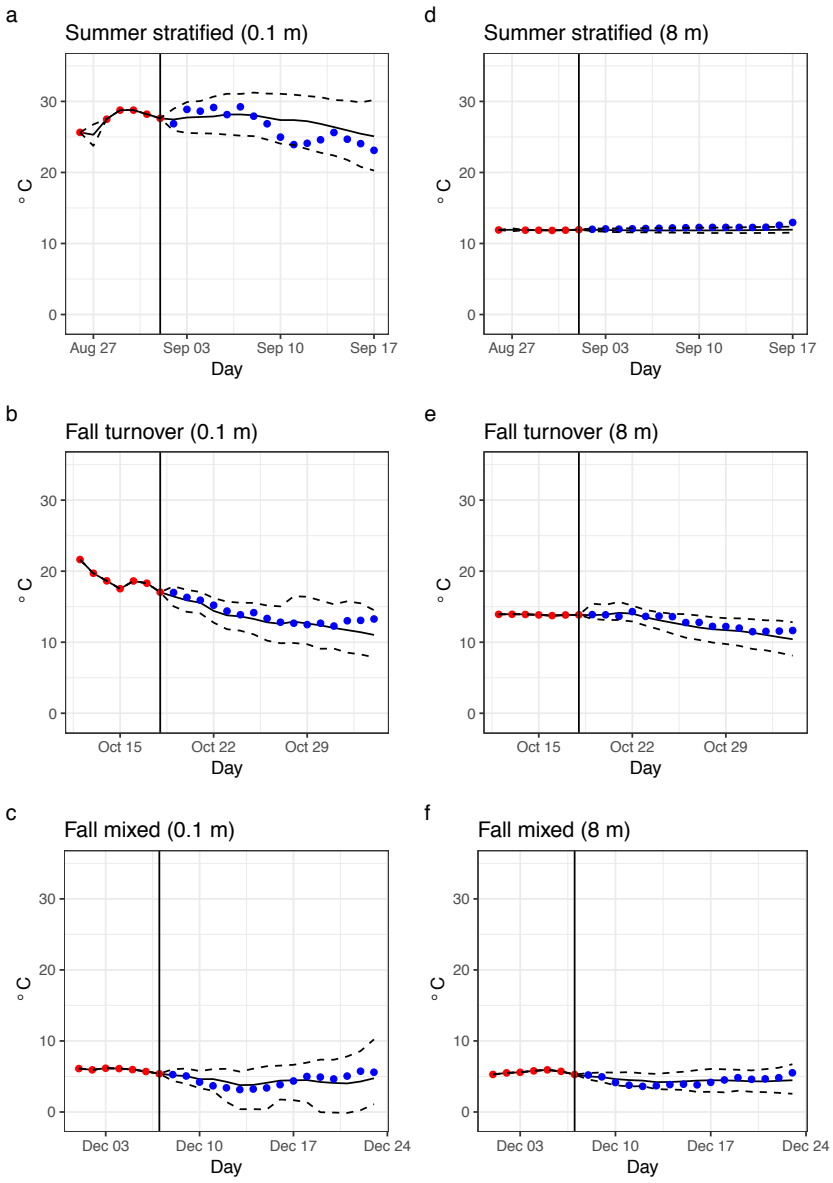

Figure 3. Forecasts successfully predicted water temperature across different lake thermal conditions (summer stratification, fall turnover, and fall mixing) in Falling Creek Reservoir. The forecasts for water temperature at the surface (0.1 m: a, b, c) and at the bottom of the reservoir (8.0 m: d, e, f) in the three thermal regime periods are shown. The vertical black line designates the day when that particular forecast was initialized. The red circles show sensor observations that were used in the data assimilation prior to the forecast generation and the blue circles show the observations during the forecast horizon and represent an independent assessment of forecast performance with data that were not used in data assimilation. The mean water temperature forecast (solid) and 95% forecast intervals (dashed) from each day's 441 forecast ensemble members are shown.

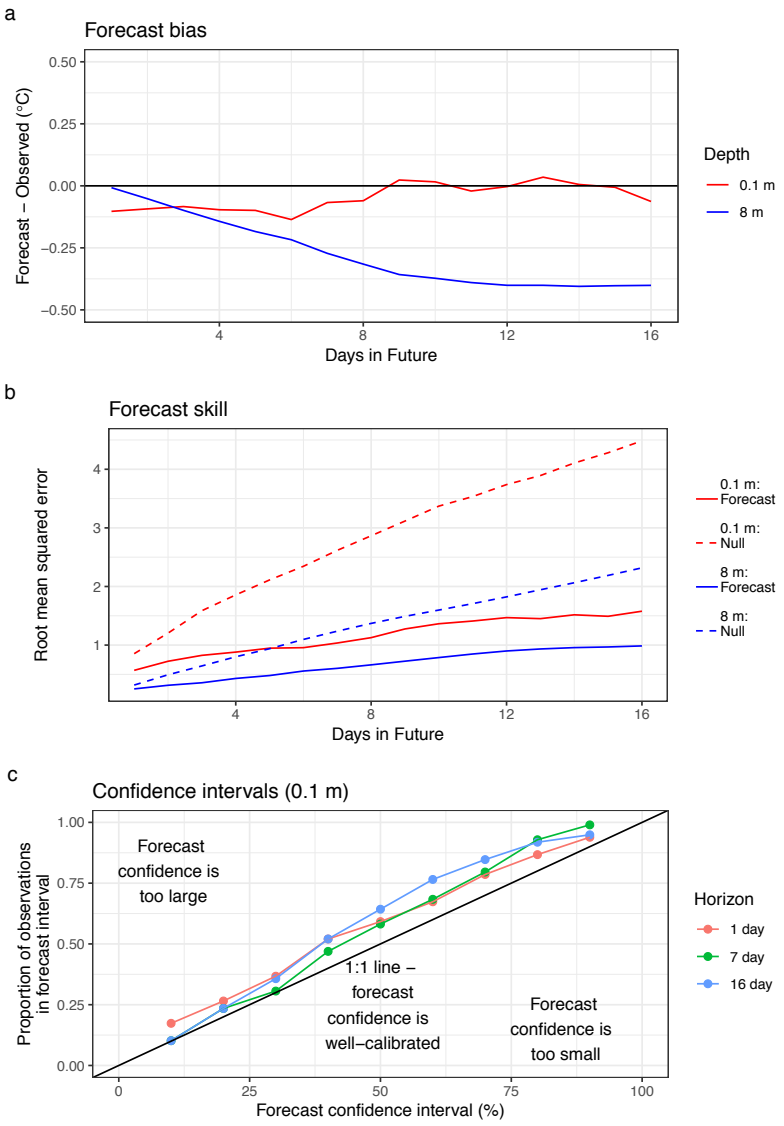

Figure 4. Water temperature forecasts for 0.1 m (red) and 8.0 m (blue) had low bias, high skill (relative to a null persistence model), and well-calibrated confidence intervals. (a) The absolute difference between the forecasted and observed water temperatures at two depths averaged across all forecasts generated during the 100-day forecasting period. (b) The root mean squared error (RMSE) of the forecasts at two depths averaged across the entire forecasting period in comparison to a null persistence model. The null model assumed that water temperature did not change over the 16-day forecast horizon. (c) The proportion of observations during the forecasting period that were within the forecast confidence interval that is specified on the x-axis. For example, the 90% forecast confidence interval is the 5th to 95th percentile of the 441 forecast ensemble members and ideally should contain 90% of water temperature observations. Values for 1, 7, and 16-day forecast horizons for the 0.1 m depth are shown.

364 averaged across all forecast time horizons (Figure 4A). Aggregated over the entire forecasting
365 period, the forecast bias for 0.1 m remained consistently low regardless of forecast horizon,
366 while the 8-m forecasts showed a small increase in bias from -0.27°C at the 7-day horizon to -
367 0.41°C at the 16-day horizon (Figure 4A). Forecast skill declined as the forecast horizon
368 increased, though it was notable that the forecasted water temperature consistently had lower
369 RMSE than a null persistence model, especially for 0.1 m temperatures throughout the 16-day
370 horizon (Figure 4B). Forecast confidence intervals generally were well-calibrated, though the
371 forecast confidence at the 16-day horizon for 0.1 m depth tended to have more observations
372 within the 40-80% forecast confidence intervals than expected (Figure 4C).
373 In a comparison of forecasting performance during the different thermal structure
374 periods, the RMSE for 0.1 m depth water temperature forecasts improved after fall turnover for
375 all forecast horizons, while the forecast bias was generally similar for all depths during the
376 summer stratified and fall mixing periods (Table 1). Surface water temperature forecasts after
377 fall turnover had 41-52% lower RMSE than forecasts before turnover for 1, 7, and 16-day
378 horizons, while RMSE for forecasts at 8.0 m did not substantially vary among forecast horizons
379 (Table 1). Bias at the two depths was similar between stratified and mixed conditions across
380 forecast horizons (Table 1). The proportion of observations within the 90% confidence intervals
381 increased by 2-30% in the post-turnover mixed period for both depths at all forecast horizons,
382 except for 0.1 m depth at the 1-day horizon (Table 1).
383 The forecasts successfully predicted the onset of fall turnover on 21 October ~8 days in
384 advance (Figure 5). At two weeks prior to turnover (7 October), the predicted chance of turnover
385 occurring on 18, 19, 20, or 21 October ranged from 31.5-45.1%. By 13 October, the predicted
386

Table 1. Forecast evaluation metrics for the forecasting period before fall turnover (during thermally-stratified conditions) and after fall turnover (during mixed conditions) at the water's surface (0.1 m) and near-sediment (8 m) depths. RMSE (the root mean squared error between observations and forecasted temperatures); bias (the difference between observations and forecasted temperatures); and the proportion of observations within the 90% confidence intervals (CI) are presented.

Forecast time horizon	Metric	Pre-turnover stratified period		Post-turnover mixed period	
		0.1 m	8.0 m	0.1 m	8.0 m
1 day	RMSE	0.70	0.21	0.41	0.29
	Bias	-0.23	-0.11	0.00	0.12
	Proportion of observations within 90% CI	0.91	0.92	0.98	0.89
7 day	RMSE	1.25	0.67	0.72	0.48
	Bias	-0.02	-0.39	-0.17	-0.11
	Proportion of observations within 90% CI	0.98	0.83	1.00	1.00
16 day	RMSE	2.11	0.97	1.01	0.97
	Bias	-0.71	-0.69	0.53	-0.02
	Proportion of observations within 90% CI	0.91	0.70	1.0	0.91

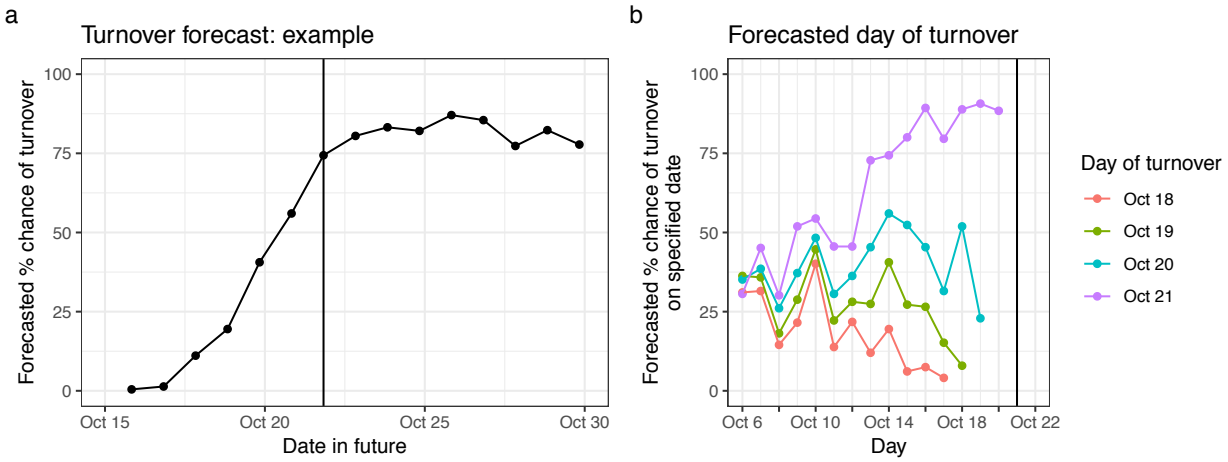

Figure 5. The forecasting system was able to anticipate the day of fall turnover 8 days in advance. (a) An example of a 16-day forecast starting on 15 October that shows the percentage of forecast ensemble members that predict the difference in temperature between 1 m and 8 m to be less than 1°C (the operational definition of turnover). The vertical line designates the day when turnover occurred (21 October). (b). The % chance of turnover for a particular day changed over time. For example, the forecast initiated on 6 October shows a relatively low and equal probability of turnover occurring on any day between 18 – 21 October. In subsequent forecasts, the probability of turnover occurring on the date when it was ultimately observed (21 October) separates from the other dates ~8 days before turnover occurred.

chance of turnover occurring on 21 October increased up to 72.8%, 27.4% higher than any of the other potential days. The chance of turnover on 21 October continued to increase over the following 8 days, while the chance of turnover occurring on the preceding days decreased to below 52% on 18 October.

3.4 *Uncertainty partitioning*

Aggregated over the three 16-day forecasts when uncertainty was partitioned (1 September 1, 18 October, and 1 December), the total forecast uncertainty was approximately two times higher for forecasts at 0.1 m than 8.0 m, but the relative importance of different components to total forecast uncertainty were similar between depths (Figure 6; Supporting Information Figures 2, 3, and 4 show the three forecasts separately). For both depths, the meteorology downscaling uncertainty and NOAA GEFS meteorology driver uncertainty were the largest contributors to the total forecast uncertainty (Figure 6). The meteorology downscaling contribution remained fairly large and constant over time, whereas the meteorology driver data uncertainty grew from near zero at the 7-day horizon to be the largest contributor of uncertainty at the 16-day horizon for the 0.1 m depth and near-equal contributor with meteorology downscaling at the 8.0 m depth. Initial condition uncertainty only contributed to total forecast uncertainty if there was a data gap on the day the forecast was generated, which decreased to near zero at the 8-day horizon for 0.1 m forecasts and the 13 day-horizon for 8.0 m forecasts (Figure 6B,D). As a result of the contributions of each of these components, total uncertainty in the 8.0 m forecasts remained fairly constant for the first 7 days in the forecast and then increased until the end of the 16-day horizon, while the total uncertainty in the 0.1 m forecasts exhibited a linear increase over time.

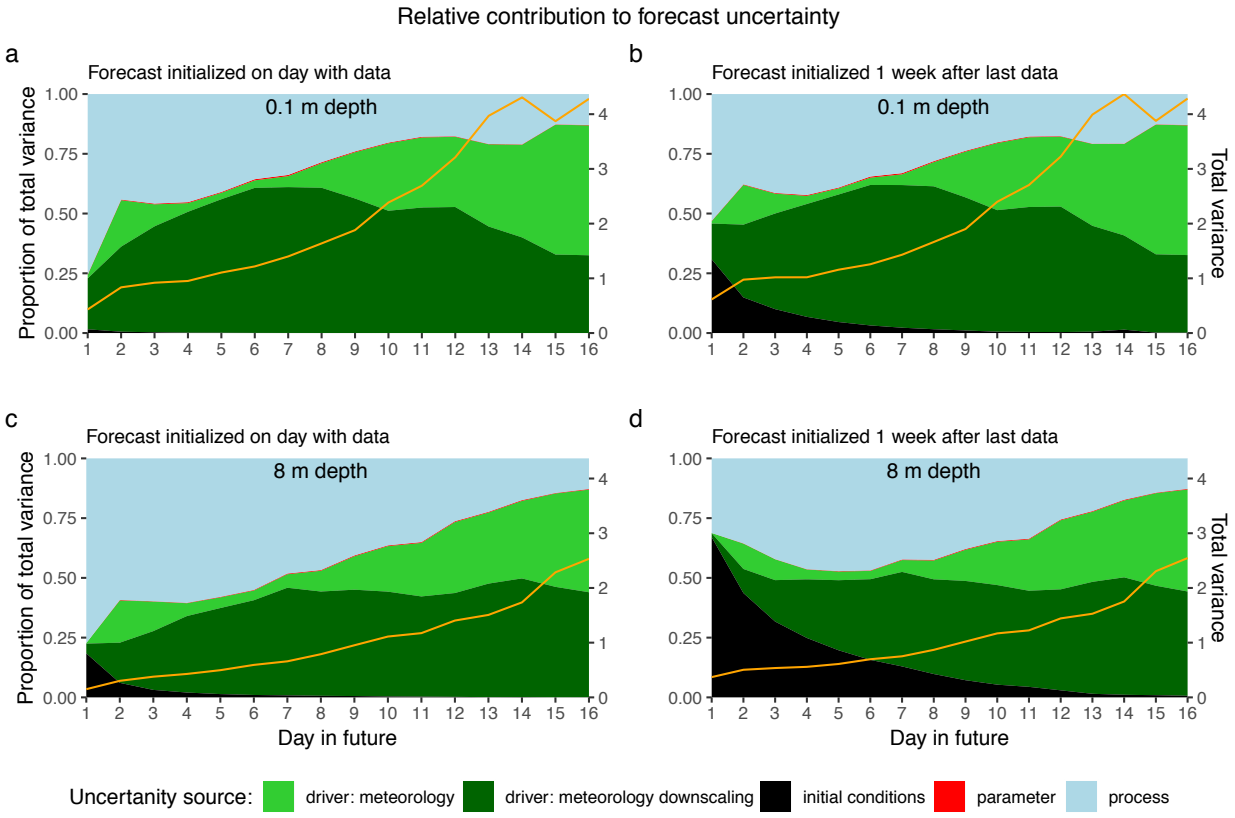

Figure 6. The relative contribution of the individual sources of uncertainty (left axis) to the total forecast uncertainty (right axis, orange line) varies through the 16-day forecast horizon. The relative contributions and total forecast variance for each day are averaged across the three 16-day forecasts shown in Figure 3: one during summer stratification (Supporting Information Figure 2), one during fall turnover (Supporting Information Figure 3), and one during the fall mixed period (Supporting Information Figure 4). Two depths are shown (0.1 m – a, b; 8.0 m – c, d) and the relative contributions of initial condition uncertainty without (left) and with (right) gaps in water temperature sensor observations are shown in the two columns.

4 Discussion

Overall, FLARE was able to forecast the water temperature on average within 0.91±0.3°C (RMSE±1 S.D.) at all depths of the reservoir over a 16-day horizon during a 100-day period that encompassed both stratified and mixed thermal conditions. Importantly, the forecasting system was able to both predict observed temperatures *and* identify the transition between stratified and mixed periods with high accuracy. In general, the forecasting system performance was similar between stratified and mixed periods (Table 1), suggesting that the system is likely robust in a range of reservoir conditions, though additional forecasts are necessary to provide a full assessment of FLARE performance. Generally, 1-D hydrodynamic models used for hindcasting aim to predict water temperature within an RMSE of 2°C [e.g., *Bruce et al. 2018, Read et al. 2014*], so the level of accuracy associated with FLARE future forecasts exceeds expectations.

We found that process uncertainty was the most important source of uncertainty early in the 16-day forecast but that driver data uncertainty dominated by the end of the forecasting period. This finding is consistent between the surface and deep depths at the reservoir and across summer stratified to fall mixed conditions. Importantly, this finding does not mean that process uncertainty declines through the forecast horizon, rather, the total uncertainty grows while the relative importance of process uncertainty diminishes. Within meteorological driver data uncertainty, the role of uncertainty in the NOAA ensemble forecast is comparable to the role of uncertainty in downscaling the coarse-scale NOAA forecast to the local site using data from the meteorological station located at the reservoir. This highlights that future work should focus on evaluating whether more advanced downscaling methods, such as neutral networks [*Kumar et al.*

461 *2012*], can build better relationships between the NOAA forecasts and the local meteorological
462 station.

463 Our uncertainty results are comparable to Dietze [*2017b*], who partitioned the uncertainty
464 in net ecosystem exchange carbon flux forecasts in a forest over 16-day horizons. In both studies,
465 the NOAA ensemble driver data uncertainty dominated total forecast uncertainty at the end of
466 the 16-day horizon. Dietze [*2017b*] found that process uncertainty was more important than the
467 meteorological driver data uncertainty early in the 16-day forecast horizon, though that study did
468 not include NOAA forecast downscaling uncertainty. Overall, while that study [*Dietze 2017b*]
469 and ours are only two examples of forecast uncertainty partitioning, they do provide insight from
470 very different ecosystems (reservoir water temperatures vs. forest carbon fluxes) and together
471 suggest that meteorological driver data are an important contributor of uncertainty in near-term
472 iterative forecasts.

473 We found that initial condition uncertainty in the forecasts was relatively small and well-
474 constrained by assimilating high-frequency temperature sensor data using the ensemble Kalman
475 filter. However, high-frequency temperature data are rarely available in real-time for most lakes
476 and reservoirs because of cost and logistics [*Marce et al. 2016*]. Here, we show that daily
477 iterative forecasting using only weekly observational data to constrain initial conditions increases
478 the role of initial condition uncertainty in the total forecast uncertainty but its role declines to
479 zero through the 16-day forecast horizon. Moreover, the total uncertainty early in the forecast
480 horizon is still low relative to the end of the 16-day horizon regardless of using a daily or weekly
481 water temperature sampling frequency. This highlights that automated temperature sensors may
482 not be a requirement for developing water temperature forecasts, depending the level of
483 acceptable forecast uncertainty over the 1 to 10-day forecast horizon. Thus, it is still possible to

484 generate accurate water temperature forecasts using only manually-collected, weekly
485 temperature profiles.

486 We anticipate that FLARE may be of interest to managers of other lakes and reservoirs
487 where water temperature forecasts up to 16 days in advance can provide decision support. For
488 Falling Creek Reservoir, an 8-day window to anticipate fall turnover provides sufficient time for
489 managers to buy additional chemicals in preparation for the water quality impairment that occurs
490 during turnover, change staffing schedules, and reprioritize management operations [*WVWA,*
491 *unpubl. data*]. Given FCR's small surface area and shallow depth, it would be expected that the
492 forecasting system would be more accurate in larger, deeper lakes that are less sensitive to
493 meteorological forcing. However, the difference in forecasts generated from the null persistence
494 model and FLARE would likely be smaller in those bigger systems. We anticipate that FLARE
495 would be most useful for managing water bodies that experience dynamic mixing throughout the
496 year and systems where water quality, fish habitat, and reservoir withdrawals are tightly coupled
497 to water temperature.

498 While the FLARE forecasting system presented here was able to predict water
499 temperature over the 16-day time horizon with low bias and an RSME substantially better than a
500 null persistence model, there are important areas for potential improvement. First, FLARE is
501 using a physics-based, 1-D hydrodynamic model for predictions. Other approaches, such as
502 machine learning or hybrids between machine learning and physics-based models [e.g., *Jia et al.*
503 *2019*], could reduce forecast uncertainty. However, machine learning-based methods must be
504 able to fully quantify forecast uncertainty to be comparable to our process model-based
505 approach. Second, our parameter uncertainty is likely unrealistically low because we only
506 included three parameters in the EnKF data assimilation. While the ability to estimate parameters

using EnKF is well-established, the EnKF method is not specifically designed to estimate parameter distributions like Bayesian Monte-Carlo Markov Chain methods [*Dietze 2017a*]. A current limitation to implementing a Bayesian Monte-Carlo Markov Chain approach is the computation time to execute the GLM simulation within a daily iterative workflow. Future work that uses emulators of GLM may be able to speed computation and allow for more robust estimation of the joint distribution model of parameters that represent both prior knowledge and observed data [e.g., *Fer et al. 2018*]. Finally, our implementation of FLARE at FCR used historical inflow data in the forecast. Ideally, the forecasts would use a watershed hydrology model to link the precipitation forecast to inflow driver data.

Overall, our study demonstrates the utility of a workflow for lake and reservoir water temperature forecasting that can be applied to other waterbodies. In addition, FLARE builds the foundation for future water quality data assimilation and forecasting because ecosystem models can easily be coupled to the hydrodynamic model, enabling predictions of dissolved oxygen, algal blooms, and biogeochemical cycling with uncertainty [e.g., *Hipsey et al. 2013, Page et al. 2018, Zwart et al. 2019*]. Importantly, FLARE provides a method for partitioning uncertainty in forecasts that identifies how to prioritize future research to increase confidence in forecasts. Given the pressing need for tools to anticipate the increasing variability of freshwater ecosystems, near-term iterative forecasting systems such as FLARE provide the ability to anticipate future change for stakeholders, managers, and policy-makers.

Acknowledgments, Data, and Author Contributions

Data and Code Availability: All data used in this study are available in the Environmental Data Initiative repository [*Carey et al. 2019a, Carey et al. 2019b, Carey et al. 2018*]. Code for

FLARE can be found at: https://github.com/CareyLabVT/FLARE/releases/tag/v1.0_beta.1.03

Author Contributions: RQT, CCC, and RJF co-developed the forecasting system framework and workflow. RQT developed the data assimilation system, RJF and VD developed the cyberinfrastructure, CCC and BJB deployed the water quality sensors, BJB and VD maintained the data workflows, and LKP and RQT developed and applied the meteorology downscaling technique. CCC and RQT wrote the manuscript; all authors provided feedback and approved the final version.

Funding and Support: This work was supported by the U.S. National Science Foundation (CNS-1737424, DEB-1753639, and EF-1702506); the Virginia Tech Global Change Center; and Fralin Life Institute. We thank the Smart Reservoir team for their helpful feedback on the project; the Western Virginia Water Authority for their long-term support and access to field sites; and Mary Lofton, Ryan McClure, and Whitney Woelmer for their critical help in the field. The authors declare no conflicts of interest.

References

Bowman, D.C. (2019). rNOMADS: An R Interface to the NOAA Operational Model Archive and Distribution System.

Brookes, J.D., Carey, C.C., Hamilton, D.P., Ho, L., van der Linden, L., Renner, R. and Rigosi, A. (2014), Emerging challenges for the drinking water industry. Environ Sci Technol 48(4), 2099-2101, https://doi.org/10.1021/es405606t

Bruce, L.C., Frassl, M.A., Arhonditsis, G.B., Gal, G., Hamilton, D.P., Hanson, P.C., Hetherington, A.L., Melack, J.M., Read, J.S., Rinke, K., Rigosi, A., Trolle, D., Winslow, L., Adrian, R., Ayala, A.I., Bocaniov, S.A., Boehrer, B., Boon, C., Brookes, J.D., Bueche, T.,

553 Busch, B.D., Copetti, D., Cortés, A., de Eyto, E., Elliott, J.A., Gallina, N., Gilboa, Y., Guyennon,
554 N., Huang, L., Kerimoglu, O., Lenters, J.D., MacIntyre, S., Makler-Pick, V., McBride, C.G.,
555 Moreira, S., Özkundakci, D., Pilotti, M., Rueda, F.J., Rusak, J.A., Samal, N.R., Schmid, M.,
556 Shatwell, T., Snorthheim, C., Soulignac, F., Valerio, G., van der Linden, L., Vetter, M., Vinçon-
557 Leite, B., Wang, J., Weber, M., Wickramaratne, C., Woolway, R.I., Yao, H. and Hipsey, M.R.
558 (2018), A multi-lake comparative analysis of the General Lake Model (GLM): Stress-testing
559 across a global observatory network. Environ. Model. Softw. 102, 274-291,
560 https://doi.org/10.1016/j.envsoft.2017.11.016

561 Butcher, J.B., Nover, D., Johnson, T.E. and Clark, C.M. (2015), Sensitivity of lake thermal and
562 mixing dynamics to climate change. Clim. Change 129(1-2), 295-305,
563 https://doi.org/10.1007/s10584-015-1326-1

564 Çalışkan, A. and Elçi, Ş. (2008), Effects of Selective Withdrawal on Hydrodynamics of a
565 Stratified Reservoir. Water Resour Manage 23(7), 1257-1273, https://doi.org/10.1007/s11269-
566 008-9325-x

567 Carey, C.C., Bookout, B.J., Lofton, M.E. and McClure, R.P. (2019a), Time series of high-
568 frequency meteorological data at Falling Creek Reservoir, Virginia, USA 2015-2018.,
569 Environmental Data Initiative,
570 https://doi.org/10.6073/pasta/68de79f732a9f3d2a686dda2eeb8197d

571 Carey, C.C., Bookout, B.J. and Woelmer, W.M. (2019b), Time series of high-frequency sensor
572 data measuring water temperature, dissolved oxygen, conductivity, specific conductivity, total
573 dissolved solids, chlorophyll a, phycocyanin, and fluorescent dissolved organic matter at discrete

574 depths in Falling Creek Reservoir, Virginia, USA in 2018, Environmental Data Initiative,
575 https://doi.org/10.6073/pasta/128777b056a4b2bfba99ff8e780e2d59

576 Carey, C.C., Gerling, A.B., McClure, R.P., Lofton, M.E. and Bookout, B.J. (2018), Discharge
577 time series for the primary inflow tributary entering Falling Creek Reservoir, Vinton, Virginia,
578 USA 2013-2018., Environmental Data Initiative,
579 https://doi.org/10.6073/pasta/64ff214b987da2997f5c823b156b3334

580 Carey, C.C., Ibelings, B.W., Hoffmann, E.P., Hamilton, D.P. and Brookes, J.D. (2012), Eco-
581 physiological adaptations that favour freshwater cyanobacteria in a changing climate. Water Res
582 46(5), 1394-1407, https://doi.org/10.1016/j.watres.2011.12.016

583 Carey, C.C., McClure, R.P., Gerling, A.B., Doubek, J.P., Chen, S., Lofton, M.E. and Hamre,
584 K.D. (2019c), Time series of high-frequency profiles of depth, temperature, dissolved oxygen,
585 conductivity, specific conductivity, chlorophyll a, turbidity, pH, oxidation-reduction potential,
586 photosynthetic active radiation, and descent rate for Beaverdam Reservoir, Carvins Cove
587 Reservoir, Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in
588 Southwestern Virginia, USA 2013-2018,
589 https://doi.org/10.6073/pasta/8f19c5d19d816857e55077ba20570265

590 Casamitjana, X., Serra, T., Colomer, J., Baserba, C. and Pérez-Losada, J. (2003), Effects of the
591 water withdrawal in the stratification patterns of a reservoir. Hydrobiologia 504(1-3), 21-28,
592 https://doi.org/10.1023/b:Hydr.0000008504.61773.77

593 Chen, S., Little, J.C., Carey, C.C., McClure, R.P., Lofton, M.E. and Lei, C. (2018), Three-
594 Dimensional Effects of Artificial Mixing in a Shallow Drinking-Water Reservoir. Water Resour.
595 Res. 54(1), 425-441, https://doi.org/10.1002/2017wr021127

596 Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M.,
597 Pascual, M., Pielke Jr, R., Pizer, W., Pringle, C., Reid, W.V., Rose, K.A., Sala, O., Schlesinger,
598 W.H., Wall, D.H. and Wear, D. (2001), Ecological Forecasts: An Emerging Imperative. Science
599 293(5530), 657-660, https://doi.org/10.1126/science.293.5530.657

600 Cooke, G.D., Welch, E., B., Peterson, S. and Nichols, S.A. (2005), Restoration and Management
601 of Lakes and Reservoirs, CRC Press, Boca Raton, FL.

602 Delpla, I., Jung, A.V., Baures, E., Clement, M. and Thomas, O. (2009), Impacts of climate
603 change on surface water quality in relation to drinking water production. Environ Int 35(8),
604 1225-1233, https://doi.org/10.1016/j.envint.2009.07.001

605 Dietze, M.C. (2017a), Ecological Forecasting, Princeton University Press, Princeton.

606 Dietze, M.C. (2017b), Prediction in ecology: a first-principles framework. Ecol Appl 112(1),
607 6252-6213, https://doi.org/10.1002/eap.1589

608 Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S.,
609 Keitt, T.H., Kenney, M.A., Laney, C.M., Larsen, L.G., Loescher, H.W., Lunch, C.K.,
610 Pijanowski, B.C., Randerson, J.T., Read, E.K., Tredennick, A.T., Vargas, R., Weathers, K.C. and
611 White, E.P. (2018), Iterative near-term ecological forecasting: Needs, opportunities, and
612 challenges. Proc Natl Acad Sci U S A 115(7), 1424-1432,
613 https://doi.org/10.1073/pnas.1710231115

614 Effler, S.W. and Matthews, D.A. (2008), Implications of redox processes for the rehabilitation of
615 an urban lake, Onondaga Lake, New York. Lake Reserv Manag 24(2), 122-138,
616 https://doi.org/10.1080/07438140809354056

617 Evensen, G. (2003), The Ensemble Kalman Filter: theoretical formulation and practical
618 implementation. Ocean Dyn. 53(4), 343-367, https://doi.org/10.1007/s10236-003-0036-9

619 Evensen, G. (2009), Data Assimilation, Springer Berlin Heidelberg, Berlin, Heidelberg.

620 Fer, I., Kelly, R., Moorcroft, P.R., Richardson, A.D., Cowdery, E.M. and Dietze, M.C. (2018),
621 Linking big models to big data: efficient ecosystem model calibration through Bayesian model
622 emulation. Biogeosciences 15(19), 5801-5830, https://doi.org/10.5194/bg-15-5801-2018

623 Gerling, A.B., Browne, R.G., Gantzer, P.A., Mobley, M.H., Little, J.C. and Carey, C.C. (2014),
624 First report of the successful operation of a side stream supersaturation hypolimnetic
625 oxygenation system in a eutrophic, shallow reservoir. Water Res 67, 129-143,
626 https://doi.org/10.1016/j.watres.2014.09.002

627 Gerling, A.B., Munger, Z.W., Doubek, J.P., Hamre, K.D., Gantzer, P.A., Little, J.C. and Carey,
628 C.C. (2016), Whole-Catchment Manipulations of Internal and External Loading Reveal the
629 Sensitivity of a Century-Old Reservoir to Hypoxia. Ecosystems 19(3), 555-571,
630 https://doi.org/10.1007/s10021-015-9951-0

631 Hipsey, M.R., Bruce, L.C., Boon, C., Busch, B., Carey, C.C., Hamilton, D.P., Hanson, P.C.,
632 Read, J.S., de Sousa, E., Weber, M. and Winslow, L.A. (2019), A General Lake Model (GLM
633 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory

634 Network (GLEON). Geosci. Model Dev. 12(1), 473-523, https://doi.org/10.5194/gmd-12-473-
635 2019

636 Hipsey, M.R., Bruce, L.C. and Hamilton, D.P. (2013). Aquatic Ecodynamics (AED) Model
637 Library Science Manual,
638 http://aed.see.uwa.edu.au/research/models/AED/Download/AED_ScienceManual_v4_draft.pdf.

639 Jia, X., Willard, J., Karpatne, A., Read, J., Zwart, J., Steinbach, M. and Kumar, V. (Year),
640 Physics Guided RNNs for Modeling Dynamical Systems: A Case Study in Simulating Lake
641 Temperature Profiles. Proceedings of the 2019 SIAM International Conference on Data Mining.
642 https://doi.org/10.1137/1.9781611975673.63

643 Jöhnk, K.D., Huisman, J.E.F., Sharples, J., Sommeijer, B.E.N., Visser, P.M. and Stroom, J.M.
644 (2008), Summer heatwaves promote blooms of harmful cyanobacteria. Glob. Change Biol. 14(3),
645 495-512, https://doi.org/10.1111/j.1365-2486.2007.01510.x

646 Klug, J.L., Richardson, D.C., Ewing, H.A., Hargreaves, B.R., Samal, N.R., Vachon, D., Pierson,
647 D.C., Lindsey, A.M., O'Donnell, D.M., Effler, S.W. and Weathers, K.C. (2012), Ecosystem
648 effects of a tropical cyclone on a network of lakes in northeastern North America. Environ Sci
649 Technol 46(21), 11693-11701, https://doi.org/10.1021/es302063v

650 Kumar, J., Brooks, B.-G.J., Thornton, P.E. and Dietze, M.C. (2012), Sub-daily Statistical
651 Downscaling of Meteorological Variables Using Neural Networks. Procedia Comput. Sci. 9,
652 887-896, https://doi.org/10.1016/j.procs.2012.04.095

653 Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J.S. and Schimel, D.S. (2011),
654 Ecological forecasting and data assimilation in a data-rich era. Ecol Appl 21(5), 1429-1442,
655 https://doi.org/10.1890/09-1275.1

656 Marce, R., George, G., Buscarinu, P., Deidda, M., Dunalska, J., de Eyto, E., Flaim, G., Grossart,
657 H.P., Istvanovics, V., Lenhardt, M., Moreno-Ostos, E., Obrador, B., Ostrovsky, I., Pierson, D.C.,
658 Potuzak, J., Poikane, S., Rinke, K., Rodriguez-Mozaz, S., Staehr, P.A., Sumberova, K., Waajen,
659 G., Weyhenmeyer, G.A., Weathers, K.C., Zion, M., Ibelings, B.W. and Jennings, E. (2016),
660 Automatic high frequency monitoring for improved lake and reservoir management. Environ Sci
661 Technol 50(20), 10780-10794, https://doi.org/10.1021/acs.est.6b01604

662 McClure, R.P., Hamre, K.D., Niederlehner, B.R., Munger, Z.W., Chen, S., Lofton, M.E.,
663 Schreiber, M.E. and Carey, C.C. (2018), Metalimnetic oxygen minima alter the vertical profiles
664 of carbon dioxide and methane in a managed freshwater reservoir. Sci Total Environ 636, 610-
665 620, https://doi.org/10.1016/j.scitotenv.2018.04.255

666 Mi, C., Sadeghian, A., Lindenschmidt, K.-E. and Rinke, K. (2019), Variable withdrawal
667 elevations as a management tool to counter the effects of climate warming in Germany's largest
668 drinking water reservoir. Environ Sci Eur 31(1), 19, https://doi.org/10.1186/s12302-019-0202-4

669 Millennium Ecosystem Assessment (2005), Ecosystems and Human Well-being: Synthesis,
670 Island Press, Washington, DC.

671 Page, T., Smith, P.J., Beven, K.J., Jones, I.D., Elliott, J.A., Maberly, S.C., Mackay, E.B., De
672 Ville, M. and Feuchtmayr, H. (2018), Adaptive forecasting of phytoplankton communities.
673 Water Res 134, 74-85, https://doi.org/10.1016/j.watres.2018.01.046

674 Read, J.S., Hamilton, D.P., Jones, I.D., Muraoka, K., Winslow, L.A., Kroiss, R., Wu, C.H. and
675 Gaiser, E. (2011), Derivation of lake mixing and stratification indices from high-resolution lake
676 buoy data. Environ. Model. Softw. 26(11), 1325-1336,
677 https://doi.org/10.1016/j.envsoft.2011.05.006

678 Read, J.S., Winslow, L.A., Hansen, G.J.A., Van den Hoek, J., Hanson, P.C., Bruce, L.C. and
679 Markfort, C.D. (2014), Simulating 2368 temperate lakes reveals weak coherence in stratification
680 phenology. Ecol Model 291, 142-150, https://doi.org/10.1016/j.ecolmodel.2014.07.029

681 Schmidt, S.R., Gerten, D., Hintze, T., Lischeid, G., Livingstone, D.M. and Adrian, R. (2018),
682 Temporal and spatial scales of water temperature variability as an indicator for mixing in a
683 polymictic lake. Inland Waters 8(1), 82-95, https://doi.org/10.1080/20442041.2018.1429067

684 Sharma, S., Gray, D.K., Read, J.S., O'Reilly, C.M., Schneider, P., Qudrat, A., Gries, C.,
685 Stefanoff, S., Hampton, S.E., Hook, S., Lenters, J.D., Livingstone, D.M., McIntyre, P.B., Adrian,
686 R., Allan, M.G., Anneville, O., Arvola, L., Austin, J., Bailey, J., Baron, J.S., Brookes, J., Chen,
687 Y., Daly, R., Dokulil, M., Dong, B., Ewing, K., de Eyto, E., Hamilton, D., Havens, K., Haydon,
688 S., Hetzenauer, H., Heneberry, J., Hetherington, A.L., Higgins, S.N., Hixson, E., Izmest'eva,
689 L.R., Jones, B.M., Kangur, K., Kasprzak, P., Koster, O., Kraemer, B.M., Kumagai, M., Kuusisto,
690 E., Leshkevich, G., May, L., MacIntyre, S., Muller-Navarra, D., Naumenko, M., Noges, P.,
691 Noges, T., Niederhauser, P., North, R.P., Paterson, A.M., Plisnier, P.D., Rigosi, A., Rimmer, A.,
692 Rogora, M., Rudstam, L., Rusak, J.A., Salmaso, N., Samal, N.R., Schindler, D.E., Schladow, G.,
693 Schmidt, S.R., Schultz, T., Silow, E.A., Straile, D., Teubner, K., Verburg, P., Voutilainen, A.,
694 Watkinson, A., Weyhenmeyer, G.A., Williamson, C.E. and Woo, K.H. (2015), A global database

695 of lake surface temperatures collected by in situ and satellite methods from 1985-2009. Sci Data
696 2, 150008, https://doi.org/10.1038/sdata.2015.8

697 Snortheim, C.A., Hanson, P.C., McMahon, K.D., Read, J.S., Carey, C.C. and Dugan, H.A.
698 (2017), Meteorological drivers of hypolimnetic anoxia in a eutrophic, north temperate lake. Ecol
699 Model 343, 39-53, https://doi.org/10.1016/j.ecolmodel.2016.10.014

700 Stocker, F.T., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y.,
701 Bex, V. and Midgley, P.M. (2013), Climate Change 2013: The Physical Science Basis.
702 Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel
703 on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York,
704 NY, USA.

705 Ummenhofer, C.C. and Meehl, G.A. (2017), Extreme weather and climate events with ecological
706 relevance: a review. Philos. Trans. Royal Soc. B 372(1723), 20160135-20160113,
707 https://doi.org/10.1098/rstb.2016.0135

708 Weber, M., Rinke, K., Hipsey, M.R. and Boehrer, B. (2017), Optimizing withdrawal from
709 drinking water reservoirs to reduce downstream temperature pollution and reservoir hypoxia. J
710 Environ Manage 197, 96-105, https://doi.org/10.1016/j.jenvman.2017.03.020

711 West, J.M., Julius, S.H., Kareiva, P., Enquist, C., Lawler, J.J., Petersen, B., Johnson, A.E. and
712 Shaw, M.R. (2009), U.S. natural resources and climate change: concepts and approaches for
713 management adaptation. Environ Manage 44(6), 1001-1021, https://doi.org/10.1007/s00267-009-
714 9345-1

715 Zhang, H., Hendricks Franssen, H.-J., Han, X., Vrugt, J.A. and Vereecken, H. (2017), State and
716 parameter estimation of two land surface models using the ensemble Kalman filter and the
717 particle filter. Hydrol. Earth Syst. Sci. 21(9), 4927-4958, https://doi.org/10.5194/hess-21-4927-
718 2017

719 Zwart, J.A., Hararuk, O., Prairie, Y.T., Jones, S.E. and Solomon, C.T. (2019), Improving
720 estimates and forecasts of lake carbon dynamics using data assimilation. Limnol. Oceanogr.:
721 Methods 17(2), 97-111, https://doi.org/10.1002/lom3.10302

722

723