Ecological Applications, 14(4), 2004, pp. 1272–1293

BEYOND THEORY TO APPLICATION AND EVALUATION: DIFFUSION APPROXIMATIONS FOR POPULATION VIABILITY ANALYSIS

ELIZABETH E. HOLMES[1]

National Marine Fisheries Service, Northwest Fisheries Science Center, 2725 Montlake Blvd. E., Seattle, Washington 98112 USA

Abstract. Census data on endangered species are often plagued by problems that make quantitative population viability analysis (PVA) a challenge. This paper addresses four such problems: sampling error, density dependence, nonstable age structure, and population supplementation that masks the true population status. Estimating trends and extinction risks using such corrupted data presents serious parameter estimation difficulties. Here I review diffusion approximation (DA) methods for estimating population status and risks from time series data. A variety of parameterization methods are available for DA models; some correct for data corruption and others do not. I illustrate how stochastic Leslie matrix models can be used to evaluate the performance of a proposed DA model and to select among different DA parameterization methods for a given application. Presenting the uncertainty in estimated risks is critical, especially when the data are highly corrupted and estimated parameters are more uncertain. Using a Bayesian approach, I demonstrate how the level of data support for different risk levels can be calculated using DA parameter likelihood functions.

Key words: Dennis model; diffusion approximation; extinction; matrix models; population models; population viability analysis; process error; salmon; stochasticity.

INTRODUCTION

Limited data are a common stumbling block for quantitative population viability analysis (PVA). Sufficient data for developing detailed life-history models are often unavailable; indeed often the only available or planned data are a simple time series of counts (Morris et al. 2002). In the last decade, diffusion approximation (DA) methods have been developed that use count data alone for the estimation of PVA risk metrics, such as the probability of crossing extinction thresholds, mean passage times, and average long-term rates of population growth or decline (Lande and Orzack 1988, Dennis et al. 1991). These methods have since been used to estimate extinction risks for numerous species of conservation concern (Dennis et al. 1991, Nicholls et al. 1996, Gerber et al. 1999, Morris et al. 1999, McClure et al. 2003). The appeal of DA methods from an applied standpoint is their simplicity and their reliance on simple census data alone, and they have become one of the basic quantitative tools presented in recent books on PVA methods (Morris and Doak 2002, Lande et al. 2003).

Diffusion approximation methods stem from theory concerning the behavior of stochastic age-structured population models with no density dependence:

$$\begin{pmatrix} n_{1,t+1} \\ n_{2,t+1} \\ n_{3,t+1} \\ \vdots \\ n_{k,t+1} \end{pmatrix} = \mathbf{A}_t \begin{pmatrix} n_{1,t} \\ n_{2,t} \\ n_{3,t} \\ \vdots \\ n_{k,t} \end{pmatrix} \quad (1)$$

where $\mathbf{A}_t$ is a stochastic population transition matrix, e.g., a Leslie matrix, for time t. The $\mathbf{A}_t$'s are assumed to be drawn from some unspecified stationary statistical distribution, although they need not be drawn randomly. For example, $\mathbf{A}_{t+1}$ might be drawn conditioned on $\mathbf{A}_t$ if there is temporal correlation in year-to-year environmental conditions (i.e., good years follow good years). For such models, the asymptotic behavior of any weighted census of the population, $N_t = \Sigma_i\, w_i n_{i,t}$, is a stochastic exponential process (Tuljapurkar and Orzack 1980, Tuljapurkar 1989):

$$N_t = N_0 \exp(\mu t + \varepsilon_t) \quad (2)$$

where ε_t is distributed normally with mean 0 and variance $\sigma^2 t$ for t big. The statistical distribution of future population sizes, here $\log N_t/N_0$, is distributed normally with mean μt and variance $\sigma^2 t$, for t big. The weights w_i can be equal to 1 so that N_t is simply the total number of individuals alive at time t, but other weightings, such w_i equals the fraction of age i individuals that eventually become reproductive, may be more appropriate when calculating extinction metrics. The parameter μ in Eq. 2 determines the rate at which the median log population size, $\log(N_t/N_0)$, increases through time, while σ^2 determines the rate at which the distribution spreads, or in other words, the variability of potential

Manuscript received 13 March 2002; revised 1 August 2003; accepted 23 September 2003; final version received 18 December 2003. Corresponding Editor: L. B. Crowder.

[1] E-mail: oeli.holmes@noaa.gov

population sizes at time t. Throughout this paper, σ^2 will be referred to as process error. See Caswell (2001: section 14.3.3) for a review of the theory of products of random matrices in the context of stochastic age-structured models, and Morris and Doak (2002: chapter 3) for a discussion of stochastic exponential models.

Diffusion approximation methods assume that Eq. 2 holds for all $t > 0$ including small t and that the ε_t are independently and identically distributed. This allows one to model the population as a diffusion process (Lande and Orzack 1988):

$$\partial p/\partial t = -\mu\partial p/\partial x + (\sigma^2/2)\partial^2 p/\partial x^2 \quad (3)$$

where p is the probability density of $(N_t/N_0) = x$. The diffusion model has the property that $\log N_t/N_0$ is distributed normally with mean μt and variance $\sigma^2 t$, like the stochastic exponential process it is used to approximate. See Dennis et al. (1991) for a fuller discussion of the diffusion approximation.

This approximation opens a toolbox of parameterization methods for linear models with normal error. It also provides analytical estimates of passage probabilities, i.e., the probability of crossing a particular threshold within a given time frame, for example extinction estimates (to one individual) or quasi-extinction estimates to some critical population size >1. Strictly speaking, however, an age-structured population process is not a diffusion process. For a time series from an age-structured population, $\log N_{t+\tau}/N_t$ where $\tau = 1$ is not normally distributed; only $\log N_{t+\tau}/N_t$ where τ is big, is normally distributed. Also in an age-structured population time series, the temporal independence assumptions at $\tau = 1$ are violated since the ratios, $\log N_{t+1}/N_t$ and $\log N_{t+2}/N_{t+1}$, will be correlated, even if the environment is uncorrelated.

Despite these assumption violations, the diffusion model approximates many types of stochastic age-structured population processes, as seen both from simulated and real data (Lande and Orzack 1988, Dennis et al. 1991, Holmes and Fagan 2002; W. F. Fagan, J. Rango, A. Folarin, J. Sorensen, J. Lippe, and N. E. McIntyre, *unpublished manuscript*). There are, however, well-known cases where the diffusion approximation performs poorly, such as when year-to-year variability is high and populations are small (Ludwig 1996*b*, 1999), when demographic stochasticity introduces strong nonlinearities (Wilcox and Possingham 2002), and when density dependence is extreme (Sabo et al. 2004). Thus, it is critical to carefully evaluate the appropriateness of the diffusion model given the life history of the species at hand and given the types of extinction or quasi-extinction metrics one would like to estimate. In addition, careful selection of parameterization methods for estimating the DA parameters is essential since many real population time series will contain extraneous variability whether due to age-structure cycles, density-dependent feedback, or sampling error, and this will confound proper parameterization. In this paper, I discuss and illustrate the application of DA methods with the biologist faced with assessing risks for a real population in mind. I discuss the available parameterization methods and present a quantitative approach for assessing the performance of the diffusion approximation and its parameterization.

Once a parameterization method is chosen, evaluation of which risk metrics to use is needed. Some metrics will be robust and others may be useful only over certain time frames. In particular, estimates of the probability of crossing population thresholds (e.g., extinction metrics) can be extremely variable for certain parameter ranges (Fieberg and Ellner 2000, Ellner et al. 2002). This is not a problem specific to the DA approach; it is a feature of stochastic population processes in certain parameter ranges. Nonetheless, conservation biologists are often asked to make an assessment of extinction or quasi-extinction risk; indeed, such an assessment may be mandated by law when working with species being considered under the U.S. Endangered Species Act. One of the positive features of DA methods is that uncertainty is well characterized and can be estimated from the data. Thus, while we may be uncertain about the point estimates of the risk metrics, we can be precise about the level of uncertainty. In this paper, I discuss two different methods for representing uncertainty: confidence intervals and posterior probability distributions.

These methods for evaluating and presenting DA risk metrics are illustrated using three simulated populations of different salmonid species. These concrete examples are intended to illustrate analyses and models that can be used to assess the performance of a proposed DA approach given the life histories and data corruption one expects in a particular PVA application. The paper is divided into four steps, reflecting the steps of evaluating a proposed DA application: (1) evaluating whether a diffusion approximation is reasonable for the population at hand, (2) evaluating and selecting a parameterization method, (3) evaluating the sensitivity of proposed risk metrics to parameterization errors, and (4) estimating parameters and risk metrics given a census time series from the concerned population and presenting the uncertainty in these estimates.

Evaluating the Diffusion Approximation for the Population

Simulated data

The first step in evaluating a proposed DA application is to develop a model that captures the basic life history and that has realistic environmental variability, including the temporal correlation in the environment. Although a precise model for the population will likely not be available, given that diffusion approximations are typically used where data are limited, often enough natural history information will be available to roughly parameterize a model for the purpose of evaluating the

performance of proposed methods. Such an evaluation is illustrated here using stochastic matrix models for Snake River spring/summer chinook, Snake River fall chinook, and upper Columbia River steelhead, which are currently listed under the U.S. Endangered Species Act. Although this paper uses matrix models, the simulated data could be produced by any variety of models or by using a bootstrapping method to simulate from actual time series data, e.g., Hinrichsen (2002).

The basic model used here tracks individuals from age 1 to spawners:

$$\begin{pmatrix} \text{age 1} \\ \text{age 2} \\ \text{age 3} \\ \vdots \\ \text{spawners} \end{pmatrix}_{t+1} = \mathbf{A}_t \begin{pmatrix} \text{age 1} \\ \text{age 2} \\ \text{age 3} \\ \vdots \\ \text{spawners} \end{pmatrix}_{t} \quad (4)$$

where $\mathbf{A}_t$ is the Leslie matrix model describing the transition probabilities from age i to $i + 1$. The stochastic matrices used for the three salmonid populations are given in Table 1, along with parameter definitions and estimates.

Stochasticity was added to survivorships and fecundity by multiplying the survivorship and fecundity terms by a random variable ε_i with the maximum possible survivorship capped at 1.0. The ε_1 term represents variability in spawner to age 1 ratios. For all matrices, ε_1 variables were drawn from a lognormal distribution specified by exp(normal(mean = 0, variance = 0.13)). The variance was estimated from spawner to smolt (~age 1.5) data for spring/summer chinook in the Snake River (P. Levin, *unpublished data*). The ε_2 variables associated with age 1 to age 2 survivorship for spring/summer chinook and steelhead were drawn from a lognormal distribution specified by exp(normal(mean = 0, variance = 0.08)). The variance was estimated from parr to smolt (~age 0.5 to age 1.5) survivorship data for spring/summer chinook in the Snake River (Achord et al. 2003). The ε_1 and ε_2 variables were set to covary slightly (correlation = 0.2) since both are associated with basin conditions, although one is associated more with spawning ground conditions and the other is associated more with downstream migration. The ε_o variables associated with ocean survivorship were drawn from a lognormal distribution specified by exp(normal(mean = 0, variance = 0.02)). The variance for ε_o was chosen such that the resulting σ^2 estimates at the population level reflected approximately that estimated from actual time series. The ε_o variables associated with different ages were set to covary strongly (correlation coefficient = 0.8) since they are all associated with ocean conditions.

In addition, studies of ocean condition indices and studies of cohort (i.e., all young-of-year t) survivorship data for Columbia River salmonids suggest that there is also strong year-to-year correlation in ocean survivorship, i.e., good ocean survivorship one year tends to be followed by good ocean survivorship the next year. To model this, the mean ocean survivorship across ages was chosen such that it covaried strongly (correlation coefficient = 0.7) with the mean ocean survivorship in the previous year. The level of correlation was chosen from the levels seen in time series of cohort survivorship data. Temporally correlated juvenile survivorship (in-stream survivorship) was not included since evidence of this was not seen in the study by Achord et al. (2003).

Although current densities of salmon in these three populations are well below historical levels, density dependence in parr to smolt survivorship has been shown from studies on Snake River spring/summer chinook (Achord et al. 2003). Achord et al. postulate that juveniles continue to show density dependence at low density because of nutrient limitation from the associated low densities of spawner carcasses. Density dependence was incorporated into the egg to age-1 survivorship term (s_1) by fitting a linear model to the juvenile survivorship data in Achord et al. (their Fig. 3), giving the following:

$$\tilde{s}_{1,t} = 0.7393 \times s_1 - 0.2607 \times s_1 \times \left(N_{s,t-1} \Big/ \frac{1}{10} \sum_{j=t-1}^{t-11} N_{s,j} \right) \quad (5)$$

where s_1 is the mean egg to age-1 survivorship in Table 1 and $N_{s,t}$ is the number of spawners in year t. The minimum $\tilde{s}_{1,t}$ was set at $0.39s_1$ and the maximum at $1.17s_1$. The effect of Eq. 5 was an increase in $\tilde{s}_{1,t}$ of up to 17% over s_1 when spawner density was lower than the previous 10-year mean spawner density and up to 61% lower than s_1 survivorship when spawner density was higher than the 10-year average.

Using the stochastic matrices and Eq. 4, 1000 population time series were generated for each species. The time series were started from a stable age structure drawn at random from the stochastic distribution of stable age structures. The distribution of $\log N_t/N_0$ from the 1000 time series describes the statistical distribution of population size at time t.

Performance of the diffusion approximation for simulated salmonid time series

The goal was to determine whether a diffusion model exists that describes the future distribution and quasi-extinction risks for the number of current and future spawners within a salmon population. A weighted total population at time t was defined as $N_t = \Sigma_i w_i n_{i,t}$ where w_i equals the average fraction of age i fish that will eventually become spawners. This was used instead of simply $N_t = \Sigma_i n_{i,t}$ since the age structure is heavily dominated by age-1 individuals due to low age-1 to age-2 survivorship and using $\Sigma_i n_{i,t}$ would mean effectively tracking the dynamics of age-1 individuals alone

TABLE 1. Matrix models and parameter estimates.

A) Matrix models

Snake River spring/summer chinook (adapted from Kareiva et al. [2000] and modified to separate out the spawner class):

$$\mathbf{A}_t = \begin{bmatrix} 0 & 0 & 0 & 0 & p\tilde{s}_{1,t}m\varepsilon_1 \\ s_2\varepsilon_2 & 0 & 0 & 0 & 0 \\ 0 & s_o\varepsilon_o & 0 & 0 & 0 \\ 0 & 0 & (1-b_4)s_o\varepsilon_o & 0 & 0 \\ 0 & 0 & b_4s_o\varepsilon_o & b_5s_o\varepsilon_o & 0 \end{bmatrix}.$$

Snake River fall chinook (adapted from Kareiva et al. [2000]):

$$\mathbf{A}_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & p\tilde{s}_{1,t}m\varepsilon_1 \\ (1-h_2)s_o\varepsilon_o & 0 & 0 & 0 & 0 & 0 \\ 0 & (1-h_3)s_o(1-b_3)\varepsilon_o & 0 & 0 & 0 & 0 \\ 0 & 0 & (1-h_4)s_o(1-b_4)\varepsilon_o & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-h_5)s_o(1-b_5)\varepsilon_o & 0 & 0 \\ 0 & (1-h_3)s_ob_3\varepsilon_o & (1-h_4)s_ob_4\varepsilon_o & (1-h_5)s_ob_5\varepsilon_o & (1-h_6)s_o\varepsilon_o & 0 \end{bmatrix}.$$

Upper Columbia steelhead:

$$\mathbf{A}_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & p\tilde{s}_{1,t}m\varepsilon_1 \\ s_2(1-b_2)\varepsilon_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_o(1-b_3)\varepsilon_o & 0 & 0 & 0 & 0 \\ 0 & 0 & s_o(1-b_4)\varepsilon_o & 0 & 0 & 0 \\ 0 & 0 & 0 & s_o(1-b_5)\varepsilon_o & 0 & 0 \\ s_2b_2\varepsilon_2 & s_ob_3\varepsilon_o & s_ob_4\varepsilon_o & s_ob_5\varepsilon_o & s_o\varepsilon_o & 0 \end{bmatrix}.$$

B) Parameter estimates

Parameter	Spring/summer chinook	Fall chinook	Steelhead
p, mouth to spawning ground survival†	0.4815	0.36	0.58
s_1, egg to age-1 survival	0.018	0.0044167	0.047
s_2, mean age-1 to age-2 survival in-stream	0.044	NA	0.009
s_o, ocean survivorship	0.8	0.8	0.8
b_2, fraction of age 2 that spawn	0	0	0.009
b_3, fraction of age 3 that spawn	0	0.081	0.333
b_4, fraction of age 4 that spawn	0.216	0.65	0.693
b_5, fraction of age 5 that spawn	1.0	0.863	0.923
b_6, fraction of age 6 that spawn	NA	1.0	1.0
h_2, in ocean harvest of age 2	NA	0.0123	NA
h_3, in ocean harvest of age 3	NA	0.0465	NA
h_4, in ocean harvest of age 4	NA	0.1368	NA
h_5, in ocean harvest of age 5	NA	0.1838	NA
h_6, in ocean harvest of age 6	NA	0.1953	NA
m, average female eggs per female spawner	2747	1500	2500

Notes: The parameter estimates for both chinook models are based on estimates for Marsh Creek (1980–1998; M. McClure, *unpublished data*). Parameter estimates for steelhead were based on estimates for Methow River steelhead for pre-1998 conditions (T. Cooney, *unpublished data*).

† The parameter p is the fraction that survive harvest during upstream migration × fraction that survive upstream migration to spawning ground × survival on the spawning ground.

rather than the individuals with reproductive potential across all ages.

First the predicted vs. observed distribution of log N_t/N_0 was tested. The predicted distribution is normal with a mean and variance that increase linearly with time, such that $(1/t)$ times the mean and variance of log N_t/N_0 equal constants. Fig. 1(a–c) shows the mean and variance of log N_t/N_0, divided by t for the simulated salmon populations. This figure illustrates that the mean and variance increased linearly with t as predicted by theory. The mean and variance of log N_t/N_0 divided by t reached a constant after ~5 years (Fig. 1a–c). The

distribution of log N_t/N_0 approached normality on the same time frame, as determined by the difference between the cumulative distribution of log N_t/N_0 vs. the cumulative distribution of a normal with mean μt and variance $\sigma^2 t$ (not shown). This evaluation indicates that for time periods past five years in the future, there exists a diffusion model that properly describes the statistical distribution of population sizes at time $t > 5$. The parameters, μ and σ^2, for this diffusion model are given respectively by the two constants: $(1/t)$ times the mean and variance of log N_t/N_0 as t increases. The next step is to ask whether and over what time frames, a diffusion model gives reasonable approximations for the extinction or quasi-extinction metric of interest.

Even though the distribution of log N_t/N_0 for t not too small may be properly characterized by a diffusion model, the probability of crossing thresholds may not be since the variance of log N_t/N_0 is higher than $\sigma^2 t$ for $t < 5$ and since there is temporal correlation in the population process due to age structure in the population model. Thus the extinction or quasi-extinction metric itself needs to be tested. For this paper, the extinction metric of interest was the probability of 90% decline within different time frames. Although estimating extinction to one individual is a popular risk metric, and unfortunately sometimes mandated, the reader is cautioned about using the diffusion approximation to estimate extinction to very low numbers since factors that drive dynamics at very low population sizes (such as demographic stochasticity) and the catastrophes often associated with ultimate extinction will likely be poorly represented in a time series of a relatively larger population declining to low numbers (e.g., Fagan et al., *unpublished manuscript*).

The diffusion approximation was tested by comparing the actual probabilities from the 1000 simulations to those predicted by the diffusion model with parameters μ and σ^2 from Fig. 1(a–c). For the diffusion model, the probability of decline to a threshold population size N_{crit} within t years is (Lande and Orzack 1988, Dennis et al. 1991):

$$\begin{aligned}\Pr(\text{decline to } N_{crit} \text{ in 0 to } t) \\ = \pi' \times \left\{ \Phi\left[\frac{-\log(N_0/N_{crit}) + |\mu|t}{\sigma\sqrt{t}}\right] \right. \\ + \exp\left[\frac{2\log(N_0/N_{crit})|\mu|}{\sigma^2}\right] \\ \left. \times \Phi\left[\frac{-\log(N_0/N_{crit}) - |\mu|t}{\sigma\sqrt{t}}\right]\right\} \qquad (6)\end{aligned}$$

where

$$\pi' = \begin{cases} 1 & \mu \le 0 \\ \exp[-2\mu \log(N_0/N_{crit})/\sigma^2] & \mu > 0. \end{cases}$$

For 90% decline, $N_0/N_{crit} = N_0/(0.1N_0) = 10$. $\Phi(y)$ is the value of the cumulative distribution function at y of a standard normal distribution with mean 0 and variance 1. Eq. 6 gives the probability of 90% decline at any time within 0 to t; however in the simulations, population size is only observed at discrete yearly intervals. In this application, the difference between Eq. 6 and the technically correct comparison, the probability of observing quasi-extinction at a discrete set of yearly intervals, is minor, but this is not always the case. If needed, the probability of observing quasi-extinction at discrete intervals can be calculated numerically by simulating the diffusion process.

Fig. 1(d–f) shows the actual (gray lines) vs. diffusion approximation estimates (black lines) of the probability of 90% decline. The plots illustrate that the diffusion model correctly described the probability of 90% decline within different time frames for the simulated time series of current and future spawners. This was true despite the fact that these time series were not strictly speaking a diffusion process. These results are for a particular passage metric of interest, i.e., the probability of 90% decline. The results should not be overgeneralized to say that a simple diffusion model would correctly describe all extinction and quasi-extinction metrics for salmonids. It is important to test particular metrics of interest such as done here for the probability of 90% decline. Also it is important to consider what segment of the population to track. ''Future and current spawners'' integrates over multiple age classes without over representing any one age class and smoothes out the year-to-year boom–bust cycles that salmon are prone to. This effectively limits the nonprocess error in the time series and enables Eq. 6, which is based on a diffusion model with no nonprocess error, to work well.

Parameter Estimation for Diffusion Approximations

This first evaluation indicated that there exists a diffusion model that correctly approximates the behavior of these simulated salmon populations for the purposes of estimating the distribution of future reproductive population sizes and for estimating the probability of hitting a critical level of 10% of current size. The next question is how well the parameters of this diffusion model can be estimated given realistic data constraints. Here I review currently available methods for parameterizing a diffusion model: maximum likelihood methods assuming low data corruption (Dennis et al. 1991), the same methods but using a running sum transformation of the data (Holmes 2001), maximum likelihood methods using a Kalman filter (Harvey 1989, Lindley 2003), slope methods (Holmes 2001, Holmes and Fagan 2002), and asymptotically unbiased estimators for random matrix products (Heyde and Cohen 1985). Code for these methods is given in Supplement 1. The different methods have their pluses and minuses depending on the quality and length of data available. Below these methods are described and in the next

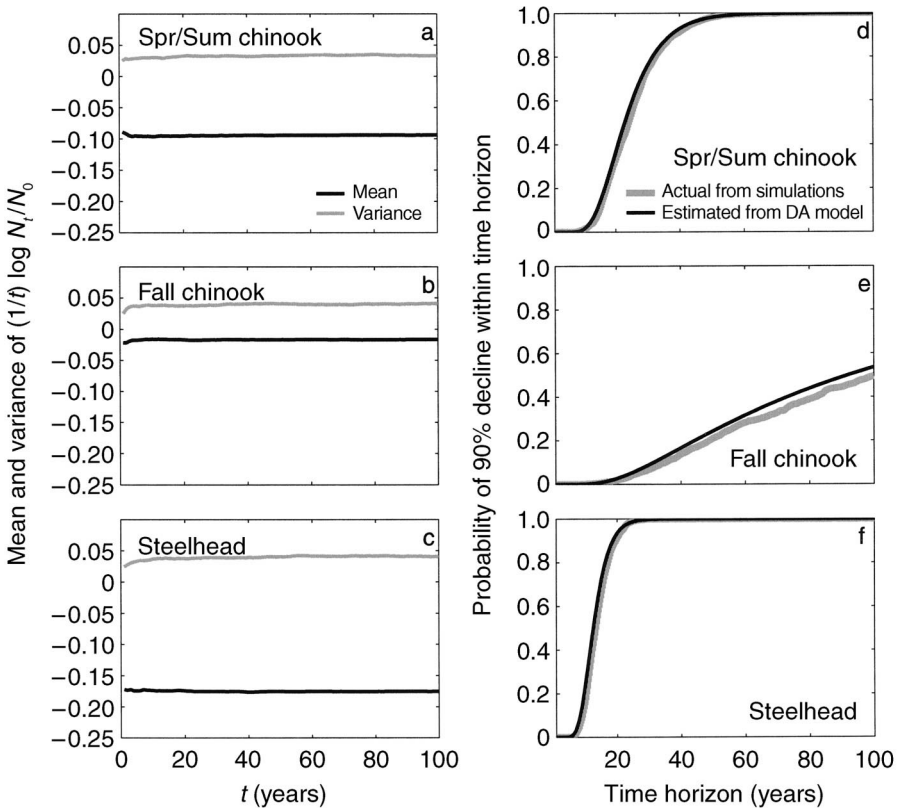

FIG. 1. Behavior of the population trajectories from the age-structured salmonid models versus the behavior of the diffusion approximation. For a diffusion model, the mean and variance of log population size $(N_t/N_0) \times (1/t)$ should be a constant. The age-structured models had this property beginning at approximately $t = 5$ (a–c). The diffusion model also correctly estimated the probability of 90% decline ($N_t/N_0 = 1/10$) within different time frames (d–f). ''Spr/Sum chinook'' indicates spring/summer chinook.

section, I evaluate the performance of these different methods given the data constraints faced by a PVA of Columbia River endangered and threatened salmonid populations (McClure et al. 2003).

All these methods start with a time series of censuses, $O_1\ O_2\ O_3 \cdots O_k$, of the population. For the purposes of this paper, I assume O_t represents yearly censuses. The census need not enumerate the entire population; index counts, such as a segment of the population or a specific age or stage class, can also be used.

Maximum likelihood methods

I use the following state-space model to model the relationship between the diffusion model we wish to estimate from the data, the actual population counts, and the observed counts:

$$X_t = X_{t-1}\exp(\mu + \varepsilon_p) \tag{7a}$$

where $\varepsilon_p \sim$ normal(mean = 0, variance = σ^2),

$$N_t = X_t\exp(\varepsilon_n) \tag{7b}$$

where $\varepsilon_n \sim$ g(mean = 0, variance = σ_n^2),

$$O_t = N_t\exp(\varepsilon_{se}) \tag{7c}$$

where $\varepsilon_{se} \sim$ f(mean = β, variance = σ_{se}^2).

In the above equations, ''$\sim$'' means ''is distributed as'' and g(·) and f(·) are some unspecified statistical distributions. X_t represents the diffusion process, which correctly approximates the age-structured population process for t not too small. N_t are the true population counts, which include process error (the X_t part) and nonprocess error, σ_n^2, which is variability due for example to density-dependent feedback or age-structure cycles. Nonprocess error is apparent in the time series, but does not multiply through time as ε_p does in Eqs. 7. Observations of N_t are made and these observations are corrupted by some sampling error, ε_{se}. Eqs. 7 reduce to:

$$X_t = X_{t-1}\exp(\mu + \varepsilon_p)$$
$$O_t = X_t\exp(\varepsilon_n + \varepsilon_{se}) \tag{8}$$

where $\varepsilon_n + \varepsilon_{se} \sim$ h(mean = β, variance = σ_{np}^2) and where h(·) is an unspecified statistical distribution. In this paper, I refer to this state-space model as a cor-

rupted diffusion process. The total nonprocess error, σ^2_{np}, equals nonprocess error in the age-structured population process, σ^2_n, plus nonprocess error from sampling error in the censuses of the population, σ^2_{se}.

If we assume that there is no nonprocess error in our observations, $\sigma^2_{np} = 0$ and $O_t = X_t$. Then the maximum likelihood estimates of μ and σ^2 are (Dennis et al. 1991):

$$\hat{\mu}_d = \text{mean of } \log(O_{t+1}/O_t)$$
$$\hat{\sigma}^2_d = \text{variance of } \log(O_{t+1}/O_t)$$
$$\text{for } t = 1, 2, 3, \ldots, k - 1. \quad (9)$$

Eq. 9 assumes that there are no missing data; cf. Dennis et al. (1991) for methods when there are missing data. Although the assumption of no nonprocess error in the data will never strictly hold for real data, analysis of a variety of nonsalmon time series suggests that it can be a reasonable approximation (Fagan et al., *unpublished manuscript*).

The estimates $\hat{\mu}_d$ and $\hat{\sigma}^2_d$ are the maximum likelihood estimates for the diffusion approximation (Eq. 8) not of the age-structured process itself. This creates problems if the life history of the species of concern is such that there is high nonprocess error in the time series due to age-structure cycling. If there is also sampling error in the data, this adds additional nonprocess error variability. Thus Eq. 9 will produce biased estimates of σ^2 because it attributes the nonprocess error in the data to process error, and the expected value of $\hat{\sigma}^2_d$, $E(\hat{\sigma}^2_d)$, is $\sigma^2 + 2\sigma^2_{np}$. In this case, it is more appropriate to use the full state-space model (Eq. 8) with σ^2_{np} not assumed to be zero. Kalman filters are widely used in time series analysis for maximum likelihood estimation for state-space models of the form of Eq. 8 with ε_{np} assumed to be normally distributed (Harvey 1989: section 3.4). Lindley (2003) gives a specific case of such an algorithm for population processes of the form of Eq. 8.

This approach for maximum likelihood estimation for Eq. 8 can be summarized as follows. Assume $h(\cdot)$ in Eq. 8 is normal with $\beta = 0$ (if β is non-zero it can be factored out) and define $y_t \equiv \log O_t$, then the distribution of y_t given all the data prior to y_t is:

$$y_t \mid \{y_1, y_2, \ldots, y_{t-1}\} \sim \text{normal}[\text{mean} = E(y_t \mid \{y_1, y_2, \ldots, y_{t-1}\}), \text{variance} = F_t] \quad (10)$$

where $E(\cdot)$ is "expected value" and F_t denotes the variance of $y_t \mid \{y_1, y_2, \ldots, y_{t-1}\}$. The likelihood of a particular set of μ, σ^2_{np} and σ^2 given the data, $\{y_1, y_2, \ldots, y_T\}$, is

$$L(\mu, \sigma^2_{np}, \sigma^2 \mid \{y_1, y_2, \ldots, y_T\}) = \prod_{t=1}^{T} p(y_t \mid \{y_1, y_2, \ldots, y_{t-1}\}) = \prod_{t=1}^{T} \exp\left\{-\frac{[y_t - E(y_t \mid \{y_1, y_2, \ldots, y_{t-1}\})]^2}{2F_t}\right\} \times (2\pi F_t)^{-1/2}$$

and the log likelihood is

$$\log L(\mu, \sigma^2_{np}, \sigma^2 \mid \{y_1, y_2, \ldots, y_T\}) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log F_t - \frac{1}{2}\sum_{t=1}^{T}\frac{v_t^2}{F_t} \quad (11)$$

where $v_t = y_t - E(y_t \mid \{y_1, y_2, \ldots, y_{t-1}\})$. The Kalman filter calculates v_t and F_t given a particular set of μ, σ^2_{np}, and σ^2. The Kalman filter for this application is given in Appendix A or see Harvey (1989: chapter 3). The maximum likelihood estimates are the set of μ, σ^2_{np}, and σ^2 that maximize Eq. 11 and are denoted here as $\hat{\mu}_{ka}$ and $\hat{\sigma}^2_{ka}$.

A variant of the maximum likelihood methods uses a running sum transformation of the data in place of O_t in Eq. 9. A running sum is defined in this case as

$$R_t = \sum_{i=0}^{L-1} a_i O_{t+i} \quad (12)$$

where the a_i are weightings. The running sum estimate for μ is

$$\hat{\mu}_{run} = \text{mean of } \log(R_{t+1}/R_t)$$
$$\text{for } t = 1, 2, 3, \ldots, k - L. \quad (13)$$

Often the purpose of the running sum is to transform the data into a population-level count (see examples in Dennis et al. 1991 and Holmes 2001). But the transformation has an independent benefit when data are corrupted by extraneous variability. In this case, $\hat{\mu}_{run}$ can have lower variability and bias than $\hat{\mu}_d$ (Holmes 2001, Holmes and Fagan 2002). The process error, σ^2, cannot be estimated using R_t in Eq. 9. The running sum transformation filters out not only the extraneous variability but also the process error that we are trying to estimate.

Slope method

The slope method was developed to deal with situations where O_t is corrupted with high nonprocess variability, σ^2_{np}, that is up to one to two orders of magnitude $>\sigma^2$ and has an unknown statistical distribution (Holmes 2001, Holmes and Fagan 2002). It also uses the corrupted diffusion model, but estimates the parameters by examining how the distribution of $\log O_{t-\tau}/O_t$ changes with τ. From Eq. 8:

$$\text{mean of } \log(O_{t+\tau}/O_t) = \mu\tau$$
$$\text{variance of } \log(O_{t+\tau}/O_t) = \sigma^2\tau + 2\sigma^2_{np} \quad (14)$$

where the τ are chosen such that consecutive $O_{t+\tau}/O_t$ ratios do not overlap. The slopes of the linear regressions of the mean and variance of $\log(O_{t+\tau}/O_t)$ vs. τ

give estimates of μ and σ^2, respectively. An unbiased estimator for σ^2 can be devised using Eq. 14 with a long enough time series, however when the available time series is short (such as <20 years) such an estimator is prone to negative variance estimates. To stabilize the estimates, regression can be done on the running sum transformed data, $R_{t+\tau}/R_t$ instead (Holmes and Fagan 2002):

$$\hat{\mu}_{\text{slp}} = \text{slope of mean}\left[\log\left(\frac{R_{t+\tau}}{R_t}\right)\right] \text{ vs. } \tau, \text{ intercept} = 0$$

$$\hat{\sigma}^2_{\text{slp}} = \text{slope of var}\left[\log\left(\frac{R_{t+\tau}}{R_t}\right)\right] \text{ vs. } \tau, \text{ intercept free} \quad (15)$$

for $t = 1, 2, 3, \ldots, k - L$ and maximum $> \cong 5$. This slope estimate for σ^2 is biased; from numerical simulations $E(\hat{\sigma}^2_{\text{slp}}) \approx 0.5\sigma^2 + 0.15\sigma^2_{\text{np}}$. If nonprocess error is an order of magnitude greater than the process error, the result is mean estimates that are approximately 200% of the correct value.

Asymptotically unbiased estimator for σ^2

The previous methods are all based on the diffusion approximation for $\log N_{t+\tau}/N_t$, which has variance $\sigma^2\tau$ for very small τ. The σ^2 estimates from these methods will tend to be biased to some degree even if the time series is infinitely long. This problem is worst for $\hat{\sigma}^2_{\text{d}}$, which relies on the variance at $\tau = 1$ for estimation of σ^2. Heyde and Cohen (1985) provide an estimator for σ^2 that is not based on the diffusion approximation but rather on products of random matrices:

$$\hat{\sigma}^{1/2}_{\text{hc}} = \frac{1}{2}\left(\frac{\pi}{2}\right)^{1/2}\left[\frac{1}{\log(T-1)}\sum_{j=1}^{T-1} j^{-3/2}\left|\log\left(\frac{\mathbf{N}_{j+1}}{\mathbf{N}_1}\right) - j\mu\right| + \frac{1}{\log(T-2)}\sum_{j=1}^{T-2} j^{-3/2}\left|\log\left(\frac{\mathbf{N}_{j+2}}{\mathbf{N}_2}\right) - j\mu\right|\right]. \quad (16)$$

This estimator is asymptotically unbiased, i.e., unbiased as the length of the census time series goes to infinity, although not necessarily unbiased for short time series as shall be seen. The asymptotically unbiased estimator for μ is the same as $\hat{\mu}_{\text{d}}$.

Correcting for inputs

All these methods assume that there are no external inputs into (or subtractions from) the population. If inputs are occurring, they will mask the true μ and thus must be corrected to prevent bias in the estimation of μ. In Appendix B, a general method to correct for the input problem is derived. The formulation of the correction depends on the life history of the species at hand and the ages at which individuals are externally input into the population. Thus a specific correction will have to be derived following the methods in Appendix B for the species of concern. Here the calculation is illustrated for salmon populations experiencing regular hatchery fish introductions.

Hatchery-reared juvenile fish are regularly released into Pacific salmon stocks as part of remediation for impacts or to provide fishing opportunities. If these hatchery-born fish return to reproduce in the wild, their offspring are indistinguishable from the offspring of wild-born fish, for typical census purposes at least. The goal of the input correction is to estimate what μ would be if no supplementation were occurring. As described in Appendix B, correcting for inputs requires determining the relationship between the true mean number of offspring per wild-born spawner at year t, denoted $R_{0,t}$, to the apparent mean number, denoted $\tilde{R}_{0,t}$, based on the population growth rate with inputs occurring. The species in this study reproduce once and die, which makes $R_{0,t}/\tilde{R}_{0,t}$ a relatively straightforward calculation (cf. Appendix B):

$$\tilde{R}_{0,t} = R_{0,t}(1 + S_{\text{h},t}/S_{\text{w},t}) = R_{0,t}/f_{\text{w},t} \quad (17)$$

where $f_{\text{w},t}$ is an estimate of the fraction of returning spawners at year t that were wild-born. For notational simplicity, it is assumed here that hatchery fish reproduce at the same rate as wild-born fish. If hatchery fish reproduce at lower effectiveness, $S_{\text{h},t}$ can be reexpressed as wild-equivalents by multiplying $S_{\text{h},t}$ by the reproductive effectiveness of hatchery fish relative to wild-born fish.

To solve for the no supplementation population growth rate at year t, denoted λ_t, in terms of the observed with supplementation growth rate, denoted $\tilde{\lambda}_t$, the relationship $\log(\lambda) \approx \log(R_0)/T$ was used. Here T is mean generation time (Caswell 2001: 126–130). Combining this relationship with Eq. 17 gives

$$\log(\lambda_t) = \frac{1}{T}\log\left(\frac{R_{0,t}}{\tilde{R}_{0,t}}\right) + \log(\tilde{\lambda}_t) = \frac{1}{T}\log(f_{\text{w},t}) + \log(\tilde{\lambda}_t). \quad (18)$$

As discussed in Appendix B, $\tilde{\lambda}_t$ could be estimated from the ratio of wild spawners at year $t + 1$ vs. year t, e.g., $(f_{\text{w},t+1}S_{t+1}/f_{\text{w},t}S_t)$, or by the ratio of age-1 fish, e.g., $(F_tS_t/F_{t-1}S_{t-1})$, where F_t is the fecundity of spawners. However, for Pacific salmon populations, $f_{\text{w},t}$ is often approximate, and yearly estimates of F_t are unusual. Instead, I will use here the total spawner ratios, S_{t+1}/S_t, to estimate $\tilde{\lambda}_t$. This is an approximation since it tacitly assumes that either $f_{\text{w},t}$ or F_t remain relatively constant from year to year, but reflects an approximation that is required given real data constraints. The input-corrected μ estimate is then

$$\hat{\mu}_{\text{ic}} = \text{mean of}\left[\frac{1}{T}\log(\hat{f}_{\text{w},t}) + \log\left(\frac{S_{t+1}}{S_t}\right)\right]. \quad (19)$$

Overview of parameterization performance

Table 2 gives an overview of the four different parameterization methods in terms of whether they cor-

TABLE 2. Overview of estimator performance and limitations.

Estimators	Corrects for sampling error	Expected value	No. years needed (rules of thumb)
μ estimators			
Dennis (same as Heyde-Cohen)	no	μ, sensitive to age-structure perturbations in data	10+
Kalman	yes	μ	10+
Runsum	yes	μ	15+
Slope	yes	μ	20+
σ^2 estimators			
Dennis	no	$\sigma^2 + 2\sigma^2_{np}$	10+
Kalman	yes	σ^2†	20–50+†
Slope	yes	$\approx 0.5\sigma^2 + 0.15\sigma^2_{np}$	20+
Heyde-Cohen	no	$\approx \sigma^2 + 2\sigma^2_{np}$ for short series	100+

† Although the mean estimates are equal to the true value, i.e., are unbiased, the distribution of estimates can be highly skewed for short time series with low σ^2 and high σ^2_{np} (total nonprocess error), such that the vast majority of estimates can be 0 and a few rare extremely high estimates bring the mean up to σ^2.

rect for extraneous variability in the time series and some rough guidelines regarding their data needs. The table emphasizes that there is a trade-off between data needs and the ability to correct for error in the data. When the data contain high extraneous variability, the reduction in parameter bias warrants the increase in the variability of parameter estimates. The flip side of this is that using a method that corrects for error on time series with low errors will lead to an unnecessary loss of precision.

PERFORMANCE OF PARAMETERIZATION METHODS GIVEN DATA CONSTRAINTS

Although Table 2 can help select likely candidate methods for a particular application, a more quantitative evaluation of the performance of particular parameterization methods is needed since actual census data often have a level extraneous error somewhere between none and severe and thus the trade-off between precision and bias is unclear. Such an evaluation is illustrated here using the stochastic matrix models for the three salmonid populations. This evaluation incorporates the following data constraints faced by a PVA of these populations (McClure et al. 2003): (1) counts of only the spawning segment of the populations, (2) time series limited to 20 years, (3) severe age-structure perturbations in the beginning of some time series due to reproductive collapses during dam construction (Williams et al. 2001), and (4) high sampling error.

To examine the robustness of the parameter estimation methods to these constraints, 20-year time series of spawner only counts were generated using the matrix models. The time series were generated using either an initial stable age structure (drawn randomly from the stable set) or an initial age structure with no age-1 individuals. The simulated spawner time series, S_t, were then corrupted with lognormal sampling error:

$$O_t = S_t \exp(\varepsilon_{se}) \quad \text{with } \varepsilon_{se} \sim \text{normal}(\text{mean} = \alpha, \text{variance} = \sigma^2_{se}). \tag{20}$$

The observed spawner count at year t is O_t. The DA parameters, μ and σ^2, were estimated from each 20-year simulated time series using the Dennis, Kalman, running sum, Heyde-Cohen, or slope methods. For the running sum and slope methods observed, spawner counts were transformed by adding four consecutive counts: $R_t = \Sigma_{i=0}^{3} O_{t+i}$. Parameter estimates were compared with the correct DA parameters determined from Fig. 1.

Realistic levels of sampling error to add were estimated from studies on sampling error in spawner and redd surveys (Jones et al. 1998, Dunham and Rieman 2001) and from an examination of the average nonprocess error in Columbia River redd count data. Redds are the egg nests made by spawning salmon, and redd counts were the most common data type for the Columbia River PVA in McClure et al. (2003). Using a lognormal model for observation errors, Jones et al. (1998) found levels of total within plus between observer sampling error variability of $\sigma^2_{se} = 0.09$ to 0.78. These levels of lognormal sampling variability are consistent with those found in another study of variability between different observers' counts of redds (Dunham and Rieman 2001). To find the average total nonprocess error (which includes sampling variability plus other nonprocess error) in actual time series, I examined a collection of 44 20-year redd-count time series from different spawning areas in the Snake River basin, Idaho (data in supplements of McClure et al. 2003). I estimated an average total nonprocess error of $\sigma^2_{np} = 0.56$ in this redd-count data. The empirical studies on sampling errors combined with the average total nonprocess error within redd-count time series suggest that year-to-year sampling error variability is likely to be in the range of $\sigma^2_{se} = 0.1$ to 0.75, with 0.1 on the low

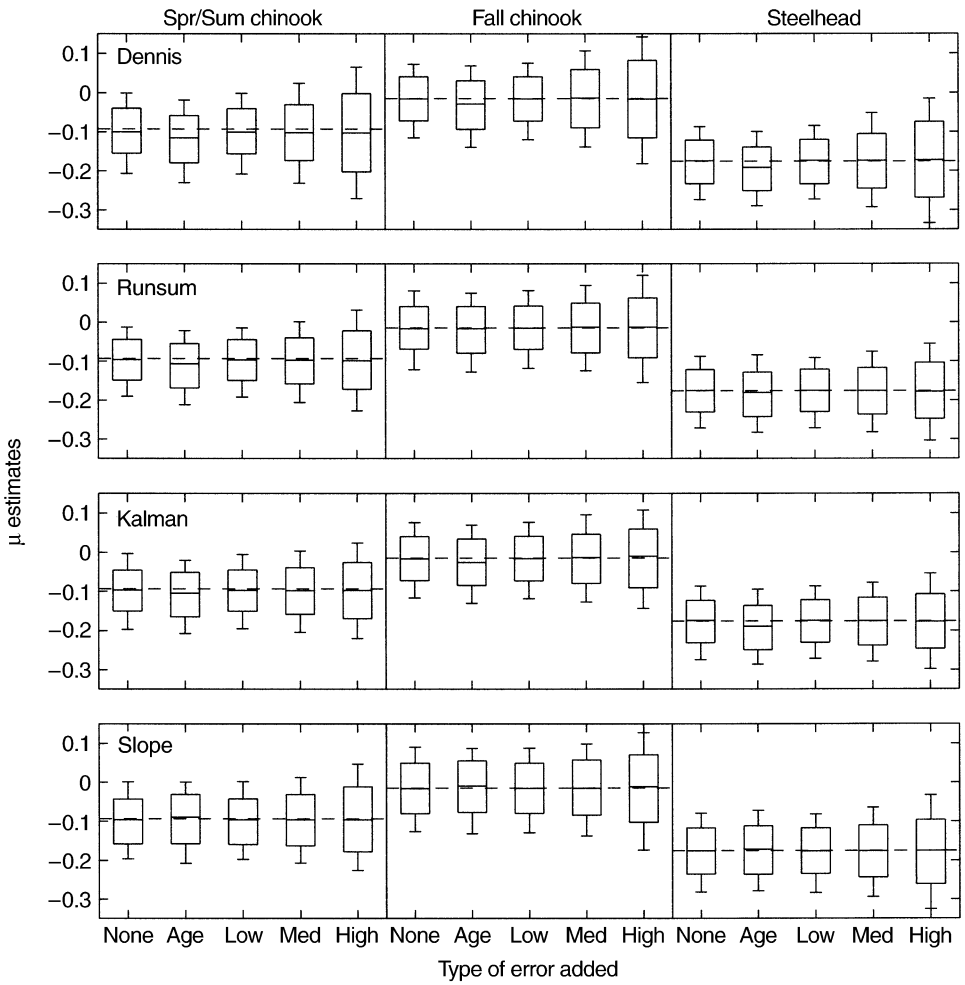

FIG. 2. Performance of the μ estimates from the Dennis, runsum, Kalman, and slope methods. One thousand 20-year spawner time series were generated from the spring/summer chinook, fall chinook, and steelhead stochastic matrix models. The column labels indicate the type of error added: ''None'' indicates stable starting age structure and no sampling error; ''Age'' indicates no sampling error but perturbed age structure; ''Low,'' ''Med,'' and ''High'' indicate stable age structure plus lognormal sampling error with $\sigma^2_{se} = 0.1$, 0.5, or 0.75, respectively. The box plots summarize the estimates among all 1000 simulations. The line in the box shows the median estimates, and the box encloses the middle 75% of the estimates; the whiskers enclose 95% of the estimates.

end given minimum observer variability alone and 0.75 on the high end given that this is more than observed levels of total nonprocess error. These levels of sampling error are ~10–100 times larger than the median process error estimates from actual salmonid time series in the Columbia River basin.

To model these sampling error levels, lognormal sampling error was added to the spawner counts using Eq. 20 with low ($\sigma^2_{se} = 0.1$), medium ($\sigma^2_{se} = 0.5$), or high ($\sigma^2_{se} = 0.75$) sampling error. The bias in error, α, drops out since the parameter estimates always use the difference between log counts (e.g., $\log O_{t+1} - \log O_t$). Jones et al. (1998) found that observers have a nonlinear tendency to underestimate as the number of objects to be counted increases. However, they found this for counts of tens of thousands of salmon; Dunham and Rieman (2001) did not find this pattern for redd counts in the hundreds, which are more typical of the counts for the endangered and threatened salmon populations simulated in this study. Thus, I assumed that the sampling error was independent of total spawner numbers.

Fig. 2 shows the performance of the μ estimators with either no sampling error and a stable starting age structure, with only an age-structure perturbation and no sampling error, or a stable initial age structure and low, medium, or high sampling error. In the box plots, the middle line is the median and the box encloses 75% of the estimates from the 1000 simulations. All methods gave unbiased estimates of μ for the simulations started with a stable age structure. The Dennis and slope estimates of μ were most variable, and the runsum and Kalman estimates were generally least variable. The age-structure perturbation led to an underestimation of μ that was most apparent for the Dennis

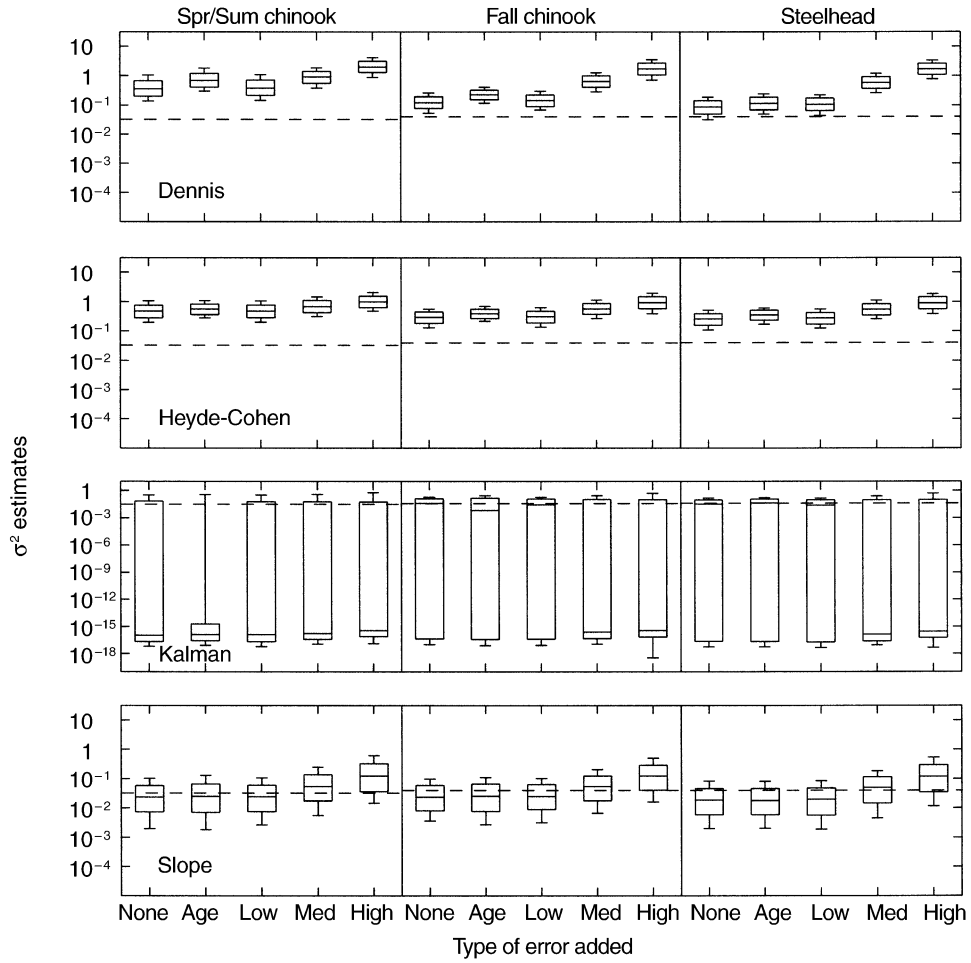

FIG. 3. Performance of the σ^2 estimates (log scale) from the Dennis, Heyde-Cohen, Kalman, and slope methods for the spring/summer chinook, fall chinook, and steelhead simulations. See Fig. 2 for details.

estimate, while the other three methods showed only slight underestimation. For the runsum and Kalman estimates, 75% of the estimates were within ± 0.05 of the correct value for low and medium sampling errors. For high sampling error, the range of the estimates increased to 75% of estimates between ± 0.08 of the correct value. These errors can be translated into errors in percentage yearly growth rate by multiplying by 100.

Fig. 3 shows the performance of the σ^2 estimators. The Dennis method overestimated σ^2 for all cases, with and without sampling error. For example, with no sampling error and a stable starting age structure, the median $\hat{\sigma}^2_d$ estimates were 0.35, 0.12, and 0.08 vs. the correct values of 0.033, 0.042, and 0.040 for spring/summer chinook, fall chinook, and steelhead, respectively. The bias was due to the tendency of the salmon simulations, like real salmon populations, to exhibit boom and bust cycles. The result was high nonprocess error in the time series even without added sampling error. With a perturbed age structure, the $\hat{\sigma}^2_d$ more highly overestimated σ^2, and with sampling error added in addition, the overestimation was severe. The median σ^2_d were 2.02, 1.74, and 1.71 for the three species with the high added sampling error. The Heyde-Cohen estimates were similarly highly biased with high nonprocess error.

In contrast to the overestimation by these estimators, the Kalman estimator was prone to underestimating the correct σ^2; median Kalman estimates were close to 10^{-16} for spring/summer chinook across all corruption levels and for fall chinook and steelhead for the higher sampling error levels. Even when the median estimates were close to the true values (i.e., for fall chinook and steelhead with low sampling error), σ^2 estimates close to zero were still common and comprised 20–30% of the estimates. Overall for the 20-year time series especially the time series with high sampling error levels, the data were usually most likely under a scenario with all the variability attributed to nonprocess error and thus $\sigma^2 = 0$. In contrast, the mean estimate of σ^2 (rather than the median) was uniformly close to the true σ^2 (similar to the results by Lindley [2003]). This difference between the mean and median estimates occurred because the distribution of estimates was highly

skewed, with most estimates essentially zero except for a few very high estimates, which brought up the mean. This difference between the median Kalman estimates and the correct DA values diminished if a longer time series was used, and largely disappeared for 100-year time series.

The slope method gave estimates that were overall closest to the true σ^2 for the 20-year time series, although with no sampling error added to the time series, they tended to underestimate the true value as predicted. With a stable initial age structure and no sampling error, the median estimates were 0.025, 0.022, and 0.018 relative to the correct values of 0.033, 0.042, and 0.040 for the three species, respectively. The estimates were not appreciably changed by the age-structure perturbation nor low sampling error. As sampling error increased to the medium and high levels, however, the slope estimator began to increase and for the highest sampling error the median estimates were 0.12, 0.12, and 0.12, for the three species. These estimates were still substantially closer to the true values than the Dennis and Heyde-Cohen estimators at high sampling error levels.

Testing the input corrector

To study the robustness of the input correction (Eq. 19), simulations were run as above but with a random number of age-1 hatchery fish added each year. Between 1 and 1×10^5 fish were added each year; the number drawn from a uniform random distribution. All simulations were started with a stable age structure. Each simulation produced wild and hatchery spawner counts of which observations, $O_{w,t}$ and $O_{h,t}$, were made with medium sampling error. The parameter μ was estimated using $\hat{\mu}_d$ with either no input correction (Eq. 13 with $O_t = O_{w,t} + O_{h,t}$) or with input correction (Eq. 19). The wild fraction was estimated from the $O_{w,t}$ and $O_{h,t}$ counts:

$$\hat{f}_{w,t} = \frac{O_{w,t}}{O_{w,t} + O_{h,t}}. \tag{21}$$

The slope method (Eq. 15 with $R_t = \Sigma_{i=0}^{3} O_{t+i}$) was used to estimate σ^2.

The top panel of Fig. 4 contrasts the input-corrected vs. the uncorrected μ estimates. Without correction for inputs, the median estimates of μ were 0.00, 0.02, and 0.00 indicating stable or increasing populations with the inputs relative to the correct values of -0.093, -0.015, and -0.18, indicative of declining wild dynamics for the three species, respectively. With correction, the median estimates were -0.096, -0.023, and -0.18. Thus, the input correction successfully extracted the correct value from population trajectories that otherwise appeared to come from a process with much higher μ due to hatchery inputs. The lower panel illustrates that the $\hat{\sigma}^2_{slp}$ estimate with no explicit hatchery correction gave a relatively unbiased σ^2 estimate despite the variable hatchery inputs.

Sensitivity of Risk Metrics to Parameter Misestimation

The errors in parameter estimation seen in Figs. 2 and 3 translate into errors in the estimated risk metrics. Two risk metrics were evaluated for their sensitivity to parameterization errors: the long-term population growth rate, $\lambda = \exp(\mu)$ and the probability of a 90% decline within a given time frame (Eq. 6). Fig. 5 compares the estimates of λ. The figure shows the probability density function, i.e., $P(\alpha < \lambda < \alpha + d\alpha)/d\alpha$. Since the variability of μ estimates was not dramatically different among the estimators, the variability in λ using only $\hat{\mu}_{run}$ is shown. The figure illustrates that 50% of the estimates were within ± 0.04 of the correct value even for high sampling error; although, much lower and higher estimates were not uncommon. With low sampling error 10% of estimates were more than ± 0.07 greater than the correct values for the three species. For high sampling error, this increased to ± 0.10. This level of variability indicates that unless the population is declining very rapidly, a 20-year time series is unlikely to be sufficient to reject a null hypothesis that the population is stable or increasing ($\lambda > 1$) at the $P = 0.05$ level using a standard t test.

Fig. 6 compares the estimates of the probability of 90% decline from simulated time series with a stable initial age structure and no sampling error added. Three methods for estimating the parameters were contrasted. The first used the Dennis estimators, which assumed zero nonprocess error (as in Dennis et al. 1991). The second method used the Kalman estimates with normal nonprocess error. The third used the slope method for σ^2 and the running sum method for μ (as in Holmes and Fagan 2002). Due to high overestimation of σ^2 by $\hat{\sigma}^2_d$, the median probability estimates using the Dennis estimators were highly biased relative to the observed probabilities from the simulated spring/summer chinook and fall chinook time series (Fig. 6, left panels). The probability of 90% decline for steelhead was dominated by the very low μ, and was less affected by overestimation of σ^2. The majority of Kalman estimates for σ^2 were severe underestimates for spring/summer chinook (Fig. 3), and the median probability estimates were correspondingly biased for this species (Fig. 6, middle panels). For steelhead and fall chinook, the median Kalman estimates for σ^2 were close to the true values when no sampling error was added (Fig. 3), and in this case the median probability estimates were correspondingly close to the true probabilities. The slope estimates of σ^2 were generally close to the true values for all species, and the median probability estimates using $\hat{\sigma}^2_{slp}$ matched the observed probabilities across species.

Estimation of the probability of 90% decline using the parameter estimates, $\hat{\sigma}^2_{slp}$ and $\hat{\mu}_{run}$, was also tested for sensitivity to sampling error levels. Estimation was generally robust to low to medium sampling error (ex-

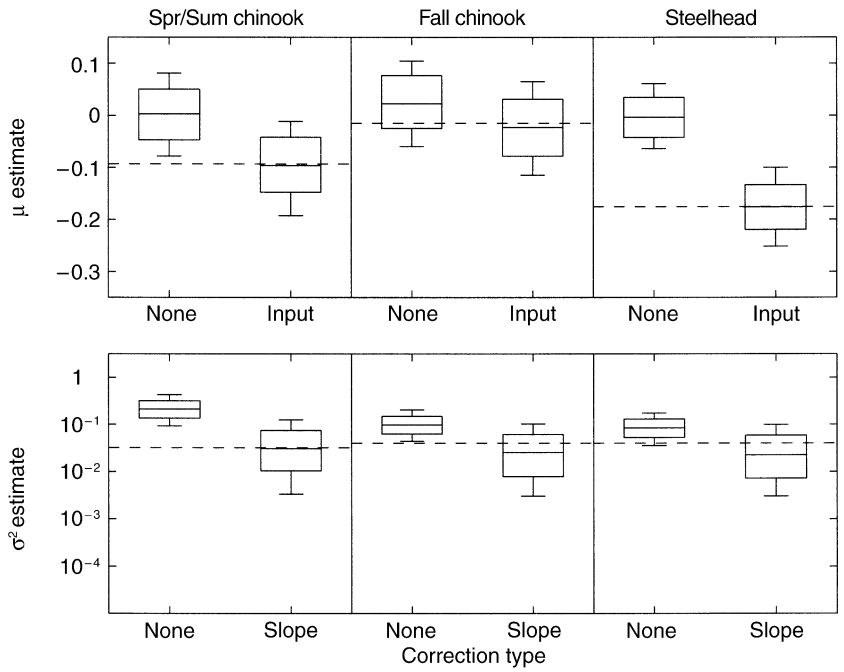

FIG. 4. Error in μ and σ^2 estimates (log scale for σ^2) due to yearly hatchery inputs into the salmon simulations. The starting age structure was perturbed, and low sampling error ($\sigma^2_{se} = 0.1$) was added. The parameter μ was estimated using $\hat{\mu}_d$ with either no input correction (Eq. 13) or with input correction (Eq. 19). The σ^2 parameter was estimated using $\hat{\sigma}^2_d$ vs. $\hat{\sigma}^2_{slp}$ with no explicit hatchery correction.

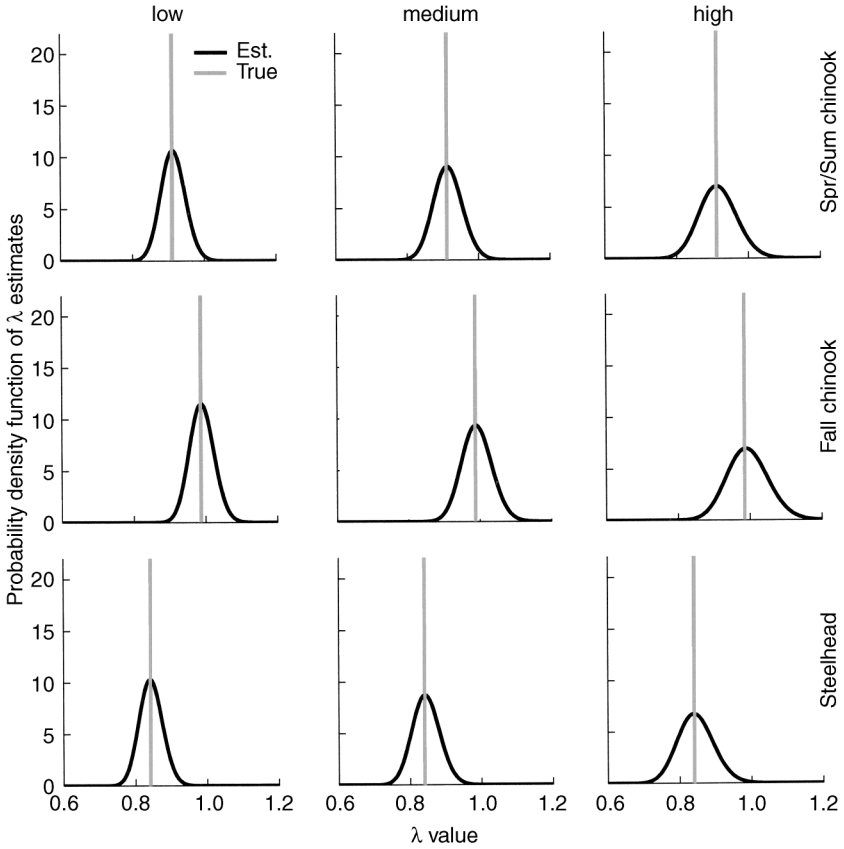

FIG. 5. Variability of λ estimates using $\hat{\mu}_{run}$. See Fig. 2 for simulation details.

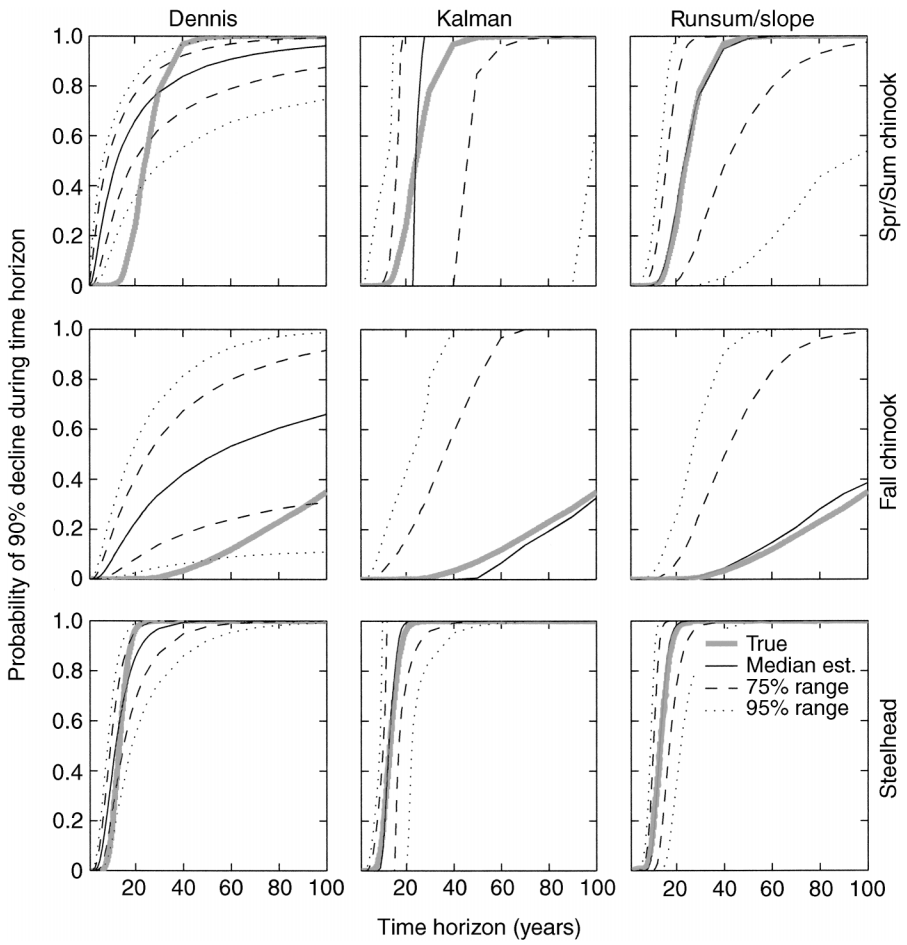

FIG. 6. Estimated vs. true probability of 90% decline within a given time horizon. Probability estimates were calculated using Eq. 6 with $\hat{\mu}_d$ and $\hat{\sigma}^2_d$ (Dennis), Eq. 6 with $\hat{\mu}_{ka}$ and $\hat{\sigma}^2_{ka}$ (Kalman), or Eq. 6 with $\hat{\mu}_{run}$ and $\hat{\sigma}^2_{slp}$ (runsum/slope). The parameters were estimated from simulated spawner counts. The true probability was calculated from 1000 100-year time series generated from each stochastic matrix model. Results are shown for simulations with no sampling error added to the spawner counts. Simulations were started with a stable initial age structure. The solid line shows the median errors observed in the simulations, and the dashed and dotted lines encompass 75% and 95% of the estimates, respectively.

cept for fall chinook), but a mismatch between median predictions and the true probabilities became apparent with high sampling error (Fig. 7) when the difference between $\hat{\sigma}^2_{slp}$ and σ^2 increased (Fig. 3). This was especially true for fall chinook with the predicted probabilities bearing little relationship to the true low probabilities. Variability in the estimated probabilities was highest for fall chinook and lowest for steelhead. These differences in the variability of estimates between species was likely due to differences between their rates of decline; λ was closest to 1.0 for fall chinook (with $\lambda = 0.98$) and farthest for steelhead (with $\lambda = 0.84$). Fig. 7 illustrates that in some cases, e.g., the fall chinook with λ close to 1.0, estimated probabilities of severe declines were so variable as to provide little useful risk information. In other cases, however, e.g., the spring/summer chinook and steelhead with λ values farther from 1.0, the median estimated probabilities correctly tracked the true probabilities. Nonetheless the probability estimates were variable and accurately representing the variability is critical to the use of probability metrics in applications, as discussed in the next section.

APPLICATION TO ACTUAL DATA

The previous sections have focused on using simulated data to evaluate the appropriateness of the diffusion approximation for a given application and to select parameterization methods. Once this is completed, the next step is to use the diffusion approximation to estimate risk metrics using the actual time series data from the population of concern. This final step is illustrated using a 38-year time series from spring chinook in the Upper Columbia River basin (T. Cooney, *unpublished data*). To analyze this data, two approaches were used. The first focuses on point es-

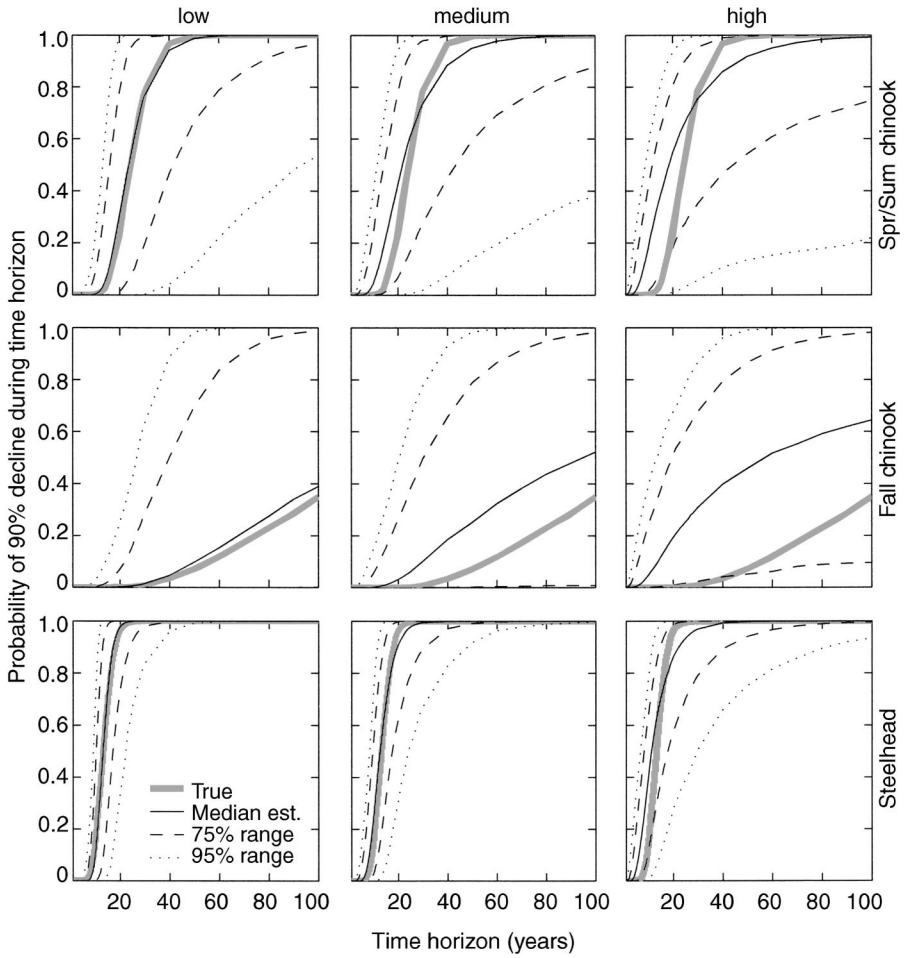

FIG. 7. Effect of sampling error on estimated vs. actual probability of 90% decline within a given time horizon. Probabilities were calculated using Eq. 6 with $\hat{\mu}_{run}$ and $\hat{\sigma}^2_{slp}$. Results are shown for simulations with low, medium, or high sampling error added to the spawner counts. Simulations were started with a stable initial age structure. The solid line shows the median errors observed in the simulations, and the dashed and dotted lines encompass 75% and 95% of the estimates, respectively.

timates and confidence intervals. The second uses a Bayesian approach and estimates the probability of the parameters and risk metrics given the data. This approach gives a measure of the data support for different potential true risk levels. Preliminary diagnostic tests for the data, namely testing $\log N_{t+1}/N_t$ for normality, outliers, and serial correlation, are not discussed here, but are reviewed in Dennis et al. (1991).

Point estimates and confidence intervals

The approach here consists of four steps: (1) select estimators and make the needed assumptions about their expected values, (2) calculate the point estimates, (3) specify their expected distributions, and (4) use parametric bootstrapping to estimate confidence intervals.

1) *Estimators and their expected values*.—The previous analyses suggest that given the expected levels of process and nonprocess error in salmonid data:

$$E(\hat{\sigma}^2_{slp}) \approx \sigma^2$$
$$E(\hat{\sigma}^2_{d}) \approx 2\sigma^2_{np} \text{ (since } \sigma^2_{np} \gg \sigma^2)$$
$$E(\hat{\mu}_{run}) = \mu. \quad (22)$$

2) *Point estimates*.—From the time series, the following estimates were calculated: $\hat{\mu}_{run} = -0.07$, $\hat{\sigma}^2_{slp} = 0.01$, and $\hat{\sigma}^2_{d} = 0.50$. Using the assumptions in Eq. 22, the point estimates of the risk metrics can be calculated. The point estimate of λ for this population is $\hat{\lambda} = \exp(\hat{\mu}_{run}) = 0.93$, indicating a 7% per year decline. The point estimate for a 90% decline within 25 years uses Eq. 6 with $\mu = \hat{\mu}_{run}$ and $\sigma^2 = \hat{\sigma}^2_{slp}$: $P(N_{25} < 0.1 \times N_0) = 0.44$.

3) *Their estimated distributions*.—The statistical distributions of the parameter estimates are themselves estimated using the known distribution of the parameter estimates (Holmes and Fagan 2002) and in place of the expected or true value of the parameters, which appear

in these distributions, using the point estimates for the parameters (from step 2):

$$\hat{\sigma}^2 \sim \text{gamma}\left(\frac{2\hat{\sigma}^2_{\text{slp}}}{df}, \frac{df}{2}\right)$$

is the estimated $\hat{\sigma}^2$ distribution given

$$\hat{\sigma}^2_{\text{slp}} \sim \text{gamma}\left[\frac{2E(\hat{\sigma}^2_{\text{slp}})}{df}, \frac{df}{2}\right]$$

and using $\hat{\sigma}^2_{\text{slp}}$ as an estimate of $E(\hat{\sigma}^2_{\text{slp}})$;

$$\hat{\sigma}^2_{\text{np}} = \frac{1}{2}\hat{\sigma}^2_{\text{d}} \sim \text{gamma}\left(\frac{\hat{\sigma}^2_{\text{d}}}{n-1}, \frac{n-1}{2}\right)$$

is the estimated $\hat{\sigma}^2_{\text{np}}$ distribution given

$$\hat{\sigma}^2_{\text{d}} \sim \text{gamma}\left[\frac{2E(\hat{\sigma}^2_{\text{dp}})}{n-1}, \frac{n-1}{2}\right]$$

and using $E(\hat{\sigma}^2_{\text{d}}) = \hat{\sigma}^2_{\text{d}}$;

$$\hat{\mu} \sim \text{normal}\left(\hat{\mu}_{\text{run}}, \frac{\hat{\sigma}^2}{n-L} + \frac{2\hat{\sigma}^2_{\text{np}}}{(n-L)^2}\right)$$

is the estimated $\hat{\mu}$ distribution given

$$\hat{\mu}_{\text{run}} \sim \text{normal}\left[\mu, \frac{\sigma^2}{n-L} + \frac{2\sigma^2_{\text{np}}}{(n-L)^2}\right] \qquad (23)$$

and using the estimated σ^2 and σ^2_{np}. The parameter n is the length of the time series, L is the number of counts summed together to form a running sum, and $df \approx 0.333 + 0.212n - 0.387L$. The distribution gamma(β,ω) denotes a gamma distribution with scale β and shape ω, a $\chi^2(\alpha)$ denotes a chi-square distribution with α degrees of freedom and a normal(a, b) denotes a normal distribution with mean a and variance b.

4) *Confidence intervals via parametric bootstrapping*.—Theconfidence intervals were estimated by generating thousands of parameter estimates from their estimated distributions (Eq. 23). First $\hat{\sigma}^2$ and $\hat{\sigma}^2_{\text{np}}$ were drawn using their estimated distributions, and then a $\hat{\mu}$ was generated using the $\hat{\sigma}^2$ and $\hat{\sigma}^2_{\text{np}}$ draws. For each parameter set, λ and $P(N_{25} < 0.1 \times N_0)$ were calculated. This was repeated 10 000 times. The 95% confidence intervals contain 95% of the estimates: λ 95% confidence intervals were (0.85, 1.04) and $P(N_{25} < 0.1 \times N_0)$ 95% confidence intervals were (0.04, 0.94).

Posterior probability distributions

The point estimates do not take into account the uncertainty in the parameter estimates, and the confidence intervals, which do, are difficult to use in a decision-making framework. For example, the 95% confidence intervals on λ include 1.0. However, this population has declined 93% over the last 38 years, and the data are most consistent with a $\lambda < 1.0$, even though it is possible for such a decline to have occurred by chance in a population with long-term nondeclining dynamics (i.e., true $\lambda > 1$). Statistical decision theory (e.g., Berger 1985) presents a framework for incorporating uncertainty in the true risk levels and the costs of different management decisions given different true risk levels. Wade (2000) and Dorazio and Johnson (2003) provide recent discussions of this Bayesian decision framework in conservation biology and resource management contexts. This framework relies on calculating the probability of different risk levels (i.e., the posterior probability distributions) given the uncertainty in the underlying true parameters, in this case in μ, σ^2, and σ^2_{np}.

Calculation of the posterior probability distribution for a given risk metric, Ψ, involves calculating the probability that the risk metric is some particular value, ϕ, given the data. This is calculated by integrating ''the probability of the data given a particular set of parameter values, θ, times the probability of that θ set'' over all sets of parameter sets, θ, for which the risk metric equals ϕ:

$$p(\Psi = \phi \,|\, \text{data}) = \int_{\text{all } \theta \text{ for which } \Psi = \phi} \frac{L(\theta \,|\, \text{data})\pi(\theta)}{\eta(\text{data})}\, d\theta \qquad (24)$$

where $p(\Psi = \phi \,|\, \text{data})$ is the probability density at $\Psi = \phi$ given the data (i.e., the posterior probability density function for Ψ), $L(\theta \,|\, \text{data})$ is the probability of the data given the parameters θ (i.e., the likelihood function), $\pi(\theta)$ is the initial assumption regarding the probability of different true process parameters (i.e., the prior), and η(data) is a normalizing constant. There are many texts on Bayesian statistics. The sections in Hilborn and Mangel (1997) on Bayesian methods are particularly accessible for ecological applications.

The following algorithm calculates $p(\Psi = \phi \,|\, \text{data})$ using $\hat{\mu}_{\text{run}}$, $\hat{\sigma}^2_{\text{d}}$, $\hat{\sigma}^2_{\text{slp}}$ as the data. Similar algorithms can be found in Hilborn and Mangel (1997:256–260). The code to run this algorithm and produce Figs. 8 and 9 is given in Supplement 1.

Step 1.—Specifyprior distributions for μ, σ^2, and σ^2_{np}. For this example, I used a uniform distribution of μ on $(-0.2, 0.2)$ and a uniform distribution of σ^2_{np} on (0, 1). As will be seen, the data provide much information on μ and σ^2_{np}, and so the posterior distributions are quite different than the priors. The same cannot be said for σ^2. My subjective prior having examined hundreds of salmonid time series, is that σ^2 is between 0.001 and 0.1. To express this, I used a prior for σ^2 of a gamma distribution with scale of 0.25 and shape of 5. I compared this with an alternate prior of a uniform distribution on (0, 0.5). This prior says the prior belief is that $\sigma^2 > 0.1$ by a factor of four to one, which seems unlikely. A standard reference prior for normal variances is uniform on log σ^2 (cf. Lee 1989: section 2.7), but for the corrupted diffusion model, $L(\sigma^2 \,|\, \text{data})$ goes to a constant as log σ^2 goes to negative infinity since $\sigma^2 = 0$ is perfectly plausible from the model's standpoint. Thus this prior will lead to infinite probability densities when Eq. 24 is integrated.

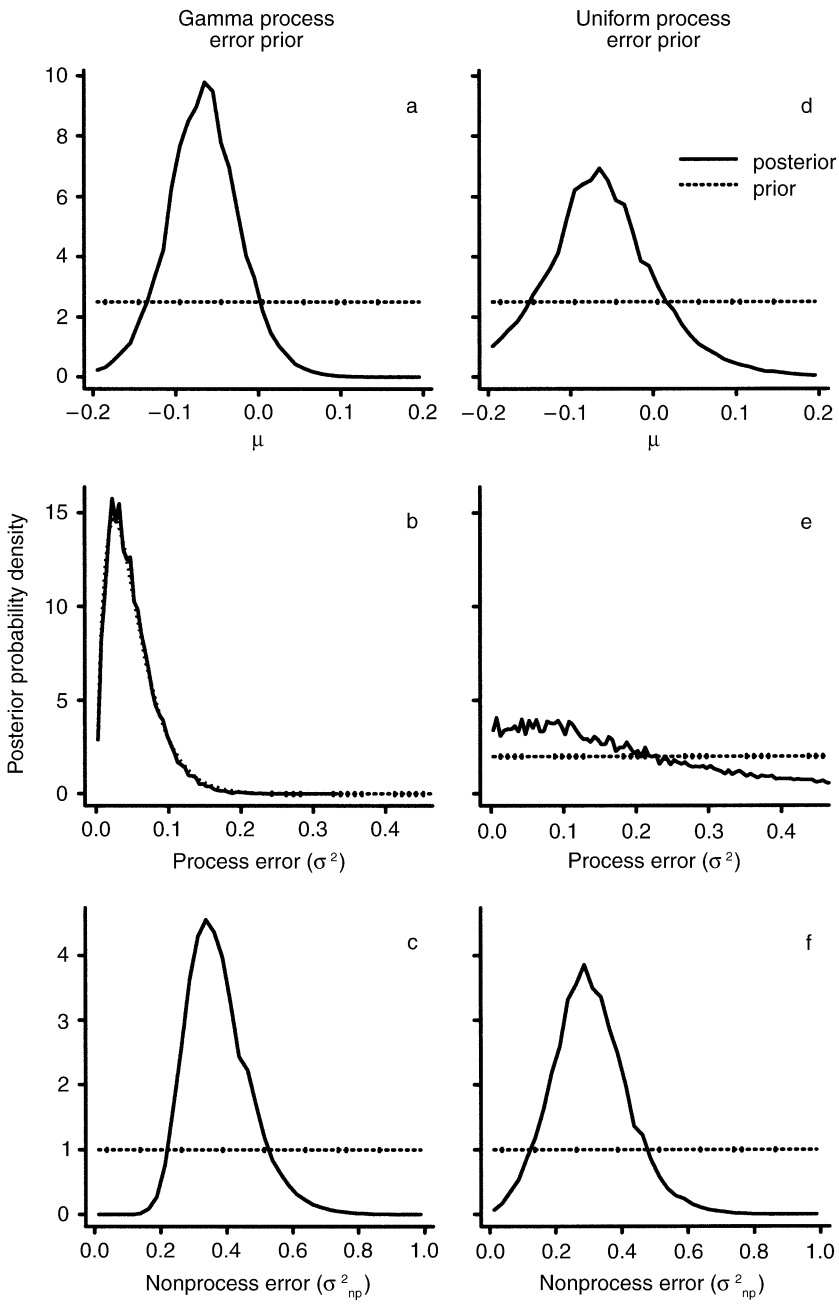

FIG. 8. Priors and estimated posterior probability distributions for μ, σ^2, and σ^2_{np}. This analysis used a 1960–1998 spawner count time series from Methow River spring chinook in the Upper Columbia River basin. The right panels show the estimated distributions using an informative prior, which put most prior probability on σ^2 between 0 and 0.1. The left panels show the estimated distributions using a uniform prior which put an even prior probability on σ^2 between 0 and 0.5.

Step 2.—Randomly draw values of μ, σ^2, and σ^2_{np} from their priors.

Step 3.—Calculate the Ψ using the values of μ, σ^2 from step 2. If the risk of interest is the probability of 90% decline in a given time frame, this means putting these parameters into Eq. 6 to calculate Ψ. If the risk metric of interest is λ, $\Psi = \exp(\mu)$.

Step 4.—Calculate the total likelihood, $\mathbf{L}$, of the μ, σ^2, and σ^2_{np} parameters generated in step 2 given the actual $\hat{\mu}_{run}$, $\hat{\sigma}^2_d$, $\hat{\sigma}^2_{slp}$ estimates from the data:

$$\mathbf{L} = L(\mu, \sigma^2, \sigma^2_{np} \mid \hat{\mu}_{run}) \times L(\mu, \sigma^2, \sigma^2_{np} \mid \hat{\sigma}^2_{slp}) \times L(\mu, \sigma^2, \sigma^2_{np} \mid \hat{\sigma}^2_d).$$

Note that $\hat{\mu}_{run}$ is independent of $\hat{\sigma}^2_d$ and $\hat{\sigma}^2_{slp}$. To simplify

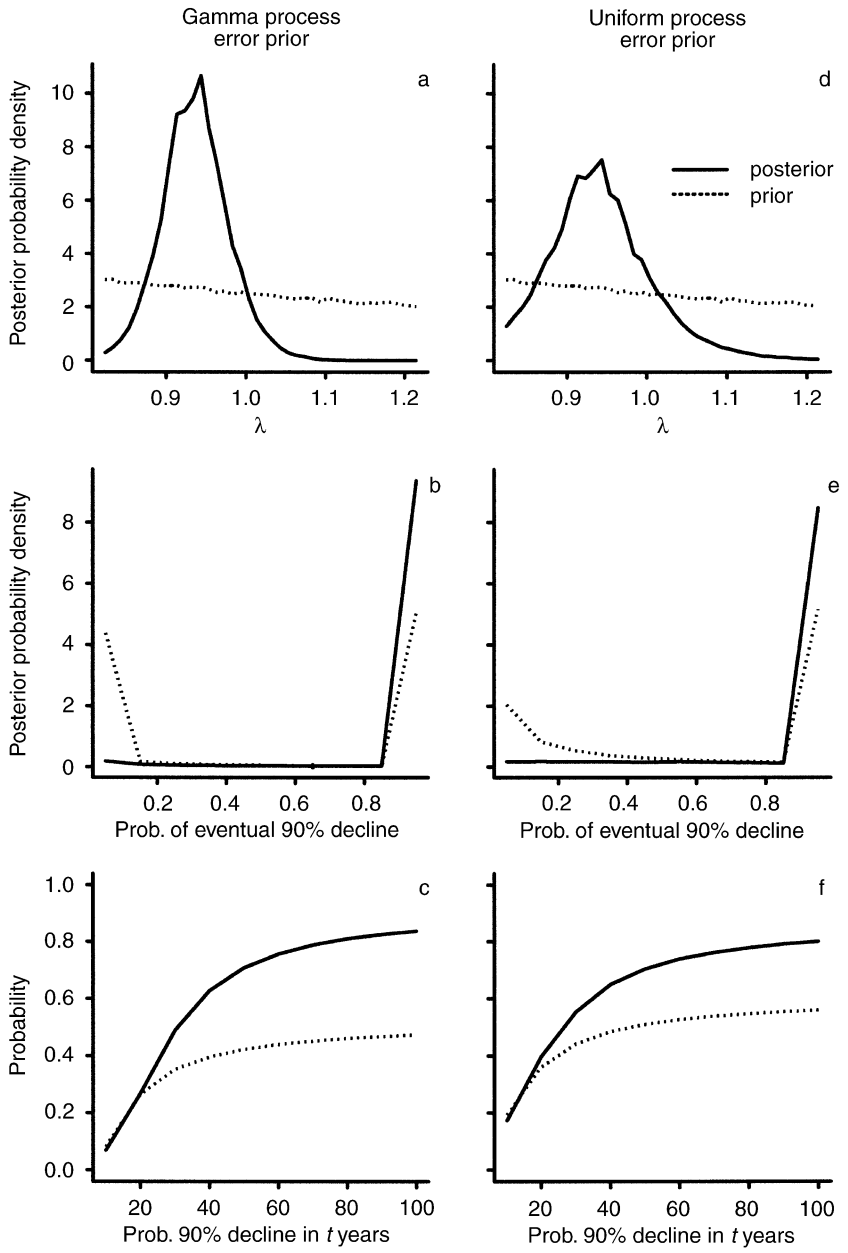

FIG. 9. Priors and estimated posterior probability distributions for λ, the probability that the population ever experiences a 90% decline at some point in the future, and the expected value of the probability of 90% decline within different time frames.

this step, I also assumed that $\hat{\sigma}^2_d$ and $\hat{\sigma}^2_{slp}$ are independent, given a particular set of parameters. This is not entirely true. In simulations, I have found correlation coefficients of ~0.2.

The likelihoods are calculated from the probability of $\hat{\mu}_{run}$, $\hat{\sigma}^2_d$, $\hat{\sigma}^2_{slp}$ given the μ, σ^2, and σ^2_{np} parameters generated in step 2. The distributions of $\hat{\mu}_{run}$, $\hat{\sigma}^2_d$, and $\hat{\sigma}^2_{slp}$ are given in Eq. 23, and from these the likelihoods are calculated. The $L(\mu, \sigma^2, \sigma^2_{np} | \hat{\mu}_{run})$ equals the probability density at $\hat{\mu}_{run}$ of a normal with mean μ and variance $(\sigma^2 + 2\sigma^2_{np})/(n - L)$. The $L(\mu, \sigma^2, \sigma^2_{np} | \hat{\sigma}^2_{slp})$ equals the probability density at $\hat{\sigma}^2_{slp}$ of a gamma with scale β and shape ω where $\beta \approx (0.5\sigma^2 + 0.15\sigma^2_{np})/\omega$ and $\bar{\omega} \approx (0.333 + 0.212n - 0.387L)/2$. Recall from Eq. 15 that $E(\hat{\sigma}^2_{slp}) \approx (0.5\sigma^2 + 0.15\sigma^2_{np})$. The $L(\mu, \sigma^2, \sigma^2_{np} | \hat{\sigma}^2_d)$ equals the probability density at $\hat{\sigma}^2_d$ of a gamma with scale β and shape ω where $\beta = (\sigma^2 + 2\sigma^2_{np})/\bar{\omega}$ and $\bar{\omega} = (n - 1)/2$. Recall the introduction of $\hat{\sigma}^2_d$ that $E(\hat{\sigma}^2_d) = (\sigma^2 + 2\sigma^2_{np})$. See any statistics text for the probability density functions for normal and gamma distributions. The gamma distribution arises given that $q\hat{\sigma}^2/E(\sigma^2) \sim \chi^2(q)$; thus $\hat{\sigma}^2 \sim \chi^2(q)E(\sigma^2)/q$, which equals

a gamma distribution with scale = $2E(\sigma^2)/q$ and shape = $q/2$.

Step 5.—Repeat steps 2–4 thousands of times, recording the parameters (step 2), the risk values (step 3), and total likelihood (step 4) for each iteration.

Step 6.—Divide the outputs of interest, e.g., risk values and parameters, into discrete intervals and calculate the proportion of the total likelihood (**L** summed over all the iterations) within each interval. Divide the probabilities by the size of the interval. This gives the posterior probability distribution for the output of interest.

Fig. 8 shows the posterior probability distributions for μ, σ^2, and σ^2_{np} relative to the priors for the 1960–1998 Methow River spring chinook data. The plots (Fig. 8d–f) show the posterior distributions using the uniform prior on σ^2. The posterior on σ^2 is not uniform and gives more weight to $\sigma^2 < 0.2$ than $\sigma^2 > 0.2$. The plots (Fig. 8a–b) show the posterior distributions using my quite informative prior on σ^2 based on previous analyses. The posterior distribution for σ^2 is the same as the prior in this case. Together, this indicates that the data are not providing much information beyond that σ^2 is more likely to be <0.2. For μ and σ^2_{np}, the data provide much more information, and the posteriors are quite different than the uniform priors. Putting a strong prior on small σ^2 decreases the uncertainty in μ and σ^2_{np}, and their posterior distributions are more peaked, but even with a prior that gives weight to higher σ^2, there is still considerable information on μ and σ^2_{np}.

Fig. 9 shows the posterior probability distributions for the following risk metrics: $\lambda = \exp(\mu)$, the probability that the population ever experiences a 90% at some point in the future (this is π' in Eq. 6), and the expected value of the probability of 90% decline within different time frames. The expected value of the probability of extinction multiplies the probability of extinction given a particular set of true parameters by the probability of those parameters (specified by the posterior probability distribution) and integrates this over all possible parameter values. See Ludwig (1996*a*) for an extensive discussion of the expected value of the probability of extinction, which he terms the Bayesian probability. A simple algorithm to calculate the expected probability is to divide 0 to 1 into regular intervals, take the value of the posterior probability density of *P*(90% decline in *x* years) from step 6 at the center of those intervals and multiply that by the width of the interval, and then sum these together to calculate the expected value of *P*(90% decline in *x* years).

Most of the area of posterior probability distribution of λ is in the region $\lambda < 1$ (Fig. 9a,d) indicating that the data give high support to long-term declining dynamics ($\lambda < 1$). The support is considerably higher when strong prior information on σ^2 is included (Fig. 9a), but is still high when less specific prior information on σ^2 is used (Fig. 9d). Fig. 9(b,e) shows the posterior distributions of the probability of eventual 90% decline. The prior (based on the parameter priors) gave roughly equal weight to a 0 or 1 probability. The posterior distribution was strongly skewed toward a probability of 1 for eventual 90% decline indicating that the data strongly support a high probability of eventual 90% decline. This occurred because either $\mu < 0$, which means 90% decline is certain eventually, or if $\mu > 0$, σ^2 was large since that is the only way we could have observed a 7% yearly decline over the last 38 years by chance with $\mu > 0$. Thus the probability of eventual 90% decline was still high. Given that data strongly support eventual probability of a 90% decline, when is it likely to occur? Fig. 9(c,f) shows the expected value of the probability of 90% decline within different time frames. Fifty percent of the population trajectories generated by sampling from the posterior distributions of the parameters would have declined 90% within 30 years. Over 65% would have declined 90% within 50 years.

Overall, λ appeared to be the most straightforward metric to use. It behaved well and is not overly sensitive to estimation of σ^2, which is imprecise. The λ metric is also transparent. If $\lambda < 1$, the population is declining; if $\lambda < 0.9$, it is collapsing. If the population needs to be recovered from low levels, λ must be increased above 1.0. Risk metrics involving probabilities are much more problematic. They are highly variable and sensitive to estimation of σ^2. This section gave an example of how to incorporate uncertainty into probability metrics and how to express that uncertainty, however much work remains to be done in this area. It is tempting to argue that because of their high uncertainty, metrics involving extinction or quasi-extinction probabilities are bad and should simply be abandoned; indeed others have argued as much. However, a measure of extinction risk is integral to the legal definition of ''jeopardy'' under the U.S. Endangered Species Act. While this is true, research into how to quantify and express the extinction risk in a way that properly characterizes what we do and do not know about the probability of population declines to critical levels is essential to science-based decision making for threatened and endangered species.

Concluding Remarks

The diffusion approximation approach has a solid foundation on theory concerning the behavior of stochastic age-structured models and opens a large toolbox of quantitative methods for linear stochastic models. In addition, the fact that many population processes with a multitude of parameters can be approximated by a three parameter model means that parameters can be estimated with increased statistical power (relative to trying to estimate a large number of parameters with limited data). However, the diffusion approximation approach makes simplifying assumptions, which will always be violated to some degree for real populations. This paper describes a method for evaluating the ap-

propriateness of the diffusion approximation for a particular PVA application and for selecting among the currently available parameterization methods. The salmonid populations modeled in this paper exhibit substantial deviations from the diffusion model, such as nonstable age structure, high temporal autocorrelation in ocean survivorship, and juvenile density dependence, yet the diffusion model does approximate the behavior of the simulated population trajectories, in terms of probabilities of crossing severe decline thresholds and of the predicted distribution of future population sizes.

The analyses also illustrate that careful estimation of the diffusion parameters is critical and that poor performance of estimators should not be confused with poor performance of the diffusion approximation. In particular, one should be cautious about assuming that nonprocess error within the data is low even if sampling error is known to be low. In the simulated salmonid populations, density dependence in the juvenile stage and perturbation of the age structure created feedbacks that led to nonprocess error variability within the data. The result was high nonprocess error variability even with no sampling error added. This led to overestimation of the process error with parameterization methods that are not designed to deal with high nonprocess error. Two estimators designed to deal with high nonprocess error were tested: a Kalman filter maximum likelihood estimator and a slope estimator. For the 20-year salmon time series with small process error, the majority of Kalman filter maximum likelihood estimates of σ^2 were essentially zero. In other words, a model with nonprocess error alone had the highest likelihood of producing the data. An alternative estimator, the slope estimator for σ^2, performed better for the simulated salmon populations and gave σ^2 estimates that were within 50–150% of the true value. For low to medium sampling error levels relative to those expected within actual salmon censuses (but still 5–50 times higher than the expected process error), $\hat{\sigma}^2_{\text{slp}}$ gave median estimates that did not appear to substantially degrade quasi-extinction estimates; however for high sampling error, σ^2 estimates were an order of magnitude too high and this did degrade the estimates. The robustness of $\hat{\sigma}^2_{\text{slp}}$ to nonprocess error due to density dependence is encouraging; similar robustness was also found in a cross-validation of diffusion-approximation methods using a large collection of real salmon time series (Holmes and Fagan 2002).

The purpose of this paper is to illustrate a method for investigating the appropriateness of diffusion approximation for a specific population and to evaluate its parameterization given the data constraints for that population. New parameterization methods are being continually refined, and no doubt better methods than those presented here will eventually be available. The methods described here can be used to evaluate and compare these new methods to existing ones. Experience suggests that the best parameterization method is application specific; that is, it may not be possible to find a method that is best for all situations. In particular, the relative performances of the different parameterization methods for the simulated salmonid time series in this paper should not be overgeneralized to their performance for other population time series. The maximum likelihood estimates assuming zero nonprocess error did especially poorly for the salmonid time series. However, a study of the performance of these estimators on a large collection of real time series found that they provided good estimates for time series of a wide variety of species, just not for salmon (Fagan et al., *unpublished manuscript*). Presumably, the level of nonprocess error in many nonsalmonid time series is not so large as to significantly bias σ^2 estimates. Similarly, maximum likelihood estimation assuming nonzero nonprocess error (the Kalman estimates) produced severe underestimates of σ^2 for the 20-year salmonid time series. However this was not a problem with longer simulations (100+ years), and for nonsalmonid population processes, it may perform well for short time series. It is also worth noting that the Kalman algorithm does estimate the parameters with the highest likelihood of producing the data, assuming all parameters are equally likely. In the real world, all parameters are not equally likely; specifically process error equaling essentially zero is unlikely given that year-to-year variability in survivorship and fecundity is occurring. Incorporating informative priors on σ^2 into the Kalman algorithm may be a way to formally avoid unrealistic parameter estimates.

Regardless of the PVA model used, a certain amount of variability in estimated parameters and risk metrics is an unavoidable aspect of the analysis of stochastic population processes, simply due to the nature of these processes. One of the strengths of DA methods is that the statistical distributions of the estimated parameters are known. As a result, the uncertainty in the estimated risks can be calculated. This is often not the case for other PVA approaches, such as Leslie matrix models or individual-based simulations, where uncertainty in the estimated model parameters is often poorly known, if known at all. Even though the uncertainty in DA risk metrics can be calculated, this uncertainty is definitely high. In this situation, examining either the likelihood functions or the posterior probability distributions for the risk metrics, rather than simply the point estimates and confidence intervals, will help to clarify the level of data support for different risks and to choose risk metrics that are most informative. Statistical decision theory (e.g., Berger 1985) provides a framework for integrating estimates of the data support for different risk levels with the consequences of different true risk levels. Taking an approach that assesses the degree of data support for questions of conservation concern emphasizes that while uncertainty does exist, being uncertain does not mean we know nothing.

Acknowledgments

I thank R. Hinrichsen, C. Jordan, P. Kareiva, P. Levin, M. McClure, P. McElhany, and B. Sanderson for helpful discussions throughout the development of these methods. T. Cooney provided the impetus and initial simulations that led to the development of the input corrections. T. Cooney, P. Kareiva, and M. Marvier supplied the salmon matrices. I also thank S. Heppell, B. Morris, and an anonymous reviewer for reviews that helped improve this paper, and J. Butzerin for her copy-editing.

Literature Cited

Achord, S., P. S. Levin, and R. W. Zabel. 2003. Density-dependent mortality in Pacific salmon: the ghost of impacts past? Ecology Letters **6**:335–342.

Berger, J. O. 1985. Statistical decision theory and Bayesian analysis. Springer Verlag, New York, New York, USA.

Caswell, H. 2001. Matrix population models. Sinauer, Sunderland, Massachusetts, USA.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. Ecological Monographs **61**:115–143.

Dorazio, R. M., and F. A. Johnson. 2003. Bayesian inference and decision theory—a framework for decision making in natural resource management. Ecological Applications **13**: 556–563.

Dunham, J., and B. Rieman. 2001. Sources and magnitude of sampling error in redd counts for bull trout. North American Journal of Fisheries Management **21**:343–352.

Ellner, S. P., J. Fieberg, D. Ludwig, and C. Wilcox. 2002. Precision of population viability analysis. Conservation Biology **16**:258–261.

Fieberg, J., and S. Ellner. 2000. When is it meaningful to estimate extinction probability? Ecology **81**:2040–2047.

Gerber, L., D. DeMaster, and P. Kareiva. 1999. Grey whales and the value of monitoring data in implementing the U.S. Endangered Species Act. Conservation Biology **13**:1215–1219.

Harvey, A. C. 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge, UK.

Heyde, C. C., and J. E. Cohen. 1985. Confidence intervals for demographic projections based on products of random matrices. Theoretical Population Biology **27**:120–153.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Hinrichsen, R. A. 2002. The accuracy of alternative stochastic growth rate estimates for salmon populations. Canadian Journal of Fisheries and Aquatic Sciences **59**:1014–1023.

Holmes, E. E. 2001. Estimating risks in declining populations with poor data. Proceedings of the National Academy of Sciences (USA) **98**:5072–5077.

Holmes, E. E., and W. F. Fagan. 2002. Validating population viability analysis for corrupted data sets. Ecology **83**:2379–2386.

Jones, E. L., T. J. Quinn, and B. W. Van Alen. 1998. Observer accuracy and precision in aerial and foot survey counts of pink salmon in a Southeast Alaska stream. North American Journal of Fisheries Management **18**:832–846.

Kareiva, P., M. Marvier, and M. McClure. 2000. Recovery and management options for spring/summer chinook salmon in the Columbia River Basin. Science **290**:977–979.

Lande, R., S. Engen, and B. Saether. 2003. Stochastic population models in ecology and conservation: an introduction. Oxford University Press, Oxford, UK.

Lande, R., and S. H. Orzack. 1988. Extinction dynamics of age-structured populations in a fluctuating environment. Proceedings of the National Academy of Sciences (USA) **85**:7418–7421.

Lee, P. M. 1989. Bayesian statistics: an introduction. Halsted Press, New York, New York, USA.

Lindley, S. T. 2003. Estimation of population growth and extinction parameters from noisy data. Ecological Applications **13**:806–813.

Ludwig, D. 1996*a*. Uncertainty and the assessment of extinction probabilities. Ecological Applications **6**:1067–1076.

Ludwig, D. 1996*b*. The distribution of population survival times. American Naturalist **147**:506–526.

Ludwig, D. 1999. Is it meaningful to estimate extinction probabilities? Ecology **80**:298–310.

McClure, M. M., E. E. Holmes, B. L. Sanderson, and C. E. Jordan. 2003. A large-scale, multi-species risk assessment: anadromous salmonids in the Columbia River Basin. Ecological Applications **13**:964–989.

Morris, W., D. Doak, M. Groom, P. Kareiva, J. Fieberg, L. Gerber, P. Murphy, and D. Thomson. 1999. A practical handbook for population viability analysis. The Nature Conservancy, Sunderland, Massachusetts, USA.

Morris, W. F., P. L. Bloch, B. R. Hudgens, L. C. Moyle, and J. R. Stinchcombe. 2002. The use of population viability analysis in endangered species recovery planning: past trends and recommendations for future improvement. Ecological Applications **12**:708–712.

Morris, W. F., and D. F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Press, Sunderland, Massachusetts, USA.

Nicholls, A. O., P. C. Viljoen, M. H. Knight, and A. S. Van Jaarsveld. 1996. Evaluating population persistence of censused and unmanaged herbivore populations from the Kruger National Park, South Africa. Biological Conservation **76**:57–67.

Sabo, J. L., E. E. Holmes, and P. Kareiva. 2004. Efficacy of simple viability models in ecological risk assessment: does density dependence matter? Ecology **85**:328–341.

Tuljapurkar, S. D. 1989. An uncertain life: demography in random environments. Theoretical Population Biology **35**: 227–294.

Tuljapurkar, S. D., and S. H. Orzack. 1980. Population dynamics in variable environments. I. Long-run growth rates and extinction. Theoretical Population Biology **18**:314–342.

Wade, P. R. 2000. Bayesian methods in conservation biology. Conservation Biology **14**:1308–1316.

Wilcox, C., and H. Possingham. 2002. Do life history traits affect the accuracy of diffusion approximations for mean time to extinction? Ecological Applications **12**:1163–1179.

Williams, J. G., S. G. Smith, and W. D. Muir. 2001. Survival estimates for downstream migrant yearling juvenile salmonids through the Snake and Columbia River hydropower system, 1966–1980 and 1993–1999. North American Journal of Fisheries Management **21**:310–317.

APPENDIX A

The Kalman filter for maximum likelihood estimation given corrupted observations is available in ESA's Electronic Data Archive: *Ecological Archives* A014-023-A1.

APPENDIX B

The development of a general method for correcting for age-specific inputs into the censused population is available in ESA's Electronic Data Archive: *Ecological Archives* A014-023-A2.

SUPPLEMENT 1

The Splus code for estimating parameters and calculating the posterior probability distributions for DA risk metrics is available in ESA's Electronic Data Archive: *Ecological Archives* A014-023-S1.

SUPPLEMENT 2

The Matlab code for running a stochastic matrix model and generating diagnostic plots is available in ESA's Electronic Data Archive: *Ecological Archives* A014-023-S2.