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

Andrew T. Tredennick[1]

[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 38 years of population counts assuming Gompertz population growth with an environmental effect of annual accumulated snow water equivalent. We made forecasts for an additional 7 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 0% at a forecast horizon of three 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; population model; Yellowstone National Park, USA; uncertainty.*

Introduction

Forecasting the impacts of environmental change on populations, communities, and ecosystems is a major challenge of the 21$^{\text{st}}$ century.

2

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:

$$\left[\boldsymbol{\theta}_{\mathrm{P}}, \kappa, z_{(t)}, z_{(t-1)} \big| y_{(t)}, x_{(t)}\right] \propto \prod_{t=2}^{58} \underbrace{\left[z_{(t)} \big| \boldsymbol{\theta}_{\mathrm{P}}, z_{(t-1)}, x_{(t)}\right]}_{\text{process}} \prod_{t=1}^{48} \underbrace{\left[y_{(t)} \big| \kappa, z_{(t)}\right]}_{\text{data}} \underbrace{\left[\boldsymbol{\theta}_{\mathrm{P}}, \kappa, z_{(t=1)}\right]}_{\text{parameters}}, \tag{1}$$

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, and $x_{(t)}$ is the standardized value of mean annual SWE at time t. Note that the product associated with the "process model" applies

3

₅₂ over 10 more years than the product associated with the "data model." The extra ten years are

₅₃ forecasts of the latent state ten years into the future, for which no likelihood can be calculated.

₅₄ Our process model represents the population dynamics of the Yellowstone bison using the

₅₅ stochastic model

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

₅₆ where the determinstic model $\log(z_{(t-1)}) + r + b_0\log(z_{(t-1)}) + b_1 x_{(t)}$ is a Gompertz model of

₅₇ population growth (Ricker 19xx) that predicts the mean of $z_{(t)}$ on the log scale as function of the

₅₈ true state of the log population at time $t-1$ ($z_{(t-1)}$), the intrinsic, per capita rate of increase (r),

₅₉ the strength of density-dependence (b_0), and the effect (b_1) of mean annual snow water equiva-

₆₀ lent at time t ($x_{(t)}$). The quantity σ_{P}^2 is the process variance on the log scale, which accounts for

₆₁ all the drivers of the true state that are not found in the determinstic model.

₆₂ We used a negative binomial likelihood to link the observations (**y**) to the estimated latent states

₆₃ (**z**). The likelihood is $y_{(t)} \sim \text{NB}\left(z_{(t)}, \kappa\right)$. As mentioned earlier, replicate observations were

₆₄ not available for seven of the 48 years. In those seven years we could not estimate observation

₆₅ error. Therefore, in any year for which observation error could not be estimated, we set obser-

₆₆ vation variance to the maximum value observed across the time series. This is a conservative

₆₇ appproach that ensures our estimates of the latent states in those years are not overly confident.

₆₈ Our state-space model requires prior distributions for all parameters and for the initial condi-

₆₉ tion of z. We used an informative prior for the intrinsic growth rate, r, based on the population

₇₀ growth (λ) reported by Hobbs et al. (2015). The Hobbs et al. (2015) model includes estimates

₇₁ of two population growth rates, one for populations infected with brucellosis and one for pop-

₇₂ ulations not infected with brucellosis. Although the population growth rate was depressed

₇₃ in brucellosis-infected populations, the posterior estimate of the difference between the two

₇₄ growth rates overlapped zero, indicating little support for differentiating among the two. We

₇₅ therefore chose to use the estimates for the brucellosis-free population: mean = 1.11 and s.d.

₇₆ = 0.0241. We converted λ, the population growth rate, to r, the per capita rate of increase, by

⁷⁷ log-transforming λ: $r = \log(\lambda) = 0.1$. To get the standard deviation on the same scale, we

⁷⁸ simulated 100,000 numbers from the distribution $\log(\text{Normal}(1.11, 0.0241))$ and calculated

⁷⁹ the standard deviation of those numbers, which equaled 0.02. Thus, our prior distribution

⁸⁰ for r was $\text{Normal}(0.1, 0.02)$. We defined the prior distribution of the initial condition $z_{(t=1)}$ as

⁸¹ $z_{(t=1)} \sim \text{Normal}\left(y_{(t=1)}, \sigma^2_{o,(t=1)}\right).$

⁸² We chose all other prior distributions to be vague. However, no prior distribution is completely

⁸³ uninformative, so we made sure that our choice of priors did not have large impacts on posteri-

⁸⁴ ors by trying several choices of priors and their associated parameters. We then observed their

⁸⁵ effects on the posteriors (Hobbs and Hooten 2015), which were small. Our final chosen priors

⁸⁶ were: $1/\sigma^2_p = \tau_p \sim \text{Gamma}(0.01, 0.01)$ and $b_{\in(0,1)} \sim \text{Normal}(0, 1000)$.

⁸⁷ We estimated the posterior distributions of parameters and states using Markov chain Monte

⁸⁸ Carlo (MCMC) methods implemented using JAGS 4.2.0 (Plummer) and the `rjags` package

⁸⁹ (Plummer2) to connect JAGS to R (R citation). Posterior samples were collected from three

⁹⁰ chains, each with unique initial values that were variable relative to the posterior distributions.

⁹¹ We used 50,000 iterations to adapt the chains for optimal sampling, then accumulated 100,000

⁹² samples from each chain after an initial burn-in period of 50,000 itareations. We checked for

⁹³ convergence and stationarity of the MCMC chains by visually inspecting the trace plots for each

⁹⁴ parameter and state, and by calculating multivariate scale reduction factors and the Gelman

⁹⁵ metric (citations) using the `coda` package in R. R code and data necessary to reproduce our

⁹⁶ analysis has been archived on Figshare (*link*) and released on GitHub (https://github.com/

⁹⁷ atredennick/bison_forecast/releases).

⁹⁸ **Model forecasts**

⁹⁹ **Partitioning forecast uncertainty**

¹⁰⁰ To understand the "forecasting problem" in ecology, Dietze (2017) proposed a first-principles

¹⁰¹ approach to partitioning forecast uncertainty. Consider a dynamical model designed to pre-

¹⁰² dict some state y in the future (y_{t+1}) based on the current state (y_{t+1}), an external covariate (x),

¹⁰³ parameters (θ), and process error (ϵ). We can then write a general form of the model as:

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

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:

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

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.

Results

Discussion

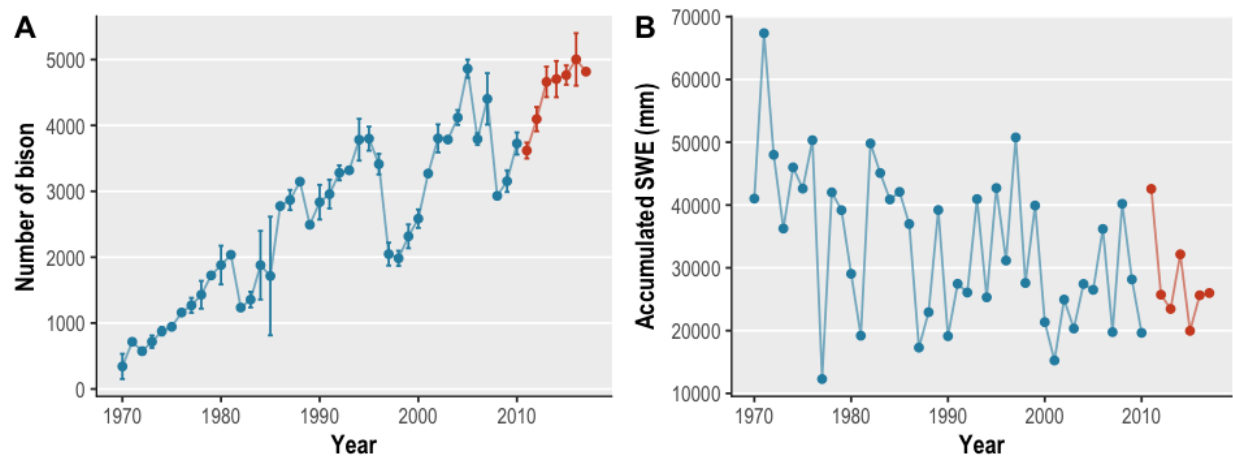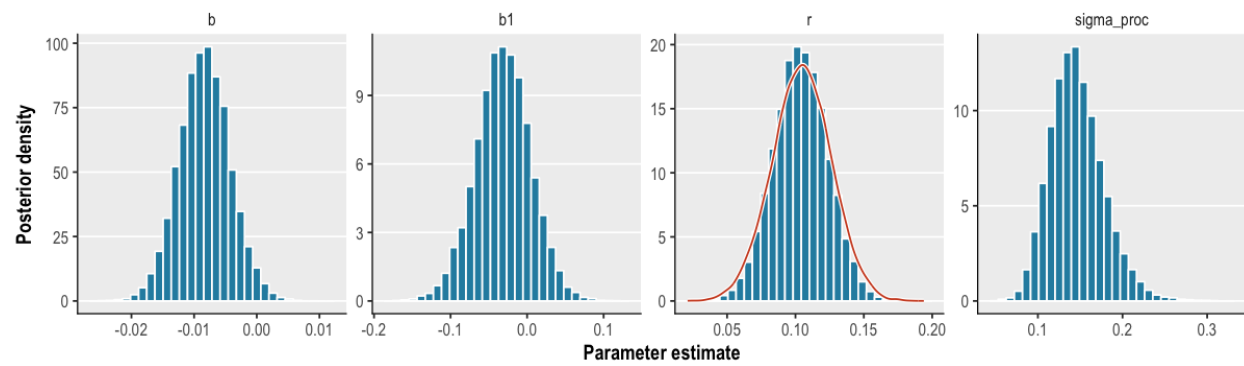# Acknowledgements

Tables

Figures

Figure 1: The data.

Figure 2: The posteriors.

Figure 3: The forecasts.