WILEY ECOLOGICAL APPLICATIONS

Increased adoption of best practices in ecological forecasting enables comparisons of forecastability across systems

Journal:	*Ecological Applications*
Manuscript ID	Draft
Wiley - Manuscript type:	Articles
Date Submitted by the Author:	n/a

Complete List of Authors:	Lewis, Abigail; Virginia Tech, Biological Sciences Woelmer, Whitney; Virginia Tech, Biological Sciences Wander, Heather; Virginia Tech, Biological Sciences Howard, Dexter; Virginia Tech, Biological Sciences Smith, John ; Virginia Tech, Statistics McClure, Ryan; Virginia Tech, Biological Sciences Lofton, Mary; Virginia Tech, Biological Sciences Hammond, Nicholas; Virginia Tech, Geosciences Corrigan, Rachel; Virginia Tech, Forest Resources and Environmental Conservation Thomas, R. Quinn; Virginia Tech, Forest Resources and Environmental Conservation Carey, Cayelan; Virginia Tech, Biological Sciences

Substantive Area:	Statistics and Modeling < Theory < Substantive Area
Organism:	
Habitat:	
Geographic Area:	Global < Geographic Area
Key words/phrases:	Data assimilation, Decision support, Ecological predictability, Forecast automation, Forecast horizon, Forecast skill, Forecast uncertainty, Iterative forecasting, Near-term forecast, Null model, Open science, Uncertainty partitioning

Abstract:	Near-term iterative forecasting is a powerful tool for ecological decision support and has the potential to transform our understanding of ecological predictability. However, to this point, there has been no cross-ecosystem analysis of near-term ecological forecasts, making it difficult to synthesize diverse research efforts and prioritize future developments for this emerging field. In this study, we analyzed 178 near-term ecological forecasting papers to understand the development and current state of near-term ecological forecasting literature and compare forecast skill across ecosystems and variables. Our results indicate that near-term ecological forecasting is widespread and growing: forecasts have been produced for sites on all seven continents and the rate of forecast

publication is increasing over time. As forecast production has accelerated, a number of best practices have been proposed and application of these best practices is increasing. In particular, data publication, forecast archiving, and workflow automation have all increased significantly over time. However, adoption of proposed best practices remains low overall: for example, despite the fact that uncertainty is often cited as an essential component of an ecological forecast, only 45% of papers included uncertainty in their forecast outputs. As the use of these proposed best practices increases, near-term ecological forecasting has the potential to make significant contributions to our understanding of predictability across scales and variables. In this study, we found that forecast skill decreased in predictable patterns over 1–7 day forecast horizons. Variables that were closely related (i.e., chlorophyll and phytoplankton) displayed very similar trends in predictability, while more distantly related variables (i.e., pollen and evapotranspiration) exhibited significantly different patterns. Increasing use of proposed best practices in ecological forecasting will allow us to examine the forecastability of additional variables and timescales in the future, providing a robust analysis of the fundamental predictability of ecological variables.

VIRGINIA TECH
College of Science

Abigail Lewis
Department of Biological Sciences
926 W. Campus Dr.
Blacksburg, Virginia 24061
(262) 565-7269
E-mail: aslewis@vt.edu
www.biol.vt.edu

Juan Corley, Ph.D.
Editor-in-Chief
Ecological Applications
April 19, 2021

Dear Dr. Corley,

Please find attached our manuscript entitled, "Near-term ecological forecasts increasingly follow best practices, allowing comparisons of forecastability among studies" for consideration as an *Ecological Applications* research article. We analyze the current state of near-term ecological forecasting and use data from 29 studies to compare the relative forecastability of multiple ecological variables. To the best of our knowledge, our study is the first to quantify and compare the rates at which the predictability of multiple ecological variables decrease over increasing forecast horizons (the amount of time into the future for which predictions are made).

We anticipate this manuscript will be of broad interest to the readers of *Ecological Applications* by providing an overview of current techniques in ecological forecasting and recommendations for prioritizing future developments in the field. Our analysis reveals that near-term ecological forecasting is a widespread practice within ecology: forecasts have been produced for sites on all seven continents and the rate of forecast publication has increased dramatically over the past four decades. As forecast publication has accelerated, the use of many forecasting best practices (e.g., archiving forecasts, publishing driver data) has also increased over time. However, several areas in need of development remain: for example, despite the fact that uncertainty is often considered essential to the definition of an ecological forecast, uncertainty is reported in less than half of published forecasts. The most commonly used best practice in this analysis was "assess and report forecast skill," which enabled us to compare forecast skill across papers. We found that forecast skill decreased in consistent patterns over 1–7 day forecast horizons, but the magnitude of decline varied by forecast variable. This analysis makes an important contribution to our understanding of ecological predictability, providing support for previous theoretical predictions. We specifically targeted *Ecological Applications* for this manuscript because our analysis shows that it is the premier journal within ecology for publishing near-term ecological forecasts and thus will have a broad audience among its readership.

We recommend the following reviewers, with whom we do not have any conflicts of interest:

- Dr. Christine Rollinson, Morton Arboretum, Lisle, Illinois, USA, crollinson@mortonarb.org
- Dr. Peter Adler, Utah State University, Logan, Utah, USA, peter.adler@usu.edu
- Dr. Andrew Tredennick, Western EcoSystems Technology, Laramie, Wyoming, USA, atredenn@gmail.com
- Dr. David Harris, Chewy, Dania Beach, FL dave@harris-research.me

This manuscript has not been accepted for publication before, nor is it under consideration for another journal or book. The research met U.S. legal requirements for responsible research. Each named author has substantially contributed to conducting the underlying research and drafting this manuscript, and all co-authors have approved this submission. No co-authors hold any conflict of interest. As mentioned in the text, all datasets analyzed in this manuscript have been published to the staging environment of the Ecological Data Initiative (EDI) data portal and included as a supplemental file for manuscript review. If accepted for publication, we will archive the datasets to the full data portal where they will be assigned a DOI.

We hope you find this manuscript suitable for publication in *Ecological Applications* and look forward to hearing from you.

Sincerely,

Abigail Lewis

Abigail Lewis, on behalf of the coauthors

Virginia Tech
926 West Campus Drive
Blacksburg, Virginia 24061
aslewis@vt.edu
(262) 565-7269

Running head: Best practices and forecastability

Title: Increased adoption of best practices in ecological forecasting enables comparisons of forecastability across systems

Authors: Abigail S. L. Lewis[1a], Whitney M. Woelmer[1b], Heather L. Wander[1c], Dexter W. Howard[1d], John W. Smith[2e], Ryan P. McClure[1f], Mary E. Lofton[1g], Nicholas W. Hammond[3h], Rachel S. Corrigan[4i], R. Quinn Thomas[4j], Cayelan C. Carey[1k]

[1]Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA

[2]Department of Statistics, Virginia Tech, Blacksburg, Virginia, USA

[3]Department of Geosciences, Virginia Tech, Blacksburg, Virginia, USA

[4]Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia, US

---

[a] Corresponding author. Address: Department of Biological Sciences, Derring Hall Room 2025, 926 West Campus Drive, Mail Code 0406, Blacksburg, VA 24061, USA. Phone: +1-540-231-0788. Fax: +1-540-231-9307. Email: aslewis@vt.edu, ORCID ID: 0000-0001-9933-4542

[b] Email: wwoelmer@vt.edu, ORCID ID: 0000-0001-5147-3877

[c] Email: hwander@vt.edu, ORCID ID: 0000-0002-3762-6045

[d] Email: dwh1998@vt.edu, ORCID ID: 0000-0002-6118-2149

[e] Email: wsjohn2@vt.edu, ORCID ID: 0000-0002-1564-3290

[f] Email: ryan333@vt.edu, ORCID ID: 0000-0001-6370-3852

[g] Email: melofton@vt.edu, ORCID ID: 0000-0003-3270-1330

[h] Email: hammondnw@vt.edu, ORCID ID: 0000-0003-2975-8280

[i] Email: rachelc1@vt.edu, ORCID ID: 0000-0001-6101-8085

[j] Email: rqthomas@vt.edu, ORCID ID: 0000-0003-1282-7825

[k] Email: cayelan@vt.edu, ORCID ID: 0000-0001-8835-4476

Open data statement

Data and metadata are provided as private-for-peer review in a supplement and via the following link in the Environmental Data Initiative staging repository (in the text below, we refer to this data publication as forthcoming):

https://portal-s.edirepository.org/nis/mapbrowse?scope=edi&identifier=196&revision=5

Submitted as a Research Article to *Ecological Applications*

Declarations of interest: none

ABSTRACT

Near-term iterative forecasting is a powerful tool for ecological decision support and has the potential to transform our understanding of ecological predictability. However, to this point, there has been no cross-ecosystem analysis of near-term ecological forecasts, making it difficult to synthesize diverse research efforts and prioritize future developments for this emerging field. In this study, we analyzed 178 near-term ecological forecasting papers to understand the development and current state of near-term ecological forecasting literature and compare forecast skill across ecosystems and variables. Our results indicate that near-term ecological forecasting is widespread and growing: forecasts have been produced for sites on all seven continents and the rate of forecast publication is increasing over time. As forecast production has accelerated, a number of best practices have been proposed and application of these best practices is increasing. In particular, data publication, forecast archiving, and workflow automation have all increased significantly over time. However, adoption of proposed best practices remains low overall: for example, despite the fact that uncertainty is often cited as an essential component of an ecological forecast, only 45% of papers included uncertainty in their forecast outputs. As the use of these proposed best practices increases, near-term ecological forecasting has the potential to make significant contributions to our understanding of predictability across scales and variables. In this study, we found that forecast skill decreased in predictable patterns over 1–7 day forecast horizons. Variables that were closely related (i.e., chlorophyll and phytoplankton) displayed very similar trends in predictability, while more distantly related variables (i.e., pollen and evapotranspiration) exhibited significantly different patterns. Increasing use of proposed best practices in ecological forecasting will allow us to examine the forecastability of additional

variables and timescales in the future, providing a robust analysis of the fundamental predictability of ecological variables.

KEY WORDS

Data assimilation, decision support, ecological predictability, forecast automation, forecast horizon, forecast skill, forecast uncertainty, iterative forecasting, near-term forecast, null model, open science, uncertainty partitioning

INTRODUCTION

Nearly 90 years ago, Hodgson (1932) published what was arguably the first near-term ecological forecast, using demographic trends to predict herring age structure one year into the future. Hodgson concluded by stating "... we hope that before long these prognostications will be issued with the same confidence as those which are broadcast daily by the Meteorological Office, and, once they are received with confidence by the trade, they should be of considerable financial value" (p. 118). During the past 90 years, advances in data availability, computational power, and statistical methodologies have enabled a substantial increase in the development and application of near-term ecological forecasts (Luo et al. 2011, Petrovskii and Petrovskaya 2012, Hampton et al. 2013, LaDeau et al. 2017). Near-term ecological forecasting has become an increasingly powerful tool for ecological decision support (Dietze 2017a, Henden et al. 2020, Carey et al. 2021) and has the potential to provide new insights into fundamental questions about ecological functioning and predictability (Petchey et al. 2015, Dietze 2017b, Dietze et al. 2018). However, to this point, there has been no systematic analysis of the development or current state of near-term ecological forecasting literature, making it difficult to synthesize diverse research efforts and prioritize future developments for this emerging field.

Throughout the development of near-term ecological forecasting, there have been numerous calls for the adoption of standardized best practices (e.g., Clark et al. 2001, Pielke and Conant 2003, Harris et al. 2018, Dietze et al. 2018, White et al. 2019, Hobday et al. 2019, Carey et al. 2021). Developing and adhering to best practices advances the contributions of forecasting to both basic and applied research, as it allows for comparisons of forecast skill across forecast horizons (the amount of time into the future for which predictions are made) and increases the reliability of forecast products as decision support tools (Armstrong 2001). Recent interest in establishing best practices for ecological forecasting follows similar efforts in meteorology and economics, disciplines in which forecasting is well-established (Armstrong 2001, Hyndman and Athanasopoulos 2018).

While proposed best practices for near-term ecological forecasting differ among papers, a number of common themes related to forecast development, assessment, archiving, and decision support can be identified (Box 1). As ecological forecasting has developed over the past several decades, we expect that adherence to these proposed best practices is increasing. However, without a comprehensive review of published ecological forecasts, it is difficult to assess which of the proposed best practices have been adopted and which should be prioritized for further advancement of the field. Ideally, best practices should evolve using a community-driven approach to enable buy-in and robustness to many applications (following Hanson et al. 2016); consequently, the list of proposed best practices in Box 1 is not exhaustive, and some of the practices may not be appropriate for every forecasting application. However, these practices provide a framework to begin analyzing the state of the field.

Adoption of these proposed best practices in near-term ecological forecasting may be particularly important to improving our understanding of predictability across ecosystems and scales. As the number of published near-term ecological forecasts has increased over the past several decades (Luo et al. 2011, Dietze et al. 2018), we now have an unprecedented opportunity to compare across studies and analyze the relative forecastability of environmental variables at varying forecast horizons. Quantifying ecological predictability is a fundamental goal in ecology (Sutherland et al. 2013, Godfray and May 2014, Houlahan et al. 2017) and provides valuable information regarding the nature of ecological processes (Petchey et al. 2015). Taking a data-driven approach to this problem complements and extends existing theoretical and modeling-based work that has predicted how various factors (e.g., forecast horizon, computational irreducibility) influence the relative predictability of ecological variables (Beckage et al. 2011, Petchey et al. 2015).

In this study, we performed a systematic analysis of near-term ecological forecasting papers to examine the use of our proposed best practices over time (Box 1). To illustrate how proposed best practices can enable insights into fundamental ecological understanding, we then compared forecast skill across ecosystems and variables. We discuss the implications of our findings for further development and adoption of best practices within the near-term ecological forecasting research community.

METHODS

We systematically reviewed literature on near-term ecological forecasting to determine how proposed best practices have been implemented over time and compare forecastability across ecosystems. First, we searched the Web of Science™ Core Collection [v.5.34] database (Clarivate Analytics, Philadelphia, USA) and reviewed abstracts to identify papers that reported near-term ecological forecasts (described in *Literature search* below). Two reviewers then independently read and analyzed each selected paper using a standardized matrix of criteria (*Matrix analysis*) and recorded forecast skill when reported. Once collated, we analyzed the full dataset to understand the development and current state of ecological forecasting (*Dataset description* and *Assessment of forecasting best practice adoption*). Finally, we analyzed forecast skill data to assess how forecast performance varied with forecast horizon for ecological variables with sufficient data (*Comparing forecast skill across ecosystem and models*).

Literature search

Creating an all-encompassing search query to identify near-term ecological forecasts presented three challenges: first, the term "near-term" was neither universally defined nor used in all papers that report near-term forecasts; second, there was no one search term that can match all papers describing ecological variables; and third, many papers used the word "forecast" when talking about implications of their research, despite not actually reporting forecasting results in the paper. To address these challenges, we began by querying the Web of Science Core Collection [v.5.34] for "forecast*" in the title, abstract, or keywords of papers published in 301 ecological journals, then manually screened abstracts of all resulting papers. We conducted the Web of Science search on 18 May 2020 and limited the search to articles and proceedings papers (hereafter, 'papers') published in English. This yielded 2711 results (Fig. 1).

We screened the abstracts of all 2711 papers and selected those that met three criteria:

1. Papers had to include at least one forecast, which we defined as a prediction of future conditions from the perspective of the model; forecasts could be developed retroactively (i.e., "hindcasts") but could only use driver data that were available before the forecast date (e.g., forecasted or time-lagged driver variables).
2. The forecast had to be near-term, which we defined as predicting ≤10 years into the future.
3. The forecast had to be ecological, which we defined as predicting a biogeochemical, population, or community response variable. This definition therefore excluded physical (e.g., streamflow or water temperature) and meteorological forecasts. Forecasts of human disease were only included if there was an animal vector.

If the abstract indicated that the paper met all three criteria, it was moved to the second round of screening. Here, a second reviewer read the full paper to ensure that at least one forecast in the paper met all three criteria.

By the end of this screening process, we identified 142 near-term ecological forecasting papers out of the 2711 Web of Science results (Fig. 1a, 1b). The initial Web of Science search did well at identifying studies with ecological focal variables, as 74% of the initial search results were marked as 'ecological' during our review process. However, only 36% of papers from this search actually included forecasts (predicting future conditions from the perspective of the forecast model). Furthermore, of the ecological forecasts identified in this search, only 21% met our near-term criteria by including forecast horizons that were ≤10 years; the majority of forecasts predicted ecological changes over multidecadal timescales (Fig. 1b).

Because ecological forecasts may be published in journals that are not categorized as "ecological" by Web of Science, we then searched all papers that were cited by the 142 near-term ecological forecast papers we identified, as well as all papers that cited these studies. From the citing and cited papers, we selected those that were published in English and included "forecast*" in the title, abstract, or keywords, then screened the abstracts using our three criteria described above. Finally, a second reviewer read all papers that passed the abstract screening to confirm that at least one forecast in the paper met all three criteria. Searching the papers that cite and are cited by the near-term ecological forecasting papers from our initial search yielded proportionally more ecological forecasting papers than the initial Web of Science search. Of the 472 search results, 112 (24%) of these papers were identified as near-term ecological forecasts after two rounds of review (Fig. 1a, 1c). Furthermore, this search highlighted predominantly near-term forecasts; 73% of the ecological forecasts identified in this search included forecast horizons that were ≤10 years (Fig. 1c). After combining our initial search with the citing and cited papers, 254 papers were included in our dataset for matrix review (Fig. 1a).

Matrix analysis

We analyzed each of the 254 papers using a standardized matrix of questions (Lewis et al. Environmental Data Initiative repository forthcoming data publication). This matrix was co-developed over several months of iteration and discussion by all authors within an Ecological Forecasting graduate seminar at Virginia Tech (January–May 2020). The final matrix used for this study included 57 fields of information about the forecast paper's model(s), evaluation, cyberinfrastructure, archiving, and decision support (Lewis et al. Environmental Data Initiative repository forthcoming data publication).

During the graduate seminar, we read and analyzed 10 papers as a group, ensuring that all reviewers understood how to interpret and answer questions in a consistent manner. Reviewers also screened several papers individually and checked their responses with another reviewer prior to the start of this analysis, helping to ensure consistency between reviewers. For the matrix analysis described in this paper, all 254 papers were read and analyzed independently by two reviewers, and reviewers then compared any differing answers to reach consensus on a final set of responses for each paper.

During the matrix analysis, 76 papers were determined to not meet our criteria of being near-term ecological forecasts, despite having passed the initial rounds of screening. These papers typically used one or more data sources that became available after the forecast issue date, which was difficult to identify without reading the entire text, including supplementary information, in detail. These papers were excluded from the analysis, leaving 178 papers in the final dataset (Fig. 1a).

Dataset description

To characterize the current state of near-term ecological forecasting, we began by analyzing the distribution of forecasts across geographical locations, variables, and time scales, as described below.

We classified the spatial scale of each forecast into five categories: point (localized to one discrete site, such as pollen forecasts for a city or algal forecasts for a lake), multipoint (several distinct forecast locations, such as three different lakes), regional (localized to a broad geographic region, such as coral bleaching forecasts that span a sea), national (spanning all of one nation, such as nationwide production of an agricultural crop), or global (such as coral bleaching stress in world oceans), and we calculated the percentage of forecasting papers within each of these categories. We recorded latitude and longitude of the forecast site(s) for point or multipoint forecasts or of the approximate centroid of the site for regional and national forecasts.

Forecast variables were divided into two categories: organismal (population and community; e.g., white-tailed deer populations) and biogeochemical (e.g., evapotranspiration), and each paper was classified within one of 11 ecosystem types: forest, grassland, freshwater, marine, desert, tundra, atmosphere, agricultural, urban, global, other, where “other” included any ecosystem types that could not be classified within one of the other 10 categories (e.g., plant phenology across the entire United States). We recorded the number of years of data used to create each forecasting paper (summed across model development, training, evaluation, etc.) and calculated the percentage of papers that use long-term datasets in their analysis, using the definition of long-term as any dataset with more than ten years of data (Lindenmayer et al. 2012).

Assessment of forecasting best practice adoption

To analyze how adherence to the proposed best practices has changed over time, we performed binary logistic regressions assessing how adoption of each best practice (binary yes/no) varied with publication year. Hodgson (1932) was excluded from this analysis as a temporal outlier, leaving a dataset of papers published between 1980 and 2020. We used the following criteria in the matrix analysis to assess which proposed best practices (Box 1) were included in each forecasting paper:

Forecast Requirements

1. "Include uncertainty": uncertainty was included in forecast outputs
2. "Assess and report forecast skill": any form of forecast evaluation was reported (this includes figures that compare forecasts and observations, as well as any skill score)

Decision Support

3. "Identify an end user": A specific end user was mentioned
4. "Make iterative forecasts": Forecasts were made repeatedly, incorporating new data over time. For this practice, we included all types of data assimilation, including those that only updated the initial conditions of the forecast. As a separate analysis, we also determined whether data assimilation methods that updated the parameters of the model (not just initial conditions) have increased over time

5. "Automate forecasting workflows": at least one source of new driver and/or observation data was made available to the model in real time (<24 hours from collection) without any manual effort when the system was working as intended

Research

6. "Make data available": Data availability was specified
7. "Archive forecasts": Text specified that forecasts were archived and available
8. "Use null model comparisons": Forecasts were compared to a persistence or climatology null model
9. “Partition uncertainty”: At least two different sources of uncertainty were quantified and compared

All analyses were performed using R version 4.0.3 (R Core Team 2020).

Comparison of forecast skill across ecosystem and models

To compare forecast performance across forecast variables, sites, and scales, it is necessary to identify a skill metric that is not dependent on the units or range of the forecast variable. For reasons discussed below, we chose R^2 as our metric of forecast performance in this analysis. Petchey et al. (2015) recommend using the length of time until a forecast performs no better than a relevant threshold value as one way of comparing between papers. However, this type of analysis would require that a threshold value be determined equitably for each forecast variable, which would be challenging across the numerous variables in our dataset. Performance of null models offers one objective way of determining these threshold values, but null models were not commonly reported in this dataset. Another means of comparing forecast performance would be to directly compare forecast skill using a standardized statistical score. Commonly used forecast skill metrics include root mean squared error (RMSE), mean absolute error (MAE), the coefficient of determination (R^2), and bias (Petchey et al. 2015, Dietze 2017a). To fully assess probabilistic forecasts, the continuous ranked probability score (CRPS) and ignorance can also be used (Roulston and Smith 2002, Gneiting et al. 2005). Among these, only R^2 allows comparisons between forecasts that have different native units or forecasts for the same variable in very different ranges. Furthermore, more papers reported Pearson's r or R^2 (n = 56, 42%) than any other forecast performance metric in this dataset: for comparison, only 34% included RMSE and 20% included MAE. While the fact that R^2 is bias-corrected makes it an imperfect metric of forecast performance, it remains widely reported and uniquely suited to inter-study comparisons.

We recorded all R^2 and Pearson's r data reported in papers in the dataset. Pearson's r values were squared to yield R^2 (following Rousso et al. 2020). Because a majority of papers that reported R^2 data had forecast horizons between one and seven days (n = 35; 56%), we restricted our analysis to papers within this range of time horizons. We selected all forecast variables that had at least three papers represented over this interval and plotted forecast performance (in R^2) as a function of forecast horizon for these variables. Because some papers reported R^2 individually for each plot, site, or year and others reported one overall evaluation per model, we averaged all R^2 across sites and years for forecasts that used the same model within each paper.

We used indicator variable analysis (Draper and Smith 1998) to compare the slope of R^2 values over 1–7 day horizons among forecast variables by performing a 50% quantile regression predicting R^2 based upon indicator ("dummy") predictors for all forecast variables, as well as terms for the interaction between all forecast variables and forecast horizon. Quantile regression was performed using the package "quantreg" in R (Koenker et al. 2021). Indicator analysis compares the slope and intercept of the first indicator ("reference" indicator) to all subsequent indicators (Draper and Smith 1998). In this case, chlorophyll was used as the reference indicator to enable comparisons between phytoplankton and chlorophyll, two closely related forecast variables. We analyzed which terms were significant in the model to determine how patterns in forecast performance over time differed among forecast variables: significance was determined using the "wild" bootstrapping method to account for heteroskedasticity (Feng et al. 2011).

RESULTS

Dataset description

The number of ecological forecasts published each year has increased substantially over time: more papers were published in the last seven years of the dataset (2014–2020) than in the first 82 years (1932–2013; Fig. 2). Forecast sites for these papers were located on all seven continents (Fig 3a). The majority of forecast sites were located in the northern hemisphere (n = 211, 91%), especially the United States, China, and Western Europe (Fig. 3a). The geographic scale of the forecasts was most often either point (n = 66, 37%), or regional (n = 66, 37%; Fig. 3b).

More forecasts predicted organismal (population and community) variables than biogeochemical variables. Very few papers included forecasts for both biogeochemical and organismal focal variables (organismal: n = 146, 82%; biogeochemical: n = 43, 24%; both: n = 11, 6%; Fig. 3c). The majority of papers predicted ecological processes in either marine (n = 49, 28%), freshwater (n = 41, 23%), or agricultural (n = 34, 19%) ecosystems (Fig. 3). In particular, many papers predicted fish taxa (n = 25), phytoplankton taxa (n = 20), chlorophyll (n = 14), evapotranspiration (n = 14), pollen (n = 10), and crop yield (n = 9).

Papers in this dataset included forecasts at a wide range of forecast horizons and were developed using diverse time steps, forecast horizons, and datasets. Among the forecasts surveyed in this analysis, 75% of papers predicted within one year into the future (n = 130; Fig. 4). In particular, many papers either predicted 2–7 days into the future on a daily time step (n = 39, 23% of all papers) or one year into the future on a yearly time step (n = 30, 17%; Fig. 4). The median temporal duration of data used to create a forecasting paper (summed over model development, training, evaluation, etc.) was 15 years (min. = 17 days, mean = 19.2 years, max. = 145 years; Fig. 5), and 60% of papers (n = 107) used more than 10 years of data in the forecast paper.

The 178 papers included in this analysis were published in 114 unique journals and conference proceedings (103 journals, 11 conferences). The journal with the greatest number of papers represented in the dataset was *Ecological Applications*, which published a total of 14 near-term ecological forecasting papers.

Adoption of proposed best practices is low but increasing over time

Overall rates of proposed best practice use are low but may be increasing. On average, papers used two of the proposed best practices (median and mode = 2, mean = 2.37), but there was considerable variation: 12 papers did not use any of the best practices, and one paper used seven of the best practices. The probability that a paper would use a best practice in the year 2020 (as calculated by logistic regression) did not exceed 50% for any practice except “Assess and report forecast skill” (Fig. 6). All but one of our proposed best practices have been increasingly adopted over time. However, the increase in adoption with time was only statistically significant ($p < 0.05$) for three practices: “Automate forecasting workflows,” “Archive forecasts,” and “Make data available” (Fig. 6; Table 1).

Both of the *Forecast Requirement* best practices (“Include uncertainty” and “Assess and report forecast skill”) show a positive trend in adoption, though neither had a statistically significant relationship with publication year (Fig. 6; Table 1). Overall, “Include uncertainty” was demonstrated in 45% of papers and “Assess and report forecast skill” was demonstrated in 75% of papers. When uncertainty was included in forecasts, the most commonly included uncertainty sources were observation uncertainty (48%), process uncertainty (40%), and parameter uncertainty (35%). Driver uncertainty was included in 23% of papers that report uncertainty, and initial condition uncertainty was included in 18%. Of the papers that reported uncertainty (n = 80), 55% did not specify a data-driven origin of this uncertainty (e.g., ensemble model parameters, forecasted meteorological driver data) in the text. Over 70% of forecasts that did not report forecast evaluation in the text predicted at forecast horizons of at least one year; in comparison, 47% predicted at forecast horizons of at least one year in the dataset as a whole. As noted in the Methods, the most commonly reported metric of forecast performance was R^2.

Within the *Decision Support* tier, all three proposed best practices ("Identify an end user", "Make iterative forecasts", "Automate forecasting workflows") showed positive relationships with publication year, but only "Automate forecasting workflows" significantly increased over time (Fig. 6, Table 1). Overall, 20% of papers identified an end user, 39% of papers made iterative forecasts, and 11% of papers included automated forecasting workflows. Of the papers that mentioned a specific end user (n = 35), 31% mentioned that the end user aided in forecast development and 46% mentioned that forecasts were in use by the end user. Data assimilation for iterative forecasts most often updated initial conditions but not parameters of the model (67% of iterative forecasts only updated initial conditions). However, data assimilation methods that updated the parameters of the model (not just initial conditions) have increased significantly over time (Table 1).

Two of the four *Research* tier practices have increased significantly over time ("Archive forecasts" and "Make data available"; Table 1). "Partition uncertainty" has not significantly increased over time, though the only three papers that partitioned uncertainty were published in or after 2014. "Use null model comparisons" was the only practice that has decreased in adoption over time (Fig. 6). Overall, there was a wide range in the percentage of papers that used Research tier best practices. "Archive forecasts" was demonstrated in 8% of papers, "Make data available" was demonstrated in 25% of papers, "Partition uncertainty" was demonstrated in 2% of papers, and "Use null model comparisons" was demonstrated in 12% of papers. For papers that described forecast archiving (n = 15), the most common repository for archived forecasts was Zenodo (used in 20% of papers that archive forecasts); other papers used websites specific to the forecasting project to archive their forecasts. All of the papers that partitioned uncertainty (n = 3) quantified the influence of process, initial condition, and parameter uncertainty, and only one partitioned driver uncertainty. Process uncertainty dominated total uncertainty for two papers, while initial condition uncertainty dominated in the third paper. Of the papers that used null models in this study (n = 21), 62% used persistence null models and 48% used climatology null models. Two papers used both persistence and climatology null models.

Declines in forecast performance over increasing forecast horizons differ between variables

Forecast skill data (R^2) from at least three papers on 1–7 day forecast horizons were available for four forecast variables: chlorophyll, phytoplankton taxa, pollen, and evapotranspiration (ET). Forecast performance decreased with increasing forecast horizon for all forecast variables (Fig. 7, Table 2). The slope and intercept of forecast skill over increasing forecast horizons differed between variables, as revealed in our indicator analysis: the intercepts for pollen and ET were significantly lower than for chlorophyll, the reference indicator. In comparison to chlorophyll, forecast skill for pollen and ET decreased significantly more slowly over time. Unsurprisingly, the intercept and slope of phytoplankton were not significantly different from the intercept and slope of chlorophyll, the reference indicator (Fig. 7, Table 2).

DISCUSSION

Our systematic analysis of 178 near-term ecological forecasting papers demonstrates that the field of near-term ecological forecasting is widespread and growing: forecasts have been produced on all seven continents, and the rate of forecast publication is increasing over time. Although the overall implementation of proposed best practices is low, best practice use is increasing. In particular, the use of automated forecasting workflows, making data available, and archiving forecasts are all increasing significantly over time. Using this dataset of published studies we were able to compare forecast skill across scales and variables, and we found that forecast skill decreased in consistent patterns over 1–7 day forecast horizons. Variables that were closely related (i.e., chlorophyll and phytoplankton) displayed very similar trends in predictability over increasing forecast horizons, while more distantly related variables (i.e., pollen, evapotranspiration) exhibited significantly different patterns.

Near-term ecological forecasting: state of the field

As publication of near-term ecological forecasts continues to accelerate, evaluating the state of the field now can provide critical insight to help prioritize areas of improvement moving forward. Below we discuss aspects of near-term ecological forecasting that are well-developed, those that are improving over time, and areas that may need improvement based upon the results of this analysis.

Well-developed practices in near-term ecological forecasting: high rates of forecast assessment

Only one out of the nine proposed best practices was demonstrated in more than half of the papers in this analysis: assess and report forecast skill. In this study, high rates of forecast skill assessment and reporting allowed us to compare forecast performance across scales and variables. While R^2 was the most commonly reported forecast skill metric and served as an effective skill score for this preliminary analysis, it would be preferable to use a metric of forecast performance that is not bias-corrected. Other scale-independent metrics of forecast performance include percentage errors (e.g., mean absolute percentage error; MAPE) or scaled errors (e.g., mean absolute scaled error; MASE; Hyndman and Athanasopoulos 2018). Percentage errors are not universally applicable because they penalize a lack of precision more heavily in a range closer to zero (in the units of the forecast), which is not valid for a number of forecast variables (e.g., temperature in units of Fahrenheit or Celsius; Hyndman and Athanasopoulos 2018). Scaled errors may present the most effective means of comparing forecasts with different ranges and units, however, they require choosing a relevant null model (Hyndman and Athanasopoulos 2018), which is currently not common in near-term ecological forecasting literature (Fig. 6).

Of the papers that did not assess and report any metric of forecast skill, many (77%) predicted at forecast horizons greater than or equal to one year, suggesting that part of the reason these papers did not assess forecast skill may be the long lag before data would be available for forecast evaluation. In cases when the forecast horizon is too far into the future to evaluate results, researchers could consider making and evaluating additional forecasts at short time horizons to provide at least an intermediate evaluation of their forecasting system (Harris et al. 2018). Assessing hindcasts of historical data may also provide a means of evaluating the forecasting system, given sufficient historical data.

Developments in near-term ecological forecasting: high rates of forecast assessment, increasing automation and use of open science practices

Over time, near-term ecological forecasting is becoming increasingly automated, creating forecast products that enable real-time decision support (Dietze et al. 2018). Forecast automation can be beneficial to decision support because it decreases the amount of manual effort required to create each forecast once the automated system is set up and thereby increases the sustainability of iterative forecasting workflows (White et al. 2019, Hobday et al. 2019, Carey et al. 2021). However, it is important to note that automated forecasting workflows may still require significant human effort to maintain cyberinfrastructure over time (Carey et al. 2021). While the increase in use of iterative forecasts over time was not statistically significant, the percentage of papers that use iterative workflows to update model parameters rather than just the initial conditions of the forecast has increased significantly (Table 1). Updating model parameters as new data are incorporated allows the forecasting system to learn over time and potentially make more accurate predictions in the future (Luo et al. 2011, Niu et al. 2014, Zwart et al. 2019).

Another area of promise is in the adoption of open scientific practices: both data publication and forecast archiving have increased significantly over the past 40 years. These advances likely reflect a broader movement for open scientific practices that has gained momentum over the past decade in response to intersecting needs for greater reproducible science, knowledge dissemination, and collaboration (e.g., Reichman et al. 2011, Fecher and Friesike 2013, Beardsley 2014, Wilkinson et al. 2016, Munafò et al. 2017, Powers and Hampton 2019). Further increases in the use of open scientific practices have the potential to increase the reproducibility of published forecasting literature while fostering collaboration and accelerating the development of the field.

Priorities for the future development of near-term ecological forecasting: uncertainty, end user engagement, and null models

One of the most notable gaps identified in this analysis is the lack of specified uncertainty in published forecasts. Meaningful representations of uncertainty are considered so critical to forecast interpretation and evaluation that many definitions of ecological forecasts include uncertainty as an essential component (e.g., Clark et al. 2001, Luo et al. 2011, Harris et al. 2018, Dietze et al. 2018, Carey et al. 2021). However, only 45% of papers included uncertainty in their forecasts. Lack of forecast uncertainty can be problematic in decision support because when uncertainty is not specified in a forecast output, forecast users create their own, often inaccurate, expectations of forecast uncertainty (Morss et al. 2008, Joslyn and Savelli 2010).

In addition to aiding decision support, uncertainty partitioning has the potential to contribute to our understanding of the limits of ecological predictability. Partitioning uncertainty into its respective components (e.g., initial condition, driver, parameter, and process uncertainty) allows researchers to understand the constraints to predictability for a given ecological variable and analyze how those constraints differ between forecast scales and horizons (Dietze 2017b). For example, it is well established that forecasting meteorological conditions is a fundamentally chaotic problem, dominated by initial condition uncertainty (Kalnay 2003). Due to stabilizing feedbacks in ecological systems (e.g., carrying capacity, functional redundancy), other components of uncertainty are hypothesized to dominate ecological forecasts (Dietze 2017b). This hypothesis is partially supported by our dataset: initial condition uncertainty was the dominant source of forecast uncertainty in only one of three papers that partitioned uncertainty in this analysis. However, because uncertainty partitioning is a relatively new development in ecological forecasting, the small number of studies that partition uncertainty currently prevents us from making conclusions about the limiting factors for predictability across ecosystems and forecast horizons.

While not all of the near-term ecological forecasts described in our dataset were designed for decision support, approximately 20% of papers mentioned a specific end user for their forecasts. Of these, nearly half specify that the forecasting system was in use by the specified end user (e.g., drinking water management organization, coral reef conservation agency, etc.). Designing forecasts for end users involves a variety of ethical considerations, including equity for end users, communication of forecast skill, and stakeholder education (Hobday et al. 2019). However, it was rare for a paper to include any explicit mention of ethical considerations made in designing the forecast (5% of papers overall; 25% of forecasts that are in use by an end user).

Given the power of forecasts to inform decision support, education on how to carefully navigate these decisions may be useful in improving the utility of forecasts for stakeholder use.

In this study, we found that the use of null model comparisons remains low and has not increased in adoption over time, despite the importance of this practice for contextualizing model skill (Harris et al. 2018, Dietze et al. 2018, White et al. 2019). For meteorological forecasting, comparing forecasts to a climatological null model has proved useful as a means of analyzing the limit of predictive skill and the comparative performance of weather forecasts across spatial and temporal scales (Buizza and Leutbecher 2015). Parallel advances in ecological forecasting may be enabled through increased use of null model comparisons in the future (Petchey et al. 2015, Hyndman and Athanasopoulos 2018).

Published forecasts provide insight into forecastability across ecosystems and models

Analyzing forecastability across variables, we found that chlorophyll and phytoplankton taxa were more predictable than pollen and evapotranspiration at the shortest time horizons (chlorophyll: 1–5 days; phytoplankton: 1–7 days). However, the predictability of chlorophyll and phytoplankton decayed faster over increasing forecast horizons than that of evapotranspiration and pollen. Similar patterns in forecast performance for chlorophyll and phytoplankton likely result from the fact that these two ecological variables are closely related. Greater predictability of chlorophyll and phytoplankton than evapotranspiration and pollen at short time horizons likely indicates a greater degree of autocorrelation in these processes (Reynolds 2006), but predictability quickly decays over time due to bloom dynamics (e.g., Rigosi et al. 2011, Recknagel et al. 2016). The consistency of these patterns across 3–10 different papers for each forecast variable suggests that the relationship between forecast performance and forecast horizon could be a robust indicator of the predictability of other ecological processes.

While this is a preliminary analysis limited to four ecological variables, it is among the first comparative tests that have analyzed forecastability across scales and variables, building on two previously published studies. Ward et al. (2014) analyzed the ability of multiple time-series models to predict 2379 vertebrate population datasets. They found that increased forecast performance (measured using MASE) was correlated with long lifespans and large body size for fish and high trophic level for birds over 1–5 year forecast horizons. Additionally, Rousso et al. (2020) performed a systematic review of cyanobacterial bloom forecasting literature and analyzed the relationship between forecast performance (R^2) and forecast horizon for three types of models: artificial neural networks, decision trees, and genetic programming. They found that forecast performance decreased over 1–30 day forecast horizons, and forecasts created using greater amounts of historical data had superior forecast performance. Altogether, these first analyses of the forecastability of ecological variables highlight the growing applicability of forecasting to inform our understanding of ecological predictability.

Accelerating forecast publication and increased adoption of proposed best practices will increase the statistical strength of future analyses to detect trends in forecast performance over increasing forecast horizons, including possible non-linear patterns. In particular, increased assessment and reporting of forecast skill ensures that published papers can be included in a meta-analysis of predictability; increased data publication allows reevaluation of forecasts; increased forecast archiving addresses publication biases in forecast results; increased use of null models allows researchers to analyze how the maximum length of time until a forecast performs no better than the null differs among variables; and increased uncertainty partitioning allows researchers to compare how uncertainty sources differ across scales and variables.

Literature search process: observations and limitations

While the rates of adoption of these proposed best practices (Box 1) are low overall, they are not necessarily unexpected. Ecological forecasting is an emerging discipline and many of these methods are still in development; notably, our list of proposed best practices was derived from papers that were all published within the last three years of the dataset (Harris et al. 2018, Dietze et al. 2018, White et al. 2019, Hobday et al. 2019, Carey et al. 2021). Ideally, best practices should evolve using a community-driven approach to enable buy-in and robustness to many applications (following Hanson et al. 2016); consequently, our list is not exhaustive and some of the proposed practices may not be appropriate for every forecasting application. Armstrong (2001) proposed as many as 139 principles for forecasting at large, all of which could be relevant to ecological forecasting applications. If one had to wait to publish a forecast until it satisfied all potential best practices, it is likely that no forecasting papers would ever be published. Increasing the number of published ecological forecasts benefits the field of ecological forecasting even if forecasts do not follow all proposed best practices because the research community gains increasingly more information about the predictability of ecological variables and the tools and techniques needed to make effective forecasts (Dietze et al. 2018). Still, as near-term ecological forecasting continues to grow, assessing the rate of best practice adoption now allows researchers to identify and prioritize areas for growth and education, simultaneously advancing the basic and applied value of ecological forecasting.

Results from our literature search process highlight the decentralized nature of near-term ecological forecasting and the challenges associated with systematically reviewing this literature. The 178 near-term ecological forecasting papers in this analysis came from 114 unique journals and conference proceedings, and no one journal published more than 15 near-term ecological forecasts papers in this analysis. Decentralized forecast publications may present a barrier to those interested in this literature, particularly because there is no one search term that comprehensively surveys the current near-term ecological forecasting literature. Many papers do not explicitly use the terms "near-term" or "ecological" when describing forecasts for a particular application, and only 5% of the results from our initial search for the term "forecast*" in ecology-related journal articles were identified as near-term ecological forecasts after two rounds of review (Fig. 1). By systematically reviewing and synthesizing near-term ecological forecasting literature published to date, we aim to begin addressing this gap.

Importantly, this comprehensive analysis of near-term ecological forecasting literature is limited to published forecast results. Operational forecasting systems that have not been described in peer-reviewed literature were not included (e.g., the U.S. National Oceanic and Atmospheric Administration, NOAA, has multiple operational forecasting systems for harmful algal blooms, fisheries, and coral reef bleaching that are available via websites). We anticipate that this may affect results in at least three ways: first, because unpublished operational forecasting systems are often used for decision support, the percentage of forecasting systems that connect to a specific end user is likely underrepresented in published literature. Second, both the need to refine forecasting models prior to paper submission and reviewer requests for forecast revisions may make it difficult to publish genuine forecasts. Because of this, most papers in this study are likely hindcasts or forecast reanalyses. Third, because of publication biases (Dickersin 1990, Mlinarić et al. 2017), we anticipate that average forecast performance is artificially inflated in published literature relative to unpublished operational forecasts. As coordination within the near-term ecological forecasting discipline increases, surveying and comparing operational forecasts may become increasingly possible over time.

Future needs in near-term ecological forecasting

Looking to the future, advancing the field of near-term ecological forecasting will involve a suite of technological, organizational, and educational advances. First, the low level of adoption of the proposed best practices suggests that increased coordination within the ecological forecasting research community for developing common forecasting standards, best practices, and vocabulary will advance near-term ecological forecasting. Second, we find that incorporation and partitioning of uncertainty and null models are critical gaps in ecological forecasting literature where education may be needed. The creation of additional educational resources will enable more forecasts to be created and facilitate the adoption of best practices in ecological forecasting. Third, our analysis strongly suggests that long-term data are an important resource for near-term ecological forecast development and assessment. In our dataset, 60% of published near-term ecological forecasting studies used >10 years of ecological data when developing, calibrating, and assessing their forecasts (Fig. 5). Long-term support for data collection will likely be necessary to advance the field. Finally, our analysis indicates that near-term ecological forecasting may be disproportionately centered in the northern hemisphere, particularly the United States of America, Western Europe, and China. This result follows the disproportionate representation of these geographic regions across all sciences (UNESCO 2015). Lack of forecast locations in other countries, particularly in the southern hemisphere, is a detriment to the field as a whole, as the full diversity of ecological systems is not represented in ecological forecasting research to date.

While there are a variety of challenges and opportunities facing the development of near-term ecological forecasting, the literature indicates that the field has grown significantly over the past 90 years. Near-term ecological forecasting is now widespread and the rate of forecast publication continues to increase over time. Moving forward, near-term ecological forecasting is well-positioned to transform ecological management and provide critical insight into the predictability of ecological systems.

ACKNOWLEDGMENTS

We thank members of the Ecological Forecasting Initiative (EFI) for productive discussions throughout the process of writing this paper and Arpita Das for helping compile R^2 values. We are also grateful for the assistance of the Virginia Tech libraries in planning this review and accessing the papers in the analysis. This project was supported by U.S. National Science Foundation (NSF) grants 1737424, 1926050, 1926388, 1933016, and 1933102, as well as NSF Graduate Fellowship Program fellowships to ASL and WMW (DGE-1651272).

Authorship contributions: This paper originated in an Ecological Forecasting graduate seminar at Virginia Tech led by CCC and RQT; all authors participated in the seminar and co-developed the matrix used for this analysis. Conceptualization of the project was led by ASL, CCC, JS, RPM, RQT, and WMW. ASL, CCC, MEL, and WMW led the development of methods for the project. All authors contributed to abstract screening and paper review. Data analysis was led by ASL, with input from CCC, HLW, and WMW. ASL, CCC, DWH, HLW,

MEL, RPM, and WMW contributed to initial development of the manuscript, which was written by ASL. All authors edited and approved the final manuscript.

REFERENCES

Armstrong, J. S. 2001. Principles of forecasting: a handbook for researchers and practitioners. Springer Science & Business Media, New York.

Beardsley, T. M. 2014. Are you ready for open data? BioScience 64:855–855.

Beckage, B., L. J. Gross, and S. Kauffman. 2011. The limits to prediction in ecological systems. Ecosphere 2:art125.

Buizza, R., and M. Leutbecher. 2015. The forecast skill horizon. Quarterly Journal of the Royal Meteorological Society 141:3366–3382.

Carey, C. C., W. M. Woelmer, M. E. Lofton, R. J. Figueiredo, B. J. Bookout, R. S. Corrigan, V. Daneshmand, A. G. Hounshell, D. W. Howard, A. S. L. Lewis, R. P. McClure, H. L. Wander, N. K. Ward, and R. Q. Thomas. 2021. Advancing lake and reservoir water quality management with near-term, iterative ecological forecasting. Inland Waters. doi: 10.1080/20442041.2020.1816421

Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657–660.

Dickersin, K. 1990. The existence of publication bias and risk factors for its occurrence. JAMA 263:1385–1389.

Dietze, M. C. 2017a. Ecological forecasting. Princeton University Press, Princeton.

Dietze, M. C. 2017b. 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.

Draper, N. R., and H. Smith. 1998. Applied regression analysis. John Wiley & Sons.

Fecher, B., and S. Friesike. 2013. Open science: one term, five schools of thought.

Feng, X., X. He, and J. Hu. 2011. Wild bootstrap for quantile regression. Biometrika 98:995–999.

Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman. 2005. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Monthly Weather Review 133:1098–1118.

Godfray, H. C. J., and R. M. May. 2014. Open questions: are the dynamics of ecological communities predictable? BMC Biology 12:22.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter. 2013. Big data and the future of ecology. Frontiers in Ecology and the Environment 11:156–162.

Hanson, P. C., K. C. Weathers, and T. K. Kratz. 2016. Networked lake science: how the Global Lake Ecological Observatory Network (GLEON) works to understand, predict, and communicate lake ecosystem response to global change. Inland Waters 6:543–554.

Harris, D. J., S. D. Taylor, and E. P. White. 2018. Forecasting biodiversity in breeding birds using best practices. PeerJ 6:e4278.

Henden, J.-A., R. A. Ims, N. G. Yoccoz, E. J. Asbjørnsen, A. Stien, J. P. Mellard, T. Tveraa, F. Marolla, and J. U. Jepsen. 2020. End-user involvement to improve predictions and management of populations with complex dynamics and multiple drivers. Ecological Applications 30:e02120.

Hobday, A. J., J. R. Hartog, J. P. Manderson, K. E. Mills, M. J. Oliver, A. J. Pershing, and S. Siedlecki. 2019. Ethical considerations and unanticipated consequences associated with ecological forecasting for marine resources. ICES Journal of Marine Science 76:1244–1256.

Hodgson, W. C. 1932. The forecasting of the East Anglian herring fishery. Journal of Animal Ecology 1:108–118.

Houlahan, J. E., S. T. McKinney, T. M. Anderson, and B. J. McGill. 2017. The priority of prediction in ecological understanding. Oikos 126:1–7.

Hyndman, R. J., and G. Athanasopoulos. 2018. Forecasting: principles and practice. OTexts.

Joslyn, S., and S. Savelli. 2010. Communicating forecast uncertainty: public perception of weather forecast uncertainty. Meteorological Applications 17:180–195.

Kalnay, E. 2003. Atmospheric modeling, data assimilation and predictability. Cambridge University Press, New York.

Koenker, R., S. Portnoy, P. T. Ng, B. Melly, A. Zeileis, P. Grosjean, C. Moler, Y. Saad, V. Chernozhukov, I. Fernandez-Val, and B. D. Ripley. 2021. quantreg: Quantile Regression.

LaDeau, S. L., B. A. Han, E. J. Rosi-Marshall, and K. C. Weathers. 2017. The next decade of big data in ecosystem science. Ecosystems 20:274–283.

Lindenmayer, D. B., G. E. Likens, A. Andersen, D. Bowman, C. M. Bull, E. Burns, C. R.

Dickman, A. A. Hoffmann, D. A. Keith, M. J. Liddell, A. J. Lowe, D. J. Metcalfe, S. R. Phinn, J. Russell-Smith, N. Thurgate, and G. M. Wardle. 2012. Value of long-term ecological studies. Austral Ecology 37:745–757.

Luo, Y., K. Ogle, C. Tucker, S. Fei, C. Gao, S. LaDeau, J. S. Clark, and D. S. Schimel. 2011. Ecological forecasting and data assimilation in a data-rich era. Ecological Applications 21:1429–1442.

Mlinarić, A., M. Horvat, and V. Šupak Smolčić. 2017. Dealing with the positive publication bias: why you should really publish your negative results. Biochemia Medica 27:447–452.

Morss, R. E., J. L. Demuth, and J. K. Lazo. 2008. Communicating uncertainty in weather forecasts: a survey of the U.S. public. Weather and Forecasting 23:974–991.

Munafò, M. R., B. A. Nosek, D. V. M. Bishop, K. S. Button, C. D. Chambers, N. Percie du Sert, U. Simonsohn, E.-J. Wagenmakers, J. J. Ware, and J. P. A. Ioannidis. 2017. A manifesto for reproducible science. Nature Human Behaviour 1:1–9.

Niu, S., Y. Luo, M. C. Dietze, T. F. Keenan, Z. Shi, J. Li, and F. S. C. Iii. 2014. The role of data assimilation in predictive ecology. Ecosphere 5:art65.

Pennekamp, F., A. C. Iles, J. Garland, G. Brennan, U. Brose, U. Gaedke, U. Jacob, P. Kratina, B. Matthews, S. Munch, M. Novak, G. M. Palamara, B. C. Rall, B. Rosenbaum, A. Tabi, C. Ward, R. Williams, H. Ye, and O. L. Petchey. 2019. The intrinsic predictability of ecological time series and its potential to guide forecasting. Ecological Monographs 89:e01359.

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.

Petrovskii, S., and N. Petrovskaya. 2012. Computational ecology as an emerging science. Interface Focus 2:241–254.

Pielke, R. A., and R. T. Conant. 2003. Best practices in prediction for decision-making: lessons from the atmospheric and earth sciences. Ecology 84:1351–1358.

Powers, S. M., and S. E. Hampton. 2019. Open science, reproducibility, and transparency in ecology. Ecological Applications 29:e01822.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Recknagel, F., R. Adrian, J. Köhler, and H. Cao. 2016. Threshold quantification and short-term forecasting of Anabaena, Aphanizomenon and Microcystis in the polymictic eutrophic Lake Müggelsee (Germany) by inferential modelling using the hybrid evolutionary algorithm HEA. Hydrobiologia 778:61–74.

Reichman, O. J., M. B. Jones, and M. P. Schildhauer. 2011. Challenges and opportunities of open data in ecology. Science 331:703–705.

Reynolds, C. S. 2006. The ecology of phytoplankton. Cambridge University Press.

Rigosi, A., R. Marcé, C. Escot, and F. J. Rueda. 2011. A calibration strategy for dynamic succession models including several phytoplankton groups. Environmental Modelling & Software 26:697–710.

Roulston, M. S., and L. A. Smith. 2002. Evaluating probabilistic forecasts using information theory. Monthly Weather Review 130:1653–1660.

Rousso, B. Z., E. Bertone, R. Stewart, and D. P. Hamilton. 2020. A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes. Water Research 182:115959.

Sutherland, W. J., R. P. Freckleton, H. C. J. Godfray, S. R. Beissinger, T. Benton, D. D. Cameron, Y. Carmel, D. A. Coomes, T. Coulson, M. C. Emmerson, R. S. Hails, G. C. Hays, D. J. Hodgson, M. J. Hutchings, D. Johnson, J. P. G. Jones, M. J. Keeling, H. Kokko, W. E. Kunin, X. Lambin, O. T. Lewis, Y. Malhi, N. Mieszkowska, E. J. Milner-Gulland, K. Norris, A. B. Phillimore, D. W. Purves, J. M. Reid, D. C. Reuman, K. Thompson, J. M. J. Travis, L. A. Turnbull, D. A. Wardle, and T. Wiegand. 2013. Identification of 100 fundamental ecological questions. Journal of Ecology 101:58–67.

Ward, E. J., E. E. Holmes, J. T. Thorson, and B. Collen. 2014. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. Oikos 123:652–661.

White, E. P., G. M. Yenni, S. D. Taylor, E. M. Christensen, E. K. Bledsoe, J. L. Simonis, and S. K. M. Ernest. 2019. Developing an automated iterative near-term forecasting system for an ecological study. Methods in Ecology and Evolution 10:332–344.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. ’t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B.

Mons. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3:160018.

Zwart, J. A., O. Hararuk, Y. T. Prairie, S. E. Jones, and C. T. Solomon. 2019. Improving estimates and forecasts of lake carbon dynamics using data assimilation. Limnology and Oceanography: Methods 17:97–111.

Box 1: Proposed best practices for ecological forecasting, drawn from peer-reviewed literature: Dietze et al. (2018), Harris et al. (2018), White et al. (2019), Hobday et al. (2019), and Carey et al. (2021). *Forecast Requirements* include traits of forecasting systems that are essential to the development of a forecast. *Decision Support* practices are those that are particularly helpful if the forecast will be applied as a decision support tool. *Research* practices include characteristics of a forecasting system that are particularly suited to enabling the advancement of fundamental ecological understanding across studies. Importantly, these last two tiers are not mutually exclusive: decision support practices can also be important for ecological understanding and vice versa.

Forecast Requirements

1. Include uncertainty
    a. Meaningful representations of uncertainty are important to forecast interpretation and evaluation, so much so that in many definitions uncertainty is identified as an essential component of an ecological forecast (Clark et al. 2001, Luo et al. 2011, Harris et al. 2018, Dietze et al. 2018, Carey et al. 2021).
2. Assess and report forecast skill
    a. Assessing and reporting forecast skill allows end users to understand the reliability of the forecasting system (Harris et al. 2018, Hobday et al. 2019) and provides the near-term ecological forecasting research community with increased insight into the tools and techniques needed to produce effective forecasts (Dietze et al. 2018). Furthermore, assessing and reporting forecast skill contributes to our understanding of ecological predictability by facilitating comparisons of forecast skill across scales and variables (Beckage et al. 2011, Petchey et al. 2015).

Decision Support

3. Identify an end user
    a. One of the goals of ecological forecasting is to aid in decision-making. The first step in this process is to identify an end user and consider their needs throughout forecast development (Dietze et al. 2018, Hobday et al. 2019, Carey et al. 2021).
4. Make iterative forecasts
    a. Iterative forecasts incorporate new data as they become available, providing updated predictions that can aid in continuous decision-making and forecast improvement (Dietze et al. 2018).
5. Automate forecasting workflows
    a. End-to-end automation of the forecasting workflow (from data processing to forecast communication) allows for more frequent forecast outputs and more sustainable forecasting infrastructure (Dietze et al. 2018, White et al. 2019). This practice is particularly relevant for forecasts with horizons of days to months that are rerun often to provide updated information.

Research

6. Make data available
    a. To ensure full forecast reproducibility and allow future research to build off of existing forecasting workflows, all data and code used to create forecasts should be findable, accessible, interoperable, and reusable (FAIR; Wilkinson et al. 2016, Harris et al. 2018, Dietze et al. 2018, White et al. 2019, Carey et al. 2021).
7. Archive forecasts

    a. Archiving forecasts as they are created provides a means of demonstrating when forecasts were made and tracking forecast improvement over time (Dietze et al. 2018, White et al. 2019).

8. Use null model comparisons
    a. Comparing forecasts to simple, standard baselines (e.g., persistence or climatology null models) allows researchers to compare forecastability across systems (Petchey et al. 2015) and evaluate the amount of information contained in the forecasts (Harris et al. 2018, Dietze et al. 2018, White et al. 2019).

9. Partition uncertainty
    a. Partitioning uncertainty into its respective components (e.g., initial conditions, drivers, parameters, process) is helpful when prioritizing improvements to the forecasting system (Dietze et al. 2018, Carey et al. 2021), and it allows for the comparison of limitations to predictability across time horizons and ecosystems.

Table 1: Logistic regression results for each proposed best practice based on n = 177 papers (one paper from 1932 was excluded from this analysis). Statistically significant p values are in bold. In addition to the nine proposed best practices, statistics are included for the use of iterative forecasting to update model parameters.

	Estimate	Standard error	Z value	P value
Include uncertainty				
Intercept	-7.615	35.157	-0.217	0.83
Year	0.004	0.017	0.211	0.83
Assess and report forecast skill				
Intercept	-59.665	38.486	-1.550	0.12
Year	0.030	0.019	1.579	0.11
Identify an end user				
Intercept	-86.809	52.611	-1.650	0.10
Year	0.042	0.026	1.624	0.10
Make iterative forecasts				
Intercept	-14.147	36.152	-0.391	0.70
Year	0.007	0.018	0.379	0.71
Make iterative forecasts (updating model parameters)				
Intercept	-182.741	85.680	-2.133	**0.03**
Year	0.090	0.043	2.111	0.04
Automate forecasting workflows				
Intercept	-237.502	100.550	-2.362	**0.02**
Year	0.117	0.050	2.344	0.02
Make data available				
Intercept	-252.217	69.775	-3.615	**<0.001**
Year	0.125	0.035	3.602	<0.001
Archive forecasts				
Intercept	-308.891	136.546	-2.262	**0.02**
Year	0.152	0.068	2.247	0.03
Use null model comparisons				
Intercept	33.795	50.388	0.671	0.50
Year	-0.018	0.025	-0.710	0.48
Partition uncertainty				
Intercept	-384.450	355.406	-1.082	0.28
Year	0.189	0.176	1.071	0.28

Table 2: Indicator variable analysis results comparing the slope of R^2 values over 1–7 day horizons among forecast variables. Chlorophyll was treated as the reference variable for the analysis. Statistically significant p values are in bold.

	Estimate	SE	t	P value
Chlorophyll (reference; n = 68 data points from 8 papers)				
Intercept	0.98	0.025	39.53	**<0.001**
Horizon	-0.08	0.012	-6.78	**<0.001**
Phytoplankton (n = 33 data points from 8 papers)				
Intercept	0.04	0.055	0.71	0.48
Horizon	0.02	0.016	1.06	0.29
Pollen (n = 110 data points from 3 papers)				
Intercept	-0.29	0.043	-6.62	**<0.001**
Horizon	0.05	0.014	3.94	**<0.001**
Evapotranspiration (n = 113 data points from 10 papers)				
Intercept	-0.26	0.059	-4.40	**<0.001**
Horizon	0.05	0.015	3.46	**<0.001**

FIGURE LEGENDS

Figure 1: Systematic literature analysis methods. a: Flow chart of literature review process. b and c: Venn diagrams illustrating the number of studies that met each of our three criteria after two rounds of review (abstract and paper reviews) for our original Web of Science search (b) and a search of citing and cited papers (c).

Figure 2: Number of near-term ecological forecasts published per year. Five papers from the final year (2020) are not plotted because data for this year are incomplete: only papers indexed on Web of Science by the date of our search (18 May 2020) were included in this study.

Figure 3: Description of ecological forecasting papers identified in this study. a: Map of all forecasts: regional and national studies are shown in large transparent points near the center of the forecast region, while point and multipoint forecasts are shown in small opaque points. b: Bar chart illustrating the spatial extent of the forecast for each paper—point, multipoint (several distinct points), regional (a broad region that does not follow national bounds), national, or global (for details about how spatial extent was determined, see Lewis et al. Environmental Data Initiative repository forthcoming data publication). c: Bar chart illustrating the class—organismal (population or community) or biogeochemical—of the forecast variable for each paper. Fill colors illustrate ecosystem type. Forecasts that could not be matched to one of our nine ecosystem types have been labeled “other.”

Figure 4: Relationship between time step and time horizon of forecasting papers. Colors and numbers within each square indicate the number of papers corresponding to that combination of

time step and time horizon. White areas indicate combinations of time step and horizon that were not represented in the dataset. Papers that did not have a defined horizon (e.g., predicting end of summer harvest) or that did not specify time step/horizon were omitted (n = 10).

Figure 5: Histogram illustrating the total number of years of data used to develop each forecasting paper, summed across model development, training, evaluation, etc. Vertical lines represent the median (left) and mean (right) number of years used.

Figure 6: Best practice adoption over time. Points demarcate whether or not an individual paper demonstrated the best practice (1 = observed, 0 = not observed), and lines represent logistic regression results. Significance of the year term in the regression is indicated using asterisks: * indicates $p < 0.05$, *** indicates $p < 0.001$. One paper from 1932 was excluded from this analysis.

Figure 7: Relationship between forecast performance and forecast horizon (Horiz) for four forecast variables: chlorophyll (Chl), phytoplankton (Phyto), pollen, and evapotranspiration (ET). Different papers are indicated by points of different colors and shapes. Within a paper, forecasts using the same model were averaged (across sites, years, etc.). Rightmost panel: model predictions from the quantile regression indicator analysis.

Figure 1

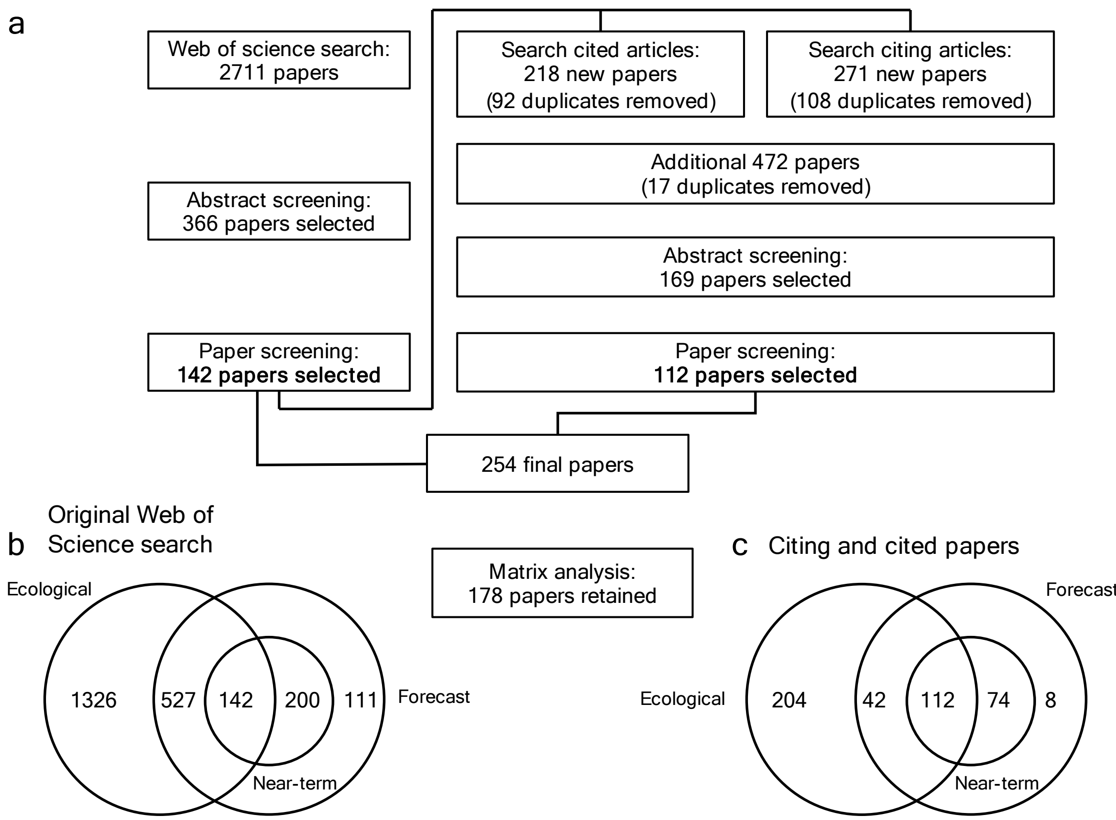
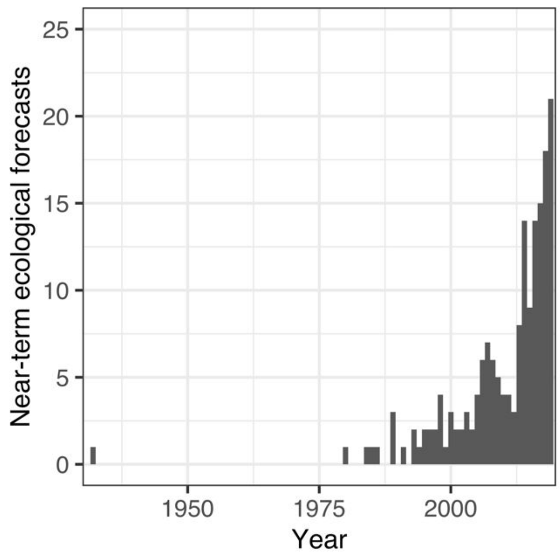
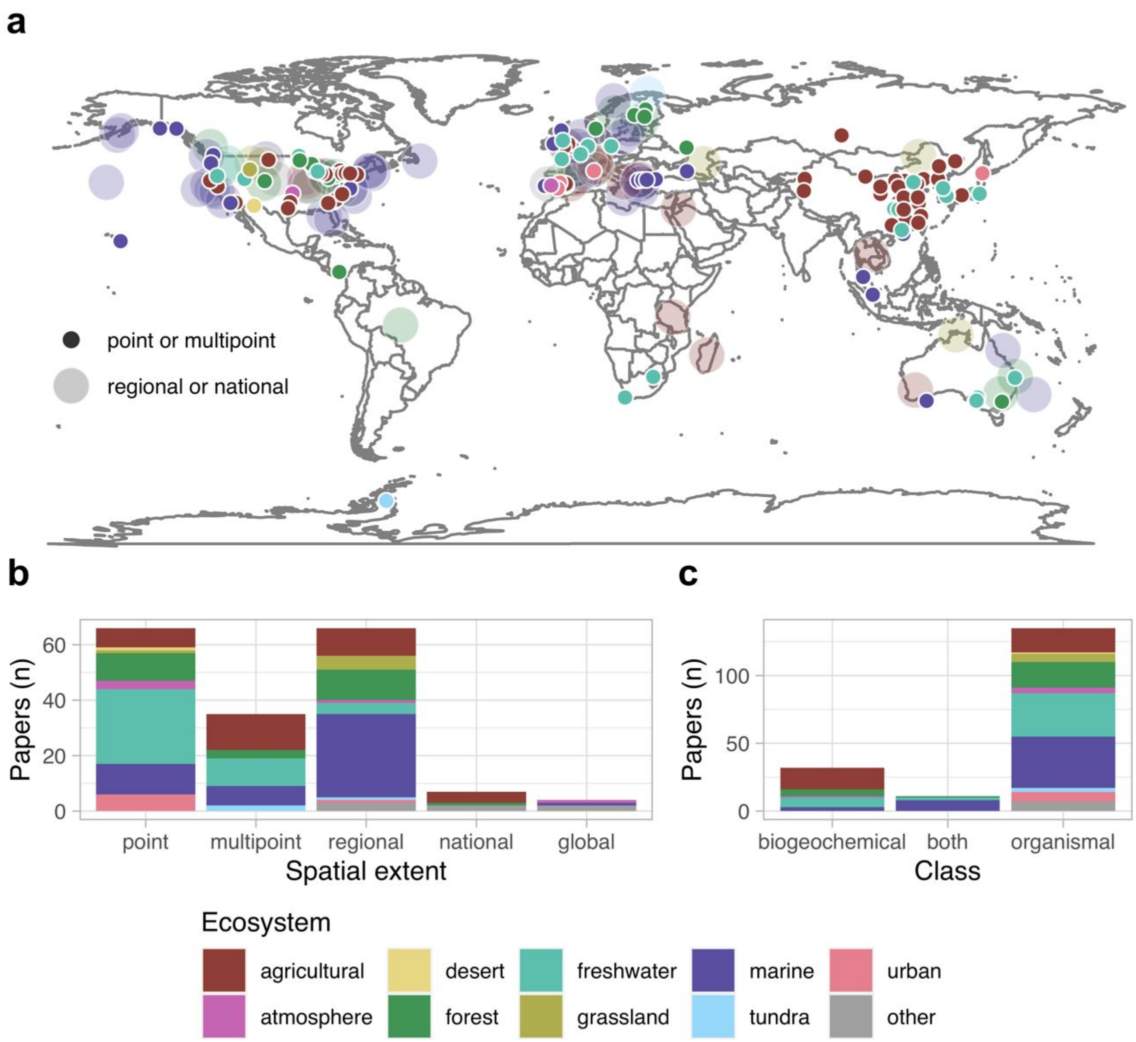

a
Web of science search:
2711 papers
Search cited articles:
218 new papers
(92 duplicates removed)
Search citing articles:
271 new papers
(108 duplicates removed)
Additional 472 papers
(17 duplicates removed)
Abstract screening:
366 papers selected
Abstract screening:
169 papers selected
Paper screening:
142 papers selected
Paper screening:
112 papers selected
254 final papers
b
Original Web of
Science search
Matrix analysis:
178 papers retained
c
Citing and cited papers
Ecological
1326
527
142
200
111
Forecast
Near-term
Forecast
Ecological
204
42
112
74
8
Near-term

Figure 2


Near-term ecological forecasts
0
5
10
15
20
25
1950
1975
2000
Year

873 Figure 3



874

Figure 4

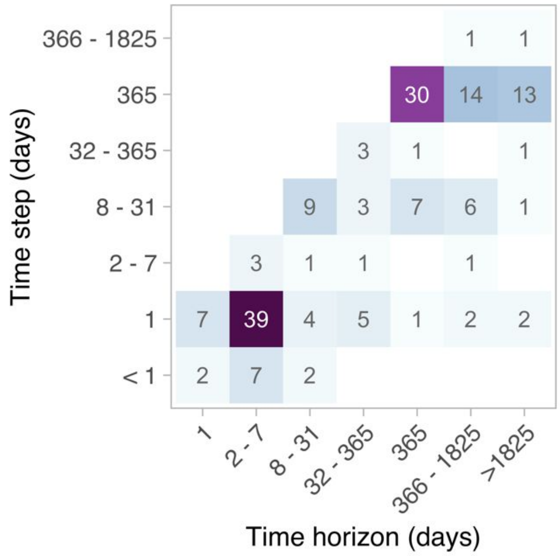
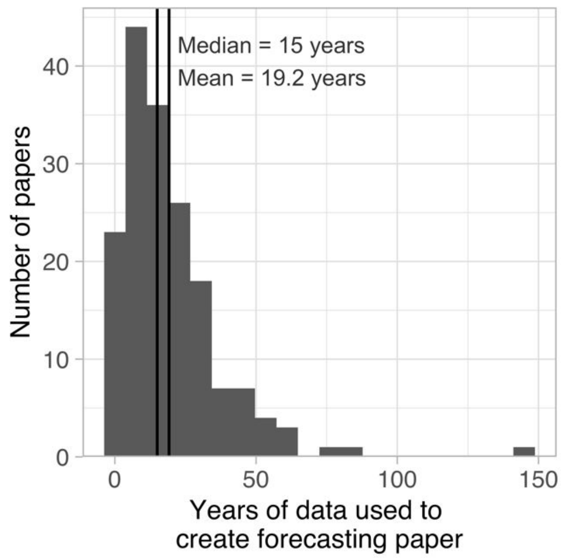
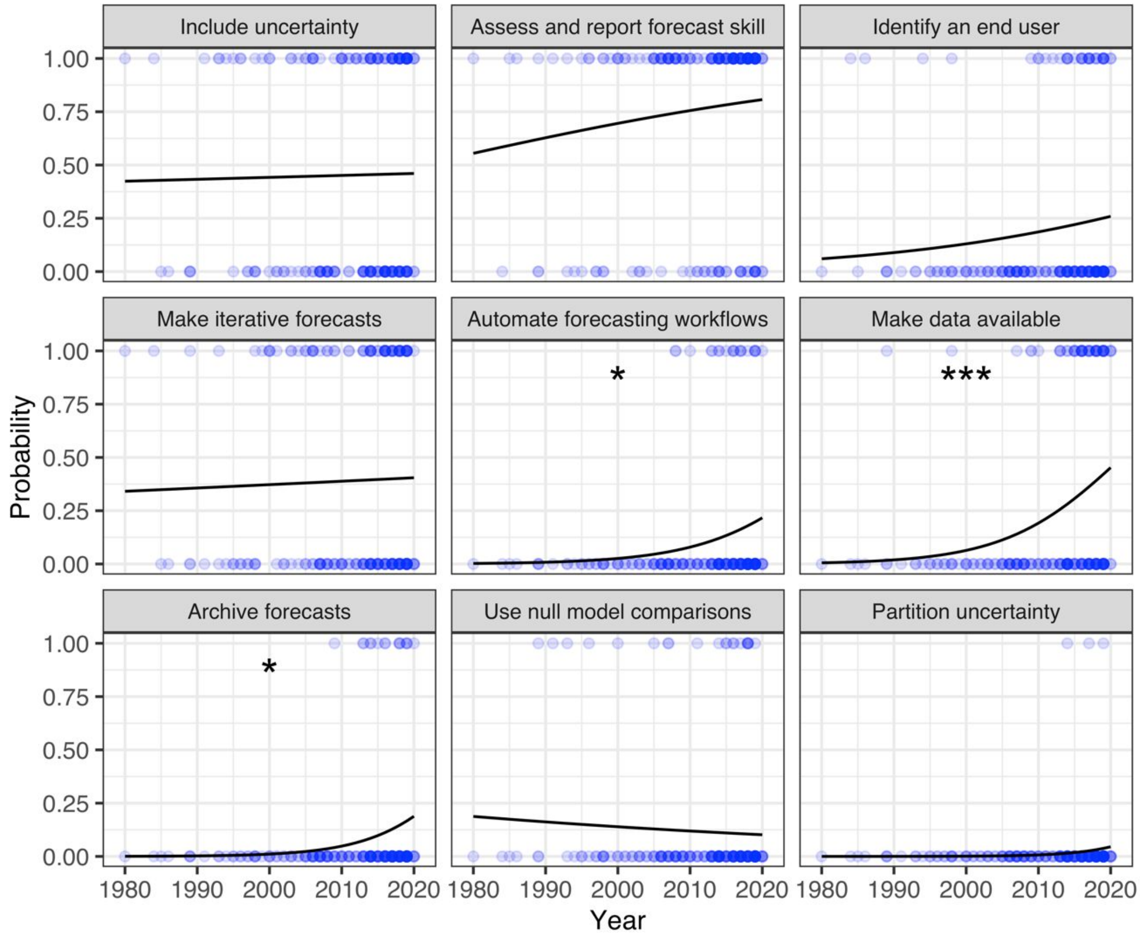

Time step (days)
366 - 1825
365
32 - 365
8 - 31
2 - 7
1
< 1
1 1
30 14 13
3 1 1
9 3 7 6 1
3 1 1 1
7 39 4 5 1 2 2
2 7 2
1
2 - 7
8 - 31
32 - 365
365
366 - 1825
>1825
Time horizon (days)

Figure 5

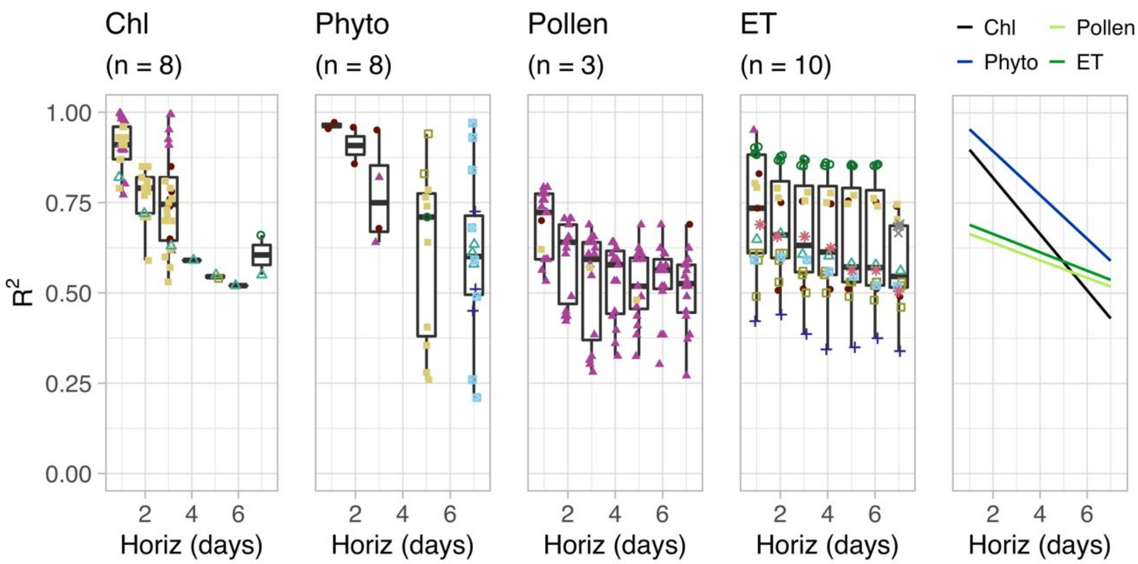

Median = 15 years
Mean = 19.2 years
Number of papers
0
10
20
30
40
0
50
100
150
Years of data used to
create forecasting paper

879 Figure 6



880

881 Figure 7



882

JOURNAL PUBLICATION CITATION: Lewis, A. S. L., W. M. Woelmer, H. L. Wander, D. W. Howard, J. W. Smith, R. P. McClure, M. E. Lofton, N. W. Hammond, R. S. Corrigan, R. Q. Thomas, C. C. Carey. 2021. Increased adoption of best practices in ecological forecasting enables comparisons of forecastability across systems. *Ecological Applications.*

Data S1

Data and code to reproduce best practice analysis and forecastability comparison

---

Author(s) [of the material provided in DataS1.zip]

Abigail S. L. Lewis
Department of Biological Sciences, Virginia Tech
926 W. Campus Dr., Blacksburg VA 24060
aslewis@vt.edu

Whitney M. Woelmer
Department of Biological Sciences, Virginia Tech
wwoelmer@vt.edu

Heather L. Wander
Department of Biological Sciences, Virginia Tech
hwander@vt.edu

Dexter W. Howard
Department of Biological Sciences, Virginia Tech
dwh1998@vt.edu

John W. Smith
Department of Statistics, Virginia Tech
wsjohn2@vt.edu

Ryan P. McClure
Department of Biological Sciences, Virginia Tech
ryan333@vt.edu

Mary E. Lofton
Department of Biological Sciences, Virginia Tech
melofton@vt.edu

Nicholas W. Hammond
Department of Geosciences, Virginia Tech
hammondnw@vt.edu

Rachel S. Corrigan
Department of Forest Resources and Environmental Conservation, Virginia Tech
rachelc1@vt.edu

R. Quinn Thomas
Department of Forest Resources and Environmental Conservation, Virginia Tech
rqthomas@vt.edu

Cayelan C. Carey
Department of Biological Sciences, Virginia Tech
cayelan@vt.edu

File list

```
Venn diagrams.Rmd
Final import and analysis - for EDI.Rmd
R2_analysis_EDI.Rmd
R2_dataset_for_EDI.csv
complete_dataset_with_source.csv
abstract_review_all_papers_EDI.csv
```

Description

`Venn diagrams.Rmd` – Code used to analyze abstract review data and create Venn diagrams for Fig. 1.

`Final import and analysis - for EDI.Rmd` – Code used to analyze matrix analysis results and create five figures: (1) the number of near-term ecological forecasts published per year, (2) a general description of ecological forecasting papers identified in this study, (3) the relationship between time step and time horizon of forecasting papers, (4) the total number of years of data used to develop each forecasting paper, (5) best practice adoption over time.

`R2_analysis_EDI.Rmd` – Code used to process R^2 data, create a figure showing forecast performance (R2) over increasing forecast horizons for chlorophyll, phytoplankton, pollen, and evapotranspiration, and conduct a logistic regression to determine how R2 changes with forecast horizon for these variables. File also includes code for basic summary statistics related to R^2

`R2_dataset_for_EDI.csv` – Dataset of R^2 values, forecast horizons, forecast variables, model groups, site or year groups, and paper information

`complete_dataset_with_source.csv` – Dataset of matrix analysis results, including 57 fields of information for each of 178 papers

`abstract_review_all_papers_EDI.csv` – Dataset of abstract review results, including whether or not each paper met the criteria of being a forecast, being near-term, and being ecological