Intro paragraph

As lake and reservoir ecosystems continue to exhibit unprecedented change, ecological forecasting provides a unique opportunity to preemptively manage freshwater ecosystems. Because lakes and reservoirs are more commonly being pushed outside the envelope of historical conditions (i.e., increased summer surface water temperatures, O’Reilly et al. 2015), being able to forecast how they will respond is a novel and critical tool for managers. Ecological forecasting is a new and currently developing field which aims to develop forecasts of “the state of ecosystems, ecosystem services, and natural capital, with fully specified uncertainties” (Clark et al. 2001) (see Box 1 for further definitions). Forecasts provide managers with an estimate of how their system may respond to current or future conditions, and allow them to take preemptive management actions to adapt to or prevent a water quality impairment. Producing forecasts which are in the near-term, are iterative, and which quantify uncertainty in their output adds to their utility as both decision-support tools for managers and in advancing the science of ecological forecasting. The near-term, iterative (Box 1) nature of ecological forecasts allows for continued improvement of forecasting techniques, as forecast output is continually compared to observed conditions, and fed back into the forecasting workflow to improve model output. Quantified uncertainty (Box 1) is also a critical component of ecological forecasts as it both allows forecasters to learn what contributes to forecast uncertainty and improve upon their forecasts and for allowing managers to make more informed decisions based on the likelihood of a forecasted event. For example, if an algal bloom is forecasted for next week at a 60% chance, managers may be more likely than if an algal bloom is only 15% likely. These components of an ecological forecast are what give it unique appeal as a decision-support tool for managers.

**Define and make relevant to lakes and reservoirs**

**Box 1: Glossary of forecasting terms in lakes and reservoirs**

Data assimilation – aka data-model fusion

Ensemble—a set of forecast outputs which propagate alternative competing hypotheses based on probabilities of uncertainty sources (e.g., producing forecasts using multiple model structures due to uncertainty in model selection)

Forecast—probabilistic estimate of a future state of an ecosystem or ecosystem service with fully specified uncertainty (e.g., water temperature one week in the future) and which can be verified for accuracy

Human-centered design—an approach to system design that includes the human user in all aspects of system development, also referred to as human-in-the-loop

Iterative—a crucial part of the forecasting process, the iterative nature of an ecological forecast means that forecast output is both created on a regular, repeated interval and validated with observed conditions in order to assimilate observed conditions and forecasted output into future iterations

Near-term—in the near future (days to months), allowing for quick, iterative validation of accuracy

Prediction—probabilistic forecast based on current trends and conditions (e.g., next week’s chlorophyll-a concentration based on current conditions)

Projection—probabilistic forecast based on an explicit scenario (e.g., Dissolved oxygen concentration in response to implementation of a hypolimnetic oxygenation system)

Quantified uncertainty—propagation throughout the forecast workflow and quantification of the different sources of uncertainty (parameter, driver, process, etc.) within a forecast output to provide accurate decision support and attempt to minimize future uncertainty within forecast output

FAIR—Data principles which require data to be findable, accessible, interoperable, and reusable (Wilkinson et al. 2016). Best practice within forecast is to publish data which follows these FAIR principles.

|  |  |  |
| --- | --- | --- |
| **Box 1. Glossary of forecasting terms in lakes and reservoirs** | | |
| Term | Definition | Example |
| Data assimilation | The regular addition of new data within the forecasting workflow to improve forecast accuracy | Regularly incorporating the most recent observation of chlorophyll-a data into a training dataset which is used to calculate a forecast |
| Ensemble | A set of forecast outputs which propagate alternative competing hypotheses based on probabilities of uncertainty sources | Producing forecasts using multiple model structures (e.g., multiple regression models with different covariates) due to uncertainty in model selection |
| FAIR | Data principles which require data to be findable, accessible, interoperable, and reusable (Wilkinson et al. 2016). Best practice within forecasting is to publish data which follows these FAIR principles. | Publishing forecast driver data within a repository that follows FAIR principles, such as the Environmental Data Initiative |
| Human-centered design | An approach to system design that includes the human user in all aspects of system development, also referred to as human-in-the-loop | Including managers, as the forecast end-user, throughout the forecast development (e.g., consulting managers in deciding which water quality variables are most important for their management regime) |
| Iterative | A crucial part of the forecasting process, the iterative nature of an ecological forecast means that forecast output is both created on a regular, repeated interval and validated with observed conditions in order to assimilate observed conditions and forecasted output into future iterations | Producing chlorophyll-a forecasts on a regular, daily timestep, and comparing forecast output with observed conditions to improve forecast skill |
| Near-term | Pertaining to the near future (days to months), allowing for quick, iterative validation of accuracy | A forecast of water temperature 2 days in the future |
| Prediction | probabilistic forecast based on current trends and conditions | A forecast of next week’s chlorophyll-a concentration based on current conditions |
| Projection | probabilistic forecast based on an explicit scenario | A forecast of dissolved oxygen concentration in response to activation of a hypolimnetic oxygenation system |
| Quantified uncertainty | Propagation throughout the forecast workflow and quantification of the different sources of uncertainty (parameter, driver, process, etc.) within a forecast output to provide accurate decision support and attempt to minimize future uncertainty within forecast output | Uncertainty as to the true values of parameters within model structure |

Lesson Learned Y: Developing forecast decision-support tools that are usable for managers

Understanding decision-making processes of managers is critical for developing forecast output that is effectively integrated into management workflows. While the science of producing accurate ecological forecasts is rapidly developing, the utility of forecast output as a management tool is a crucial measure of forecast success, as well. Further, it is common for stakeholders to be reticent to the integration of new decision support tools into their already developed workflows (Callahan et al. 1999, Pagano et al. 2001), so we note the importance of co-developing tools along with stakeholders. As the Smart Reservoir project was developing a method for forecast dissemination, managers were invited to join the forecasting team in receiving a daily email of nowcasts of water quality variables, as well as water temperature and turnover forecasts. We found that sending the emails at 7 A.M. daily was most useful to the managers because this is when they were online checking the status of Falling Creek Reservoir and ready to make management decisions.

An important aspect to consider in forecast dissemination is when to release forecast output to managers or the general public. Especially while forecasts are being actively developed as research products, many forecasters are hesitant to release forecast output as they are still considered research products, and not thoroughly vetted as decision-support tools to be used to make management decisions. However, as noted above, stakeholders can be reticent to incorporate new tools into their workflow, and so exposure to forecasts could increase willingness to use them. In order to garner awareness and support for the forecasting project and to incorporate aspects of FAIR data principles (See Lesson Learned XX), the Smart Reservoir project publishes all forecast output on our public Github Repository (link), while explicitly labeling forecasts as experimental research products. As a way to build trust and awareness with the managers of Falling Creek Reservoir, forecasts are also disseminated daily to both the Smart Reservoir team and managers, as mentioned above. Because the ideal application of the forecast output is to be used as a decision-support tool, we found that integration of the daily emails into the workflow of the managers helped to increase awareness and familiarity with forecast output.

Finally, producing decision-support tools of forecasts which effectively represent their inherent uncertainty presents a uniquely challenging opportunity. Despite the fact that probabilistic uncertainty is a notoriously difficult concept for most individuals to grasp, as well as to represent graphically (Bonneau et al., in *Scientific Visualization* 2014, Potter et al. 2012), forecasters have an ethical obligation to properly represent uncertainty in their forecast output (Hobday et al. 2019). However, many studies show that different representations of uncertainty result in differential comprehension by users (Kinkeldy et al 2017, Mckenzie et al 2016). For example, using summary statistics such as a mean and a confidence interval to represent ensemble forecast output may yield different comprehension and resulting decision making than portraying forecast output with all ensemble members showing the entire range of uncertainty. Further, different types of forecasts may require different representations of uncertainty meaning that considering how forecast output will be used by managers may influence how uncertainty should be portrayed. For example, for high impact events such as an algal bloom or turnover event it may be more appropriate to disseminate forecast output as the probability of a ‘yes’ or ‘no’ event rather than providing a continuous value with confidence intervals, which may be more appropriate for forecasts of dissolved oxygen concentrations. In all of these scenarios, it is crucial to include managers in the co-development of forecast decision-support tools to better understand how they comprehend and interpret the inherent uncertainty of forecast output.