Can you predict the future? A tutorial for the National Ecological Observatory Network Ecological Forecasting Challenge

Freya Olsson [1,2], **Carl Boettiger** [3], **Cayelan C. Carey** [1,2], **Mary E. Lofton** [1,2], **and R. Quinn Thomas** [1,2,4]

1 Center for Ecosystem Forecasting, Virginia Tech, Blacksburg, Virginia, USA **2** Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA **3** Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, California, USA **4** Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, USA

Summary

This tutorial introduces participants to key concepts in ecological forecasting and provides hands-on materials for submitting forecasts to the National Ecological Observatory Network (NEON) Forecasting Challenge (hereafter, Challenge), hosted by the Ecological Forecasting Initiative Research Coordination Network. The tutorial has been developed and used with >300 participants and provides the ecological understanding, workflows, and tools to enable ecologists with minimal forecasting experience to participate in the Challenge via a hands-on R-based tutorial. This tutorial introduces participants to a near-term, iterative forecasting workflow that includes obtaining observations from NEON, developing a simple forecasting model, generating a forecast, and submitting the forecast to the Challenge, as well as evaluating forecast performance once new observations become available. The overarching aim of this tutorial is to lower the barrier to ecological forecasting and empower participants to develop their own ecological forecasts.

Statement of need

Ecological forecasting is an emerging field that aims to improve natural resource management and ecological understanding by providing future predictions of the state of ecosystems (Dietze, 2017; Lewis et al., 2023). Generating ecological forecasts requires a suite of quantitative and computational skills, including accessing real-time data, building ecological models, quantifying uncertainty associated with predictions, generating forecasts, and updating models with new observations as they become available (Dietze, 2017). While resources to educate ecologists on these skills, individually, are available (Ernest et al., 2023), there are few hands-on demonstrations of how to implement a full workflow to generate a near-term forecast. In response to this gap, we designed this tutorial for ecologists who are interested in learning about ecological forecasting through hands-on instruction but may not have prior experience in this domain. The tutorial is also designed for individuals that are interested in submitting forecasts to the Challenge but may not know how to start generating forecasts. Altogether, this tutorial provides a framework that can be modified to generate forecasts, thereby increasing participation in the Challenge and expanding our understanding of environmental predictability.

Background and development

The NEON Forecasting Challenge aims to create a community of practice that builds capacity for ecological forecasting by leveraging recently-available NEON data (Thomas et al., 2023). The Challenge revolves around five themes (Aquatics, Terrestrial, Phenology, Beetles, Ticks) that span aquatic and terrestrial systems, and population, community, and ecosystem processes across 81 NEON sites across the U.S. The motivation of the Challenge is for teams and individuals to forecast the conditions at NEON sites before the data are collected. Challenge forecasts are automatically evaluated against observations when they become available. By collating forecasts from many different models and sites, the Challenge organisers and participants aim to quantify how ecological predictability varies over space and time.

This 90-minute tutorial was initially developed for a workshop at the 2022 Global Lakes Ecological Observatory Network (GLEON) All-Hands' conference, which had participants with a range of forecasting and coding experience. The GLEON workshop was given on the first day of the five-day conference. This timing enabled forecasts submitted at the workshop to be evaluated throughout the conference, allowing participants to see near-real time forecast performance, and for a "winner" to be declared on the final day. The tutorial has since been taught in nine workshop/classroom settings (Table 1).

Table 1: Table 1 Implementation of the tutorial by the authors across a range of settings. The participants in these workshops covered a wide range of forecasting and coding experience.

Name of meeting or group delivered to	Approximate number of participants	Modality	Challenge theme
GLEON* All-Hands conference	50	In-person	Aquatics
InventWater PhD training programme	15	Synchronous on-line	Aquatics
GLEON All-Hands' Virtual conference	70	Asynchronous on-line (pre-recorded)	Aquatics
AEMON-J**/DSOS*** Hacking Limnology	70	Synchronous on-line	Aquatics
Global Change Ecology Lab, University of Edinburgh	10	Synchronous on-line	Terrestrial
NEON Technical Working Group on Ecological Forecasting	10	Synchronous on-line	Terrestrial
Ecological Society of America conference	50	In-person	Terrestrial
Graduate environmental data science class	40	In-person	Terrestrial/Aquatics
Ecological Forecasting Initiative conference	20	In-person	Aquatics

*Global Lake Ecological Observatory Network; **Aquatic Ecosystem MOdeling Network - Junior; ***Data Science Open Science

Audience

The audience for this tutorial includes individuals who: 1) want to participate in the Challenge but are not sure how to start; 2) want a 'hands-on' way to learn about ecological forecasting ; and/or 3) are involved in the broader forecasting enterprise (e.g., researchers collecting data used for forecasting) and want to submit forecasts themselves. We encourage users to modify the tutorial as needed, as all materials are open-source.

Features

Learning objectives

The overarching objectives of the tutorial are:

1. Build an understanding of foundational ecological forecasting concepts;
2. Apply forecasting concepts to submit a simple forecast to the Challenge; and
3. Learn about additional forecasting resources.

These objectives can be adapted depending on the context of the tutorial. If the participant/instructor's goals are geared towards understanding forecasting concepts then the emphasis of the presentation and hands-on workshop can be modified accordingly.

Instructional design

The R-based tutorial is in a public GitHub repository (Olsson et al. (2024); https://github.com/eco4cast/NEON-forecast-challenge-workshop) that includes an introductory presentation (Microsoft PowerPoint or PDF format), Rmarkdown documents, rendered versions of the markdown files (.md), as well as pre-tutorial instructions for participants. The tutorial allows for both in-person and virtual participation (Table 1) and can be completed synchronously in an instructor-led workshop/course or asynchronously in a self-paced tutorial.

The tutorial

The tutorial is designed as a 90-minute, standalone session that includes pre-tutorial materials, an introductory presentation (20-30 minutes), a guided demonstration of forecast code (30-40 minutes), and discussion (20-30 minutes). If more time is available, the tutorial has additional content that includes more advanced topics (Figure 1). Details on each of these sections is detailed in the workshop's README.md at the workshop GitHub repository (Olsson et al. (2024); https://github.com/eco4cast/NEON-forecast-challenge-workshop).

1. Introductory presentation: introduces the participants to forecasting concepts, the Challenge, NEON data, and tools that will be used in the R coding portion of the tutorial. This presentation can be tailored to the audience based on their familiarity with forecasting concepts and NEON data (Figure 1).

2. Coding walk through: participants walk through a pre-written Rmarkdown forecast workflow script. This code is written primarily using `tidyverse` syntax to optimise readability (Wickham et al., 2019). The tutorial utilises the functions from the `neon4cast` R package, developed by Challenge organisers to ease access to weather covariate data and simplify forecast submission (see https://github.com/eco4cast/neon4cast).

3. Open time for discussion: the remaining time can be used for multiple purposes (detailed in the R markdown and README) depending on the interests of the participants (e.g. debug code issues, modify forecast models or form teams to submit additional forecasts to the Challenge).

4. Optional extension: to extend the tutorial beyond 90 minutes (Figure 1), we provide additional materials that show participants how to automate forecast submission (directory `Automate_Forecast`) and introduce participants to forecast evaluation and synthesis (Thomas et al. (2023)). These materials could also be used in a self-paced manner after the workshop.

Reuse, implementation, and modification

The primary tutorial focuses on the Aquatics theme (specifically, water temperature) of the Challenge as an example of how to generate a forecast, though the tools and workflows are applicable to all Challenge themes. For example, we adapted the materials to the Terrestrial theme (see Table 1) based on the interests of the participants and have further modified the materials for other forecasting challenges.

Several lessons learned have emerged from earlier implementations of the tutorial (Table 1). First, engagement in the Challenge post-tutorial is best when there is an opportunity for follow-up discussion, troubleshooting, and continuation of team collaboration beyond 90 minutes. Second, we found that providing installation instructions and preparatory material in advance promoted best engagement during synchornous and in-person workshops. Third, the introductory presentation can be adapted to meet the needs and experience level of the participants (Table 1). Finally, the tutorial requires a stable and relatively fast internet connection. We found that slow internet speeds limited access to downloading weather forecasts used to generate ecological forecasts. This issue can be addressed by either having the rendered Rmarkdown document available so that individuals can follow along even if connectivity becomes an issue or providing access to remote computational environments.

•Open time refers to unstructured time available for partipants to troubleshoot code, explore additional forecasting methods, and/or modify and expand the forecasting model.

Figure 1: Figure 1 Potential workshop/course structures using this tutorial. The original 90-minute workshop setup is shown as the "regular tutorial" which can be expanded and modified according to the time available, the anticipated skills and background of the participants and the goals of participation. The alternate modes of delivery were delivered from administering the tutorial to audiences of mixed coding and forecasting experience shown in Table 1.

Acknowledgments

enthusiasm, interest, and feedback that helped us iteratively improve the materials (like our forecasts!).

References

Dietze, M. C. (2017). *Ecological Forecasting*. Princeton University Press. https://doi.org/10.1515/9781400885459

Ernest, S. K. M., Ye, H., & White, E. P. (2023). Ecological Forecasting and Dynamics: A graduate course on the fundamentals of time series and forecasting in ecology. *Journal of Open Source Education*, *6*(66), 198. https://doi.org/10.21105/jose.00198

Lewis, A. S. L., Rollinson, C. R., Allyn, A. J., Ashander, J., Brodie, S., Brookson, C. B., Collins, E., Dietze, M. C., Gallinat, A. S., Juvigny-Khenafou, N., Koren, G., McGlinn, D. J., Moustahfid, H., Peters, J. A., Record, N. R., Robbins, C. J., Tonkin, J., & Wardle, G. M. (2023). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution*, *14*(3), 746–756. https://doi.org/10.1111/2041-210X.13955

Olsson, F., Thomas, R. Q., Boettiger, C., & Lofton, M. E. (2024). *OlssonF/NEON-forecast-challenge-workshop: EFI NEON Forecast Challenge Workshop*. Zenodo. https://doi.org/10.5281/zenodo.8316965

Thomas, R. Q., Boettiger, C., Carey, C. C., Dietze, M. C., Johnson, L. R., Kenney, M. A., McLachlan, J. S., Peters, J. A., Sokol, E. R., Weltzin, J. F., Willson, A., & Woelmer, W. M. (2023). The NEON Ecological Forecasting Challenge. *Frontiers in Ecology and the Environment*, *21*(3), 112–113. https://doi.org/10.1002/fee.2616

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686