Received: 1 December 2023 | Accepted: 14 December 2023

DOI: 10.1111/gcb.17123

INVITED COMMENTARY

Global Change Biology WILEY

Context matters in ecological forecasting: Lessons in predicting species distributions

Elise F. Zipkin | **Jeffrey W. Doser**

Ecology, Evolution, and Behavior Program, Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA

Correspondence
Elise F. Zipkin, Department of Integrative Biology, Michigan State University, East Lansing, MI 48824 USA.
Email: ezipkin@msu.edu

Funding information
Division of Environmental Biology, Grant/Award Number: 2213565

Keywords
forecasts, global change, prediction, species distributions, uncertainty

Forecasting is an increasingly common and important practice in the field of ecology (Dietze et al., 2018). Understanding how and why an individual, population, and/or community of species is likely to respond to novel environmental conditions is critical for evaluating and mitigating ongoing biodiversity loss in a rapidly changing world (Clark et al., 2001). However, forecasting the future state of a species (i.e., status, trend, and dynamics) is a tricky process, as there are numerous hidden factors that influence species trajectories in addition to the obvious unknowns about the future state of the planet. Accurate forecasts require disentangling multiple sources of uncertainty related to biological model parameters, environmental variability, and whether a species' response to environmental conditions, such as climate and land cover, is likely to remain constant—or at least be comparable—across time and space (Zylstra & Zipkin, 2021).

Clare et al. (2024) demonstrate the complexities associated with making predictions on future species distributions using historical avian occurrence data. With a truly unique dataset, their analyses reveal how multiple sources of uncertainty work together to inhibit our ability to understand what the future holds for biodiversity. The team uses the notable Grinnell surveys in which hundreds of sites across California were surveyed between 1896 and 1940, providing unprecedented historical data on the occurrence and distributions of birds more than a century ago. The study area was then resampled using contemporary avian point count surveys between 2003 and 2016. To make forecasts using the historical data, the authors evaluated a variety of model structures, examining the effects of key sources of uncertainty individually and in connection with one another. Their results reveal that models that incorporated multiple sources of uncertainty generally performed best. Specifically, models that included latent variables with species-specific associations allowed for temporal changes in species–environment relationships, and leveraged past information on whether a species had previously occurred at a site outperformed simpler variants. This paper is an important addition to the growing literature on ecological forecasting, helping to find solutions to the vexing yet critically important question: How do we assess biodiversity risk into the future and how do we make conservation decisions when there are high degrees of uncertainty about that future? Clare et al. (2024) reveal that we lack important knowledge on the drivers of present species distributions, impeding our abilities to make predictions, with important implications for how we can and should consider ecological forecasts and the timeframes over which we generate predictions (e.g., short-term vs. long-term forecasts).

The resurveys of sites nearly a century apart allowed Clare et al. (2024) to assess forecasting accuracy when the "future" (i.e., 2003–2016) climate and environment are perfectly known. Much of the forecasting literature is focused on using contemporary data to evaluate what the future might hold next year, in 20 years, or even at the end of the century (e.g., Neupane et al., 2022). In these cases, there are high levels of uncertainty about future environmental conditions, as the extent of climate and other types of change far into

This article is a Invited Commentary on Clare et al, https://doi.org/10.1111/gcb.17019.

the future are undetermined, with wide possibilities. However, even in the unrealistic case in which future environmental conditions are perfectly known, Clare et al.'s (2024) models based on avian data from decades earlier do not perform all that well in predicting contemporary species ranges.

There are several reasons why this may be the case. First, despite numerous studies on birds and other taxonomic groups, our understanding of how and why species respond to environmental conditions is ultimately limited. Due to natural constraints of data collection, species distribution analyses generally focus on estimating the effects of environmental variables over limited spatial and temporal extents. Over and over, we find that species–environment relationships in one time period or spatial location do not hold outside that range (Rousseau & Betts, 2022). Such nonstationarity can arise because environmental variables in new spatial locations or time periods may take novel values outside the range of the existing data, or because of missing interactions with other abiotic and/or biotic variables that have themselves changed over the prediction area or time period (Rollinson et al., 2021). It is also possible that the true underlying species–environment relationship has simply changed over space and/or time. In other words, it is unlikely that there is a universal relationship between environmental variables and species distributions—the context matters. As a result, any forecast that does not incorporate nonstationarity into predictions is, at best, likely to underestimate uncertainty and at worst, wholly misrepresent the range of possible future scenarios. The most obvious solution for this is to collect data over a wider range of environmental and climate conditions, which will allow more accurate characterizations of the nuanced factors that affect species distributions. New and ongoing data collection efforts should consider where and what types of data will be most beneficial for improving parameter accuracy and precision to maximize the return on investment.

Second, the most accurate forecasts are likely to be those that use long-term time series data collected across multiple spatial locations to forecast shortly into the future. In most cases, it is unrealistic to make predictions far into the future without extensive spatiotemporal data. Because of the difficulties in correctly assessing the factors influencing species occurrences, predictions of species distributions are likely to be most accurate when they take advantage of the auto-correlated nature of species habitat use over space and time using a random effects structure (e.g., Wright et al., 2021). Clare et al. (2024) found species that used a site in the previous century were more likely to use that site in the present. Analyses with longer time series data (i.e., many points in time rather than two distinct time periods) may thus provide an improved ability to leverage autocorrelation in predictions of species distributions. Recent advances in spatiotemporal statistics (i.e., graphical Gaussian process models) have been shown to substantially reduce uncertainty in air pollution forecasts compared with current state-of-the-art spatial dynamic models (Dey et al., 2022). Such approaches could similarly be applied to reduce uncertainty in ecological forecasts.

So where does that leave us? For forecasting research to be useful to the many that are grappling with how to make ecological predictions in a rapidly changing world, specific guidance can help to determine where to focus efforts. Lots of models do poorly in terms of predicting future outcomes; does this mean it is just not wise to make predictions on species distributions one year, a decade, or a century into the future? Are forecasts likely to be so uncertain that they are meaningless? We echo the guidance of Clare et al. (2024) to use near-term and long-term forecasting in complementary ways. Near-term forecasts can be used to guide specific management and conservation actions, which can be updated as new data and evidence are collected. The advantage of short-term forecasts is that they can be assessed with the subsequent true outcome and then tweaked as goals change and the efficacy of interventions are evaluated. Long-term forecasts can be used to characterize uncertainty further into the future, which can help guide longstanding conservation planning and legislative actions that are based on such uncertainty in possible future outcomes. While many questions remain, Clare et al.'s (2024) in-depth exploration into the causes of uncertainty in population forecasts will undoubtedly improve the utility of long-term ecological forecasting initiatives by placing the focus on prediction uncertainty and its many sources rather than a single, likely inaccurate, best guess.

AUTHOR CONTRIBUTIONS

Elise F. Zipkin: Conceptualization; writing – original draft; writing – review and editing. **Jeffrey W. Doser:** Conceptualization; writing – original draft; writing – review and editing.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed for the current article.

ORCID

Elise F. Zipkin https://orcid.org/0000-0003-4155-6139
Jeffrey W. Doser https://orcid.org/0000-0002-8950-9895

REFERENCES

Clare, J. D. J., de Valpine, P., Moanga, D. A., Tingley, M. W., & Beissinger, S. R. (2024). A cloudy forecast for species distribution models: Predictive uncertainties abound for California birds after a century of climate and land-use change. *Global Change Biology*, *30*, e17019.

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

Dey, D., Datta, A., & Banerjee, S. (2022). Graphical Gaussian process models for highly multivariate spatial data. *Biometrika*, *109*, 993–1014.

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

Neupane, N., Zipkin, E. F., Saunders, S. P., & Ries, L. (2022). Grappling with uncertainty in ecological projections: A case study using the migratory monarch butterfly. *Ecosphere*, *13*, e03874.

Rollinson, C. R., Finley, A. O., Alexander, M. R., Banerjee, S., Dixon Hamil, K. A., Koenig, L. E., Lock, D. H., DeMarche, M. L., Tingley, M. W., Wheeler, K., Youngflesh, C., & Zipkin, E. F. (2021). Working across space and time: Nonstationarity in ecological research and application. *Frontiers in Ecology and the Environment*, *19*, 66–72.

Rousseau, J. S., & Betts, M. G. (2022). Factors influencing transferability in species distribution models. *Ecography*, *2022*, e06060.

Wright, W. J., Irvine, K. M., Rodhouse, T. J., & Litt, A. R. (2021). Spatial Gaussian processes improve multi-species occupancy models when range boundaries are uncertain and nonoverlapping. *Ecology and Evolution*, *11*, 8516–8527.

Zylstra, E. R., & Zipkin, E. F. (2021). Accounting for sources of uncertainty when forecasting population responses to climate change. *Journal of Animal Ecology*, *90*, 558–561.

How to cite this article: Zipkin, E. F., & Doser, J. W. (2024). Context matters in ecological forecasting: Lessons in predicting species distributions. *Global Change Biology*, *30*, e17123. https://doi.org/10.1111/gcb.17123