Received: 24 May 2023 | Revised: 18 October 2023 | Accepted: 31 October 2023

DOI: 10.1002/ecs2.4752

ECOSPHERE
AN ESA OPEN ACCESS JOURNAL

ARTICLE

Methods, Tools, and Technologies

Data assimilation experiments inform monitoring needs for near-term ecological forecasts in a eutrophic reservoir

Heather L. Wander[1] | **R. Quinn Thomas**[1,2] | **Tadhg N. Moore**[1,2] | **Mary E. Lofton**[1] | **Adrienne Breef-Pilz**[1] | **Cayelan C. Carey**[1]

[1]Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA

[2]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, USA

Correspondence
Heather L. Wander
Email: hwander@vt.edu

Funding information
U.S. National Science Foundation, Grant/Award Numbers: DBI-1933016, DBI-1933102, DEB-1753639, DEB-1926050, DEB-1926388, OAC-2004323

Handling Editor: Jean-François Lapierre

Abstract

Ecosystems around the globe are experiencing changes in both the magnitude and fluctuations of environmental conditions due to land use and climate change. In response, ecologists are increasingly using near-term, iterative ecological forecasts to predict how ecosystems will change in the future. To date, many near-term, iterative forecasting systems have been developed using high temporal frequency (minute to hourly resolution) data streams for assimilation. However, this approach may be cost-prohibitive or impossible for forecasting ecological variables that lack high-frequency sensors or have high data latency (i.e., a delay before data are available for modeling after collection). To explore the effects of data assimilation frequency on forecast skill, we developed water temperature forecasts for a eutrophic drinking water reservoir and conducted data assimilation experiments by selectively withholding observations to examine the effect of data availability on forecast accuracy. We used in situ sensors, manually collected data, and a calibrated water quality ecosystem model driven by forecasted weather data to generate future water temperature forecasts using Forecasting Lake and Reservoir Ecosystems (FLARE), an open source water quality forecasting system. We tested the effect of daily, weekly, fortnightly, and monthly data assimilation on the skill of 1- to 35-day-ahead water temperature forecasts. We found that forecast skill varied depending on the season, forecast horizon, depth, and data assimilation frequency, but overall forecast performance was high, with a mean 1-day-ahead forecast root mean square error (RMSE) of 0.81°C, mean 7-day RMSE of 1.15°C, and mean 35-day RMSE of 1.94°C. Aggregated across the year, daily data assimilation yielded the most skillful forecasts at 1- to 7-day-ahead horizons, but weekly data assimilation resulted in the most skillful forecasts at 8- to 35-day-ahead horizons. Within a year, forecasts with weekly data assimilation consistently outperformed forecasts with daily data assimilation after the 8-day forecast horizon during mixed spring/autumn periods and 5- to 14-day-ahead horizons during the summer-stratified period, depending on depth. Our results suggest that lower

frequency data (i.e., weekly) may be adequate for developing accurate forecasts in some applications, further enabling the development of forecasts broadly across ecosystems and ecological variables without high-frequency sensor data.

KEYWORDS
data collection frequency, FLARE, high-frequency sensors, initial conditions, observations, uncertainty, water temperature

INTRODUCTION

In the face of increasing ecological variability due to climate and land use change, in which ecosystems are experiencing changes in the magnitude and fluctuations of environmental variables (e.g., Gilarranz et al., 2022; Malhi et al., 2020), ecological forecasting is increasingly being used for understanding and predicting future ecological change (Carey, Woelmer, et al., 2022; Lewis et al., 2022). Here, we define ecological forecasts as predictions of future environmental conditions with quantified uncertainty (see Carey, Woelmer, et al., 2022; Lewis et al., 2022). Because of their broad utility, forecasts are increasingly being developed by the research community to predict population, community, and ecosystem dynamics (Lewis et al., 2022). For example, an ongoing, community-based forecasting challenge organized by the Ecological Forecasting Initiative's Research Coordination Network has received thousands of ecological forecast submissions of National Ecological Observatory Network (NEON) data (e.g., lake water temperature, tick abundances, forest net ecosystem production, beetle communities) before the data have been collected (Thomas, Boettiger, et al., 2023).

Many near-term (daily to decadal) ecological forecasts are produced using the iterative, near-term forecasting cycle, in which models are updated as new observational data become available to generate forecasts into the future with quantified uncertainty (Dietze et al., 2018). The process of updating forecast models with newly available data, termed data assimilation (DA), is a critical component of the iterative, near-term forecast cycle (Dietze et al., 2018; Luo et al., 2011). DA allows for iterative updating of ecological hypotheses and models as forecasts are continuously assessed and updated with the most recent ecosystem observations (Dietze et al., 2018; White et al., 2019). DA can also improve forecast accuracy by updating forecast model initial conditions (i.e., starting values given to the model), states, and/or parameters at the time step that the new observations become available (Cho et al., 2020; Gottwald & Reich, 2021; Luo et al., 2011; Niu et al., 2014). Despite the usefulness of DA for improving forecasts, however, the optimal frequency of observations for updating ecological models to produce skillful forecasts is not well characterized.

While there are a number of best practices proposed for applying the near-term, iterative forecast cycle in ecology (e.g., Clark et al., 2001; Harris et al., 2018; Lewis et al., 2022; White et al., 2019), few recommendations exist for choosing the optimal frequency of DA to produce accurate forecasts. Specifically, determining the appropriate frequency of observations for DA across a range of ecological variables is needed to improve the scalability of ecological forecasting, particularly if accurate forecasts can be developed using lower frequency observations. For example, if weekly or fortnightly DA yielded similarly accurate lake-dissolved oxygen forecasts as daily DA, then water quality forecasting systems could be developed for lakes that have weekly or fortnightly routine monitoring program data without needing expensive high-frequency sensors, thereby enabling forecasts to be generated for many waterbodies globally. The benefits of lower frequency DA are also evident for remote sensing applications, as the frequency at which different satellites acquire data is often greater than two days (Herrick et al., 2023).

Currently, many automated ecological forecasting systems rely on high-frequency sensors to assimilate data at each time step and generate accurate forecasts (e.g., Baracchini, Wüest, et al., 2020; Corbari et al., 2019; Marj & Meijerink, 2011; Page et al., 2018; Tanut et al., 2021), but it is possible that high-frequency sensor data collection may not be needed for DA. Moreover, deployment of high-frequency sensors is not always feasible for all ecological variables (e.g., zooplankton abundance, biogeochemical concentrations; Marcé et al., 2016) and some remote locations have additional logistical constraints for maintaining autonomous sensor operation (Steere et al., 2000). Furthermore, some remotely sensed variables may only be available as satellite orbits and weather conditions (e.g., cloud cover) allow (e.g., Herrick et al., 2023). Thus, identifying how best to integrate observational data collected at different temporal frequencies into forecast models has emerged as a critical need for ecological forecasters (LaDeau et al., 2017).

Studies on the frequency of DA for environmental forecasts have generally shown that more temporally frequent DA improves forecast accuracy, but not always, which may be related to the sensitivity of forecasts to model initial conditions. For example, DA occurring every 24 h using in situ snow data (e.g., snow depth, density, snow water equivalent) resulted in better predictions of these snow variables in an alpine snowpack model compared with DA occurring every 3 h (Piazzi et al., 2018). Conversely, DA "experiments" performed for National Oceanic and Atmospheric Administration's (NOAA) Global Ensemble Forecasting System using meteorological observations collected at different frequencies showed that DA occurring every 2 h resulted in more accurate air temperature and wind speed forecasts than DA occurring every 6 h (He et al., 2020). These differences are likely because uncertainty in meteorological forecasts is primarily driven by the forecast model's initial conditions. Thus, more frequent DA, which constrains the model's initial conditions, will almost always improve the skill of meteorological forecasts (e.g., Clark et al., 2016; He et al., 2020; Simonin et al., 2017). In contrast, for forecasts of environmental systems in which model initial conditions are less important sources of uncertainty, and model process uncertainty and model driver data uncertainty dominate total uncertainty (e.g., Dietze, 2017a; Heilman et al., 2022; Lofton et al., 2022; Thomas et al., 2020), it is unknown whether more frequent DA can improve forecast skill by generating initial conditions that are more consistent with observations.

To the best of our knowledge, there have been only a few ecological DA experiments that have tested the effects of different observation frequencies on forecast skill (e.g., Massoud et al., 2018; Piazzi et al., 2018; Weng & Luo, 2011; Ziliani et al., 2019), and none that have considered how the frequency of data used for assimilation affects forecast skill across both spatial and temporal scales. Weng and Luo (2011) assimilated eight different carbon datasets (e.g., root biomass, litter fall, soil respiration), each with different collection frequencies, to identify the relative importance of these data sources in constraining long-term carbon dynamics, but did not consider how different frequencies of the same dataset could affect forecast skill. Piazzi et al. (2018) assimilated multiple snow observations at two different frequencies (3 and 24 h) for predicting different snow-related variables (e.g., depth, density, and snow water equivalent), and Ziliani et al. (2019) performed DA tests using 1- to 20-s assimilation of water depth data to assess water level forecast skill, but neither considered the effect of less frequent assimilation (e.g., >24 h). Massoud et al. (2018) performed DA tests using a wider range of temporal frequencies (e.g., ~3- to 34-day abundance data) to predict plankton community dynamics, but did not consider the effects of DA across spatial scales (i.e., how DA affects forecast skill across multiple sites or depths in an aquatic ecosystem). As a result, further work is needed to quantify the utility of increased observation and DA frequency over both time and space to forecast performance in ecological systems with varying sensitivities to initial conditions.

Among ecosystems, freshwater lakes and reservoirs are particularly important systems for developing near-term forecasts because they provide essential ecosystem services, including drinking water, food, irrigation, and recreation (Carpenter et al., 2011; Meyer et al., 1999; Williamson et al., 2016). Because freshwaters are experiencing greater variability and adverse water quality issues in response to land use and climate change (e.g., O'Reilly et al., 2015; Paerl & Paul, 2012; Woolway et al., 2021), some water managers have used forecasts to preemptively address poor water quality events (reviewed by Lofton et al., 2023). To date, iterative, near-term freshwater forecasts have been developed for a number of water quality variables, including water temperature (e.g., Carey, Woelmer, et al., 2022; Thomas, McClure, et al., 2023), dissolved oxygen (e.g., Wang et al., 2016), and phytoplankton (e.g., Page et al., 2017; Woelmer et al., 2022). These forecasts have been developed using DA with observations collected by high-frequency sensors at intervals ranging from 4 min to 24 h. However, most manual-sampling water quality monitoring programs collect observations on weekly to fortnightly scales (e.g., Francy et al., 2015; Kirchner & Neal, 2013; Romero et al., 2002), currently precluding the scaling of existing forecasting systems broadly and underscoring the need to determine whether less frequent observations can be used to produce accurate forecasts.

To quantify how DA at different frequencies affects forecast skill up to 35 days into the future, we performed DA experiments in which we separately assimilated daily, weekly, fortnightly, and monthly data into reservoir water temperature forecasts. Water temperature forecasts are used to inform management decisions on water extraction depth and preemptive water quality interventions (Georgakakos et al., 2005; Kehoe et al., 2015; Mi et al., 2020), and thus our study has much utility for both informing how best to forecast complex ecosystem dynamics and manage drinking water supplies. Our research questions were: (1) Which frequency of DA generates the most skillful water temperature forecasts? (2) How does forecast skill vary across time (specifically focusing on the mixed vs. stratified seasons within a year) and space (i.e., reservoir depth)? (3) How does DA frequency influence total forecast uncertainty and what is the relative contribution of initial condition uncertainty to total forecast uncertainty? We expected that less frequent DA would result in decreased forecast skill and

increased total uncertainty. In addition, we expected that forecast skill would be better at deeper depths, especially during thermally stratified periods (e.g., Mercado-Bettín et al., 2021; Thomas et al., 2020).

METHODS

Forecasting system overview

We applied the Forecasting Lake And Reservoir Ecosystems (FLARE) forecasting system (Thomas et al., 2020) to Beaverdam Reservoir (BVR), Virginia, USA, to produce daily water temperature forecasts for 1–35 days into the future (hereafter referred to as forecast horizon) during 1 January 2021–31 December 2021. FLARE is an open source forecasting system that incorporates real-time water quality sensor data, DA, ensemble-based forecasts, and uncertainty quantification to predict near-term water quality conditions (Thomas et al., 2020).

Forecast generation via FLARE can be summarized in four steps (Figure 1). First, 10-min resolution water temperature data were collected by sensors deployed in the reservoir (Figure 1, step 1). Second, these data were transferred to the cloud and stored in a GitHub repository, where they were downloaded daily and made available for DA (Figure 1, step 2). Simultaneously, 1- to 35-day-ahead NOAA meteorological forecasts were downloaded daily as driver data for the reservoir hydrodynamic model to generate the water temperature forecasts. Third, during the forecast generation step, DA was used to update initial conditions and parameters with the most recent observations using an ensemble Kalman filter, a numerical approach that allows for the updating of model states and parameters using data (Evensen, 2003) (Figure 1, step 3a). Following DA, the reservoir hydrodynamic model was initialized with the updated model states and parameters to produce 1- to 35-day-ahead forecasts for each 0.5-m depth interval across the water column (Figure 1, step 3b). Finally, forecast skill was assessed by comparing observed versus predicted water temperatures for each daily forecast at each depth (Figure 1, step 4). We repeated steps 3a–4 for daily, weekly, fortnightly, and monthly intervals of DA throughout the year as part of the DA experiments to compare forecast skill over time.

Study site and monitoring

BVR is a small (0.28 km^2), shallow ($Z_{max} = 11$ m), dimictic, eutrophic reservoir in southwestern Virginia, USA (37.31° N, 79.82° W; Figure 2). BVR is managed by the Western Virginia Water Authority as a secondary drinking water supply and is located in a deciduous forest catchment (Doubek et al., 2019). During a typical year, BVR is stratified from mid-March to late October and mixed from November to early March (Hounshell et al., 2021). BVR experiences summer hypolimnetic anoxia and cyanobacterial blooms, both of which are controlled by water temperature and thermal stratification (Doubek et al., 2019; Hamre et al., 2018), making forecasts of water temperature important for water quality management.

Water quality monitoring of BVR includes both manual sampling and high-frequency sensors. From 2014 to present, manual water quality sampling occurred weekly to fortnightly during the summer-stratified period and fortnightly to monthly during the remainder of the year (Carey, Lewis, McClure, et al., 2022). Starting in June 2020, high-frequency sensors were deployed in the reservoir, enabling a range of DA frequencies to be compared in this study. We deployed NexSens T-Node FR Temperature Sensors (NexSens Technology, Fairborn, OH, USA) at 1-m intervals from the surface to sediments and a YSI EXO2 sonde (YSI Incorporated, Yellow Springs, OH, USA) that monitored water temperature at 1.5 m at the deepest site in BVR (Figure 1; see Carey et al., 2023, for sensor information). These sensors collected data every 10 min, which were transmitted every 3–9 h via secure sensor gateways to a Git repository in the cloud (Carey et al., 2023; Daneshmand et al., 2021). We removed observations collected during periods of sensor maintenance, as well as depth-adjusted the data using an offset calculated from CS451 Stainless Steel Pressure Transducer (Campbell Scientific, Logan, UT, USA) to account for water level changes (Wander et al., 2023b). Because of this range in latency, or the time that it takes for data to become available for modeling after they are initially collected, we used the daily mean water temperature in our forecasting application. Following quality checks, these data were integrated into the FLARE forecasting system to produce depth-specific daily water temperature forecasts.

Hydrodynamic model configuration

For modeling reservoir hydrodynamics, we used the General Lake Model (GLM) v.3.3.0 (Hipsey et al., 2022) to forecast water temperature in BVR. GLM is an open source, 1-D process-based hydrodynamic model commonly used within the freshwater research community to simulate water quality in lakes and reservoirs (Hipsey et al., 2019). GLM uses a Lagrangian approach for

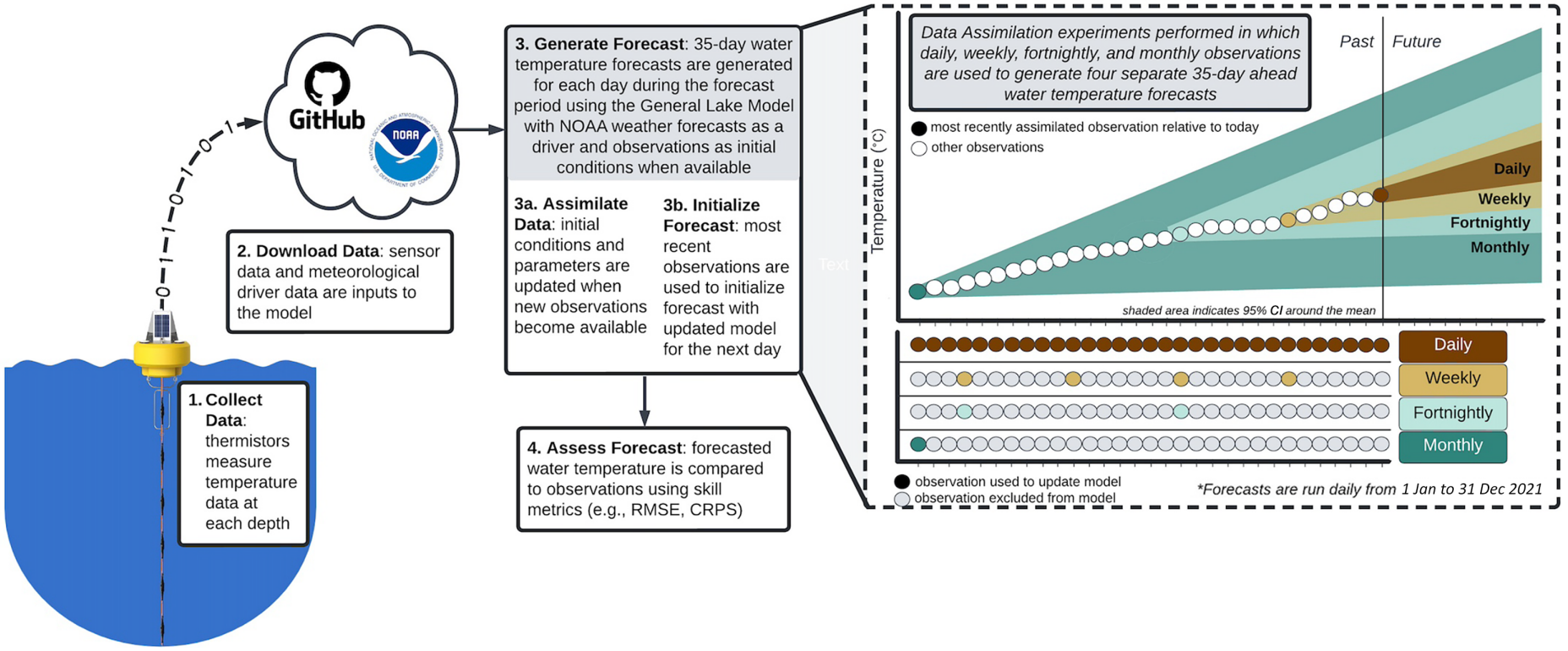

FIGURE 1 Forecasting Lake And Reservoir Ecosystems workflow showing the step-by-step process for generating daily water temperature forecasts, starting with data collection from thermistors deployed in the reservoir (step 1), then data access for running the forecast model (step 2), then generation of forecasts with data assimilation (step 3), and ending with forecast skill assessment (step 4). During the data assimilation steps (3a–b), data assimilation experiments were performed with four different data assimilation frequencies (daily, weekly, fortnightly, and monthly; see dashed line box). Steps 1–4 occurred throughout the entire forecast period (1 January–31 December 2021). CRPS, continuous ranked probability score; NOAA, National Oceanic and Atmospheric Administration; RMSE, root mean square error. Buoy figure via NexSens Technology Inc., CC by 2.0 https://creativecommons.org/licenses/by/2.0/.

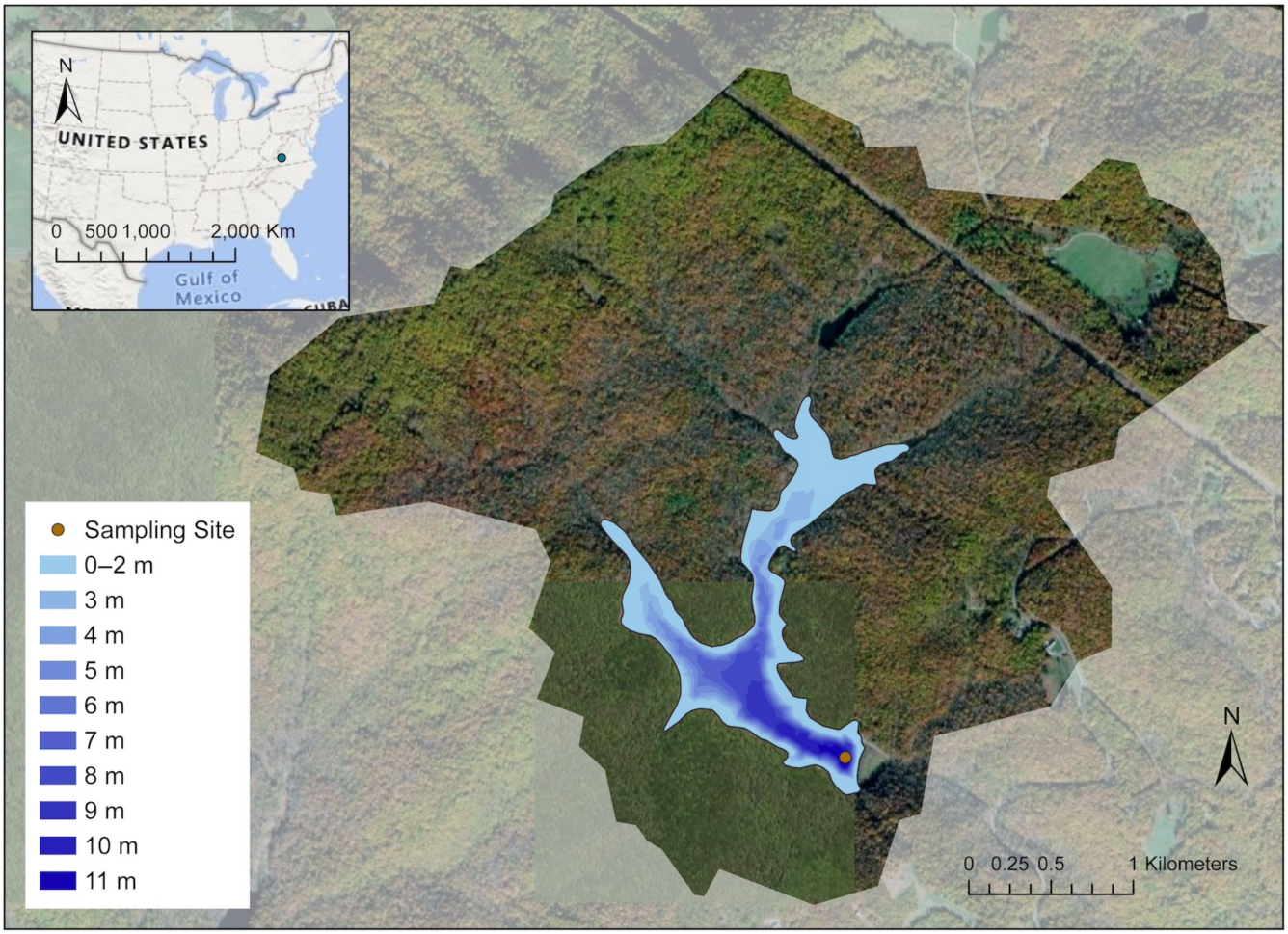

FIGURE 2 Map of Beaverdam Reservoir, Vinton, VA, USA (37.31° N, 79.82° W). The map shows the surrounding forested watershed; the point represents the reservoir monitoring site where high-frequency sensor data were collected.

simulating different water layers and has been applied to a variety of lakes worldwide for modeling (e.g., Bruce et al., 2018; Read et al., 2014) and forecasting hydrodynamics (e.g., Thomas et al., 2020; Thomas, McClure, et al., 2023).

We configured GLM for BVR using historical bathymetric data (Carey, Lewis, Howard, et al., 2022) and water temperature observations for initial conditions (Carey et al., 2023). We configured GLM with two sediment zones to simulate epilimnetic (surface) and hypolimnetic (bottom) sediment temperature dynamics following Carey, Hanson, et al. (2022). GLM requires meteorological and reservoir inflow observations as driver data to run the model. Because we were applying GLM for forecasting, meteorological forecasts, not observed meteorology, were used as driver data in the model, as described below. Additionally, we set the inflow to equal outflow in this study given limited inflow data for validation and the relatively short forecast horizons (≤35 days). We therefore kept water level constant throughout the study period in the model. We initiated the model using its default parameter set (Hipsey et al., 2019; Wander et al., 2023b) and performed calibration via a 35-day spin-up period with DA to tune parameters before the start of our focal forecasting period (described below).

FLARE configuration for DA and uncertainty

We configured FLARE for BVR following its application to other lakes and reservoirs (Thomas et al., 2020; Thomas, McClure, et al., 2023). We initialized the ensemble Kalman filter with 256 ensemble members, each with a vector of modeled water temperatures for different depths in the reservoir and parameter values (Thomas et al., 2020). We set the number of forecast ensemble members to 256 to ensure an adequate representation of uncertainty and prevent the ensemble Kalman filter from developing erroneous correlations among ensemble

members that can occur with low ensemble sizes (Duc et al., 2021; Machete & Smith, 2016).

We chose three parameters (a longwave radiation scaling parameter, and epilimnetic and hypolimnetic sediment temperature parameters) to assimilate following a global sensitivity analysis of all model parameters, as described in Thomas et al. (2020). These parameters were identified as important for water temperature simulations using GLM in a similar, nearby reservoir (Carey, Hanson, et al., 2022; Thomas et al., 2020): (1) the longwave radiation scaling factor (hereafter, *longwave*); (2) epilimnetic sediment temperature parameter (hereafter, *epi_sed_temp*); and (3) hypolimnetic sediment temperature parameter (hereafter, *hypo_sed_temp*).

The ensemble Kalman filter adds normally distributed random noise to both the predictions (vector of water temperature at each depth) and observations. Based on prior FLARE applications (Thomas et al., 2020), we set the SD of the predictions (process error) to 0.75°C and the SD of the observations to 0.1°C. The fact that the observation SD is lower than the modeled SD reflects greater confidence in the measurements than the predictions. As a result, model predictions should match observations when observations are available for assimilation. These perturbed predictions and observations were used in the Kalman gain to update the predictions using the difference between observed versus modeled water temperature, and parameters using the correlation with the updated water temperature predictions. In addition, the correlations between the parameter values and the model states with observations (i.e., water temperatures at the depths with sensor observations) were used to adjust parameters to be consistent with the most recent data used in DA. We used model default values provided in Hipsey et al. (2019) for the remaining GLM parameters, which were set to constant values that did not vary across depth and time.

We used state augmentation to tune the three parameters in the ensemble Kalman filter (Thomas et al., 2020). The three tuned parameters were initially calibrated during a spin-up period from 27 November to 31 December 2020 and were subsequently updated via DA throughout the forecasting period. To avoid the common issue of artificially low parameter uncertainty in sequential DA (Dietze, 2017a), we specified the SD of a normal distribution for each parameter (1.0°C for the sediment temperature parameters and 0.02 for the longwave radiation scaling factor). Initial exploration of parameter fitting in this study indicated that the application of FLARE over the full year resulted in low parameter uncertainty, necessitating us to specify the SD a priori rather than estimating it using DA. The distributions we chose were adapted from a prior application of FLARE that estimated the SD of parameter distributions across six small, dimictic lakes (Thomas, McClure, et al., 2023). Although BVR was not one of the lakes forecasted in Thomas, McClure, et al. (2023), BVR is similarly small and dimictic and, therefore, likely has similar parameter distributions using the same forecasting framework.

FLARE uses a numerical ensemble-based approach to simulate and propagate forecast uncertainty (Thomas et al., 2020). We represented the contribution of uncertainty from meteorological driver data, initial conditions, model process, and model parameters using the 256-member ensemble, following Thomas et al. (2020). First, to represent the contribution of meteorological driver data uncertainty, we assigned each of the 256 FLARE ensemble members one of the 30 ensemble members from the 1- to 35-day-ahead meteorological forecasts (NOAA's Global Ensemble Forecasting System) to drive GLM for forecasting. Second, we represented the contribution of uncertainty in the initial conditions of the forecasts using the spread in model states among the 256 ensemble members on the first day of each forecast. This spread was determined by either using the prior day's forecast as a starting point for the next day's forecast (when no data were available for DA) or the updated states following DA (when data were available for DA). We set the observation uncertainty SD to 0.1°C, determined from the SD of temperature observations and following prior applications of FLARE (Thomas et al., 2020). Third, we represented the contribution of model process uncertainty by adding random noise to the water temperature predictions from each of the 256 FLARE ensemble members at each daily time step in a 1- to 35-day-ahead forecast horizon. The random noise for each modeled depth within an ensemble member was drawn from a normal distribution with a SD of 0.75°C, as used in a previous application of FLARE that reported well-calibrated forecast uncertainty (Thomas et al., 2020). The random noise was spatially correlated so that it was most similar for nearby depths and most different for further-apart depths. The strength of the spatial correlation was determined by the exponential decay of the correlation strength with distance (Thomas, McClure, et al., 2023). Fourth, we represented parameter uncertainty using the SD of the distributions for the three tuned GLM parameters described above. A unique parameter value drawn from each of the three distributions was assigned to each of the 256 FLARE ensemble members. The parameter value assigned to an ensemble member was only updated when DA occurred. Parameters not tuned by the ensemble Kalman filter (i.e., those that were not identified from the global sensitivity analysis) had fixed values and uncertainty in these parameters was not calculated.

To determine whether there was a relationship between the magnitude of initial conditions uncertainty and the sensitivity of forecast skill to more frequent DA (following Clark et al., 2016; He et al., 2020; Simonin et al., 2017), we quantified the contribution of initial conditions uncertainty to total forecast uncertainty in our DA forecasts for all DA frequencies. For this analysis, we isolated the magnitude of initial conditions uncertainty by generating the water temperature forecasts for all 365 days with and without initial conditions uncertainty and compared the variance among all 256 ensemble members. We also calculated the proportion of initial conditions uncertainty within total forecast uncertainty for all depths, horizons, and stratified versus mixed periods.

DA experiments

To quantify the effect of DA at different frequencies on forecast skill, we conducted DA experiments in BVR from 1 January to 31 December 2021 ($n = 365$ days). As noted above, we used a spin-up period from 27 November to 31 December 2020 ($n = 35$ days) during which water temperature observations were used to update model parameters and initial conditions (i.e., DA occurred), but no forecasts were generated. During the one-year forecast period in 2021, we forecasted daily water temperature at 23 depths in the reservoir (spanning 0.1- to 11-m depth at 0.5-m intervals) and assessed forecast performance relative to observations across each forecast's daily predictions for 1- to 35-day-ahead horizons and depth intervals. We focused on three focal layers (surface, middle, and bottom) when reporting results to account for small changes in water level over the year. The surface layer encompassed 0–2.5 m and represented the epilimnion, the middle layer encompassed 2.6–8.4 m, which represented the metalimnion, and the bottom layer encompassed 8.5–11.0 m, which represented the hypolimnion.

We performed DA experiments using four different DA frequencies, daily, weekly, fortnightly, and monthly (see data assimilation experiments box in Figure 1 and Appendix S1: Figure S1 for visualization of DA frequencies), to represent different data collection latencies that are commonly used by water quality monitoring programs (e.g., Engelhardt & Kirillin, 2014; Francy et al., 2015; Kirchner & Neal, 2013; Liu et al., 2019; Romero et al., 2002). We assimilated water temperature data across different temporal frequencies by downsampling from the high-frequency observations collected by our sensors. For example, for the weekly DA frequency, observations were selected every seven days starting on 4 January 2021 and ending on 31 December 2021. In this example, DA only occurred once per week; the forecasts that were generated on the six other days in the same week did not include DA (i.e., no DA occurred during 5 January–10 January 2021 even though forecasts were still generated daily during this interval; Figure 1). Fortnightly and monthly DA occurred every 14 and 30 days, respectively, throughout the year.

We generated 365 daily forecasts starting on 1 January 2021 for each of the four DA frequencies. While we recognize that we are producing hindcasts for a historical period, because the model was forced with only forecasted drivers and out-of-sample forecast evaluation occurred, we refer to these retrospective forecasts or hindcasts as forecasts throughout for consistency (following Jolliffe & Stephenson, 2012).

Analysis

Question 1

For all $n = 1460$ forecasts produced (365 forecasts generated daily over a year for four different DA frequencies), we used root mean square error (RMSE) and continuous ranked probability score (CRPS; Gneiting et al., 2005) to quantify forecast skill. We defined skillful water temperature forecasts as those with an RMSE < 2°C, a commonly used threshold for lake and reservoir hydrodynamic modeling following Bruce et al. (2018), Read et al. (2014), and many others. We note that there are many ways to quantify skill beyond the 2°C RMSE threshold used here, such as the correlation coefficient, Nash–Sutcliffe model efficiency coefficient, percent relative error, normalized mean absolute error, or others (Bennett et al., 2013). We focused on RMSE for results reporting because it is a commonly used metric by lake modelers to determine the deviation between observed versus modeled values (Bruce et al., 2018; Read et al., 2014), and all CRPS results are reported in Appendix S1. Mean full water column RMSE was calculated for each of the 35 days across all forecast horizons for each DA frequency regardless of whether data were assimilated the day the forecast was generated. We calculated RMSE by taking the square root of the mean difference between predicted and observed water temperature values for each depth and day. To determine how parameter updating affected overall forecast skill, we also ran FLARE with constant, untuned parameters throughout the forecasting period, rather than updating them with DA. In the constant parameter forecasts, initial conditions were still updated at the daily time step to isolate the effect of parameter tuning on forecast skill.

Question 2

Using RMSE and CRPS, we compared forecast skill across depths and seasons to identify how the frequency of DA affected forecast accuracy over space and time. To quantify spatial forecast performance, we calculated RMSE and CRPS for each depth (1–11 m) at each forecast horizon (1–35 days ahead) and DA frequency in BVR. To quantify temporal forecast performance, we compared forecast skill at each horizon aggregated within thermally stratified versus mixed periods in BVR. The stratified period began on the first day that the water density difference between the reservoir surface (0.1 m) and the maximum depth observed for the reservoir on each day (e.g., between 9 and 11 m) was ≥ 0.1 kg/m^3 for at least three consecutive days (following Ladwig et al., 2021). Conversely, the mixed period began on the first day when surface and bottom-water density differences were <0.1 kg/m^3 for at least three consecutive days. Altogether, we compared forecast skill between stratified versus mixed periods; among layers (surface, middle, and bottom); and among forecast horizons (focusing in on 1, 7, and 35-day-ahead forecasts) for each of the four DA frequencies.

Question 3

We quantified total forecast uncertainty for each day in the 1- to 35-day forecast horizon using the variance of the 256-member FLARE ensemble. The relative contribution of initial condition uncertainty to total forecast uncertainty was calculated for each forecasted day by comparing the variance in the 256-member FLARE ensemble between the set of forecasts with initial condition uncertainty included and the set without initial condition uncertainty.

All statistical analyses were conducted in R v.4.2.0 (R Core Team, 2022). All R code and data files used to run these analyses are archived and available in the Zenodo repository (Wander et al., 2023a, 2023b).

RESULTS

BVR water temperature dynamics

BVR exhibited typical annual water temperature dynamics during the forecasting period in 2021. Water temperature throughout the water column ranged from 1.4 to 29.9°C during the year. The summer-stratified period began on 12 March and ended on 7 November 2021, and the reservoir was mixed from 1 January to 11 March and 8 November to 31 December (Figure 3). Thermocline deepening occurred throughout the summer-stratified period, starting at 1.5 m in March with stratification onset and deepening to 9.5 m in November before fall turnover (Figure 3). During the winter, there were three brief periods of ice cover of one to three days in duration in January and February when inverse stratification occurred (Figure 3; Carey & Breef-Pilz, 2022). We removed these few ice-cover days from the analysis and grouped mixed ($n = 118$ days) versus summer-stratified data ($n = 241$ days) for analysis.

DA frequency altered forecast output and parameters over time

We were able to successfully forecast water temperature throughout the water column over the year using DA to update model states and parameters (Figures 4 and 5). Across all depths, DA constrained uncertainty by updating initial conditions with the most recent water temperature observations. Forecast uncertainty for the lower DA frequencies was strongly dependent on the time since last assimilation (Figure 4). Mean forecast variance at the one-day horizon across 2021 for forecasts with daily DA was 0.60°C, while mean forecast variance at the one-day horizon for forecasts with monthly DA was 2.38°C.

We observed that DA frequency altered the parameter evolution of the forecasts (Figure 5). The daily DA frequency resulted in more variable parameter estimates through time for all three tuned parameters, reflecting the more frequent adjustment that occurred each time data were assimilated. Importantly, parameter evolution for forecasts with daily DA yielded very different mean estimates than the weekly, fortnightly, and monthly DA forecast frequencies (Figure 5). For example, the evolution of the longwave radiation scaling parameter (longwave) over the 365-day forecast period showed that forecasts with weekly, fortnightly, and monthly DA frequencies converged at a mean of ~0.91 by December 2021, whereas the mean longwave parameter for forecasts with daily DA was at ~0.85 by the end of the year (Figure 5a). Similarly, the parameter controlling the surface layer sediment temperature (epi_sed_temp) in daily DA forecasts began to diverge from the other DA frequencies in April (Figure 5c). The nondaily DA frequencies (i.e., weekly, fortnightly, monthly DA) surface sediment layer temperature parameter (epi_sed_temp) values ranged from 13.29 to 17.0°C, whereas the daily DA frequency epi_sed_temp values ranged from 13.73 to 20.0°C during April–December. For the parameter controlling the bottom-layer sediment temperature

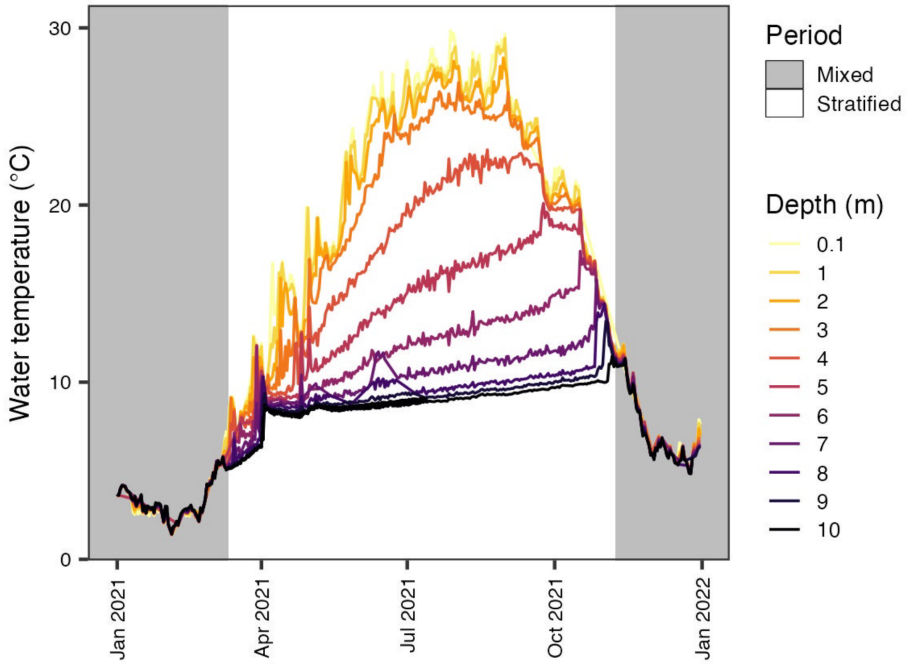

FIGURE 3 Observed water temperature for all depths with high-frequency sensors during the forecasting period of 1 January–31 December 2021 in Beaverdam Reservoir. The gray background indicates the mixed period (1 January–11 March, 8 November–31 December 2021), while the white background indicates the thermally stratified period (12 March–7 November 2021), defined by a >0.1 kg/m^3 density differential between surface and bottom layers.

(hypo_sed_temp), daily DA forecasts exhibited much more variable values (ranging from 8.20 to 10.98°C) than forecasts for any other DA frequency (range 10.24–11.21°C; Figure 5b) from April to December.

Question 1: Which frequency of DA generates the most skillful water temperature forecasts?

Aggregated among depths and time periods, weekly DA resulted in the most skillful water temperature forecasts of the four DA frequencies for the greatest number of 1- to 35-day-ahead horizons (Figure 6). Among horizons, we observed that the frequency of DA needed to produce skillful forecasts varied (Figure 6). At shorter horizons (1–7 days ahead), daily DA resulted in the most skilled forecasts, but at longer horizons (8–35 days ahead), weekly DA resulted in the most skilled forecasts (Figure 6).

The skill of all forecasts degraded as the forecast horizon increased, but the decrease in performance was greatest for daily DA forecasts, such that forecasts generated using monthly, fortnightly, and weekly DA all outperformed daily DA forecasts by the 19-day forecast horizon (Figure 6), when aggregating across all depths and time periods. The daily DA forecasts exceeded the 2°C RMSE metric of skill on the 28-day-ahead horizon, whereas the weekly, fortnightly, and monthly forecasts never exceeded that metric for any of the 1- to 35-day-ahead horizons. These results were consistent across forecast evaluation metrics, including the CRPS metric that evaluates the full ensemble forecast (Appendix S1: Figure S2).

Question 2: How does forecast skill vary across time and space?

Aggregated across depths, horizons, and DA frequencies over the year, forecast skill overall was high (i.e., RMSE < 2°C), with a mean water temperature forecast RMSE of 1.50 ± 1.86°C (1 SD). Forecast skill was generally best at bottom depths regardless of horizon or DA frequency. Aggregated bottom-water forecast skill was 1.13 ± 1.76°C, followed by aggregated mid-water column forecast skill (1.58 ± 1.94°C), and aggregated surface forecast skill (1.78 ± 1.88°C). As expected, forecast skill generally decreased with horizon, with a mean 1-day-ahead forecast RMSE of 0.81 ± 1.20°C, mean 7-day RMSE of 1.15 ± 1.59°C, and mean 35-day RMSE of 1.94 ± 2.19°C.

On average, forecast skill was slightly better (as indicated by smaller RMSE) during the stratified period than during the mixed period, aggregated among all depths and horizon regardless of DA frequency

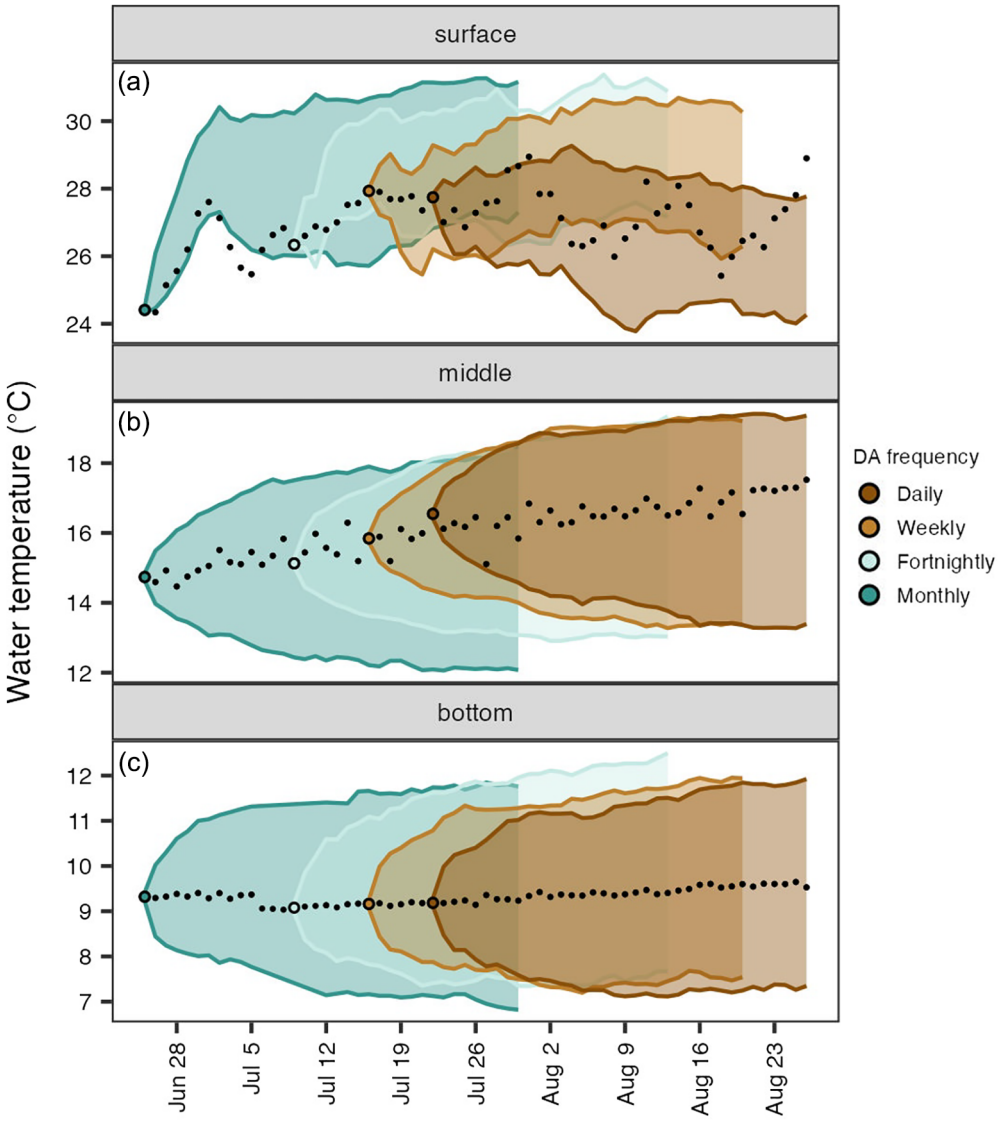

FIGURE 4 Example of water temperature forecasts at the surface (a), middle (b), and bottom depths (c) generated for 1–35 days into the future in Beaverdam Reservoir. Data assimilation (DA) frequencies are depicted by colors; shading shows 95% CIs around the mean predicted temperature for each day. Black points represent water temperature observations. Colored points represent the most recent day that data were assimilated for each DA frequency. In this example, data were most recently assimilated on the day that the forecasts were generated: 25 June for the monthly DA scenario, 9 July for the fortnightly DA scenario, 16 July for the weekly DA scenario, and 22 July for the daily DA scenario. For this figure, the surface was represented by 1 m; middle was represented by 5 m; and bottom was represented by 9-m individual forecasts.

(aggregated mixed RMSE = 1.56 ± 1.58°C, stratified RMSE = 1.43 ± 2.14°C; Figure 7). Forecast skill was more variable among forecast horizons than depths in the mixed period, whereas forecast skill was variable across both depths and horizons in the stratified period (Figure 7). In the stratified period, forecast skill was best at bottom depths, with relatively similar skill across all forecast horizons (Figure 7f). In the mixed period, forecast skill varied very little among depths aggregated across horizons (Figure 7a,c,e), with consistently greater decreases in skill with increasing horizon than in the stratified period.

While daily DA always resulted in the best forecast skill for 1-day-ahead horizons, lower frequency DA typically outperformed daily DA as the forecast horizon increased. The horizon at which weekly DA forecasts outperformed daily DA forecasts varied with depth, where weekly DA forecasts outperformed daily DA forecasts at the surface at horizons >5 days ahead, on average. In contrast, weekly DA for mid- and bottom-water depths resulted in more skillful forecasts on average than daily DA only at horizons >8 days ahead (Figure 7). Across all depths and stratified/mixed periods, monthly DA forecasts outperformed daily DA forecasts by 12- to 31-day-ahead horizons (Figure 7). Bottom-water stratified forecasts were the only forecasts for which skillful (RMSE < 2°C) forecasts were produced for all DA frequencies and horizons,

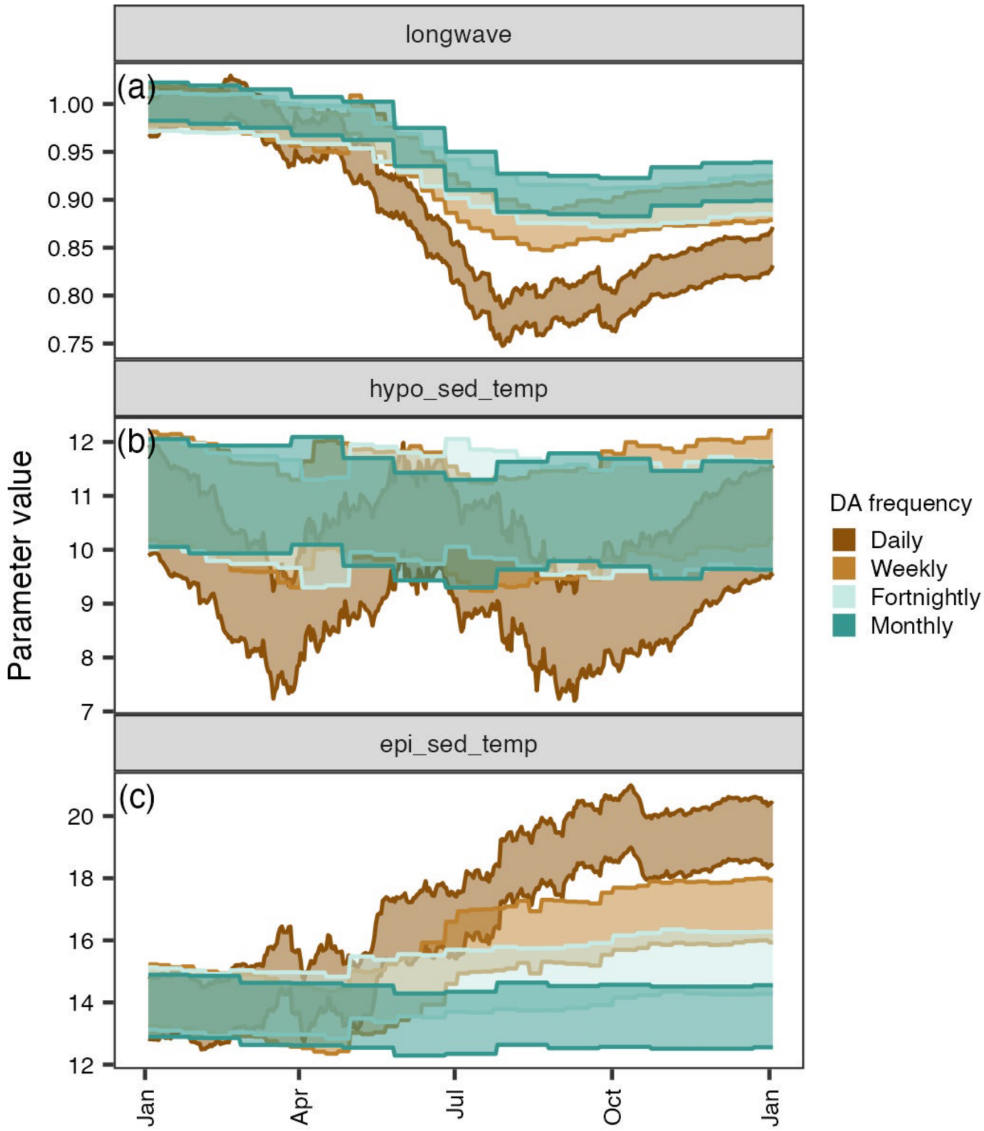

FIGURE 5 Parameter evolution during the forecast period (1 January–31 December 2021) for daily, weekly, fortnightly, and monthly data assimilation (DA) frequencies at 1-day-ahead forecast horizons. Longwave (a) is the longwave radiation scaling parameter, hypo_sed_temp (b) is the hypolimnetic sediment temperature parameter, and epi_sed_temp (c) is the epilimnetic sediment temperature parameter.

as skill never exceeded 1.01°C RMSE for the duration of the 35-day forecast horizon (Figure 7f).

Question 3. How does DA frequency influence total forecast uncertainty and what is the relative contribution of initial condition uncertainty to total forecast uncertainty?

Lower frequency DA forecasts consistently had more total uncertainty (Figure 8). We found that the differences between uncertainty for daily and monthly DA were largest at 1-day-ahead horizons and largely converged by the end of the 35-day horizon (Figure 8). For surface depths, total uncertainty was similar between the mixed and stratified periods across the 35-day horizon, but total uncertainty was on average higher in the stratified than mixed period for mid- and bottom-water depths. Both RMSE and total variance were similar for forecasts run with and without initial conditions uncertainty included (Appendix S1: Figures S3 and S4).

Forecasts with less frequent DA had a greater contribution of initial condition uncertainty to total forecast uncertainty during the first few days of the forecast horizon. However, overall, initial conditions uncertainty contributed a minimal proportion of the total uncertainty for forecasts generated with daily DA (Figure 9). At the 1-day-ahead forecast horizon, daily DA initial conditions uncertainty contributed 0.01% of total uncertainty, whereas initial

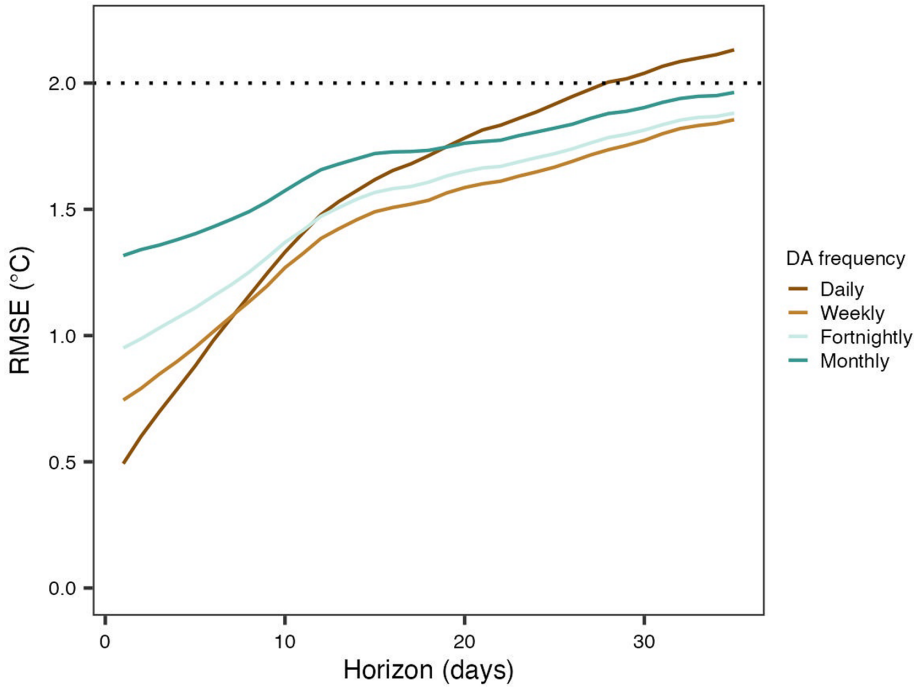

FIGURE 6 Root mean square error (RMSE) of mean forecasted water temperature compared with observations for 1- to 35-day-ahead forecast horizons in Beaverdam Reservoir, aggregated for all depths in the water column and days within the 365-day forecast period. RMSE for each forecast horizon was averaged from forecasts generated during 1 January–31 December 2021. Colored lines represent different data assimilation (DA) frequencies. The dotted line depicts the 2°C threshold for skillful water temperature forecasts.

conditions uncertainty contributed a mean of 54%–71% of total forecast uncertainty in forecasts across all other DA frequencies (Figure 9). The role of initial conditions uncertainty was minimal (<1%) across all DA frequencies after the 10-day horizon throughout the entire water column in the mixed period and surface depths only in the stratified period (Figure 9a–c,e). Conversely, initial conditions uncertainty made up a larger proportion of total forecast uncertainty for stratified mid- and bottom-water forecasts for forecast horizons between 10 and 20 days (~11%; Figure 9d,f).

Finally, we found that forecasts generated with daily parameter tuning had substantially greater skill than forecasts generated with constant parameters (but updated initial conditions) throughout the forecast period. Across all horizons, depths, and periods, forecasts generated with daily DA had a mean RMSE of 1.95 ± 1.81°C, versus a mean RMSE of 3.12 ± 2.03°C for forecasts generated with constant parameters (Appendix S1: Figure S5).

DISCUSSION

Across a year of water temperature forecasts in our focal reservoir, we found that weekly DA generally resulted in the most skillful water temperature forecasts. However, skill varied among depths, forecast horizons, and time of year, suggesting that DA frequency should be chosen based on the specific forecast application. For example, if water temperature forecasts are specifically needed to guide decision-making that involves short time horizons (e.g., <8 days ahead), daily DA might be most advantageous (Figures 6 and 7). Conversely, if water temperature forecasts are needed for the surface or mid-water column depths at 20- to 35-day-ahead horizons, then weekly to monthly DA may be sufficient (Figure 7). Despite the usefulness of DA for improving forecast skill, more frequent DA did not always lead to more skillful water temperature forecasts, in part because initial conditions uncertainty only comprised a significant proportion of total forecast uncertainty within the first few days of the forecast horizon (Figure 9). Below, we interpret our results for each research question and make recommendations for considering which DA frequency might be appropriate for different ecological forecast applications.

Question 1: Which frequency of DA generates the most skillful water temperature forecasts?

In this study, we found that less frequent DA (e.g., weekly, fortnightly, and monthly DA) sometimes led to more skillful water temperature forecasts than

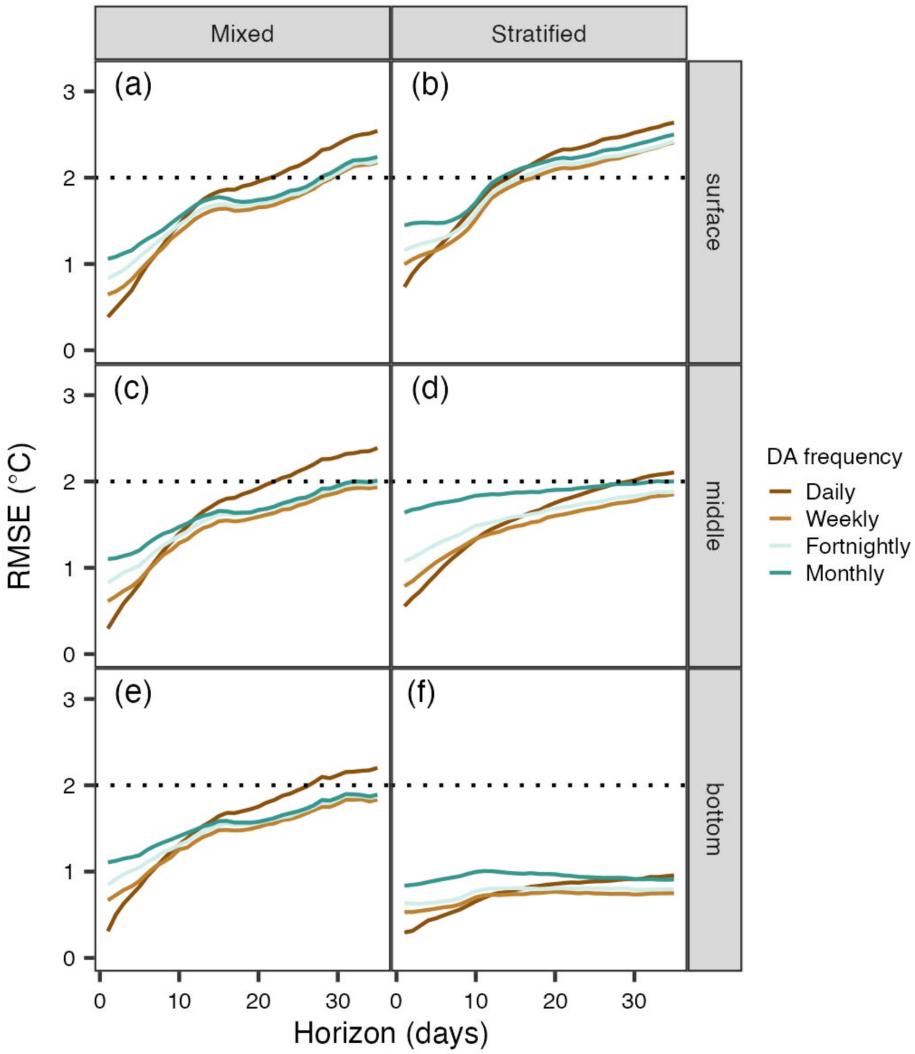

FIGURE 7 Root mean square error (RMSE) of mean forecasted water temperature compared with observations for 1- to 35-day-ahead forecast horizons in Beaverdam Reservoir during the mixed (a, c, e) versus stratified (b, d, f) periods for aggregated surface (0–2.5 m; a, b), middle (2.6–8.4 m; c, d), and bottom (8.5–11.0 m; e, f) depths. RMSE for each forecast horizon was averaged across the 365-day forecast period (1 January–31 December 2021). Colored lines correspond to different data assimilation (DA) frequencies; dotted horizontal lines depict the 2°C threshold for skillful forecasts.

daily DA for all depths during the mixed period. This pattern of weekly DA outperforming daily DA forecast skill during the mixed period is likely because daily DA led to parameter overfitting, as indicated by the greater short-term variability in parameter estimates over time (Figure 5). Because water temperatures are fairly stable at deeper depths, and thus daily observations can consistently predict tomorrow's water temperature accurately, parameter overfitting was less problematic for daily DA at hypolimnetic depths (Figure 7f). As a result, hypolimnetic forecast skill was best with daily DA during stratified conditions, but this pattern did not extend to other depths or the mixed period (Figure 7).

At the surface, monthly DA produced forecasts that did not capture all observations within the SD of the forecast mean (Figure 4). This was likely because observations were assimilated only once a month and therefore did not provide enough information to accurately forecast water temperature patterns >7 days into the future. Additionally, given the greater sensitivity of the surface waters to meteorological forcing, changes in day-to-day water temperature were greater in magnitude at the surface compared with the mid-water column and bottom layers, necessitating more frequent DA to accurately capture these temperature dynamics. In comparison, daily DA forecasts generally captured observations skillfully at the surface until the last eight days of the 35-day forecast horizon (Figures 4 and 6). This decline in forecast skill using daily DA was likely a result of overfitting of the longwave scaling parameter and

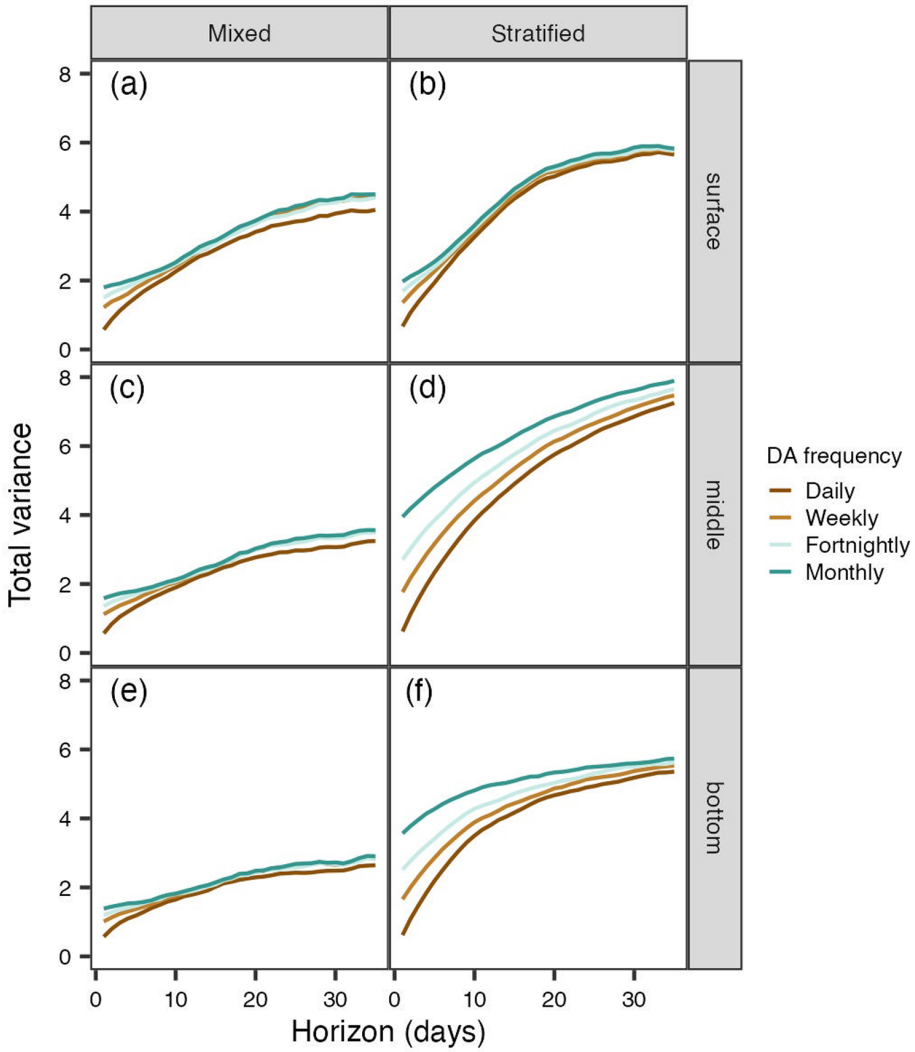

FIGURE 8 Mean water temperature forecast variance across horizons (1–35 days ahead) in Beaverdam Reservoir during the mixed (a, c, e) versus stratified (b, d, f) periods for aggregated surface (0–2.5 m; a, b), middle (2.6–8.4 m; c, d), and bottom (8.5–11.0 m; e, f) depths. Variance for each forecast horizon was averaged from all 365 forecasts generated during the forecast period (1 January–31 December 2021). Colored lines correspond to different data assimilation (DA) frequencies.

sediment temperature parameters (Figure 5). Because the values of these parameters varied following daily DA compared with those obtained using weekly, fortnightly, and monthly DA (Figure 5), parameter overfitting may have led to observations falling outside the predicted range of uncertainty around the mean forecasted water temperature at longer horizons (Figure 4). Finally, the range of uncertainty for surface forecasts within the first seven days of the forecast horizon was substantially smaller than that observed for mid-water column and bottom layers (Figure 4), further supporting that DA was more useful for <7-day-ahead surface water temperature forecasts.

Our work is consistent with studies that have found that the optimal DA frequency often matches that of the forecast model time step (e.g., Derot et al., 2020; Woelmer et al., 2022). For example, during both the mixed and stratified periods, daily DA was always better for 1-day-ahead forecasts, but was often outperformed by weekly DA at 8-day-ahead forecast horizons (Figures 6 and 7). Because water temperatures were homogenous among all depths during the mixed period, water temperature variability among all depths was likely driven by air temperature variability, ultimately making it more challenging to predict water temperature across depths as the forecast horizon increased. During the stratified period, however, less frequent DA could still generate accurate surface and mid-depth water temperature forecasts (Figure 7). This pattern is in contrast with other water temperature forecasting studies that have found daily DA necessary for improving the skill of forecasts in the middle of the water column around the thermocline (Baracchini, Chu,

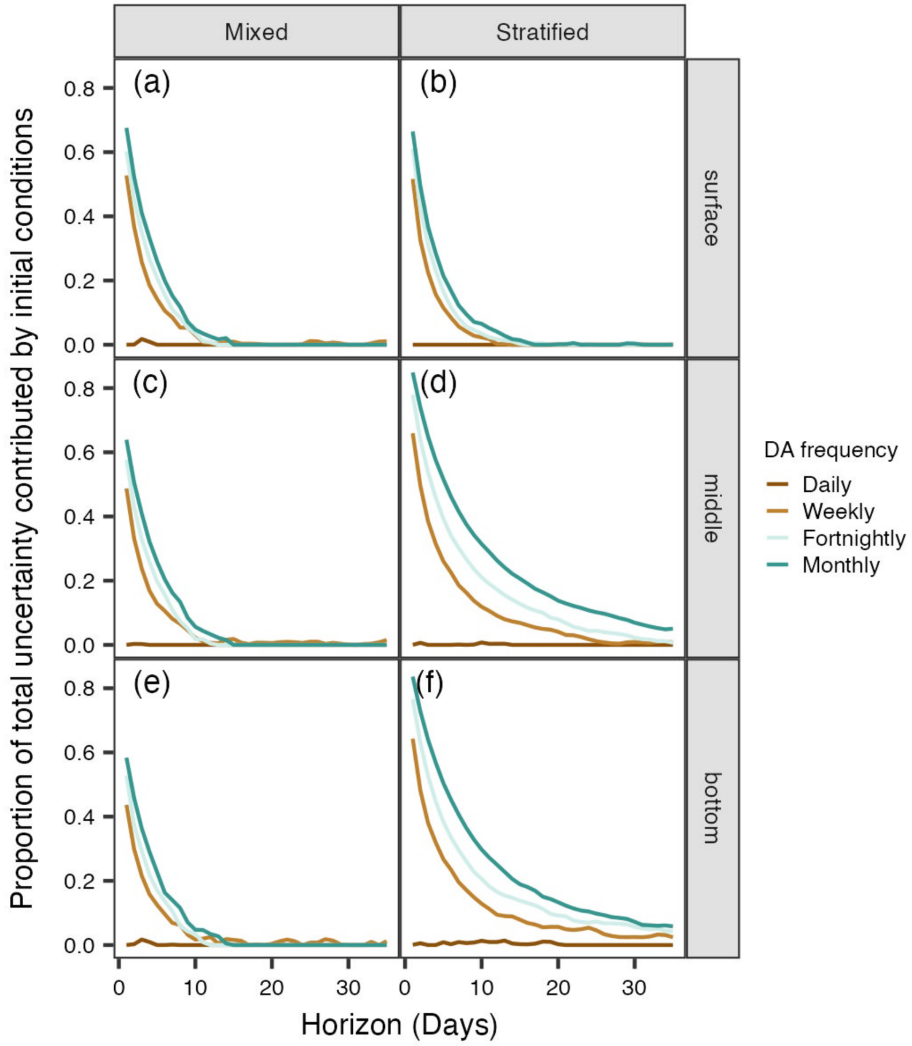

FIGURE 9 Proportion of total forecast uncertainty that is contributed by initial conditions uncertainty, averaged across all forecasts generated with each data assimilation (DA) frequency across 1 January–31 December 2021. Colored lines depict DA frequencies. Panels a, c, and e represent mixed period forecasts, and panels b, d, and f represent stratified period forecasts. Aggregated surface (0–2.5 m), middle (2.6–8.4 m), and bottom (8.5–11.0 m) depths are indicated by gray facet labels to the right of each panel.

et al., 2020), but is likely explained by the overfitting of both the daily longwave radiation and the epilimnetic sediment temperature parameters (Figure 5). The decreased importance of daily DA at bottom depths during the stratified period is likely because of the fairly consistent temperatures exhibited at bottom depths associated with thermal stratification (Figures 3 and 7).

We note two considerations for future work. First, many high-frequency sensors provide observations at subdaily timescales, providing the opportunity to test the effects of more frequent DA on forecast skill than the timescales tested in this study (e.g., Piazzi et al., 2018; Ziliani et al., 2019). Surface water temperature dynamics, in particular, can exhibit cyclical patterns within a day that may be better simulated with subdaily DA (Hollan & Simons, 1978; Tasnim et al., 2021). However, FLARE is configured to generate forecasts and assimilate data at a daily time step because the underlying process model is ideally designed to run simulations for one day or longer (Hipsey et al., 2019), so more frequent DA would require a different forecasting framework.

Second, we recognize that RMSE is a deterministic metric that uses the ensemble mean to evaluate forecast skill. Alternative forecast skill metrics that incorporate uncertainty across all ensemble members include CRPS and the ignorance score (Simonis et al., 2021; Smith et al., 2015). These metrics are useful for probabilistic forecast assessment because forecast skill is determined using all ensemble members rather than just the ensemble mean, which may provide useful information for management (McSharry et al., 2005). However, given that the CRPS results followed similar patterns as RMSE

(Appendix S1: Figure S2), evaluating the full distribution of forecasts does not change the overall interpretation of forecast skill across depths and horizons.

Question 2: How does forecast skill vary across time and space?

Overall, we observed generally high forecast skill across all depths and times of year for most forecast horizons. Across DA frequencies, depths, and times of year, RMSE was only consistently above the 2°C skill threshold for daily DA at 28- to 35-day horizons (Figure 6). By the end of the 35-day forecast horizon, daily DA forecast skill for most depths and times of the year was >2°C, except for 9-m stratified forecasts, which had a mean RMSE of 1.29 ± 1.8°C across DA frequencies (Figure 7). As noted above, the higher forecast skill at bottom-water depths is likely because fluctuations in bottom-water temperatures were minimal during stratification (Figure 3).

Our findings are similar to other water temperature lake and reservoir forecasting studies that used one-dimensional hydrodynamic models to simulate thermal dynamics. First, the pattern of increased forecast skill in the bottom waters is consistent with Mercado-Bettín et al. (2021) and Thomas et al. (2020), who both found that the bottom-water forecasts were more skillful than surface water forecasts. This is likely because bottom waters are not changing as much as surface waters throughout the year due to less atmospheric exchange. However, Clayer et al. (2023) found that surface water temperatures were more accurately simulated than bottom-water temperatures, suggesting that the complex lake characteristics that control bottom-water temperatures were not captured as well as the air temperature dynamics controlling surface water temperatures. Second, our finding that forecast skill was greater in the stratified period rather than mixed period is similar to the results of Thomas et al. (2020), likely due to the fact that water temperature dynamics were changing less day-to-day in stratified than mixed periods (Figure 3). Because of the variability in water temperature dynamics among seasons and depths, determining the conditions in which we can most accurately forecast water temperature can improve our understanding of ecosystem processes and functioning. Moreover, accurately forecasting water temperature is critical for forecasting additional lake and reservoir variables that are strongly driven by water temperature, such as phytoplankton biomass, dissolved oxygen concentrations, and greenhouse gas emissions (e.g., McClure et al., 2021).

Question 3: How does DA frequency influence forecast uncertainty?

We found that initial conditions uncertainty contributed a substantial proportion of total uncertainty for weekly, fortnightly, and monthly DA, but only during the first few days of the forecast horizon. After 14–16 days into the future, the contribution of initial conditions uncertainty to total uncertainty decreased to <1% for all depths during the mixed period and surface depths in the stratified period across all DA frequencies (Figure 9). The contribution of initial conditions uncertainty to total uncertainty for mid-water column and bottom-layer forecasts using weekly, fortnightly, and monthly DA exhibited a more gradual decline with increasing forecast horizon than forecasts for the surface layer in the stratified period (Figure 9). This result suggests that there may be a greater benefit of using more frequent observations for DA for forecasts of the mid-water column and deeper depths than the surface during the stratified period. However, more frequent DA may not always improve forecast performance, especially when initial conditions uncertainty is not the dominant source of uncertainty, as seen at longer horizons. Given that the contribution of initial conditions uncertainty to total uncertainty is very low at longer forecast horizons, one or more of the remaining quantified sources of uncertainty (e.g., process, parameter, and/or driver data) must dominate total forecast uncertainty at these horizons (Figure 9). Conversely, the dominant source of uncertainty for weather forecasting is typically initial conditions uncertainty given the inherent instability of atmospheric processes (Dietze, 2017b), which is why more frequent DA often substantially improves meteorological forecast skill.

Other lake and reservoir water quality forecasting studies have found that model driver data and process uncertainty were the dominant sources of total forecast uncertainty (Lofton et al., 2022; McClure et al., 2021; Thomas et al., 2020). Therefore, constraining other sources of uncertainty using an ensemble approach or different forecasting models would likely further improve water temperature forecast skill. Additionally, using a different DA technique that uses a Bayesian approach to estimate a posterior distribution, rather than assuming that the parameters and model states are normally distributed, may also reduce uncertainty (e.g., particle filter; Wang et al., 2023). Because the dominant source of uncertainty in ecological forecasts will likely differ depending on the variable being forecasted, different DA techniques may not improve forecast skill equally among all ecological variables.

Recommendations for setting up DA for other forecasting systems

Determining whether an ecological forecasting application requires high-frequency sensors is necessary for increasing the scalability of ecological forecasting across ecosystems and variables. While high-frequency sensor data may improve forecast skill in some cases, sensor deployment is often costly, which limits the application of high-frequency data in some forecasting systems. Moreover, even if high-frequency sensors are deployed, identifying the minimum frequency of data required to make skillful ecological forecasts can be a useful exercise because high-frequency sensors malfunction and require maintenance, which can result in data gaps (e.g., Herrick et al., 2023). Many water quality forecasting applications to date have relied on high-frequency sensor data for assimilation to produce skillful forecasts of different variables (e.g., Cho & Park, 2019; Derot et al., 2020; Page et al., 2018). We note that despite its limitations, daily parameter tuning through DA did improve forecast skill by 1.17 ± 0.22°C compared with forecasts generated using constant parameters (Appendix S1: Figure S5).

In this study, we found that daily and weekly DA exhibited a trade-off in terms of which DA frequency produced the most skillful water temperature forecasts (Figure 6), where the horizon at which daily DA was outperformed by weekly DA differed among depths and times of year (Figure 7). Our findings indicate that high-frequency sensors may not be needed for accurate water temperature forecasts for the middle and bottom layers, as weekly DA forecasts remained skillful (<2°C RMSE) through the full 1- to 35-day-ahead horizon at these depths and periods (Figure 7).

The minimum frequency of DA needed to set up fully operational forecasting systems is likely to vary based on the ecosystem and forecast variable of interest. Depending on the water quality forecast application, different frequencies of data collection may be necessary to fully understand and predict water quality dynamics over time. For example, we do not know how more frequent DA (e.g., hourly DA) may affect forecast skill, given that some waterbodies exhibit fluctuations in surface water temperature over 24 h (e.g., Hollan & Simons, 1978; Serruya, 1975). Alternatively, George and Hurley (2004) found that fortnightly observations were required to discern gradual trends in phytoplankton productivity, but monthly data were adequate for capturing declines in phytoplankton biomass over a 30-year period. Despite many successful applications of high-frequency DA in the literature for forecasting (e.g., Cho et al., 2020; Gottwald & Reich, 2021; Luo et al., 2011; Niu et al., 2014), not all ecological variables benefit from frequent DA, as not all variables are similarly forecastable. For example, Massoud et al. (2018) found that dynamic application of DA to plankton forecasts improved forecasting system efficiency only when observations were highly variable (i.e., when large fluctuations in abundance were observed), but did not consistently improve forecasts when DA was applied at every model time step.

In addition to the frequency of data collection, data latency can also affect the frequency of DA. High-frequency data are not always available, particularly in remote sensing applications when satellite orbit schedules and cloud coverage limit access to high-frequency observations (Herrick et al., 2023). Even for forecasting systems with high-frequency sensor data, data latency may reduce forecast skill if data are not immediately transmitted to forecasting workflows (e.g., they require a manual download; Dietze et al., 2018). In cases with high data latency of the forecast variable (e.g., microscope counts of phytoplankton requiring laboratory analysis), data fusion approaches that assimilate multiple data sources may improve forecast skill (e.g., Baracchini, Wüest, et al., 2020; Chen et al., 2021). For example, some studies have assimilated both in situ measurements and remote sensing data to forecast reservoir water quality variables, including chlorophyll *a* and conductivity (Abdul Wahid & Arunbabu, 2022; Chen et al., 2021).

Finally, understanding the contributions of different sources of uncertainty can be useful for determining the DA frequency that generates the most skillful forecasts. Specifically, knowing the relative contribution of initial conditions uncertainty can inform the sampling frequency needed to improve ecological forecast skill. For forecasts with total uncertainty dominated by process, parameter, or driver uncertainty, improving forecast skill may require modifying processes used for forecasting the ecological variable of interest, further constraining parameters by collecting more data, or improving weather forecast driver data (e.g., Grönquist et al., 2021).

Study limitations

Our results suggest that weekly DA may suffice for some lake and reservoir water temperature forecasting applications, with the caveat that more frequent DA often improved water temperature forecast performance at short forecast horizons. However, we only assessed forecast skill for a single reservoir and ecological variable for only one year, and therefore note the limitations of extending these results to other systems and variables. For example, parameter overfitting due to daily updating

of model parameters reduced overall forecast skill, which helps explain why weekly rather than daily DA resulted in more skillful water temperature forecasts in the mixed period (see Lin et al., 2021), but it is unknown how generalizable this parameter overfitting may be at other sites. Finally, because we did not partition the contribution of the other sources of uncertainty that were summed to quantify total forecast uncertainty, we can only identify the relative role that DA has on reducing initial conditions uncertainty.

Conclusions

This study emphasizes the importance of DA for improving ecological forecast skill and has implications for forecasting efforts among a wide range of ecosystems and ecological variables. We argue that weekly observations of water temperature are likely "good enough" to set up a skillful forecasting system for many reservoir management applications, while daily DA would be most useful for applications requiring high forecast accuracy in the bottom waters or at short (<5–7 days) forecast horizons. Because water temperature dynamics control many biological, chemical, and physical lake processes (Magnuson et al., 1979; Read et al., 2019; Yvon-Durocher et al., 2012), water temperature must be accurately forecasted before we can forecast other water quality variables. Therefore, determining ways to improve water temperature forecasts will have broad utility for advancing the development of many additional water quality forecasting systems.

Because near-term, iterative forecasts are particularly well suited to address ecological questions (Carey, Woelmer, et al., 2022; Dietze et al., 2018; White et al., 2019), determining how best to design and deploy ecological near-term, iterative forecasting systems is a pressing need (Diez et al., 2012; Ibáñez et al., 2013; Moustahfid et al., 2021). With the increasing deployment of high-frequency sensor networks (e.g., NEON and Global Lake Ecological Observatory Network (GLEON); Mantovani et al., 2020; Marcé et al., 2016; Park et al., 2020) comes a growing need to understand how best to use these sensor data for forecasting. In response, we advocate for using DA experiments across ecosystems and ecological variables to determine how best to integrate observational data into iterative forecasting systems.

AUTHOR CONTRIBUTIONS

Heather L. Wander, R. Quinn Thomas, and Cayelan C. Carey co-developed the design of this study. R. Quinn Thomas developed the FLARE framework for data assimilation experiments and forecasting workflow used in this study. Tadhg N. Moore helped develop early iterations of the forecasting workflow. Adrienne Breef-Pilz oversaw sensor data collection. Heather L. Wander wrote the initial draft of the manuscript with Cayelan C. Carey and Mary E. Lofton. All coauthors reviewed the manuscript and approved its final version.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation (DEB-1753639, DBI-1933016, DBI-1933102, DEB-1926050, DEB-1926388, and OAC-2004323) and greatly enabled by the Virginia Reservoirs LTREB monitoring program (DEB-2327030). We thank the Western Virginia Water Authority for their long-term access to Beaverdam Reservoir and the CIBR-FLARE team, Carey Lab, and Reservoir Group for helpful feedback throughout this project.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data (Carey et al., 2023; Wander et al., 2023a) and code (Wander et al., 2023b) are available from the Environmental Data Initiative and Zenodo repositories: https://doi.org/10.6073/pasta/4182de376fde52e15d493fdd9f26d0c7; https://doi.org/10.5281/zenodo.7925097; https://zenodo.org/records/7958470.

ORCID

Heather L. Wander https://orcid.org/0000-0002-3762-6045
R. Quinn Thomas https://orcid.org/0000-0003-1282-7825
Tadhg N. Moore https://orcid.org/0000-0002-3834-8868
Mary E. Lofton https://orcid.org/0000-0003-3270-1330
Adrienne Breef-Pilz https://orcid.org/0000-0002-6759-0063
Cayelan C. Carey https://orcid.org/0000-0001-8835-4476

REFERENCES

Abdul Wahid, A., and E. Arunbabu. 2022. "Forecasting Water Quality Using Seasonal ARIMA Model by Integrating In-Situ Measurements and Remote Sensing Techniques in Krishnagiri Reservoir, India." *Water Practice and Technology* 17: 1230–52.

Baracchini, T., P. Y. Chu, J. Šukys, G. Lieberherr, S. Wunderle, A. Wüest, and D. Bouffard. 2020. "Data Assimilation of In Situ and Satellite Remote Sensing Data to 3D Hydrodynamic Lake Models: A Case Study Using Delft3D-FLOW v4.03 and OpenDA v2.4." *Geoscientific Model Development* 13: 1267–84.

Baracchini, T., A. Wüest, and D. Bouffard. 2020. "Meteolakes: An Operational Online Three-Dimensional Forecasting Platform for Lake Hydrodynamics." *Water Research* 172: 115529.

Bennett, N. D., B. F. Croke, G. Guariso, J. H. Guillaume, S. H. Hamilton, A. J. Jakeman, S. Marsili-Libelli, et al. 2013. "Characterising Performance of Environmental Models." *Environmental Modelling & Software* 40: 1–20.

Bruce, L. C., M. A. Frassl, G. B. Arhonditsis, G. Gal, D. P. Hamilton, P. C. Hanson, A. L. Hetherington, et al. 2018. "A Multi-Lake Comparative Analysis of the General Lake Model (GLM): Stress-Testing across a Global Observatory Network." *Environmental Modelling & Software* 102: 274–291.

Carey, C. C., and A. Breef-Pilz. 2022. "Ice Cover Data for Falling Creek Reservoir and Beaverdam Reservoir, Vinton, Virginia, USA for 2013-2022." Environmental Data Initiative Repository. https://doi.org/10.6073/pasta/917b3947d91470eecf979e9297ed4d2e.

Carey, C. C., A. Breef-Pilz, B. J. Bookout, R. P. McClure, and J. H. Wynne. 2023. "Time Series of High-Frequency Sensor Data Measuring Water Temperature, Dissolved Oxygen, Conductivity, Specific Conductance, Total Dissolved Solids, Chlorophyll a, Phycocyanin, Fluorescent Dissolved Organic Matter, and Turbidity at Discrete Depths in Beaverdam Reservoir, Virginia, USA in 2016-2022." Environmental Data Initiative Repository. https://doi.org/10.6073/pasta/4182de376fde52e15d493fdd9f26d0c7.

Carey, C. C., P. C. Hanson, R. Q. Thomas, A. B. Gerling, A. G. Hounshell, A. S. L. Lewis, M. E. Lofton, et al. 2022. "Anoxia Decreases the Magnitude of the Carbon, Nitrogen, and Phosphorus Sink in Freshwaters." *Global Change Biology* 28: 4861–81.

Carey, C. C., A. S. Lewis, R. P. McClure, A. B. Gerling, A. Breef-Pilz, and A. Das. 2022. "Time Series of High-Frequency Profiles of Depth, Temperature, Dissolved Oxygen, Conductivity, Specific Conductance, Chlorophyll a, Turbidity, pH, Oxidation-Reduction Potential, Photosynthetic Active Radiation, and Descent Rate for Beaverdam Reservoir, Carvins Cove Reservoir, Falling Creek Reservoir, Gatewood Reservoir, and Spring Hollow Reservoir in Southwestern Virginia, USA 2013-2021." Environmental Data Initiative Repository. https://doi.org/10.6073/pasta/c4c45b5b10b4cb4cd4b5e613c3effbd0.

Carey, C. C., A. S. L. Lewis, D. W. Howard, W. M. Woelmer, P. A. Gantzer, K. A. Bierlein, J. C. Little, and WVWA. 2022. "Bathymetry and Watershed Area for Falling Creek Reservoir, Beaverdam Reservoir, and Carvins Cove Reservoir." Environmental Data Initiative Repository. https://doi.org/10.6073/pasta/352735344150f7e77d2bc18b69a22412.

Carey, C. C., W. M. Woelmer, M. E. Lofton, R. J. Figueiredo, B. J. Bookout, R. S. Corrigan, V. Daneshmand, et al. 2022. "Advancing Lake and Reservoir Water Quality Management with Near-Term, Iterative Ecological Forecasting." *Inland Waters* 12: 107–120.

Carpenter, S. R., E. H. Stanley, and M. J. Vander Zanden. 2011. "State of the World's Freshwater Ecosystems: Physical, Chemical, and Biological Changes." *Annual Review of Environment and Resources* 36: 75–99.

Chen, C., Q. Chen, G. Li, M. He, J. Dong, H. Yan, Z. Wang, and Z. Duan. 2021. "A Novel Multi-Source Data Fusion Method Based on Bayesian Inference for Accurate Estimation of Chlorophyll-a Concentration over Eutrophic Lakes." *Environmental Modelling & Software* 141: 105057.

Cho, H., and H. Park. 2019. "Merged-LSTM and Multistep Prediction of Daily Chlorophyll-a Concentration for Algal Bloom Forecast." *IOP Conference Series: Earth and Environmental Science* 351: 012020.

Cho, K. H., Y. Pachepsky, M. Ligaray, Y. Kwon, and K. H. Kim. 2020. "Data Assimilation in Surface Water Quality Modeling: A Review." *Water Research* 186: 116307.

Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, et al. 2001. "Ecological Forecasts: An Emerging Imperative." *Science* 293: 657–660.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez. 2016. "Convection-Permitting Models: A Step-Change in Rainfall Forecasting." *Meteorological Applications* 23: 165–181.

Clayer, F., L. Jackson-Blake, D. Mercado-Bettín, M. Shikhani, A. French, T. Moore, J. Sample, et al. 2023. "Sources of Skill in Lake Temperature, Discharge and Ice-Off Seasonal Forecasting Tools." *Hydrology and Earth System Sciences* 27: 1361–81.

Corbari, C., R. Salerno, A. Ceppi, V. Telesca, and M. Mancini. 2019. "Smart Irrigation Forecast Using Satellite LANDSAT Data and Meteo-Hydrological Modeling." *Agricultural Water Management* 212: 283–294.

Daneshmand, V., A. Breef-Pilz, C. C. Carey, Y. Jin, Y.-J. Ku, K. C. Subratie, R. Q. Thomas, and R. J. Figueiredo. 2021. "Edge-to-cloud Virtualized Cyberinfrastructure for Near Real-time Water Quality Forecasting in Lakes and Reservoirs." *2021 IEEE 17th International Conference on eScience (eScience)*. IEEE, pp. 138–148.

Derot, J., H. Yajima, and F. G. Schmitt. 2020. "Benefits of Machine Learning and Sampling Frequency on Phytoplankton Bloom Forecasts in Coastal Areas." *Ecological Informatics* 60: 101174.

Dietze, M. C. 2017a. *Ecological Forecasting*. Princeton, NJ: Princeton University Press.

Dietze, M. C. 2017b. "Prediction in Ecology: A First-Principles Framework." *Ecological Applications* 27: 2048–60.

Dietze, M. C., A. Fox, L. M. Beck-Johnson, J. L. Betancourt, M. B. Hooten, C. S. Jarnevich, T. H. Keitt, et al. 2018. "Iterative Near-Term Ecological Forecasting: Needs, Opportunities, and Challenges." *Proceedings of the National Academy of Sciences of the United States of America* 115: 1424–32.

Diez, J. M., I. Ibáñez, A. J. Miller-Rushing, S. J. Mazer, T. M. Crimmins, M. A. Crimmins, C. D. Bertelsen, and D. W. Inouye. 2012. "Forecasting Phenology: From Species Variability to Community Patterns." *Ecology Letters* 15: 545–553.

Doubek, J. P., K. L. Campbell, M. E. Lofton, R. P. McClure, and C. C. Carey. 2019. "Hypolimnetic Hypoxia Increases the Biomass Variability and Compositional Variability of Crustacean Zooplankton Communities." *Water* 11: 2179.

Duc, L., K. Saito, and D. Hotta. 2021. "Analysis and Design of Covariance Inflation Methods Using Inflation Functions. Part 2: Adaptive Inflation." *Quarterly Journal of the Royal Meteorological Society* 147: 2375–94.

Engelhardt, C., and G. Kirillin. 2014. "Criteria for the Onset and Breakup of Summer Lake Stratification Based on Routine Temperature Measurements." *Fundamental and Applied Limnology* 184: 183–194.

Evensen, G. 2003. "The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation." *Ocean Dynamics* 53: 343–367.

Francy, D. S., J. L. Graham, E. A. Stelzer, C. D. Ecker, A. M. Brady, P. Struffolino, and K. A. Loftin. 2015. "Water Quality,

Cyanobacteria, and Environmental Factors and Their Relations to Microcystin Concentrations for Use in Predictive Models at Ohio Lake Erie and Inland Lake Recreational Sites, 2013-14." US Geological Survey Scientific Investigations Report 2015–5120. https://doi.org/10.3133/sir20155120.

Georgakakos, K. P., N. E. Graham, T. M. Carpenter, and H. Yao. 2005. "Integrating Climate-Hydrology Forecasts and Multi-Objective Reservoir Management for Northern California." *EOS. Transactions of the American Geophysical Union* 86: 122–27.

George, D. G., and M. A. Hurley. 2004. "The Influence of Sampling Frequency on the Detection of Long-Term Change in Three Lakes in the English Lake District." *Aquatic Ecosystem Health and Management* 7: 1–14.

Gilarranz, L. J., A. Narwani, D. Odermatt, R. Siber, and V. Dakos. 2022. "Regime Shifts, Trends, and Variability of Lake Productivity at a Global Scale." *Proceedings of the National Academy of Sciences of the United States of America* 119: e2116413119.

Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman. 2005. "Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation." *Monthly Weather Review* 133: 1098–1118.

Gottwald, G. A., and S. Reich. 2021. "Supervised Learning from Noisy Observations: Combining Machine-Learning Techniques with Data Assimilation." *Physica D: Nonlinear Phenomena* 423: 132911.

Grönquist, P., C. Yao, T. Ben-Nun, N. Dryden, P. Dueben, S. Li, and T. Hoefler. 2021. "Deep Learning for Post-Processing Ensemble Weather Forecasts." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379: 20200092.

Hamre, K. D., M. E. Lofton, R. P. McClure, Z. W. Munger, J. P. Doubek, A. B. Gerling, M. E. Schreiber, and C. C. Carey. 2018. "In Situ Fluorometry Reveals a Persistent, Perennial Hypolimnetic Cyanobacterial Bloom in a Seasonally Anoxic Reservoir." *Freshwater Science* 37: 483–495.

Harris, D. J., S. D. Taylor, and E. P. White. 2018. "Forecasting Biodiversity in Breeding Birds Using Best Practices." *PeerJ* 6: e4278.

He, H., L. Lei, J. S. Whitaker, and Z.-M. Tan. 2020. "Impacts of Assimilation Frequency on Ensemble Kalman Filter Data Assimilation and Imbalances." *Journal of Advances in Modeling Earth Systems* 12: e2020MS002187.

Heilman, K. A., M. C. Dietze, A. A. Arizpe, J. Aragon, A. Gray, J. D. Shaw, A. O. Finley, S. Klesse, R. J. DeRose, and M. E. K. Evans. 2022. "Ecological Forecasting of Tree Growth: Regional Fusion of Tree-Ring and Forest Inventory Data to Quantify Drivers and Characterize Uncertainty." *Global Change Biology* 28: 2442–60.

Herrick, C., B. G. Steele, J. A. Brentrup, K. L. Cottingham, M. J. Ducey, D. A. Lutz, M. W. Palace, M. C. Thompson, J. V. Trout-Haney, and K. C. Weathers. 2023. "lakeCoSTR: A Tool to Facilitate Use of Landsat Collection 2 to Estimate Lake Surface Water Temperatures." *Ecosphere* 14: e4357.

Hipsey, M. R., C. Boon, L. C. Bruce, Q. Thomas, M. Weber, L. Winslow, J. S. Read, and D. P. Hamilton. 2022. "AquaticEcoDynamics/glm-aed: v3.3.0." https://doi.org/10.5281/zenodo.7047527.

Hipsey, M. R., L. C. Bruce, C. Boon, B. Busch, C. C. Carey, D. P. Hamilton, P. C. Hanson, et al. 2019. "A General Lake Model (GLM 3.0) for Linking with High-Frequency Sensor Data from the Global Lake Ecological Observatory Network (GLEON)." *Geoscientific Model Development* 12: 473–523.

Hollan, E., and T. J. Simons. 1978. "Wind-Induced Changes of Temperature and Currents in Lake Constance." *Archiv für Meteorologie, Geophysik und Bioklimatologie, Series A* 27: 333–373.

Hounshell, A. G., R. P. McClure, M. E. Lofton, and C. C. Carey. 2021. "Whole-Ecosystem Oxygenation Experiments Reveal Substantially Greater Hypolimnetic Methane Concentrations in Reservoirs during Anoxia." *Limnology and Oceanography Letters* 6: 33–42.

Ibáñez, I., E. S. Gornish, L. Buckley, D. M. Debinski, J. Hellmann, B. Helmuth, J. HilleRisLambers, A. M. Latimer, A. J. Miller-Rushing, and M. Uriarte. 2013. "Moving Forward in Global-Change Ecology: Capitalizing on Natural Variability." *Ecology and Evolution* 3: 170–181.

Jolliffe, I. T., and D. B. Stephenson. 2012. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Hoboken, NJ: John Wiley & Sons.

Kehoe, M. J., K. P. Chun, and H. M. Baulch. 2015. "Who Smells? Forecasting Taste and Odor in a Drinking Water Reservoir." *Environmental Science & Technology* 49: 10984–92.

Kirchner, J. W., and C. Neal. 2013. "Universal Fractal Scaling in Stream Chemistry and Its Implications for Solute Transport and Water Quality Trend Detection." *Proceedings of the National Academy of Sciences of the United States of America* 110: 12213–18.

LaDeau, S. L., B. A. Han, E. J. Rosi-Marshall, and K. C. Weathers. 2017. "The Next Decade of Big Data in Ecosystem Science." *Ecosystems* 20: 274–283.

Ladwig, R., L. A. Rock, and H. A. Dugan. 2021. "Impact of Salinization on Lake Stratification and Spring Mixing." *Limnology and Oceanography Letters* 8: 93–102.

Lewis, A. S. L., W. M. Woelmer, H. L. Wander, D. W. Howard, J. W. Smith, R. P. McClure, M. E. Lofton, et al. 2022. "Increased Adoption of Best Practices in Ecological Forecasting Enables Comparisons of Forecastability." *Ecological Applications* 32: e2500.

Lin, E., Y. Yang, X. Qiu, Q. Xie, R. Gan, B. Zhang, and X. Liu. 2021. "Impacts of the Radar Data Assimilation Frequency and Large-Scale Constraint on the Short-Term Precipitation Forecast of a Severe Convection Case." *Atmospheric Research* 257: 105590.

Liu, H., Y. D. Chen, T. Liu, and L. Lin. 2019. "The River Chief System and River Pollution Control in China: A Case Study of Foshan." *Water* 11: 1606.

Lofton, M. E., J. A. Brentrup, W. S. Beck, J. A. Zwart, R. Bhattacharya, L. S. Brighenti, S. H. Burnet, et al. 2022. "Using Near-Term Forecasts and Uncertainty Partitioning to Inform Prediction of Oligotrophic Lake Cyanobacterial Density." *Ecological Applications* 32: e2590.

Lofton, M. E., D. W. Howard, R. Q. Thomas, and C. C. Carey. 2023. "Progress and Opportunities in Advancing Near-Term Forecasting of Freshwater Quality." *Global Change Biology* 29: 1691–1714.

Luo, Y., K. Ogle, C. Tucker, S. Fei, C. Gao, S. LaDeau, J. S. Clark, and D. S. Schimel. 2011. "Ecological Forecasting and Data

Assimilation in a Data-Rich Era." *Ecological Applications* 21: 1429–42.

Machete, R. L., and L. A. Smith. 2016. "Demonstrating the Value of Larger Ensembles in Forecasting Physical Systems." *Tellus Series A: Dynamic Meteorology and Oceanography* 68: 28393.

Magnuson, J. J., L. B. Crowder, and P. A. Medvick. 1979. "Temperature as an Ecological Resource." *American Zoologist* 19: 331–343.

Malhi, Y., J. Franklin, N. Seddon, M. Solan, M. G. Turner, C. B. Field, and N. Knowlton. 2020. "Climate Change and Ecosystems: Threats, Opportunities and Solutions." *Philosophical Transactions of the Royal Society B: Biological Sciences* 375: 20190104.

Mantovani, C., L. Corgnati, J. Horstmann, A. Rubio, E. Reyes, C. Quentin, S. Cosoli, J. L. Asensio, J. Mader, and A. Griffa. 2020. "Best Practices on High Frequency Radar Deployment and Operation for Ocean Current Measurement." *Frontiers in Marine Science* 7: 210.

Marcé, R., G. George, P. Buscarinu, M. Deidda, J. Dunalska, E. de Eyto, G. Flaim, et al. 2016. "Automatic High Frequency Monitoring for Improved Lake and Reservoir Management." *Environmental Science & Technology* 50: 10780–94.

Marj, A. F., and A. M. J. Meijerink. 2011. "Agricultural Drought Forecasting Using Satellite Images, Climate Indices and Artificial Neural Network." *International Journal of Remote Sensing* 32: 9707–19.

Massoud, E. C., J. Huisman, E. Benincà, M. C. Dietze, W. Bouten, and J. A. Vrugt. 2018. "Probing the Limits of Predictability: Data Assimilation of Chaotic Dynamics in Complex Food Webs." *Ecology Letters* 21: 93–103.

McClure, R. P., R. Q. Thomas, M. E. Lofton, W. M. Woelmer, and C. C. Carey. 2021. "Iterative Forecasting Improves Near-Term Predictions of Methane Ebullition Rates." *Frontiers in Environmental Science* 9: 756603.

McSharry, P. E., S. Bouwman, and G. Bloemhof. 2005. "Probabilistic Forecasts of the Magnitude and Timing of Peak Electricity Demand." *IEEE Transactions on Power Systems* 20: 1166–72.

Mercado-Bettín, D., F. Clayer, M. Shikhani, T. N. Moore, M. D. Frías, L. Jackson-Blake, J. Sample, et al. 2021. "Forecasting Water Temperature in Lakes and Reservoirs Using Seasonal Climate Prediction." *Water Research* 201: 117286.

Meyer, J. L., M. J. Sale, P. J. Mulholland, and N. L. Poff. 1999. "Impacts of Climate Change on Aquatic Ecosystem Functioning and Health." *JAWRA Journal of the American Water Resources Association* 35: 1373–86.

Mi, C., T. Shatwell, J. Ma, Y. Xu, F. Su, and K. Rinke. 2020. "Ensemble Warming Projections in Germany's Largest Drinking Water Reservoir and Potential Adaptation Strategies." *Science of the Total Environment* 748: 141366.

Moustahfid, H., L. C. Hendrickson, A. Arkhipkin, G. J. Pierce, A. Gangopadhyay, H. Kidokoro, U. Markaida, et al. 2021. "Ecological-Fishery Forecasting of Squid Stock Dynamics under Climate Variability and Change: Review, Challenges, and Recommendations." *Reviews in Fisheries Science & Aquaculture* 29: 682–705.

Niu, S., Y. Luo, M. C. Dietze, T. F. Keenan, Z. Shi, J. Li, and F. S. C. Iii. 2014. "The Role of Data Assimilation in Predictive Ecology." *Ecosphere* 5: 1–16.

O'Reilly, C. M., S. Sharma, D. K. Gray, S. E. Hampton, J. S. Read, R. J. Rowley, P. Schneider, et al. 2015. "Rapid and Highly Variable Warming of Lake Surface Waters around the Globe." *Geophysical Research Letters* 42: 10–773.

Paerl, H. W., and V. J. Paul. 2012. "Climate Change: Links to Global Expansion of Harmful Cyanobacteria." *Water Research* 46: 1349–63.

Page, T., P. J. Smith, K. J. Beven, I. D. Jones, J. A. Elliott, S. C. Maberly, E. B. Mackay, M. De Ville, and H. Feuchtmayr. 2017. "Constraining Uncertainty and Process-Representation in an Algal Community Lake Model Using High Frequency In-Lake Observations." *Ecological Modelling* 357: 1–13.

Page, T., P. J. Smith, K. J. Beven, I. D. Jones, J. A. Elliott, S. C. Maberly, E. B. Mackay, M. De Ville, and H. Feuchtmayr. 2018. "Adaptive Forecasting of Phytoplankton Communities." *Water Research* 134: 74–85.

Park, J., K. T. Kim, and W. H. Lee. 2020. "Recent Advances in Information and Communications Technology (ICT) and Sensor Technology for Monitoring Water Quality." *Water* 12: 510.

Piazzi, G., G. Thirel, L. Campo, and S. Gabellani. 2018. "A Particle Filter Scheme for Multivariate Data Assimilation into a Point-Scale Snowpack Model in an Alpine Environment." *The Cryosphere* 12: 2287–2306.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Read, J. S., X. Jia, J. Willard, A. P. Appling, J. A. Zwart, S. K. Oliver, A. Karpatne, et al. 2019. "Process-Guided Deep Learning Predictions of Lake Water Temperature." *Water Resources Research* 55: 9173–90.

Read, J. S., L. A. Winslow, G. J. A. Hansen, J. Van Den Hoek, P. C. Hanson, L. C. Bruce, and C. D. Markfort. 2014. "Simulating 2368 Temperate Lakes Reveals Weak Coherence in Stratification Phenology." *Ecological Modelling* 291: 142–150.

Romero, J. R., I. Kagalou, J. Imberger, D. Hela, M. Kotti, A. Bartzokas, T. Albanis, et al. 2002. "Seasonal Water Quality of Shallow and Eutrophic Lake Pamvotis, Greece: Implications for Restoration." *Hydrobiologia* 474: 91–105.

Serruya, S. 1975. "Wind, Water Temperature and Motions in Lake Kinneret: General Pattern." *Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen* 19: 73–87.

Simonin, D., C. Pierce, N. Roberts, S. P. Ballard, and Z. Li. 2017. "Performance of Met Office Hourly Cycling NWP-Based Nowcasting for Precipitation Forecasts." *Quarterly Journal of the Royal Meteorological Society* 143: 2862–73.

Simonis, J. L., E. P. White, and S. M. Ernest. 2021. "Evaluating Probabilistic Ecological Forecasts." *Ecology* 102: e03431.

Smith, L. A., E. B. Suckling, E. L. Thompson, T. Maynard, and H. Du. 2015. "Towards Improving the Framework for Probabilistic Forecast Evaluation." *Climatic Change* 132: 31–45.

Steere, D. C., A. Baptista, D. McNamee, C. Pu, and J. Walpole. 2000. "Research Challenges in Environmental Observation and Forecasting Systems." In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking – MobiCom '00* 292–99. Boston, MA: ACM Press.

Tanut, B., R. Waranusast, and P. Riyamongkol. 2021. "High Accuracy Pre-Harvest Sugarcane Yield Forecasting Model

Utilizing Drone Image Analysis, Data Mining, and Reverse Design Method." *Agriculture* 11: 682.

Tasnim, B., J. A. Jamily, X. Fang, Y. Zhou, and J. S. Hayworth. 2021. "Simulating Diurnal Variations of Water Temperature and Dissolved Oxygen in Shallow Minnesota Lakes." *Water* 13: 1980.

Thomas, R. Q., C. Boettiger, C. C. Carey, M. C. Dietze, L. R. Johnson, M. A. Kenney, J. S. McLachlan, et al. 2023. "The NEON Ecological Forecasting Challenge." *Frontiers in Ecology and the Environment* 21: 112–13.

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 Resources Research* 56: e2019WR026138.

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." *Frontiers in Ecology and the Environment* 21: 220–26.

Wander, H. L., R. Q. Thomas, T. N. Moore, M. E. Lofton, A. Breef-Pilz, and C. C. Carey. 2023a. "Data Assimilation Experiments Inform Monitoring Needs for Near-Term Ecological Forecasts in a Eutrophic Reservoir: Data, Forecasts, and Scores (Version 3) [Data Set]." Zenodo. https://doi.org/10.5281/zenodo.7925097.

Wander, H. L., R. Q. Thomas, T. N. Moore, M. E. Lofton, A. Breef-Pilz, and C. C. Carey. 2023b. "hlwander/BVRE-forecast-code: Data Assimilation Experiments Inform Monitoring Needs for Near-Term Ecological Forecasts in a Eutrophic Reservoir: Code (v1.2)." Zenodo. https://doi.org/10.5281/zenodo.7958470.

Wang, S., N. Flipo, and T. Romary. 2023. "Which Filter for Data Assimilation in Water Quality Models? Focus on Oxygen Reaeration and Heterotrophic Bacteria Activity." *Journal of Hydrology* 620: 129423.

Wang, X., J. Zhang, and V. Babovic. 2016. "Improving Real-Time Forecasting of Water Quality Indicators with Combination of Process-Based Models and Data Assimilation Technique." *Ecological Indicators* 66: 428–439.

Weng, E., and Y. Luo. 2011. "Relative Information Contributions of Model vs. Data to Short- and Long-Term Forecasts of Forest Carbon Dynamics." *Ecological Applications* 21: 1490–1505.

White, E. P., G. M. Yenni, S. D. Taylor, E. M. Christensen, E. K. Bledsoe, J. L. Simonis, and S. K. M. Ernest. 2019. "Developing an Automated Iterative Near-Term Forecasting System for an Ecological Study." *Methods in Ecology and Evolution* 10: 332–344.

Williamson, C. E., E. P. Overholt, J. A. Brentrup, R. M. Pilla, T. H. Leach, S. G. Schladow, J. D. Warren, et al. 2016. "Sentinel Responses to Droughts, Wildfires, and Floods: Effects of UV Radiation on Lakes and Their Ecosystem Services." *Frontiers in Ecology and the Environment* 14: 102–9.

Woelmer, W. M., R. Q. Thomas, M. E. Lofton, R. P. McClure, H. L. Wander, and C. C. Carey. 2022. "Near-Term Phytoplankton Forecasts Reveal the Effects of Model Time Step and Forecast Horizon on Predictability." *Ecological Applications* 32: e2642.

Woolway, R. I., E. Jennings, T. Shatwell, M. Golub, D. C. Pierson, and S. C. Maberly. 2021. "Lake Heatwaves under Climate Change." *Nature* 589: 402–7.

Yvon-Durocher, G., J. M. Caffrey, A. Cescatti, M. Dossena, P. del Giorgio, J. M. Gasol, J. M. Montoya, et al. 2012. "Reconciling the Temperature Dependence of Respiration across Timescales and Ecosystem Types." *Nature* 487: 472–76.

Ziliani, M. G., R. Ghostine, B. Ait-El-Fquih, M. F. McCabe, and I. Hoteit. 2019. "Enhanced Flood Forecasting through Ensemble Data Assimilation and Joint State-Parameter Estimation." *Journal of Hydrology* 577: 123924.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wander, Heather L., R. Quinn Thomas, Tadhg N. Moore, Mary E. Lofton, Adrienne Breef-Pilz, and Cayelan C. Carey. 2024. "Data Assimilation Experiments Inform Monitoring Needs for Near-Term Ecological Forecasts in a Eutrophic Reservoir." *Ecosphere* 15(2): e4752. https://doi.org/10.1002/ecs2.4752