Quantifying the limits to forecasting the impacts of environmental change on the Yellowstone bison population

Andrew T. Tredennick[1], others...

[1]*Department of Wildland Resources and the Ecology Center, 5230 Old Main Hill, Utah State University, Logan, Utah 84322 USA*

Running Head: Limits to bison forecasting

Corresponding Author:
Andrew Tredennick
Department of Wildland Resources and the Ecology Center
Utah State University
5230 Old Main Hill
Logan, Utah 84322 USA
Phone: +1-970-443-1599
Fax: +1-435-797-3796
Email: atredenn@gmail.com

Abstract

Making ecological forecasts requires bringing together data with models that can project states under future environmental conditions. An interesting, but often overlooked, complication of ecological forecasting is that the state of interest (e.g., population abundance) is often impacted by an environmental covariate (e.g., precipitation) that itself must be forecast. In many cases the limits to ecological forecasting lie in the environmental covariates we choose, because even if those covariates produce better understanding of the system, uncertainty in their forecast will propogate through to impact the uncertainty of the focal state. Thus, understanding when and by how much forecasts are limited by the uncertainty of environmental covariates is an urgent need. Here we use new methods to explicitly quantify the different sources of uncertainty on forecasts of the Yellowstone bison population. We fit a Bayesian state-space model to 41 years of population counts assuming Gompertz population growth with an environmental effect of annual accumulated snow water equivalent. We made forecasts for an additional seven years and found (1) that including known values of mean annual snow water equivalent increased out-of-sample forecast accuracy and (2) that xx - xx% of forecast error was associated with uncertainty in future projections of snow water equivalent. In comparison, only x% of forecast error was associated with uncertainty in initial conditions (e.g., population state at $t - 1$), and fell to x% at a forecast horizon of seven years. We used simulations to show how forecast error decreased with increased precision of snow water equivalent forecasts. Our findings highlight the necessity of partitioning forecast error into quantifiable components. Only then can key limitations be identified. In our case, more precise observations of bison population abundance will not lead to better forecasts. To better forecast the impacts of environmental change on the Yellowstone bison population using our particular model, we need better forecasts of the environmental change itself.

Key words: Bayesian state-space model; Bison bison*; ecological forecasting; popu-*

lation model; Yellowstone National Park, USA; uncertainty.

Introduction

Forecasting the impacts of environmental change on populations, communities, and ecosystems is a major challenge of the 21st century.

Materials and Methods

Data

Aerial counts of the Yellowstone bison population from 1970-2017 (Fig. 1A) were used to estimate model parameters and states. Counts typically occurred four times a year. We used summer counts only because the summer aggregation of bison tends to produce more accurate counts (Hobbs et al. 2015). Replicate counts were taken in 41 of the 48 years, and from these replicates we estimated the mean annual total population in the park and the annual standard deviation of counts (i.e., observation error).

We used annual accumulated snow water equivalent (SWE, mm) as an environmental covariate in our population model. SWE data came from the West Yellowstone SNOTEL station (Site #924). We downloaded the SNOTEL data using the `snotelr` package in R and then summed SWE across days for each year.

State-space model

We used a Bayesian state-space model to obtain posterior distributions of all unknown parameters, latent states, and forecasts of future states. The fully specified model, including

a model of the ecological process, a model linking the process to the data, and parameter models, takes the form:

$$\left[\boldsymbol{\theta}_{\mathrm{p}}, \kappa, z_{(t)} | y_{(t)}, x_{(t)}\right] \propto \prod_{t=2}^{48} \underbrace{\left[z_{(t)} | \boldsymbol{\theta}_{\mathrm{p}}, z_{(t-1)}, x_{(t)}\right]}_{\text{process}} \prod_{t=1}^{41} \underbrace{\left[y_{(t)} | \kappa, z_{(t)}\right]}_{\text{data}} \underbrace{\left[\boldsymbol{\theta}_{\mathrm{p}}, \kappa, z_{(t=1)}\right]}_{\text{parameters}}, \tag{1}$$

where the notation $[a|b, c]$ reads as the probability or probability density of a conditional on b and c (Hobbs and Hooten 2015). $\boldsymbol{\theta}_{\mathrm{p}}$ is a vector of parameters in the process model, $z_{(t)}$ is the latent, or unobservable and true, state of the bison population abundance at time t, $y_{(t)}$ is the observed state of bison population abundance at time t, $x_{(t)}$ is the standardized value of mean annual SWE at time t, and κ is a variance term associated with the model likelihood (see below). Note that the product associated with the "process model" applies over seven more years than the product associated with the "data model." The extra seven years are forecasts of the latent state seven years into the future, for which no likelihood can be calculated.

Our process model represents the population dynamics of the Yellowstone bison using the stochastic model

$$\log(z_{(t)}) \sim \text{Normal}\left(\log(z_{(t-1)}) + r + b_0 \log(z_{(t-1)}) + b_1 x_{(t)}, \sigma^2_{\mathrm{p}}\right), \tag{2}$$

where the determinstic model $\log(z_{(t-1)}) + r + b_0 \log(z_{(t-1)}) + b_1 x_{(t)}$ is a Gompertz model of population growth that predicts the mean of $z_{(t)}$ on the log scale as function of the true state of the log population at time $t-1$ ($z_{(t-1)}$), the intrinsic, per capita rate of increase (r), the strength of density-dependence (b_0), and the effect (b_1) of mean annual snow water equivalent at time t ($x_{(t)}$). The quantity σ^2_{p} is the process variance on the log scale, which accounts for all the drivers of the true state that are not found in the determinstic model.

We used a negative binomial likelihood to link the observations ($\mathbf{y}$) to the estimated latent

states ($\mathbf{z}$). The likelihood is $y_{(t)} \sim \text{NB}\left(z_{(t)}, \kappa\right)$. As mentioned earlier, replicate observations were not available for seven of the 48 years. In those seven years we could not estimate observation error. Therefore, in any year for which observation error could not be estimated, we set observation variance to the maximum value observed across the time series. This is a conservative appproach that ensures our estimates of the latent states in those years are not overly confident.

Our state-space model requires prior distributions for all parameters and for the initial condition of z. We used an informative prior for the intrinsic growth rate, r, based on the population growth (λ) reported by Hobbs et al. (2015). The Hobbs et al. (2015) model includes estimates of two population growth rates, one for populations infected with brucellosis and one for populations not infected with brucellosis. Although the population growth rate was depressed in brucellosis-infected populations, the posterior estimate of the difference between the two growth rates overlapped zero, indicating little support for differentiating among the two. We therefore chose to use the estimates for the brucellosis-free population: mean = 1.11 and s.d. = 0.0241. We converted λ, the population growth rate, to r, the per capita rate of increase, by log-transforming λ: $r = \log(\lambda) = 0.1$. To get the standard deviation on the same scale, we simulated 100,000 numbers from the distribution $\log(\text{Normal}(1.11, 0.0241))$ and calculated the standard deviation of those numbers, which equaled 0.02. Thus, our prior distribution for r was $\text{Normal}(0.1, 0.02)$. We defined the prior distribution of the initial condition $z_{(t=1)}$ as $z_{(t=1)} \sim \text{Normal}\left(y_{(t=1)}, \sigma^2_{\text{o},(t=1)}\right)$.

We chose all other prior distributions to be vague. However, no prior distribution is completely uninformative, so we made sure that our choice of priors did not have large impacts on posteriors by trying several choices of priors and their associated parameters. We then observed their effects on the posteriors (Hobbs and Hooten 2015), which were small. Our final chosen priors were: $1/\sigma^2_{\text{p}} = \tau_{\text{p}} \sim \text{Gamma}(0.01, 0.01)$ and $b_{\in(0,1)} \sim \text{Normal}(0, 1000)$.

We estimated the posterior distributions of parameters and states using Markov chain

Monte Carlo (MCMC) methods implemented using JAGS 4.2.0 (Plummer) and the `rjags` package (Plummer2) to connect JAGS to R (R citation). Posterior samples were collected from three chains, each with unique initial values that were variable relative to the posterior distributions. We used 50,000 iterations to adapt the chains for optimal sampling, then accumulated 100,000 samples from each chain after an initial burn-in period of 50,000 itareations. We checked for convergence and stationarity of the MCMC chains by visually inspecting the trace plots for each parameter and state, and by calculating multivariate scale reduction factors and the Gelman metric (citations) using the `coda` package in R. R code and data necessary to reproduce our analysis has been archived on Figshare (*link*) and released on GitHub (https://github.com/atredennick/bison_forecast/releases).

Model forecasts

We forecasted total bison abundance for the years 2011-2017 under two scenarios. First, we used known values of accumulated annual SWE from the West Yellowstone SNOTEL station. Sevond, we assumed no knowledge of the SWE in 2011-2017 by setting $x_{t \in (2011,\ldots,2017)} = 0$, which means we assume average SWE in all future years. These two scenarios allowed us to compare forecast error with and without an environmental covariate.

Partitioning forecast uncertainty

To understand the "forecasting problem" in ecology, Dietze (2017) proposed a first-principles approach to partitioning forecast uncertainty. Consider a dynamical model designed to predict some state y in the future (y_{t+1}) based on the current state (y_{t+1}), an external covariate (x), parameters (θ), and process error (ϵ). We can then write a general form of the model as:

$$y_{t+1} = f(y_t, x_t|\theta) + \epsilon, \tag{3}$$

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). Using a Taylor Expansion, Dietze shows that forecast variance ($Var[y_{t+1}]$) is:

$$Var[y_{t+1}] = \underbrace{\left(\frac{\delta f}{\delta y}\right)^2}_{\text{stability}} \underbrace{Var[y_t]}_{\text{IC uncert.}} + \underbrace{\left(\frac{\delta f}{\delta x}\right)^2}_{\text{driver sens.}} \underbrace{Var[x_t]}_{\text{driver uncert.}} + \underbrace{\left(\frac{\delta f}{\delta \theta}\right)^2}_{\text{param sens.}} \underbrace{Var[\theta]}_{\text{param. uncert.}} + \underbrace{Var[\epsilon]}_{\text{process error}}, \tag{4}$$

where each additive term follows a pattern of *sensitivity* times *variance*. Thus, 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, large sensitivity to the covariate x can be compensated for if the uncertainty of the covariate is low.

We used the posterior distributions of model parameters to conduct numerical simulations that sequentially introduced the sources of uncertainty shown in equation 4 (parameter, driver, initial condition, process) (Dietze et al. 2017). For example, to isolate the impact of initial conditions uncertainty on forecast uncertainty, we used the posterior median of all model parameters and made $k = 1000$ seven-year forecasts with 1000 unique initial conditions ($z^k_{(t=2010)}$) sampled from the posterior distribution of the latent population state in the year 2010, $z_{t=2010}$. We then conducted similar simulations with parameter uncertainty, then driver uncertainty, and then process uncertainty. Uncertainty was incorporated by sampling from the posterior distributions, except for process error. Uncertainty was included for process error by simulating the dynamic process (eq. 2) using the posterior median estimate for process variance (σ^2_{p}).

We partitioned forecast uncertainty by calculating the variance of the 1000 forecasts in

each year for each simulation described above. We then divided the variance of each simulation in each year by the maximum forecast variance (simulation with uncertainty from all sources). This resulted in estimates of the proportion of total forecast variance in each year that can be attributed to each source of uncertainty described in equation 4.

Results

The posterior distributions of model parameters show that density-dependence is evident but weak (b_0 in Fig. 2), accumulated snow water equivalent has a weak negative effect on bison population growth (b_1 in Fig. 2), and that the prior distribution of per capita growth rate entirely informed the posterior distribution (r in Fig. 2). The complete overlap of the prior and posterior distributions for r is not unexpected because we used a strong prior informed by a previous study using this same data (though with a different model) (Hobbs et al. 2015).

Model forecasts are very uncertain and grew over time (Fig. 3A). Indeed, by 2017 forecasts ranged from ~2,000 bison to over 10,000 bison. Forecast uncertainty mostly comes from driver error, which accounts for about 70% of total forecast variance (Fig. 3B). Initial conditions error is consistently low across all years, as is parameter error. Process error contributes about 15-25% across all years.

Discussion

Something profound...

Acknowledgements

Tables

Figures

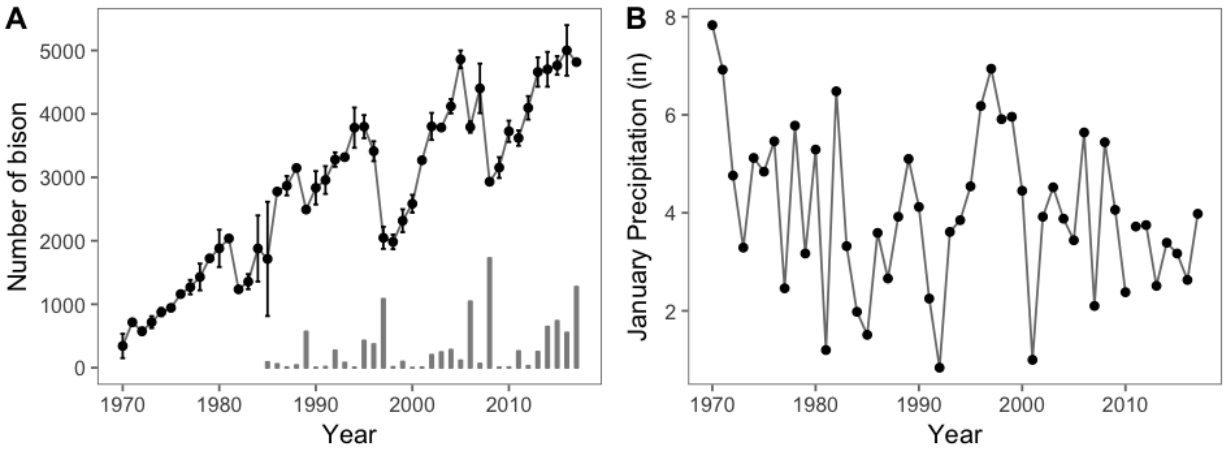

Figure 1: Time series of (A) annual summer counts of Yellowstone bison (mean ± 1 S.D.) and (B) accumulated annual snow water equivalent from the West Yellowston SNOTEL station. These data were used to estimate parameters and states in our model. Blue points are training data, orange points are validation data for which we made forecasts.

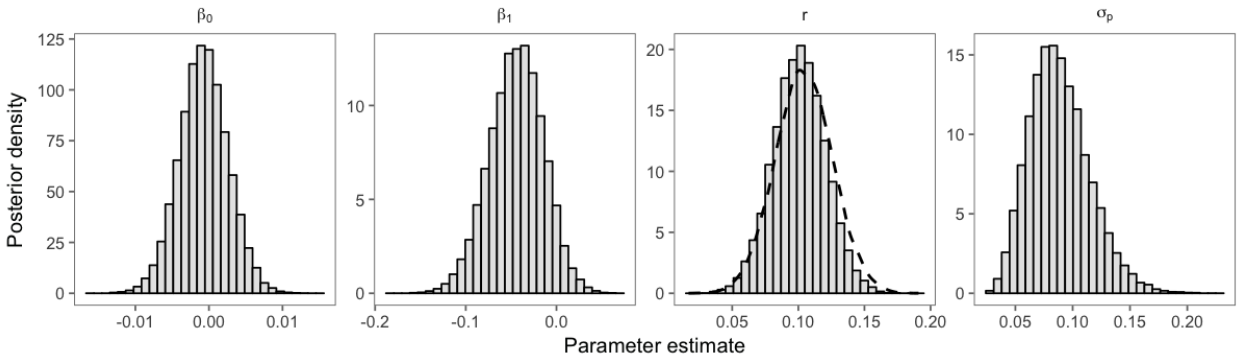

Figure 2: Posterior distributions of select model parameters: b = density-dependence, b_1 = effect of SWE, r = per capita growth rate, σ_{p} = process error. The orange line in the panel for r is the informed prior distribution. Prior distributions are not shown for other parameters because they are not visible on the scale of the posteriors.

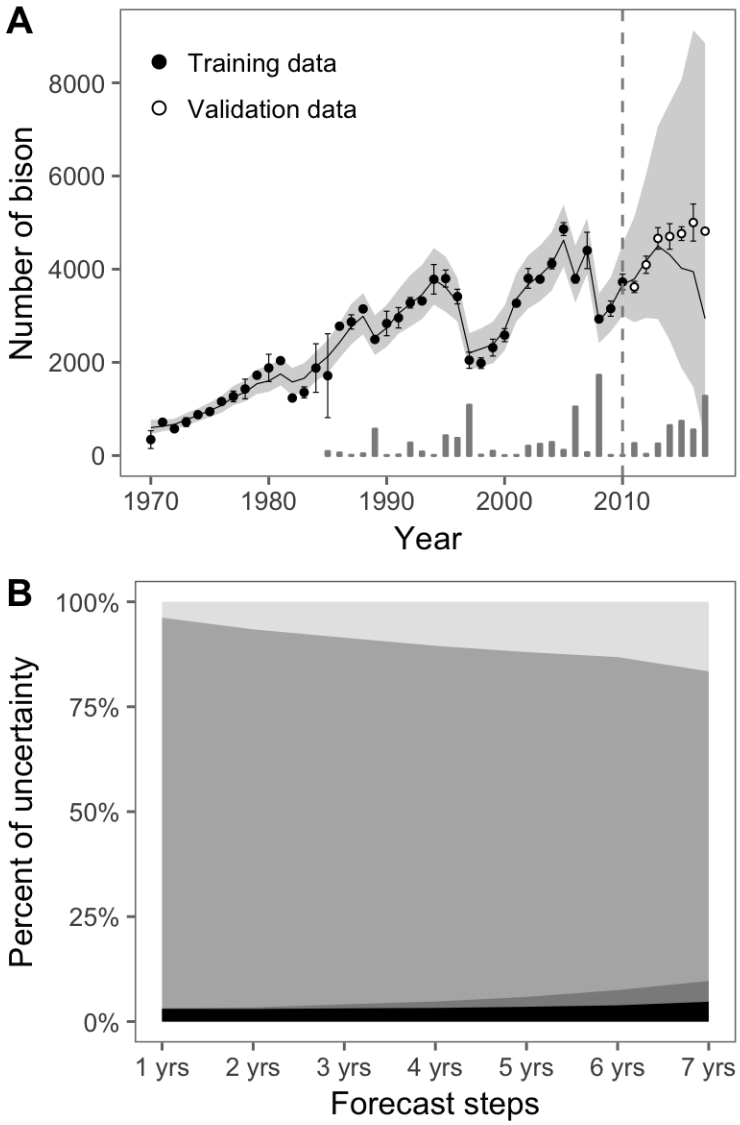

Figure 3: (A) Posterior predictions (before dashed vertical line) and forecasts (after dashed vertical line) of the Yellowstone bison population. Solid line is the median of the posterior predictive distribution and the shaded area is the 95% Bayesian credible interval. Forecasts shown here were made using known values of accumulated snow water equivalent in each year. (B) Partitioned forecast variance using a numerical approach.