Appendix B. Prior specification and parameter estimation.

B-1. Prior specification

In order for prior distributions to play a minimal role in posterior inference, we want prior distributions for lnα, β, ϕ, τ^2, σ^2 to be as uninformative as possible. In an extreme case where $\tau^2 \to 0$, fitting the GP prior is equivalent to fitting the Ricker model. Therefore, we used uninformative reference priors developed for the Ricker model (Millar 2002); for the location parameters $\Pr(\ln\alpha, \beta) \propto 1$ and for the scale parameter $\Pr(\sigma^2) \propto \sigma^{-2}$.

However, priors for ϕ and τ^2 had to be specified informatively because our preliminary results showed that the posterior distributions did not converge when uninformative priors were used. Since ϕ controls the number of inflection points in realizations of $f(A)$ (section 1-2), the prior for ϕ was determined in light of available biologically plausible models for density dependence, including Allee effects models. These models are usually simple smooth functions with at most 1 inflection point over A. We allowed $f(A)$ to have, on average, 2 inflection points over the range, 0 to $\max(A)$, to capture somewhat more complex shapes. We determined the relationship between ϕ and the expected number of inflection points in $f(A)$ by simulation. We generated $f(A)$ in $A/\max(A) \in [0,1]$ from the GP prior in Eq 6 and counted the number of inflection points in $A\exp[f(A)]$. We found that $\phi = 8$ generates 2 inflection points on average over A. With this information, we used an informative gamma prior distribution with the mean $\mathrm{E}(\phi) = 8$, which is given by $\Pr(\phi) \propto \phi\exp(-\phi/4)$. Note that this prior can be easily overwhelmed by data which clearly show different number of inflection points over A.

For τ^2, we note that the total variance in y is $\mathrm{V}(y) = \mathrm{V}[f(A)] + \sigma^2$, and obtain a ballpark for the variance in y from its observed range (i.e., $r_y = \max(y) - \min(y)$). Assigning the prior variance in y to uncertainty in $f(A)$, we used a gamma prior, $\Pr(\tau^2) \propto \tau^2\exp\left(2\tau^2/r_y\right)$

which has $\mathrm{E}(\tau^2) = r_y$. We conducted prior sensitivity analysis using different parameterization of the gamma prior to determine possibility that the inference of the presence of Allee effects, π, is an artifact of prior specification, and confirmed that π does not change qualitatively across different priors.

B-2. Parameter estimation

With *n* years of data for data, $(\boldsymbol{A}, \boldsymbol{J}) = (\{A_1, \ldots, A_n\}, \{J_1, \ldots, J_n\})$, we have $\boldsymbol{y} = \{\ln(J_1/A_1), \ldots, \ln(J_n/A_n)\}$. With data, $f(\boldsymbol{A})$ is a sample from an *n* dimensional multivariate normal distribution:

(B.1) $$f(\boldsymbol{A})|\ln\alpha, \beta, \phi, \tau^2 \sim \mathrm{MVN}\left[\mu(\boldsymbol{A}), \Sigma_{f,f}(\boldsymbol{A}, \boldsymbol{A}')\right]$$

where $\mu(\boldsymbol{A}) = ln\alpha\mathbf{1_n} + \beta\boldsymbol{A}/\max(\boldsymbol{A})$ and $\mathbf{1_n}$ is an $n \times 1$ vector of 1's. The i, j^{th} element in the covariance is given by $\Sigma_{f,f}(\boldsymbol{A}, \boldsymbol{A}')_{i,j} = \tau^2\exp\left[-\phi|\left(A_i - A_j{}'\right)/\max(\boldsymbol{A})|^2\right]$. Obtaining posterior inference for parameters in this model is computationally cumbersome, but marginalizing over $f(\boldsymbol{A})$ can speed up the computation (Munch et al. 2005):

(B.2) $$\boldsymbol{y}|\boldsymbol{A}, \boldsymbol{\Theta} \sim \mathrm{MVN}\left[\mu(\boldsymbol{A}), \Sigma_{f,f}(\boldsymbol{A}, \boldsymbol{A}') + \sigma^2\boldsymbol{I_n}\right]$$

$$\boldsymbol{\Theta} \sim Pr(\boldsymbol{\Theta}).$$

$\boldsymbol{I_n}$ is $n \times n$ identity matrix. Parameters in the model are collected in the vector $\boldsymbol{\theta} = \{\ln\alpha, \beta, \phi, \tau^2, \sigma^{-2}\}$ and $\Pr(\boldsymbol{\Theta}) \propto 1 \cdot \phi\exp(-\phi/4) \cdot \tau^2\exp\left(2\tau^2/r_y\right) \cdot \sigma^{-2}$. In light of this, the SB model is essentially identical to an *n* dimensional multivariate normal likelihood.

We used Metropolis sampling to obtain posterior distributions for τ^2, ϕ, σ^2, at log scale, and Gibbs sampling for $\ln\alpha$ and β. The assessment for convergence in the updated posterior distributions was conducted using multiple over-dispersed chains (Gelman et al. 2003). Obtaining inference for an unknown function $f(A)$ was deferred until the posterior distributions for $\boldsymbol{\Theta}$ were updated. It follows simply from the fact that the inference for $f(A)$ is sampled from a posterior predictive GP specified with the posterior mean function and the posterior covariance function (see Appendix D) (Munch et al. 2005).

LITERATURE CITED

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2003. Bayesian Data Analysis (Second edition). Chapman and Hall/CRC press.

Millar, R. B. 2002. Reference priors for Bayesian fisheries models. *Canadian Journal of* Fisheries and Aquatic Sciences 59:1492–1502.

Munch, S. B., A. Kottas, and M. Mangel. 2005. Bayesian nonparametric analysis of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences 62:1808–1821.