Near-term Ecological Forecasting Initiative **Short Course**

College of Arts & Sciences Building
685 Commonwealth Ave
Boston University
Boston, MA 02215

June 26 - July 1, 2022

https://ecoforecast.org/nefi2022/

Course Description

Iterative ecological forecasts are continually-updated, actionable predictions, with uncertainties, of the state of ecosystems and their services. The backbone of this course covers topics related to the statistics of model-data fusion and forecasting: Bayesian statistics; machine learning; uncertainty partitioning, propagation, and analysis; fusing multiple data sources; assessing model performance; and a suite of iterative data assimilation techniques. While the goals of some ecological forecasts are simply to advance basic science, many are performed to inform the public, managers, and decision makers. Therefore the course also covers the basics of structured decision making and expert elicitation.

Course Objectives

After taking this course, you should be able to generate a forecast from a simple model that makes use of more than one data stream. You should be able to categorize and estimate the uncertainties in the model, propagate this uncertainty into the forecast, and assess the performance of the model. You should be able to assess and critique forecast data products and scientific publications about ecological forecasts, and understand how such forecasts can be used to inform decision making.

Diversity and Inclusion

We intend to create a learning environment where participants from diverse backgrounds feel welcomed and supported. We intend to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture. Your suggestions are encouraged and appreciated. Please let us know ways to improve the effectiveness of the course for you personally, or for other participants. We also acknowledge that the course is being held in the territory of the Massachusett People who have stewarded this land for hundreds of generations and we strive to make ecology more inclusive of indigenous knowledge. We expect all participants and instructors to support diversity through at a minimum adhering to the below Code of Conduct.

Code of Conduct

Everyone (participants and instructors) is expected to follow the Ecological Forecasting Initiative's Code of Conduct, which is available in full at https://ecoforecast.org/efi-code-of-conduct/. In summary, we agree to (1) treat each other with kindness, respect, and consideration; (2) empower and create space for one another; (3) welcome and support a diverse community; and (3) express gratitude and recognition. Harassment and discrimination will not be tolerated. Please let an instructor know if you have been subjected to or witnessed unacceptable behavior.

Textbook

Dietze, M. C. (2017). Ecological Forecasting. Princeton, NJ: Princeton University Press.

Course Organizers

Name	Institution	email	cell
Mike Dietze, PI	Boston Univ.	dietze@bu.edu	410-533-5130
Shannon LaDeau, coPI	Cary Institute	ladeaus@caryinstitute.org	919-323-5983
Jenny Bhatnagar, coPI	Boston Univ.	jmtalbot@bu.edu	
Jake Zwart	USGS	jzwart@usgs.gov	
Kira Sullivan Wiley	Pew Charitable Trusts	kira.s.wiley@gmail.com	
John Foster, TA	Boston Univ.	fosterj@bu.edu	651-558-7362
Kathryn Wheeler, TA	Boston Univ.	kiwheel@bu.edu	828-773-9737
Zoey Werbin, TA	Boston Univ.	zrwerbin@bu.edu	954-496-2100
Charlotte Malmborg, TA	Boston Univ	malmborg@bu.edu	802-310-4177

Thursday Round Table

Discussion of issues that span academia, agencies, and non-traditional research both generally (e.g. career options, publishing, funding, etc) and in the specific context of trying to build a community of practice in ecological forecasting. Where is ecological forecasting going, how do we get it there?

Schedule

Sun 1-5pm, CAS-426	CAS-426	STO-453	STO-453	STO-453	CAS-426
Time	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
9:00	Introductions (0:40)	State-space (0:30)	Analytical DA (1:15)	Ensemble DA (1:15)	Model Assessment (1:00)
9:30	Lightning talks (0:35)	Hands On			
10:00	Project descriptions & polling (0:20)		Break (0:15)	Break (0:15)	Break (0:15)
10:30	Break	Break	Hands On (1:15)	Hands On (1:00)	Forecast Infrastructure (1:15)
11:00	Characterizing Uncertainty (0:30)	Dynamic Models (0:30)			
11:30	Hands On	Project	PROACT (1:00)	Human dimensions of ecological forecasting (1:00)	Hands On (1:00)
12:00					
12:30	Lunch	Lunch	Lunch	Lunch	Lunch
13:00					
13:30	Hierarchical Bayes (1:15)		Machine Learning (1:15)	Project	Project
14:00		Propagating Uncertainty (1:15)			
14:30	Break (0:15)		Break (0:15) at 15:45		Break (0:15)
15:00	hands on (1:15)	Break (0:15)	Project		Project Presentations (1:00)
15:30		Hands On (1:30)		Break (0:15)	
16:00	Break (0:15)			Round Table (1:00)	Wrap up (0:45)
16:30	Expert Elicitation (0:30)				

Hands-on Activities
The goal of the course isn't just to expose you to the concepts of ecological forecasting, but to give you hands-on experience in applying the tools used. Therefore most lectures are paired with hands-on activities. Most of these activities are independent (you don't need to complete one to move on to the next) and many are open-ended -- there's more you could play with than we expect you to be able to complete in the allocated time. The instructors will often state what the minimal goal of each activity is. Beyond that, we encourage you to complete what you can, play with the activities, and know that you can come back to the activities that you find most important and interesting at a later date. We will split you into randomized break out groups for each of the hands on activities. Each group with have TAs and instructors to assist with the activity.

Group Projects
The goal of the group project is to take what you've learned in lecture and the hands-on activities and see if you can apply it to a new problem. Students will work in small groups for the final project and make a 10-15 min presentation about what they've accomplished this week on Friday afternoon. The goals of the project are:

- **Calibration**: Fit a simple dynamic model to time-series data, preferably using a Bayesian state-space framework and accounting for the complexities of the data. Make sure to hold out some portion of the available data for validation.
- **Forecast**: Predict or project the dynamics of the system forward from the calibration period into the validation period while accounting for uncertainties. If time permits, partition the forecast uncertainties to determine which dominates.
- **Validation**: Assess model performance against held-out data
- **Analysis:** For the next time point beyond the calibration period, use data assimilation to update the state of the system and re-forecast. If time permits, iteratively assimilate all validation data. How does the uncertainty in your forecast change with lead time and what is the forecast horizon for this system?

The set of projects available to participate in has been pre-selected based on the Ecological Forecasting Initiative's NEON Forecasting Challenge. The set of available projects will be described Monday morning. Everyone will be asked to rank their interest in the available projects and then will be assigned to a team by the instructors

R / RStudio
The course activities will be done in R and will be provided as R Markdown (Rmd) documents. Use of RStudio is not required, but activities are prepared assuming you will be working in that IDE.

Git / Github
Course activities and in-house R packages will be distributed via Github, so at a minimum you

will be expected to be able to clone / download those repositories (we'll be sending a follow up email in a few weeks with more details on exactly what repositories to install). Bare-bones basics on how to use Github to retrieve course materials is located at the top of our basic R tutorial: https://github.com/EcoForecast/EF_Activities/blob/master/Exercise_01_RPrimer.Rmd

While not required just to download, you'll probably want to create an account at https://github.com/ if you don't have one already.

You are also strongly encouraged to use Github (or equivalent) when collaborating on code for your course projects. We also have a tutorial on using Github for collaborative coding, which is best completed with a partner:
https://github.com/EcoForecast/EF_Activities/blob/master/Exercise_04_PairCoding.Rmd

Code & data management
These are covered in chapters 3 and 4 of the book, which we won't be covering in class.

You may also enjoy:
Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., & Teal, T. K. (2016). Good Enough Practices in Scientific Computing. PLoS Computational Biology, 13, 1–20. http://doi.org/10.1371/journal.pcbi.1005510

Wilson, G., Aruliah, D. A., Brown, C. T., Chue Hong, N. P., Davis, M., Guy, R. T., Haddock, S. H. D., Huff, K. D., Mitchell, I. M., Plumbley, M. D., Waugh, B., White, E. P., & Wilson, P. (2014). Best Practices for Scientific Computing. PLoS Biology, 12(1), e1001745. https://doi.org/10.1371/journal.pbio.1001745

Bayesian statistics
In addition to basic concepts, we're also going to assume basic familiarity with Bayesian software. Specifically, we will be using JAGS (called through R) for the course tutorials, so you will want to make sure you have install the JAGS software ahead of time (http://mcmc-jags.sourceforge.net/). As a point of reference, we'd like everyone to be able to implement a basic Bayesian linear regression model in JAGS (this is the milestone we aim to reach in the optional Bayes primer).

Software
As mentioned previously we'll be using R / RStudio, JAGS, and Git / Github so you'll want to make sure that those are installed on your laptop prior to the course.

Furthermore, all the course activities are in the following repository:

https://github.com/EcoForecast/EF_Activities

Which you'll want to 'git clone' (or fork to your local repo) before the course starts. A number of activities do rely on additional packages. Most of these are quick & easy, but do make sure you have 'rjags' installed and working.

Readings

From the *Ecological Forecasting* textbook, we're going to assume Ch 3-5 as prior knowledge and will focus the course on Chapters 2, 6, 8, 9, 11, 13, 14, and 17. We're not requiring that folks have read these chapters ahead of time, but if you do want to start diving in ahead of time that's where we recommend you focus. Any advanced reading you're able to do will make following lectures and doing the activities that much easier.

In addition to the textbook we've provided the following readings as additional references:

Optional Expert Elicitation Reading

Morgan, M. Granger. 2014. Expert elicitation for decision making. Proceedings of the National Academy of Sciences,111 (20) 7176-7184; DOI: 10.1073/pnas.1319946111

Optional ML Readings

Joseph, M. B. (2020). Neural hierarchical models of ecological populations. Ecology letters, 23(4), 734-747. https://onlinelibrary.wiley.com/doi/full/10.1111/ele.13462

Nearing, G. S., Klotz, D., Sampson, A. K., Kratzert, F., Gauch, M., Frame, J. M., Shalev, G., & Nevo, S. (2021). Technical Note: Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks [Preprint]. Catchment hydrology/Modelling approaches. https://doi.org/10.5194/hess-2021-515

Klotz, D., Kratzert, F., Gauch, M., Keefe Sampson, A., Brandstetter, J., Klambauer, G., Hochreiter, S., & Nearing, G. (2021). Uncertainty Estimation with Deep Learning for Rainfall–Runoff Modelling [Preprint]. Catchment hydrology/Modelling approaches. https://doi.org/10.5194/hess-2021-154

Workflow Reading

Ecological Forecasting Initiative's Reproducible Forecasting Workflows task view https://projects.ecoforecast.org/taskviews/reproducible-forecasting-workflows.html

Optional Workflow Reading

Patil, P., Peng, R. D., & Leek, J. T. (2016). A statistical definition for reproducibility and replicability [Preprint]. https://doi.org/10.1101/066803