The NEON Ecological Forecasting Challenge

R. Quinn Thomas[1], Carl Boettiger[2], Cayelan C Carey[1], Michael C Dietze[3], Leah R Johnson[1], Melissa A Kenney[4], Jason S Mclachlan[5], Jody A Peters[5], Eric R Sokol[6], Jake F Weltzin[7], Alyssa Willson[5], Whitney M Woelmer[1], and Challenge Contributors[8]

[1]Virginia Tech
[2]University of California -Berkeley
[3]Boston University
[4]University of Minnesota
[5]University of Notre Dame
[6]National Ecological Observatory Network
[7]U.S. Geological Survey
[8]Affiliation not available

December 11, 2022

Title:

The NEON Ecological Forecasting Challenge

Authors:

R. Quinn Thomas[1], Carl Boettiger[2], Cayelan C. Carey[1], Michael C. Dietze[3], Leah R. Johnson[1], Melissa A. Kenney[4], Jason S. McLachlan[5], Jody A. Peters[5], Eric R. Sokol[6], Jake F. Weltzin[7], Alyssa Willson[5], Whitney M. Woelmer[1], and Challenge contributors[8].

Affiliations:

1. Virginia Tech, 2. University of California - Berkeley, 3. Boston University, 4. University of Minnesota, 5. University of Notre Dame, 6. National Ecological Observatory Network, 7. U.S. Geological Survey, 8. See Supplemental Material for additional authors

Main text:

The 21st century has been - and will continue to be - characterized by major changes to the environment and the ecosystem services upon which society depends. Anticipating and responding to these changes requires that we make explicit forecasts of future conditions in real time (Dietze *et al.* 2018). Ecological forecasting, like weather and epidemiology forecasting, involves integrating data and models to generate *quantitative* predictions of the future state of ecological systems before observations are collected. It thus represents a rigorous out-of-sample test of our foundational understanding of ecological dynamics that are embedded in our models (Dietze *et al.* 2018), while also providing predictions that can be used for decision making. The iterative cycle of creating ecological forecasts, evaluating them with new observations, updating models, and then making new forecasts that are integrated into adaptive management has the potential to accelerate learning across many disciplines within ecology. This iterative cycle builds on openly available data, often published soon after collection, as is increasingly common in ecological observatory networks, such as the National Ecological Observatory Network (NEON). To accelerate the development of theory, training, and technology in ecological forecasting, we have designed and launched the NEON Ecological Forecasting Challenge (Figure 1), an open platform for the ecological and data science communities to forecast NEON data before they are collected.

The ecological forecasting community is interested in using forecasts to advance theory in ecology (Lewis *et al.* 2022) and in the translation of forecasts for natural resource management (Enquist *et al.* 2017). For example, a fundamental question asks how limits to predictability (e.g., the time in the future where the forecast is no better than a simple baseline) differs across population, community, and ecosystem dynamics (Lewis *et al.* 2022). This question (and many more) is an example of what we can uniquely learn by analyzing a catalog of forecasts developed for a range of ecological systems, scales, and environmental gradients. The Ecological Forecasting Initiative Research Coordination Network (EFI-RCN) - funded by the National Science Foundation (NSF) - invites the ecological research community to help build this catalog by forecasting NEON data. NEON is a powerful platform to support a forecasting

challenge because it provides standardized data with reported uncertainties that span a range of environmental conditions across the United States, and that includes a broad range of terrestrial and freshwater population, community, and ecosystem dynamics.

The NEON Ecological Forecasting Challenge was designed based on input from the academic, government, and private sectors through workshops, working groups, and early forecast contributors. While there are similarities to data science competitions (Makridakis *et al.* 2021), we call it a “Challenge” because we are empowering the community to do more than just submit forecasts by also collaboratively developing software, training materials, and best practices. In May 2020, we launched the design of the Challenge at a virtual conference with over 200 attendees (Peters and Thomas, 2021). Attendees prioritized five forecasting “themes” that (1) draw on NEON data, (2) address open science questions, and (3) have potential to support decision-making for natural resource management: 1) freshwater temperature, oxygen, and chlorophyll-a; 2) terrestrial carbon fluxes and evapotranspiration; 3) plant canopy phenology; 4) tick populations; and 5) beetle communities. Following the identification of the themes at the virtual meeting, smaller design teams were formed that developed detailed protocols for each theme that defined the timing of forecast submissions (when and how often forecasts are due) and forecast horizons (how far in the future will be forecasted). With these in place, they developed code to convert NEON data products into standardized time-series that are ready for modeling and forecast evaluation (i.e., the “targets”). In parallel to the design teams, the EFI-RCN Standards working group was assembled to define the format of forecast submissions and metadata across themes (Dietze *et al.* 2021). The EFI-RCN steering committee worked with each design team to ensure the protocols for evaluating forecasts are consistent across themes. This shared forecast evaluation approach emphasizes the importance of representing uncertainty in forecasts.

To support the Challenge, we have developed software and workflows for provisioning model inputs and processing model outputs that leverages modern cloud storage and computing (Figure 1). We developed new software to improve the efficiency of downloading NEON data while facilitating the analysis of data that exceeds computer memory (Boettiger *et al.* 2021). Other end-user tools convert NEON data into easy-to-use time-series, process submitted forecasts, score the probabilistic forecasts, and visualize submissions. Every day we are downloading, processing, and sharing NOAA numerical ensemble weather forecasts for all NEON sites that can be used as inputs to forecasting models, eliminating the need for users to have the knowledge to do so themselves. All of these technologies are open source and generalized to be usable beyond the Challenge.

What does participating in the Challenge look like? Foremost, we hope everyone who is interested feels empowered to submit forecasts as individuals or teams. To reduce barriers, we have generated documentation, workflow code examples, videos, and curated resources to train teams in computational skills needed for model development and submission. Participants can contribute forecasts from all sites and themes or any subset of interest. Submissions using any type of modeling framework (e.g., empirical, process-based, machine learning) are welcome.

We have created a set of baseline, or null, forecasts (e.g., historical average models) that serve as templates and benchmarks for the community submissions.

The Challenge provides a foundation for training in ecological modeling and forecasting. Teaching ecological modeling in undergraduate classrooms improves students' system thinking skills and quantitative literacy (Carey *et al.* 2020). Similarly, we have found that integrating ecological forecasting into classrooms expands students' understanding of complex ecological concepts (Moore *et al.* 2022). Challenge submissions are an ideal project for students in undergraduate and graduate courses and workshops, as we have discovered by training students in several university courses how to submit forecasts in collaborative teams . The rapid, iterative feedback inherent to the Challenge design inspires student engagement and improvement – submissions are evaluated daily as new submissions are accepted and as new data become available.

The NEON Ecological Forecasting Challenge is a community-driven enhancement of the NEON program that can help transform ecology into a more predictive science while providing feedback to NEON that improves data generation and delivery. Forecasting is part of the mission of NEON and the NEON Ecological Forecasting Challenge empowers the ecological forecasting community to lead the charge in accomplishing that mission. However, the Challenge extends well beyond NEON. The Challenge engages researchers who have not previously considered forecasting, who are in subfields without a history of model intercomparisons that promote rigor, or who are learning how to be more open and predictive in their data analysis and modeling. It is a testing ground for the development of new forecasting techniques that can be rapidly applied across broad environmental gradients with products and outcomes that have the potential to support decision-making in fields related to public and private environmental management and conservation. It fosters the development of a more quantitative, ecological workforce and can be an inspiration and blueprint for other environmental observatory networks across the globe. Overall, the Challenge catalyzed an interdisciplinary, community-scale project that has led to new forecast innovations, reduced latency of NEON data delivery, implemented a curriculum to give forecasting training and experience, and sparked engagement by forecast users about the potential of ecological forecasting.

We launched a beta round of the Challenge that ran throughout 2021 and resulted in 2,516 forecast submissions contributed by 54 different teams ranging in composition from undergraduate and graduate students to researchers at private companies. Contributions at this stage of the Challenge were critical to refining the protocols and to identifying the need for new software tools and education materials. The Challenge is now fully operational and actively seeking new contributions. If you are interested in becoming involved, using it in a course, or learning more, see www.neon4cast.org (archived at Thomas *et al.* 2021).

Acknowledgements

The EFI-RCN is supported by NSF (DEB-1926388) with computational support provided by NSF-funded Jetstream2 (OAC-2005506). NEON is a program sponsored by NSF and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the NSF through the NEON Program. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References:

Boettiger C, Thomas RQ, Laney C, and Lunch C. 2021. neonstore: NEON Data Store. https://CRAN.R-project.org/package=neonstore.

Carey CC, Farrell KJ, Hounshell AG, and O'Connell K. 2020. Macrosystems EDDIE teaching modules significantly increase ecology students' proficiency and confidence working with ecosystem models and use of systems thinking. *Ecology and Evolution* **10**: 12515–27.

Dietze MC, Fox A, Beck-Johnson LM, *et al.* 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proc Natl Acad Sci U S A* **115**: 1424–32.

Dietze MC, Thomas RQ, Peters JA, *et al.* 2021. A Community Convention for Ecological Forecasting: Output Files and Metadata. *EcoEvoRxiv*. https://doi.org/10.32942/osf.io/9dgtq

Enquist, CA *et al.* 2017. Foundations of translational ecology. *Frontiers in Ecology and the Environment* **15**: 541–550.

Lewis ASL, Rollinson CR, Allyn AJ, *et al.* 2022. The power of forecasts to advance ecological theory. *Methods Ecol Evol*. https://doi.org/10.1111/2041-210X.13955

Makridakis S, Spiliotis E, and Assimakopoulos V. 2021. The M5 competition: Background, organization, and implementation. *International Journal of Forecasting* **38**: 1325-1336

Moore TN, Thomas RQ, Woelmer WM, and Carey CC. 2022. Integrating Ecological Forecasting into Undergraduate Ecology Curricula with an R Shiny Application-Based Teaching Module. *Forecasting* **4**: 604–33.

Peters JA and Thomas, RQ. 2021. Going Virtual: What we Learned from the Ecological Forecasting Initiative Research Coordination Network Virtual Workshop. *Bull Ecol Soc Am* **102**: e01828

Thomas RQ, Boettiger C, Carey CC, *et al.* 2021. Ecological forecasting initiative: NEON ecological forecasting challenge documentation V1.0. https://doi.org/10.5281/zenodo.4780155

Figures:

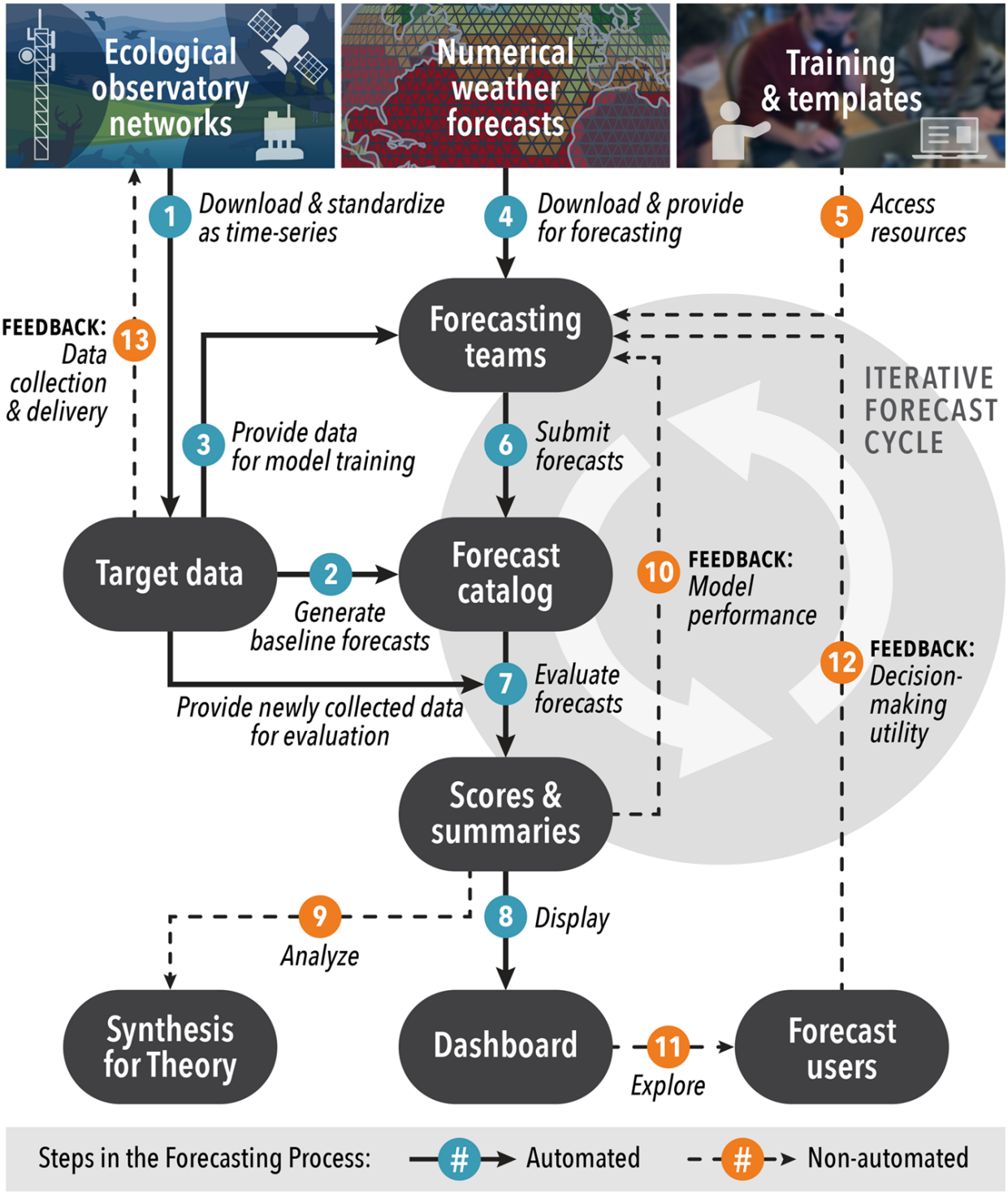

Figure 1. The workflow of the National Ecological Observatory Network (NEON) Ecological Forecasting Challenge, hosted by the Ecological Forecasting Initiative. It includes automated steps that are run daily and non-automated steps that are coordinated by the research network.

Supplemental Material

List of design team members and forecast contributors that opted in to be included in the group author list (“Challenge contributors”), reviewed, and approved of the manuscript. Each author was a member of the team that designed the protocols for one of the five themes and/or was a member of a team that submitted a forecast to the NEON Ecological Forecasting Challenge in the first round (2021).

Name	Institution
Carolina C. Barbosa	University of Wyoming
Nia Bartolucci	Boston University
Michael Benson	Indiana University
Uttam Bhat	University of California Santa Cruz
Matt Bitters	University of Colorado
George Burba	LI-COR Biosciences, Water for Food Global Institute, Carbon Dew
Sarah Burnet	University of Idaho
Pricilla Ceja	Humboldt State University
Min Chen	University of Wisconsin, Madison
Nicholas Clark	School of Veterinary Science, University of Queensland
Jamie R. Cleverly	James Cook University
Jon Cranko Page	University of New South Wales
Tad Dallas	University of South Carolina
Martin De Kauwe	University of Bristol
Ankur Desai	University of Wisconsin, Madison
Garret Dettmann	Virginia Tech
Maria Diuk-Wasser	Columbia University
Pilar Fernandez	Washington State University
Raphaela Floreani Buzbee	UC Berkeley
Kathryn Fuller	Western Sydney University
William Hammond	University of Florida
Dalei Hao	Pacific Northwest National Laboratory
Chris Jones	NC State University
Mira Kelly-Fair	Boston University
Kesondra Key	Indiana University
Shannon LaDeau	Cary Institute
Robert Ladwig	University of Wisconsin-Madison
David LeBauer	University of Arizona

Abigail S. L. Lewis	Virginia Tech
Catherine Lippi	University of Florida
Haoran Liu	University of Wisconsin, Madison
Mary Lofton	Virginia Tech
Charlotte Malmborg	Boston University
Ryan McClure	Carnegie Institution for Science
Brett A. Melbourne	University of Colorado Boulder
Casey Middleton	University of Colorado Boulder
Wynne Moss	University of Colorado
Naresh Neupane	Georgetown University
Kari Norman	Université de Montréal
Freya Olsson	Virginia Tech
Debasmita Pal	Michigan State University
L. Claire Powers	University of Colorado Boulder
Andrew Richardson	Northern Arizona University
Arun Ross	Michigan State University
Daniel Ruiz Carrascal	Columbia University
Sadie Ryan	University of Florida
Joshua Seabaugh	University of Colorado Boulder
Bijan Seyednasrollah	Northern Arizona University
Muhammed Shikhani	Helmholtz Centre for Environmental Research - UFZ Germany
Yiluan Song	University of California, Santa Cruz
Anna Spiers	University of Colorado Boulder
Laura Super	University of British Columbia
Thilina Surasinghe	Bridgewater State University
Benjamin Tonelli	University of California, Los Angeles
Eric Vance	University of Colorado Boulder
Kathryn Wheeler	Boston University
Jacob Wynne	Virginia Tech
Kelsey Yule	NEON Biorepository at Arizona State University
Luke Zachmann	Conservation Science Partners, Inc.
Yelu Zeng	China Agricultural University
Kai Zhu	University of Michigan
Jacob Zwart	U.S. Geological Survey