Annual Report for the NSF-sponsored Ecological Forecasting Initiative Research Coordination Network

January 26, 2023

This is a condensed and updated report from our annual report to the National Science Foundation in December 2022 reflecting activities and accomplishments from the third year of the Research Coordination Network funding (DEB-1926388).

The overarching goal of the project is to create a community of practice that builds capacity for ecological forecasting by leveraging NEON data products. Through meetings, working groups, and collaborative code development we aim to achieve the following objectives:

Objective 1. Define community standards and best practices for developing, sharing, and archiving forecasts and models

Objective 2. Increase the number and diversity of NEON-enabled forecasts by developing and hosting the NEON Ecological Forecasting Challenge

Objective 3. Create educational materials to empower scientists at all career stages to forecast using NEON data products

Objective 4. Support the creation of software to produce NEON-enabled forecasts at intensive and collaborative coding-focused workshops

Objective 5. Align forecast outputs and decision support with the needs of forecast users at mission-driven agencies to guide decision making, and

Objective 6. Synthesize forecasts to examine how limits to forecastability vary across ecological systems and scales.

Major Activities

All major activities of the Ecological Forecasting Initiative Research Coordination Network (EFI-RCN) occurred through virtual meetings. This was based on a survey sent to the Ecological Forecasting Initiative members that asked for meeting format preferences, where only 20% preferred an in-person meeting over a virtual meeting. Over the project-year, we effectively organized the community through multiple means that include large virtual meetings, smaller virtual working groups, collaborative software development, and sessions at society meetings.

Our first major activity during the third year of the EFI-RCN was running and expanding the NEON Ecological Forecasting Challenge, where teams are challenged to forecast NEON data before it is collected (www.neon4cast.org). The “Challenge” was created by volunteer efforts of the RCN where teams have designed the protocols, created the cyberinfrastructure to support the “Challenge”, created training materials, and evaluated the forecasts. Based on the outcome of our May 2020 RCN meeting, we launched the Challenge for five different themes, each organized by a theme-specific design team. The five themes were aquatic oxygen/temperature, beetle community diversity, tick populations, terrestrial carbon fluxes, and canopy phenology. During the third year we expanded the Challenge to include all 81 NEON sites, which required software development and refinement to efficiently process and evaluate the increased volume of forecasts and sites. We worked closely with NEON to reduce the data latency (time between data collection and public data availability). Now three of the five themes (all of the themes that use sensor data) have data latency less than 5 days, enabling the forecasts that are submitted and evaluated to be genuine near-term forecasts of the future. The other two themes require identification of tick and beetle samples so longer latencies are expected. We built an explicit partnership with the USGS this year by co-designing the aquatics forecasting challenge with the USGS so they can submit forecasts of NEON data that help them address their objective of improving forecasting capacities of chlorophyll-a in river ecosystems. Finally, the Challenge was used as the foundation of a summer training program in ecological forecasting at Boston University, a graduate course at the University of Notre Dame, the annual Fluxcourse, a workshop at the Global Lakes Ecological Observatory Network meeting, a mini-course at UCLA, and a workshop at the LTER All Scientists Meeting, bringing the total courses/workshops that have used the Challenge up to 11. To date, 90 teams have submitted 6,874 unique forecasts across the five themes, resulting in 9 million pairs of forecasts and observations that have been evaluated. The NEON Ecological Forecasting Challenge is a resounding success in advancing theory, technology, and training by providing a focal point for the community that leveraged open data. We are empowering the ecological forecasting community to lead the charge in accomplishing NEON’s forecasting mission.

Our second major activity during the third year of the EFI-RCN was organizing a large virtual meeting focused on sharing research and building a network in ecological forecasting (May 23-25, 2022). The meeting, titled “EFI 2022 Virtual Conference” was held with the goals to 1) provide opportunities to network with other members, including networking activities focused on early career individuals, 2) share research via short presentations or posters on the state of the field, 3) provide training in forecasting skills via longer workshops and tutorials. The meeting had wide participation: 284 people registered with almost 100 being grad students. It included

early, mid, and late career academics, NGO, government and industry participants. There was an even split of men and women. The majority of registrants were from the US but 35 different countries were represented. Using the virtual meeting software “Gather.town”, the meeting had 22 short talks, 21 posters, and 5 workshops, 3 facilitated networking sessions and multiple impromptu networking opportunities. We hosted an early career award with awards for best poster and best talk.

Our third major activity was organizing ecological forecasting sessions, symposia, and town hall meetings at major society meetings. We sponsored three organized sessions at the Ecological Society of America (ESA) meeting: “Ecological Forecasting: Applications, Discoveries, and Opportunities”, “A comparative perspective on integrating ecological forecasting with decision-making”, and “Making Useful Ecological Predictions with Indigenous Communities”. We supported a session on ecological forecasting at the American Geophysical Union meeting: “Ecological Forecasting in the Earth System”. Finally, we also organized an ecological forecasting session at the Joint Aquatic Sciences Meeting (“Advancing Near-term, Iterative Ecological Forecasting in Aquatic Ecosystems”) and taught a workshop on the NEON Ecological Forecasting Challenge at the Global Lakes Ecological Observatory Network (GLEON).

Our fourth major activity was general community-building activities that included the maintenance of a website, working group meetings, regular newsletters, blog posts, and an active Twitter feed. The Ecological Forecasting Initiative’s working groups revolve around Cyberinfrastructure & Methods & Tools; Forecasting Standards; EFI Student Association; Theory; Education; Diversity, Equity, and Inclusion; a Translation and Actionable Science working group which merged the Social Science, and Partners & Knowledge Transfer working groups; the Forecast Challenge design teams and teams working on manuscripts, and a book group related to diversity, equity, and inclusion. There were almost 100 Zoom calls for these working group calls during the third year of the project. Progress by the working groups and design teams, an overview of the Challenge and how it was used in educational settings, and awards presented for forecasting papers and presentations were shared in four blog posts on the ecoforecast.org website and through eight newsletters that were sent to a listserv with over 1,000 members. Finally, our Twitter feed (@eco4cast) has gained over 100 additional followers this year giving us over 2,000 followers. This year we had 86 tweets and 555 mentions.

Our fifth major activity focused on efforts targeted toward educational materials and webinars and activities to promote and increase diversity, equity, and inclusion. Specifically, we supported 10 students from minority backgrounds to travel to the Data Science in Indian Country Geoscience Alliance 2022 meeting. At the meeting, we ran a workshop for 20 people who were majority BIPOC, called “Ecological Forecasting and Native American Land and Water Resource Management”. The workshop was adapted from a multi-week exercise led by Georgia Smies (Salish Kootenai College, SKC) and co-developed with Jason McLachlan (U of Notre Dame) and grad student Helena Kleiner (U of Notre Dame). In the multi-week exercise data science students at SKC developed models in the R programming language to highlight the synergies and tensions between the values of western science embedded in established Environmental Protection Agency (EPA) monitoring protocols and the values and priorities of TEK.

We hosted additional panel seminars and series that revolved around topics pertinent to the field of ecological forecasting. In particular, we co-hosted a Statistical Methods Seminars with the ESA Statistical Ecology Section that provided an overview of a statistical method and corresponding R package(s). We also hosted the second session in a virtual science communications series called “Sharing is Caring: Communicating Science Beyond Academic Publications”.

Key Outcomes or Other Achievements Associated with Specific Objectives

Objective 1.

We revised a preprint defining standards for saving forecast output and generating metadata describing the forecast. These standards are supported by software developed to help generate and validate the metadata. We used the NEON Ecological Forecasting Challenge as a test-bed for the standard revision. We also collaborated with the SpatioTemporal Asset Catalogs community to ensure interoperability in the standards across schemas. Finally, we were awarded a CSSI NSF grant (2209866), in collaboration with the GeoCODES community, that builds on these standards to improve the discoverability and re-use of ecological forecasts.

Objective 2.

We hosted and revised the NEON Ecological Forecasting Challenge (www.neon4cast.org). The process of hosting included generating a website describing the forecast challenge themes and rules; keeping the cyberinfrastructure running and constantly downloading NOAA and NEON data, processing submitted forecasts, scoring forecasts, and generating a dashboard describing the results. Our server ran the NSF-funded Jetstream Cloud computation environment. In total 90 teams have submitted 6,874 unique forecasts across the five themes, resulting in 9 million pairs of forecasts and observations that have been evaluated. The Challenge is now fully built out and includes all 81 NEON sites. The Challenge will continue to run for the full length of the RCN project and serves as a focal point for the network. The Challenge has provided feedback to NEON that improves their data products, including reducing data latencies, expanding data products, and revising data products to correct anomalous data.

Objective 3.

First, the EFI 2022 Virtual Meeting (May 23-25, 2022) was a very broad reaching effort to get students involved in ecological forecasting. There were 100 students that registered for the meeting and specific networking sessions were provided for students. We awarded prizes for best talks and posters. We provided training in ecological forecasting at the GLEON annual meeting. We supported 10 students from minority backgrounds to travel to the Geoscience Alliance 2022 meeting titled “Data Science in Indian Country”. At the meeting, we ran a workshop called “Ecological Forecasting and Native American Land and Water Resource Management”. We co-hosted with the ESA Statistical Ecology Section an on-going webinar series on 9 different statistical methods in ecology. In total, over 2,300 people joined the virtual events in person, and almost 17,500 have since watched the recordings from the events on the

EFI YouTube channel. Finally, we submitted a Sloan Foundation proposal to diversify and increase equitable access to Earth and data science graduate pathways for Native American students through collaborations with Tribal Colleges & Universities, Minority Serving Institutions, research universities, and professional organizations (American Indian Higher Education Consortium, the Geoscience Alliance, the Native Food, Energy, Water Systems Alliance, and EFI) by developing classroom and online Earth and data science courseware that centers Native American values, strengthens collaborations between TCU, MSI, and research universities, provides undergraduate research projects, and facilitates engagement in the EFI community.

Objective 4.

We created new software to download NOAA weather forecasts from Amazon Web Services and process them for all 81 NEON sites, thus continuing our development of a cloud-native cyberinfrastructure that supports the NEON Ecological Forecasting Challenge (R package:gefs4cast). Many of the teams that have submitted forecasts to the Challenge use the NOAA weather forecasts as inputs to their forecasting models. We also developed a new package for efficiently evaluating forecasts using new observations in a cloud-native cyberinfrastructure environment (R package: score4cast). Finally, we updated all of our packages to work with the new 'arrow' package in R which allows us to efficiently increase the scale of the project. The arrow package allows for the analysis of datasets that are larger than typical computer memory. The functions in these packages are being used by other ecological forecasting projects beyond the Challenge. For example, the NSF Project 1933016 CIBR: Cyberinfrastructure Enabling End-to-End Workflows for Aquatic Ecosystem Forecasting is using the software in forecast evaluation. We refined software (R package: neonstore) to provide a high-level user interface for downloading and storing NEON data products. While each of NEON data products consists of hundreds or thousands of individual files, unlike 'neonUtilities', this package will avoid repeated downloading, provides persistent storage, and improves performance. 'neonstore' can also construct a local 'duckdb' database of stacked tables, making it possible to work with tables that are far too big to fit into memory. All of our software is publicly available on GitHub.

Objective 5.

We co-developed the expanded aquatics theme in the NEON Ecological Forecasting Challenge with the USGS. This involved expanding the number of sites in the Challenge to include river sites and a new variable (chlorophyll-a) that were specially identified by the USGS as priorities.

Objective 6.

We published a paper that describes how forecasts can be used to advance theory in ecology (Lewis et al. 2022), and another that describes how the use of best practices in ecological forecasting has changed over time (Lewis et al. 2021, which came out in the March 2022 issue of Ecological Applications). We had a paper accepted that explores how predictability varies across NEON lake sites (Thomas et al. 2022), and had another paper accepted that provides an overview of the NEON Forecasting Challenge (Thomas et al. 2022).

Opportunities for training and professional development

Multiple opportunities for training and professional development have been provided by the RCN to the broader EFI community as well as specifically for graduate students and postdoctoral scholars.

First, our RCN has an active graduate student association (EFISA) that meets monthly, resulting in 10 meetings over the past year. The group currently has over 50 individuals from 42 institutions officially registered as EFISA members and an #efisa Slack channel with over 100 participants. Graduate students have developed their own operating principles and procedures and defined the goal of the association to provide a community of students who have expertise in a diverse array of fields, but who all share a common interest of improving and contributing to the development and application of ecological forecasts. As a result, knowledge exchange and community-building are two of their foremost goals. Along with these aims, the EFISA strives to serve the student community by developing technical and professional development skills, providing peer support, and advancing career aims.

In May 2022, we hosted a virtual meeting with over 100 graduate students that registered. The meeting provided opportunities for graduates to share research through talks and posters, network with others in the ecological forecasting community through facilitated networking events, and gain new skills through multi-hour workshops.

One post-doc (Freya Olsson at Virginia Tech) is leading the training, recruitment, and synthesis of the aquatic theme in the NEON Ecological Forecasting Challenge. Likewise, Gerbrand Koren (Assistant Professor at Utrecht University) is championing the terrestrial carbon and water flux theme.

Students and post-docs are leading manuscripts that are currently being developed. One manuscript, led by graduate student Abby Lewis (Virginia Tech), describes how a forecasting approach can shed new light on existing ecological theories while also allowing researchers to address novel questions. A second manuscript, submitted by Alyssa Willson (U of Notre Dame), compiles and synthesizes the materials from the Education and Inclusive Pedagogy June 2022 meetings. This manuscript also assesses the type of ecological forecasting-related courses already available in all types of US institutions (e.g., community colleges, R1 institutions, MSI, HBCUs, etc). A third manuscript, submitted by Mary Lofton (Virginia Tech), reviews the state of freshwater water quality forecasting. A fourth paper was recently submitted by Kathryn Wheeler (formerly a graduate student at Boston University, now a post-doc at MIT) who led a large group analyzing 18 forecasts submitted by 13 teams to the phenology theme of the NEON Forecast Challenge for the spring greenup in 2021.

Graduate students have gained professional leadership experience by serving on the RCN steering committee. Whitney Woelmer (Virginia Tech), Alyssa Willson (U Notre Dame), and Lynda Bradley (Emory University) have provided important guidance for making the RCN serve the needs of early career scientists and extensively helped to coordinate the RCN meetings in May, including a student led networking event and another student focused networking event.

We hosted one book group this past year on the book “Weapons of Math Destruction”. The group met for 5 weeks during the summer and had 8 participants ranging from graduate students to mid-career individuals from academia and government agencies.

The Challenge was used as the foundation of a summer training program in the Near-term Ecological Forecasting Summer Course at Boston University, a graduate course at the University of Notre Dame, the annual Fluxcourse, a workshop at the Global Lakes Ecological Observatory Network meeting, a mini-course at UCLA, and a workshop at the LTER All Scientists Meeting, bringing the total courses/workshops that have used the Challenge up to 11.

Through a collaboration between the University of Notre Dame, Boston University and California State Polytechnic University, Humboldt (CPH; formerly Humboldt State University), Tempest McCabe (Boston U) mentored five undergraduates (3 women, and 2 men; all five Hispanic) in the Louis Stokes Alliance for Minority Participation (LSAMP) program at CPH. Students came from a wide range of majors (botany, environmental resources engineering, environmental science & management, and wildlife). Over the course of the semester, the students went through the 24 NEON/EFI-produced “Fundamentals of Ecological Forecasting” videos and assessed what they learned about ecological forecasting from the videos. McCabe also worked with students to search for and apply for summer research positions. Students, Brenden Chavez and Daniel Chaidez participated in summer research. Chavez worked with the McLachlan lab at U of Notre Dame to compare how negative demographic feedback processes are represented in two forest ecosystem models. Chaidez participated in summer research at Washington State University where he built a model using Stella to predict the cost and efficacy of two methods of reducing sunburn on apple production on agricultural modeling. Both students prepared posters to present at the Geoscience Alliance and Society for the Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS), respectively.

The multi-week exercise co-developed by Georgia Smies (SKC) and Jason McLachlan and Helena Kleiner (U of Notre Dame) focused on using R skills to develop data analysis skills needed for EPA assessments and incorporating TEK values and priorities was taught at SKC, modified and presented as a 1-day workshop at the Geoscience Alliance, and will be used to demonstrate ethical issues for practitioners to consider with ecological forecasting. In addition to being adapted for the one-day workshop, the material is currently being adapted as an online exercise, which will be featured in a manuscript (in prep) on teaching the ethical dimensions of ecological forecasting.

The Statistical Methods webinar series that was co-hosted with the Ecological Society of America Statistical Ecology section provided an overview of 9 statistical methods and advice for using and avoiding pitfalls for R packages used for those statistical methods. Webinar topics included: Structural Equation Models and the piecewiseSEM R package, Analysis of Bioacoustic Data, Spatial Occupancy Models and the spOccupancy R package, Hidden Markov Models, NIMBLE, Multi-Species (Species Interactions) Occupancy Modeling, Integrated Step-Selection Analysis, Movement Ecology, Generalized Joint Attribute Modeling (GJAM).

References

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. (2022). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution*, *00*, 1–11. https://doi.org/10.1111/2041-210X.13955

Thomas, R. Q., McClure, R., Moore, T., Woelmer, W., Boettiger, C., Figueiredo, R., et al. (2022). Near-term forecasts of NEON lakes reveal gradients of environmental predictability across the U.S. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10510642.1 (accepted at Frontiers in Ecology and Environment)

Lewis, A. S. L., Woelmer, W. M., Wander, H. L., Howard, D. W., Smith, J. W., McClure, R. P., et al. (2021). Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. *Ecological Applications*, *32*(2), e02500. https://doi.org/10.1002/eap.2500

Pre-Print: Willson, A.M., H. Gallo, J.A. Peters, A. Abeyta, N. Bueno Watts, C.C. Carey, T.N. Moore, et al. 2022. “Assessing Opportunities and Inequities in Undergraduate Ecological Forecasting Education.” *Zenodo*, July. https://doi.org/10.5281/zenodo.6886540.

Pre-print: Lofton, M.E., D.W. Howard, R.Q. Thomas, C. C Carey. Progress and opportunities in advancing near-term forecasting of freshwater quality. ESS Open Archive https://doi.org/10.1002/essoar.10512426.1

Pre-print (Accepted at Frontiers in Ecology and Environment). Thomas, R.Q., C. Boettiger, C.C. Carey, M.C. Dietze, L.R. Johnson, M.A. Kenney, J.S. Mclachlan, J.A. Peters, E.R. Sokol, J.F. Weltzin, A. Willson, W.M. Woelmer, and Challenge Contributors. 2022. The NEON Ecological Forecasting Challenge. ESS Open Archive. https://www.doi.org/10.22541/essoar.167079499.99891914/v1