Whitney Woelmer, Master’s Prospectus   
05 April 2019

**Overall Introduction**

Freshwater ecosystems provide crucial services to society. Whether it be naturally-formed lakes or human-built impoundments, freshwater ecosystems provide water for irrigation, industry, recreation, tourism, food production, and most importantly, drinking water (Carpenter et al., 2011). But as the human population has grown and our need for fresh water has increased, these resources have been degraded (Smith, 2003). My research interests combine the need to understand how water quality in freshwater ecosystems is changing and how we can better predict and adapt to improve the future state of freshwater ecosystems.

Climate change and land use change are two factors that have put the health of freshwater ecosystems in jeopardy (Brookes and Carey, 2011; Carpenter et al., 2011; Jeppesen et al., 2010; Smith, 2003). Eutrophication, excessive nutrient enrichment of a freshwater ecosystem, is increasing in both severity and occurrence (Smith, 2003), and is just one type of water quality impairment as a result of global change. Harmful algal blooms (HABs), a result of eutrophication, are known to produce unsightly water quality problems that can clog filters and produce deadly toxins, jeopardizing human health and costing millions of dollars to water treatment plants (Cooke and Kennedy, 2001). The economic impact of HABs ranges from decreased tourism to costly mitigation for water treatment plants, and is estimated to be over $2 billion in the U.S. alone annually (Dodds et al., 2009). The short generation time of phytoplankton species, major contributors to HABs, further motivates their study because these negative impacts are often realized on a rapid time scale.

The influence of human activity can quickly cascade across catchments through the transport of water through and among stream networks downstream to lakes and reservoirs. As precipitation patterns are expected to change, leading to fewer but more intense storms (Prein et al., 2017), it is likely that we will see cascading effects of decreased water quality in freshwater ecosystems. For example, with increased storm frequency, we are likely to see rapid transport of sediment and nutrients from the landscape to downstream freshwater ecosystems, which can lead to increases in nutrient concentrations, turbidity, HABs, and other poor water quality events within lakes and reservoirs (Jeppesen et al., 2011). Consequently, we need to better understand how rapid temporal and spatial changes in nutrients and biological communities can occur as a result of changing climate and land use. By studying nutrient and phytoplankton dynamics at a high-resolution temporal and spatial scale, we can gain an understanding of how our freshwater ecosystems are responding to global change.

As the uncertainty of water quality increases due to climate and land use change, anticipating future water quality events is increasingly important and needed to ensure the provisioning of critical ecosystem services. For society to coexist with changing ecosystems and for managers to anticipate and potentially mitigate poor water quality events before they happen, we need forecasts of future water quality and other ecosystem variables. For example, if water managers in Toledo, OH had known in advance that a major HAB event in the western basin of Lake Erie in 2014 was going to occur, preventative measures could have been taken that would have lessened the $65 million loss in economic damage and social trust that resulted from the shutdown of water utilities for half a million people over several days (Bingham et al. 2015).

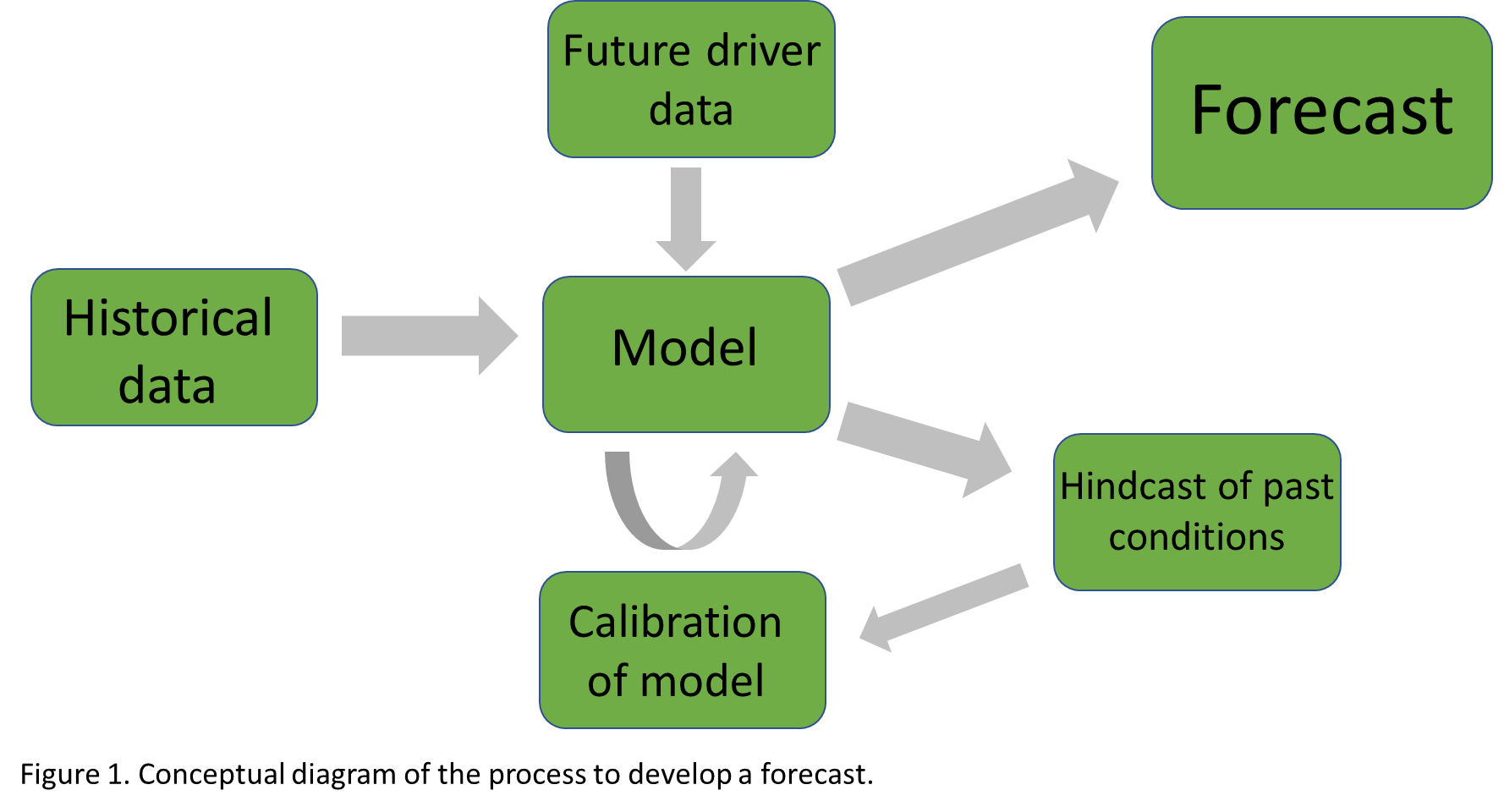
Together, my two chapters will better inform the scientific community on the drivers and dynamics of water quality in small, eutrophic drinking water reservoirs and produce forecasts of how water quality will respond to changing conditions in the future.

**Proposed Research**

*Chapter 1: Developing near-term forecasts of phytoplankton in a drinking water reservoir***Introduction**

Given the unprecedented level of anthropogenic degradation already experienced by freshwater lakes, reservoirs, streams, and wetlands due to land use and climate change (Vorosmarty et al. 2005), understanding not only the current state of our freshwater ecosystems, but predicting how they will respond tomorrow, next week, and next year is of utmost importance. In particular, knowing future water quality in advance would provide a substantial benefit to managers and freshwater ecologists by allowing them to make preventative measures when a poor water quality event is predicted, rather than managing an ecosystem after poor water quality has already occurred. Being able to anticipate water quality impairment could help save millions of dollars in water treatment costs, as well as protecting the trust required between a community and their water utility by avoiding a shutdown of water resources (e.g., Cvetkovich and Löfstedt, 1999; Ross et al. 2014). A recent survey has shown that less than half of Americans are confident in their water utility’s ability to treat their drinking water (AP-GfK, *Tap Water Confidence Poll 2016*). 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.

Forecasting as a technique outside of ecology has been developing for decades in many disciplines and has substantial breadth. Across disciplines, a forecasts are developed by using historical data to produce hindcasts of past conditions, which is used to inform calibration of the model that is then used to produce a forecast (Figure 1.) Uses of forecasting can be found in many fields and applications, ranging from the well-known and widely-used weather forecasts to epidemiological forecasts of population mortality (Lee et al., 1992) and global Alzheimer’s trends (Brookmeyer et al., 2007) to economic forecasts of bankruptcy declaration in the United States (Shumway et al., 2001). Many of these fields have been improving their forecasting abilities for decades (Shumway et al., 2001). For example, the accuracy of 36-hour weather forecasts has increased from ~25% in the 1950s to ~80% in the early 2000s (Dietze, 2017). This slow yet substantial increase in forecast accuracy over the past 50 years shows that progress cannot be expected to be immediate, and that there is merit in developing forecasts even when they have low accuracy.



In comparison to weather and economic forecasts, the development of forecasts of ecosystems and ecosystem services is still in its relative infancy (Dietze, 2017). The current shift in the scientific community towards publishing data in publicly accessible venues (Wilkinson, 2016), as well as the development of high-frequency sensors which result in massive amounts of data (Hampton et al., 2013), have both enabled the field of ecological forecasting to recently expand. However, as the development of ecological forecasting is still in very early stages, there is not yet a consensus as to the best approach for making ecological forecasts. In particular, it is unclear if data-driven empirical models or numerical simulation process-based models yield more accurate forecasts (Dietze, 2017). To better understand the current state of studies using ecological forecasting and their approaches, I conducted a high-level literature review. Studies were found using the Google Scholar database with the search terms ‘ecological forecasting’ and ‘forecasts.’ Studies which met the following requirements were selected as using ecological forecasting methods: they must 1) use models, 2) quantify uncertainty to make a probabilistic forecast or forecasts, and 3) run the model outside a specified training period (Table 1).

Empirical modeling methods were favored a majority of times (65%, n= 17, Table 1) in current ecological forecasting studies selected in the literature review. Empirical modeling approaches were likely popular because of their data-driven nature; they are inherently developed for a single particular ecosystem because they are based on past trends within that ecosystem. Input time series data for empirical models are commonly available through routine monitoring and empirical models are generally simple and quick to develop and implement (Hipel and Macleod, 1994). However, because empirical models are built on the historical conditions of a system, if future conditions are outside the realm of past conditions, these models may no longer be able to accurately make predictions (Dietze, 2017). In contrast to empirical time series models, process-based models were favored in just over a third of the studies examined (35%, n=17, Table 1). It is not surprising that process-based models are currently less frequently used in the ecological forecasting literature given that they often require more input data, as well as more time and expertise to properly calibrate the multitude of parameters to accurately represent a specific waterbody. However, these models may be especially useful in the current era of global change given their ability to explain underlying mechanisms which cause a given response and are likely more generalizable to other systems because they are based on a set of fundamental equations rather than the historical pattern of a single system.

The studies examined in the literature review spanned the realm of basic and applied research, showing multiple motivations for ecological forecasting. While a majority of the studies in our literature review were conducted to inform the field of basic ecology (73%, n = 15), all of the studies whose forecasting products were directly tied to a management tool or product were focused on forecasting various aspects of aquatic ecosystems, including water quality variables as well as species distributions in marine environments and habitat quality (Table 1). This finding illustrates not only the interest in the scientific community to understand changing aquatic systems, but the need by managers and stakeholders for probabilistic forecasts in order to respond and adapt to changing conditions in aquatic ecosystems.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 1. Summary of literature review targeting studies using ecological forecasting methods. | | | | | |
| Authors & Year | Approach | Best Model | Ecosystem | Forecast Product | Use of forecast product in decision making (Y/N) |
| Araújo et al., 2005 | Uses multiple empirical models to assess uncertainty in projections | Empirical | Terrestrial | Bird species ranges | N |
| Brown et al., 2013 | Coupled process-based and empirical model to predict water quality | Process-based and empirical | Temperate estuary | Chesapeake Bay water quality | Y |
| Dean et al., 2004 | Combines climate projections and process-based model to make probabilistic projections | Process-based | Central highland forests | Carbon sequestration in forests | N |
| Estes et al., 2013 | Comparison of empirical and process-based models | Empirical | Agricultural dryland | Productivity and suitability of crops in South Africa | N |
| Gonzalez-Benecke and Martin, 2017 | Future projections of forest growth | Process-based | Southern-temperate forest | Loblolly Pine projections | N |
| Hazen et al., 2017 | Coupled empirical models to produce habitat suitability forecasts | Empirical | Marine | 8-day forecasts of Blue Whale density | Y |
| Lindegren et al., 2010 | Forecast impacts of climate and fishing pressure on marine food webs | Process-based | Marine | Baltic cod dynamics | N |
| Martínez-Meyer et al., 2004 | Used machine-learning techniques to predict species distributions during past geological time periods based on current distributions | Empirical | Conterminous United States during the present and the Pleistocene Era | Ecological niches of mammal species | N |
| Perretti et al., 2013 | Compared forecast efficiency of mechanistic and empirical models | Empirical | Simulated and laboratory data for beetle species | Species abundance | N |
| Stow et al., 2003 | Compared forecast efficiency of two process-based and one empirical model | Process-based and empirical | Temperate estuary | Neuse River Estuarine water quality | Y |
| Thomas, M.K. et al., 2018 | Used an empirical approach to determine forecast efficiency at multiple time scale (hours, months, years) | Empirical | Freshwater | Phytoplankton dynamics | N |
| Thuiller et al., 2004 | Examined the sensitivity of an empirical model to restrictions in input driver data | Empirical | Temperate forest | Tree species distributions | N |
| White and Nemani, 2004 | Used a process-based model to determine the relative importance of meteorology and vegetation phenology | Process-based | Temperate forest | Soil water concentrations | N |
| Woodbury et al., 1998 | Quantified forecast uncertainties from climate driver data, forest condition, and quantitative relationships using an empirical model | Empirical | Temperate forest | Loblolly pine growth | N |

My first chapter will focus on addressing the knowledge gap in ecological forecasting regarding how best to forecast water quality. We will focus on the dynamics of HABs (harmful algal blooms) as one metric of water quality in our system, which will be measured by changes in phytoplankton abundance. I will produce hindcasts and near-term iterative forecasts of phytoplankton in a drinking water reservoir using both an empirical and a process-based approach. Model performance will be assessed by a suite of metrics addressing both the ability of the model to capture overall phytoplankton dynamics, as well as extreme events (blooms). My work will specifically address the following questions:

* + Question 1: How well can an empirical model and a process-based model hindcast observed phytoplankton dynamics?
  + Question 2: How well can an empirical model and a process-based model forecast near-term phytoplankton dynamics over a 16-day period (assessed by comparing quantified uncertainties of a probabilistic forecast with observed dynamics)?
  + Question 3: How does an ensemble model approach improve 16-day forecasts of phytoplankton over a single model approach?

Chapter 1 will thus span the field of applied and basic forecasting science by**: 1) informing 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 as part of setting up the empirical model, and 2) providing essential information to the freshwater research community regarding modeling approaches to forecasting water quality.**

**Methods**

*Study Site*

Falling Creek Reservoir (FCR) is a small (~12 ha), shallow (maximum depth = ~9.3 m) dimictic drinking water reservoir located in southwestern Virginia (37.30333333˚N, -79.83722222˚E). FCR is owned and operated by the Western Virginia Water Authority (WVWA). The watershed of FCR is almost entirely forested, although the reservoir continues to exhibit incidences of poor water quality as a result of historical eutrophication (Gerling et al., 2016). The major water source to FCR comes from a single tributary which flows from Beaverdam Reservoir (BVR; Figure 1).

A close up of a logo

Description automatically generated

Figure 1. Map of Falling Creek Reservoir and Beaverdam Reservoir and their watersheds.

*Historical and sensor dataset*

An extensive, routine monitoring dataset of water quality in FCR has been collected since 2013 in collaboration with the WVWA and Virginia Tech. This dataset includes meteorological, physical, chemical, and biological data collected both at the deep hole of the reservoir and at the major inflow to FCR (EDI Portal: Carey et al. 2018, Carey et al. 2019). The inflow dataset also includes discharge to the reservoir measured every 15 minutes at a weir installed at the stream site. 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 streamed wirelessly to a staging server and pushed to Github multiple times per day. Sensor data include multiple meteorological (e.g., precipitation, shortwave radiation), physical (e.g., water temperature, dissolved oxygen), chemical (e.g. total and soluble nutrients), and biological variables (e.g., chlorophyll-a concentrations).

*Model Development—Question 1: How well can an empirical model and a process-based model hindcast observed phytoplankton dynamics?*

We developed both an empirical and a process-based model to forecast chlorophyll-a, a proxy for phytoplankton, in the surface water (1.0 m) at FCR during the summer stratified period (May-October), when phytoplankton populations are at their highest. The training period for both models was 2013-2016 during which we have regularly, weekly coverage of both response and driver data. Both models will be validated using data from 2018.

Empirical: ARIMA

We chose an autoregressive integrated moving average (ARIMA) generalized linear model for our empirical approach to forecast chlorophyll-a in FCR. The autoregressive term in ARIMA was determined by selecting the previous timestep of chlorophyll-a with the highest Pearson’s r correlation coefficient with the current measurement of chlorophyll-a. From a pool of 53 potential meteorological, physical, chemical, and biological driver variables, we first focused on driver variables that have biological significance for phytoplankton growth and which are also predictable using physical models (e.g., temperature, discharge). We excluded variables that were correlated with each other through the use of a Pearson’s correlation analysis (r > 0.5 & r < -0.5). Using these variables, we developed all possible ARIMA combinations with the selected driver variables, and the best model was determined by AICc (corrected Akaike’s Information Criterion).

Process-based: GLM

We used the General Lake Model (Hipsey et al., 2019, 2017), coupled with Aquatic  
Ecodynamics (Hipsey et al., 2013), hereafter, GLM-AED, as our process-based model. GLM-AED is a 1-dimensional, open-source lake ecosystem model. GLM-AED requires a number of input data driver files, including the hypsographic curve of the lake, initial physical, chemical, and biological conditions, daily stream inflow volume and water chemistry, and hourly weather. Hourly climate driver data from 2013-2017 were downloaded from the North American Land Data Assimilation System (NLDAS-2) website, and other driver data files were derived from the historical and sensor dataset for FCR.

GLM-AED was calibrated to best fit phytoplankton dynamics during the summer stratified period from 2013-2016. Parameters were systematically adjusted to improve the model fit with observational data, first for temperature, then for dissolved oxygen, nitrogen, phosphorus, and carbon, and finally for phytoplankton (following Bennett et al., 2013).

Model Assessment

Model performance will be assessed using a number of metrics currently being used in the ecological forecasting field. As one metric alone cannot determine if a model performs better than another, I will use a holistic approach comparing several model performance metrics.

Comparisons of observed and model descriptors will include the mean, standard deviation, quantiles (Dietze, 2017), and kurtosis (Bennett et al., 2013). Model performance will also be examined through visual assessment of figures showing observed and model predictions over time, and predicted versus observed data (Dietze, 2017). Model outputs will also be compared with each other using the performance metrics found in Table 2.

|  |  |  |
| --- | --- | --- |
| **Table 2. Model Assessment Metrics for comparing forecast model outputs** | | |
| **Metric** | **Equation** | **Ideal Value** |
| RMSE (root mean squared error) |  | 0 |
| R2 (variation from the 1:1 line) |  | 1 |
| Bias |  | 1 |
| Probability of bloom detection |  | 1 |
| Probability of false bloom detection (false alarm rate) |  | 0 |

*Forecasting Framework: FLARE—Question 2: How well can an empirical model and a process-based model forecast near-term phytoplankton dynamics over a 16-day period?*

After addressing Question 1, both models will be integrated into an existing forecasting framework, Forecasting Lake and Reservoir Ecosystems (FLARE), to produce iterative near-term 16-day forecasts of chlorophyll-a (Thomas et al., in prep). FLARE is designed to pull 2-week weather predictions from the National Oceanic and Atmospheric Administration (NOAA) GEFS server to force both ARIMA and GLM-AED on a daily time step. Additionally, ARIMA’s autoregressive term will be determined from measured chlorophyll-a streamed via our sensor-cloud network to provide yesterday’s chlorophyll-a concentration as well as for comparison with the previous days’ forecasts. Lastly, inflow discharge data for both models will be estimated using an autoregressive linear model of discharge based on the previous day’s mean discharge at the major inflow to the reservoir. Driver uncertainty will be calculated for both ARIMA and GLM-AED forecasts by propagating the uncertainty associated with the NOAA weather forecasts. Parameter uncertainties for both ARIMA and GLM-AED will be calculated by sampling from a distribution of key parameters within each model. Forecast effectiveness will be assessed using a suite of performance metrics calculated by comparing the forecast model outputs to the observed high-frequency chlorophyll-a sensor in FCR (Table 2). Assessment metrics will be calculated for various conditions (e.g., summer stratified period, fall mixed period, winter ice period, and following storm events) to determine which approach is most robust under different reservoir and meteorological conditions.

*Integrated Model Averaging—Question 3: How does an ensemble model approach improve 16-day forecasts of phytoplankton over a single model approach?*

To develop the most informative forecast of chlorophyll-a, we will explore the use of a weighted model integration that can include both empirical and process-based model outputs. This third question is motivated by the preliminary finding that while both of our current model outputs capture observed chlorophyll-a dynamics at lower concentrations, bloom events are often underestimated, leading to a “false negative” (see Preliminary Results section below). This result strongly suggests that a weighted average of two or more models may lead to a forecast that is able to more accurately predict when large bloom events are likely to occur. The use of integrated model averaging, or “ensemble” modeling, which has been adopted in the field of weather forecasting, may be a more effective way to forecast water quality because of its ability to capitalize on the different strengths of each model type. This will allow the ensemble forecast to capture as much variability as possible in chlorophyll-a and the different strengths of the two models.I anticipate that for low probability yet high impact events such as phytoplankton blooms, an ensemble approach may be necessary to capture these peaks in phytoplankton concentration that may operate under different mechanisms. Incorporating multiple models will enable us to develop an early-warning system which will alert users when conditions that indicate a peak in phytoplankton is likely to occur.

**Preliminary Results**

Question 1: Hindcasts of chlorophyll-a

Using the summer period of 2013-2016 as our training period, we developed and calibrated both an empirical model (ARIMA) and process-based model (GLM-AED). Our best-fitting ARIMA model over 2013-2016 included discharge to the reservoir and shortwave radiation, where ‘t’ is the current timestep and ‘t-1’ is the previous timestep, with 1 standard error of each parameter term included:

**Chlorophyll-at = 1.65(±0.26) + 0.45(±0.08)Chlorophyll-at-1 – 3.05(±1.39)Discharget**

**– 0.0025(±)Shortwavet + Ɛ** (eqn. 1)

ARIMA hindcasted chlorophyll-a over 2013-2016 with an R2 = 0.49 and RMSE = 1.71 ug/L. The ARIMA model was able to successfully capture fluctuations at lower chlorophyll-a concentrations (<10 ug/L) (Figure 2). However, when chlorophyll-a reached values above ~10 ug/L, the model the model was unable to recreate these observed dynamics. GLM-AED was calibrated over 2013-2016 and hindcasted chlorophyll-a with an R2 = 0.05 and RMSE = 3.07. While GLM-AED does not capture the same fluctuations in chlorophyll-a dynamics as the empirical model, it does capture some large peaks in chlorophyll-a that the ARIMA does not.

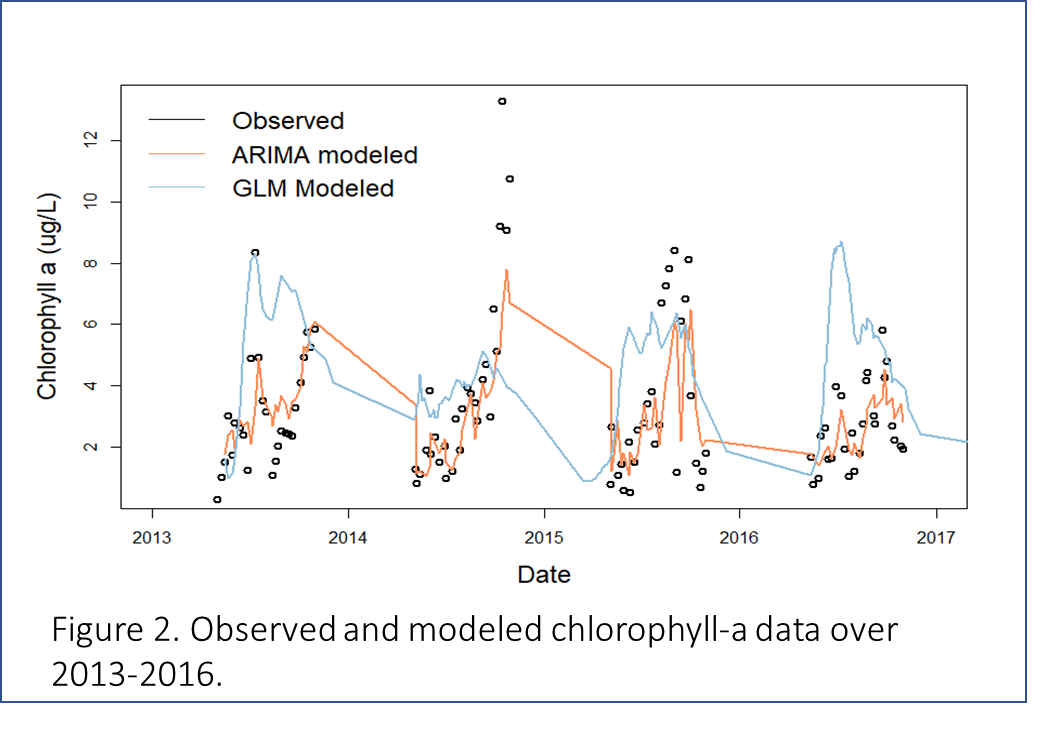


Table 3 shows the proposed timeline of remaining steps to finish work required for Chapter 1 and to address the steps needed to address Questions 2 and 3.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 3. Proposed timeline for Chapter 1** *Black boxes indicate completed tasks, gray boxes indicate proposed completion* | | | | | |
|  | Fall 2018 | Spring 2019 | Summer 2019 | Fall 2019 | Spring 2020 |
| Develop ARIMA GLM-AED |  |  |  |  |  |
| Calibrate GLM-AED |  |  |  |  |  |
| Produce hindcasts |  |  |  |  |  |
| Validate model (Q1) |  |  |  |  |  |
| Adapt FLARE code for ARIMA (Q2) |  |  |  |  |  |
| Create integrated model average (Q3) |  |  |  |  |  |
| Run forecasts in FLARE |  |  |  |  |  |
| Submit manuscript, target journal: *Ecological Applications* |  |  |  |  |  |

**Chapter 2. Spatial heterogeneity of nutrients and phytoplankton along a double reservoir continuum**

Reservoirs are ecosystems marked by physical heterogeneity and are classically defined as having a gradient of riverine, transitional, and lacustrine zones (Thornton et al. 1990). The nutrient chemistry and biological communities change substantially as water moves downstream along this gradient, as demonstrated by the large decrease in nutrient concentrations between the inflow and outflow of reservoirs (Harrison et al., 2009; Kling et al., 2000; Powers and Tank, 2015). However, studies examining how concentrations of nitrogen, phosphorus, and chlorophyll-a change along the reservoir gradient atmultiple sitesbetween the inflow and outflow are rare, and those that do measure nutrients and biology throughout the reservoir zones along the gradient report inconsistent patterns (see below). Further, because managers generally extract water from reservoirs for drinking from the deep hole only, most reservoir studies only focus on the lacustrine zone (e.g., Calijuri, M.C. et al., 2002; Hamre et al., 2018; Liu et al., 2019; Rigosi et al., 2011), leading to a critical gap in the understanding of how these water quality variables change along a dynamic reservoir gradient.

For my second chapter, I am particularly interested in examining which sites along a reservoir gradient that disproportionally affect nutrient concentrations and chlorophyll-a, which has many implications for water quality in the lacustrine zone that is extracted for drinking. Many interacting biogeochemical processes occur along the reservoir gradient that need to be considered in order to understand the inconsistent patterns in nutrient and chlorophyll-a concentrations reported in previous studies. If the location of sampling sites is adjacent to reservoir “hot spots” of nutrient processing (*sensu* (Mcclain et al., 2003) that may occur within and at the intersection between zones, snapshot sampling at the inflow and outflow alone may not reflect longitudinal trends of the reservoir continuum. Thus, finely-resolved spatial sampling is needed to fully understand longitudinal reservoir nutrient and chlorophyll-a dynamics.

My proposed work builds on past studies looking at nitrogen concentrations along a reservoir gradient that report a number of contradictory patterns. Some have found that nitrogen concentrations are highest in the riverine zone due to high nutrient inputs from the watershed (Soares et al., 2012), while others have found that the transitional zone, where turbulence is decreased yet nutrients are often still abundant, is a hotspot for nitrogen fixation across reservoirs of varying trophic status, effectively increasing the concentration of ammonium (NH4) (Scott et al., 2009). Additionally, when the hypolimnion of a reservoir becomes anoxic, it can lead to disproportionately high levels of NH4 being released from the sediment while nitrate (NO3) is concurrently denitrified within the lacustrine zone (Gerling et al., 2016). These dynamics may lead to an effective increase in NH4 yet decrease in NO3 in the lacustrine zone of the reservoir.

Similar to nitrogen, patterns in chlorophyll-a heterogeneity are also inconsistent between studies. There is some support to show that chlorophyll-a is highest in the riverine zone, decreasing along the downstream gradient (Scott et al. 2009). In contrast, others report the transitional zone as being an especially active zone for phytoplankton activity (Rychtecky and Znachor, 2011; Thornton 1990), and others still have found the lacustrine zone to have the highest chlorophyll-a in ecosystems that are not limited by nutrient availability (Soares et al. 2012). Further, Borges et al. (2008) found that the longitudinal pattern of heterogeneity of chlorophyll-a was variable among seasons in two tropical reservoirs in Brazil. This finding highlights the fact that patterns in reservoir chlorophyll-a are highly variable.

Spatial trends in reservoir phosphorus concentrations are less reported than nitrogen or chlorophyll-a. One study of a large, deep reservoir in the Czech Republic showed that both soluble and total phosphorus decreased along the reservoir gradient (Rychtecky & Znachor, 2011). Borges et al. (2008) found contrasting patterns in two reservoirs in Brazil, one showing a decrease in phosphorus along the reservoir gradient, while another showed no clear longitudinal pattern. Sedimentation processes also influence longitudinal distributions of phosphorus, with classical reservoir theory positing that sedimentation should be highest in the transitional zone, which should lead to an overall decrease of nutrients along the reservoir gradient (Thornton 1990). However, internal loading of soluble phosphorus during anoxic conditions in the hypolimnion may lead to increases in phosphorus in the lacustrine zone (Thornton 1990). While all of these studies agree that spatial heterogeneity of nutrients and phytoplankton exists along a reservoir gradient, the lack of consistency among studies necessitates additional research, especially in smaller reservoirs, which overall remain understudied relative to larger reservoirs. In smaller reservoirs, the transitions among zones may happen more rapidly, which could result in greater variability in nitrogen, phosphorus, and chlorophyll-a along the downstream gradient.

An important factor that can affect the heterogeneity of chemical and biological variables along reservoir gradients is hydrologic flow. Under low flow conditions, residence times of reservoirs and streams are increased, resulting in slower export of nutrients and more time for biotic processing (Saunders and Kalff, 2001). This can lead to an increase in chlorophyll-a concentrations while nutrient conditions remain high, but external loading of nutrients is low under this scenario and nutrients can become quickly depleted. In comparison, high flow conditions may result in less biotic uptake of nutrients by phytoplankton as water is flushed out more rapidly, but high flow also brings in high levels of external nutrients which are often limiting to phytoplankton. This can result in high nitrogen and phosphorus, but low chlorophyll-a concentrations. Low chlorophyll-a concentrations are typical of the deep hole of a reservoir, which is often the only study site in reservoir surveys, and less is known about how these varying hydrologic regimes affect nutrient and chlorophyll-a concentrations along the entire reservoir gradient. One study measured a direct decrease in spatial heterogeneity of phytoplankton with increases in residence time (Soares et al. 2012), illustrating the need to better understand how nutrient and chlorophyll-a vary not just at the lacustrine site but along the entire reservoir gradient. Further support for the importance of hydrologic flow in driving phytoplankton at the lacustrine site comes from preliminary results of Chapter 1. The empirical model from my forecasting chapter revealed the importance of incoming stream discharge and variability in hydrologic flow to phytoplankton dynamics at the deepest site of a reservoir. Gaining a better understanding of how upstream discharge can influence nutrient dynamics along the reservoir continuum before reaching the lacustrine zone, may help in identifying ways to better manage chlorophyll-a at the deep hole of the reservoir.

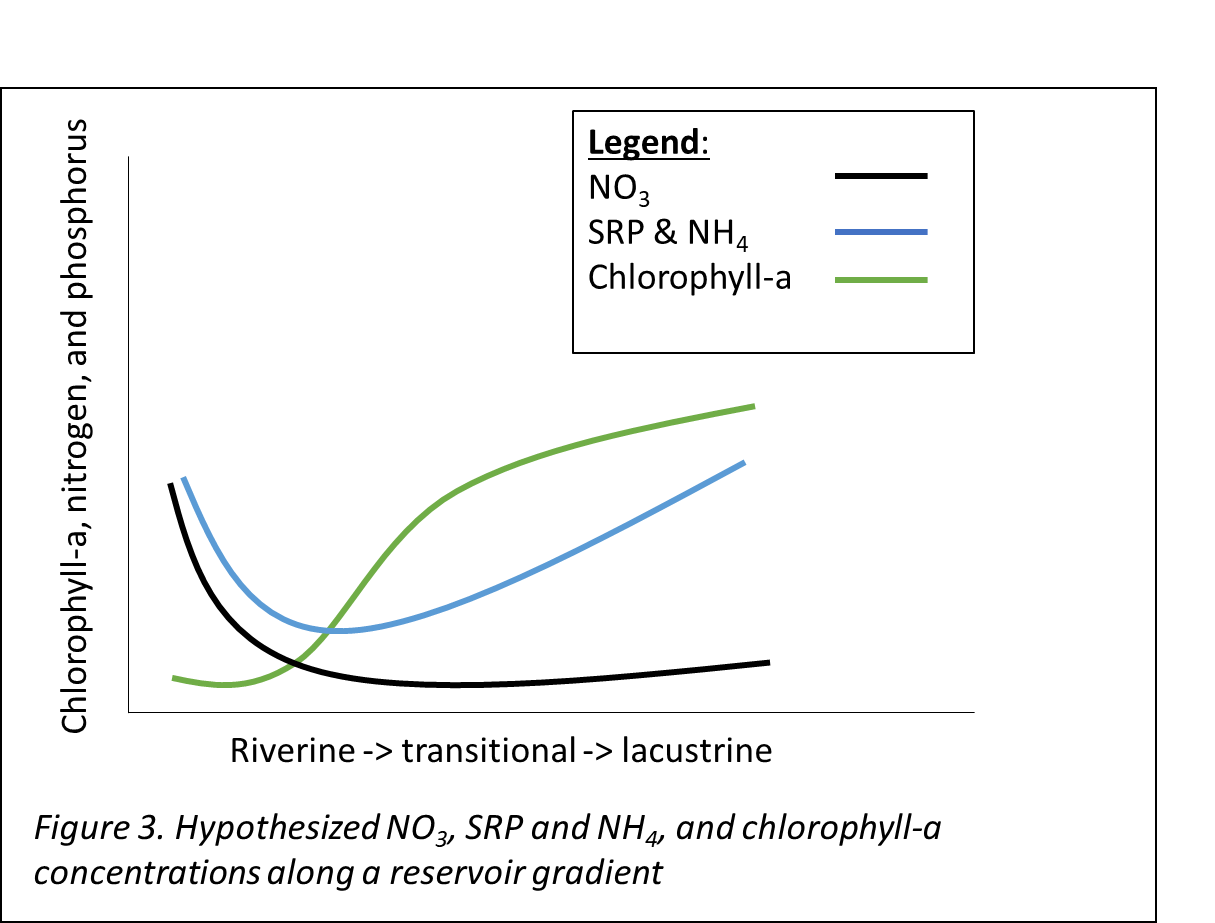
In the face of global change, understanding how water quality varies along the reservoir continuum is increasingly important. Large storms have already increased in frequency and severity and are projected to continue these patterns under future climate change (Prein et al., 2017). With these storms, we are likely to see increases in hydrologic flow which can increase external loading (Jeppesen et al., 2011) and changes to mixing regimes within reservoirs which can alter internal loading (Butcher et al., 2015; Kirillin, 2010). Both of these climate drivers may impact the heterogeneity of nutrients and chlorophyll-a along a reservoir continuum. This increased heterogeneity due to global change highlights the need to improve our understanding of how reservoir ecosystems function. Moreover, identifying areas within a reservoir where chemical or biological concentrations are disproportionately high can help prioritize management efforts to areas within freshwater systems that are more vulnerable to eutrophication or changes in trophic status.

**Proposed Work**

My second chapter will focus on the longitudinal spatial heterogeneity of nitrogen, phosphorus, and chlorophyll-a in Beaverdam Reservoir (BVR) and Falling Creek Reservoir (FCR) and the influence of hydrologic flow on this heterogeneity. Our study system also offers the opportunity to examine this question across a double reservoir continuum, as outflows from BVR feed the main inflow to FCR, and assess the effect of reservoir chain number on relative concentrations of nitrogen, phosphorus, and chlorophyll-a which has been shown to be important in many lake ecosystems (Sadro et al., 2012; Soranno et al., 1999; Stachelek and Soranno, 2019). This research will inform our understanding of how reservoirs will respond to global climate change, which is expected to include more variability in hydrologic flow and resultant residence times as a result of changing precipitation patterns.

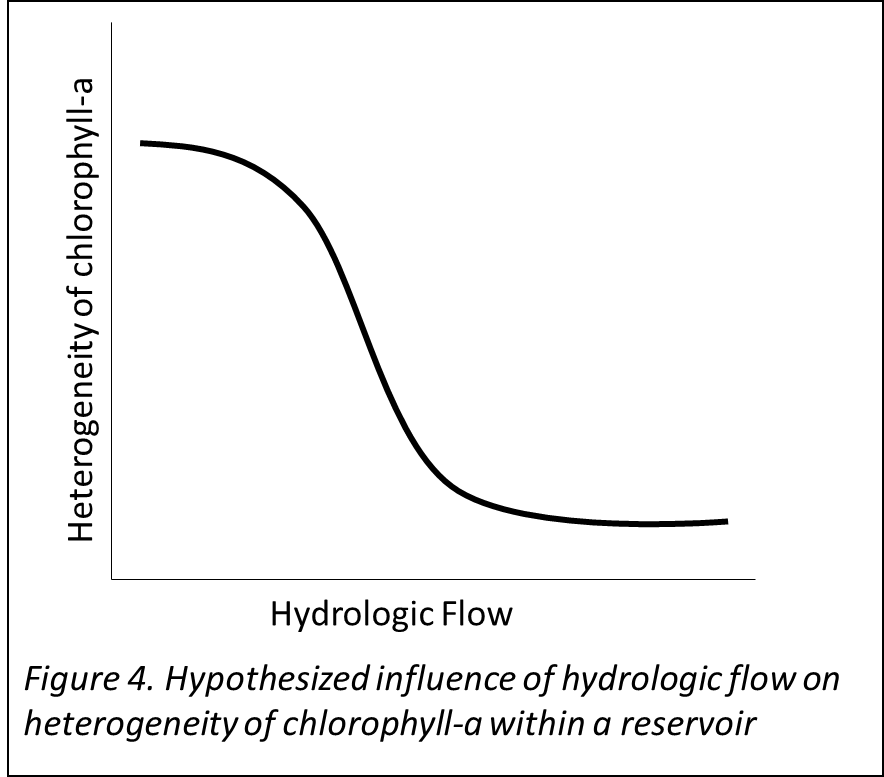
***Question 1****: What is the longitudinal heterogeneity of nitrogen, phosphorus, and chlorophyll-a along a double reservoir continuum?*

**Hypothesis 1**: I hypothesize that there will be substantial spatial heterogeneity in nitrogen, phosphorus, and chlorophyll-a within both BVR and FCR along a reservoir gradient of riverine to transitional to lacustrine (Figure 3). I expect nitrate (NO3) to follow the traditional view of nutrient decrease along the reservoir gradient due to both sedimentation and biotic uptake (Thornton 1990). In contrast, I expect to find a different pattern with soluble reactive phosphorus (SRP) and NH4, due to the importance of internal loading in our shallow, eutrophic reservoirs which experience substantial anoxia during the summer stratified period (Gerling et al. 2016). I expect that both SRP and NH4 will enter the riverine zone at relatively high concentrations from stream inputs, decrease within the transitional zone due to sedimentation, and increase in the lacustrine zone as a result of internal loading under anoxic conditions. Lastly, I expect to find low chlorophyll-a concentrations in the riverine zone due to high flow and low nutrient conditions from stream inputs, and increases in the transitional zone in response to increased NO3, followed by a further increase in the lacustrine zone as internal SRP and NH4 become abundant. Testing these hypotheses requires intensive sampling along the reservoir continuum to determine if these patterns exist, which would not be discernable from snapshot inflow and outflow sampling.



***Question 2****: How does a gradient of hydrologic flow conditions influence longitudinal heterogeneity?*

**Hypothesis 2:** I hypothesize that increases in flow will increase spatial heterogeneity in both reservoirs (Figure 4). Spatial heterogeneity will be calculated as the coefficient of variation (CV) in chlorophyll-a observations measured at the sampling sites along the reservoir continuum.



**Methods**

*Study site*

Beaverdam Reservoir (BVR) is a small (39 ha), shallow (maximum depth < 13 m), dimictic reservoir owned and operated by the Western Virginia Water Authority (WVWA) and located in Vinton, VA (37.3164747˚N, -79.8188267˚E). BVR has multiple inflow streams, but we have selected the major inflow in the western and the eastern arms for this study (Figure 1, above). BVR has an outflow pipe that flows to Falling Creek Reservoir (FCR), a smaller (~12 ha), shallower (maximum depth = ~9.3 m), dimictic reservoir, also owned and operated by WVWA.

*Proposed Field sampling, summer 2019*

We intend to sample the reservoir continuum once per month from May to September to capture a gradient in seasonal conditions. Additionally, we will add adaptive sampling events to capture a gradient of hydrologic flow conditions (n = ~8 total). We will collect nutrient chemistry (total and soluble nitrogen and phosphorus), chlorophyll-a, and physical characteristic using a YSI sonde (dissolved oxygen, conductivity, and water temperature) of the surface (~0.1m) at all sites. At stream sites, we will also measure discharge using a flowmeter. This proposed sampling leverages an ongoing sampling campaign to capture carbon transformations along the reservoir gradient from the inflows to the dam at FCR.

*Sampling Sites*

Proposed sampling locations can be seen in Figure 5. Sampling for spatial heterogeneity will begin at the two inflow sites to BVR (n = 2) to gauge nutrient and biological inputs to the reservoir. We will sample at 5 sites within BVR to capture a gradient of riverine, transitional, and lacustrine environments, including one site at the outflow pipe that feeds the stream that flows into FCR. Along the stream that flows from BVR to FCR, we will sample at 4 locations. The first location will be immediately after the surfacing of the pipe that flows out of BVR; the second location will be downstream of a confluence between a natural stream and the inflow stream created by BVR’s outflow pipe; the third location will be at a weir where historical sampling has taken place; the fourth location will be immediately before the stream enters into FCR. We will also sample a second inflow to FCR at a wetland site at the northern end of FCR. Within FCR, we will sample 4 sites along a gradient of riverine to lacustrine, as well as at the spillway outflow of FCR. In total, we propose 18 sites along the continuum from BVR to FCR (Figure 5).



*Data analysis*

To determine the heterogeneity of sites within BVR and FCR, we will calculate the coefficient of variation among the sites, among other statistical metrics. We will also develop generalized linear models to analyze drivers of nutrient and phytoplankton dynamics along a stream-reservoir gradient, which will include potential driver variables of physical, chemical, and meteorological variables.

Table 4 outlines my proposed timeline for completing the steps required for Chapter 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4. Proposed timeline for Chapter 2**  *Gray boxes indicate proposed completion* | | | | |
|  | Spring 2019 | Summer 2019 | Fall 2019 | Spring 2020 |
| Set up transects in BVR and FCR |  |  |  |  |
| Summer field sampling |  |  |  |  |
| Nutrient analyses |  |  |  |  |
| Nutrient statistical analyses |  |  |  |  |
| Manuscript writing and edits |  |  |  |  |
| Submit manuscript, possible target journals: L&O, Ecosystems, Biogeosciences |  |  |  |  |

**Citations**

AP-GfK, *Tap Water Confidence Poll*, in AP-GfK Poll 2016, Associated Press-GfK Public

Affairs & Corporate Communications.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Borges, P.A.F., Train, S., Rodrigues, L.C., 2008. Spatial and temporal variation of phytoplankton in two subtropical Brazilian reservoirs 63–74. https://doi.org/10.1007/s10750-008-9367-3

Brookes, J.D., Carey, C.C., 2011. Resilience to Blooms. Science 334, 46–47. https://doi.org/10.1126/science.1207349

Brookmeyer, R., Johnson, E., Ziegler-graham, K., Arrighi, H.M., 2007. Forecasting the global burden of Alzheimer ’ s disease 3, 186–191. https://doi.org/10.1016/j.jalz.2007.04.381

Brown, C.W., Hood, R.R., Long, W., Jacobs, J., Ramers, D.L., Wazniak, C., Wiggert, J.D., Wood, R., Xu, J., 2013. Ecological forecasting in Chesapeake Bay : Using a mechanistic – empirical modeling approach 125, 113–125. https://doi.org/10.1016/j.jmarsys.2012.12.007

Butcher, J.B., Nover, D., Johnson, T.E., Clark, C.M., 2015. Sensitivity of lake thermal and mixing dynamics to climate change. Clim. Change 129, 295–305. https://doi.org/10.1007/s10584-015-1326-1

Calijuri, M.C., Dos Santos, A.C.A., Jati, S., 2002. Temporal changes in the phytoplankton community structure in a tropical and eutrophic reservoir ( Barra Bonita , S . P .— Brazil ).

Carpenter, S.R., Stanley, E.H., Vander Zanden, M.J., 2011. State of the World’s Freshwater Ecosystems: Physical, Chemical, and Biological Changes. Ssrn. https://doi.org/10.1146/annurev-environ-021810-094524

Cooke, G.D., Kennedy, R.H., 2001. Managing Drinking Water Supplies. Lake Reserv. Manag. 17, 157–174. https://doi.org/10.1080/07438140109354128

Dean, C., Roxburgh, S., Mackey, B.G., 2004. Forecasting landscape-level carbon sequestration using gridded , spatially adjusted tree growth 194, 109–129. https://doi.org/10.1016/j.foreco.2004.02.013

Dodds, W.K., Bouska, W.W., Eitzmann, J.L., Pilger, T.J., Pitts, K.L., Riley, A.J., Schloesser, J.T., Thornbrugh, D.J., 2009. Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. Environ. Sci. Technol. 43, 12–19. https://doi.org/10.1021/es801217q

Estes, L.D., Bradley, B.A., Beukes, H., Hole, D.G., Lau, M., Oppenheimer, M.G., Schulze, R., Tadross, M.A., Turner, W.R., 2013. R E S E A RC H Comparing mechanistic and empirical model projections of crop suitability and productivity : implications for ecological forecasting 1007–1018. https://doi.org/10.1111/geb.12034

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, 555–571. https://doi.org/10.1007/s10021-015-9951-0

Gonzalez-benecke, C.A., Martin, T.A., 2017. Pinus taeda forest growth predictions in the 21st century vary with site mean annual temperature and site quality 4689–4705. https://doi.org/10.1111/gcb.13717

Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S., Porter, J.H., 2013. Big data and the future of ecology In a nutshell : https://doi.org/10.1890/120103

Hamre, K.D., Lofton, M.E., McClure, R.P., Munger, Z.W., Doubek, J.P., Gerling, A.B., Schreiber, M.E., Carey, C.C., 2018. In situ fluorometry reveals a persistent, perennial hypolimnetic cyanobacterial bloom in a seasonally anoxic reservoir. Freshw. Sci. 37, 000–000. https://doi.org/10.1086/699327

Harrison, J.A., Maranger, Æ.R.J., Alexander, Æ.R.B., Giblin, A.E., Emilio, Æ.P.J.Æ., Seitzinger, S.P., Sobota, Æ.D.J., Wollheim, Æ.W.M., 2009. The regional and global significance of nitrogen removal in lakes and reservoirs 143–157. https://doi.org/10.1007/s10533-008-9272-x

Hazen, E.L., Palacios, D.M., Forney, K.A., Howell, E.A., Becker, E., Hoover, A.L., Irvine, L., Deangelis, M., Bograd, S.J., Mate, B.R., Bailey, H., 2017. WhaleWatch : a dynamic management tool for predicting blue whale density in the California Current 1415–1428. https://doi.org/10.1111/1365-2664.12820

Hipsey, M.R., Bruce, L.C., Boon, C., Busch, B., Carey, C.C., Hamilton, D.P., Hanson, P.C., Read, J.S., de Sousa, E., Weber, M., Winslow, L.A., 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. Katlenburg-Lindau 12, 473–523. http://dx.doi.org/10.5194/gmd-12-473-2019

Hipsey, M.R., Bruce, L.C., Boon, C., Busch, B., Carey, C.C., Hamilton, D.P., Hanson, P.C., Read, J.S., de Sousa, E., Weber, M., Winslow, L.A., 2017. A General Lake Model (GLM 2.4) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). Geosci. Model Dev. Discuss. 1–60. https://doi.org/10.5194/gmd-2017-257

Jeppesen, E., Kronvang, B., Olesen, J.E., Audet, J., 2011. Climate change effects on nitrogen loading from cultivated catchments in Europe : implications for nitrogen retention , ecological state of lakes and adaptation 1–21. https://doi.org/10.1007/s10750-010-0547-6

Kling, G.W., Kipphut, G.W., Miller, M., Brien, W.J.O., 2000. Integration of lakes and streams in a landscape perspective : the importance of material processing on spatial patterns and temporal coherence.

Lee, R.D., Carter, L.R., Lee, R.D., Modeling, L.R.C., Lee, R.D., Carter, L.R., 1992. Modeling and Forecasting U . S . Mortality Modeling and Forecasting U . S . Mortality 1459. https://doi.org/10.1080/01621459.1992.10475265

Lindegren, M., Mollmann, Christian, Nielsen, A., Brander, K., Mackenzie, B.R., Stenseth, N.C., 2010. Ecological forecasting under climate change : the case of Baltic cod 2121–2130. https://doi.org/10.1098/rspb.2010.0353

Liu, X., Feng, J., Wang, Y., 2019. Chlorophyll a predictability and relative importance of factors governing lake phytoplankton at different timescales. Sci. Total Environ. 648, 472–480. https://doi.org/10.1016/j.scitotenv.2018.08.146

Martínez-meyer, E., Peterson, A.T., Hargrove, W.W., 2004. Ecological niches as stable distributional constraints on mammal species , with implications for Pleistocene extinctions and climate change projections for biodiversity 305–314.

Mcclain, M.E., Boyer, E.W., Dent, C.L., Gergel, S.E., Grimm, N.B., Groffman, P.M., Hart, S.C., Harvey, J.W., Johnston, C.A., Mayorga, E., Mcdowell, W.H., Pinay, G., 2003. Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems 301–312. https://doi.org/10.1007/s10021-003-0161-9

Moss, B., Kosten, S., Meerhoff, M., Battarbee, R.W., Mazzeo, N., Havens, K., Lacerot, G., Liu, Z., Meester, D., Paerl, H., Scheffer, M., Moss, B., Kosten, S., Meerhoff, M., Battarbee, R.W., Mazzeo, N., Havens, K., Lacerot, G., Liu, Z., Meester, L.D., Moss, B., Kosten, S., Meerhoff, M., Battarbee, R.W., Jeppesen, E., Mazzeo, N., 2017. Allied attack : climate change and eutrophication Allied attack : climate change and eutrophication 2041. https://doi.org/10.5268/IW-1.2.359

Perretti, C.T., Munch, S.B., Sugihara, G., 2013. Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data 110, 5253–5257. https://doi.org/10.1073/pnas.1216076110

Powers, S.M., Tank, J., 2015. Control of nitrogen and phosphorus transport by reservoirs in agricultural landscapes. https://doi.org/10.1007/s10533-015-0106-3.CITATIONS

Prein, A.F., Liu, C., Ikeda, K., Trier, S.B., Rasmussen, R.M., Holland, G.J., Clark, M.P., 2017. Increased rainfall volume from future convective storms in the US. Nat. Clim. Change 7. https://doi.org/10.1038/s41558-017-0007-7

Rigosi, A., Marcé, R., Escot, C., Rueda, F.J., 2011. A calibration strategy for dynamic succession models including several phytoplankton groups. Environ. Model. Softw. 26, 697–710. https://doi.org/10.1016/j.envsoft.2011.01.007

Rychtecky, P., Znachor, P., 2011. Spatial heterogeneity and seasonal succession of phytoplankton along the longitudinal gradient in a eutrophic reservoir 175–186. https://doi.org/10.1007/s10750-010-0571-6

Sadro, S., Nelson, C.E., Melack, J.M., 2012. The Influence of Landscape Position and Catchment Characteristics on Aquatic Biogeochemistry in High-Elevation Lake-Chains 363–386. https://doi.org/10.1007/s10021-011-9515-x

Saunders, D.L., Kalff, J., 2001. Nitrogen retention in wetlands , lakes and rivers 205–212.

Scott, J.T., Stanley, J.K., Doyle, R.D., Forbes, M.G., Brooks, B. W., 2009. River – reservoir transition zones are nitrogen fixation hot spots regardless of ecosystem trophic state 61–68. https://doi.org/10.1007/s10750-008-9696-2

Shumway, T., Shumway, T., 2001. Forecasting Bankruptcy More Accurately : A Simple Hazard Model \* 74, 101–124.

Smith, V., 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environ. Sci. Pollut. Res. 10, 126–139. https://doi.org/10.1065/espr2002.12.142

Soares, M.C.S., Marinho, M.M., Azevedo, S.M.O.F., Branco, C.W.C., Huszar, V.L.M., 2012. Limnologica Eutrophication and retention time affecting spatial heterogeneity in a tropical reservoir. Limnologica 42, 197–203. https://doi.org/10.1016/j.limno.2011.11.002

Soranno, P.A., Webster, K.E., Riera, J.L., Kratz, T.K., Baron, J.S., Bukaveckas, P.A., Kling, G.W., White, D.S., Caine, N., Lathrop, R.C., Leavitt, P.R., 1999. Spatial Variation among Lakes within Landscapes : Ecological Organization along Lake Chains 395–410.

Stachelek, J., Soranno, P.A., 2019. Does freshwater connectivity influence phosphorus retention in lakes ? 1–14. https://doi.org/10.1002/lno.11137

Stow, C.A., Roessler, C., Borsuk, M.E., Bowen, J.D., Reckhow, K.H., 2003. Comparison of Estuarine Water Quality Models for Total Maximum Daily Load Development in Neuse River Estuary 129, 307–314.

Thomas, M.K., Fontana, S., Reyes, M., Keho, M., Pomati, F., 2018. The predictability of a lake phytoplankton community , over time-scales of hours to years 619–628. https://doi.org/10.1111/ele.12927

Thuiller, W., Brotons, L., Araujo, M.B., Lavorel, S., 2004. Effects of restricting environmental range of data to project current and future species distributions 2.

White, M.A., Nemani, R.R., 2004. Soil water forecasting in the continental United States : relative forcing by meteorology versus leaf area index and the effects of meteorological forecast errors 30, 717–730.

Wilkinson, M.D., 2016. Comment : The FAIR Guiding Principles for scienti fi c data management and stewardship 1–9. https://doi.org/10.1038/sdata.2016.18

Woodbury, P.B., Smith, J.E., Weinstein, D.A., Laurence, J.A., 1998. Assessing potential climate change effects on loblolly pine growth : A probabilistic regional modeling approach 99–116.