Ecology, 95(2), 2014, pp. 329–341
© 2014 by the Ecological Society of America

A Bayesian approach to identifying and compensating for model misspecification in population models

James T. Thorson,[1,4] Kotaro Ono,[2] and Stephan B. Munch[3]

[1]*Fisheries Resource Assessment and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. East, Seattle, Washington 98112 USA*
[2]*University of Washington, School of Aquatic and Fishery Sciences, Box 355020, Seattle, Washington 98195-5020 USA*
[3]*Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 110 Shaffer Road, Santa Cruz, California 95060 USA*

Abstract. State-space estimation methods are increasingly used in ecology to estimate productivity and abundance of natural populations while accounting for variability in both population dynamics and measurement processes. However, functional forms for population dynamics and density dependence often will not match the true biological process, and this may degrade the performance of state-space methods. We therefore developed a Bayesian semiparametric state-space model, which uses a Gaussian process (GP) to approximate the population growth function. This offers two benefits for population modeling. First, it allows data to update a specified "prior" on the population growth function, while reverting to this prior when data are uninformative. Second, it allows variability in population dynamics to be decomposed into random errors around the population growth function ("process error") and errors due to the mismatch between the specified prior and estimated growth function ("model error"). We used simulation modeling to illustrate the utility of GP methods in state-space population dynamics models. Results confirmed that the GP model performs similarly to a conventional state-space model when either (1) the prior matches the true process or (2) data are relatively uninformative. However, GP methods improve estimates of the population growth function when the function is misspecified. Results also demonstrated that the estimated magnitude of "model error" can be used to distinguish cases of model misspecification. We conclude with a discussion of the prospects for GP methods in other state-space models, including age and length-structured, meta-analytic, and individual-movement models.

Key words: Bayesian models; Gaussian process; model misspecification; population dynamics model; semiparametric models; state-space models.

Introduction

Predicting population growth is one of the pillars of applied ecology. Models of population growth have many implications for conservation and management; among other things, they are used to determine estimates of extinction risk (Ginzburg et al. 2005) and harvest rates for exploited populations (Quinn and Deriso 1999). However, although numerous population models have been developed, observed population growth may not clearly follow any single parametric model across all species (Thorson et al. 2012). This is particularly problematic because conservation and management strategies may be quite sensitive to model choice (Harwood and Stokes 2003).

Several authors have used methods that do not specify a parametric relationship between predictive and response variables (i.e., nonparametric methods) to address this uncertainty in model specification (Jost and Ellner 2000, Wood and Augustin 2002, Munch et al. 2005). Nonparametric methods are capable of capturing a wide range of functional forms and consequently are a robust tool for modeling population growth. However, most of the previously developed nonparametric methods assume that population data are observed without error or confound different sources of uncertainty.

State-space versions of population models have been developed to account for sampling imprecision (Harwood and Stokes 2003, Royle and Dorazio 2008, Knape and de Valpine 2012), and outperform simpler estimation methods in many cases (de Valpine and Hastings 2002, Punt 2003, Staples et al. 2004). State-space methods separately estimate variance in the measurement process (due, for example, to small sample sizes for the index of abundance), and variance in the population growth process (due, for example, to demographic stochasticity [Lande et al. 2003] and environmental variability). However, state-space population models are likely to be sensitive to any misspecification of the population growth function (Calder et al. 2003) or other model misspecifications, e.g., failing to account for

Manuscript received 29 January 2013; revised 22 July 2013; accepted 23 July 2013. Corresponding Editor: B. D. Inouye.
[4] E-mail: James.Thorson@noaa.gov

competitive effects (Walters and Kitchell 2001) or eco-evolutionary changes in individual characteristics (Smallegange and Coulson 2012). Intuitively, errors in model specification must be accounted for by elevated estimates of process and observation variance, which (among other things) may lead to biased estimates of extinction risk (Staples et al. 2004, Ginzburg et al. 2005).

Thus, there is a clear need to develop a unified framework to identify and account for uncertainty in model specification, process uncertainty, and observation uncertainty in population modeling (Harwood and Stokes 2003). Although the literature on nonparametric time series analysis is substantial (see e.g., Gao 2007), relatively little has been done on semiparametric inference in a state-space context. Here we adopt a Bayesian view by specifying a flexible prior on the population growth function rather than a specific parametric function (Munch et al. 2005). Bayesian estimates of functions have not been incorporated into state-space models in the ecological literature, and their performance in these circumstances is unknown.

Here, we use simulated data to demonstrate that Bayesian estimates of functions in state-space models can generically provide two benefits for population dynamics models. First, they allow data to update the assumed prior for the population growth functions. Second, they allow estimation of the relative magnitude of process and model errors. We conclude by discussing the prospects for Bayesian estimates of functions in other state-space ecological contexts such as models of early juvenile survival or individual movement.

Methods

Population growth models

Models approximating population growth can be estimated using an index of abundance and optionally a time series of known mortality, i.e., harvest. When a relative index of abundance is available (which is proportional to abundance up to a fixed but unknown value called the "detectability coefficient"), population growth models use harvest time series data to estimate the detectability coefficient. In such models, a decrease in abundance that coincides with high harvest indicates that harvest is probably a nontrivial proportion of total abundance, and therefore gives information about the absolute scale of abundance, i.e., the detectability coefficient. Alternatively, population growth models can be fitted to abundance time series in the absence of harvest data when an absolute index of abundance is available, i.e., the detectability coefficient is known (e.g., Knape and de Valpine 2012). Population growth models can also be estimated using only a relative index of abundance (when information about absolute population biomass is not sought), or using auxiliary information on demographic rates (when this is available), although we do not here consider these latter two model types.

We begin by modeling a population subject to density dependence and process uncertainty which can be represented by

$$B_{t+1} = \Big(B_t + f(B_t) - C_t\Big)e^{\varepsilon_t} \tag{1}$$

where $f(\cdot)$ represents population growth at population size B_t (either numbers or biomass), C_t is an element in the vector $\mathbf{C}$ representing catches (deaths from known sources of mortality, i.e., harvest, which may be zero), and the total variability in the population growth process ε_t in year t is assumed to be independent from other years and normally distributed with a constant variance σ_ε^2, i.e., $\varepsilon_t \sim \mathcal{N}(-\sigma_\varepsilon^2/2, \sigma_\varepsilon^2)$. Factors not explicitly modeled in the parametric growth function $f(\cdot)$, for example, environmental stochasticity, underlie the residual process uncertainty ε_t (Lande et al. 2003), which includes a bias-correction term such that $f(\cdot)$ represents the mean growth function rather than the median. We recognize that variability might be state-dependent, but for simplicity of presentation we retain this homoscedastic assumption throughout.

A parametric population growth function is conventionally specified for $f(\cdot)$ to predict population growth as a function of population size. A common growth function for $f(\cdot)$ in ecology is the theta-logistic model, which is similar to the Pella-Tomlinson function in fisheries (Quinn and Deriso 1999). Our parameterization of the Pella-Tomlinson function includes two parameters: K representing positive abundance at which population growth is expected to be zero, and Y representing maximum population growth (although we re-parameterize for computational reasons using α, defined as the ratio of $\gamma Y/K$, where γ is defined in Eq. 3). The Pella-Tomlinson function also included a shape parameter φ that determines the ratio of abundance at maximum population growth to equilibrium abundance (B_Y/K):

$$f(B_t) = \alpha B_t - \alpha K\left(\frac{B_t}{K}\right)^{\varphi} \tag{2}$$

where γ is derived from the Pella-Tomlinson shape parameter,

$$\gamma = \frac{\varphi^{\varphi/\varphi-1}}{\varphi - 1}. \tag{3}$$

Recent research suggests that B_Y/K has a mean of ~40% for many fish species (Thorson et al. 2012). To provide a specific simulation example, we therefore specify by default the φ corresponding to $B_Y/K = 0.4$. However, we note that the following simulation comparison could be applied to many other population growth functions.

In a state-space framework, we explicitly recognize that the observed population sizes are imprecise. Here, we assume that the available index data $\mathbf{I}$ (where $\mathbf{I}$ is a

vector) are noisy measurements which are roughly proportional to the actual population size:

$$I_t = qB_t e^{\tau_t} \tag{4}$$

where q is the detectability coefficient (a proportionality constant representing the fraction of the population observed) and the observation errors τ_t are assumed to be independent and normally distributed, i.e., $\tau_t \sim \mathcal{N}(-\sigma_t^2/2, \sigma_t^2)$, where $\boldsymbol{\tau}$ is the vector of τ_t for all years from 1 to the final year T. Measurement errors are specified as independent due to the presumed independence of sampling errors (e.g., sampling variance from a survey that generates the index $\mathbf{I}$ is independent each year), and as having an approximately constant log(standard deviation), as is commonly observed in fishery indices (Thorson and Ward 2013).

In the parametric state-space model (which we hereafter refer to as the "conventional" approach), we assume that the growth function $f(\cdot)$ is known without error and we try to estimate model parameters (K, α, B_t, q, σ_ε^2, σ_t^2) using the information from the index of abundance and optionally harvest. This state-space model requires integration across states B_t, as frequently accomplished using particle filters or Markov chain Monte Carlo (as explained later).

The Bayesian approach to model misspecification

In the Bayesian approach to estimating the population growth function, we relax the assumption that the dynamics are, on average, correctly described by Eq. 1. To do so, we partition ε_t in Eq. 1 into components $z(b_t) + \eta_t$, where η_t is uncorrelated process noise with mean 0 and variance σ_p^2. The function $z(b_t)$ is an unknown function that allows for persistent, smooth departures from the assumed model structure, and that hence approximates the difference between true dynamics and the putative population growth function. As is standard in Bayesian statistics, we assign z a prior distribution representing what we know, and we use the available data to update this distribution. One possible prior for z that is flexible and computationally feasible is a Gaussian process (which we hereafter refer to as the "GP" approach; Rasmussen and Williams 2006). GPs have been used widely in spatial ecology under the guise of Kriging (Banerjee et al. 2003) and have since been applied to modeling density dependence (Munch et al. 2005), detecting Allee effects (Sugeno and Munch 2013), and determining seasonal variation in growth potential (Sigourney et al. 2012). An obvious alternative would be to model z using a basis expansion (spline, Fourier, and so forth), as is commonly done in generalized additive models (Wood and Augustin 2002). These approaches are dual in the sense that specifying a basis implies a covariance function, while specifying a covariance function implies a basis expansion (Rasmussen and Williams 2006). We adopt the GP perspective because it allows for relatively straightforward incorporation of biological information (via the parametric prior) and a priori control of wiggliness in the population growth function.

The GP is specified in terms of a mean function $\mu(b)$, which gives the predicted population growth for any hypothetical abundance b, and a covariance function $C(b,b')$, which gives the covariance of deviations away from the mean function for abundance b and b'. Because b and b' can be any continuous values for abundance, the GP can be interpreted as a continuous generalization of a multivariate normal distribution. This specification is written succinctly as

$$z \sim \mathrm{GP}(\mu, C). \tag{5}$$

We set the prior mean to zero, $\mu(b) = 0$, reflecting the fact that we do not have prior information on systematic deviations from the putative model (if we did, we would modify Eqs. 1 and 2). We here assume that the average distance between the mean function $\mu(b)$ and the true but unknown function $f(b)$ is approximately constant for all b; in other words, that the variance of $z(b)$ is the same for all b (where this assumption could be further explored in a real-world application). In the case of constant variance for $z(b)$, we can rewrite $C(b,b')$ as $\sigma_m^2 R(b,b')$, where σ_m^2 is the variance in z, and R is the correlation among departures from the putative model for abundance levels b and b'. The choice of correlation function controls whether departures are differentiable and the length scale over which they become independent. We assume a squared-exponential correlation function:

$$R(b, b') = \exp\left(-\frac{(b-b')^2}{2\lambda K}\right) \tag{6}$$

where λ controls the degree of wiggliness in the estimated population growth function, and where K is included in the denominator so that λ is defined as proportion of K. We note that other correlation functions are also possible in a Gaussian process, but the squared-exponential correlation is continuous and infinitely differentiable, similar to most previously developed growth functions.

In the present application, we are interested in quantifying the error arising from model misspecification. Since the covariance function $\sigma_m^2 R(b,b')$ controls deviations from the putative model, we expect the variance σ_m^2 of deviations to go to 0 in the case when the putative model is correct. However, to retain consistency with Eq. 1, the total error variance must be the same. That is, $\sigma_\varepsilon^2 = \sigma_m^2 + \sigma_p^2$ (recall that σ_p^2 is the variance of uncorrelated errors). To facilitate quantifying the relative importance of independent process error and model misspecification errors, we introduce a new parameter, $\theta = \sigma_m^2/\sigma_\varepsilon^2$. Using this parameterization, we rewrite the covariance for $\boldsymbol{\varepsilon}$ as follows:

$$C(b,b') = \sigma_\varepsilon^2\left[\theta\exp\left(-\frac{(b-b')^2}{2\lambda K}\right) + (1-\theta)\delta_{b-b'}\right] \quad (7)$$

where $\delta_{b-b'} = 1$ if $b = b'$ and $\delta_{b-b'} = 0$ otherwise. The first term in the parentheses in the right-hand side of Eq. 7 represents model error, and the second represents residual process variability. In this parameterization, θ represents the fraction of the total uncertainty in state transitions resulting from model misspecification. When the putative model is sufficiently close to the true growth function, we expect θ to approach 0. This suggests a straightforward test for model misspecification using a comparison between the performance of models with $\theta = 0$ and θ free. We also note that, when $\theta = 0$, the GP model reduces to the conventional model because model variance σ_m^2 is 0, and thus the ε_t's are uncorrelated.

To estimate parameters for the GP model, we collect all of the realized process variability in a vector $\boldsymbol{\varepsilon} = \{\varepsilon_1, \ldots, \varepsilon_{T-1}\}^\top$ and the population sizes in a vector $\mathbf{b} = \{B_1, \ldots, B_{T-1}\}^\top$, where subscript T represents the number of years of data. Given the specification in Eqs. 1 and 5, this is equivalent to multivariate normal distribution for $\boldsymbol{\varepsilon}$ given $\mathbf{b}$:

$$\boldsymbol{\varepsilon} \sim \mathcal{N}\left(-\frac{\sigma_\varepsilon^2}{2}1,\ C(\mathbf{b},\mathbf{b}^T)\right) \quad (8)$$

where $C(\mathbf{b},\mathbf{b}^T)$ is the covariance matrix constructed using Eq. 7, i.e., the i and jth element of $C(\mathbf{b},\mathbf{b}^T)$ is Eq. 7 evaluated at population sizes b_i, b_j. The vector of 1's ($\mathbf{1}$) has length $T-1$, where T is the number of years of data.

Calculating the Gaussian process growth function

After fitting the GP model to data, we can visualize systematic departures from the putative model by examining the posterior mean of $z(b)$. However, it is easier to interpret the modified growth function, $\tilde{f}(b)$, defined such that the average dynamics in Eq. 1 are correct when $f(b)$ is replaced by $\tilde{f}(b)$. The posterior estimate of this population growth function $\tilde{f}(b)$ can be computed for any hypothetical biomass b^*. Thus, $\tilde{f}$ can be visualized by computing $\tilde{f}(b^*)$ for a large number of different b^* values and subsequently plotting $\tilde{f}(b^*)$ against b^* for all values. This process is similar for a geostatistical model, where the estimate of an interpolated surface is visualized by estimating its value at all locations on a grid and then plotting that grid (Banerjee et al. 2003). Given the preceding model specification, $\hat{f}(b^*)$ is computed using the estimates arising from the parametric prior $\hat{f}(b^*)$, given estimates of α and K, as well as the vector of process errors $\boldsymbol{\varepsilon}$:

$$\tilde{f}(b^*) = \left(b^* + \hat{f}(b^*)\right)e^{\tilde{z}(b^*)} - b^* \quad (9)$$

where $\tilde{z}(b^*)$ is the posterior mean deviation from the putative model at hypothetical abundance b^*, and $\hat{f}(b^*)$ is calculated from the posteriors distribution of α and K (Eq. 2). Here, $\tilde{z}(b^*)$ is calculated as a weighted average of the observed process errors:

$$\hat{z}(b^*) = \sum_{t=1}^{T}(\omega_t\varepsilon_t) \quad (10)$$

where the vector of weights $\boldsymbol{\omega}$ (where element ω_t is the weight given to process error in year t) is calculated using the covariance among process errors, as well as the covariance between observed errors and $z(b)$:

$$\boldsymbol{\omega} = [C(b^*,\mathbf{b}^T) - \sigma_\varepsilon^2(1-\theta)\delta_{b^*-b^T}]C(\mathbf{b},\mathbf{b}^T)^{-1} \quad (11)$$

where $C(b^*,\mathbf{b}^T)$ is the vector of covariances in z evaluated between b^* (a scalar representing a hypothetical value for abundance) and vector $\mathbf{b}$ (representing observed biomasses for all years $B_1 \ldots B_T$. Here $C(\mathbf{b},\mathbf{b}^T)^{-1}$ is the inverse of the covariance matrix obtained for the observed biomasses, i.e., the inverse of the matrix in Eq. 8. These computations for $\tilde{z}(b^*)$ and $\boldsymbol{\omega}$ are analogous to the steps used to calculate an interpolated surface in geostatistics (Banerjee et al. 2003). Calculating $\tilde{f}(b^*)$ for various fixed values of b^* yields the estimate of population growth function, i.e., an estimate of population growth that is conditioned on the available data. Given that this is a Bayesian model (and we have MCMC samples from the posterior for α, K, $\boldsymbol{\varepsilon}$, and $\mathbf{b}$), this calculation is repeated separately for each MCMC sample, and the mean $\tilde{f}$ is reported.

Model estimation

Both conventional and GP models were estimated using Bayesian methods. This required specifying priors for all estimated parameters. The detectability coefficient was given a reciprocal prior, i.e., $\Pr(q) \propto 1/q$ (which is the Jeffrey's prior for q, given lognormal measurement errors as per Eq. 4; see Millar 2002). This prior specification is identical to assuming a uniform prior on $\ln(q)$ and reflects the fact that q must be positive. As Jaynes (1968) indicates, this is the minimally informative prior that is invariant to a change of scale (i.e., $q' \rightarrow aq$). Following conventional practice, K and Y were given uniform priors. We integrated across states $\ln(\mathbf{b})$ rather than process variability $\boldsymbol{\varepsilon}$, because this parameterization has smaller correlations between $\ln(B_t)$ and $\ln(B_{t+2})$ than would exist between ε_t and ε_{t+2}, and thus improves subsequent MCMC mixing. Here, B_1 was fixed at K (the population was assumed to start at its equilibrium level), both for simplicity of presentation and following Punt (1992). Estimation of both process and observation uncertainties is one of the great challenges in state-space modeling. Typically some prior information is required to separate them and authors have used a variety of devices to do so, including fixing the process noise at some small value (West and Harrison 1997), obtaining an estimate of the observation error from repeated measurements (e.g., Sigourney et al. 2012), or fixing the ratio (Ono et al. 2012). Here, we assigned a joint prior for σ_τ and σ_ε that is marginally uniform on $[0,\infty)$ for both parameters, as is conventionally done for random effects (Gelman 2006), but with an informative prior on their ratio. Specifically we assumed a lognormal prior on

their ratio with log(mean) of 0 and log(standard deviation) of 0.5. This specification retains some flexibility in obtaining separate estimates of σ_τ and σ_ε while incorporating the prior belief that both process variability and measurement errors are significant (see *Results* and Appendix D for further discussion).

The "conventional" state-space model differs from the GP model by stipulating that $\theta = 0$ (and hence causing λ to be irrelevant). This corresponds to the assumption that there is no correlation among process errors, and yields an estimated population growth function that is identical to the prior function, i.e., the putative model (by contrast, $\theta \neq 0$ implies that process errors are correlated via the GP function). The GP requires priors for two additional parameters. Here, θ was given an exponential prior bounded between 0 and 1, i.e., $\Pr(\theta) \propto \exp(-\theta/\mu_\theta)/\mu_\theta$ for all $0 < \theta < 1$, where $\mu_\theta = 0.5$, while the wiggliness parameter λ was given a lognormal prior with log(mean) calculated as ln(0.2) and log(standard deviation) of 1. These latter priors were chosen based on simulating production curves from the prior (i.e., without reference to fitting any data), to verify that they allowed for a generally unimodal population growth function and assigns low prior weight to biologically unreasonable population growth models with multiple modes. This prior is analogous from a frequentist standpoint to the wiggliness penalty approach widely used in fitting GAMs and other nonparametric regression models (Wood and Augustin 2002). Importantly, this was done a priori and was not modified to improve model performance in the simulations. Future research could seek to develop informative priors for these based on meta-analysis.

Markov chain Monte Carlo (MCMC) was used to approximate the marginal posterior distribution for all parameters. We used conventional algorithms and convergence checks as outlined in the Appendix A.

Simulation testing

We used simulation modeling to compare the performance of conventional and GP population dynamics models. We show results for two basic scenarios: absolute index and catch data.

The first scenario uses an absolute index of abundance (i.e., q is known to equal 1). This is plausible for many terrestrial populations, e.g., given distance sampling, and can be used to estimate population growth in the absence of known harvest data (Knape and de Valpine 2012). In these simulations, we assume that $C = 0$, where C is known sources of mortality. In order to ensure that the simulated trajectories exhibit some fluctuations (e.g., cycles or chaos) in the absence of noise, we select a sufficiently high value of maximum growth (Y). This provides sufficient contrast in **B** to allow the parametric methods to estimate productivity.

The second scenario uses a relative index of abundance (i.e., q is freely estimated), such as catch data from a trawl survey of a marine population. In this case, catch is specified to deplete and then recover the population, as is necessary to estimate population growth, given a relative index and catch data (Hilborn and Walters 1991).

These scenarios are designed to achieve contrast in biomass (so that scenarios are informative about population growth for a variety of biomass levels), and are chosen for illustration purposes. Exploratory analysis (not shown) confirmed that conventional and GP models yield similar estimates of production for levels of biomass that are not generated in the data.

We also show results for the following three levels of error. In all cases, $\sigma_\varepsilon = \sigma_\tau$, although previous research has shown that the state-space formulation is robust to different error ratios (Punt 2003, Ono et al. 2012).

(1) For low error, simulation errors were fixed at $\sigma_\varepsilon = \sigma_\tau = 0.05$. This matches the level of errors assumed in Ono et al. (2012), and represents a case where state-space estimation methods are likely to perform well when the assumed population growth function matches the true data-generating process. (2) For moderate error, simulation errors were fixed at $\sigma_\varepsilon = \sigma_\tau = 0.10$. (3) For high error, simulation errors were fixed at $\sigma_\varepsilon = \sigma_\tau = 0.20$. This was used as an upper bound because preliminary analysis showed that higher levels of error did not allow an accurate estimate of model scale (i.e., q and K) for either the conventional or GP models, even when the model was correctly specified. It has additionally been used as an upper bound on errors in other state-space model studies (e.g., de Valpine and Hastings 2002).

Finally, we explored four candidate population growth functions (illustrated in Fig. 1), while specifying that $K = 1$ for each: the correct (matched) assumption, high productivity, stable growth, and depensatory growth.

Correct assumptions.—The Pella-Tomlinson function was used to simulate the data given $B_Y/K = 0.4$ (i.e., the average value for all marine fishes in Thorson et al. 2012), where this value of the shape parameter was also assumed in the conventional model and in the parametric prior on the population growth function for the GP model.

High productivity.—The Pella-Tomlinson function was used to simulate data with $B_Y/K = 0.26$, the average value for highly productive fishes (Clupeiformes) in Thorson et al. (2012), while the conventional model and the parametric prior for GP growth function used $B_Y/K = 0.40$. This scenario matches the case where a species is more productive than expected, and is included to represent mild model misspecification, i.e., it could also be dealt with by using a Pella-Tomlinson model where the shape parameter is estimated.

Stable population growth.—An ad hoc population growth function was formed by taking a power function of the Pella-Tomlinson function where $B_Y/K = 0.4$. This was intended to represent a case where population growth is relatively stable for a wide range of

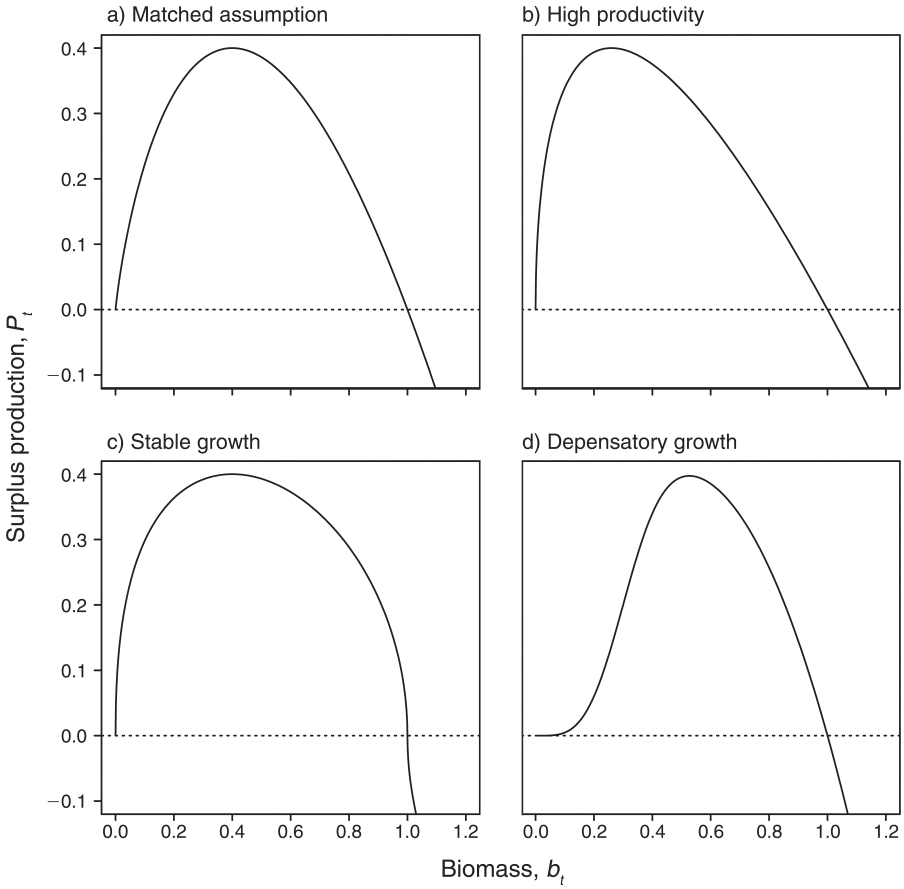

FIG. 1. Population growth functions that give population growth based on surplus production, y-axis) as a function of abundance (based on biomass, x-axis), used for simulation testing of the semiparametric model for population growth, which include (a) a Pella-Tomlinson function with $B_Y/K = 0.4$; (b) a Pella-Tomlinson function with $B_Y/K = 0.26$; (c) a "stable growth" function; and (d) a depensatory Pella-Tomlinson function, where B_Y/K is the ratio of abundance at maximum population growth to equilibrium abundance. See *Simulation testing* for equations and parameter values.

abundance, while still accommodating a unimodal growth function with a fixed carrying capacity:

$$f(B_t) = \mathrm{sign}(B_0 - B_t)\left[\gamma Y^{1/\theta}\frac{B_t}{K} - \gamma Y^{1/\theta}\left(\frac{B_t}{K}\right)^{\varphi}\right]^{\theta} \quad (12)$$

where $\theta = 0.5$.

Depensatory population growth.—A depensatory Pella-Tomlinson model was used (R. Hilborn, *personal communication*), combining the conventional Pella-Tomlinson with the assumption that fewer female individuals will encounter a mate at relatively lower population sizes (Liermann and Hilborn 1997). Given the assumption of random and infrequent encounters (i.e., a Poisson-distributed number of encounters for each female), this results in the following growth function:

$$f(B_t) = \left(\gamma y B_t - \gamma y B_0\left[\frac{B_t}{K}\right]^{\varphi}\right)\left(1 - \exp\left[\ln(0.5)\frac{B_t^{\sigma}}{K^{\sigma}}\right]\right) \quad (13)$$

where σ is the degree of depensation, and B_{50} is the abundance at which 50% of females encounter at least one mate. We parameterized this function such that $B_Y/K = 0.55$, $\sigma = 3$, and $B_{50} = 0.275$.

For the "catch data" scenario, we specified that maximum population growth $Y = 0.4$. For the "absolute index" scenario, we chose a value for maximum population growth (which varies among each combination of growth function and error level) that resulted in a large contrast in biomass. Abundance trajectories are shown in Appendix B.

We simulated 100 replicates for each combination of two scenarios, three error magnitudes, and four population growth functions. Abundance dynamics were simulated using Eq. 1; the index of abundance was simulated using Eq. 5, given $q = 1$; and population growth was simulated using Eqs. 6, 11, or 12. In each replicate, we simulated 40 years of data. For the "catch data" scenario, mortality due to harvest was absent, increased, and subsequently decreased, and we iteratively tuned **C** to achieve the condition that the lowest true abundance is between 20% and 30% of true unfished abundance, given a particular stochastic draw for $\boldsymbol{\tau}$ and

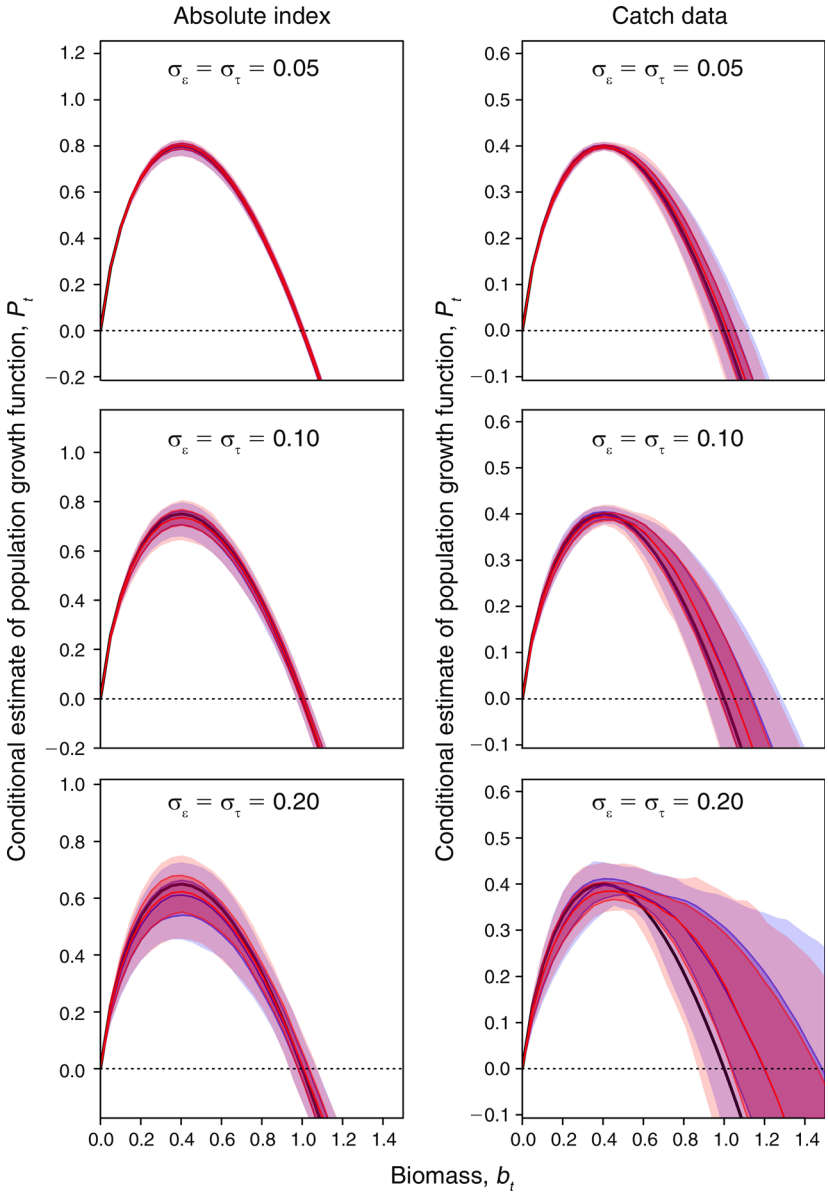

FIG. 2. The true population growth function (black) and conventional (blue) and Gaussian process (red) estimates, showing the median (solid line), 25–75% quantiles (heavy shading), and 5–95% quantiles (light shading) of the posterior medians for each simulation replicate when correctly specifying the growth function, and either the absolute index (left column) or catch data scenarios (right column). Variability in total process error σ_ε and observation error σ_τ is low (0.05), medium (0.10), or high (0.20).

ε. This pattern of harvest mortality and abundance was stipulated because population dynamics models with a relative index require either a two-way trip in abundance or an informative prior on maximum population growth or the detectability coefficient to estimate the absolute scale of the population. We opted for simplicity to forgo any exploration of informative priors, while noting that informative priors for population growth and detectability can be feasibly generated (McAllister et al. 2001, 2010). Both the conventional and GP models were then fitted to simulated data for each replicate.

RESULTS

Objective 1: Updating the parametric "prior" for the population growth function

We first demonstrate the performance of conventional and GP models at estimating the population growth function when it is correctly specified (i.e., it matches the true data-generating process). If detectability is known (Fig. 2, left column), the conventional and GP models have a median estimate of the population growth function that perfectly matches the true function for

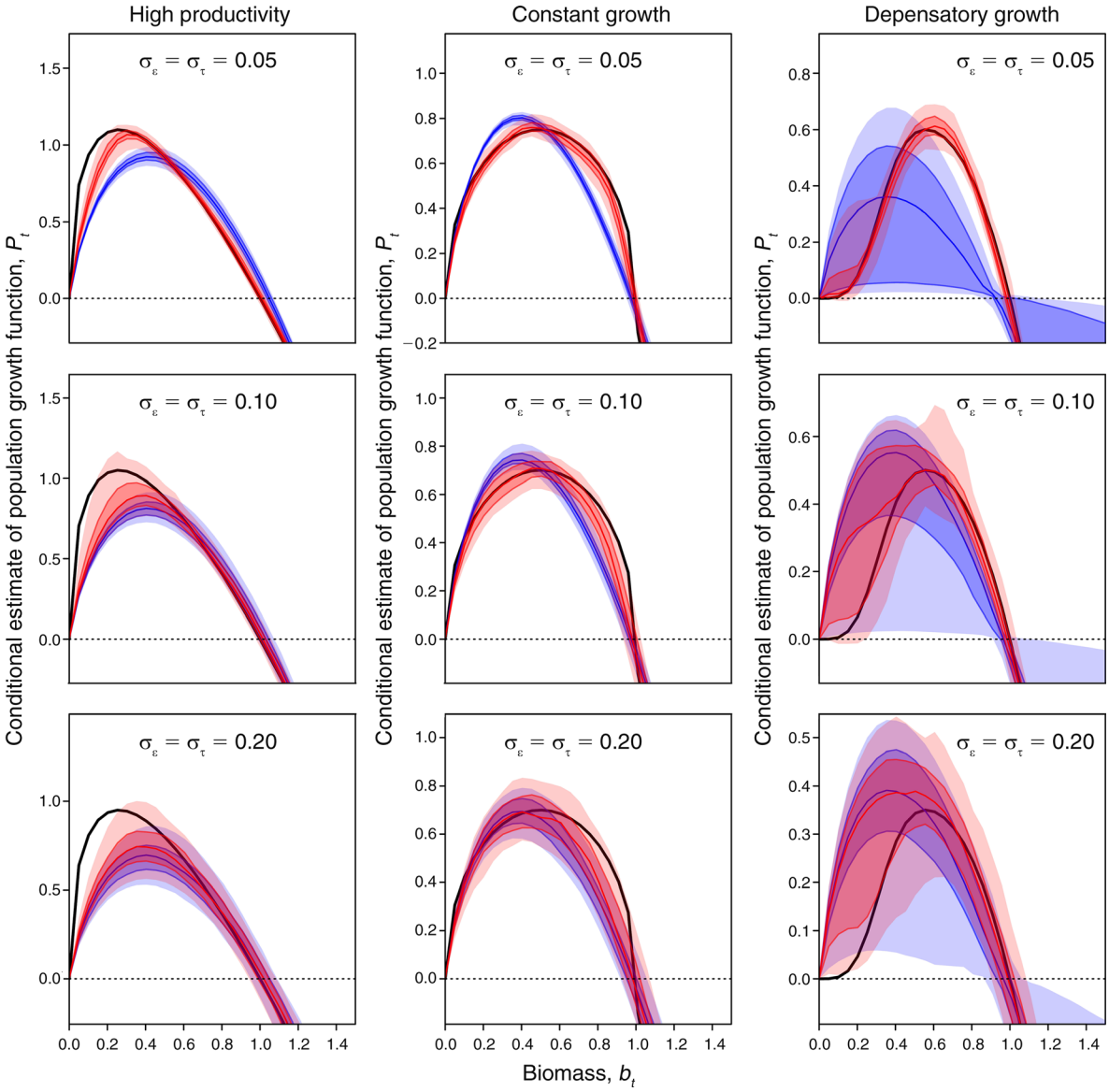

FIG. 3. The true population growth function (black) and conventional (blue) and Gaussian process (red) estimates, showing the median (solid line), 25–75% quantiles (heavy shading), and 5–95% quantiles (light shading) for each simulation replicate for the absolute index scenario (i.e., assuming that the detectability coefficient $q = 1$) given three possible mismatches for the growth function (high productivity, stable or constant growth, and depensatory growth). Variability in σ_ε and σ_τ is low (0.05), medium (0.10), or high (0.20).

all levels of error. This is as expected, given that the conventional model and the GP prior mean perfectly match the function used to simulate the data. In this case of known detectability, increases in error increase the width of the 90% simulation interval for the growth function, but there is no visible difference in performance between conventional and GP models. When detectability is freely estimated (Fig. 2, right column), conventional and GP still perform identically in their estimates of the growth function. However, both models have a positive bias in K estimates (i.e., the estimated population growth function is stretched along the x-axis) for moderate ($\sigma_\varepsilon = \sigma_\tau = 0.10$) and high ($\sigma_\varepsilon = \sigma_\tau = 0.20$) variability. The performance of the credible interval estimates follows a similar pattern for conventional and GP models (Appendix C).

We next demonstrate the performance of conventional and GP models in the more interesting cases, i.e., when the assumed population growth function differs from the true data-generating process. When detectability is known (Fig. 3), this model misspecification for the conventional model obviously causes the estimated growth function to differ from the "true" model. By contrast, the GP model adapts to approximate the true

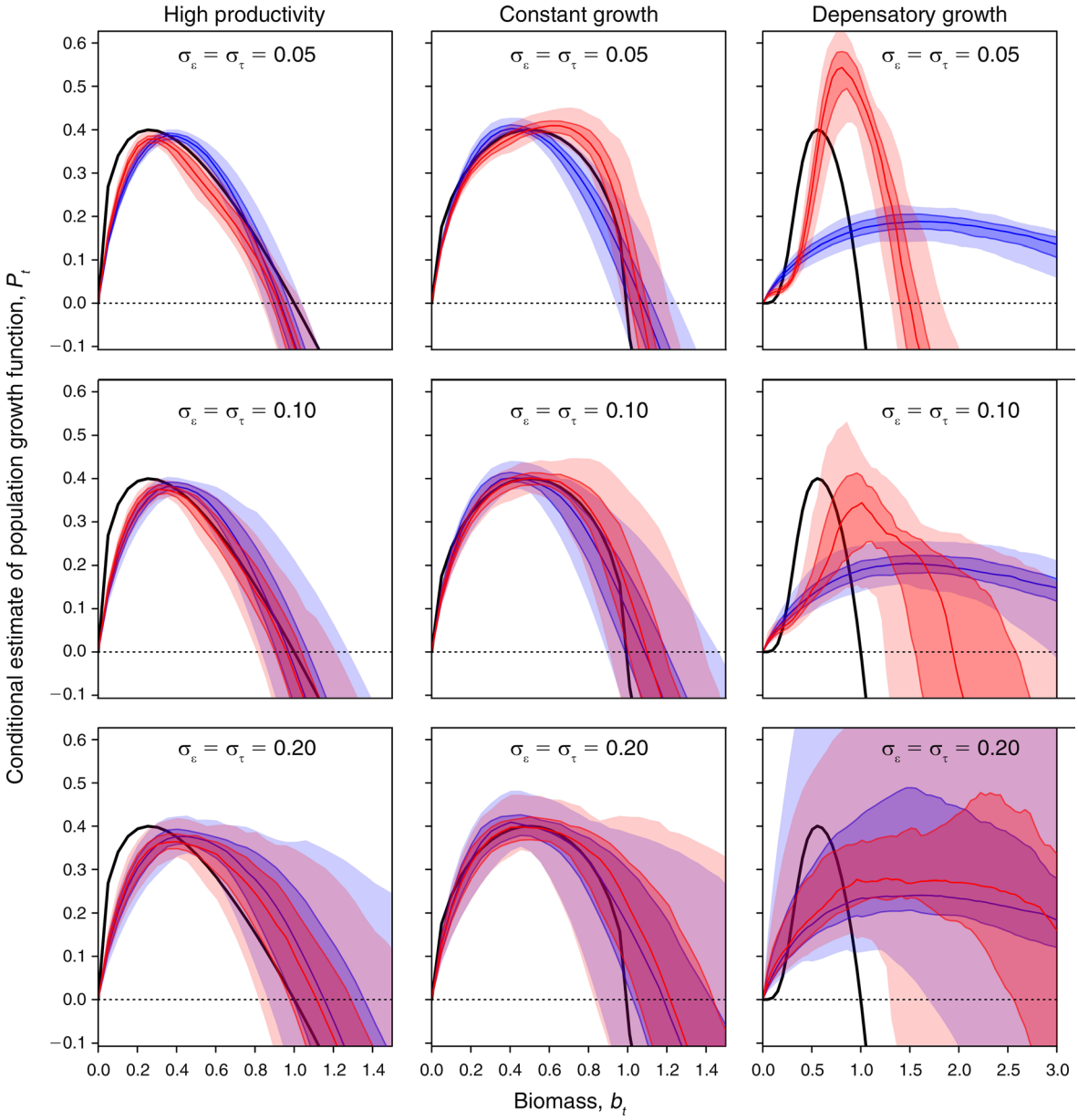

FIG. 4. The true population growth function (black) and conventional (blue) and Gaussian process (red) estimates, showing the median estimate across simulation replicates (solid line), 25–75% quantiles (heavy shading), and 5–95% quantiles (light shading) for the catch data scenario (i.e., estimating q) given three possible mismatches for the growth function and low, medium, or high variability in error, as in Fig. 3.

function. For example, the GP model is able to estimate the left-skew of the "high productivity" and the depensation of the "depensatory" scenarios, given low errors. However, the GP model more closely resembles the conventional model when errors are high. When detectability is estimated (Fig. 4), the model misspecification has a larger and more interesting effect on model performance. This is easiest seen with the depensatory model and low errors. In this case, the conventional model has a very large and positive (+200–400%) bias in K estimates, and an associated negative bias in detectability estimates. This occurs because the misspecified parametric model cannot account for the observed population dynamics. As a consequence, all abundance changes are attributed to "process errors" while specifying that harvest has little effect on abundance (i.e., that total abundance is much large than catches). The depensatory and low-error scenario also results in smaller but positive biases in K and negative biases in detectability for the GP model. This was unexpected and occurs because the "prior" for the population growth function still exerts an influence over

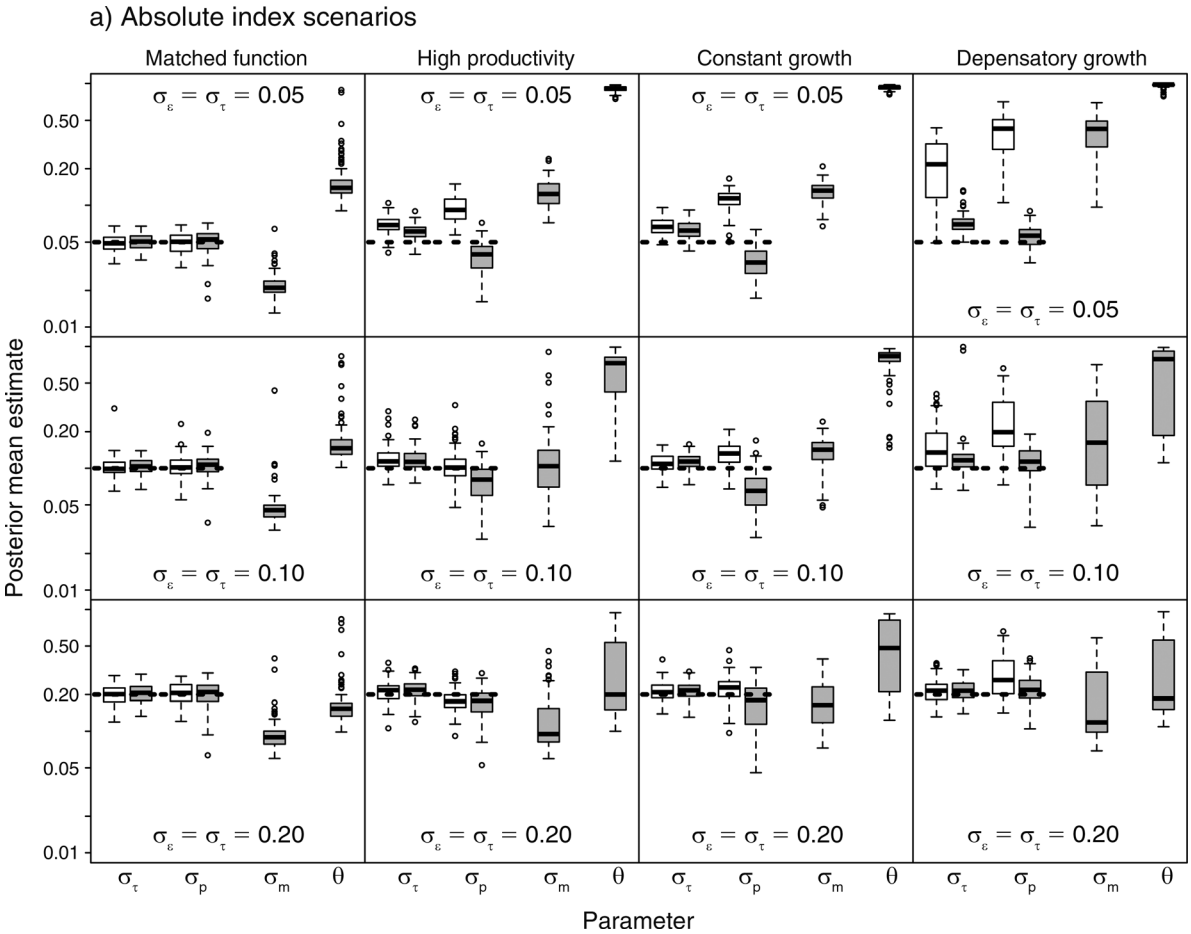

FIG. 5. Boxplots (middle line, median estimate from simulated replicates; box, interquartile range from 25th to 75th percentile; whiskers, furthest observation within 1.5 times the interquartile range from the median estimate; dots, estimates outside of the whiskers) showing the posterior mean for model parameters (σ_τ, measurement errors; σ_p, process errors; σ_m, model misspecification errors; θ, model error parameter) for conventional (white) and Gaussian process (gray) models, with the true value indicated (heavy dashed black line), for the absolute index scenarios (panel a) or the catch data scenarios (panel b), where the conventional model does not provide an estimate of σ_m and θ and therefore these boxes are missing for the conventional model. Variability in σ_ε and σ_τ is low (0.05), medium (0.10), or high (0.20).

the GP model, and is miscalibrated with regard to the "true" process. To accommodate this miscalibrated prior on the population growth function, the model systematically overestimates biomass and population growth, and hence effectively stretches the estimated population growth function by a fixed amount in both the *x*- and *y*-axes. Nevertheless, the GP performs significantly better at estimating the population growth function, and results in only a +50% bias in *K* while largely replicating the "true" pattern of depensatory population growth. As errors increase in magnitude, the difference between conventional and GP models decreases, until they have very similar performance for the depensatory function given high errors. A similar pattern is seen in the other growth functions, where the GP is able to replicate the general form of the growth function given low errors, but barely differs from the conventional model given large errors (i.e., when data are less informative), and the credible interval estimates also follow a similar pattern (Appendix C). Finally, a comparison of prior and posterior values for variance parameters shows the influence of the informative joint prior on process and measurement error variances (Appendix D).

Objective 2: Estimating the relative magnitude of measurement, process, and model errors

We also seek to demonstrate the capacity of the GP to distinguish between "process error" and "model errors" (as defined in the introduction). Posterior mean estimates for the "absolute index" scenario (Fig. 5a) show that the GP model is able to accurately estimate the true value of measurement and process errors in all configurations. Model errors (σ_m) have lower magnitude than process errors when the population growth function is correctly specified (left column), but have

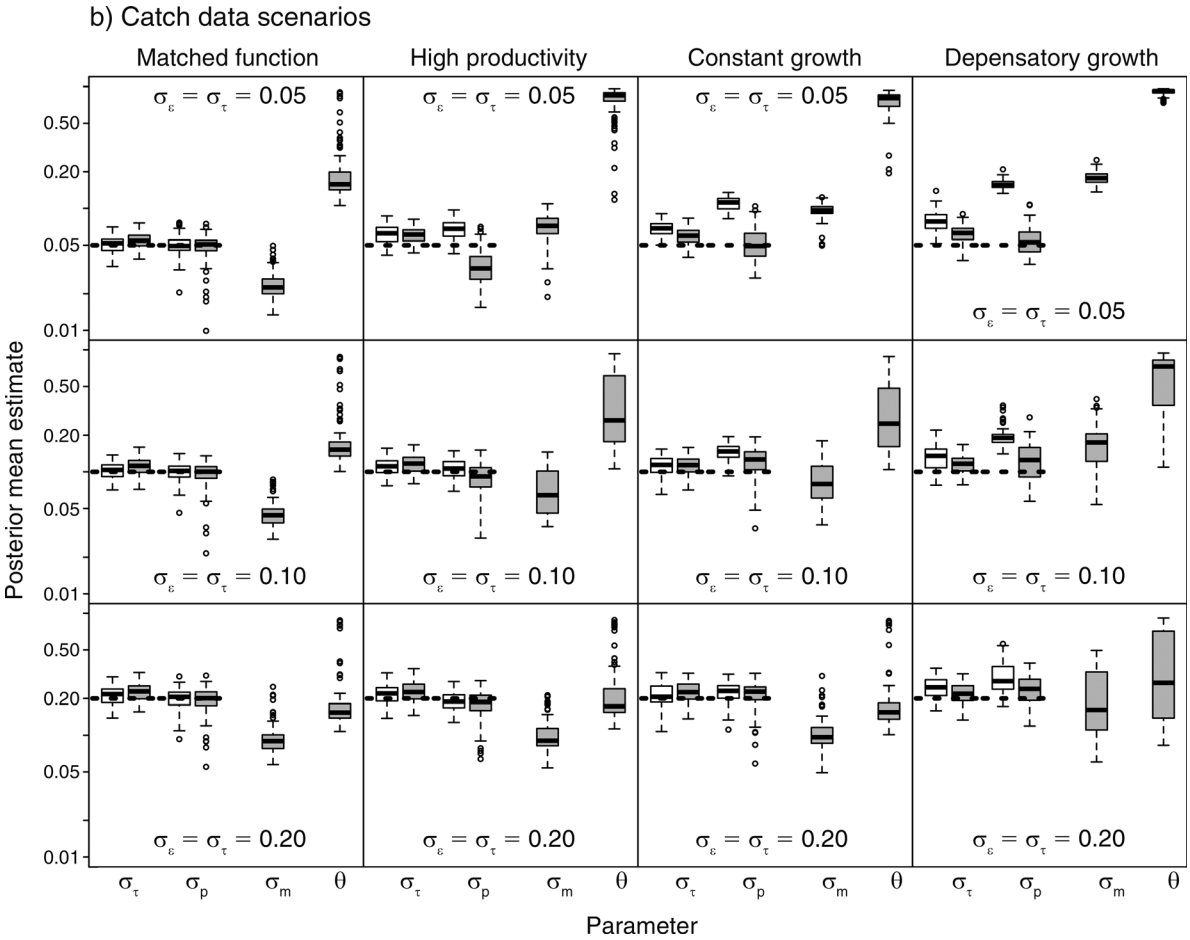

FIG. 5. Continued.

greater magnitude than process errors when the model is misspecified. Model errors have particularly large magnitude, given the depensatory model (right-most column of Fig. 5a), when model misspecification also has the greatest impact on estimates of the growth function. The model error parameter (θ) is generally close to its prior mean for the correctly specified model and is substantially greater for other models. Posterior mean estimates for the "catch data" scenario (Fig. 5b) show the same patterns of measurement, process, and model error estimates, although the model error parameter θ is closer to its prior mean (i.e., less informed by the data) for the cases with high errors.

DISCUSSION

In this simulation evaluation study, Gaussian process state-space models were robust to model misspecification and performed as well or better than a conventional state-space model. Their performance was especially good when the levels of errors were low and when a time series of absolute abundance was available. However, the GP model was not free of bias, which increased with the level of model misspecification and magnitude of variability in the data. Nevertheless, the GP method provided accurate estimates of both process and measurement errors that could help users to infer the potential bias in their estimate. Additionally, the GP approach was able to evaluate the degree of mismatch between the assumed process model and the observed data by estimating the model misspecification error σ_m.

Although appealing in terms of performance, the GP approach is more computationally demanding than the state-space model. It both takes longer to run (due to the matrix inversion necessary in Eq. 8) and cannot be readily implemented in free software such as WinBugs or JAGS (because the model cannot easily be implemented as an acyclic graph). However, the Gaussian process method can be extended to virtually any stochastic process (i.e., where variability in the process is estimated) if one can deal with these computational difficulties, and the number of possible applications is virtually infinite. We discuss for illustrative purposes three potential applications in the following section.

As a first example, state-space estimation is used for age- and/or length-structured population models (Gudmundsson and Gunnlaugsson 2012). In these cases,

Gaussian process methods could be fruitfully applied to the function representing density-dependent changes in early juvenile survival, which typically drives compensatory dynamics in these models. As one example, recruitment in common fishery models (e.g., Methot and Wetzel 2013) is approximated using deviations away from the function representing early juvenile survival. These recruitment deviations are generally stipulated to be independent and identically distributed, but could also be distributed according to a Gaussian process, in which predictions of future recruitment would deviate from the specified prior through their correlation with previously observed recruitment (Munch et al. 2005).

Gaussian process methods could also improve the many meta-analytic studies that are routinely done. For example, meta-analysis has previously been applied to estimating the presence or absence of depensation in stock–recruit functions (Liermann and Hilborn 1997) and the ratio between management targets and life history parameters (Zhou et al. 2012). Such meta-analysis may be more robust if conclusions are less influenced by whatever parametric form is assumed a priori, i.e., using a Gaussian process (Sugeno and Munch 2013).

Finally, recent research has extended state-space estimation methods to animal movement models (Jonsen et al. 2005, Patterson et al. 2008), while explicitly accounting for individual-level differences. Covariates are typically included using a linear function predicting spatial behaviors. The functional form for covariates in movement models could alternatively follow a Gaussian process, and our results imply that failing to account for model misspecification (e.g., in the link between covariates and behaviors) may have an important and complicated effect on inference for seemingly unrelated model parameters.

We therefore propose that Gaussian process models provide a conceptually unified and computationally feasible approach to semiparametric models in general, and have the potential to reduce bias in parameter estimates for population models. Furthermore, we have shown that Gaussian process models can diagnose the degree of match between an assumed population model and the observed data. For all of these reasons, we believe that these methods can improve many types of ecological models across a range of data availability or parametric structure.

Acknowledgments

This work was supported in part by the "Modeling species response to environmental change: development of integrated, scalable Bayesian models of population persistence" working group supported by the John Wesley Powell Center for Analysis and Synthesis, funded by the U.S. Geological Survey, which J. Thorson and S. Munch attended. K. Ono acknowledges funding from NMFS grant NA10OAR4320148. The manuscript was improved by editorial suggestions by J. Samhouri and two anonymous reviewers.

Literature Cited

Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2003. Hierarchical modeling and analysis for spatial data. First edition. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Calder, C., M. Lavine, P. Müller, and J. S. Clark. 2003. Incorporating multiple sources of stochasticity into dynamic population models. Ecology 84:1395–1402.

de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs 72:57–76.

Gao, J. 2007. Nonlinear time series: semiparametric and nonparametric methods. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Analysis 1:515–533.

Ginzburg, L. R., S. Ferson, and H. R. Akçakaya. 2005. Reconstructibility of density dependence and the conservative assessment of extinction risks. Conservation Biology 4: 63–70.

Gudmundsson, G., and T. Gunnlaugsson. 2012. Selection and estimation of sequential catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences 69:1760–1772.

Harwood, J., and K. Stokes. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends in Ecology and Evolution 18:617–622.

Hilborn, R., and C. J. Walters. 1991. Quantitative fisheries stock assessment—Choice, dynamics and uncertainty. First edition. Springer, Norwell, Massachusetts, USA.

Jaynes, E. T. 1968. Prior probabilities. IEEE Transactions on Systems Science and Cybernetics 4:227–241.

Jonsen, I. D., J. M. Flemming, and R. A. Myers. 2005. Robust state-space modeling of animal movement data. Ecology 86: 2874–2880.

Jost, C., and S. P. Ellner. 2000. Testing for predator dependence in predator–prey dynamics: a non-parametric approach. Proceedings of the Royal Society B 267:1611–1620.

Knape, J., and P. de Valpine. 2012. Fitting complex population models by combining particle filters with Markov chain Monte Carlo. Ecology 93:256–263.

Lande, R., S. Engen, and B.-E. Sæther. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, New York, New York, USA.

Liermann, M., and R. Hilborn. 1997. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. Canadian Journal of Fisheries and Aquatic Sciences 54:1976–1984.

McAllister, M. K., E. K. Pikitch, and E. A. Babcock. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Canadian Journal of Fisheries and Aquatic Sciences 58:1871–1890.

McAllister, M. K., R. D. Stanley, and P. Starr. 2010. Using experiments and expert judgment to model catchability of Pacific rockfishes in trawl surveys, with application to bocaccio (*Sebastes paucispinis*) off British Columbia. Fishery Bulletin 108:282–304.

Methot, R. D. J., and C. R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142:86–99.

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.

Ono, K., A. E. Punt, and E. Rivot. 2012. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fisheries Research 125–126: 173–183.

Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space models of individual animal movement. Trends in Ecology and Evolution 23:87–94.

Punt, A. E. 1992. Selecting management methodologies for marine resources, with an illustration for southern African hake. South African Journal of Marine Science 12:943–958.

Punt, A. E. 2003. Extending production models to include process error in the population dynamics. Canadian Journal of Fisheries and Aquatic Sciences 60:1217–1228.

Quinn, T. J., and R. B. Deriso. 1999. Quantitative fish dynamics. First edition. Oxford University Press, Oxford, UK.

Rasmussen, C. E., and C. K. I. Williams. 2006. Gaussian processes for machine learning. First edition. MIT Press, Cambridge, Massachusetts, USA.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. First edition. Academic Press, London, UK.

Sigourney, D. B., S. B. Munch, and B. H. Letcher. 2012. Combining a Bayesian nonparametric method with a hierarchical framework to estimate individual and temporal variation in growth. Ecological Modelling 247:125–134.

Smallegange, I. M., and T. Coulson. 2012. Towards a general, population-level understanding of eco-evolutionary change. Trends in Ecology and Evolution 28:143–148.

Staples, D. F., M. L. Taper, and B. Dennis. 2004. Estimating population trend and process variation for PVA in the presence of sampling error. Ecology 85:923–929.

Sugeno, M., and S. B. Munch. 2013. A semiparametric Bayesian approach to estimating maximum reproductive rates at low population sizes. Ecological Applications 23: 699–709.

Thorson, J. T., J. Cope, T. Branch, and O. Jensen. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences 69: 1556–1568.

Thorson, J. T., and E. Ward. 2013. Accounting for space–time interactions in index standardization models. Fisheries Research 147:426–433.

Walters, C., and J. F. Kitchell. 2001. Cultivation/depensation effects on juvenile survival and recruitment: implications for the theory of fishing. Canadian Journal of Fisheries and Aquatic Sciences 58:39–50.

West, M., and J. Harrison. 1997. Bayesian forecasting and dynamic models. Springer, New York, New York, USA.

Wood, S. N., and N. H. Augustin. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecological Modelling 157:157–177.

Zhou, S., S. Yin, J. Thorson, T. Smith, and M. Fuller. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Canadian Journal of Fisheries and Aquatic Sciences 69:1292–1301.

Supplemental Material

Appendix A

Markov chain Monte Carlo methods (*Ecological Archives* E095-029-A1).

Appendix B

Abundance trajectories for simulated data (*Ecological Archives* E095-029-A2).

Appendix C

Credible interval estimates (*Ecological Archives* E095-029-A3).

Appendix D

Priors and posteriors for variance parameters (*Ecological Archives* E095-029-A4).