Whitney Woelmer, Master’s Prospectus Outline

Overall Introduction

1. Importance of freshwater ecosystems and the many services they provide (setting up why their water quality is so important, as well as what services are the highest priority- e.g., clean water for drinking, etc.)
2. Threats to freshwater systems in a changing world
   1. Changing climate
   2. Land use shifts
   3. Increasing human population
3. Understanding systems and anticipating their response is more important than ever
   1. to allow society to coexist with changing ecosystems
   2. for managers to preemptively manage/anticipate impending poor water quality events
   3. Lake Erie 2014 example, major economic loss

**Proposed Research**

*Chapter 1: Developing near-term forecasts of phytoplankton in drinking water reservoirs***Background/Introduction/Objectives**

* Re-iterate need for forecasting
* Why is ecological forecasting developing now?
  1. Data availability
  2. Changing ecosystems
* What is forecastable
  1. “Forecastable ecosystem attributes are ones for which uncertainty can be reduced to the point where a forecast reports a useful amount of information.” Clark et al 2001
     + Too much uncertainty means not easily forecastable
     + “strong nonlinearity and stochasticity” clark
* Current state of ecological forecasting: mini-lit review. **Papers selected which: 1) use models, 2) quantify uncertainty to make a probabilistic forecast, and 3) run the model outside a given training period. (it is hard to find studies that make *probabilistic* forecasts…maybe this can also be an emphasis of the current state and what is missing, rather than a need to exclude papers that don’t include probabilities/quantified uncertainties)** 
  1. Estes et al 2013: empirical does better than mechanistic at predicting observed productivity and suitability of crops in south Africa
     1. Paper aimed at comparing empirical and mechanistic approaches
     2. Uses four models (2 mechanistic, 2 empirical)
     3. Uncertainty: false positive and true positive estimates?
     4. Tested outside the training period, yes
  2. Lindegren 2013: process-based models to forecast Baltic cod
     1. Models: yes, mechanistic
     2. Uncertainty: uses CI, accounts for climate and process uncertainty in projections
     3. Tested outside training period: yes, with multiple scenarios, which include recent past as validation
  3. Fenocci et al 2019: use process-based (GLM) to hindcast +validate phytos in large lake
     1. Model: yes, GLM-AED
     2. ~~Uncertainty: discuss, but do not quantify parameter uncertainty~~
     3. Tested outside training period:
     4. Table 1 includes more studies
  4. Thuiller et al 2004, Effects of restricting environmental range of data to project current and future species distributions: uses GAM for projections of species distributions
     1. Model: empirical
     2. Uncertainty: yes, probabilistic forecasts
     3. Tested outside training period: yes, calibrated on 70% of data, and projected on 30% of data
  5. Araujo, Thuiller, and Pearson 2006, Climate warming and the decline of amphibians and reptiles in Europe: uses multiple empirical model approaches to develop species distributions (70% of data for calibration, 30% for validation) and then develops projections using climate scenarios
     1. Model: multiple empirical models (GLinearM, GAM, regression trees, ANN)
     2. ~~Uncertainty: discuss model uncertainty which is their reasoning for using so many models to produce so many projections, but do not produce probabilistic forecasts~~
     3. Tested outside training period: yes, 70% calibration, 30% validation, then projections
  6. White and Nemani 2004: mechanistic model framework to forecast soil ecology
     1. Model: mechanistic, TOPS
     2. Uncertainty: yes, They identify meteorological forecast error and the amount of error required to induce statistically different outputs
     3. Tested outside training period: validated from 1982-1997, then ran the model on different input conditions over the same time period (I think?) to see how much they varied from the control (no change in inputs)
  7. Araujo et al 2005, Reducing uncertainty in projections of extinction risk from climate change: uses multiple empirical model approaches to model species ranges and compares to observed and makes projections
     1. Model: empirical models
     2. Uncertainty: sort of? Uncertainty is looked at in terms of the variability over the different projections and is ‘reduced’ through grouping of the various model outputs
     3. Tested outside training period: yes, 70% calibration and 30% validation
     4. Uses ‘ensemble forecasting’ deciding the best forecast based on the likelihood of a number of forecasts
  8. Martinez-Meyer et al 2004, Ecological niches as stable distributional constraints on mammal species , with implications for Pleistocene extinctions and climate change projections for biodiversity: empirical models for current species niches and validation on past data (Pleistocene distributions)
     1. Model: empirical models, GARP
     2. ~~Uncertainty: no?~~
     3. Tested outside training period: yes calibrated using current data and validated during Pleistocene times
  9. Thomas et al 2018, The predictability of a lake phytoplankton community , over time-scales of hours to years:
     1. Model: machine learning, random forests
     2. Uncertainty: sort of? They assess model performance with different parameters and input variables, so the uncertainty between these models is quantified? “Our approach quantifies the decline in predictability with increasing time lag, identifies the predictors that contribute to predictive power and points towards realistic trade-offs and parameterisations through the examination of partial effects.”
     3. Tested outside training period: OOB (out of bag) prediction, so the forests are built with only a subset of the data and can therefore be ‘validated’ using the data not included
     4. “as time lag increased, including environmental predictors led to larger improvements in predictability” over including just the AR term
  10. Perretti et al 2013, Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data: kinda confused, they used ‘mechanistic control models’ to produce a simulated time series?
      1. Model: use several empirical model approaches, but still confused about their mention of using mechanistic models to produce time series? I guess this is just the training/validation data and it isn’t real but that doesn’t matter?
      2. Uncertainty: YES process noise and observation error! This paper does an awesome job of talking and quantifying uncertainty!!
      3. Tested outside training period: yes, 50-yr training period and 50-yr validation period
      4. **Take home:** the empirical models did better than mechanistic
  11. Payne et al 2017, Lessons from the First Generation of Marine Ecological Forecast Products
      1. More of a review of marine ‘forecasting’
      2. Table 1 includes other forecast studies and what type of model they used
  12. Stow et al 2003, Comparison of Estuarine Water Quality Models for Total Maximum Daily Load Development in Neuse River Estuary
      1. Model: 2 process-based models and one probabilistic Bayesian model (empirical)
      2. Uncertainty: yes
      3. Tested outside training period: yes, trained with data pre-2000 and then validated on 2000 estuarine chlorophyll data
      4. Summary: used both process-based and empirical models to predict riverine chl at various spatial scales, but they report that none do well. Their model assessment metrics are a little hard to compare because they don’t include all of the usual metrics (r2) but their RMSE values look pretty good to me
  13. Brown et al 2013, Ecological forecasting in Chesapeake Bay: Using a mechanistic–empirical modeling approach
      1. Model: mechanistc and empirical
      2. Uncertainty: yes! They have a whole distribution of probabilities for species distributions over a geographic range, not as much discussion on uncertainty partitioning but the empirical forecasts are probabilistic
      3. Tested outside training period: yes, because the model is tested for accuracy each time a prediction is made and the future becomes present
      4. Summary: uses a mechanistic model to make predictions of multiple variation (water temps, chla, zoop, etc.) and then feeds that output into an empirical model to predict species relative abundance for some habs and other nuisance species
  14. Shaman and Karspeck 2012, Forecasting seasonal outbreaks of influenza
      1. Model: can’t tell if this is a process based or empirical model…
      2. Uncertainty:
      3. Tested outside training period:
  15. Seeing lots more empirical studies, not sure if 15 is enough?
  16. Hazen et al 2017, WhaleWatch : a dynamic management tool for predicting blue whale density in the California Current
      + Model: multiple empirical models
      + Uncertainty: yes? They mention uncertainty
      + Tested outside training period: yes, models developed and then 8-day forecasts made which were then validated as time passed
  17. Case studies in Dietze for more???
  18. Maybe before getting into approaches, start off with the state of where we are in the forecasting community- could you skim a few different forecasting studies and see what the researchers are using- empirical statistical models vs process-based models? It would be super cool if you could add in a mini-lit review of “true” forecasting studies and be able to say something like, “85% of all (n=20) forecasting studies are using process-based models; 100% of freshwater studies (n=8) are using process-based models”
* Forecasting approaches
  1. Rastetter 2017
  2. Empirical
     + Pros
       - Used in majority of studies currently
       - Necessary data commonly available through routine monitoring
         * Many data sources can be used and built to different empirical models
       - Relatively easy and quick to build
       - Inherently suited for a particular system, but easily built for others
     + Cons
       - Specific to a given system
       - Historical range of conditions may not cover the range of conditions expected under global change
  3. Process-based
     + Pros
       - Useful tools for simulating changes within a system
         * Can be used to explain underlying mechanisms
     + Cons
       - Require a multitude of input driver data
       - Many parameters
       - Require a lot of time and expertise to calibrate
         * Expertise on the model, as well as the system
* Model assessment
  1. Table of OBS, ARIMA, and GLM descriptors
     + Mean
       - Forecasts are said to be in consensus if they are insignificantly different from the sample mean (Gregory et al ., 2001)
     + Measure of variance (standard deviation)
     + quantiles
  2. Table of ARIMA vs. GLM performance metrics
     + RMSE
     + R2 (variation from the 1:1 line)
     + Bias
       - SD(model)/SD(data)
       - Base R bias calc mean(observed - model)
     + Slope of the regression btw model and OBS
     + Pearson’s correlation btw model and OBS
     + Area under the curve?? (used in Araujo, Thuiller, and Pearson 2006, Climate warming and the decline of amphibians and reptiles in Europe and probably others that I wasn’t paying attention to)
     + Stow et al 2003 also reports a list of metrics used to compare
     + Kurtosis? (bennett et al 2013)
       - Measure of how peaked the data is
       - Calculate this for both models and for observed as a comparison
  3. Figures—visual assessment
     + Plot chlorophyll over time and model predictions over time on same plot
     + Predicted vs. observed + 95% confidence interval
     + Predicted vs. observed variables with known relationship to chlorophyll
       - TP? Temp? Turbidity?, etc.
* Weighted model averaging as a way to incorporate strengths of multiple models
  1. ‘‘Observation and theory get on best when they are mixed together, both helping one another in the pursuit of truth’’ (Eddington 1935).
* FLARE framework
* Objectives of my work
  1. Question 1: How well can an empirical model and a process-based model capture/recreate etc observed chlorophyll dynamics over a 5-year period?
  2. Question 2: How well (measured through quantified uncertainties/a probabilistic forecast) can an empirical model and a process-based forecast near-term chlorophyll-a dynamics over a 16-day period?
  3. Question 3: What information does an ensemble model approach provide for chlorophyll forecasting that cannot be explained from a single-model approach? near-term chlorophyll-a?

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.665.3952&rep=rep1&type=pdf>

start with Bennett et al 2013

Lakes and reservoirs provide a suite of critical ecosystem services, including recreation, food production, and drinking water1. Despite the importance of these services, fresh waters are increasingly threatened by rapidly changing land use and climate2, two factors that have led to the contamination of drinking water for millions globally by severe harmful algal blooms (HABs). Consequently, there is a pressing need to not only understand the current state of our freshwater ecosystems, but to predict how they will respond tomorrow, next week, and next year. Therefore, the ability to forecast the future state of our drinking water sources is of utmost importance to society and freshwater ecology as a discipline.

Because the field of ecological forecasting is relatively new, there is not yet a consensus as to the best approach for predicting water quality3. Ecologists commonly use time series statistical models to predict future water quality based on past and current conditions (e.g., today’s phosphorus load, water temperature)4, while another approach uses process-based models (built from coupled differential equations) such as the General Lake Model (GLM), a one-dimensional hydrodynamic model, to simulate different metrics of water quality5. A simplified time series model enables application in many waterbodies without the intensive monitoring technology or the extensive set of parameters required by GLM, yet GLM may give a more informed forecast with lower uncertainty, allowing managers to make decisions with more confidence. However, it remains unknown as to which of these approaches is most effective for forecasting water quality at both the single and multiple waterbody scale. Further, by developing multiple models to inform forecasts of phytoplankton, we can draw from an ensemble of possible outcomes to produce a suite of forecasts with quantified uncertainty.

I propose to build on my past research studying the drivers of historical water quality to develop new methods for predicting future water quality. My project will **(1) develop hindcasts of phytoplankton** over 4-5 years using an ensemble model approachin a monitored drinking water reservoir, **(2)** **produce near-term 16-day forecasts of phytoplankton** using the developed models**,** and **(3) determine the most effective method for developing forecasts** in a drinking water reservoir. This chapter 1 will thus**: 1) inform managers and decision-makers about which variables are most important for routine monitoring and management of water quality in the face of land use and climate change, and 2) provide essential information to the research community that will guide ecological forecasting applications across many different ecosystems.**

**Methods**

* **Study Site**
  + Falling Creek Reservoir (FCR) is a small (~12 ha), shallow (maximum depth < 9.3m) dimictic drinking water reservoir located in southwestern Virginia. FCR is owned and operated by the Western Virginia Water Authority (WVWA), and lies within a forested watershed, although the reservoir continues to exhibit incidences of poor water quality as a result of historical eutrophication of the reservoir (Gerling et al 2016). The major water source to FCR comes from a single tributary which flows from Beaverdam Reservoir (BVR; Figure 1).
* Historical and sensor dataset
  + An extensive, routine monitoring dataset has been cultivated in FCR since 2013 in collaboration with the WVWA and Virginia Tech University including meteorological, physical, chemical, and biological data. The major inflow to FCR has also been routinely monitored for chemical and physical data, as well as discharge. More recently, as part of the Smart and Connected Communities (SCC) project, FCR has been outfitted with numerous high-frequency sensors to capture real-time changes in water quality. These data are
* **Objective I**
  + **Develop hindcasts of phytoplankton over 4-5 years using an ensemble model approach** I will use the historical meteorological, hydrological, biological, and chemical datasets to develop two models: an autoregressive integrated moving average (ARIMA) time series model and a calibration of GLM to assess the drivers of chlorophyll-a and HABs in FCR. Models will be developed over a training period covering 2013-2016 for ARIMA and 2013-2017 for GLM. Explain why different training periods, and give a bit more detail as to how modeling will happen? I would add in three sections here: 1) available data description; 2) ARIMA details and 3) GLM details. To develop the ARIMA, we will select variables which are predictable from the entire pool of available data and create model iterations containing all of the variables. The best model will be selected based on Akaike’s Information Criterion (AICc) and parsimony, choosing the simplest model. What is global pool of drivers? How were these chosen? GLM will be calibrated over the 2013-2017 time period. Both models will be validated using data from 2018, which has been withheld from the training dataset. Model comparisons will be made using R2 and root mean squared error (RMSE) to assess model fit of chlorophyll-a to observed chlorophyll-a.
* **Objective II: Produce near-term 16-day forecasts of phytoplankton.** 
  + Plugging into the FLARE framework
    - Use near-term NOAA forecasts + sensor-cloud networks to pull current data and forecasts of driver variables to go into ARIMA and GLM models
  + For both models, I will sample from a distribution of parameters using Bayesian techniques to generate uncertainties for each forecast
  + These forecasts will be direct deliverables to WVWA
* **Objective III: Determine the most effective method for developing forecasts.**
* The forecast model outputs will then be compared to the observed high-frequency chlorophyll-a sensor data to determine which approach is most robust under different conditions.
* Model fit will be assess using R2 and RMSE
* If model strengths between ARIMA and GLM are distinct, a weighted model average may be developed to capture the relative strengths of each model and cover as much of the variability in chlorophyll-a as possible
* TEXT TO BE INCORPORATED INTO THIS SECTION: Once both models are calibrated, I will assess how effective each model is at predicting near-term future states of water quality. For GLM, I will use 2-week weather predictions to run the calibrated model to predict chlorophyll-a. For the time series model, if the model indicates, for example, that the previous month’s P load is the most important driver of current chlorophyll-a, I will feed the model with current P loads to predict future chlorophyll-a. For both models, I will sample from a distribution of parameters using Bayesian techniques to generate uncertainties for each forecast. The forecast model outputs will then be compared to the observed high-frequency chlorophyll-a sensor data to determine which approach is most robust under different conditions. I predict that output from the time series model will have a larger uncertainty than the GLM model, given that GLM is a process-based model driven by coupled equations, rather than statistical relationships alone.

**Preliminary Results**

* Metrics of comparing models
  + Kurtosis? (bennett et al 2013)
    - Measure of how peaked the data is
    - Calculate this for both models and for observed as a comparison
  + R2 (deviance from the 1:1 line)
  + RMSE
* Objective I: hindcasts of phytoplankton
  + ARIMA
    - Autoregressive model includes chlorophyll-a at the previous timestep, discharge to the reservoir, and shortwave radiation
    - Over 2013-2016, R2 = 0.44, RMSE = 1.71 ug/L
    - Captures quite detailed dynamics at lower chlorophyll-a concentrations (<10 ug/L—in CTD units, maybe need to scale this for general comparison if most instruments read higher and that’s the magnitude people are used to thinking about chlorophyll?)
    - Misses some larger peaks
    - 2018 validation: R2 = 0.44, RMSE = 1.02 ug/L
  + GLM
    - Calibration still in progress
    - Over 2013-2017, R2 = 0.001, RMSE = 3.42 ug/L
      * This includes winter months which ARIMA does not model
    - Does a better job at capture larger peaks in chlorophyll-a
    - 2018 validation: in progress
* Objective II
  + Plug both models into the FLARE framework
  + Fully quantified uncertainties for both ARIMA and GLM
* Objective III
  + Weighted model averaging to incorporate strengths of both models (e.g., detailed changes in concentration and peaks)

Chapter 2 Assessing the dynamics of stream-reservoir linkages across reservoirs

*Question 1: How do stream nutrient and water loads affect reservoir chla?*

*Question 2: How do major storm events influence these dynamics?*

***Question 4: How does within-reservoir nutrient processing change along a double reservoir continuum?***

***Question 3: How do stream dynamics differ between neighboring watersheds?***

***How do upstream reservoir dynamics affect downstream reservoir dynamics?***

***What is the long-term effect of mgmt. on FW ecosystem functioning?***

**Background**

* Importance of watershed characteristics on waterbody condition
  + Terrestrial-aquatic linkages
  + Stream-lake linkages
    - Further proposed reading: Wurtsbaugh et al 2005, Robinson et al 2007, Goodman et al 2011, Xu & Xu 2018, Stachelek & Soranno 2019, Marcarelli & Wurtsbaugh 2009, Sadro et al 2012, Schmadel et al 2018, Jones 2010, Hotchkiss et al 2018., ETC
* Importance of storm events in stream dynamics
* In the new age of data availability, can we relate stream dynamics of a highly monitored stream to neighboring streams lacking data availability?

**Methods**

* Study site
  + Introduce BVR
    - 4 selected inflows (see Figure 1)
    - Relationship to FCR
      * Watershed history
      * Geographical similarity
      * Outflow to FCR
* Field sampling, summer 2019
  + Weekly (bi-monthly?) sampling of 4 major inflows to develop baseline
    - Nutrient chemistry (total and soluble nitrogen, phosphorus, and carbon)
    - Discharge (using a flowmeter)
    - Physical characteristics (dissolved oxygen, conductivity, temperature)
  + Periodic event-based sampling
    - ISCO sampler to capture high-frequency dynamics during major flow events
  + Weekly sampling at deep-hole of BVR
    - Biological, physical, and chemical sampling
  + Grab samples at strategically chosen sites along the reservoir gradient (?) (see red arrows on Figure 1)
    - CTD + nutrient grab samples

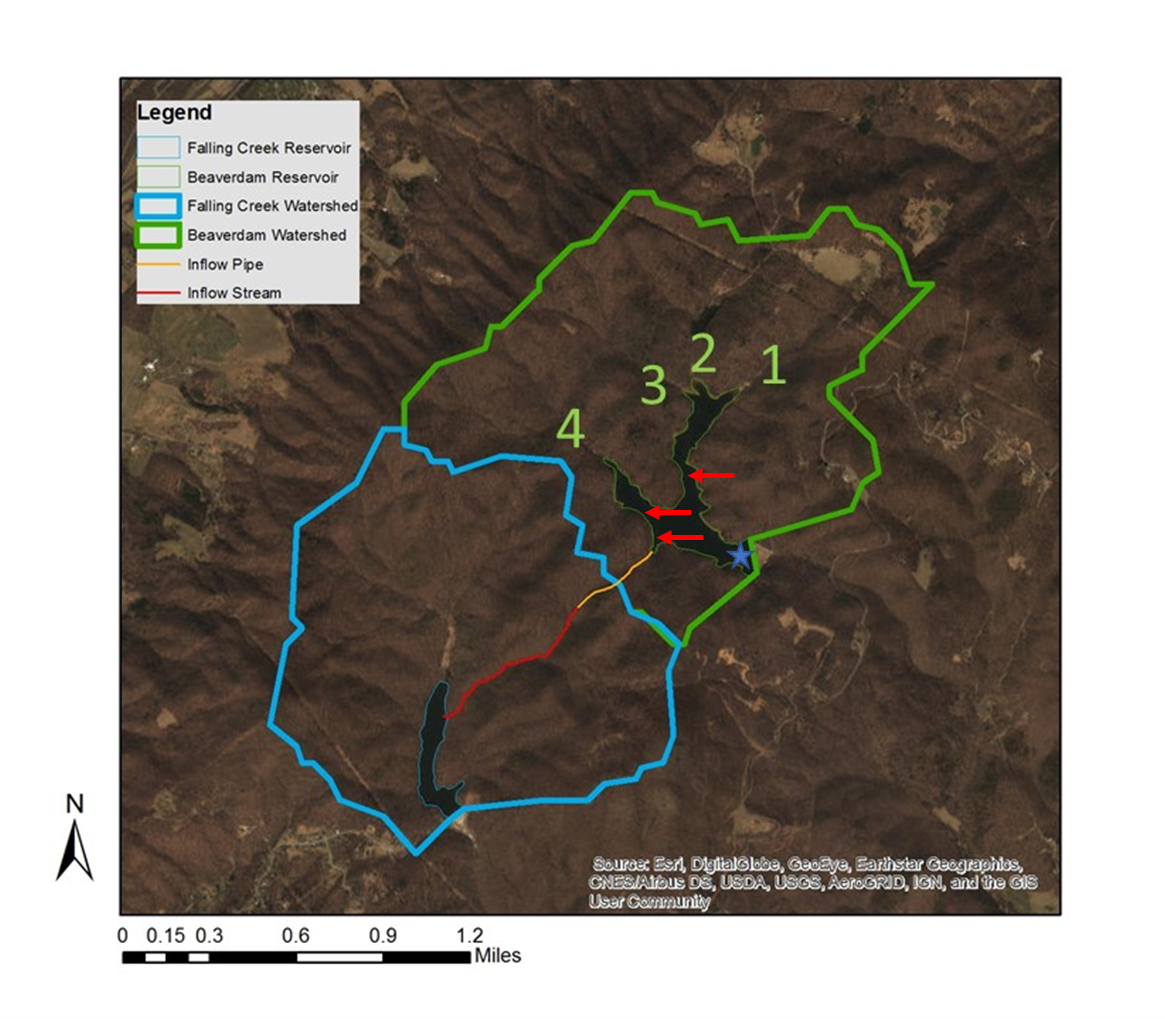


Figure 1. Map of Falling Creek Reservoir and Beaverdam Reservoir, along with watershed boundaries and proposed sampling sites

* Data analysis
  + Generalized linear models to analyze drivers of nutrient and phytoplankton dynamics along a stream-reservoir gradient

References

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