Bayesian Statistics and Biological Forecasting
BIOS 40552/60552
Fall 2022

Time: MW 14:00-15:15
Location: Jordan 322

Professor: Jason McLachlan
Galvin 176
jmclachl@nd.edu

Group work time: Friday 2-3pm (Galvin 298)
Office Hours: By appointment

Course Web Page: **https://canvas.nd.edu/courses/52021**

Course Description:

Ecological forecasting is the process of predicting the state of ecosystems, ecosystem services, and natural capital, with fully specified uncertainties, and contingent on explicit scenarios for climate, land use, human population, technologies, and economic activity. The backbone of this course covers topics related to the statistics and informatics of model-data fusion and forecasting: data management, workflows, Bayesian statistics, uncertainty partitioning, propagation, and analysis, fusing multiple data sources, assessing model performance, scenario development, decision support, and a suite of data assimilation techniques.

Prerequisite:

- It will help to have some experience coding in R (or another language), but if you've got moxie you can pick this up along the way.
- Familiarity with basic statistics (e.g., Biostats) will help.

Course Objectives:

After taking this course, you should be able to generate an operational 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 explain the following: the model, the data, the forecast, and the best practices employed for managing all three. You should be able to assess and critique forecast data products and scientific publications about ecological forecasts.

Technology:

Slack:

Students are encouraged to use Slack as part of the course, such as to ask question, discuss hands-on activities, and coordinate the semester project. The primary course slack will be the #ndbio4cast channel within the "JMc courses" Slack
Invite link: https://join.slack.com/t/nd-u3x9788/shared_invite/zt-1ev55z1jn-IvTrRKMuPpvwy3yfvac5Zg

In addition to the course-specific Slack, you are strongly encouraged to join the Ecological Forecasting Initiative's Slack, ecoforecast.slack.org , as part of the semester project. In particular the #neon4cast_courses channel is a shared space for students participating in the forecasting challenge (across universities) and there are individual #neon4cast channels for each specific challenge. If your interest in forecasting grows you might also check out some of the working group channels or the EFI student association #students.
Invite link: https://join.slack.com/t/ecoforecast/shared_invite/zt-l5od1yk1-Ulc777kMCWT0H~YClpt4Jg

Laptops/tablets:

Students are encouraged to bring laptops and/or tablets for the purpose of taking notes, bringing papers for discussions, etc. **Laptops are required from most hands-on activities**.

R:

Most hands-on activities and final projects will employ R for data management, analysis, and modeling. R instructions will be given assuming the user is running the RStudio editor. Therefore **it would also be good to install R** http://www.r-project.org **and RStudio** http://www.rstudio.com. A few other pieces of software will be used for individual in-class activities and introduced on a case-by-case basis. The first hands-on activity for the class is a R tutorial, which includes references to other R resources

Git/Github:

Similarly, most hands-on activities and final projects will make extensive use of Git (for version control) and Github (https://github.com) (for project management), which not coincidentally is also integrated with RStudio. Make sure that you have integrated RStudio and GitHub, as per the first class exercise.

If you are unfamiliar with Git and Github good place to start is the GitHub 5 minute illustrated tutorial. In addition, there are three fun tutorials for learning git:

- Learn Git is a great web-based interactive tutorial.
- LearnGitBranching
- TryGit.

Also note that Hands-on Activity 4 will focus on learning how to use Git collaboratively

Course Format & Outline:

Topics and readings are subject to change. Changes will be announced in class and posted on the course website.

Day	Topic	Activity/Reading (O = optional) (PM = Project Milestone)
8/24	Overview	
8/29	Dynamic Models	Dietze Chapter 1 1. R primer Chapter 1 O: Google R tutorials O: Git/Github tutorials
8/31	From models to forecasts	Chapter 2 O: (Currie 2019)
9/5	Predictability	2. From Models to Forecasts PM: Project Preference O: Dietze et al (2018)
9/7	Data large and small	Chapter 3 O: Data One Best Practices
9/12	Forecast challenge: Data Streams	3. Big Data
9/14	Informatics of Model-Data Fusion	Chapter 4 O: (Read et al. 2016)
9/19	Workflows	Pair Coding
9/21	Forecast Challenge [Quinn Thomas]	NEON Forecast Challenge PM: Dynamic models
9/26	Midterm Exercise	NEON forecast example
9/28	Intro to Bayes	Chapter 5 O: (Ellison 2004) O: Bayesian Regression
10/3	BUGS/JAGS	JAGS Primer PM: Set up repository
10/5	Characterizing Uncertainty	Chapter 6 O: Fitting Uncertainties O: Hierarchical Bayes
10/10	State-space models	Chapter 8 6. State Space
10/12	Fusing data sources	Chapter 9 PM: Pull & Visualize data
10/24	Data-Fusion	8. Tree Rings O: (Clark et al. 2007)
10/26	Propagating Uncertainty	Chapter 11
10/31	Uncertainty Partitioning	Uncertainty Partitioning PM: Fit historical time-series O: (Dietze 2017)
11/2	Data Assimilation: Analytical	Chapter 13 Disciplinary assessment due.

		O: (Wikle and Berliner 2007)
11/7	Kalman Filter	9. Kalman Filter
11/9	Data Assimilation: Monte Carlo	Chapter 14 PM: Initial ensemble forecast O: (Evensen 2009)
11/14	Ensemble KF	10. Particle Filter
11/16	Assessing Model Performance	Chapter 16
11/21	Model Assessment	11. Model Assessment
11/28	Scenarios & Decision Support	Chapter 17 PM: Teamwork Reflection
11/30	Stakeholder exercise	12. Decision support
12/5	Forecast Ethics	Essays due before class.
12/7	Final Thoughts/ Project presentations	Chapter 18 PM: Present Iterative Forecast O: (Dietze et al. 2018)
12/15	Project Report Due	PM: Final Project Report

Grading:

Hands-on activities	48
Midterm Exercise	7
Disciplinary assessment	8
Essay	7
Final Project	30*
	100

* Project grade:
30 pts = 2 pts per milestone (through 11/28) + 4 pts for presentation + 10 pts for report

This class follows the binding Code of Honor at Notre Dame. The graded work you do in this class must be your own. In the case where you collaborate with other students make sure to fairly attribute their contribution to your project.

Hands-on Activities **~ weekly**

The requirements for the hand-on activities will vary from week-to-week but generally involve the exploration and application of a new tool or technique. These assignments need to be **submitted on CANVAS by the following Sunday**.

Research Project **numerous milestones, Submitted by 12/15**

The goal of the final research product is to generate an automated forecast for a specified ecosystem state. This year projects will be selected from among the forecast areas of the Ecological Forecasting Initiative's NEON Forecasting Challenge. At the start of the project you will begin by developing an uncalibrated process model that could be informed by one or more publicly available data streams. You will calibrate and validate the model against training data, assess and quantify the sources of uncertainty that contribute to the forecast, and then produce an initial forecast. After this you will generate of updates to your forecast given new data and assess the skill of your previous forecasts against this new data. Forecasts will be embedded in a workflow and the instructor must be able to further update the forecast using this workflow (i.e. the analysis needs to be transparent and repeatable).

Students will work in small groups for the final project and students are strongly encouraged to work on the project throughout the semester as they are introduced to each topic. The overall project is broken up into a series of smaller milestones and due dates throughout the semester, as detailed in the course project document. Each milestone will be worth 2 points. The 'final' for the course will be a brief classroom presentation (10 min) on 12/7 (4 points) and a five-page write-up describing and assessing your forecast due 12/15 (10 points).

Midterm exercise 9/26/2022

By the end of September, you will have worked through the logistical steps of implementing a biological forecast. Your forecast team will complete the NEON Ecological Forecast Challenge example (described here) and submit an .Rmd and knitted .html description of the exercise with sufficient comments describing the significance and interpretation of each step in the forecasting process to a smart but naïve reader (imagine yourself as a high school senior). The Example code and details of the assignment are under the Assignments tab in Canvas. For this exercise, you should work alone.

Disciplinary assessment 11/2/2022

Every discipline faces different challenges in making scientific inference from data (for example, data are sparse and noisy in some fields, while integrating enormous streams of big data are the challenge in others). Identify recent paper in your field that you feel is at the cutting edge of integrated statistical analysis and use concrete examples from this paper to critically evaluate the state of your field in the characterization, integration, propagation, analysis, and reduction of uncertainty. Frame this in the setting of forecasting (5 page paper)

Forecast Ethics Essay 12/5/2022

Forecasting is a relatively new discipline, which means: (A) The inevitable ethical challenges each scientific discipline faces are only emerging now; & (B) There is a short window of opportunity to define and establish high ethical standards for forecasting. You will write a 3 - 5 page essay exploring an issue of importance to establishing these standards in the field. The essay will include references to literature both in forecasting and in related fields of scientific ethics. These essays will be shared with the class and the subject of a class discussion later in the semester.

Attendance, make-ups, and late work:

Students are expected to attend class. If you have to miss class, let me know.
Points will be deducted at 10% per day for any unexcused late work.

Textbook:

The textbook for the class should be available from the campus bookstore or online (Amazon, Princeton University Press, etc)

Dietze MC. *Ecological Forecasting*. Princeton, NJ: Princeton University Press; 2017.

Readings (many here are optional):

Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke Jr, W. Pizer, C. Pringle, and W. V Reid. 2001. Ecological forecasts: an emerging imperative. Science 293:657–60.

Clark, J. S., M. S. Wolosin, M. C. Dietze, I. Ibáñez, S. LaDeau, M. Welsh, and B. Kloeppel. 2007. Tree growth inference and prediction from diameter censuses and ring widths. Ecological Applications 17:1942–53.

Currie, D. J. 2019. Where Newton might have taken ecology. Global Ecology and Biogeography 28:18–27.

Dietze, M. C. 2017. Prediction in ecology: a first-principles framework. Ecological Applications 27:2048–2060.

Dietze, M. C., A. Fox, L. M. Beck-Johnson, J. L. Betancourt, M. B. Hooten, C. S. Jarnevich, T. H. Keitt, M. A. Kenney, C. M. Laney, L. G. Larsen, H. W. Loescher, C. K. Lunch, B. C. Pijanowski, J. T. Randerson, E. K. Read, A. T. Tredennick, R. Vargas, K. C. Weathers, and E. P. White. 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences 115:1424–1432.

Ellison, A. M. 2004. Bayesian inference in ecology. Ecology Letters 7:509–520.

Evensen, G. 2009. The Ensemble Kalman Filter for Combined State and Parameter Estimation. IEEE Control Systems Magazine:83–104.

Hobday, A. J., J. R. Hartog, J. P. Manderson, K. E. Mills, M. J. Oliver, A. J. Pershing, and S. Siedlecki. 2019. Food for Thought Ethical considerations and unanticipated consequences associated with ecological forecasting for marine resources.

Ibanez, I., J. M. Diez, L. P. Miller, J. D. Olden, C. J. B. Sorte, D. M. Blumenthal, B. A. Bradley, C. M. D'Antonio, J. S. Dukes, R. I. Early, E. D. Grosholz, and J. J. Lawler. 2014. Integrated assessment of biological invasions. Ecological Applications 24:25–37.

Luo, Y., J. T. Randerson, G. Abramowitz, C. Bacour, E. Blyth, N. Carvalhais, P. Ciais, D. Dalmonech, J. B. Fisher, R. Fisher, P. Friedlingstein, K. Hibbard, F. Hoffman, D. Huntzinger, C. D. Jones, C. Koven, D. Lawrence, D. J. Li, M. Mahecha, S. L. Niu, R. Norby, S. L. Piao, X. Qi, P. Peylin, I. C. Prentice, W. Riley, M. Reichstein, C. Schwalm, Y. P. Wang, J. Y. Xia, S. Zaehle, and X. H. Zhou. 2012. A framework for benchmarking land models. Biogeosciences 9:3857–3874.

Memarzadeh, M., and C. Boettiger. 2019. Resolving the measurement uncertainty paradox in ecological management. Arxiv preprint.

Miller, S., A. Rassweiler, L. Dee, K. M. Kleisner, T. Mangin, R. Oliveros-ramos, J. Tam, F. P. Chavez, M. Ñiquen, S. E. Lester, M. Burden, S. Gaines, and C. Costello. 2019. Optimal harvest responses to environmental forecasts depend on resource knowledge and how it can be used. Canadian Journal of Fisheries and Aquatic Sciences.

Morgan, M. G. 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. Proceedings of the National Academy of Sciences of the United States of America 111:7176–84.

Petchey, O. L., M. Pontarp, T. M. Massie, S. Kéfi, A. Ozgul, M. Weilenmann, G. M. Palamara, F. Altermatt, B. Matthews, J. M. Levine, D. Z. Childs, B. J. McGill, M. E. Schaepman, B. Schmid, P. Spaak, A. P. Beckerman, F. Pennekamp, and I. S. Pearse. 2015. The ecological forecast horizon, and examples of its uses and determinants. Ecology Letters 18:597–611.

Powers, S. M., and S. E. Hampton. 2019. Open science, reproducibility, and transparency in ecology. Ecological Applications 29:1–8.

Read, E., M. Rourke, G. Hong, P. Hanson, L. Winslow, S. Crowley, C. Brewer, and K. Weathers. 2016. Building the team for team science. Ecosphere 7:1–9.

Wikle, C., and L. Berliner. 2007. A Bayesian tutorial for data assimilation. Physica D: Nonlinear Phenomena 230:1–16.