

---

June 27, 2025  
University of Wyoming  
Botany, Dept 3165  
1000 E. University Ave  
Laramie, Wyoming, 82071

Dear Dr. Kathy Cottingham and *Ecology* Editorial Board,

Thank you for receiving our manuscript, *Accounting for missing data in autoregressive models of ecological time series* for consideration as a Statistical Innovations article in *Ecology*. Missing time series data are a ubiquitous phenomenon in ecological studies for reasons such as faulty environmental sensors, inaccessibility of field sites, and funding challenges. Data gaps are especially problematic when fitting autoregressive time series models, since missing values violate the key statistical assumption of observations that are equally spaced in time. Our study provides a much needed comparison of multiple previously proposed approaches for handling missing data in time series.

Here, we evaluate statistical methods for handling missing data in time series by fitting models to two common types of ecological data: daily observations of a continuous variable with a Gaussian error distribution (analogous to sensor data) and annual observations of a discrete variable with a Poisson error distribution (analogous to population census data). Using both simulated and empirical datasets, we artificially introduce different amounts and types of missingness and then quantify the performance of six approaches for dealing with missing data in terms of their ability to precisely recover model parameters and accurately forecast held-out data. Our results show that when data are missing completely at random, parameters were recovered well even with as high as 50% missingness and high levels of temporal autocorrelation in missingness. Conversely, parameter estimates and forecasts were unreliable when data were missing not at random. The best performing missing data approaches were 1) the Kalman filter and Bayesian data augmentation for time series with Gaussian error and 2) Bayesian data augmentation and complete case data deletion for time series with Poisson error. Overall, we believe our study is novel and of interest to the broad audience of *Ecology* as it can be a resource for ecologists, environmental scientists, and statisticians in search of robust, reproducible methods for confronting time series models with times series containing missing values. In particular:

- Our paper provides a robust test of competing statistical approaches for handling missing data in time series, using both simulated and real data with different error distributions.
- It highlights common aspects of missing data that are especially challenging to time series, such as highly auto-correlated data gaps or data that are missing not at random.
- Our results highlight a positive message that missing values in time series do not necessarily have a catastrophic effect on model accuracy, precision, or forecast ability. In fact, parameter estimation can be fairly robust as long as the error distribution, the type of missingness, and amount of autocorrelation are taken into account.

We have no competing interests to report and the submitted work is not under consideration elsewhere. The data and code accompanying this manuscript are currently available in a Github repository, and will be archived upon acceptance. Our manuscript currently stands at 18 pages of text, 8 pages of figures and captions, and 6 pages of references (32 pages total) which slightly exceeds the recommended page limit for *Statistical innovations* articles; however, we hope you will review the manuscript as is given the following justification: (**Section 1**) As articulated above, this manuscript provides tools to address missing data in time series, a phenomenon that is common across all sub-disciplines, taxonomic groups, and temporal and spatial scales of study in ecology. We are not aware of previously published work that provides comprehensive and practical guidance to ecologists for dealing with this phenomenon, and as

---

---

such believe that this manuscript will be of broad interest. **(Section 2)** Removing material from the current version of the manuscript would require either curtailing the missing data approaches included in our comparison, limiting the type of missing data we evaluated, or only examining one type of time series data. We believe that any of these reductions would diminish both the broad appeal and the utility of this work.

Thank you for considering our manuscript for publication in *Ecology*.

Sincerely,

Alice E. Stears, Melissa H. DeSiervo (co-first authors)  
Lauren G. Shoemaker, Topher Weiss-Lehman (co-last authors)  
and all other coauthors

---