A simulation-based approach for quantifying and partitioning uncertainty to improve forecasts of dynamic processes

Andrew T. Tredennick[1]*, Peter B. Adler*[1]*, and others*[2]

[1] Department of Wildland Resources and the Ecology Center, Utah State University, Logan, UT, United States

[2] Department of somewhere

Abstract

Making informed ecosystem management decisions in the face of rapid environmental change requires forecasts from models of ecological processes. However, forecasts from ecological models are often associated with high degrees of uncertainty, making it difficult for such forecasts to inform decision-making processes. To make progress toward the goal of reliable and informative ecological forecasts, we need to know from where forecast uncertainty arises. Such knowledge can guide investment in future research that will most improve forecast skill. There is a rich history of analytical expressions that partition the variance of future dynamics, but these expressions suffer from necessary assumptions such as linear dynamics and small-variance approximations that exclude interactions. Similarly, the earth systems modeling community has developed methods for paritioning uncertainty of model projections, but these operate at a different modeling grain than most of ecology. Building on these approaches, we develop a simulation-based approach for

quantifying and partitioning forecast uncertainty from Bayesian state-space models that overcomes the limitations of previous analytical approaches. Our approach is similar to an Analysis of Variance, where the total variance of a forecast is paritioned among its constituent parts, namely initial conditions uncertainty, parameter uncertainty, driver uncertainty, process error, and their interactions. We apply the approach to near-term forecasts of the Yellowstone bison population and measles in China. We also provide functions written in the statistical programming language R, which will allow others using Bayesian state-space models to employ our approach in their own research.

Keywords: Bayesian state-space model, forecast, Markov chain Monte Carlo, measles, prediction, population model, uncertainty, Yellowstone bison

Introduction

A fundamental challenge facing society is to predict the ecological impacts of global environmental changes such as nitrogen deposition, climate change, and habitat fragmentation. Each of these global change drivers have now exceeded their historical ranges of variability (Steffen et al. 2015), ushering in a no-analog era in which the past cannot predict the future. We can, however, look to the past to parameterize models that allow us to forecast the future states of ecological systems (Clark et al. 2001, Dietze et al. 2018). Ecologists are in an excellent position to meet this forecasting challenge because we have spent decades gaining understanding of the processes that regulate populations, communities, and ecosystems. However, we lack a systematic understanding of the current limits to ecological forecasts. As a result, we do not know how to allocate research effort to improve our forecasts.

Making poor forecasts is inevitable as ecology matures into a more predictive science. The key is to learn from our failures so that forecasts become more accurate over time. The success of meteorological forecasting tells us that basic research on the causes of forecast uncertainty is an

essential component of this learning process (Bauer et al. 2015).

Various approaches have been used to characterize and partition forecast uncertainty (Sobol' 1993, Cariboni et al. 2007). For example, consider a dynamic model designed to predict some state y in the future (y_{t+1}) based on the current state (y_t), an environmental driver(s) (x), parameters (θ), and process error (ε). We can then write a general form of the model as:

$$y_{t+1} = f(y_t, x_t|\theta) + \varepsilon_{t+1}, \tag{1}$$

which states that y at time $t+1$ is a function of y and x at time t conditional on the model parameters (θ) plus process error (ε). Ignoring covariance among factors and assuming linear dynamics, Dietze (2017), following Sobol' (1993) and Cariboni et al. (2007), suggests that forecast variance ($Var[y_{t+1}]$) is approximately:

$$Var[y_{t+1}] \approx \underbrace{\left(\frac{\delta f}{\delta y}\right)^2}_{\text{stability}} \underbrace{Var[y_t]}_{\text{IC uncert.}} + \underbrace{\left(\frac{\delta f}{\delta x}\right)^2}_{\text{driver sens.}} \underbrace{Var[x_t]}_{\text{driver uncert.}} + \underbrace{\left(\frac{\delta f}{\delta \theta}\right)^2}_{\text{param sens.}} \underbrace{Var[\theta]}_{\text{param. uncert.}} + \underbrace{Var[\varepsilon_{t+1}]}_{\text{process error}}, \tag{2}$$

where each additive term follows a pattern of *sensitivity* times *variance* and "IC uncert." refers to "*I*nitial *C*onditions uncertainty." The variance attributable to any particular factor is a function of how sensitive the model is to the factor and the variance of that factor. For example, the atmosphere is a chaotic system, meaning its dynamics are internally unstable and sensitive to initial conditions uncertainty. This is why billions of dollars are spent each year to measure meterological variables – meterologists learned that the key to reducing forecast error ($Var[y_{t+1}]$) was to reduce the uncertainty of initial conditions ($Var[y_t]$). In contrast, ecologists are attempting to make actionable forecasts with little knowledge of which term in Eq. 2 dominates forecast error. Knowing which term dominates forecast error in different ecological settings will advance our fundamental understanding of the natural world and immediately impact practical efforts to monitor, model, and

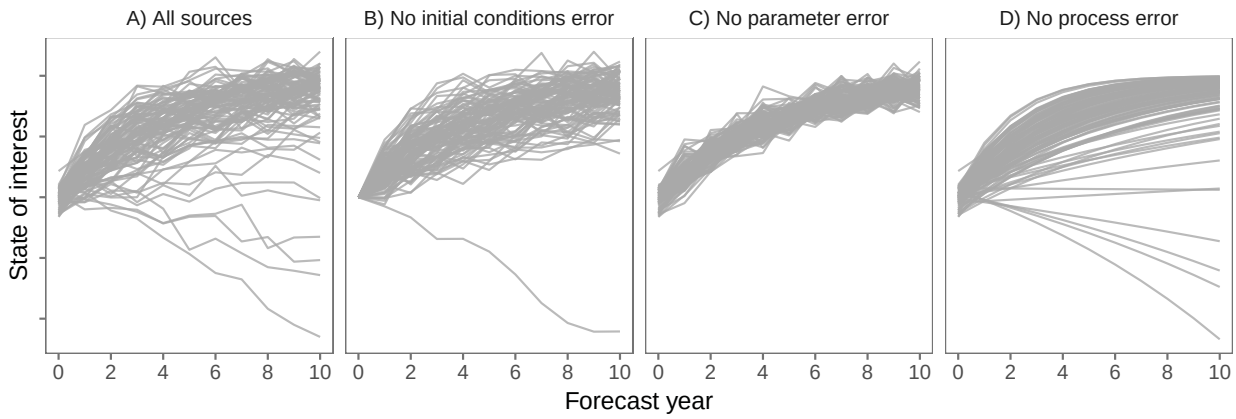

Figure 1: Example of forecast uncertainty with different sources of error set to zero. Each line represents one realization, out of 200, from an order-one autoregressive model (AR1 process). Contrary to the analytical expression (Eq. 2), initial conditions uncertainty and parameter uncertainty clearly interact. The spread of lines in (A) is not wholly because of initial conditions uncertainty (B) or parameter uncertainty (C), it is their combined influence that causes the spread of realizations in (A). At least in this example, process error (D) does appear to be independent, but we used a small value of process error to highlight other interactions. Source code: `generate_forecast_fxns.R`.

predict ecological dynamics.

While having an analytical expression such as Eq. 2 is satisfying, arriving at the expression involves strict assumptions. First, Eq. 2 only holds when the underlying dynamics are linear, which may not be the case for many populations and models. Second, Eq. 2 only includes additive effects of each factor because the Taylor series decomposition requires small-variance approximations that eliminate interactions. But, interactions among the factors are probably commone. For example, in a simple simulation of an AR(1) process, we show that initial conditions uncertainty and parameter error interact to generate the full spread of forecast variance (Figure 1). Therefore, progress in quantifying and partioning forecast uncertainty requires a more flexible approach than that provided by Eq. 2.

A Brief History of Quantifying and Partitioning Forecast Uncertainty

Error propagation

Methods for quantifying and partitioning forecast uncertainty all share common roots in statistical error propagation. Error propagation is concerned with translating the effects of variable (or parameter) uncertainty into the uncertainty of a function based those variables. That is, for the output q of some function $q = f(x_1, x_2, \ldots, x_n)$, we are interested in the variance of q given the variance of input variables ($\mathbf{x}$). It is this generic model formulation that leads to the canonical expression of error propagation

$$\sigma_q^2 = \left(\frac{\delta q}{\delta x_1}\sigma_{x_1}\right)^2 + \left(\frac{\delta q}{\delta x_2}\sigma_{x_2}\right)^2 + \cdots + \left(\frac{\delta q}{\delta x_n}\sigma_{x_n}\right)^2 \tag{3}$$

$$= \sum_{i=1}^{n}\left(\frac{\delta q}{\delta x_i}\sigma_{x_i}\right)^2. \tag{4}$$

This expression states that the variance of the output of function q (σ_q^2) is equal to the sum of squares of the input variables weighted by the sensitivity of the output variable to each input variable, quantified as the partial derivative.

Weather forecasting

Chaos.

Earth system models

Carbon.

Dynamical models

Lotka-Volterra.

A Simulation-Based Approach for Partitioning Forecast Uncertainty

Analytical expressions of forecast uncertainty must rely on simplifying assumptions. Two important assumptions are (1) that different sources of uncertainty do not interact and (2) that the system of equations is linear. These analytical expressions are important for guiding our intuition, but these strict assumptions limit our ability to partition forecast uncertainty in practice. Thus, we present a simulation approach that is entirely model-based and requires no assumptions, other than those embedded in the model itself. We are building on the ideas put forth by Dietze (2017), who suggested a simulation approach for quantifying the terms in Eq. 2. Here we test the general idea using simulated data and extend the approach to consider interactions among sources of uncertainty.

As a starting point, consider a generic Bayesian state-space model

$$\begin{aligned}
\textbf{Data Model:} \quad & y_t \sim \left[y_t \mid z_t, \sigma_o^2\right], && t = 1, \ldots, T, \\
\textbf{Process Models:} \quad & z_t \sim \left[z_t \mid \mu_t, \sigma_p^2\right], \\
& \mu_t = g\left(z_{t-1}, \mathbf{x}_t', \theta\right), && t = 2, \ldots, T, \\
\textbf{Parameter Models:} \quad & \phi \sim \left[\theta, \sigma_p^2, \sigma_o^2, z_{t=1}\right], && (5)
\end{aligned}$$

where y_t is the observed state at time t, z_t is the latent state at time t, μ_t is the determinstic prediction of z at time t from the process model g, which is a function of z at time *t-1*, a vector of covariates ($\mathbf{x}$) at time t, and a set of unknown parameters, θ. σ_o^2 is observation error and σ_p^2 is process error. The notation $[a \mid b, c]$ reads, "the probability of a given b and c" (Gelfand and Smith

1990), and $\boldsymbol{\phi}$ refers to the prior probability distributions for all unknown parameters and the initial conditions for the latent state, $z_{t=1}$.

For our purposes, we are interested in the probability distributions of the true state $\mathbf{z}$ at future points in time, conditional on previous observations ($\mathbf{y}$). This is referred to as the forecast distribution or the predictive process distribution (Hobbs and Hooten 2015 pp. 199–200), which, for one time step ahead of the final observation ($T+1$), is defined as

$$\begin{aligned}[z_{T+1}|y_1,\ldots,y_T] = \int\int\cdots\int \left[z_{T+1}|z_T,\theta,\sigma_{\mathrm{p}}^2\right] \\ \times \left[z_1,\ldots,z_T,\theta,\sigma_{\mathrm{p}}^2|y_1,\ldots,y_T\right] d\theta d\sigma_{\mathrm{p}}^2 dz_1\ldots dz_T.\end{aligned} \tag{6}$$

The model in Eq. 5 can be fit using Markov chain Monte carlo (MCMC) algorithms, which makes calculating the forecast distribution a relatively simple task. To obtain $[z_{T+1}|y_1,\ldots,y_T]$, we can change the indexing of t in Eq. 3 to $t=2,\ldots,T+1$ and then sample $z_{T+1}^{(k)}$ from $\left[z_{T+1}|g(z_T^{(k)},\theta^{(k)}),\sigma_{\mathrm{p}}^{2(k)}\right]$ given the current values for $\theta^{(k)}$ and $z_T^{(k)}$ on each $k=1,\ldots,K$ iteration of the MCMC algorithm (Hobbs and Hooten 2015, Williams et al. 2018). Making forecasts further into the future than $T+1$ requires extending $T+1$ to $T+2,\ldots,T+q$ and iteratively sampling $\left[z_{T+q}|g(z_{T+q-1}^{(k)},\theta^{(k)}),\sigma_{\mathrm{p}}^{2(k)}\right]$ (Hobbs and Hooten 2015).

The forecast distribution (Eq. 6) has all of the quantities that contribute to forecast uncertainty by incorporating their uncertainty explicitly across the K MCMC iterations. For example, initial conditions uncertainty is incorporated because $z_{T+1}^{(k)}$ is a function of $z_T^{(k)}$, resulting in a total of K point forecasts for z that comprise the posterior distribution of z at each time t. If we wish to ignore initial conditions uncertainty, we can make all K point forecasts starting from the mean of z_T

$$z_T^{(*)} = E(z_T|y_1,\ldots,y_T) \approx \frac{\sum_{k=1}^{K} z_T^{(k)}}{K}, \tag{7}$$

which we call $z_T^{(*)}$ (as opposed to $z_T^{(k)}$), while retaining the uncertainty around all other parameters. Our iterative sampling to obtain the forecast distribution then becomes a conditional statement,

$$\textbf{No initial conditions uncert.:} \quad z_{T+q} \sim \left[z_{T+q}|g(z^{(k)}_{T+q-1}, \theta^{(k)}), \sigma_p^{2(k)}\right], \quad q > 1$$
$$z_{T+q} \sim \left[z_{T+q}|g(z^{(*)}_{T}, \theta^{(k)}), \sigma_p^{2(k)}\right], \quad q = 1. \tag{8}$$

Eliminating initial conditions uncertainty is the only case that requires a conditional statement, making the other scenarios (no parameter uncertainty and no process uncertainty) even easier to implement,

$$\textbf{No parameter uncert.:} \quad z_{T+q} \sim \left[z_{T+q}|g(z^{(k)}_{T+q-1}, \theta^{(*)}), \sigma_p^{2(k)}\right], \tag{9}$$
$$\textbf{No process uncert.:} \quad z_{T+q} \sim \left[z_{T+q}|g(z^{(k)}_{T+q-1}, \theta^{(k)}), 0\right], \tag{10}$$

where $\sigma_p^2 = 0$ in the case of no process uncertainty (Eq. 10).

Need to think about diff. between including process error ($E(\sigma_p^2)$) vs. process error uncertainty ($\sigma_p^{2(k)}$).

It is important to note that although our discussion has centered on obtaining forecast distributions within the MCMC algorithm used to fit the model, it is only feasible to do this for the full forecast distribution (Eq. 6). In all other cases, where states or parameters must be averaged over the K MCMC iterations, forecast simulations must be done *post hoc* using saved MCMC chains (Box 1). In other words, estimating $z_1, z_2, \ldots, z_T$ is done within the MCMC fitting algorithm, while estimating $z_{T+1}, z_{T+2}, \ldots, z_{T+q}$ is done after fitting the model, but with the full MCMC output.

add Box of pseudo-code?

With these basics in mind, we now develop our approach for partitioning and quantifying uncertainty, which is similar to an Analysis of Variance (ANOVA). Let the variance of the forecast distribution be $V(X)$, where $X = F$ (full forecast distribution), I (no initial conditions uncertainty), PA (no parameter uncertainty), or PS (no process uncertainty). $V(I)$, $V(PA)$, and $V(PS)$ are the main effects of each factor on the forecast distribution such that

$$V(F) = V(I) + V(PA) + V(PS) + V(I,PA) + V(I,PS) + V(PA,PS) + V(I,PA,PS), \tag{11}$$

where the notation $V(X,Y)$ represents the interactive effect of X and Y. The interaction between I and PA, for example, is then

$$V(I,PA) = V(F) - [V(I) + V(PA)]. \tag{12}$$

And so on...

Application: Yellowstone Bison Population

Application: Measles in China

Acknowledgements

This research was funded by National Science Foundation grant DEB-1353078 (awarded to PBA).

References

Bauer, P., A. Thorpe, and G. Brunet. 2015. The quiet revolution of numerical weather prediction. Nature 525:47–55.

Cariboni, J., D. Gatelli, R. Liska, and A. Saltelli. 2007. The role of sensitivity analysis in ecological modelling. Ecological Modelling 203:167–182.

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.

Dietze, M. C. 2017. 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.

Gelfand, A. E., and A. F. Smith. 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85:398–409.

Hobbs, N. T., and M. B. Hooten. 2015. Bayesian Models: A Statistical Primer for Ecologists. Princeton University Press, Princeton.

Sobol', I. 1993. Sensitivity Estimates for Nonlinear Mathematical Models.

Steffen, W., K. Richardson, J. Rockström, S. Cornell, I. Fetzer, E. Bennett, R. Biggs, and S. Carpenter. 2015. Planetary boundaries: Guiding human development on a changing planet. Science 348:1217.

Williams, P. J., M. B. Hooten, J. N. Womble, G. G. Esslinger, and M. R. Bower. 2018. Monitoring dynamic spatio-temporal ecological processes optimally. Ecology 99:524–535.