Using Data to Improve Ecological Forecasts:

Pre-class Handout



### Name:

# Macrosystems EDDIE: Using Data to Improve Ecological Forecasts

# Learning Objectives:

By the end of the module, you will be able to:

* Define data assimilation
* Generate an ecological forecast for primary productivity
* Describe how to assess ecological forecast accuracy
* Describe how data assimilation affects forecast accuracy and uncertainty
* Explain how updating models with data collected at different time scales (e.g., daily, weekly) and with different levels of associated uncertainty affects ecological forecasts

# Why macrosystems ecology and ecological forecasting?

**Macrosystems ecology** is the study of ecological dynamics at multiple interacting spatial and temporal scales (e.g., Heffernan et al. 2014). For example, *global* climate change can interact with *local* land-use activities to control how an ecosystem changes over the next decades. Macrosystems ecology recently emerged as a new sub-discipline of ecology to study ecosystems and ecological communities around the globe that are changing at an unprecedented rate because of human activities (IPCC 2013). The responses of ecosystems and communities are complex, non-linear, and driven by feedbacks across local, regional, and global scales (Heffernan et al. 2014). These characteristics necessitate novel approaches for making predictions about how systems may change to improve both our understanding of ecological phenomena as well as inform resource management.

**Forecasting** is a tool that can be used for understanding and predicting macrosystems dynamics. To anticipate and prepare for increased variability in populations, communities, and ecosystems, there is a pressing need to know the future state of ecological systems across space and time (Dietze et al. 2018). Ecological forecasting is an emerging approach which provides an estimate of the future state of an ecological system with uncertainty, allowing society to prepare for changes in important ecosystem services. Ecological forecasts are a powerful test of the scientific method because ecologists make a hypothesis of how an ecological system works; embed their hypothesis in a model; use the model to make a forecast of future conditions; and then when observations become available, assess the accuracy of their forecast, which indicates if their hypothesis is supported or needs to be updated. This process of iteratively updating the forecast with ecological data is called data assimilation, which can greatly improve forecast accuracy over time. Macrosystems ecologists are increasingly using ecological forecasts to predict how ecosystems are changing over space and time (Dietze and Lynch 2019).  
  
In this module, you will generate an ecological forecast for a NEON site and explore how to use ecological data to improve forecast accuracy. This module will introduce you to the concept of data assimilation within an ecological forecast; how data assimilation can be used to improve forecast accuracy; how the level of uncertainty and temporal frequency of observations affects forecast output; and how data assimilation can affect decision-making using ecological forecasts.

# References:

Dietze, M. C., et al. 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences, 115(7), 1424–1432.

Dietze, M., & Lynch, H. (2019). Forecasting a bright future for ecology. Frontiers in Ecology and the Environment, 17(1), 3.

Heffernan, J. B., et al. 2014. Macrosystems ecology: Understanding ecological patterns and processes at continental scales. Frontiers in Ecology and the Environment 12:5–14.

IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. (T. F. Stocker, et al., eds.). Cambridge Univ. Press, NY.

## Module overview:

* Introduction to materials: this handout and PowerPoint in class
* **Activity A**: Students build a model to simulate primary productivity for their chosen NEON site and generate a forecast.
* **Activity B**: Students generate multiple forecasts while assimilating data at different temporal frequencies and with different amounts of observation uncertainty.
* **Activity C**: Student then explore the effects of assimilating sensor data with different levels of observation uncertainty on forecast accuracy and make management decisions using an ecological forecast.

## Today’s focal question:

***How can we use data to improve ecological forecasts?***

To be useful for management, ecological forecasts need to be both accurate enough for managers to be able to rely on them for decision-making *and* include a representation of forecast uncertainty, so managers can properly interpret the probability of future events. To improve forecast accuracy, we can update forecasts with observational data once they become available, a process known as **data assimilation**. Recent improvements in environmental sensor technology and an increase in the number of sensors deployed in ecosystems have resulted in an increase in the availability of data for assimilation to help develop and improve forecasts for natural resource management.

One important test case in which data assimilation may greatly improve the use of ecological forecasts for management is algal blooms in lakes. Algal blooms are increasing in frequency and severity across many freshwater lakes and can substantially impact water quality. Forecasts of the likelihood of future algal blooms may be useful tools in helping water resource managers to mitigate the effect of these blooms. Forecasts can give managers time to take preemptive actions to prevent blooms, such as applying algaecide, or to make plans to reduce the bloom's impact on human health, such as closing recreational beaches.

In this module, we will address the following question: ***How can we use data to improve a forecast of lake algal biomass?***

We will address this question by building an ecological model to predict chlorophyll-a (a metric of algal biomass and primary productivity), generating ecological forecasts, and using data science approaches to integrate the most recently collected data into a forecast, a process known as data assimilation. Assimilating the most recent observations into a forecast model allows the forecaster to update the initial conditions, or the starting conditions, of the model, with the goal of improving forecast accuracy. For example, if our task is to generate a forecast of chlorophyll-a concentrations in a lake for tomorrow, it is likely that our forecast will be more accurate if we use today’s measurement of chlorophyll-a as the initial condition for our model, rather than last week’s measurement.

We will explore how assimilating different types of data (e.g., chlorophyll-a, nutrient concentrations, water temperature) at different temporal frequencies (e.g., daily, weekly) affects forecast output. Finally, we will assimilate different types of data into forecasts and examine how data assimilation affects water resource management decisions.

## Optional pre-class information on ecological forecasts:

* Read a short 1-page commentary: Dietze, M. and Lynch, H. 2019. Forecasting a bright future for ecology. *Frontiers in Ecology and the Environment*, *17*(1), 3. <https://doi.org/10.1002/fee.1994>
* Watch a short video on [Ecological Forecasting: The Science of Predicting Ecosystems](https://youtu.be/Lgi_e7N-C8E?t=63)

After reading through the information above, fill out the following questions in advance of working through the Shiny app.

# Module introduction questions:

1. What is meant by the term 'data assimilation' in the context of ecological forecasting?  
     
   **Answer:**
2. How do you think the process of integrating the most recently observed data into models can improve forecasts?  
     
   **Answer:**

# Exploration

1. Choose one of the example ecological forecasts below and use the website to answer the questions below.

* [USA-NPN Pheno Forecast](https://www.usanpn.org/data/forecasts/): The USA National Phenological Network (NPN) Pheno Forecast delivers short-term (6 day) threshold-based forecasts of phenological events in plants and pest insects.
* [Smart & Connected Water Systems](https://smartreservoir.org/forecasts/): A project which is developing a smart water system that integrates novel high-frequency sensors, cyberinfrastructure, and ecosystem forecasting techniques to improve the management of drinking water supply lakes and reservoirs.
* [EcoCast](https://coastwatch.pfeg.noaa.gov/ecocast/): EcoCast is a fisheries sustainability tool that helps fishers and managers evaluate how to allocate fishing effort to optimize the sustainable harvest of target fish while minimizing bycatch of protected or threatened animals.
* [Grassland Production Forecast](https://grasscast.unl.edu/): Grass-Cast uses almost 40 years of historical data on weather and vegetation growth - combined with seasonal precipitation forecasts - to predict if rangelands are likely to produce above-normal, near-normal, or below-normal amounts of vegetation.
* [Portal Project - Rodent Abundances](https://portal.naturecast.org/index.html): Forecasting a time series of rodent abundances from The Portal project, a long-term experimental monitoring project in desert ecology, 12 months into the future.
  1. Which ecological forecast did you select?  
       
     **Answer:**
  2. What ecological variable(s) are being forecasted?  
       
     **Answer:**
  3. How can this forecast help the public and/or managers?  
       
     **Answer:**
  4. What do you think are some possible sources of uncertainty in this forecast? Try to name at least two possible sources of uncertainty.  
       
     **Answer:**
  5. What data would you need to collect to update the initial conditions (starting conditions) of the forecast? Hint: think about what data you would need to know today’s value of the target forecast variable.

**Answer:**

# Next navigate to “Site Selection: Select a NEON site and visualize the data” in the Shiny app!