**Methods**

*Study Site*Falling Creek Reservoir is a small (), eutrophic drinking water reservoir serving the town on Vinton, VA. It is owned and operated by the Western Virginia Water Authority and has been monitored regularly by Virginia Tech since 2013.

*Historical Weekly Dataset and Model Development*The weekly training dataset covered four years (2013-2016) and was developed from weekly measurements of chlorophyll-a, meteorological variables, as well as discharge at the major inflow to the reservoir. These were all chosen as covariates because they are either available as forecasts from National Oceanic and Atmospheric Administration (NOAA) or able to be forecasted using simple linear regression models (e.g., discharge). Chlorophyll-a was estimated by taking weekly profiles using a CTD (SeaBird SERIAL NUMBER). We measured discharge using a pressure transducer at a weir installed at the major inflow to FCR. Flow was measured every 15 minutes, and discharge to the reservoir was calculated, as described in Gerling et al. (2014). A meteorological station measuring X, Y, Z collected data every ten minutes using THESE INSTRUMENTS (serial numbers) from YYYY-present. Any weeks were data was missing (n=XX) were linearly interpolated. Variables that did not follow a normal distribution were transformed to meet the assumptions of a linear model. An autoregressive lag of chlorophyll-a was chosen at one timestep (i.e. one week) and was determined using the package ‘asta’ in R (some citation).

The weekly training dataset was limited to May to October, as this is when sampling consistently occurred on a weekly basis. Using the training dataset described above, we fit multiple linear regression models using the function ‘dredge’ in the package ‘MuMIn’, and selected the best and most parsimonious model using AICc. Our weekly autoregressive model is as follows

Chl-a(t) = β1 + β2Chl-a(t-1) + β3SW mean(t) + β4mean flow(t) + Ɛ (Eq. 1)

Where the response, Chl-a(t), is the chlorophyll-a concentration at the forecasted timestep, t. Chl-a(t-1) is the autoregressive term or chlorophyll-a at the previous timestep, 1 week earlier. SW mean(t) is the mean shortwave on the forecasted timestep. Mean flow(t) is the mean discharge on the forecasted timestep. β1 , β2, β3, and β4 are parameters and Ɛ is an error term.

*Daily Dataset and Model Development*Beginning in August, 2018, an EXO sonde (serial numbers) was installed at the deepest site at FCR, measuring chlorophyll-a fluorescence at 15-minute intervals. Using the same potential covariates as listed above for the weekly model (i.e., meteorological and discharge variables), and following the same model selection protocol, we developed all possible combinations of autoregressive linear models. Because the weekly model was limited to XX number of datapoints in its training dataset, we trained our daily model on the same number of datapoints, which results in a training dataset from August 15, 2018 to December 15, 2018. Our best selected daily model is as follows

Chl-a(t) = β1 + β2Chl-a(t-1) + β3RelHum mean(t) + Ɛ (Eq. 1)

Where the response, Chl-a(t), is the chlorophyll-a concentration at the forecasted timestep, t. Chl-a(t-1) is the autoregressive term or chlorophyll-a at the previous timestep, ‘n’ days earlier. RelHum mean(t) is the mean relative humidity on the forecasted timestep. β1 , β2, and β3 are parameters and Ɛ is an error term.

*Forecasting Framework*Using the linear autoregressive model described above (Eq. 1), the model was adapted to produce forecasts with uncertainty using the Forecasting Lake and Reservoir Ecosystems (FLARE) forecasting framework (Thomas et al. 2020?). Using FLARE, real-time sensor data is automatically uploaded to a data repository (GitHub, link?). Sensor data includes a suite of physical, chemical, and biological data, including chlorophyll-a fluorescence measured using an EXO sonde (serial numbers and company), meteorological variables (At FCR, weather data arewere collected on the minute resolution from a meteorological station (with sensors measuring air temperature, wind speed, relative humidity, shortwave and longwave radiation, and precipitation; see Carey et al. 2019x) and discharge at the major inflow to the reservoir.

Something about the Bayesian framework

Number of ensembles

*Weekly Forecasts*

Using the weekly model developed above, weekly forecasts, providing 1-week ahead and 2-week ahead forecasts, were produced every day from January 01, 2019 to December 22, 2019. The time period from August 15, 2018 to December 15, 2018 (number of datapoints = K) when the driver data was available on a daily basis was used as a spin-up period. While in forecasting mode, new driver data was assimilated weekly on Monday (in order to follow the weekly timestep), and the model was re-fit at each time step under a Bayesian framework to allow the parameter values and model fit to evolve over time. All ensemble members (n=420) and parameter values were saved for archiving at each timestep.

*Daily Forecasts*

Using the daily model developed above, daily forecasts, from 1- to 16-days ahead, were produced from January 01, 2019 to December 22, 2019. To allow a spinup period similar to the weekly model, forecasts from December 15, 2018 to January 01, 2019 were used as a spinup period and not included in final forecast analyses. While in forecasting model, new driver data was assimilated daily when available, and the model was re-fit at each time step under a Bayesian framework to allow the parameter values and model fit to evolve over time. All ensemble members (n=420) and parameter values were saved for archiving at each timestep.

*Forecast Assessment*

We developed a null-persistence model in order to test the robustness of our forecasts. Our null model assumes that the chlorophyll-a concentration at the next timestep will be unchanged from the current timestep, with process error from the Bayesian model added. We calculated an ensemble of null models in order to compare to our ensemble forecasts by sampling 420 times (the number of ensembles) from the distribution of the process error term from the Bayesian model output at each timestep and adding this value to the observed chlorophyll-a concentration (e.g., for the 8-day forecast, the observed chlorophyll-a concentration from 8 days prior is the null model, plus the process error from the model fit, sampled for each model ensemble).

We compared our models using RMSE, NSE, R2, AUC, and CRPS (?).

*Uncertainty Analysis*