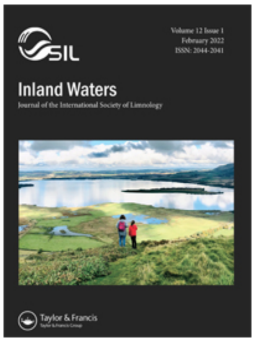

Inland Waters

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tinw20

Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting

Cayelan C. Carey, Whitney M. Woelmer, Mary E. Lofton, Renato J. Figueiredo, Bethany J. Bookout, Rachel S. Corrigan, Vahid Daneshmand, Alexandria G. Hounshell, Dexter W. Howard, Abigail S. L. Lewis, Ryan P. McClure, Heather L. Wander, Nicole K. Ward & R. Quinn Thomas

To cite this article: Cayelan C. Carey, Whitney M. Woelmer, Mary E. Lofton, Renato J. Figueiredo, Bethany J. Bookout, Rachel S. Corrigan, Vahid Daneshmand, Alexandria G. Hounshell, Dexter W. Howard, Abigail S. L. Lewis, Ryan P. McClure, Heather L. Wander, Nicole K. Ward & R. Quinn Thomas (2022) Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting, Inland Waters, 12:1, 107-120, DOI: 10.1080/20442041.2020.1816421

To link to this article: https://doi.org/10.1080/20442041.2020.1816421

Published online: 18 Jan 2021.

INLAND WATERS
2022, VOL. 12, NO. 1, 107–120
https://doi.org/10.1080/20442041.2020.1816421

Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting

Cayelan C. Carey [a*], Whitney M. Woelmer [a*], Mary E. Lofton [a*], Renato J. Figueiredo [b*], Bethany J. Bookout [a*], Rachel S. Corrigan [c*], Vahid Daneshmand [b*], Alexandria G. Hounshell [a*], Dexter W. Howard [a*], Abigail S. L. Lewis [a*], Ryan P. McClure [a*], Heather L. Wander [a*], Nicole K. Ward [a*] and R. Quinn Thomas [c*]

[a]Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA; [b]Department of Electrical & Computer Engineering, University of Florida, Gainesville, Florida, USA; [c]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, USA

ABSTRACT
Near-term, iterative ecological forecasts with quantified uncertainty have great potential for improving lake and reservoir management. For example, if managers received a forecast indicating a high likelihood of impending impairment, they could make decisions today to prevent or mitigate poor water quality in the future. Increasing the number of automated, real-time freshwater forecasts used for management requires integrating interdisciplinary expertise to develop a framework that seamlessly links data, models, and cyberinfrastructure, as well as collaborations with managers to ensure that forecasts are embedded into decision-making workflows. The goal of this study is to advance the implementation of near-term, iterative ecological forecasts for freshwater management. We first provide an overview of FLARE (Forecasting Lake And Reservoir Ecosystems), a forecasting framework we developed and applied to a drinking water reservoir to assist water quality management, as a potential open-source option for interested users. We used FLARE to develop scenario forecasts simulating different water quality interventions to inform manager decision-making. Second, we share lessons learned from our experience developing and running FLARE over 2 years to inform other forecasting projects. We specifically focus on how to develop, implement, and maintain a forecasting system used for active management. Our goal is to break down the barriers to forecasting for freshwater researchers, with the aim of improving lake and reservoir management globally.

ARTICLE HISTORY
Received 17 February 2020
Accepted 18 August 2020

KEYWORDS
data assimilation; FAIR data principles; FLARE; human-centered design; quantified uncertainty; real-time forecast

Introduction

Water quality in lakes and reservoirs around the world is becoming increasingly variable as a result of human activities (reviewed by Jiménez Cisneros et al. 2014). For example, the prevalence and duration of hypolimnetic anoxia is increasing in many lakes because of climate and land use change (Jenny et al. 2016), yet waterbodies are simultaneously experiencing more powerful storms that initiate mixing and increase oxygen availability (Prein et al. 2017), resulting in large day-to-day changes in oxygen concentrations (e.g., Perello et al. 2017). The increasing variability of many water quality metrics outside the envelope of historical conditions makes it challenging to anticipate future water quality, putting a substantial strain on managers responsible for provisioning critical lake and reservoir ecosystem services on a daily basis (Brookes et al. 2014, Khan et al. 2015).

The emerging discipline of ecological forecasting provides a novel approach for preemptively managing lakes and reservoirs in the face of increasing water quality variability. Ecological forecasting, or the prediction of future ecosystem properties with quantified uncertainty (sensu Clark et al. 2001, Luo et al. 2011, Dietze et al. 2018; Table 1), provides a useful tool for managers. Forecasts provide managers with probabilistic estimates of future water quality conditions in their focal lake or reservoir, thus allowing them to take preemptive management actions to mitigate or prevent water quality impairment.

CONTACT Cayelan C. Carey cayelan@vt.edu Department of Biological Sciences, Virginia Tech, 926 W. Campus Drive, Blacksburg, Virginia 24061, USA
*The FLARE forecasting system that inspired this manuscript was developed by RQT, CCC, RJF, and VD, with substantial help from BJB, MEL, RPM, and WMW. CCC led the overall development of this manuscript with RQT, WMW, MEL, and RJF. All authors collaboratively wrote the manuscript, provided feedback, and approved its final version.
Supplemental data for this article can be accessed here: https://doi.org/10.1080/20442041.2020.1816421.

Table 1. Glossary of ecological forecasting terms.

Term	Definition	Example
Data assimilation	The process of statistically comparing a forecast with new observations as they become available to update the forecasted states and, in some applications, model parameters for the next time step in the forecasting workflow	Using an ensemble Kalman filter (Evensen 1994, Evensen 2009) to update yesterday's forecast of what DO would be today with today's observed DO concentration. The updated DO concentrations serve as initial conditions for tomorrow's forecast
Ensemble	A set of forecast outputs that propagate alternative competing predictions, enabling quantification of different uncertainty sources in the forecast	100 unique forecasts of lake phosphorus concentration for next week created from 100 different time series models (e.g., each model had a different combination of model predictors), which can be used to assess uncertainty in model selection
FAIR	Best practice data principles that require data to be findable, accessible, interoperable, and reusable (FAIR; Wilkinson et al. 2016)	Archiving forecast driver data and forecasted states within a repository that follows FAIR principles, such as the Environmental Data Initiative repository
Forecast	A prediction of a future ecosystem property (or properties) with quantified uncertainty	Dissolved organic carbon concentrations next week will be 5.2 mg L^{-1}, with 95% predictive intervals ranging from 4.6 to 5.8 mg L^{-1}
Forecast uncertainty	A quantitative estimate of a forecast's uncertainty. See *partitioned uncertainty* for a list of potential sources of uncertainty that could be included, often derived from ensemble methods	Estimating that the 95% predictive intervals around next month's salmonid biomass forecast would be 3 (SD 0.4) kg ha^{-1}
Hindcast	A prediction with specified uncertainty developed for a time period that has already occurred (i.e., the observational data needed to evaluate the forecast are already available; Jolliffe and Stephenson 2003)	Weekly zooplankton biomass density forecasts created for last summer to test a forecasting workflow
Human-centered design	An approach to system design that includes the human user in all aspects of system development (following Kurosu 2011)	Including stakeholders throughout the forecast development (e.g., iteratively working with managers to develop the forecast delivery system and visualizations)
Iterative forecasting	The process of repeatedly generating forecasts on a regular interval that are compared with observations when they are available to update the forecast for future time steps	Producing chlorophyll *a* forecasts every day and comparing forecast output with observed chlorophyll *a* from a sensor. The comparison includes updating the forecast using data assimilation to inform tomorrow's forecast
Near-term	Pertaining to the near future (hours to a decade), allowing for quick, iterative validation of forecast accuracy with new data as they become available	A forecast of water temperature for the next 2 days in the future
Partitioned uncertainty	An analysis of forecasts to determine the individual contribution of different sources of uncertainty to the total forecast uncertainty (e.g., driver data, initial conditions, observations, parameters, process, model selection, scenarios). See Dietze (2017) for a comprehensive explanation of each uncertainty source	60% of the total forecast uncertainty is due to uncertainty in future weather (driver data uncertainty)
Prediction	A quantitative hypothesis of future conditions based on the output of a model	An estimate that next week's hypolimnetic iron concentration will be 0.5 mg L^{-1}
Projection	A probabilistic forecast based on an explicit scenario	A forecast of DO concentration next week in response to activation of a hypolimnetic oxygenation system by a manager
Real-time	Integrating data into the forecast workflow on the time scale the data are collected through the use of automated sensors, wireless data transfer, and cyberinfrastructure	Producing a water temperature forecast for next week using water temperature data that was collected today and assimilated into the forecast workflow using an automated sensor network and cyberinfrastructure

These definitions are adapted from Dietze (2017) and Luo et al. (2011) for a lake and reservoir management focus, unless otherwise specified.

Producing ecological forecasts that are near-term and iterative both adds to their utility as decision support tools for managers and advances the science of ecological forecasting. While not all ecological forecasts are near-term or iterative (see Table 1 for definitions), our focus here is on the forecasts that do meet these criteria because they enable continuous forecast improvement. Near-term, iterative forecasts are continually compared to observed data as they become available. Forecasts are updated with observations via data assimilation techniques and then fed back into the forecasting workflow to generate future forecasts (Dietze 2017). The time step of the near-term forecast model determines how rapidly forecasts are updated with observational data. Hourly or daily iterative forecasts require a tightly integrated cyberinfrastructure connecting sensor data with models for continuous and ongoing data assimilation (Table 1).

Quantified uncertainty in predictions (Table 1) propagated from different sources is also a critical component of ecological forecasts. We note that although the term forecast is applied broadly in the literature, "forecasts" without quantified uncertainty are merely predictions (Luo et al. 2011, Dietze 2017; Table 1). Different contributors to total forecast uncertainty include (1) driver data uncertainty (uncertainty in the model forcing inputs, often meteorological and/or hydrological for lake ecosystem simulations), (2) initial conditions uncertainty (uncertainty in the forecast's starting conditions), (3) model selection uncertainty (the uncertainty introduced when using multiple models for forecasting), (4) observation error (the uncertainty in accurately measuring the variables being simulated by a model), (5) parameter uncertainty (uncertainty in the model parameters), (6) process uncertainty (the uncertainty in the ability of a model to correctly simulate the complex interacting processes occurring in an ecosystem), and (7) scenario uncertainty (the uncertainty introduced when forecasting multiple future trajectories (definitions adapted from Dietze 2017). Identifying the factors contributing to total forecast uncertainty can be used to prioritize improvements to the forecast workflow (Dietze 2017). In addition, knowing the uncertainty associated with a forecast may allow managers to make more informed decisions about the likelihood of a forecasted event (Morss et al. 2008, Berthet et al. 2016).

If managers had access to iterative, near-term water quality forecasts, they could act today to prepare for or preempt future water quality impairment. For example, knowing in advance that a phytoplankton bloom would occur next week in a drinking water supply would allow managers to initiate interventions in the reservoir (e.g., add algaecides, activate a mixing system), optimize water treatment for toxin removal (e.g., alter levels of potassium permanganate and activated carbon; Dugan et al. 2018), or change the depth at which drinking water is extracted for water treatment. Forecasts of other water quality variables that could be useful for water managers include water temperature to determine dam withdrawal schedules (Pike et al. 2013, Weber et al. 2017), organic matter concentrations to preempt potentially carcinogenic disinfection by-product (DBP) formation (Tomlinson et al. 2016), hypolimnetic dissolved oxygen (DO) to prevent the release of metals and nutrients from sediments (Gerling et al. 2016, Munger et al. 2019), and the concentrations of metals and other contaminants to ensure drinking water safety.

Despite the many potential benefits of ecological forecasts, near-term iterative forecasting is rarely used as a lake and reservoir management tool. Progress has been made in developing lake and reservoir forecasts (e.g., Page et al. 2018, Baracchini et al. 2020), but multiple challenges remain, and most freshwater forecasts to date have been for river discharge (e.g., Bal et al. 2014, Hague and Patterson 2014, Caissie et al. 2017, Ouellet-Proulx et al. 2017). Several examples of lake and reservoir water quality hindcasts (see Table 1) with data assimilation exist (Recknagel et al. 2014, 2016, Rowe et al. 2016), but few studies have successfully implemented iterative water quality forecasts of future conditions using data assimilation (e.g., Kim et al. 2012, Xiao et al. 2017, Xie et al. 2012, Baracchini et al. 2020). Some studies have developed forecasting systems with data assimilation for river water temperatures and discharge operating in real time (e.g., Anderson et al. 2010, Pike et al. 2013, Ouellet-Proulx et al. 2017), but real-time automated forecasts of future conditions that include biological and chemical attributes of water quality are rare.

Freshwater forecasters need to develop cyberinfrastructure and integrated model-data systems that create real-time forecasts for future conditions, not just hindcasts, with full quantification of multiple sources of forecast uncertainty (Table 1). The most commonly considered sources of uncertainty in freshwater forecasts are driver data and parameters. Initial conditions uncertainty and process uncertainty are occasionally quantified; however, model selection uncertainty and scenario uncertainty are rarely quantified (e.g., Huang et al. 2013, Pike et al. 2013, Kim et al. 2014, Page et al. 2018). Furthermore, among existing forecasts, few have developed real-time, automated forecasts for stakeholders (but see Rowe et al. 2016, Baracchini et al. 2020). Although managers were identified as the end user in all of these freshwater forecasting studies, the use of human-centered design methods (Table 1) to develop and implement forecasts remains rare, or at least is undocumented in the literature.

The goal of this paper is to advance the implementation of near-term, iterative ecological forecasts for freshwater management. We first provide an overview of a water quality forecasting system our team developed for a drinking water reservoir in partnership with a water utility as a potential open-source option for interested users. This automated forecasting system is near-term, iterative, and real-time, generating 16-day horizon forecasts of water temperature and DO that are updated with sensor data and delivered to managers every day. We present results from a case study in which the forecasting system was applied to a drinking water reservoir to guide the management of DO. Because no current lake or reservoir forecasting system provides managers with 16-day forecasts showing the future possible effects of a suite of management interventions (to the best of our knowledge), the DO forecasts highlight the potential of scenario-based forecasts for water quality management. Finally, we share our lessons learned from operating the forecasting system over 2 years (2018–present), specifically focusing on how to develop, implement, and maintain a forecasting system used for active management. While best practices for ecological forecasting have been proposed (e.g., Harris et al. 2018, White et al. 2019), these studies provide no guidance on creating forecasts used by managers for decision-making. Our overarching goal is to break down the barriers to using forecasting as a tool to improve lake and reservoir management.

Methods

Study site

Our near-term iterative water quality forecasting system was developed for Falling Creek Reservoir (FCR), a shallow (Z_{max} = 9.3 m) and small (0.119 km^2) dimictic drinking water reservoir located in southwest Virginia, USA (37.30°N, 79.84°W), operated by the Western Virginia Water Authority (WVWA; Gerling et al. 2016). To combat high nutrient and metal concentrations released from the sediments during anoxic periods, the WVWA deployed a hypolimnetic oxygenation (HOx) system to increase oxygen concentrations and subsequent water quality of FCR (Gerling et al. 2014). The HOx system is able to successfully increase oxygen concentrations in the hypolimnion and decrease the sediment release of nutrients and metals (Gerling et al. 2016, Munger et al. 2016, 2019), but oxygen addition is not always needed to maintain oxic conditions, thereby unnecessarily increasing treatment costs during those periods.

We focused on water temperature and oxygen forecasts because they were identified a priori by the water utility as useful for their management. Specifically, because it is challenging to know how much oxygen to add to the hypolimnion in advance, forecasts of oxygen concentrations under varying management scenarios would be particularly useful for WVWA managers to determine the minimum required HOx system oxygen inputs that maintain suitable water quality while minimizing cost. Forecasts of water temperature can also be used to identify when fall turnover will occur, which is historically accompanied by severe water quality impairment as hypolimnetic metals and nutrients are mixed throughout the water column.

Forecasting system description

We developed a near-term, real-time iterative forecasting system (FLARE: Forecasting Lake and Reservoir Ecosystems; Thomas et al. 2020a) that integrates water quality sensor data with models to make 16-day forecasts of future water quality (temperature and DO) conditions in FCR (Fig. 1). FLARE updates the forecasts daily with sensor observations using an ensemble Kalman filter and quantifies the contributions of driver data, initial conditions, parameters, and process to total forecast uncertainty.

Sensor data and connectivity

Sensor observations are used both as model driver data and for evaluating and updating forecasts with data assimilation (Fig. 1). At FCR, weather data are collected at 1-minute resolution from a meteorological station (sensors measure air temperature, wind speed, relative humidity, shortwave and longwave radiation, and precipitation; Carey et al. 2020c) located on the reservoir dam. A pressure transducer at a weir on the primary inflow stream to the reservoir measures inflow discharge and water temperature every 10 min (Carey et al. 2020a). At the deepest site of the reservoir near the dam, we measured water temperature at 1 m depth intervals from the surface to the sediments and DO using sensors deployed at 1.6, 5, and 9 m depths (Carey et al. 2020b). Every morning, sensor gateways

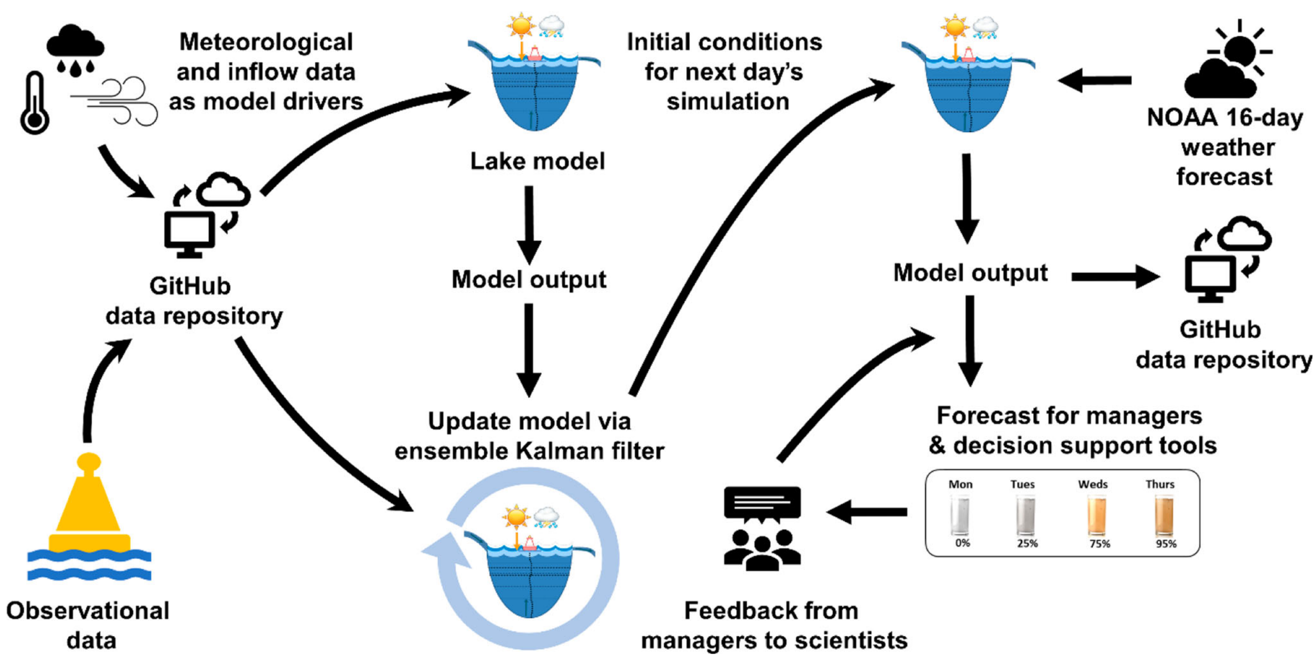

Figure 1. Conceptual overview of the daily iterative FLARE workflow from data collection to the creation of model output using real-time initial conditions and the US National Oceanic and Atmospheric Administration (NOAA) weather forecasts, to the creation of decision support tools for managers. The managers receive automated forecasts every morning and contribute feedback that is incorporated into subsequent iterations of the decision support tools. Color version available online.

(small computers deployed in the field that connect the sensors to the Internet) wirelessly transmit the data to a Git repository in the cloud, where the data are accessible for modeling using a virtual network (Subratie et al. 2020).

Model, data assimilation, and uncertainty

The meteorological and inflow data transmitted to GitHub are used as drivers of the 1-dimensional hydrodynamic General Lake Model (GLM; Hipsey et al. 2019). GLM is coupled to Aquatic EcoDynamics (AED) modules to simulate lake and reservoir thermal structure, biogeochemical cycling, and plankton food webs. The default modules include oxygen, carbon, silica, nitrogen, phosphorus, organic matter, phytoplankton, and zooplankton (Hipsey et al. 2013). Within GLM-AED, these modules can be turned on or off as needed to match the complexity of a forecasting application. To forecast FCR's thermal structure and DO, the only AED module that was turned on in our model setup was oxygen.

The daily data assimilation for FLARE includes 5 major steps that are triggered each morning (Fig. 1; described in detail by Thomas et al. 2020a). First, FCR water temperature and DO concentrations are simulated by GLM-AED for the preceding 24 h on an hourly time step, using observed meteorology and inflow data as model inputs. The model run is composed of 441 ensemble members, or individual iterations of the GLM-AED model, which differ slightly in their initial conditions and their parameters based on the outcome of prior data assimilation. Second, random noise is added to the model states (i.e., temperature and DO at each depth) for each ensemble member to represent process uncertainty. Third, the ensemble model output is then compared with the most recent observational data from the temperature and DO sensors using an ensemble Kalman filter. The ensemble Kalman filter statistically combines the model ensemble predictions and the observations to adjust the model states and the model parameters to be consistent with the observations. Fourth, the adjusted model states and model parameters for each GLM-AED ensemble are then used as initial conditions and model parameters for a 16-day forecast into the future, using meteorological forecasts from the US National Oceanic and Atmospheric Administration (NOAA) as driver data. The 16-day forecast includes the key sources of uncertainty for the system, including driver data uncertainty of both future weather and inflows. Finally, the outputs from the ensemble forecast are automatically processed to create visualizations, which are emailed to water managers every morning (Fig. 2). The ensemble output is also archived for future analysis in a GitHub repository for versioning control. We refer interested readers to Thomas et al. (2020a) for detailed information on FLARE setup and performance in forecasting thermal structure; here, we focus on DO forecasting and its application for management.

Oxygenation scenario forecasts

We used the FLARE system to develop scenario-based DO forecasts that inform HOx system activation and operation. These DO forecasts build on our existing infrastructure for water temperature forecasting detailed by Thomas et al. (2020a) by simulating a depth profile of DO and including 2 additional GLM-AED parameters in the oxygen module that are calibrated by the ensemble Kalman filter (see Supplemental Material for detailed DO forecasting methods, oxygen module configuration, parameters, and DO forecast performance evaluation).

With these DO forecasts, we can create scenarios simulating different levels of HOx system operation to compare potential future hypolimnetic oxygen concentrations. This information can be used to determine if the HOx system should be on or off, and what minimum level of oxygen managers should add to the hypolimnion to maintain water quality. Our iterative forecast cycle generated 2 different forecasts every day: a forecast with the oxygenation system activated at full capacity throughout the 16-day horizon and a forecast with the oxygenation system kept off throughout the 16-day horizon.

Results

The contrasting DO concentrations in the 2 oxygenation scenario forecasts highlight the utility of near-term, iterative forecasting for assisting water quality management (Fig. 3). From the full time series of DO forecasts produced for FCR (see Supplemental Material), we focus on 3 forecasts generated during varying initial DO conditions in summer 2019. In the first, the reservoir's hypolimnion is anoxic at the time of the forecast on 3 June 2019 (Fig. 3a). If the HOx system is activated and kept on, the forecast predicts an increase in oxygen concentrations above the sediments to ~200 mmol m^{-3} by 16 days into the future (blue solid line is the mean of 441 ensemble members with 95% confidence intervals represented by the dashed lines). By comparison, in the absence of HOx system activation, oxygen concentrations above the sediments will remain anoxic (red line). The HOx system was in

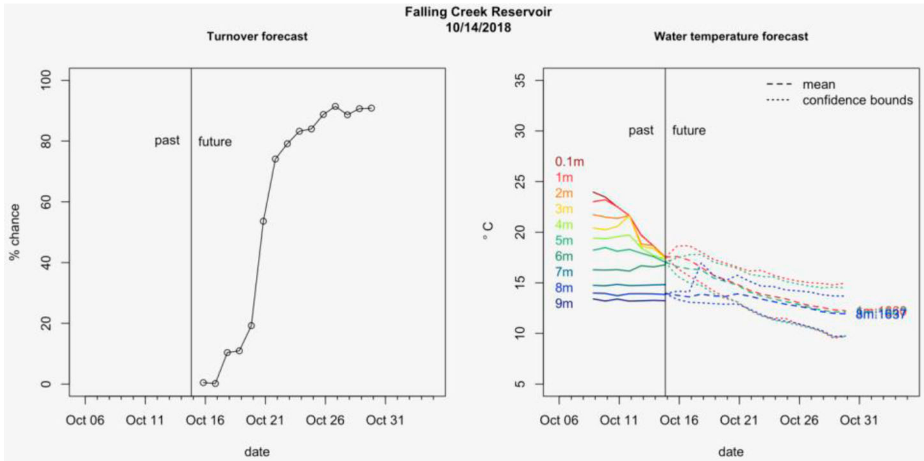

Figure 2. Iterative feedback from managers was used to co-design this forecast visualization, which was emailed daily as a decision support tool (this example was generated on 14 Oct 2018). The left panel shows the turnover forecast, or likelihood of fall turnover occurring during the next 16 days (% chance calculated across all ensembles). The right panel shows the observed sensor water temperatures during the past week, and mean forecasts of future water temperature at 3 depths with their 95% confidence intervals. Color version available online.

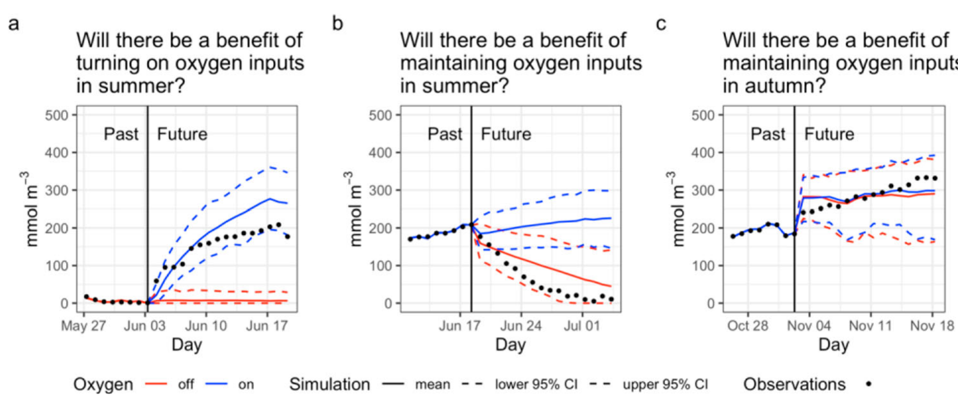

Figure 3. Scenario-based FLARE forecasts can assist management and operation of the HOx (hypolimnetic oxygenation) system. Three example cases from 2019 are shown: (a) a forecast for when the reservoir was anoxic and the HOx system was off prior to the forecast, and (b and c) the reservoir was oxic and the HOx system was on prior to the forecast on 2 different time periods. The red line shows a scenario in which the HOx system is deactivated for 16 days, and the blue line shows a scenario in which the HOx system is activated and remains on for the next 16 days. The solid line is the mean of 441 ensembles, and the dashed lines are the upper and lower 95% confidence intervals of the forecast. The black points show the oxygen observations that actually occurred for that time period. Color version available online.

fact turned on that day, and daily observations of DO concentration closely follow the forecast (black points).

Second, by comparison, the HOx system has been operational for 2 weeks at the time of the DO forecast on 18 June 2019 (Fig. 3b). If the HOx system remains activated at current levels, the hypolimnion will remain oxic (blue line); if the HOx system is turned off, the hypolimnion will go anoxic within the next 16 days (red line). The HOx system was actually turned off that day, shown by the daily DO concentrations (represented by the black points).

Third, the HOx system has been operational at the time of the forecast on 2 November 2019, but forecasts of either scenario (HOx remaining on or turned off) indicate that adding more oxygen to the hypolimnion will have a minimal effect, likely because fall turnover has already happened and the water column is fully mixed (Fig. 3c). These scenario-based forecasts highlight that the same decision (e.g., activating the HOx system) could result in markedly different consequences for DO concentrations depending on the time of year and initial conditions at the time of the forecast.

Discussion

Generating forecasts that show the 16-day outcome of 2 different management choices—activating or deactivating the oxygenation system—is a powerful approach for

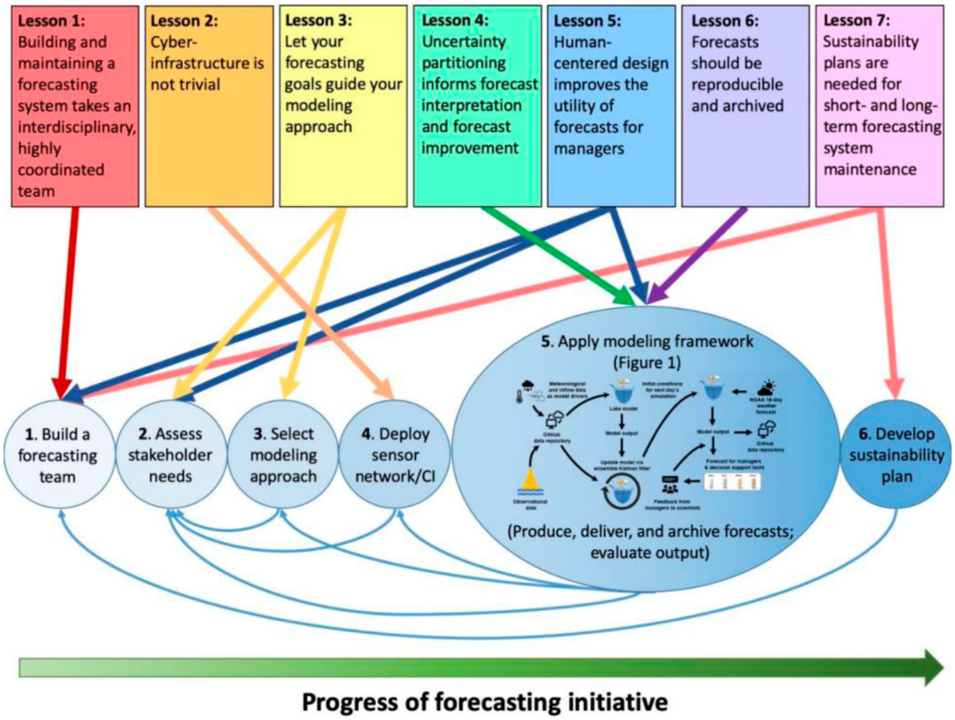

Figure 4. A conceptual framework for developing and implementing iterative, near-term water quality forecasting for lake and reservoir management. Each of the lessons learned (rainbow-colored boxes) described in the text maps onto the stages of a forecasting initiative (blue circles). We note that setting up a forecasting system is not a linear process, and thus it is important to revisit earlier stages throughout the initiative, as denoted by the blue arrows at the bottom of the figure. DO = dissolved oxygen. Color version available online.

assisting managers in their decision-making. Managers are given the tools to "see" the direct effects of implementing their decision on the future water quality of the reservoir. To the best of our knowledge, other freshwater forecasting systems (e.g., Page et al. 2018, Baracchini et al. 2020) provide future information on lake or reservoir variables but still require managers to determine how their decision-making may interact with forecasted conditions to affect future water quality. While the HOx system forecasts are specific to FCR, we envision that many other management interventions (e.g., algaecide addition, manipulation of inflows/outflows, extraction of water at different depths, epilimnetic mixing) could be implemented in forecasting frameworks to examine their potential consequences on water quality.

As a result of our experience running FLARE, we identified 7 major lessons learned that can be used to develop best practices for other teams interested in implementing near-term, iterative ecological forecasting systems for lake and reservoir management (Fig. 4). Each of these lessons learned directly map onto the near-term iterative forecasting framework described earlier for FLARE (highlighted in Fig. 4). We describe each lesson, with details on our approach and helpful issues to consider.

Lesson 1: Building and maintaining a forecasting system takes an interdisciplinary, highly coordinated team

Near-term, iterative water quality forecasting initiatives require bringing together researchers from the physical, natural, social, and computer sciences. Our FLARE research team was divided into sub-teams by their role within the iterative forecasting workflow (Fig. 1): (1) field operations – the freshwater ecologists and environmental engineers who went to the reservoir to deploy, maintain, and troubleshoot sensors; (2) cyberinfrastructure – the computer scientists working remotely to ensure data were being transferred from the sensors to the cloud; (3) modeling – the ecosystem modelers responsible for calibrating the lake model, developing the data assimilation and forecasting algorithms, and ensuring that the forecasts ran; (4) management/decision support – the social scientists and project coordinators working closely with the managers to study their workflows, collect feedback on decision support tools, and respond to questions and requests; and (5) the managers themselves, who provided feedback regarding the forecast variables and visualizations and made decisions regarding the treatment of water from the reservoir. We envision that a similar composition of personnel would be needed for other real-time, iterative water quality forecasting initiatives.

Maintaining a forecasting system operating on a daily scale increases the time urgency of completing tasks on schedule and requires near-constant communication. High task interdependence is common in interdisciplinary teams (National Research Council 2015) but is particularly heightened in forecasting because of the inherent time sensitivity of near-term, iterative, and real-time processes (e.g., Fig. 1). Communication between different FLARE sub-teams varied throughout forecasting system development, setup, and implementation stages. For example, during the forecasting system setup, the field operations and cyberinfrastructure sub-teams communicated at least weekly and sometimes daily, while the modeling team and the field operations team communicated daily to get the model calibrated with field data. As new sensors came online, these 2 sub-teams had to work closely together to update the data assimilation code and workflows. Once the system was operational, the cyberinfrastructure and modeling sub-teams communicated usually on a weekly basis to ensure software updates were implemented, but occasionally daily if the forecasts did not run, as indicated by automated alert email notifications. The modeling sub-team and managers "communicated" daily via the automated daily forecast emails delivered by the FLARE system, which was supplemented by more frequent communication as needed, based on the forecast results and water quality impairment. For example, if managers needed clarification or help with the interpretation of the forecast, they would email the modeling team with their request. These time scales of communication emerged organically over 2 years. While the frequency of communication between teams is inherently dependent on team member composition, we found that fortnightly project meetings were needed at a minimum to ensure operational forecasting and team cohesion.

Lesson 2: Cyberinfrastructure is not trivial

Developing cyberinfrastructure that seamlessly links sensors, software, data repositories, and models to create daily iterative forecasts with data assimilation can be a daunting challenge. FLARE's cyberinfrastructure is geographically distributed, with some components at the "edge" of the network (sensors and sensor gateways), some computational and storage resources in the cloud (Git repositories), and others at remote servers. Because of the physical location of resources (in our case, sensors and gateways at FCR and servers located >1000 km away at the University of Florida), the ability to perform remote management and troubleshooting is critical. Every day, data need to be iteratively acquired, transferred, and acted upon through many stages, including quality assurance/quality control (QA/AC), model runs, generation of visualization outputs, and archiving of forecast products (Fig. 1). Moreover, to avoid dependence on proprietary technology and to enable scaling to other lakes and reservoirs, it is key that all components of the cyberinfrastructure be based on open-source software. Furthermore, the cyberinfrastructure needs to provide acceptable levels of authentication and access control to ensure the security of resources and data and be relatively easy to use by all members of the FLARE team and the reservoir managers.

With these requirements in mind, 5 key components of the FLARE cyberinfrastructure emerge. First, distributed environmental sensors (i.e., the sensors measuring meteorological variables, water temperature, and DO) interface to data loggers for data capture and storage at the reservoir site. The data loggers connect to the sensor gateways, which allow remote management and orchestrate data transfers to the cloud-based Git repository. Data are retrieved from the storage Git repository for execution on cloud computing nodes to drive ensemble runs of the open-source GLM model. In FLARE, the sensor gateways, storage repository, and cloud computing nodes run the Linux operating system and are connected by the open-source IPOP (IP-over-P2P) virtual private network (Subratie et al. 2020). All code orchestrating the FLARE system is written in open-source programing languages (Bash, C, C++, Fortran, and R; Daneshmand et al. 2020, Thomas et al. 2020b).

Our team found that computer science expertise was needed throughout the development and implementation of our forecasting system workflow, hardware, and software. Creating an iterative forecasting cycle running on a daily time step required close collaboration, trust, and a shared understanding of terminology and best practices from both freshwater ecology and computer science (following the recommendations of Carey et al. 2019).

Lesson 3: Let your forecasting goals guide your modeling approach

Freshwater forecasting systems can use multiple modeling approaches that range from simple empirical models to dynamic lake ecosystem models. If the focal lake has a long time series of monitoring data and only one forecast variable is of interest (e.g., surface chlorophyll *a* concentration), empirical models that require extensive data for calibration, such as auto-regressive time series models or machine learning models, may be most appropriate for forecasting. For applications without existing long-term data or that require multiple forecast variables (e.g., chlorophyll *a* and DO concentrations at multiple depths), hydrodynamic models coupled to process-based lake ecosystem models (see Mooij et al. 2010, Hipsey et al. 2015 for an extensive list of options) may be more appropriate. These models (e.g., GLM-AED) can reasonably simulate lake conditions without the need for long monitoring time series, although some basic data (e.g., bathymetry, current water chemistry to initialize the model) will be needed for model setup and parameterization. Forecasting applications in which data assimilation calibrates parameters as more data are collected allow a lake-specific parameterization to emerge, thus improving forecast skill over time (following Thomas et al. 2020a). Coupled hydrodynamic–ecosystem models can forecast numerous variables and may be able to perform better than empirical models when initial conditions are outside the historical envelope (Dietze 2017) as well as be used for forecast scenario development (Rousso et al. 2020).

We chose to use the GLM-AED model for FLARE because we needed to simultaneously forecast lake physics (i.e., fall turnover) and chemistry (i.e., DO concentrations) for managers, but we had to overcome some challenges when implementing this modeling approach. First, the parameterization of GLM-AED can vary by focal variable, so identifying the most important forecast variables a priori is useful for optimizing calibration and overcoming model equifinality (when different parameterizations result in similar answers; Hipsey et al. 2015). Second, not all hydrodynamic–ecosystem models are set up for iterative forecasting and daily data assimilation, which requires running the model one day at a time, adjusting the model based on sensor observations and calculating process uncertainty, and then running the model for the next time step. In our case, our FLARE modeling team worked with the GLM-AED developers to add restart capacity to the GLM-AED model. Third, while the GLM-AED model allows us to model a suite of different water quality variables, not all applications need that level of model complexity, which is associated with greater computational requirements. Other model complexity considerations that will affect computational needs and model run times include model dimension (0-D, 1-D, 2-D, and 3-D), model time step (e.g., minute or hour) and forecasting time horizon (e.g., 16 days).

Each modeling choice has trade-offs; thus, we followed the recommendations of Hipsey et al. (2015), who advocate first determining the main processes that affect the primary water quality variables relevant for managers, and then choosing the simplest model possible to reasonably simulate those variables. In our particular application, we determined from consultation with the managers that water temperature and DO forecasts would be useful for their decision-making (see Lesson 5). Forecasting water temperature and DO depth profiles in a shallow reservoir necessitated a hydrodynamic model that simulated the effects of meteorology and inflows as drivers of thermal stratification and subsequent DO availability in the water column. Because GLM-AED was able to successfully recreate thermal dynamics in FCR (Thomas et al. 2020a), we did not need to use a 3-D model, which may be required for simulating waterbodies with more complex bathymetries. Finally, we ran GLM-AED on an hourly time step for a 16-day forecast horizon, which minimized computational load but still allowed us to capture sub-daily variability in thermal stratification. The 16-day forecast horizon was needed in our application because it usually takes multiple days to see an effect of activating/deactivating the HOx system.

Lesson 4: Uncertainty partitioning informs forecast interpretation and forecast improvement

Quantifying the uncertainty in water quality forecasts may improve their usability for stakeholders. Specifically, knowing the likelihood of different forecasted water quality impairment events (e.g., hypolimnetic anoxia) may improve managers' ability to interpret their risk. For example, a study of managers receiving flood forecasts revealed that the majority thought that "adding uncertainties to forecasts bring[s] useful operational information" (Berthet et al. 2016). When no uncertainties are provided with forecasts, many stakeholders will infer their own estimate of uncertainty for decision-making, which may be biased and less accurate than quantified uncertainties from forecast models (Morss et al. 2008, Berthet et al. 2016). Because little is currently known about manager interpretation of water quality forecast uncertainty, studying how managers respond to different levels of uncertainty will help improve forecast visualization and decision support tools.

Our FLARE team found that partitioning the different sources of uncertainty in our water quality forecasts was useful for prioritizing areas for further improvement of the forecasting system. Specifically, in the FLARE forecasting system, meteorological driver data have been the dominant contributor to total forecast uncertainty at 7–16-day forecast horizons, contributing more uncertainty than other sources, such as process or parameters (Thomas et al. 2020a). As a result, we are confident we have chosen an appropriate model for our application (as indicated by relatively low process uncertainty) and that this model is adequately parameterized (as indicated by relatively low parameter uncertainty). Thus, to improve the forecasting system, we should focus our efforts on improving the statistical downscaling of NOAA meteorological driver data for our study (as indicated by relatively high driver data uncertainty). Uncertainty partitioning can also be used to compare among model structures if several different models are being considered for a forecasting application (Dietze 2017) and can provide a degree of transparency to stakeholders as to the drivers of forecast uncertainty. While partitioning uncertainty among different sources is generally not common in forecasting studies (Cressie et al. 2009, Raiho et al. 2020), it has immense potential for ecological forecasters who seek a robust assessment method for forecast improvement and to improve interpretation of forecast uncertainty by managers and other stakeholders.

Lesson 5: Human-centered design improves the utility of forecasts for managers

A major goal of human-centered design (HCD; Table 1) is developing a product that can be integrated into the workflows of its end users. Stakeholders are often reticent to integrating new decision support tools in already-developed workflows (Callahan et al. 1999, Pagano et al. 2001), so we emphasize the importance of HCD methods in every step of forecast development. Although some forecasters may be hesitant to release forecasts to stakeholders while the forecast system is still in active development, an early introduction to forecasts could enable stakeholder input and increase stakeholder willingness to use forecasts for decision-making (Hobday et al. 2019).

FLARE's HCD consisted of 3 stages. The first stage was focused on eliciting managers' water quality concerns in multiple discussions to guide the choice of the focal forecast variables. For example, these meetings identified that destratification events (e.g., fall turnover) pose major management concerns because of the mixing of high concentrations of nutrients and metals throughout the water column. This information led our team to choose water temperature as a focal variable for initial forecasts. In addition, oxygen management emerged as another water quality concern, as demonstrated by the water utility's investment in an HOx system and questions about operational practices that maximized its benefit while minimizing costs. The discussions were complemented with visual observation of the water treatment plant operations in a second stage of HCD to better understand how decision support tools could be integrated into existing workflows. The discussions and treatment plant observations informed forecast delivery, which came via an automated email that was timed to occur at the beginning of each day's morning work shift at the treatment plant. Third, our team iteratively developed the daily forecast output visualizations delivered by the email through multiple rounds of feedback from managers, resulting in a "final" visualization design that provided information on the likelihood of fall turnover (Fig. 2). We found that providing multiple examples of forecast visualizations was more effective for eliciting feedback than asking the managers what they wanted in a forecast. Our visualization represents uncertainty in forecasted turnover as a percent chance (akin to how weather forecasts communicate percent chance of precipitation), which is calculated from the number of ensemble members that predict turnover divided by the total number of ensemble members ($n = 441$; Fig. 2). As noted earlier, the feedback process is iterative, and our team is currently conducting a new study on the effectiveness of communicating uncertainty to managers.

Lesson 6: Forecasts should be reproducible and archived

It is important to follow findable, accessible, interoperable, and reusable (FAIR; Wilkinson et al. 2016) data standards and methods (Table 1) when developing and publishing forecasts (Harris et al. 2018). Making data and software open source facilitates collaboration among forecast team members and reduces "duplicated efforts" among researchers (Pfenninger et al. 2017). Throughout our forecasting experience, we have followed FAIR guiding principles by making our data, software, and forecasts open source and reproducible. Our forecast code is published with digital object identifiers (DOIs; Daneshmand et al. 2020, Thomas et al. 2020b), and updated versions of the code are publicly available through GitHub (https://github.com/CareyLabVT/FLARE). All driver data are published with metadata in the Environmental Data Initiative repository (https://portal.edirepository.org). Likewise, our forecasts are shared with managers through daily emails and publicly on our website (www.smartreservoir.org). Other forecasters have set up automated pipelines that archive forecasts in the Zenodo repository (e.g., White et al. 2019), which is a promising approach for near-term, iterative forecast publishing.

We note several considerations that should be addressed when making near-term water quality forecasts FAIR. While the scientists and managers on our research team agreed to make our data, code, and forecast products publicly accessible, this objective may not be possible in all cases. Water quality forecasts for drinking water supply lakes and reservoirs may contain sensitive information about water quality conditions, the frequency of measurements, and the treatment process, making the reservoir more vulnerable to potential attacks. For example, forecasts of fish abundance and distribution could contribute to overfishing and should therefore be considered in combination with necessary legal regulations (Hobday et al. 2019). Researchers and managers developing forecasts in drinking reservoirs will therefore need to think carefully about how open their data and forecasts should be. Furthermore, forecasts of water quality within the reservoir may be difficult for the public to interpret and could add to more confusion if not disseminated with appropriate explanation. When choosing the scope of the forecasting exercise, it is important to consider who the forecast benefits, unintended consequences, and conflicts of interest. If forecasters decide not to make forecasts completely FAIR, it may still be possible to make them accessible to researchers and managers using secure authentication and appropriate licensing (Wilkinson et al. 2016).

Lesson 7: Sustainability plans are needed for short- and long-term forecasting system maintenance

Our experience with FLARE has shown that creating a sustainability plan early in a forecasting project is crucial to ensuring a functioning system over multiple time scales. Over the short-term (days to years), it is important to anticipate which components of the system should be prioritized for maintenance or updating. Both hardware (e.g., water quality sensors) and software (e.g., the code that runs the sensor gateways) will likely intermittently fail and need to be replaced or updated through the lifetime of a forecasting system. These events provide an opportunity to evaluate the need for that component. For example, forecast performance might not be affected by the loss of one water temperature sensor, but the loss of one DO sensor could severely affect the system's data assimilation and subsequent forecast quality. Conducting forecasting system "experiments" in which observational variables, model driver data, or steps in the forecasting software are selectively removed provides an important indicator of the forecasting system's robustness and fault tolerance (Hobday et al. 2019). Moreover, this information can serve as a form of sensitivity analysis of the overall system and inform how best to prioritize maintenance. For example, because the meteorological sensors at FCR provide critical hourly resolution driver data to the GLM model, our field operations team prioritizes the maintenance of those sensors over the water temperature sensors, which can easily be supplemented by manually sampling water temperature in the field. Similarly, we have to constantly reevaluate the usability of our existing forecasting software as new, potentially more robust, technology solutions become available.

Over multiple years, a major issue relating to the sustainability of maintaining forecasts is personnel continuity. Depending on the composition of the team of researchers running the forecast system (see Lesson 1), most institutional knowledge about running the forecast system could be concentrated among a few individuals. Consequently, if one of those team members abruptly leaves the project or transitions to a new position, the maintenance of the forecasting system could be in question. This issue can be alleviated by creating a collaboration plan at the launch of a forecasting project that clearly specifies the roles of each team member, how forecasting operation knowledge is shared and documented (ideally in standard operating procedure [SOP] documents accessible to all team members), and expectations for any team members leaving the project (Bennett et al. 2018). Our team found that building an editable wiki website that allows collaborative editing and serves as an up-to-date record on the sensor network components (https://github.com/CareyLabVT/SmartConnectedCommunities/wiki) was critical for knowledge transfer among team members.

Finally, over longer time scales (years to decades), sustainability of forecast delivery is critical because it has many implications for the research team and stakeholders. If, for example, funding to support the maintenance of a forecast system ends, how will that affect water managers that depend on the forecasts for decision-making? Ultimately, transitioning the forecast operations to key stakeholder groups may be an option in some cases, using a model similar to the meteorological forecasts operationalized by different countries' weather services (Hobday et al. 2019). Operationalizing forecasts by governmental agencies and utilities will ensure the long-term continuity of the system, but also has challenges (e.g., Brown et al. 2013, Dietze et al. 2018). For example, a forecasting system may need modification before it is transferred from a team of academic researchers to a governmental agency or water utility that likely has different goals and organizational values (Dietze et al. 2018). Further, agency or utility staff may not have the interdisciplinary expertise needed to maintain a forecasting system (see Lesson 1), potentially necessitating a long-term partnership with system developers (rather than a simple hand-off of infrastructure). We are working to generalize and scale FLARE beyond FCR, with the aim of operationalizing the forecasting system for other lakes and reservoirs over the next 3–5 years. The goal of this scaling effort is to improve the sustainability of our open-source forecasting system by demonstrating its utility across a range of different waterbodies, thereby encouraging its adoption by other water managers.

Conclusions

The application of real-time, near-term, iterative forecasting for lake and reservoir management is in its nascent stages, providing an exciting opportunity for this research community to make great progress in developing and running freshwater forecasting systems. While our experience and others highlight the challenges of freshwater forecasting, our goal in sharing our lessons learned is to assist new research teams as they begin this endeavor. Given the increased variability facing many freshwater ecosystems, ecological forecasting has high potential for improving preemptive management and minimizing water treatment costs. We also note that developing a real-time, near-term forecasting system may not be feasible for all managers. When deciding if a forecasting system would be useful, we advocate examining the current financial costs of water quality management and calculating if the cost savings gained from preventative measures triggered by forecasts offsets the costs of sensors and personnel. If forecasting is a potential option for your management application, we echo Dietze (2017); because of the iterative nature of near-term forecasting, in which data assimilation will improve models and forecast performance over time, there is no better time than the present to get started forecasting lake and reservoir water quality.

Acknowledgements

We thank the entire FLARE and Smart Reservoir project team for their many contributions that enabled this project, which was catalyzed by our long-term collaboration with managers at the Western Virginia Water Authority. We also thank colleagues in the Ecological Forecasting Initiative (EFI) for thoughtful feedback that improved this work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was financially supported by the Western Virginia Water Authority and U.S. National Science Foundation grants CNS-1737424, DEB-1753639, DEB-1926050, and DEB-1926388, DBI-1933016, DBI-1933102.

ORCID

Cayelan C. Carey http://orcid.org/0000-0001-8835-4476
Whitney M. Woelmer http://orcid.org/0000-0001-5147-3877
Mary E. Lofton http://orcid.org/0000-0003-3270-1330
Renato J. Figueiredo http://orcid.org/0000-0001-9841-6060
Bethany J. Bookout http://orcid.org/0000-0001-8394-926X
Rachel S. Corrigan http://orcid.org/0000-0001-6101-8085
Vahid Daneshmand http://orcid.org/0000-0003-4181-1806
Alexandria G. Hounshell http://orcid.org/0000-0003-1616-9399
Dexter W. Howard http://orcid.org/0000-0002-6118-2149
Abigail S. L. Lewis http://orcid.org/0000-0001-9933-4542
Ryan P. McClure http://orcid.org/0000-0001-6370-3852
Heather L. Wander http://orcid.org/0000-0002-3762-6045
Nicole K. Ward http://orcid.org/0000-0001-7549-0153
R. Quinn Thomas http://orcid.org/0000-0003-1282-7825

References

Anderson EJ, Schwab DJ, Lang GA. 2010. Real-time hydraulic and hydrodynamic model of the St. Clair River, Lake St. Clair, Detroit river system. J Hydraul Eng. 136(8):507–518.

Bal G, Rivot E, Baglinière J-L, White J, Prévost E. 2014. A hierarchical Bayesian model to quantify uncertainty of stream water temperature forecasts. PLoS One 9(12): e115659.

Baracchini T, Wüest A, Bouffard D. 2020. Meteolakes: an operational online three-dimensional forecasting platform for lake hydrodynamics. Water Res. 172:115529.

Bennett LM, Gadlin H, Marchand C. 2018. Collaboration and team science: a field guide. 2nd ed. Washington (DC): National Institutes of Health.

Berthet L, Piotte O, Gaume E, Marty O, Ardilouze C. 2016. Operational forecast uncertainty assessment for better information to stakeholders and crisis managers. 3rd Euro Conf Flood Risk Manag. 7:18005. doi:10.1051/e3sconf/20160718005.

Brookes JD, Carey CC, Hamilton DP, Ho L, van der Linden L, Renner R, Rigosi A. 2014. Emerging challenges for the drinking water industry. Environ Sci Technol. 48:2099–2101.

Brown CW, Hood RR, Long W, Jacobs J, Ramers DL, Wazniak C, Xu J. 2013. Ecological forecasting in Chesapeake Bay: using a mechanistic-empirical modeling approach. J Marine Syst. 125:113–125.

Caissie D, Thistle ME, Benyahya L. 2017. River temperature forecasting: case study for Little Southwest Miramichi River (New Brunswick, Canada). Hydrol Sci J. 62:683–697.

Callahan B, Edward M, Fluharty D. 1999. Policy implications of climate forecasts for water resources management in the Pacific Northwest. Policy Sci. 32(3):269–293.

Carey CC, Bookout BJ, Lofton ME, McClure RP. 2020a. Time series of high-frequency meteorological data at Falling Creek Reservoir, Virginia, USA 2015-2019. Environmental Data Initiative; [accessed 2020 Feb 11]. https://doi.org/10.6073/pasta/ea47ae493c7025d61245287649895e60

Carey CC, Bookout BJ, Woelmer WM, Lewis ASL. 2020b. Time series of high-frequency sensor data measuring water temperature, dissolved oxygen, conductivity, specific conductivity, total dissolved solids, chlorophyll *a*, phycocyanin, and fluorescent dissolved organic matter at discrete depths in Falling Creek Reservoir, Virginia, USA in 2018–2019. Environmental Data Initiative; [accessed 2020 Feb 11]. https://doi.org/10.6073/pasta/b888ac006ef4ca601f63e2703d7476b9

Carey CC, Gerling AB, McClure RP, Lofton ME, Bookout BJ, Corrigan RS, Hounshell AG, Woelmer WM. 2020c. Discharge time series for the primary inflow tributary entering Falling Creek Reservoir, Vinton, Virginia, USA 2013–2020. Environmental Data Initiative; [accessed 2020 Feb 11]. https://doi.org/10.6073/pasta/417b4e7c4f304eb84a210211211a2a28

Carey CC, Ward NK, Farrell KJ, Lofton ME, Krinos AI, McClure RP, Subratie KC, Figueiredo RJ, Doubek JP, Hanson PC, et al. 2019. Enhancing collaboration between ecologists and computer scientists: lessons learned and recommendations forward. Ecosphere. 10(5):e02753.

Clark JS, Carpenter SR, Barber M, Collins S, Dobson A, Foley JA, Lodge DM, Pascual M, Pielke R Jr, Pizer W, et al. 2001. Ecological forecasts: an emerging imperative. Science. 293:657–660.

Cressie N, Calder CA, Clark JS, Ver Hoef JM, Wikle CK. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecol App. 19:553–570.

Daneshmand V, Thomas RQ, Bookout BJ, Carey CC, Figueiredo RJ. 2020. Sensor gateway code for Forecasting Lake and Reservoir Ecosystems (FLARE), Version v.1.0. Zenodo. doi:10.5281/zenodo.3862907

Dietze MC. 2017. Ecological forecasting. Princeton (NJ): Princeton University Press.

Dietze MC, Fox A, Beck-Johnson LM, Betancourt JL, Hooten MB, Jarnevich CS, Keitt TH, Kenney MA, Laney CM, Larsen LG, et al. 2018. Iterative near-term ecological forecasting: needs, opportunities, and challenges. P Natl Acad Sci. 115(7):1424–1432.

Dugan NR, Smith SJ, Sanan TT. 2018. Impacts of potassium permanganate and powdered activated carbon on cyanotoxin release. J Am Water Works As. 110:E31–E42.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res. 99:143–162.

Evensen G. 2009. Data assimilation: the ensemble Kalman filter. Berlin: Springer-Verlag.

Gerling AB, Browne RG, Gantzer PA, Mobley MH, Little JC, Carey CC. 2014. First report of the successful operation of a side stream supersaturation hypolimnetic oxygenation system in a eutrophic, shallow reservoir. Water Res. 67:129–143.

Gerling AB, Munger ZW, Doubek JP, Hamre KD, Gantzer PA, Little JC, Carey CC. 2016. Whole-catchment manipulations of internal and external loading reveal the sensitivity of a century-old reservoir to hypoxia. Ecosystems. 19:555–571.

Hague MJ, Patterson DA. 2014. Evaluation of statistical river temperature forecast models for fisheries management. N Am J Fish Manage. 34:132–146.

Harris DJ, Taylor SD, White EP. 2018. Forecasting biodiversity in breeding birds using best practices. PeerJ. 6:e4278.

Hipsey MR, Bruce LC, Hamilton DP. 2013. Aquatic ecodynamics (AED) model library science manual, version 4. Perth, Australia: University of Western Australia.

Hipsey MR, Bruce LC, Boon C, Busch B, Carey CC, Hamilton DP, Hanson PC, Read JS, de Sousa E, Weber M, 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). Geosci Model Dev. 12:473–523.

Hipsey MR, Hamilton DP, Hanson PC, Carey CC, Coletti JZ, Read JS, Ibelings BW, Valesini FJ, Brookes JD. 2015. Predicting the resilience and recovery of aquatic systems: a framework for model evolution within environmental observatories. Water Resour Res. 51:7023–7043.

Hobday AJ, Hartog JR, Manderson JP, Mills KE, Oliver MJ, Pershing AJ, Siedlecki S. 2019. Ethical considerations and unanticipated consequences associated with ecological forecasting for marine resources. ICES J Mar Sci. 76:1244–1256.

Huang J, Gao J, Liu J, Zhang Y. 2013. State and parameter update of a hydrodynamic-phytoplankton model using ensemble Kalman filter. Ecol Model. 263:81–91.

Jenny JP, Francus P, Normandeau A, Lapointe F, Perga ME, Ojala A, Schimmelmann A, Zolitschka B. 2016. Global spread of hypoxia in freshwater ecosystems during the last three centuries is caused by rising local human pressure. Global Change Biol. 22:1481–1489.

Jiménez Cisneros BE, Oki T, Arnell NW, Benito G, Cogley JG, Doll P, Jiang T, Mwakalila SS, et al. 2014. Freshwater resources. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, editors. Climate change 2014: impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change. New York (NY): Cambridge University Press; p. 229–269.

Jolliffe IT, Stephenson DB. 2003. Forecast verification. A practitioner's guide in atmospheric science. Hoboken (NJ): John Wiley & Sons.

Khan SJ, Deere D, Leusch FDL, Humpage A, Jenkins M, Cunliffe D. 2015. Extreme weather events: should drinking water quality management systems adapt to changing risk profiles? Water Res. 85:124–136.

Kim DK, Jeong KS, McKay RIB, Chon TS, Joo GJ. 2012. Machine learning for predictive management: short and long term prediction of phytoplankton biomass using genetic algorithm based recurrent neural networks. Int J Environ Res. 6(1):95–108.

Kim K, Park M, Min JH, Ryu I, Kang MR, Park LJ. 2014. Simulation of algal bloom dynamics in a river with the ensemble Kalman filter. J Hydrol. 519:2810–2821.

Kurosu M, editor. 2011. Human centered design. Orlando (FL): Springer.

Luo Y, Ogle K, Tucker C, Fei S, Gao C, LaDeau S, Clark JS, Schimel DS. 2011. Ecological forecasting and data assimilation in a data-rich era. Ecol Appl. 21:1429–1442.

McClure RP, Hamre KD, Niederlehner BR, Munger ZW, Chen S, Lofton ME, Schreiber ME, Carey CC. 2018. Metalimnetic oxygen minima alter the vertical profiles of carbon dioxide and methane in a managed freshwater reservoir. Sci Total Environ. 636:610–620.

Mooij WM, Trolle D, Jeppesen E, Arhonditsis G, Belolipetsky PV, Chitamwebwa DBR, Degermendzhy AG, DeAngelis DL, De Senerpont Domis LN, Downing AS, et al. 2010. Challenges and opportunities for integrating lake ecosystem modelling approaches. Aquat Ecol. 44:633–667.

Morss RE, Demuth JL, Lazo JK. 2008. Communicating uncertainty in weather forecasts: a survey of the U.S. public. Weather Forecast. 23:974–991.

Munger ZW, Carey CC, Gerling AB, Hamre KD, Doubek JP, Klepatzki SD, McClure RP, Schreiber ME. 2016. Effectiveness of hypolimnetic oxygenation for preventing accumulation of Fe and Mn in a drinking water reservoir. Water Res. 106:1–14.

Munger ZW, Carey CC, Gerling AB, Doubek JP, Hamre KD, McClure RP, Schreiber ME. 2019. Oxygenation and hydrologic controls on iron and manganese mass budgets in a drinking-water reservoir. Lake Reserv Manage. 35:277–291.

National Research Council. 2015. Enhancing the effectiveness of team science. Cooke NJ, Hilton ML, editors. Washington (DC): National Academies Press.

Ouellet-Proulx S, St-Hilaire A, Boucher MA. 2017. Water temperature ensemble forecasts: implementation using the CEQUEAU model on two contrasted river systems. Water (Basel). 9(7):457.

Pagano TC, Hartmann HC, Sorooshian S. 2001. Using climate forecasts for water management: Arizona and the 1997-1998 El Niño. J Am Water Resour Ass. 37(5):1139–1153.

Page T, Smith PJ, Beven KJ, Jones ID, Elliott JA, Maberly SC, Mackay EB, De Ville M, Feuchtmayr H. 2018. Adaptive forecasting of phytoplankton communities. Water Res. 134:74–85.

Perello MM, Kane DD, Golnick P, Hughes MC, Thomas MA, Conroy JD. 2017. Effects of local weather variation on water-column stratification and hypoxia in the western, Sandusky, and central basins of Lake Erie. Water (Basel). 9:279.

Pfenninger S, DeCarolis J, Hirth L, Quoilin S, Staffell I. 2017. The importance of open data and software: Is energy research lagging behind? Energy Policy. 101:211–215.

Pike A, Danner E, Boughton D, Melton F, Nemani R, Rajagopalan B, Lindley S. 2013. Forecasting river temperatures in real time using a stochastic dynamics approach. Water Resour Res. 49:5168–5182.

Prein AF, Rasmussen RM, Ikeda K, Liu C, Clark MP, Holland GJ. 2017. The future intensification of hourly precipitation extremes. Nat Clim Change. 7:48–52.

Raiho A, Dietze M, Dawson A, Rollinson CR, Tipton J, McLachlan J. 2020. Determinants of predictability in multi-decadal forest community and carbon dynamics. bioRxiv 2020.05.05.079871.

Recknagel F, Adrian R, Köhler J, Cao H. 2016. Threshold quantification and short-term forecasting of *Anabaena, Aphanizomenon* and *Microcystis* in the polymictic eutrophic Lake Müggelsee (Germany) by inferential modelling using the hybrid evolutionary algorithm HEA. Hydrobiologia. 778:61–74.

Recknagel F, Ostrovsky I, Cao H. 2014. Model ensemble for the simulation of plankton community dynamics of lake Kinneret (Israel) induced from in situ predictor variables by evolutionary computation. Environ Model Softw. 61:380–392.

Rousso BZ, Bertone E, Stewart R, Hamilton DP. 2020. A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes. Water Res. 182:115959.

Rowe MD, Anderson EJ, Wynne TT, Stumpf RP, Fanslow DL, Kijanka K, Vanderploeg HA, Strickler JR, Davis TW. 2016. Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. J Geophys Res-Ocean. 121:5296–5314.

Subratie K, Aditya S, Daneshmand V, Ichikawa K, Figueiredo R. 2020. On the design and implementation of IP-over-P2P overlay virtual private networks. IEICE Trans Comm. E103.B(1):2–10.

Thomas RQ, Figueiredo RJ, Daneshmand V, Bookout BJ, Puckett LK, Carey CC. 2020a. A near-term iterative forecasting system successfully predicts reservoir hydrodynamics and partitions uncertainty in real time. Water Resour Res. 56:e2019WR026138.

Thomas RQ, Figueiredo RJ, Daneshmand V, Puckett LK, Carey CC. 2020b. Forecasting Lake and Reservoir Ecosystems (FLARE), Version v.1.11. Zenodo. doi:10.5281/zenodo.3862905

Tomlinson A, Drikas M, Brookes JD. 2016. The role of phytoplankton as pre-cursors for disinfection by-product formation upon chlorination. Water Res. 102:229–240.

Weber M, Rinke K, Hipsey MR, Boehrer B. 2017. Optimizing withdrawal from drinking water reservoirs to reduce downstream temperature pollution and reservoir hypoxia. J Environ Manage. 197:96–105.

White EP, Yenni GM, Taylor SD, Christensen EM, Bledsoe EK, Simonis JL, Ernest SM. 2019. Developing an automated iterative near-term forecasting system for an ecological study. Methods Ecol Evol. 10(3):332–344.

Wilkinson MD, Dumontier M, Aalsbersberg IJ, Appelton G, Baak A, Blomberg N, Boiten JW, Bonino da Silva Santos LO, Bourne PE, Bouwman J, et al. 2016. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 3(160018):1–9.

Xiao X, He J, Huang H, Miller TR, Christakos G, Reichwaldt ES, Ghadouani A, Lin S, Xu X, Shi J. 2017. A novel single-parameter approach for forecasting algal blooms. Water Res. 108:222–231.

Xie Z, Lou I, Ung WK, Mok KM. 2012. Freshwater algal bloom prediction by support vector machine in Macau storage reservoirs. Math Probl Eng. 2012:397473.