Bayesian Non-parametric Analysis of Stock-Recruitment Relationships

Stephan B. Munch*, Athanasios Kottas and Marc Mangel*

Department of Applied Mathematics and Statistics and *Center for Stock Assessment Research

Jack Baskin School of Engineering

University of California

Santa Cruz, Ca 95064

Abstract

The relationship between current abundance and future recruitment to the stock is fundamental to managing fish populations. There is general agreement about the basic attributes such a relationship should possess. However, many different models may be derived from these attributes and the data are often insufficient to distinguish among them. Although nonparametric methods may be used to circumvent this problem, these are devoid of biological underpinnings. Here we present a Bayesian nonparametric approach that allows straightforward incorporation of prior biological information and show how it may be used to estimate several fishery reference points. We applied this method to several artificial data sets generated from a variety of parametric models and compare the results to the fit of Ricker and Beverton-Holt models. We found that the Bayesian nonparametric method fit the data nearly as good as the true parametric model and always performed better than incorrect parametric alternatives. The estimated reference points agree closely with true values calculated for the underlying parametric model. Finally, we apply the method to empirical data for lingcod and several salmonids. Since this method is capable of reproducing the behavior of any of the parametric models and provides flexible, data-driven estimates of stock-recruitment relationships, it should be of great value in fisheries applications where the true functional relationship is always unknown.

The relationship between stock size and subsequent recruitment is one of the keystone concepts of fishery science, since the parameters for this function translate directly into management reference points and set the ultimate limits on sustainable fishing (Quinn and Deriso 1999, Bravington et al 2000). However, it is also one of the most problematic: data are difficult to measure and generally noisy, the relationship is surely non-linear over a range of stock sizes, and a variety of plausible biological mechanisms are consistent with very different functional relationships between spawning stock and recruitment. There is, of course, general agreement about the properties a stock-recruitment model should possess. First, any extant stock must be able to replace itself, so that recruitment should be greater than stock size over some range of stock sizes. Second, we expect there to be density dependence; recruitment may be nearly proportional to stock close to the origin but per capita recruitment is expected to decrease at large stock sizes. Third, for closed populations, recruitment should tend to zero as stock goes to zero, passing through the origin.. However, the spatial scale on which the population is closed is often unclear. Many models may be derived from this set of principles and the data are frequently insufficient to distinguish among them. Consequently, selection of a parametric model is often rather arbitrary.

Over the last decade, parametric Bayesian approaches have been increasingly applied in fisheries ([Hilborn and Mangel 1997, McAllister and Kirkwood 1998, Millar 2002). One of the great advantages of the Bayesian approach is that it allows a statement of model probability to be made in cases where there is no *a priori* biological basis for model selection. However, this approach will not work if the appropriate model is not included in the set of candidates. Moreover, even if the model is appropriately specified, important management parameters such as the slope at the origin, may be unduly influenced by points far away (Ludwig 1995). Consequently, several authors have proposed non-parametric approaches. Rothschild and Mullen (1985) and Brodziak et al. (2002) divide the stock-recruitment plane into several regions and estimate transition probabilities among them from the observed time series. Non-parametric density estimators have also been used to construct the distribution of recruitment given stock biomass (e.g. Evans and Rice 1988, Cook 2000). Non-parametric regression and spline methods that fit some locally weighted smoothing function to the stock-recruit data have also been used (e.g. Chapman 1973, Cook 1998).

The chief benefit of these non-parametric approaches is that they allow the data to speak for themselves. This is a highly desirable property, once we recognize that the available biological information is typically insufficient to specify a functional form *a priori*. However, there are several drawbacks to existing methods. First, they all require the ad hoc specification of a smoothing parameter. Although cross-validation methods may be used to circumvent this problem to some extent, they do not perform well on the relatively small data sets available in fisheries. Second, uncertainty bounds for estimates from these methods rely heavily on asymptotics; given the relatively small samples available in fisheries, these uncertainty bounds will be unreliable. Third, these methods lack biological underpinnings and this makes results hard to interpret and sometimes biologically unreasonable. Bravington et al. (2000) have made some progress in this regard by developing a smoother that is forced to pass through the origin and producing diminishing gains in recruits as spawning stock size increases. Although elegantly incorporated, these biological constraints are by no means certainties and the inclusion of hard constraints is at odds with the nonparametric philosophy of letting the data speak for themselves.

We still lack a method that allows prior information about the biology of the stock to be incorporated while simultaneously allowing the data to determine the overall shape of the fitted relationship. Here we present a method, based on Bayesian nonparametrics (see, e.g., Walker et al. 1999 and Müller and Quintan 2004), that solves the difficulty. The approach provides a non-parametric fit to the data with several advantages over previous methods. First, the distributions of the parameters that determine the smoothness of the fitted nonparametric function are driven primarily by the data, given limited amounts of prior information. Second, full specification of model uncertainty is possible for the regression function and reference points derived from it without recourse to asymptotic approximations. Third, and most importantly, biological information is consistently incorporated through specification of the prior, but does not rigidly constrain the shape of the model within the range of the data. Since this is the first application of Bayesian nonparametric methods in fisheries science, we develop the statistical background for our Bayesian nonparametric (BNP) model in some detail. We demonstrate the flexibility of the method on several artificial data sets generated from a suite of parametric models and then apply the method to three actual data series.

Methods

Model Specification

For simplicity, we assume that spawning stock size (S) is measured without error and is proportional to total egg production. This assumption may be relaxed, though at the cost of increased complexity (see **Discussion**). Recruitment (R) is a function of stock size and random processes. We assume that stochastic effects are multiplicative and lognormally distributed. Alternative choices for the distribution of stochastic effects are readily incorporated in this framework. We also use log-transformed variables, specifically, $x = \ln S$ and $y = \ln R$. Thus our stock-recruitment (SR) model becomes

$$y = \Im(x) + \varepsilon \tag{1}$$

where ε is normally distributed with mean zero and variance σ^2_ε. In the standard context, $\Im(x)$ is the log-transformed SR model (e.g. $\Im(x) = \ln a + x - be^x$ for a Ricker model). In the Bayesian nonparametric context, however, the function $\Im(x)$ is uncertain and described by a prior probability distribution over random functions. When thought of in this manner, fitting a parametric model, say Ricker or Beverton-Holt, corresponds to assigning positive prior probability to a very narrow region of the biologically plausible space of SR functions. We seek a broader prior specification, encompassing the space of biologically possible regression functions. One very general prior specification for $\Im(x)$ is the Gaussian process (GP). Other prior specifications are possible (Denison et al. 2002) but the GP prior is quite flexible, allows ready interpretation, and lends itself to relatively simple prior specification and posterior inference (see, e.g., Neal 1999 and references therein for Bayesian GP regression).

The GP prior for $\Im(x)$ is fully described by a mean function $\mu(x)$ and covariance function $C(x, x')$. It is mainly through specification of $\mu(x)$ that we may incorporate our prior beliefs about the relationship between stock size and recruitment. In keeping with a long tradition in fisheries biology we assume that the population is closed, so that there should be no recruits in the absence of any adult stock (i.e. $R = 0$ when $S = 0$, but see below) and that recruitment is proportional to stock size near the origin. Because the actual mechanisms of density dependence are unknown for most stocks, we choose to let the data tell us. Consequently, we choose as our prior mean function

(2) $$\mu(x) = \beta_0+\beta_1 x$$

which corresponds directly with Cushing′s (1973) power function model. Other choices of the prior mean function are possible. For instance, if we had strong prior evidence that the stock in question exhibited Ricker –type density dependence, we could choose $\mu(x) = \ln(\alpha) + x - \beta e^x$.

Choosing the covariance function is very important because it determines how closely related values of $\Im(x)$ are for nearby values of x. Thus, prior specification of the covariance function allow us to incorporate our biological beliefs regarding the smoothness of the underlying relationship between stock size and recruitment. A generally useful choice for the covariance function is

(3) $$C(x,x') = \tau^2 \exp\left[-\phi|x-x'|^\alpha\right] \quad \tau > 0, \phi > 0, \alpha \in (0,2]$$

Here, τ^2 is the variance and parameters ϕ and α control the interdependence of $\Im(x)$ and $\Im(x')$ as a function of the distance between x and x′. Specifically, α affects the fine scale variability or graininess in $\Im(x)$ while ϕ (for a fixed α) controls the large scale variability or range of dependence. To build intuition, we illustrate the effects of each of these parameters on a realization of $\Im(x)$ (Figure 1). Note a key assumption here: the covariance function is more properly C($\Im(x)$, $\Im(x')$) but we assume that the interdependence of values in the stock recruitment relationship arise only through the distance between log-stock sizes (i.e. we assume that the process is isotropic).

Other choices of C(x, x′) are possible, although care must be taken to ensure that C(x, x′) is a proper covariance function (i.e. a symmetrical, non-negative definite function). For example, a simple extension is

(4) $$C(x,x') = \gamma_0 + \gamma_1 xx' + \tau^2 \exp\left[-\phi|x-x'|^\alpha\right] \quad \gamma_0 > 0, \gamma_1 > 0$$

In this case, in the limit as τ goes to zero, the first two terms enforce a linear dependence of $\Im(x)$ on x where the variability in the slope is determined by γ_1.

To simplify notation, we collect the parameters specifying the prior for $\Im(x)$ in the vector $\boldsymbol{\theta} = \{\beta_0, \beta_1, \tau^2, \phi, \alpha\}$. Given the specifications in Eqs 1-3, the full Bayesian model becomes

$$
\begin{aligned}
(5) \qquad & y_i \mid \Im(x_i), \sigma^2_\varepsilon \sim N(\Im(x_i), \sigma^2_\varepsilon),\ i = 1,\ldots, n \\
& \Im(x) \mid \boldsymbol{\theta} \sim GP(\mu(x), C(x,x')) \\
& \boldsymbol{\theta},\ \sigma^2_\varepsilon \sim p(\boldsymbol{\theta})p(\sigma^2_\varepsilon)
\end{aligned}
$$

where $p(\boldsymbol{\theta})$ and $p(\sigma^2_\varepsilon)$ are independent priors for $\boldsymbol{\theta}$ and σ^2_ε. We discuss prior specification for these parameters below.

It is helpful to interpret the general model in the following way: given a specific set of observed recruitment and stock sizes, $(\mathbf{x}, \mathbf{y}) = (\{x_1,\ldots, x_n\}, \{y_1,\ldots, y_n\})$, the GP prior for $\Im(x)$ means that $\{\Im(x_1), \ldots, \Im(x_n)\}$ is n-variate normal. That is,

$$
(6) \qquad \{\Im(x_1), \ldots, \Im(x_n)\} \mid \boldsymbol{\theta} \sim N_n(\mu^n(\boldsymbol{\theta}), C^n(\boldsymbol{\theta}))
$$

where the mean vector $\mu^n(\boldsymbol{\theta}) = \{\beta_0 + \beta_1 x_1, \ldots, \beta_0 + \beta_1 x_n\}'$ and the (i, j)th entry of the covariance matrix $C^n(\boldsymbol{\theta})_{i,j} = \tau^2 \exp(-\phi|x_i - x_j|^\alpha)$.

At this point, the model is fully specified and standard Monte Carlo methods (Gamerman 1997, Robert and Casella 1999) can be used to obtain the posterior distribution for $\Im(x)$. For example, a Metropolis algorithm (Metropolis et al. 1953) could consist of steps: 1). Draw proposed values for $\boldsymbol{\theta}$, σ^2_ε,, and conditional on these values draw points from the random function $\{\Im(x_1), \ldots, \Im(x_n)\}$, 2) Evaluate the joint posterior of $(\boldsymbol{\theta}, \sigma^2_\varepsilon, \{\Im(x_1), \ldots, \Im(x_n)\} \mid \text{data})$, 3) Accept or reject the proposed values. A more detailed description of the simplified sampling algorithm used here is given below.

Posterior Inference for θ and σ^2_ε

Although it is possible to sample from the joint posterior distribution for $\Im(x)$, θ, and σ^2_ε given **x**, **y** under the model specification, posterior sampling is greatly simplified by marginalizing over $\Im(x)$ to obtain a model with one less layer in the hierarchy. To do so, first note that, as in Eq 5, **y** is conditionally n-variate Normal. The expected value of **y** taken over $\Im(x)$ is then $\mu^n(\theta)$ defined below Eqn 6.. From Eq 1, the covariances for **y** are given by

$$\text{(7)} \qquad \mathrm{Cov}\,[\,(\Im(x_i)+\varepsilon_i),(\Im(x_j)+\varepsilon_j)\,] = \mathrm{Cov}\,[\,\Im(x_i),\Im(x_j)\,] + \mathrm{Cov}[\varepsilon_i,\varepsilon_j] = C(x_i,x_j) + \sigma^2_\varepsilon\, I_n(i,j)$$

with $I_n(i,j) = 1$ for $i = j$, 0 otherwise. Thus a two level model, in which the explicit conditioning of **y** on $\Im(x)$ has been integrated out is

$$\text{(8)} \qquad \begin{aligned} \mathbf{y} \mid \theta,\ \sigma^2_\varepsilon &\sim N_n(\mu^n(\theta),\ C^n(\theta) + \sigma^2_\varepsilon I_n) \\ \theta,\ \sigma^2_\varepsilon &\sim p(\theta)p(\sigma^2_\varepsilon) \end{aligned}$$

We use the Metropolis-Hastings algorithm (Hastings 1970) to obtain draws from the joint posterior of θ, σ^2_ε, $p(\theta, \sigma^2_\varepsilon \mid \text{data})$, where data = (**x**, **y**). Because the elements of θ and σ^2_ε have restricted support, sampling is facilitated by defining a new set of parameters (η) with support on R^6. This new parameter vector is given by $\eta = \{\beta_0, \beta_1, \log \tau^2, \log \phi, \log [\alpha/(2-\alpha)], \log \sigma^2_\varepsilon\}$ and is sampled from a 6-variate Gaussian (N_6) proposal distribution with mean given by the current value of η and covariance matrix D. During the initial sampling, D is diagonal, but substantial improvements in sampling may be had with a better choice of D. We improved sampling by iteratively updating D based on the covariance of the sampled parameters after each 1,000 draws. After five such iterations we used a 'burn-in' period of another 5,000 samples which was sufficient to ensure convergence.

Since we sample $\boldsymbol{\eta}$ from a multivariate normal, the corresponding density for proposed new parameter values { $\boldsymbol{\theta}^{\text{new}}$, $\sigma^{2\ \text{new}}_{\varepsilon}$} is given by

$$(9) \qquad q(\boldsymbol{\theta}^{new}, \sigma_{\varepsilon}^{2^{new}} \big| \boldsymbol{\theta}, \sigma_{\varepsilon}^{2}, D) = \frac{2}{\tau^{2}\phi\alpha(2-\alpha)\sigma_{\varepsilon}^{2}} N_{6}(\boldsymbol{\eta}(\boldsymbol{\theta}^{new}, \sigma_{\varepsilon}^{2^{new}}) \big| \boldsymbol{\theta}, \sigma_{\varepsilon}^{2}, D)$$

and a particular draw { $\boldsymbol{\theta}^{\text{new}}$, $\sigma^{2\ \text{new}}_{\varepsilon}$} is accepted with probability

$$(10) \qquad p = \min\left\{1, \frac{\overline{p}(\boldsymbol{\theta}^{new}, \sigma_{\varepsilon}^{2^{new}})}{\tau^{2}\phi\alpha(2-\alpha)\sigma_{\varepsilon}^{2}} \frac{\tau^{2^{new}}\phi^{new}\alpha^{new}(2-\alpha^{new})\sigma_{\varepsilon}^{2^{new}}}{\overline{p}(\boldsymbol{\theta}, \sigma_{\varepsilon}^{2})}\right\}$$

where $\overline{p}(\boldsymbol{\theta}, \sigma_{\varepsilon}^{2})$ is the unnormalized posterior from the model in Eq. 8.

Posterior prediction for $\Im(x)$

Based on the observed data (**x, y**) our goal is to predict values for $\Im(x)$ and y over a range of values of x. That is, we estimate the stock – recruitment function over a grid of k points which we will denote by $\mathbf{x}_{\text{new}}$ ={$x_{\text{new }1}$, …, $x_{\text{new }k}$}. The quantities of interest then are p($\Im(\mathbf{x}_{new})$ | data, $\mathbf{x}_{\text{new}}$) and p($\mathbf{y}_{\text{new}}$ | data, $\mathbf{x}_{\text{new}}$). Note that

$$(11) \qquad \mathrm{p}(\Im(\mathbf{x}_{\text{new}}) \,|\, \text{data}, \mathbf{x}_{\text{new}}) = \int \mathrm{p}(\Im(\mathbf{x}_{\text{new}}) \,|\, \text{data}, \boldsymbol{\theta}, \sigma_{\varepsilon}^{2}, \mathbf{x}_{\text{new}}) \mathrm{p}(\boldsymbol{\theta}, \sigma_{\varepsilon}^{2} \,|\, \text{data}) \mathrm{d}\boldsymbol{\theta} \mathrm{d}\sigma_{\varepsilon}^{2}$$

and

$$(12) \qquad \mathrm{p}(\mathbf{y}_{\text{new}} \,|\, \text{data}, \mathbf{x}_{\text{new}}) = \int \mathrm{p}(\mathbf{y}_{\text{new}} \,|\, \text{data}, \boldsymbol{\theta}, \sigma_{\varepsilon}^{2}, \mathbf{x}_{\text{new}}) \mathrm{p}(\boldsymbol{\theta}, \sigma_{\varepsilon}^{2} \,|\, \text{data}) \mathrm{d}\boldsymbol{\theta} \mathrm{d}\sigma_{\varepsilon}^{2}$$

Thus the posterior predictive distributions of interest can be obtained by Monte Carlo integration of the conditional distributions for $\Im(\mathbf{x}_{new})$ and $\mathbf{y}_{\text{new}}$. To do so, we require the distribution for $\mathbf{y}_{\text{new}}$ ={$y_{\text{new }1}$, …,

$y_{new\ k}\}$ given the data, parameters, and $\mathbf{x}_{new}$. Based on Eq 1, the analogous distribution for $\Im(x)$ follows directly.

The distribution for $\mathbf{y}_{new}$ given the data, parameters, and $\mathbf{x}_{new}$ is obtained from the joint distibution of $\mathbf{y}$ and $\mathbf{y}_{new}$,

$$(13) \qquad p(\mathbf{y}_{new} \,|\, \boldsymbol{\theta}, \sigma_{\varepsilon}^2, \text{data}, \mathbf{x}_{new}) = \frac{p(\mathbf{y}_{new}, \mathbf{y} \,|\, \boldsymbol{\theta}, \sigma_{\varepsilon}^2, \mathbf{x}, \mathbf{x}_{new})}{p(\mathbf{y} \,|\, \boldsymbol{\theta}, \sigma_{\varepsilon}^2, \mathbf{x})}$$

The term in the denominator is the density $N_n(\mu^n(\theta), C^n(\theta) + \sigma^2_{\varepsilon} I_n)$ given in Eq. 8 while the term in the numerator is an (n+k)-variate normal distribution. The mean vector for this distribution is given by $\{\mu^n(\theta)', \mu^k(\theta)'\}'$ where $\mu^k(\theta) = \{\beta_0 + \beta_1 x_{new\ 1}, \ldots, \beta_0 + \beta_1 x_{new\ k}\}'$. The covariance matrix is given by $C^{n+k}(\theta) + \sigma^2_{\varepsilon} I_{n+k}$ where $C^{n+k}(\theta)$ is

$$(14) \qquad C^{n+k}(\theta) = \left(\begin{array}{c|c} \underset{(n \text{ x } n)}{C^n(\theta)} & \underset{(n \text{ x } k)}{C^{n,\,new}(\theta)} \\ \hline \underset{(k \text{ x } n)}{C^{n,\,new}(\theta)^T} & \underset{(k \text{ x } k)}{C^{new}(\theta)} \end{array}\right)$$

where $C^{new}(\theta)_{i,j} = \tau^2 \exp(-\phi |x_{new\ i} - x_{new\ j}|^{\alpha})$ and $C^{n,\,new}(\theta)^T_{\ i,j} = \tau^2 \exp(-\phi |x_{new\ i} - x_j|^{\alpha})$.

Based on standard theory for multivariate normal variables, the conditional distribution for $\mathbf{y}_{new}$ in Eqn 13 is also multivariate normal with mean vector (M_{new}) and covariance matrix (S_{new}) given by

$$(15) \qquad M_{new} = \mu^k(\theta) + C^{n,\,new}(\theta)^T (C^n(\theta) + \sigma^2_{\varepsilon} I_n)^{-1} (\mathbf{y} - \mu^n(\theta))$$

and

$$(16) \qquad S_{new} = C^{new}(\theta) + \sigma^2_{\varepsilon} I_k - C^{n,\,new}(\theta)^T (C^n(\theta) + \sigma^2_{\varepsilon} I_n)^{-1} C^{n,\,new}(\theta)$$

It can be readily seen from Eq 1that the corresponding conditional distribution for $\Im(\mathbf{x}_{new})$ is also multivariate normal with mean vector M_{new} and covariance matrix given by $S_{new} - \sigma^2_\varepsilon I_k$. Having established $p(\mathbf{y}_{new} \mid \theta,\ \sigma^2_\varepsilon, \text{data}, \mathbf{x}_{new})$ and $p(\Im(\mathbf{x}_{new}) \mid \theta,\ \sigma^2_\varepsilon, \text{data}, \mathbf{x}_{new})$ through Eqs. 13-16, draws from the posterior predictive distributions for $\mathbf{y}_{new}$ and $\Im(\mathbf{x}_{new})$ are readily obtained using Eqs. 11 and 12 and the posterior draws for θ and σ^2_ε. Hence, full inference is available for any feature of the posterior predictive distribution for the stock-recruitment relationship.

Reference points

For any method of estimating the S-R relationship to be of practical value, it is important that certain biological reference points may be calculated in a straightforward manner. Here we outline how posterior predictive distributions for unfished biomass (B_0), the steepness (h) of the SR curve (Mace and Doonan 1988), and the stock biomass at maximum sustainable yield (B_{MSY}) can be obtained in our Bayesian nonparametric framework.

Although more realistic models are certainly possible, we assume for simplicity that the stock dynamics on an annual time step are given by

$$S_{t+1} = S_t - (M+F)\, S_t + R(S_t) \tag{17}$$

where M and F are the fractions of the population removed by natural and fishing mortality over the course of a year and $R(S_t)$ is recruitment written explicitly as a function of stock at time t. Thus, at equilibrium in the absence of fishing, B_0 is the solution of $R(S_t) = MS$. Correspondingly, $x_0 = \ln[B_0]$ is the value of x at which $\Im(x)$ crosses the line $y = x + \ln(M)$. For any posterior predictive realization of $\Im(\mathbf{x}_{new})$ there may be several such points, the number of which depend strongly on the amount of variability in the data. Here, we focus attention solely on those points that correspond to locally stable equilibria, i.e. where $\Im(x)$ is greater than $x+\ln(M)$ for x less than x_0 and less than $x+\ln(M)$ for x greater than x_0. Formally, the local stability criteria is $\Im(x_0 - \varepsilon) - (x_0 - \varepsilon) > \Im(x_0 + \varepsilon) - (x_0 + \varepsilon)$ for some small ε. Specific values of x_0 and

$\Im(x_0)$ may be obtained by interpolating between the grid points of a realization from $\Im(\mathbf{x}_{new})$. For this purpose, we found that a grid of k = 150 points provided sufficient resolution. The steepness of the S-R curve is defined as recruitment at 20% of B_0 expressed as a fraction of recruitment at B_0 (Mace and Doonan 1988). Therefore, after interpolating $\Im(x_0 + \ln[0.2])$ from the realization of $\Im(\mathbf{x}_{new})$, steepness can be estimated by calculating exp[$\Im(x_0 + \ln[0.2])$ - $\Im(x_0)$]. Since yield in our model is given by FS, B_{MSY} is simply the value of S for which R(S)-MS is maximized and can be found directly from exp[$\Im(\mathbf{x}_{new})$] and exp[$\mathbf{x}_{new}$]. Posterior predictive distributions for each of these reference points were estimated from the collection of values obtained from realizations of $\Im(\mathbf{x}_{new})$ calculated at each step of the Metropolis-Hastings algorithm.

Prior Specification

In general, care must be taken with improper priors for the parameters $\boldsymbol{\theta}$ and σ^2_ε as these may lead to improper posteriors (see, e.g., Berger et al. 2001). We choose proper priors that can be specified with small amounts of prior information. Specifically we take

(18) $$\beta_0 \sim N(a_{\beta 0}, b_{\beta 0}),\ \beta_1 \sim N(a_{\beta 1}, b_{\beta 1})$$
$$\alpha \sim U(0,2)$$
$$\tau^2 \sim IG(a_\tau, b_\tau),\ \phi \sim IG(a_\phi, b_\phi),\ \sigma^2_\varepsilon \sim IG(a_\sigma, b_\sigma)$$

where IG denotes the inverse gamma density parameterized as

(19) $$F_{IG}(z \mid a, b) = \frac{b^a}{\Gamma(a)} z^{-a-1} e^{-b/z}$$

From previous studies (Myers 2001), we expect that the slope at the origin should be between 2 and 7 and that there should be no depensation. Thus we choose $a_{\beta 0}$=4.5 and $a_{\beta 1}$= 1. To fully specify the priors for β_0, β_1, τ, and σ^2_ε we note, from Eq. 8, that the marginal response variance $\mathrm{Var}(y_i) = b_\sigma + b_\tau + b_{\beta 0} + b_{\beta 1}x_i^2$. Hence using a prior guess, $\tilde{x}$, as a rough center for the predictor we divide the variance in y_i evenly among the four components by setting $b_\sigma = b_\tau = b_{\beta 0} = (r_y/4)^2$ and $b_{\beta 1} = (r_y/4)^2\, \tilde{x}^{-2}$ where r_y is a prior guess at the range of response variables and $4(r_y/4)^2$ is a guess at $\mathrm{Var}(y_i)$ inflated by a factor of 4. We set $a_\tau = a_\phi = a_\sigma = 2$, resulting in infinite variances for τ^2, ϕ,and σ^2_ε.

It is more difficult to imagine appropriate prior information for α and ϕ. Since α must remain within the interval (0, 2], we choose U(0,2) as the prior. We note that ϕ controls the correlation among points and how rapidly the correlation dies away with distance. From the spatial statistics literature where GP priors are common (Cressie 1993), for the prior mean of $\alpha = 1$, $3/\phi$ is called the "range of dependence," that is, the value of the distance $|x-x'|$ which gives $\mathrm{Corr}(\Im(x) , \Im(x')) \approx 0.05$. We use one half the range of observed x_i′s, $r_x/2$, as a guess at the range of dependence. Thus we set the prior mean of ϕ, b_ϕ, equal to $6/r_x$.

Testing the method

We demonstrate application of the method on two sets of data. The first is data simulated from a suite of 5 parametric models (Table 1): Ricker, Beverton-Holt, Shepherd, Saila-Lorda and an atypical model for a potentially open population. Model formulations and parameters used in the simulations are given in Table 1. The fifth parametric model, ′Open-Mixture′, is a linear combination of the Beverton-Holt and Saila-Lorda models with an additional supply of recruits (d) from outside the specified stock area.

We chose parameters for all models so that the maximum and equilibrium recruitment levels were nearly the same. For each of the simulations, we sampled 40 stock sizes from a triangular distribution with support on [0,750] and mode 250. Given the chosen parameters, this sampling tends to emphasize points near the stock size of maximum recruitment. For each sampled stock size, observed recruitment was sampled as $R = f(S)\, e^{\varpi}$ where $\varpi \sim N(-\frac{1}{2}V_p, V_p)$. We conducted simulations using a small observation error $V_p = 0.001$ or a larger observation error $V_p = 0.01$. For each data set so constructed, we fit the Bayesian nonparametric model with priors as described above, using a burn-in period of 10,000 draws followed by a posterior sampling period of another 10,000 draws. For quantitative measures of fit, we calculated the residual sum of squares (SS = $\Sigma(y-E[\Im(x)])^2$) and Akaike's information criterion (AIC, Burnham and Anderson 1998). For comparison, we calculated SS and AIC for Ricker and Beverton-Holt models fit to each data set as well using Bayesian approach with uninformative priors for all of the parameters. We also fit a LOESS (Cleveland 1981) smoother to each. Since the stiffness of the LOESS is arbitrary and the minimum sum of squares is zero, we do not calculate a measure of fit, but include LOESS with 'stiffness' set to 0.6 in the figures for visual comparison.

One of the chief parameters of interest that come directly from this model specification is the error variance σ^2_ε . To address how reliably our nonparametric approach could recover estimates of error variance, we simulated data from the Ricker and Beverton-Holt models in the manner described above for 100 values of V_p ranging from 0.01 to 0.5. We also calculate posterior distributions for B_0, h, and B_{MSY} based on 5000 realizations of $\Im(\mathbf{x}_{new})$ on a grid of 150 points.

We demonstrate the method on actual observations of stock and recruits for sockeye (*Oncorhynchus nerka*), chum (*Oncorhynchus keta*), and pink (*Oncorhynchus gorbuscha*) salmon from the Weaver Creek spawning channel, British Columbia (Essington et al. 2000) and northern and southern lingcod (*Ophiodon elongatus*) stocks from central California through Alaska (Jagielo et el. 2000). For each of these series, we fit the Bayesian nonparametric model as well as Ricker, Beverton-Holt models and a LOESS smoother (stiffness = 0.6).

Results

The fit of the Bayesian nonparametric model to each of the simulated data sets was quite good (Figure 2). In each case, the shape of the BNP model closely follows that of the simulated model over most of the range covered by the data. One of the most attractive features of this approach is that posterior predictive uncertainty bands for $\Im(x)$ are easily obtained from the Monte Carlo samples. Note that these bands tend to be fairly narrow in regions with sufficient data and wider in regions where data are scarce. The only substantial departures from the model occur when overcompensation is present (i.e. Ricker, Shepherd, and Saila-Lorda) and then only at the largest stock sizes. This deviation is a reflection of the choice of prior mean function and the scarcity of data at the largest stock sizes; had we assumed an overcompensating prior mean function this deviation would not occur. Comparison with the fit of the parametric models is generally quite favorable. The BNP fit (SS or AIC, Table 2) was usually close to the fit of the correct parametric model, and was always better than the fit of the incorrect model. In real fisheries, the relationship between stock and recruitment is likely to involve several different sources of density dependence and, potentially, sources of recruits from other stocks. The Open-Mixture model is a simple example. The Bayesian nonparametric method was the only one to accurately describe data from this model, highlighting its utility in situations where the underlying dynamics are unknown and complex.

A common concern when fitting non-parametric models to noisy data is that the resulting model may be over-fit. However, estimates of the error variance with the BNP model were quite good (Fig 3). The expected value of the posterior σ^2_ε was, on average, within 15% of the true value for $\sigma^2_\varepsilon > 0.2$. For $\sigma^2_\varepsilon < 0.2$, the estimated variance was substantially greater than the true value, although confidence regions from the posterior σ^2_ε typically contained the true value for σ^2_ε as small as 0.1. Overall, the estimated variance was more accurate for data generated using the Beverton-Holt model than for the Ricker model. This is a consequence of our choice of the prior mean function. Overcompensatory choices of prior mean function would produce substantially better estimates of the variance under the Ricker model. In contrast with smoother-based methods, overfitting rarely occurs. Although the error variance estimated using a parametric model is strongly influenced by proper model choice, the BNP method allows reasonably accurate estimation of the error variance in the absence of model specification. Since recruitment variability is probably the single most important variable driving variation in future stock sizes, this feature of the BNP approach will be of great utility in risk assessments.

Reference points

The ability of the BNP method to estimate reference points was generally good (Fig 4). Posterior means for B_0 and steepness were within a few percent of the true values for data sampled from both Ricker and Beverton-Holt models. The posterior mean for B_{MSY} for Ricker data was substantially biased due to a large number of samples with solutions near the origin. For Beverton-Holt data, on the other hand, the method estimated B_{MSY} to within < 1%.

Fit to actual data

For each of the salmon stocks (Fig 5), the Ricker, Beverton-Holt and BNP models all indicate recruitment increasing continuously with stock size. In the case of both chum and pink salmon, the BNP fit appears to indicate the existence of at least one inflection point (perhaps two); the curve increases from the origin, levels out at stock sizes around 25% of the observed maximum, and then increases again. Although it is impossible to say for certain, this pattern of recruitment may indicate alternative regimes, each characterized by different relationships. In contrast, the fit to the lingcod data (Fig. 6) is essentially flat for both northern and southern stocks. Uncertainty bands near the origin widen dramatically, indicating that there is no SR information in this region.

Discussion

The Bayesian nonparametric approach we have described offers a number of advantages over previous methods of defining the relation between stock size and subsequent recruitment. Regarding previous nonparametric approaches, the advantages are clear: no ad hoc smoothing parameter need be chosen, and biological information is easily incorporated through specification of the prior mean function. Moreover, choice of the prior mean function does not strictly dictate the shape of $\Im(x)$ within the range of the data.

The chief advantage of the Bayesian non-parametric over parametric approaches to inference is that only one model is required to fit the data regardless of the true underlying dynamics; dealing with model uncertainty is explicitly part of the BNP approach. Thus one could envision a scenario in which Ricker density dependence is postulated in the choice of prior mean function, but the posterior mean behaves as a Beverton-Holt. With previous methods, multiple models (parametric or non-parametric) would have to be fit and a probabalistic assessment (e.g. via AIC) of which one was better would have to be made. The Bayesian non-parametric model allows good description of data sampled from parametric models, coming nearly as close to the underlying true model as fitted parametric curves. Moreover, with an unknown true model, BNP is the only method that works effectively in all cases. It is, for example, the only model that came close to the true model in the case of the Open-Mixture stock recruitment relationship.

As with other non-parametric methods (and parametric models as well, Ludwig 1995), biologically unreasonable fits are possible outside the range of the data. However, outside the range of the data, confidence limits grow rapidly and the posterior returns to the prior mean function, indicating that there is little information added by the data to points outside of the range of the data (Ludwig 1995). Consequently, choosing a prior mean function close to the true underlying function will produce better extrapolations. Similarly, uncertainty bands near the origin tend to be large when there is no data near the origin, as in the lingcod example. This is precisely what one wants from an inferential technique; regardless of their importance as management tools, point estimates of recruitment from parametric models outside the observed range of the data are little more than elegantly concocted fictions.

While the Bayesian non-parametric approach represents a step forward in modeling the stock-recruit relationship, there are a number of issues we have not explicitly addressed. These problems in the modeling of SR data have been raised by other authors and are raised here to point out directions in which the Bayesian nonparametric approach may be expanded. The first is that we have assumed measurement errors in stock biomass are small compared with estimates of recruitment. This may be explicitly addressed in the Bayesian nonparametric framework by including an additional layer in the model hierarchy representing the unobserved, ′true′ state of the stock. Second, in some sense there is no single ′stock recruitment relationship′ and certainly no analytical formula for one, so that there are no formulae (simple or complex) relating parameters of the stock recruitment relationship to reference points. Third, we have explicitly ignored the time series nature of stock-recruit data. However, it is possible to incorporate a time series structure in the BNP framework. We have also explicitly assumed a stationary covariance function. It is straightforward to allow non-stationarity in the GP prior over stock sizes. It is possible that the nature of the relationship between stock and recruits is not constant over time. This sort of temporal non-stationarity is dealt with in dynamic linear models and could, in principle, be incorporated into the Bayesian nonparametric approach. Fourth, stock biomass may not be an appropriate indicator of reproductive output (Rochet 2000). This is a valid concern, but it cannot be addressed in a statistical framework. Rather, the realities of reproductive biology need to be address by the regular collection of relevant data. A future paper will address these and other extensions currently under development.

Stock assessment using Bayesian nonparametric methods provides a new tool for fishery management that allows one to use the data to dictate the shape of the stock-recruitment relationship. Given the great objective flexibility of this method and the certainty that the true underlying stock-recruit relationship is always unknown, the potential contribution to fishery management of this new approach is great.

Acknowledgements. This work was partially supported by the Center for Stock Assessment Research, University of California Santa Cruz.

References

Berger, J.O., V. De Oliveira, and B. Sanso. 2001. Objective Bayesian analysis of spatially correlated data. J. Am. Stat. Assoc. **96**: 1361-1374.

Bravington, M.V., T. K. Stokes, and C.M. O′Brien. 2000. Sustainable fishing: the bottom line. Mar. Freshwater Res. **51**:465-475.

Brodziak, J. K. T., W. J. Overholtz, and P. J. Rago. 2002. Reply: Does spawning stock affect recruitment of New England groundfish? Interpreting spawning stock and recruitment daa in New England groundfish. Can. J. Fish. Aquat. Sci. **59***:*193-195.

Burnham, K. P., and D. R. Anderson. 1998. Model Selection and Inference. A Practical Information-Theoretic Approach. Springer Verlag, New York.

Chapman, D.G. 1973. Spawner-recruit models and estimation of the level of maximum sustainable catch. Rapp. P.-v. Reun. Cons. int. Explor. Mer.**164:** 325-332.

Cleveland, W.S. 1981. LOWESS – A program for smoothing scatterplots by robust locally weighted regression. Am. Stat. **35**:54

Cook, R. M. 1998. A sustainability criterion for the exploitation of North Sea Cod. ICES J. Mar. Sci. **55**: 1061-1070.

Cook, R.M. 2000. A rough guide to population change in exploited fish stocks. Ecol. Lett. **3**: 394-398.

Cressie, N. 1993. Statistics for Spatial Data. Revised Edition. Wiley, New York.

Cushing, D.H. 1973. Dependence of recruitment on parent stock. J. Fish. Res. Bd. Can **30:**1965-1976.

Denison, D.G.T., C.C. Holmes, B.K. Mallick, A.F.M. Smith. 2002. Bayesian Methods for Nonlinear Classification and Regression. Wiley.

Essington, T.E., T.P. Quinn, and V.E. Ewert. 2000. Intra- and inter-specific competition and the reproductive success of sympatric Pacific salmon. Can. J. Fish. Aquat. Sci.**57**: 205-213

Evans, G.T. and J.C. Rice 1988. Predicting recruitment from stock size without mediation of a functional relation. J. Conseil. **45**: 73-81.

Jagielo, T.H., D. Wilson-Vandenberg, J. Sneva, S. Rosenfield, and F. Wallace. 2000. Assessment of lingcod (*Ophiodon elongates*). *In* Status of the Pacific Coast Groundfish Fishery through 2000 and Recommended Biological Catches for 2001: Stock Assessment and Fishery Evaluation. Pacific Fishery Management Council. Portland, OR.

Johnson, N.L. and S. Kotz. 1970. Continuous Univariate Distributions 1. John Wiley & Sons. New York.

Gamerman, D. 1997. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. CRC Press, London.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika. **57**:97-109

Hilborn, R., and M. Mangel. 1997. The Ecological Detective. Confronting models with data. Princeton University Press, Princeton, NJ.

Ludwig, D. 1995. Uncertainty and Fisheries Management. Lecture Notes in Biomathematics **100**:516-528.

Mace, P.M. and I.J. Doonan. 1988. A generalized bioeconomic simulation model for fish population dynamics. N.Z. Fish. Assess. Res. Doc. 88/4.

McAllister, M. K., and G. P. Kirkwood. 1998. Bayesian stock assessment: a review and example application using the logistic model. ICES J. Mar. Sci. **55**:1031-1060.

Metropolis, N. A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machine. J. Chem. Phys. **21:** 1087-1091.

Millar, R. B. 2002. Reference priors for Bayesian fisheries models. Can. J. Fish. Aquat. Sci. **59**:1492 1502.

Müller, P. and F.A. Quintana. 2004. Nonparametric Bayesian data analysis. *To appear in* Statistical Science.

Myers, R. A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES J. Mar. Sci. **58**:937-951.

Neal, R. M. 1999. Regression and classification using Gaussian process priors. *In* Bayesian Statistics 6. Bernardo, J.M., J.O., Berger, A.P. David, and A.F.M. Smith, Eds. Clarendon Press, Oxford.

Needle, C.L. 2002. Recruitment models: diagnosis and prognosis. Reviews in Fish Biology and Fisheries. **11**: 95-111.

Quinn, T. J. I., and R. B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press, New York.

Rochet, M.J. 2000. Does the concept of spawner per recruit make sense? ICES J. Mar. Sci. **57**: 1160-1174.

Robert, C.P. and G. Casella. 1999. Monte Carlo Statistical Methods. Springer. New York.

Rothschild, B.J. and A.J. Mullen. 1985.The information-content of stock-and-recruitment data and its nonparametric classification. J. Conseil **42**: 116-124.

Walker, S.G., P. Damien, P.W. Laud, and A.F.M. Smith 1999. Bayesian nonparametric inference for random distributions and related functions (with discussion), J. Roy. Stat. Soc. B. **61**: 485-527

Table 1.

Model	Parameterization	α	β_1	β_2	c	d
Ricker	$R = \alpha S e^{-\beta_1 S}$	5.437	0.004	N/A	N/A	N/A
Beverton-Holt	$R = \frac{\alpha S}{1+\beta_2 S}$	5.000	N/A	0.010	N/A	N/A
Shepherd	$R = \frac{\alpha S}{1+(\beta_2 S)^c}$	6.000	N/A	0.0063	1.50	N/A
Saila-Lorda	$R = \alpha S^c e^{-\beta_1 S}$	0.006	0.010	N/A	2.50	N/A
Open-Mixture	$R = \frac{\alpha S^c e^{-\beta_1 S}}{1000} + \frac{\alpha S}{1+\beta_2 S} + d$	6.200	0.0200	0.015	2.65	20

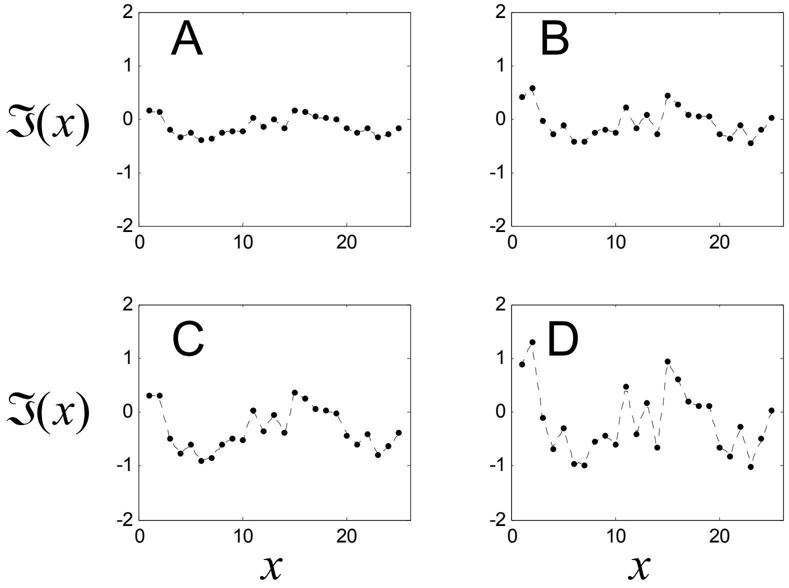
Table 2. Results for simulated data

			Residual sum of squares			AIC		
Error Variance	True Model	N	Ricker	Bev-Holt	BNP	Ricker	Bev-Holt	BNP
Low	Ricker	40	0.02	2.31	0.07	-290.58	-108.05	-241.41
	Bev-Holt	40	0.67	0.04	0.05	-157.61	-269.91	-258.55
	Shepherd	40	0.37	0.51	0.04	-181.11	-168.28	-266.79
	Saila-Lorda	40	13.42	19.89	0.19	-37.69	-21.94	-201.44
	Open-Mixture	40	1.34	0.75	0.04	-129.91	-153.06	-264.81
High	Ricker	40	0.48	3.18	0.49	-171.22	-95.32	-163.79
	Bev-Holt	40	1.07	0.46	0.44	-138.82	-172.80	-168.74
	Shepherd	40	0.65	0.79	0.33	-159.00	-151.06	-180.48
	Saila-Lorda	40	13.35	19.62	0.40	-37.89	-22.50	-172.48
	Open-Mixture	40	1.69	1.24	0.36	-120.53	-133.11	-176.53

Table 3. Results for actual stock-recruit data

		Residual sum of squares			AIC		
Population	n	Ricker	Bev-Holt	BNP	Ricker	Bev-Holt	BNP
Sockeye	34	2.62	2.62	2.44	-81.10	-81.15	-77.58
Chum	34	6.32	7.65	4.49	-51.19	-44.73	-56.86
Pink	18	4.42	4.32	2.74	-19.29	-19.70	-21.89
Lingcod (north)	19	4.63	3.41	2.39	-20.82	-26.61	-27.36
Lingcod (south)	19	6.60	6.29	4.56	-14.08	-15.01	-15.11

Figure 1. Effects of parameters of the covariance function on the sampled Gaussian process and interpolated mean function. α is 1.0 for the top four panels and 2.0 for the bottom four. In both sets, letters A-D correspond to the following values of ϕ and τ: A) 0.1, 0.1, B) 0.1, 0.5, C) 0.5, 0.1, D) 0.5, 0.5. The mean function is zero throughout.



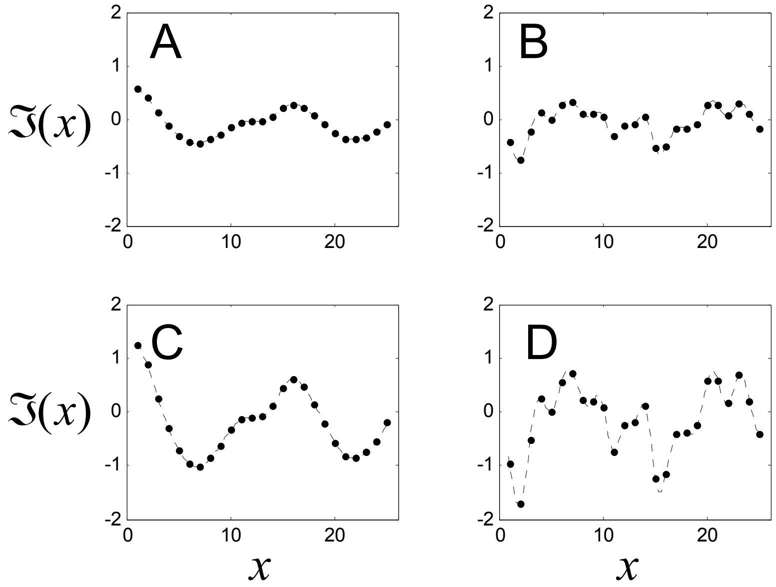


Figure 2. Comparison of parametric and Bayesian non-parametric estimates of stock-recruitment relationships when the data are generated by a known model and fit with a variety of models. Left column: Vp = 0.001, right column: Vp = 0.01. True models are A. Ricker, B. Beverton Holt, C. Shepherd, D. Saila-Lorda, E. Open-Mixture. In each panel, the true model is depicted by the blue line. Blue points are data sampled from the true model. Under the BNP method, the solid black line gives the posterior mean and the dotted lines are 95% pointwise uncertainty bands. Fitted Ricker and Beverton-Holt models are shown by the green and magenta lines. A Loess smoother is indicated by the red line.

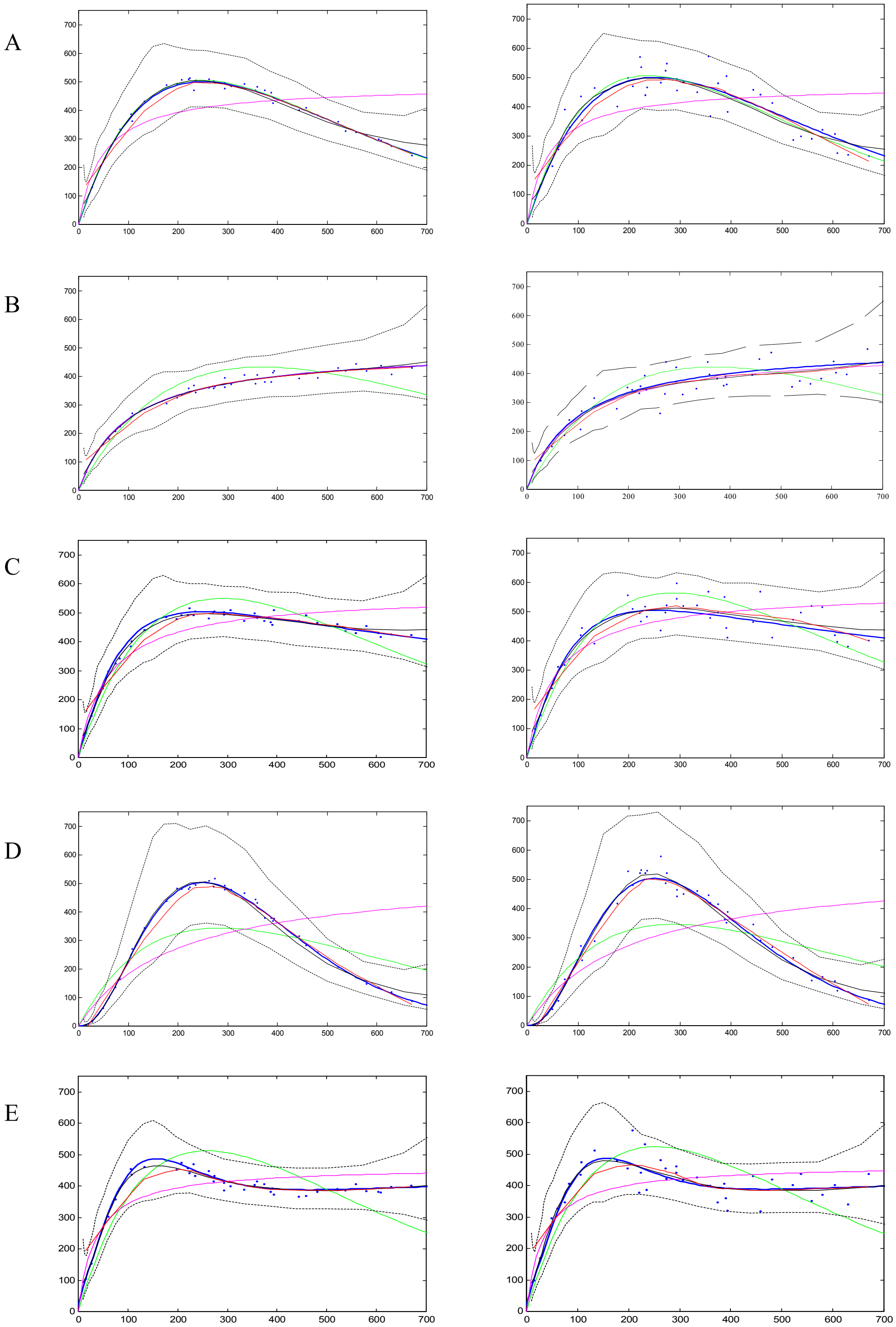


Figure 3. True σ^2_ε and posterior expectations for data simulated from Ricker (open circles) and Beverton-Holt models (filled circles). The models were parameterized as described in the text.

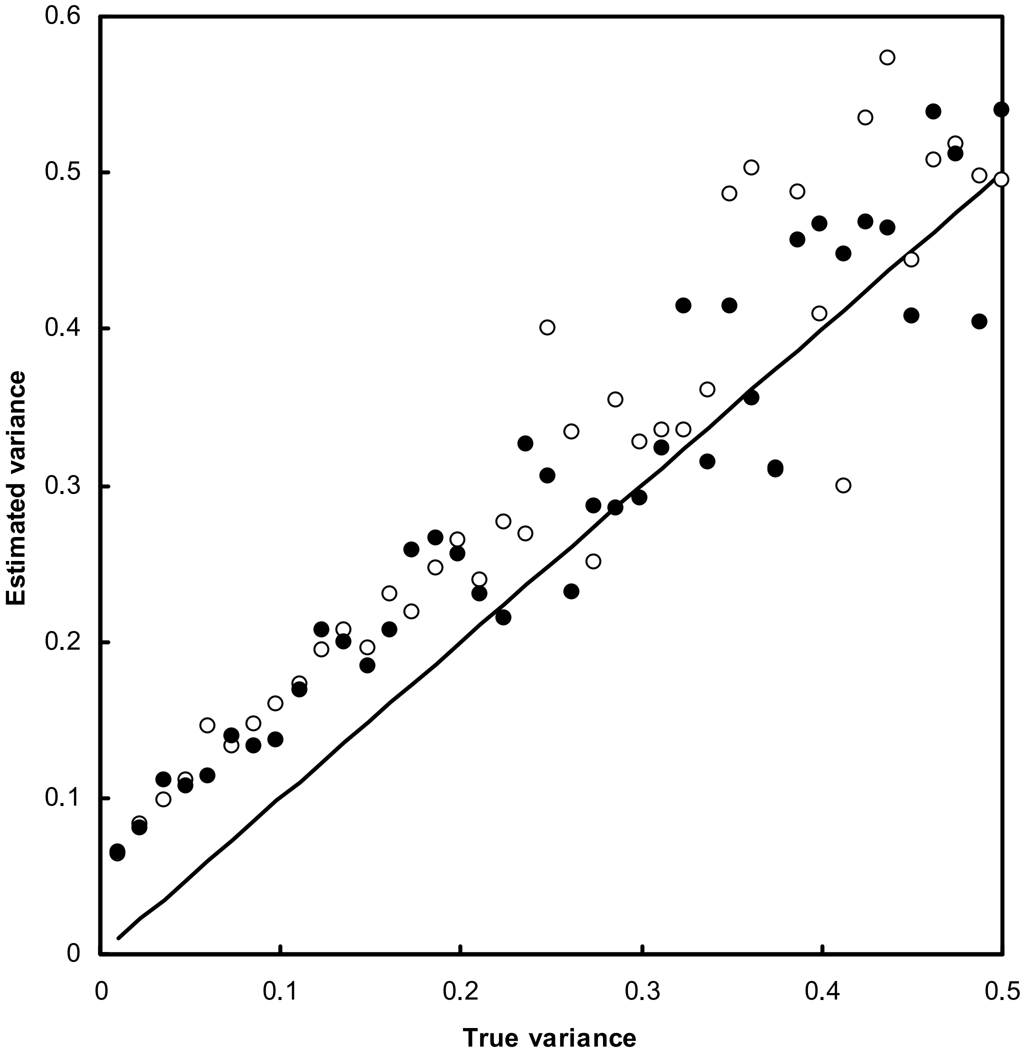


Figure 4. Posterior predictive distributions of B_0 (A,B), B_{MSY} (C,D) and steepness (E,F) for Ricker (left column) and Beverton-Holt (right column) models. Solid triangles indicate the true value, open triangles indicate the posterior mean. Solid lines are kernel density estimates for the posteriors.

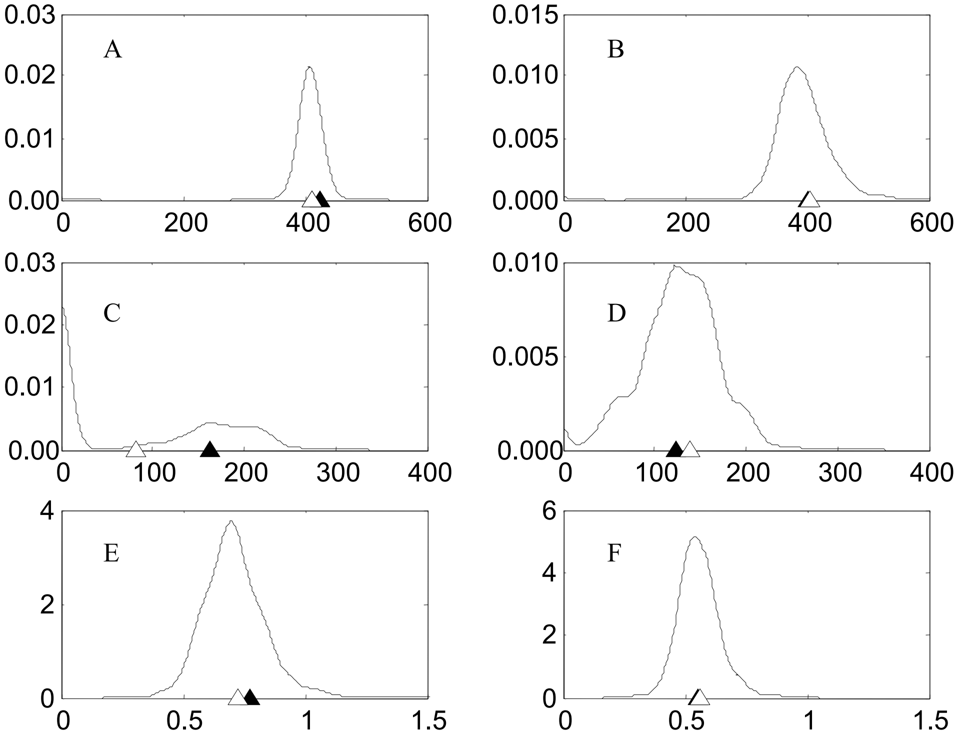


Figure 5 Stock recruit functions for three salmon species. A. sockeye, B. chum, and C. pink salmon. Blue points are observed data. Under the BNP method, the solid black line gives the posterior mean and the dotted denote 95% pointwise uncertainty bands. Fitted Ricker and Beverton-Holt models are shown by the green and magenta lines. A Loess smoother is indicated by the red line. Data from Essington et al. 2000.

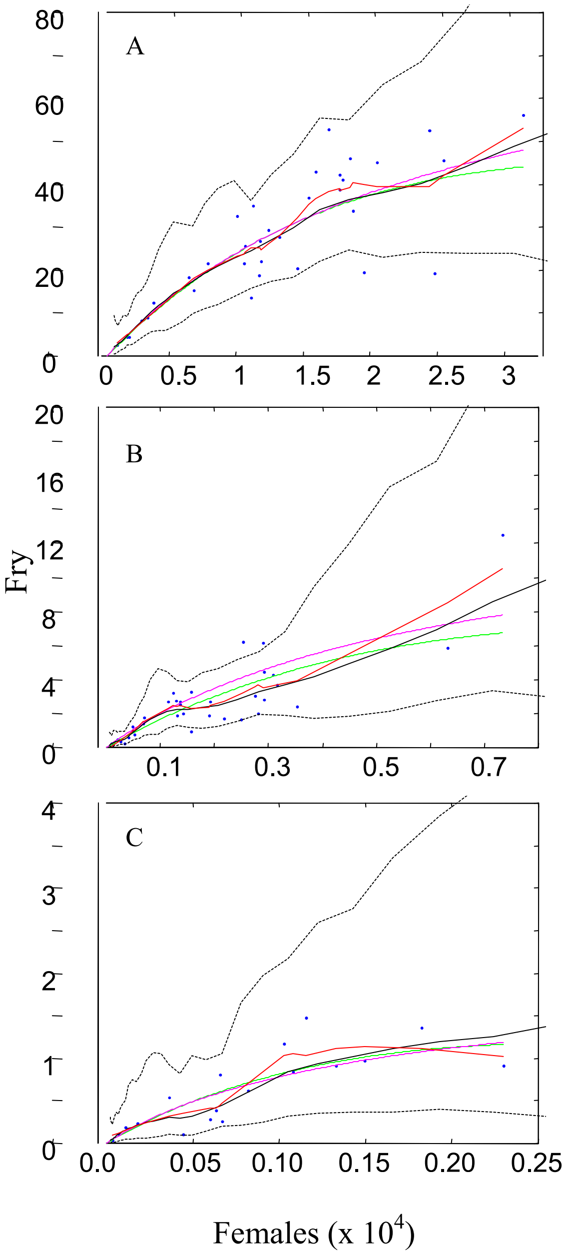


Figure 7. Lingcod stock-recruit models for northern (A) and southern (B) stocks. Blue points are observed data. Under the BNP method, the solid black line gives the posterior mean and the dotted denote 95% pointwise uncertainty bands. Fitted Ricker and Beverton-Holt models are shown by the green and magenta lines. A Loess smoother is indicated by the red line. Data from Jagielo et al. 2000.