12 October 2020

Dear Editors,

We are pleased to submit the manuscript *Evaluating Probabilistic Ecological Forecasts*, which we would like you to consider for publication as a Statistical Report in *Ecology.*

Our manuscript addresses the pressing need for statistical methods in evaluating ecological forecasts by collating substantial research from a variety of forecasting disciplines and providing an example application to a pocket mouse population. Thus, we believe our manuscript will be of broad and timely interest to the readers of *Ecology*.

As we grapple with climate change and leverage long-term data sets, the field of ecology is seeing marked growth in model-based forecasting. This development of forecast interest and application is buttressed by well-established statistical methods from a variety of fields including epidemiology, meteorology, and finance, yet we need to hone these tools from other fields to address ecological time series. We highlight the necessity of a diversity of options (as with any statistical method, there is no one-size-fits-all) and collect literature from Bayesian, frequentist, information theoretic, and tactical approaches to forecast evaluation and incorporate visual diagnostics as well as quantitative comparisons.

To exemplify the practice of probabilistic forecast evaluation in ecology, we leverage a long-term time series (312 monthly data points) of a desert pocket mouse population in Portal, AZ, USA. In comparing three models (random walk, auto-regressive, and cyclical auto-regressive), we use a rolling-origin end-sample method that allows us to robustly evaluate models. In assessing the three models visually and statistically, we note that the best long-term model (cyclic auto-regressive) was not best for the final realization, but rather, the auto-regressive model, which displayed significant bias in the long-term was best in the short-term. Similarly, while our choices for scoring rules used to evaluate performance (the Rank Probability and Log Scores) do select the same best model for short- and long-term, the relative value and ranks of the other two models differs between scores. This highlights important lessons regarding the utility of incorporating multiple evaluation metrics and the potential need to use and both long- and short-term evaluation methods.

Our manuscript therefore provides a timely and needed review of existing methods from other forecasting disciplines and details how to enact the methods. Code provided with the manuscript lays out the forecasting procedure and calculations for ecologists to adapt for their own needs. As such, we believe our manuscript to be a critical contribution to the growing field of ecological forecasting.

I confirm that this manuscript has not been published elsewhere and is not under consideration by another journal. All authors have approved the manuscript and agree with its submission to *Ecology*. The research presented here was conducted in conformance with all applicable laws.

On the behalf of my co-authors, I thank you for the opportunity to submit this manuscript and look forward to hearing from you.