Ecology Letters, (2007) **10**: 1182–1198 doi: 10.1111/j.1461-0248.2007.01105.x

REVIEW AND SYNTHESIS

A statistical approach to quasi-extinction forecasting

Elizabeth Eli Holmes,[1]* John L. Sabo,[2] Steven Vincent Viscido[1] and William Fredric Fagan[3]
[1]Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, WA 98112, USA
[2]School of Life Sciences, Arizona State University, Tempe, AZ, USA
[3]Department of Biology, University of Maryland, College Park, MD 20742, USA
**Correspondence: E-mail: eli.holmes@noaa.gov*

Abstract

Forecasting population decline to a certain critical threshold (the quasi-extinction risk) is one of the central objectives of population viability analysis (PVA), and such predictions figure prominently in the decisions of major conservation organizations. In this paper, we argue that accurate forecasting of a population's quasi-extinction risk does not necessarily require knowledge of the underlying biological mechanisms. Because of the stochastic and multiplicative nature of population growth, the ensemble behaviour of population trajectories converges to common statistical forms across a wide variety of stochastic population processes. This paper provides a theoretical basis for this argument. We show that the quasi-extinction surfaces of a variety of complex stochastic population processes (including age-structured, density-dependent and spatially structured populations) can be modelled by a simple stochastic approximation: the stochastic exponential growth process overlaid with Gaussian errors. Using simulated and real data, we show that this model can be estimated with 20–30 years of data and can provide relatively unbiased quasi-extinction risk with confidence intervals considerably smaller than (0,1). This was found to be true even for simulated data derived from some of the noisiest population processes (density-dependent feedback, species interactions and strong age-structure cycling). A key advantage of statistical models is that their parameters and the uncertainty of those parameters can be estimated from time series data using standard statistical methods. In contrast for most species of conservation concern, biologically realistic models must often be specified rather than estimated because of the limited data available for all the various parameters. Biologically realistic models will always have a prominent place in PVA for evaluating specific management options which affect a single segment of a population, a single demographic rate, or different geographic areas. However, for forecasting quasi-extinction risk, statistical models that are based on the convergent statistical properties of population processes offer many advantages over biologically realistic models.

Keywords

Extinction analysis, population models, population viability analysis, stochastic estimation, stochastic models.

Ecology Letters (2007) 10: 1182–1198

> *…a typical random mass phenomenon, unpredictable in certain details, predictable in certain numerical proportions of the whole. – George Polya (1968)*

INTRODUCTION

Population models are a standard tool in both ecology and conservation biology, but these two domains often have different goals. In ecology, one typically uses population models to infer the biological mechanisms that produced some observed data. In conservation science, one constructs a population model to forecast future trends or responses; this latter approach is referred to as population viability analysis, or PVA (Boyce 1992; Beissinger & McCullough 2002; Morris & Doak 2003). The key distinction between these approaches is that one is focused on causes, while the other is focused on consequences.

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

In both endeavours increasing use has been made of parameter-rich, mechanistic models, which can range from simple age-structured models to complex individual-based models. This approach to PVA forecasting has a number of pitfalls, most importantly the need for detailed and abundant data to parameterize the model. Such high-quality data are rarely available for populations of concern to conservationists (Morris *et al.* 2002; DeMaster *et al.* 2004). Even when detailed data are available, they typically cover too short a duration to specify the annual variability of model parameters. Thus, annual variability must either be left out of the model or assumed using (hopefully) plausible parameters. Then there is also the problem of determining the uncertainty in the model parameters. As many parameters are specified rather than estimated from data, we cannot determine their uncertainty using statistical methods. At best, we can try to assess the model's sensitivity to these parameter values. Because of these limitations, some policy-makers charged with forecasting populations have recommended against quantitative PVA – at least for the typical endangered or threatened species (DeMaster *et al.* 2004).

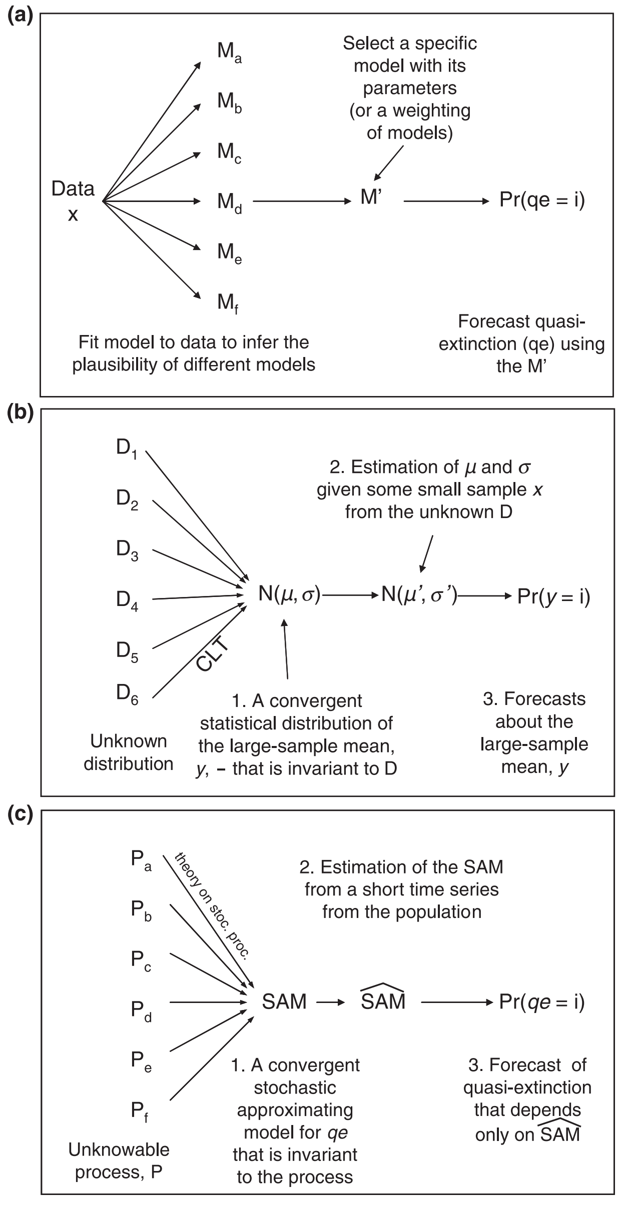

Figure 1 Reasoning using statistical vs. mechanistic modelling. (a) In a mechanistic paradigm, the model is meant to mimic the data observed. The data are thus used to choose the model via some formal or *ad hoc* selection method. Once a model is selected its parameters are estimated, often from the same data. (b) A familiar example of statistical reasoning is inference concerning a large sample mean. The Central Limit Theorem (CLT) for independent random variables says that the means of large samples converge to the normal distribution. Thus, in this case a theory concerning the statistical properties of samples is used to specify the model. Although the large-sample distribution is known, its parameters μ and σ^2 must be estimated from a small finite sample. The CLT also specifies the relationship between small samples and the distribution of large-sample means. (c) The basic steps of building a statistical model for PVA (1) theory is used to determine the common stochastic approximating model that emerge from diverse processes, (2) the theoretical stochastic model is estimated from the data and (3) a forecast is made using the estimated model.

In this paper, we argue for a different approach to forecasting, one based on the convergent statistical properties of stochastic population processes. As Polya notes in the opening quote, certain average properties of stochastic processes may be predictable even when the details of the underlying process are unpredictable and/or unknown. The key is to find statistical properties that are convergent over a broad class of plausible population dynamics. Convergent statistical properties are at the foundation of classical statistical inference. However, this type of reasoning is very different, both philosophically and practically, from inference based on a biologically realistic mechanistic model (Fig. 1a). To illustrate the distinction, we will discuss a well-worn example from statistics: the inference of large sample means (Fig. 1b).

The first step is to identify the convergent properties of the statistic of interest, assuming that some convergence exists. We avoid making a mechanistic model for the data. Instead, we consider the properties of our statistic: here the statistic is the sum of *n* random variables divided by *n*. The Central Limit Theorem (CLT) tells us that regardless of the underlying data distribution, the distribution of large-sample means is normal. This distribution is a convergent statistical property of the mean itself – although the process

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

that generated the data is unknown, the distribution of large sample means is known. Next, we need to address the problem of estimation – how can we infer the distribution of the large-sample mean given data from a small sample? Here, we rely on estimation theory and we need to be cognizant that our data may well be non-normal, unlike the large-sample mean. If so, we should use appropriate methods to adjust for the bias introduced by this non-normality. In the final step, we use the results from the first two steps to estimate the distribution and make a forecast about the large-sample mean.

Our goal is to extend this type of reasoning to the estimation of a specific property of stochastic population trajectories: the probability of decline below a pre-defined threshold. As this metric does not measure absolute extinction *per se*, we use the term quasi-extinction (Ginzburg *et al.* 1982). Quasi-extinction probabilities are a widely used risk metric by conservation and management organizations. For example, the World Conservation Union's IUCN risk criteria (Mace & Lande 1991) and the proposed quantitative criteria for the U.S. Endangered Species Act (DeMaster *et al.* 2004) rely on quasi-extinction probabilities. For this paper, we consider only quasi-extinction thresholds above the level at which demographic stochasticity and Allee effects become important. These factors accelerate the decline toward absolute extinction, which differs in many fundamental ways from the decline to a critical threshold (Gilpin & Soulé 1986; Lande *et al.* 2003; Fagan & Holmes 2006).

The present work synthesizes two bodies of theoretical research (Fig. 1c): the stochastic convergence properties of population processes and the estimation of stochastic models. We begin by identifying a simple stochastic model – a stochastic exponential overlaid with Gaussian errors – that approximates the relationships between quasi-extinction risk, threshold and forecast length. We then show that the quasi-extinction properties of this three-parameter model are similar to those of more complex, parameter-rich processes. After reviewing the theoretical basis for this approximation, we present two cross-validation studies of quasi-extinction forecasts. The first is based on simulations of three cyclic population processes, and the second is based on an analysis of time series from species of conservation concern. We conclude by discussing the merits of simple stochastic approximations in conservation risk analysis, *vis-à-vis* the important criticisms that have been raised against PVA models.

STOCHASTIC APPROXIMATIONS FOR QUASI-EXTINCTION

In this paper, we focus on estimating the expected probability of quasi-extinction (abbreviated qe):

$$\mathrm{E}[\Pr(\mathrm{qe})] = \sum_{\{\ldots,x_{-2},x_{-1},x_0\}\in\Omega} P(\mathrm{qe\ in}\{x_1,x_2,\ldots,x_T\}|\\ \{\ldots,x_{-2},x_{-1},x_0\})P(\{\ldots x_{-2},x_{-1},x_0\}) \quad (1)$$

In this equation, $\{\ldots, x_{-2}, x_{-1}, x_0\}$ are the past (observed) population trajectories and $\{x_1, x_2, \ldots, x_T\}$ are future population counts forming a forecast of length T. $P(\{\ldots, x_{-2}, x_{-1}, x_0\})$ is thus the probability of observing the 'past', and $P(\mathrm{qe}\,|\,\{\ldots, x_{-2}, x_{-1}, x_0\})$ is the probability of 'future' quasi-extinction given that past. To calculate the expected quasi-extinction probability, we sum over the set Ω of all possible past trajectories. E[Pr(qe)] is thus the *expected* probability of quasi-extinction observed by selecting a past trajectory randomly from the past trajectories that particular process could produce. It measures the propensity of a population process to produce quasi-extinctions.

Stochastic approximations for density-independent processes

Most theoretical results regarding density-independent population processes are based on random walks. A simple random walk starts at some Y_t, and at each time step 'hops' to the value $Y_{t+1} = Y_t + \xi$, where ξ is a random variable. Our starting point for population modelling is the stochastic exponential growth random walk (Lewontin & Cohen 1969):

$$\log N_{t+1} = \log N_t + \mu + \eta_t. \quad (2)$$

The mean population growth rate is μ and the random variable η_t represents year-to-year deviations from that mean. In this model, the deviations might be drawn from any smooth distribution with mean zero and finite variance σ^2. Autocorrelations in η_t are permitted, provided that the correlation between $\eta_{t+\tau}$ and η_t goes to zero as τ increases. Throughout this paper, we assume that any such correlation drops off fairly rapidly. If this were not the case, the stochastic properties would take a very long time to converge.

Although the distribution of η_t in the stochastic exponential is not specified, the behaviour of log $N_{t+\tau}/N_t$ converges for large τ: the mean and variance of log $N_{t+\tau}/N_t$ both scale linearly with τ and log $N_{t+\tau}/N_t$ has a normal distribution (section 14.3 in Caswell 2000). Thus, for large τ the stochastic exponential model (eqn 2) converges to a stochastic exponential growth model with Gaussian errors:

$$\log N_{t+1} = \log N_t + \mu + \varepsilon_t, \quad (3)$$

where ε is drawn from a normal distribution with mean zero and variance σ_{b}^2; the 'b' subscript refers to brown noise which is the term for noise in which the variance grows linearly with τ. If the errors in eqn 2 are not correlated (i.e. are i.i.d.) then $\sigma_{\mathrm{b}}^2 = \sigma^2$, but otherwise generally $\sigma_{\mathrm{b}}^2 > \sigma^2$. We will refer to this as the SEG model, which stands for Stochastic Exponential growth with Gaussian errors.

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

The SEG is a good approximation to several important types of population models. One of the most important of these is stochastic, age-structured population models with no density dependence:

$$\begin{bmatrix} N_{1,t+\tau} \\ N_{2,t+\tau} \\ N_{3,t+\tau} \\ \vdots \\ N_{k,t+\tau} \end{bmatrix} = \mathbf{A}_t \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ N_{3,t} \\ \vdots \\ N_{k,t} \end{bmatrix}. \tag{4}$$

$\mathbf{A}_t$ is the stochastic transition matrix for time t; in the context of eqn 4, it is known as the Leslie matrix. The age-specific fecundities and survival rates in $\mathbf{A}_t$ are random variables, which might have some year-to-year autocorrelation. As long as the *distributions* of these vital rates are reasonably smooth and stationary (meaning that their long-term means are not changing), any weighted sum $N_t = \sum w_i N_{i,t}$ can be approximated by the SEG model at large τ (Tuljapurkar & Orzack 1980; Tuljapurkar 1989). This will be true regardless of the specific distributions and temporal correlations adopted for the fecundity and survival rates (again, assuming that any autocorrelation goes to zero on long timescales). The approximation performs best when w_i is the reproductive value of age i– i.e. the left eigenvector of $\mathbf{A}_t$ (Engen *et al.* 2007).

Figure 2 shows an example using a Chinook salmon (*Oncorhynchus tshawytscha*) population model, following Holmes (2004). The model is specified by five classes: individuals aged 1–4 years, and returning spawners. The stochastic transition matrix is

$$\mathbf{A}_t = \begin{bmatrix} 0 & 0 & 0 & 0 & ps_1 m\varepsilon_{1,t} \\ s_2\varepsilon_{2,t} & 0 & 0 & 0 & 0 \\ 0 & s_o\varepsilon_{3,t} & 0 & 0 & 0 \\ 0 & 0 & (1-b_4)s_o\varepsilon_{4,t} & 0 & 0 \\ 0 & 0 & b_4 s_o\varepsilon_{4,t} & b_5 s_o\varepsilon_{5,t} & 0 \end{bmatrix}, \tag{5}$$

where s_i is the age-specific survival, b_i is the fraction of class i(=4 or 5) that enters the spawner class, m is the average female fecundity, p is the probability of survival during migration and $\varepsilon_{i,t}$ represents the year-to-year variability. Figure 2a shows a collection of simulated 50-year time series for the total population of this model. In this example, the total population is taken as $N_t = \sum w_i N_{i,t}$, where w_i is the reproductive value of age i (Engen *et al.* 2007). Each simulation was run for 100 years before we began plotting the trajectories; this 'burning in' is a common numerical method for starting each simulation at a randomly selected point in the stationary distribution of population abundances.

The total population exhibits the key properties of a SEG process. The mean and variance of log $N_{t+\tau}/N_t$ scale linearly with τ, and the variance passes through the origin as expected (Fig. 2c). Figure 2e shows that a two-parameter SEG model closely approximates the quasi-extinction risk of the salmon population for various thresholds and forecast lengths. SEG approximations for many other age-structured population processes can be found in Tuljapurkar & Orzack (1980); Lande & Orzack (1988); Holmes (2001); Morris & Doak (2003); and Holmes (2004).

The SEG also provides a good asymptotic approximation to stochastic metapopulation models with continuous, density-independent dynamics (Holmes & Semmens 2004). In these models, each local population i has its own stochastic growth function. In addition, each population has a stochastic dispersal function describing migration to and from other populations:

$$\begin{aligned} N_{i,t+1} &= \text{growth} - \text{dispersal out} + \text{dispersal in} \\ &= N_{i,t}e^{z_{i,t}} - d_{i,t}N_{i,t}e^{z_{i,t}} + \sum_{j\neq i} \alpha_{ji,t}d_{j,t}N_{j,t}e^{z_{j,t}} \cdot \end{aligned} \tag{6}$$

The variable $z_{i,t}$ is the stochastic growth of local population i in year t, and can be drawn from any smooth statistical distribution. Some fraction $d_{i,t}$ of individuals emigrates from local population i at year t (similarly $d_{j,t}$ individuals emigrate from local population j). No constraints need be placed on the distributions of the growth and dispersal parameters (z, d and α) or their temporal correlations. Equation 6 is quite general, allowing for sources ($z_i > 0$), sinks ($z_i < 0$), dispersal sources and dispersal targets. The model also allows for anisotropy in the dispersal, as long as all sites are connected to ensure that the transition matrix $\mathbf{A}$ (eqn 8) is ergodic.

This metapopulation model can be written succinctly as

$$\begin{bmatrix} N_{1,t+1} \\ N_{2,t+1} \\ N_{3,t+1} \\ \vdots \\ N_{k,t+1} \end{bmatrix} = \mathbf{A}_t \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ N_{3,t} \\ \vdots \\ N_{k,t} \end{bmatrix}, \tag{7}$$

where the matrix $\mathbf{A}_t$ encapsulates both dispersal and local growth:

$$\mathbf{A}_t = \begin{bmatrix} (1-d_1)e^{z_1} & \alpha_{21}d_2e^{z_2} & \alpha_{31}d_3e^{z_3} & \ldots & \alpha_{k1}d_ke^{z_k} \\ \alpha_{12}d_1e^{z_1} & (1-d_2)e^{z_2} & \alpha_{32}d_3e^{z_3} & \ldots & \alpha_{k2}d_ke^{z_k} \\ \alpha_{13}d_1e^{z_1} & \alpha_{23}d_2e^{z_2} & (1-d_3)e^{z_3} & \ldots & \alpha_{k3}d_ke^{z_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{1k}d_1e^{z_1} & \alpha_{2k}d_2e^{z_2} & \alpha_{3k}d_3e^{z_3} & \ldots & (1-d_k)e^{z_k} \end{bmatrix}. \tag{8}$$

The t subscripts on the d, α and z parameters have been dropped to remove clutter; the subscript on the matrix symbol reminds us that its elements are time-dependent. A projection of this model forward in time results in a product of ergodic random matrices (the $\mathbf{A}_t$ values), whose

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

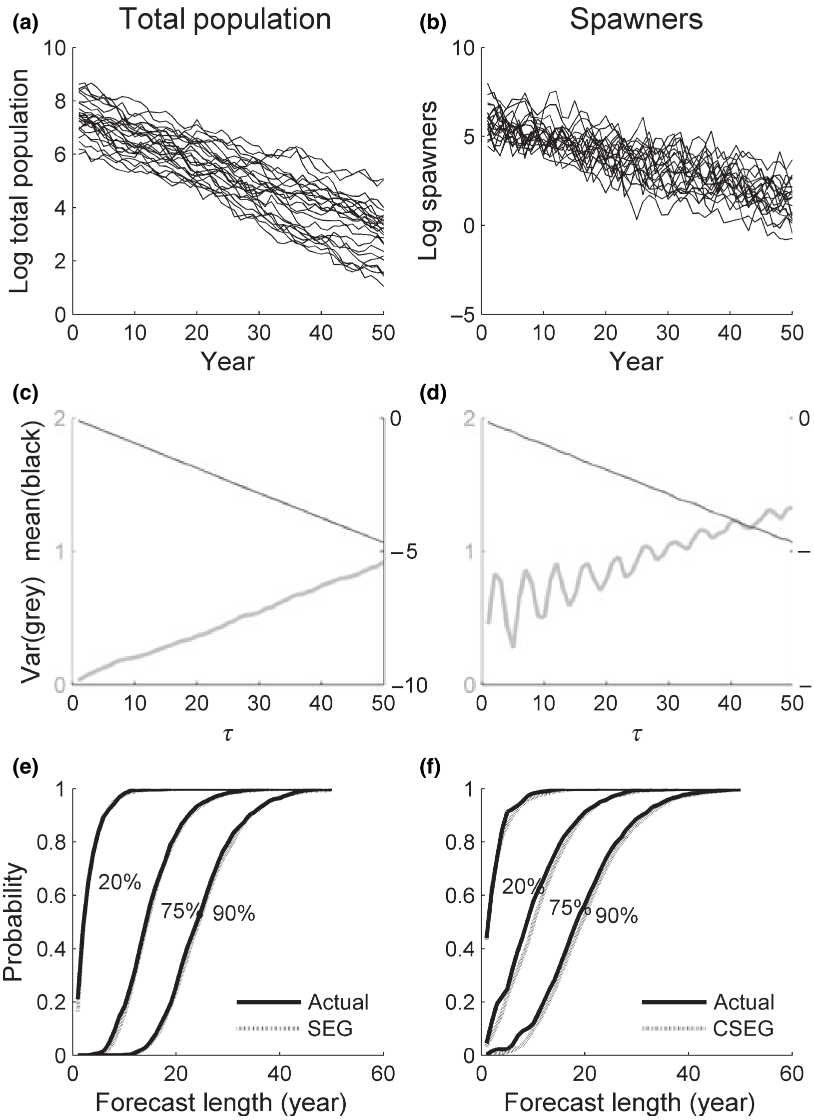

Figure 2 Example of stochastic approximations of quasi-extinction risk based on the total population and spawner counts in Chinook salmon. (a, b) Multiple realizations of the salmon population model: the log of the total population (a) and spawners (b). (c, d) Variance and mean of log $N_{t+\tau}/N_t$, where N_t is the total population at time t (c) and log $N_{5,t+\tau}/N_{5,t}$, where $N_{5,t}$ is the spawner count at time t (d). Both measures are plotted as a function of τ. These panels show the linear scaling of mean and variance. (e) Actual probabilities of the total population declining by 20%, 75% and 90% as a function of forecast length (solid black lines), and the corresponding predictions from the SEG approximation (grey lines). (f) As in (e), but for spawner counts and using the CSEG approximation instead. Because the variance has a non-zero intercept (d), a SEG approximation will not fit the spawner quasi-extinction surface. Parameter values for the salmon matrix (eqn 5) are as follows: $P = 0.4815$; $s_1 = 0.018$; $s_2 = 0.044$; $s_o = 0.8$; $b_3 = 0$; $b_4 = 0.216$; $b_5 = 1$; $m = 2747$; $\varepsilon_{i,t} \sim \exp(\text{Normal}(0, \sigma_i^2))$ with $\sigma_0^2 = 0.02$, $\sigma_1^2 = 0.13$, $\sigma_2^2 = 0.08$. The correlation coefficient between ε_1 and ε_2 was $r = 0.2$, and the correlation between ε_0 values for all age classes was $r = 0.8$. The latter represents ocean survival.

convergence properties have been well studied (Furstenberg & Kesten 1960). The theory in the cited work shows that the logarithm of the total metapopulation $\log M_{t+\tau} = \log \sum N_{i,t+\tau}$ is asymptotically normal, with a variance and mean that scale linearly with τ (Holmes & Semmens 2004):

$$\log M_{t+\tau}/M_t \underset{t\to\infty}{\rightarrow} \text{normal}(\tau\mu, \tau\sigma_b^2). \qquad (9)$$

As before, these are the key properties of a SEG model. Figure 3 shows the quasi-extinction risks for this model in a simulation with 49 heterogeneous sites connected by directional dispersal rates (of 5–25% per year). The SEG model again provides a good approximation (Fig. 3d). Note that the sites were assigned different propensities to produce dispersers and different local growth rates.

For density-independent processes, the SEG is the approximating stochastic model in the sense of Fig. 1c– it is the process to which these processes asymptotically converge and provides a reasonable approximation to qe for multiple types of models. However, we also know from large databases of real population trajectories that variance does not generally grow linearly even over time periods of 20–50 years. We address this difficulty by generalizing the SEG in the next section.

The nature of variability in real population time series

Variance patterns in population time series have been studied extensively as part of an ongoing debate over the nature of population regulation. The variance, patterns can be discussed in terms of the three general classes shown in Fig. 4 (Pimm & Redfearn 1988; Halley & Kunin 1999; Inchausti & Halley 2002): (i) *brown noise*, grows linear in τ and

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

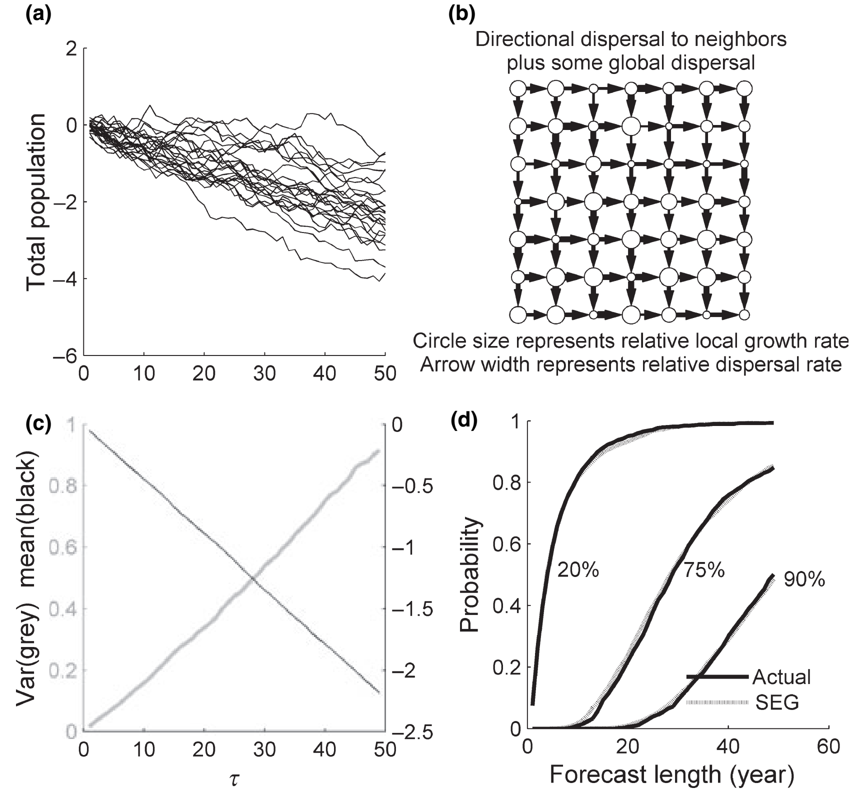

Figure 3 Example of a SEG model used to approximate the quasi-extinction risk in a stochastic metapopulation model. (a) Multiple simulated trajectories of the log total metapopulation. (b) Diagram of the metapopulation structure in this example. Each site disperses a different fraction of its population each year (the relative dispersal rates are indicated by the size of the arrows). Eighty percent of dispersers go to neighbours to the east and south, while the remaining 20% are distributed equally among all other sites in the metapopulation. Sites vary in their intrinsic population growth rates, but all are declining. Growth rates range from 0.99 to 0.80. Yearly growth rates vary (SD = 0.1) and are correlated between sites ($r = 0.9$). (c) The variance and mean of log $M_{t+\tau}/M_t$, where M_t is the total metapopulation count at time t. (d) Actual probabilities of declining by 20%, 75% and 90% as a function of forecast length (black lines) vs. those predicted from the appropriate SEG model (grey lines).

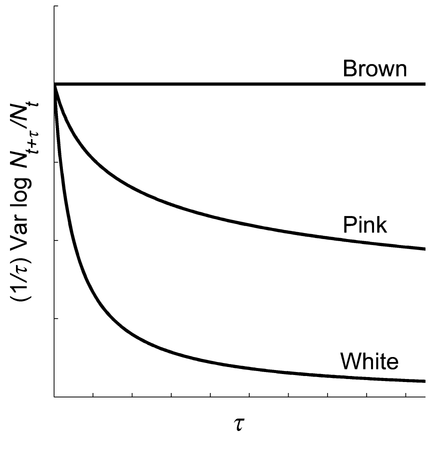

Figure 4 The three basic patterns for the $(1/\tau) \times$ variance of log $N_{t+\tau}/N_t$. Brown noise is the variance pattern typical for density-independent random walks; variance grows linearly and so $(1/\tau)$ var log $N_{t+\tau}/N_t$ is constant. In white noise, variance is constant and thus $(1/\tau)$ var log $N_{t+\tau}/N_t$ falls off geometrically with $(1/\tau)$. A white noise pattern is produced by simple error processes, $N_t = \varepsilon_{w,t}, N_{t+\tau} = \varepsilon_{w,t+\tau}$, where the $\varepsilon_{w,t}$ are drawn independently and identically from some distribution. Tightly regulated population processes will also show white noise patterns. Pink noise has an intermediate pattern. Variance does grow with τ (unlike for the white noise process), but it grows less quickly with large τ vs. small τ.

$(1/\tau)$ var log $N_{t+\tau}/N_t$ is constant; (ii) *white noise*, is constant with τ and $(1/\tau)$ var log $N_{t+\tau}/N_t$ declines geometrically; and (iii) *pink noise*, which is intermediate and grows faster at small τ, then slower as τ increases. Population time series typically exhibit a pink noise pattern falling somewhere between the pure SEG process and white noise, a fact that was first explicitly documented in the late 1980s (Pimm & Redfearn 1988; Pimm 1991). More recently, this result has been confirmed and its scope expanded by analyzing the thousands of long-term time series in the Global Population Dynamics Database (GPDD) (Inchausti & Halley 2001, 2002). Many different mechanisms can cause the intermediate pink noise pattern (see Inchausti & Halley 2002). Here, we focus on two ubiquitous processes – density-dependence and age-structure cycles – that cause a pink noise pattern, and we illustrate a simple stochastic approximation that can model the quasi-extinction surface for these processes.

This stochastic approximation is a simple extension to the SEG model, which we refer to as a Corrupted Stochastic Exponential with Gaussian errors (CSEG):

$$\log X_{t+1} = \log X_t + \mu + \varepsilon_{b,t} \tag{10a}$$

$$\log N_{t+1} = \log X_{t+1} + \varepsilon_{w,t+1}, \tag{10b}$$

where the $\varepsilon_{b,t}$ are normally distributed errors with mean 0 and variance σ_b^2 and the $\varepsilon_{w,t}$ are normally distributed errors

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

with mean β and variance σ_w^2. The CSEG is a density-independent state-space model. The population counts N_t are modelled as a SEG (X_t) overlaid with independent errors. The variance of log $N_{t+\tau}/N_t$ still scales linearly with τ, but has a non-zero intercept because of σ_w^2. This simple extension of the SEG produces a pink noise pattern. Biologically realistic values of σ_b^2 and σ_w^2 will produce levels of pink noise similar to those seen in the GPDD (Akçakaya *et al.* 2003).

In the following sections, we discuss two common population dynamics that result in a nonlinear variance pattern, and illustrate how the CSEG model can approximate their quasi-extinction probabilities. However, we should also note that the most mundane explanation for pink noise in the GPDD data is simply the presence of independent measurement errors (Akçakaya *et al.* 2003). The CSEG has also been studied extensively in this context (Holmes 2001; Holmes & Fagan 2002; Lindley 2003; Holmes 2004; Staples *et al.* 2004), where it is used as a method of separating the dynamic population variance from observation errors.

Stochastic approximations for age- or stage-specific counts

A stochastic age-structured process does not asymptotically approach a unique, stable age distribution. Rather, random perturbations cause the age structure to wander within a stochastic equilibrium (Cohen 1977). The number of individuals at age *i* fluctuates about some mean value, and those fluctuations are normally distributed (Tuljapurkar & Orzack 1980; Tuljapurkar 1989). If we consider the total population weighted by reproductive value, we are essentially averaging over these fluctuations. When we are interested in the quasi-extinction properties of a particular age, or if the available data are restricted to a particular age, these fluctuations introduce additional noise into the linear random walk and a simple SEG approximation can no longer be used.

Spawner counts from the Chinook salmon model (eqn 5) are an example of this problem. The life history of this species produces strong cycles in the spawner counts (Fig. 2b). When the variance in log $N_{5,t+\tau}/N_{5,t}$ is plotted vs. τ (Fig. 2d), we see the same linear trend displayed by the total population (Fig. 2c) – but with a non-zero intercept because of the spawner cycles. The distribution of log $N_{5,t+\tau}/N_{5,t}$ (not shown) is approximately normal for large τ, as predicted by theory. The statistical properties of the (logarithmic) spawner counts are thus the same as those of a CSEG. Figure 2f shows that the CSEG model successfully describes quasi-extinction probabilities for spawner counts across different threshold levels and forecast lengths. More CSEG approximations of age-specific counts in sea turtles, petrels, and other salmonids can be found in works by Holmes (2001, 2004).

Stochastic approximations for regulated population processes

A myriad of both intra- and interspecific forces can lead to density dependence and ultimately population regulation. In this section, we will refer to these generically as density-dependent processes. Stochastic density-dependent processes differ in fundamental ways from density-independent processes. The most basic difference is that in a stochastic density-dependent process, the population size asymptotically approaches a stationary distribution whereas in a density-independent process it does not (Dennis & Costantino 1988; Smitalova & Sujan 1992; Diserud & Engen 2000; Ives *et al.* 2003; Dennis *et al.* 2006). The variance in log $N_{t+\tau}/N_t$ is also nonlinear, exhibiting a pink noise pattern in which the variance levels off for large τ (Akçakaya *et al.* 2003). The degree of 'pinkness' depends on the strength of the density dependence. If the density dependence is weak, the variance is approximately linear in τ. Indeed, it has often been noted that weakly density-dependent processes are difficult to distinguish from a SEG (den Boer & Reddingius 1989; den Boer 1991; Dennis & Taper 1994) and a recent study illustrated that quasi-extinction for such processes can be approximated by a SEG (Staples *et al.* 2005). When density dependence is strong, on the other hand, the variance in log $N_{t+\tau}/N_t$ levels off quickly with τ and resembles white noise – although it will not necessarily be Gaussian (Diserud & Engen 2000).

The CSEG can imitate this range of variance patterns by varying the ratio of σ_b^2 to σ_w^2, and Akçakaya *et al.* (2003) showed that the level of pink noise in the GPDD can be described parsimoniously by a simple CSEG. To illustrate the CSEG approximation, consider a stochastic model with Ricker density dependence (Sabo *et al.* 2004):

$$\log N_{t+1} = \log N_t + r - |r|N_t/K + \varepsilon_t, \tag{11}$$

where ε_t is a normally distributed variable with mean zero and a process error variance of σ_p^2, r is the intrinsic rate of increase, K is the carrying capacity and N_t is the total population size. We examine two specific cases: one with weak density dependence ($r = 0.02$), and one with strong density dependence ($r = 0.2$). In both examples, K is set to 1000, σ_p^2 is 0.01 and N_0 is drawn randomly from the stochastic equilibrium.

Figure 5a and b shows sample realizations of the simulations with weak and strong density dependence respectively. Figure 5c shows that the variance in log $N_{t+\tau}/N_t$ slowly plateaus in our weak density dependence example. In the model with strong density dependence, however, the variance plateaus quickly (Fig. 5d).

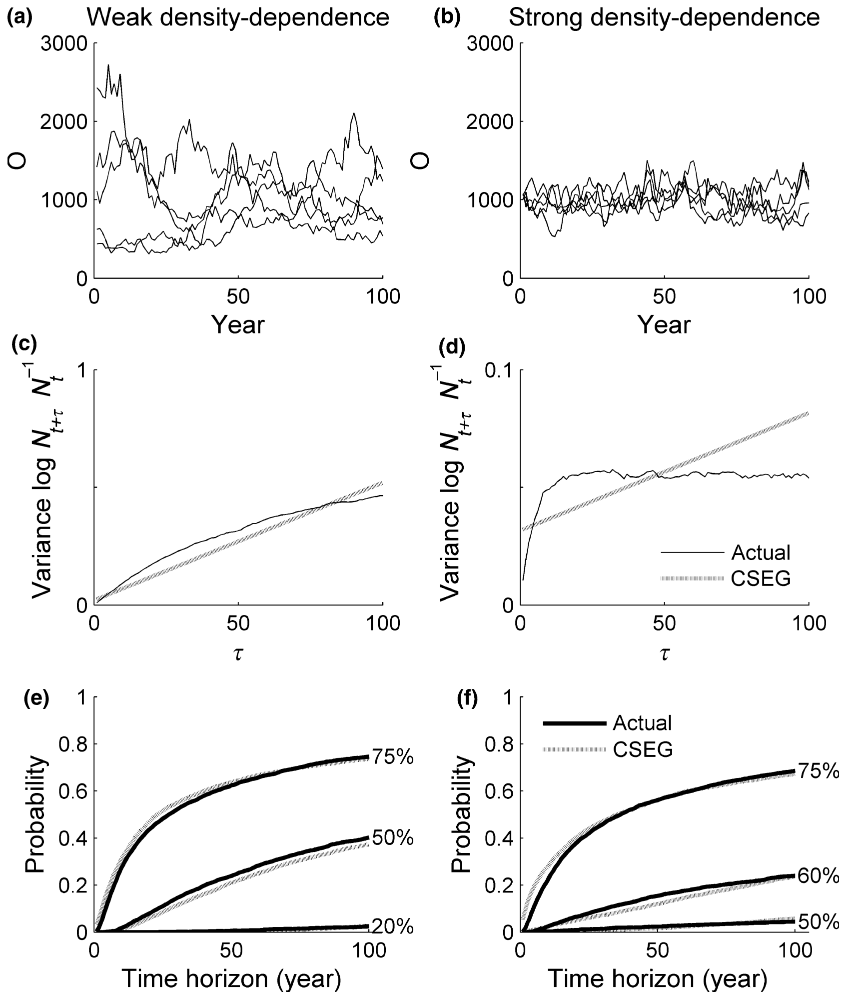

Figure 5 CSEG approximations to the quasi-extinction risk for two stochastic Ricker simulations, with weak and strong density dependencies. (a,b) Multiple realizations of each process. (c,d) The actual variance as a function of τ (black line) and the CSEG approximation (grey line). (e,f) Actual and CSEG quasi-extinction probabilities vs. forecast length for three different thresholds. The CSEG parameters are (left panels) $\mu = 0$, $\sigma_b^2 = 0.005$, $\sigma_w^2 = 0.01$; (right panels) $\mu = 0$, $\sigma_b^2 = 0.0005$, $\sigma_w^2 = 0.0158$.

The variance of the best CSEG approximation is also shown in these panels. Figure 5e and f compares the quasi-extinction probabilities of the CSEG approximation with that observed in the stochastic Ricker simulations. These panels demonstrate that the CSEG can successfully approximate the real two-dimensional quasi-extinction surface in both cases. Sabo & Gerber (2007) provide other examples of CSEG approximations to populations regulated by predator–prey interactions. In the following section, we discuss another example: a regulated, multi-species community.

ESTIMATED VS. THEORETICAL STOCHASTIC APPROXIMATIONS

The previous sections have discussed the convergent statistical properties of several important classes of population processes, illustrating the power of a simple stochastic model (the CSEG) to approximate quasi-extinction probabilities. This completes the first step of our statistical inference outline (Fig. 1c): we have selected a single approximating model that is valid for broad classes of possible generating processes. Because the CSEG model was originally developed to address the problem of measurement errors in census data (Holmes 2001; Holmes & Fagan 2002; Lindley 2003; Holmes 2004; Staples *et al.* 2004), this approximation will also cope with noise introduced by inaccurate population censuses.

While the previous sections focused on the existence of a theoretical CSEG approximation, this section focuses on the estimation of this theoretical CSEG from population count data. This is the second step illustrated in Fig. 1c. We will use two cross-validation studies, both of which focus on forecasting 80% population declines. The first study uses simulated data from three different cyclic population processes, which were chosen to illustrate strong population feedback. The noise generated by the feedback tends to mask the underlying random walk properties. Other studies have already examined the CSEG approximation for non-cyclic processes (Holmes 2001; Morris & Doak 2003; Staples *et al.* 2004). The goal of our second cross-validation

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

study is to test CSEG forecasts on real time series data for species of conservation concern.

Parameter estimation for the CSEG

The CSEG is a state-space model with Gaussian errors, and as such can take advantage of many well-developed estimation methods. One flexible maximum-likelihood method commonly used for state-space models is the Kalman filter (Shumway & Stoffer 1982; Harvey 1989; de Valpine & Hastings 2002). Lindley (2003) and Holmes (2004) illustrated its use with the CSEG. The Kalman algorithm easily deals with missing values, and allows for the incorporation of supplemental data. The extended Kalman filter allows for non-Gaussian errors as well (de Valpine & Hastings 2002; de Valpine 2002).

Other parameter estimation approaches are also available. For example, the restricted maximum likelihood (REML) method corrects the finite-sample bias common to ML parameter estimates (Staples *et al.* 2004). Standard statistical packages can be used for REML estimation (Dennis *et al.* 2006). The 'slope method' (Holmes 2001) estimates σ_b^2 simply by performing a linear regression on the variance with respect to τ . In our experience, the slope method is the most robust method when the data are generated from a strongly cyclic process, but this comes at the cost of an increased bias.

In the following cross-validations, we use the slope estimation method, because the Kalman filter and REML methods produced imprecise estimates for our strongly cyclic data. The Kalman filter was often unable to separate σ_b^2 and σ_w^2 even with 20–30 years of our simulated data. This problem was likely caused by autocorrelations in our simulated data; an extended Kalman filter might be able to correct this problem.

The principal challenge of CSEG parameterization lies in the fact that the actual data may differ from the CSEG model in very substantial ways. The CSEG is an approximation to model the quasi-extinction surface – it is not the underlying model of the data. Despite the existing work on estimation of CSEG approximations, cited above, there is still much research to be done on robust estimation methods. Research on incorporating age-structure information is helping in this regard (Hinrichsen 2002; Engen *et al.* 2007), but we also need to understand how to include non-quantitative life-history information, and how to adjust for autocorrelations generated by population dynamics.

Simulation studies of CSEG forecasts for noisy non-SEG processes

We used a Monte Carlo approach to study the performance of CSEGs estimated from various simulations of stochastic processes. Using a particular stochastic model (the 'base model'), we randomly generated 1000 time series of 10, 20 and 30 years. Each simulation was allowed to run for 100 years (the 'burn-in' period) before any observations $\{O_1, O_2, \ldots, O_k\}$ were recorded. A CSEG was estimated from each series using the slope method (Holmes 2001, 2004), and used the estimated CSEG to give an estimate of the probability of an 80% population decline within the next 50 years. To determine the *actual* probability of 80% decline for a given series, we used the base model to simulate 1000 50-year trajectories forward from the population state found at the end of the period. In this way, we obtain a CSEG estimate for the expected probability of quasi-extinction, which we compare with the actual expected probability of quasi-extinction.

To calculate the quasi-extinction probabilities, we have to define the quasi-extinction threshold more explicitly. Specifically if our threshold is a percent decline, against what population size should the percent decline be measured? PVAs often use an average of the last three to five censuses, to reduce the effect of measurement errors and other noise. Another common threshold in practice is an absolute number, for example 250 individuals; in this case, the problem of defining the baseline is avoided. In this paper we define percent declines relative to $\hat{X}_k$, the maximum-likelihood estimate of the final count after the white noise process has been removed. $\hat{X}_k$ is thus a fourth parameter that had to be estimated from the CSEG in our cross-validations.

As discussed in the introduction, the focus of this paper is on estimating the *propensity* of a given process to experience quasi-extinction. Thus, we cross-validate our CSEG estimates against the true *expected* probability of quasi-extinction:

$$\begin{aligned}\mathrm{E}[\Pr(\mathrm{qe})] = \sum_{\{O_1, O_2, \ldots, O_k\} \in \Omega} &\Pr(\{O_1, O_2, \ldots, O_k\}) \\ &\times \Pr(80\%\ \text{decline in}\ k+1\ \text{to}\ k+51 | \{O_1, O_2, \ldots, O_k\}).\end{aligned} \tag{12}$$

$\{O_1, O_2, \ldots, O_k\}$ denotes a specific simulated parameterization period, and $\mathbf{\Omega}$ signifies the random sample of 1000 such periods. Pr(80% decline in next 50 years | $\{O_1, O_2, \ldots, O_k\}$) is the actual probability of 80% decline after a given parameterization period, and $\Pr(\{O_1, O_2, \ldots, O_k\})$ is the probability of observing the given parameterization period. $\Pr(\{O_1, O_2, \ldots, O_k\})$ is simply 1/1000, as our parameterization periods are drawn randomly from the stationary distribution.

We simulated three very different cyclic population models: a stage-structured model, a density-dependent model with over-compensating dynamics and a model with four interacting species. For a stage-structured model, we used the Chinook salmon model (eqn 5) and estimated CSEGs from the spawner counts only. The parameters of

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

the underlying model (given in Fig. 2) produce strongly fluctuating spawner counts (Fig. 2b). For an over-compensating, density-dependent process we used the stochastic Ricker model (eqn 11) using parameters ($K = N_0 = 1000$, $r = 0.02$, $\sigma_p^2 = 0.04$) taken from recently published PVA analyses of density-dependent processes (McCarthy *et al.* 2003; Sabo *et al.* 2004). This model experiences large density-dependent changes in population size, but is not tightly regulated (see Fig. 5a). To generate the multi-species time series, we used a first-order, multivariate, autoregressive (MAR-1) process. In this simulation, the dynamics of each species are described by a discrete stochastic Gompertz model (Dennis *et al.* 2006):

$$\log N_{i,t+1} = g + b \log N_{i,t} + \varepsilon_{i,t}, \qquad (13)$$

where g is the growth rate, b is the strength of the density dependence and $\varepsilon_{i,t}$ is a normally distributed variable with mean zero and process error variance of σ_i^2. Multi-species dynamics are then modelled by extending eqn 13 (Ives *et al.* 2003):

$$\mathbf{X}_{t+1} = \mathbf{G} + \mathbf{B}\mathbf{X}_t + \mathbf{E}_t. \qquad (14)$$

Here $\mathbf{X}_t$ is a vector of the population sizes for each species at time t, $\mathbf{G}$ is a vector of their growth rates, $\mathbf{B}$ is the community matrix describing the strength of self-regulation (diagonal elements) and interspecies interactions (off-diagonals) and $\mathbf{E}_t$ is a process noise vector. We specified the MAR-1 process using estimates from a zooplankton community in Peter Lake, WI, USA (Ives *et al.* 2003). In estimating the CSEG and forecasting risk, we only used the time series for one of the four species (large phytoplankton). The time series of this process fluctuates tightly about an equilibrium value (similar to the process shown in Fig. 5b).

Figure 6 shows the estimated probability of decline, E[Pr(qe)], found using 10-, 20- and 30-year parameterization periods. The first thing to notice is that 10-year periods are

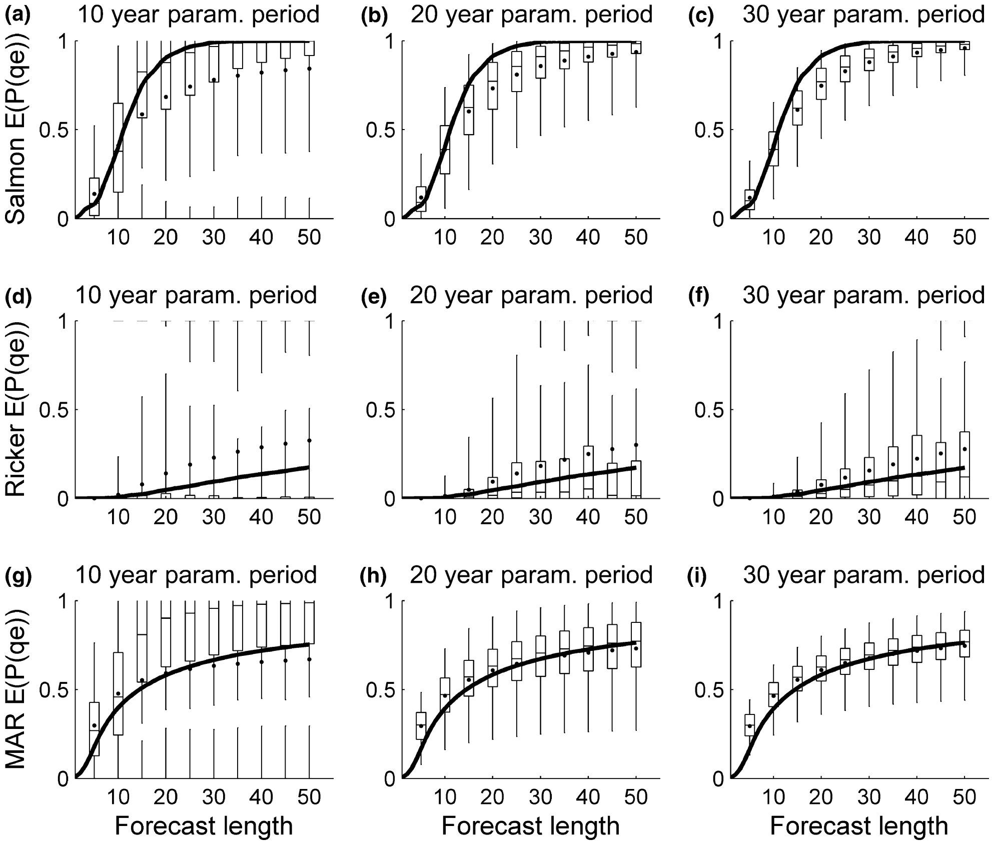

Figure 6 Cross-validation using simulations of the expected quasi-extinction probability as a function of forecast length (x-axis). The three rows refer to a Chinook salmon simulation (a-c), a stochastic Ricker model (d-f) and a four-species stochastic community (g-i). The solid black lines show the actual mean probability of 80% decline observed in the simulations. The boxes and whiskers show the range of CSEG estimates obtained using 10-year (a,d,g), 20-year (b,e,h) and 30-year (c,f,i) parameterization periods. The boxes enclose 50% of the CSEG estimates with the line in box showing the median and the whiskers show the range containing 95% of the estimates. The black dots show the mean CSEG estimate. The CSEG parameters were estimated using the slope method.

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

insufficient for estimating the quasi-extinction risk. Although the estimates themselves are not terribly biased, their distributions are spread across 0–1 and are often bimodal. Thus, from now on we refer only to results for 20- or 30-year parameterization periods. It can be seen that the mean estimated CSEG risks (black dots) track the actual risks closely (solid black lines), but are not strictly accurate; each model generates its own flavour of bias. The slope method, as expected (Holmes & Fagan 2002), produces a biased estimate of σ_b^2. Initial tests suggest that this error could have been approximately halved by estimating the bias via parametric bootstrapping. Nonetheless, it is striking that the estimated and actual risks correspond so well given that all three models produce strongly oscillatory data and have high non-process variance.

The boxes and whiskers in Fig. 6 enclose 50% and 95% of the estimates respectively. The line in the middle of each box represents the median. The inner quartiles (covering 50% of all estimates) are quite small for the MAR and salmon simulations. In these two populations, the 95% ranges are also smaller than (0,1) for both short (10–20 year) and long (30+ year) forecast lengths. In the Ricker model, the long-term forecasts have very wide and bimodal 95% ranges while the short-term (10- and 20-year) forecasts have relatively low variability (Fig. 6e,f). This occurs because the μ parameter of the CSEG (which measures trend) should have been constrained to zero. Because this process reverts to the long-term mean population very slowly, the μ estimates took on highly uncertain positive and negative values.

Estimating and forecasting are known to be difficult for processes with $\mu = 0$ (flat trend) and with slow fluctuations about the long-term mean. Small errors in the estimated value of μ lead to large errors in the estimated long-term population size and this leads to bimodal long-term extinction risk predictions. This problem has been noted frequently by others and motivates the common recommendation that policy-makers should rely on 10- to 30-year forecasts rather than 100-year. A recent paper introduced a discrete Gompertz state-space model to forecast extinction risks for density-dependent processes (Dennis *et al.* 2006). This may be a better approach for slowly reverting processes, for which the problem of bimodal qe estimates is especially apt to occur. We should note, however, that this is not necessarily a problem for the typical species being considered for an endangered, threatened or similar designation: such species are often declining.

Studies on CSEG forecasts for real-time series

Simulations are useful for studying population dynamics, but the models chosen are simplified and not necessarily representative of typical populations. In our second cross-validation study, we assembled a database of 63 time series at least 30 years long from species monitored for conservation or management reasons. These data were obtained from a literature search and through direct contact with governmental agencies across the world (Table S1). The majority are high-risk populations officially listed by one or more conservation agencies at the endangered or threatened level.

Our methods are similar to those of other cross-validations using real data (Brook *et al.* 2000; Holmes & Fagan 2002; Holmes *et al.* 2005), except that we examine the precision of our estimates as well as the bias. Each time series was divided into a 20-year parameterization period followed by a 10-, 20-, or 28-year forecast period. Very long time series were also segmented to increase the sample size; in this case the parameterization periods were allowed to overlap by 10 years. No more than five parameterization periods were taken from any one time series, so as not to over-represent any species or population. For each time series, a CSEG was estimated from the parameterization period using the slope method. The CSEG was then used to predict whether the quasi-extinction threshold would be reached during the forecast period. We validated the CSEG estimates in two different ways. First, the actual number of quasi-extinctions in the entire database was compared with the expected number derived from CSEG estimates at various thresholds and horizon lengths. This analysis assesses the systematic bias in the method and is analogous to the validation of E[Pr(qe)] estimates shown in Fig. 6. The goal of this analysis is to quantify whether we can properly estimate the risk on average. Unlike the simulations, we cannot assess the variability of the risk estimates directly. Instead, we use parametric bootstrapping (Dennis & Otten 2000; Engen *et al.* 2001) to estimate their precision. In this method, the estimated model is used to generate new data from which the parameters can be estimated again, and so on. This process creates a series of parameter estimates, which can be used to evaluate variability.

We calculated two measures of uncertainty. The first is the probability of 20%, 50%, or 80% decline within a 10-year forecast period, with 95% and 50% confidence intervals. These confidence intervals are notoriously difficult for decision makers to use – being essentially an uncertainty on an uncertainty. Population prediction intervals (PPI) have been proposed as an alternative (Engen *et al.* 2001; Sæther & Engen 2002). The PPI is the time interval that will include a decline to some quasi-extinction threshold 95% or 50% of the time. The PPI interval incorporates the uncertainty because of stochasticity as well as the uncertainty in the estimated parameters.

Figure 7 shows the estimated and actual proportion of quasi-extinctions over the entire database for 10-, 20- and 28-year forecasts following a 20-year parameterization period. The observed (line) and predicted (points) frequen-

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

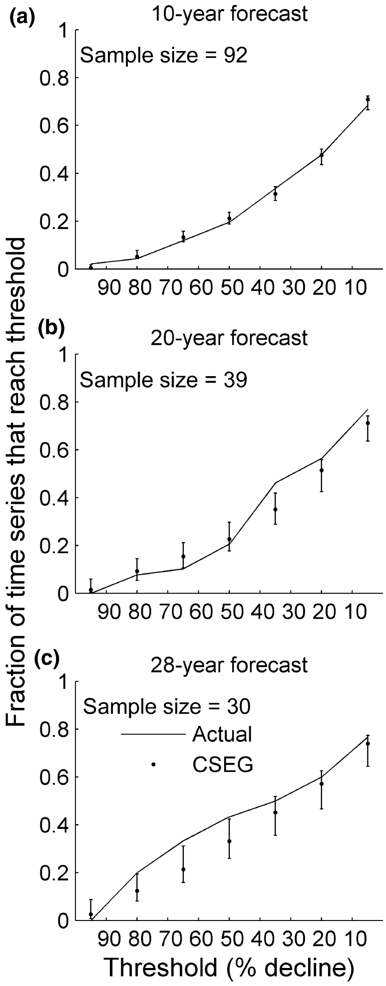

Figure 7 Cross-validation of bias using real-time series data. The panels show the expected fraction of quasi-extinction within the dataset vs. the actual fraction for (a)10-, (b)20- and (c)28-year forecasts. A 28-year forecast was used instead of 30-year since we had few time series that were 50 years or longer, but a number of 48-year time series. The sample size (number of time series) is different in each panel because there are fewer long time series in the dataset. The 95% CI error bars on the CSEG estimates were calculated via parametric bootstrapping. A 20-year parameterization period and the slope method were used for all time series.

cies match very closely. These plots show that the CSEG estimates of E[Pr(qe)] are generally unbiased. Figure 8 shows the estimated precision of the estimates – via parametric bootstrapping. The left-hand column of Fig. 8 shows the estimated 95% (grey) and 50% (black) CIs of the CSEG-estimated quasi-extinction probabilities (80%, 50% and 20% declines, all over a 10-year projection period). The 50% CIs are quite narrow, and the 95% CIs are less than (0,1). Precision, as might be expected, is best when the estimated risk is high (> 75%) or low (< 25%). Precision is worst (wide CIs) when the estimated qe is near 50%. However, the PPIs in Fig. 8 indicate that the certainty of the risk estimation is often better than the CIs might suggest. The PPIs also highlight the fact that in most cases, a low estimated probability of quasi-extinction within 10 years corresponds to a long expected quasi-extinction time (< 50% of time series experience qe before 100 years) and conversely when the estimated qe risk is > 50%, half of time series will have experienced qe well before 50 years. The dataset used in this analysis is focused on populations of conservation concern. Such populations are atypical in that they tend to be at historically low levels or declining; however, PVA is generally done for precisely these types of populations.

DISCUSSION

In this paper, we have drawn a distinction between (i) statistical models, which seek to capture the ensemble properties of stochastic processes, and (ii) mechanistic models, which seek to portray the underlying biological processes. We have shown how simple stochastic approximations can model a particular ensemble property, the risk of quasi-extinction. We have focused on the corrupted stochastic exponential Gaussian model (CSEG), and shown how the CSEG can accurately approximate quasi-extinction risks across different forecast lengths and risk thresholds, for a variety of population processes. Using simulations and 20 years of data, we illustrated that CSEG estimates of quasi-extinction risk can have relatively low bias and 95% CIs that are much smaller than (0,1) – particularly for declining and rapidly fluctuating populations. We cross-validated these results using a large dataset of abundance time series from species of conservation concern. We found that estimates had low bias across 10- to 30-year forecasts and that the 95% CIs were much less than (0,1) when forecasting severe declines (50% and 80%). Wide CIs were mainly a problem when estimating the risk of small declines (20%) and when the estimated probability of quasi-extinction was intermediate (near 0.5).

The point of this paper, however, is not that the CSEG is a panacea for all quasi-extinction estimation. Although we found that an adequate CSEG could be estimated for declining or tightly fluctuating populations, estimation was problematic for populations fluctuating slowly about an equilibrium. There is ongoing theoretical work on stochastic population processes, and there remains much to learn – especially about the statistical properties of density-dependent processes. Rather, the main point of this paper is to illustrate and advocate the use of statistical models based on the convergent properties of stochastic population trajectories. The main strengths of this approach are threefold. First, if a convergent approximation exists for several broad classes of dynamics, one needn't specify the

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

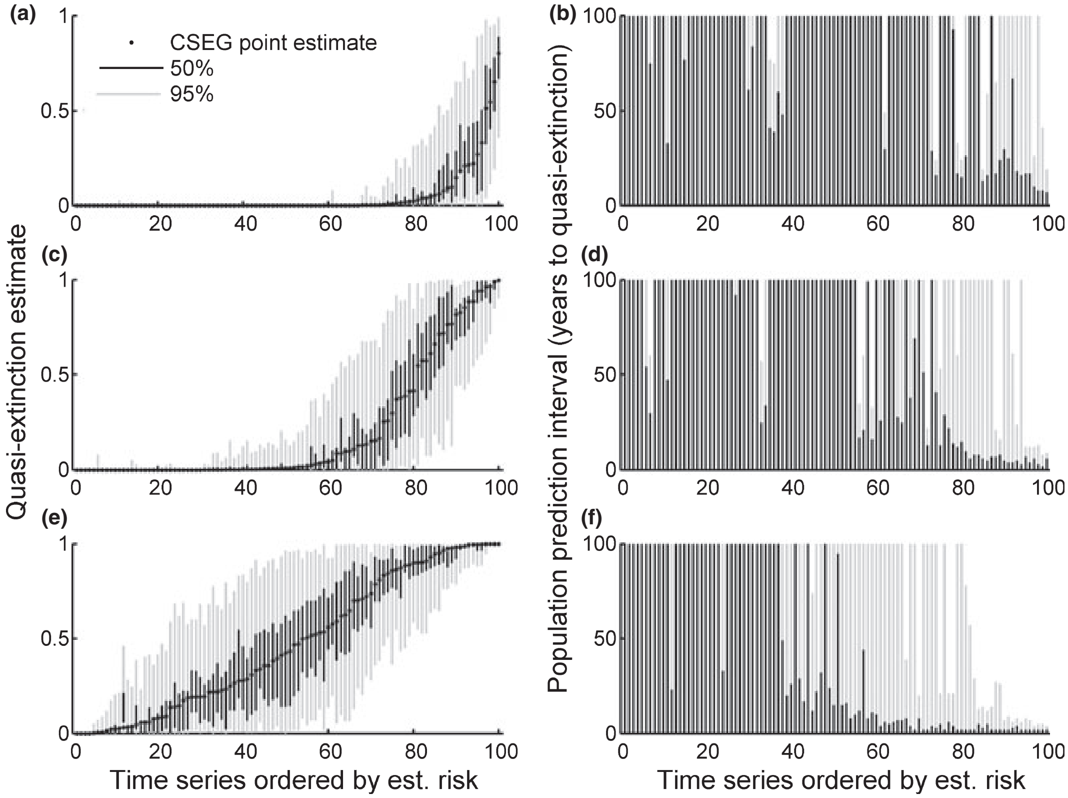

Figure 8 Two measures of quasi-extinction uncertainty: confidence intervals and population prediction intervals. The left-hand panels show the 95% and 50% confidence intervals on the estimated probabilities for 80%, 50% and 20% declines (panels a, c and e respectively). Confidence intervals were calculated using parametric bootstrapping as follows: the estimated CSEG was used to generate 1000 random 20-year parameterization periods. From each, a CSEG was estimated to give 1000 bootstrapped estimated CSEGs. The bootstrapped CSEGs were then used to generate quasi-extinction forecasts. The CIs show the variability of the bootstrapped quasi-extinction risks. The right-hand panels show the population prediction intervals (PPI) which is the 0 to T interval that includes 50% (black) or 95% (grey) of all quasi-extinctions. The y-axis maximum is 100+ years, so all PPIs with $T > 100$ years span the entire y-axis. Panels b, d, and f show PPIs for 80 %, 50 %, and 20 % declines, respectively. The population prediction intervals were determined with parametric bootstrapping, following Engen *et al.* (2001).

underlying mechanistic process to accurately forecast quasi-extinction. Second, a parameter-sparse statistical model requires much less data for estimation than a typical parameter-rich mechanistic model, and is thus less likely to be compromised by over-fitting. Finally, the parameters of a statistical model are estimated and their uncertainty can be quantified using standard statistical techniques. Our inability to rigorously quantify the uncertainty of predictions based on parameter-rich, mechanistic models is a serious problem and one too rarely acknowledged.

Confronting the criticisms of simple PVA models

PVA models that leave out biological detail have often been criticized (Taylor 1995; Ludwig 1996; Beissinger & Westphal 1998; Ludwig 1999; Fieberg & Ellner 2000; Coulson *et al.* 2001; Ellner *et al.* 2002). The critics contend that because simple PVA models ignore or incorrectly model important processes in the population dynamics, they must yield biased and/or imprecise forecasts of risk. They also assert that PVA predictions are unreliable because the data used to estimate model parameters are marred by observer errors and because sparse datasets lead to highly uncertain parameter estimates with extremely wide confidence intervals on the estimated risk. Here, we revisit these criticisms in light of our results regarding statistical PVA models.

First, in fairness to previous critiques of PVA, we note that the data used by conservation biologists are indeed riddled with observation errors. These errors can corrupt and bias estimates of the true yearly variation and, consequently, quasi-extinction risk. Much recent research has been focused on the use of state-space models to deal

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

with observation errors. CSEG models provide an explicit way of addressing this problem, and greatly reduce the effects of these errors on the quasi-extinction risk forecast (Holmes 2001; Holmes & Fagan 2002; Lindley 2003; Holmes 2004; Staples *et al.* 2004; Dennis *et al.* 2006). This research shows that the bias due to observation error can be quantified and its effects mitigated using a state-space model like the CSEG.

A second common criticism of simple PVA models is that biological detail is necessary for accurately modelling population dynamics. Examples abound of problems that can arise when a PVA model lacks age or stage structure (Fieberg & Ellner 2000; Ellner *et al.* 2002; Wilcox & Possingham 2002), individual variations (Fox & Kendall 2002), demographic stochasticity (Lande 1993; Engen *et al.* 2003, 2005), density dependence (Foley 1994; Sabo *et al.* 2004), or species interactions (Gerber *et al.* 2005; Sabo & Gerber 2007). The argument that we present in this paper is that complex details often average out when we look at ensemble properties, that these ensemble properties can have similar patterns across different types of population dynamics, and that simple approximations exist for one of these properties – quasi-extinction probabilities. The challenge is properly estimating the approximations, and many of the examples of problems occur when one simply treats the data as if they came from a SEG when they in fact show substantial differences from SEG data. Estimation is most difficult for populations fluctuating slowly about a long-term mean; however, endangered and threatened species often do not fall into that category. That details tend to average out is particularly important for PVA because we often are highly uncertain concerning even basic biological details for our population (e.g. the strength of density-dependence or age-specific vital rates). Developing models and metrics that have good convergent statistical properties ameliorates this problem. By contrast, building complex models we cannot parameterize exacerbates this problem.

Finally, there is what is known as the (0,1) criticism, which asserts that sparse data and strong environmental variations lead to quasi-extinction risk estimates whose confidence intervals either span the entire range of probabilities or are bimodal on 0 and 1 (Ludwig 1996; Fieberg & Ellner 2000; Ellner *et al.* 2002). Along similar lines, some researchers have found that risk estimates using simple PVA models are reasonably precise only for forecasts that are much shorter (e.g. 20%) than the number of years available for parameterization (Fieberg & Ellner 2000). We agree with these criticisms when too little data (i.e. only 10 years) are available. As we have seen in this study and other cross-validations (Holmes 2004), a 10-year time series is too short for estimating quasi-extinction risk. Our conclusions differ, however, when a 20-year time series is available. Using simulations, we found that in this case unimodal and relatively unbiased risk estimates could be made for up to two and a half times the length of the parameterization period in our declining and rapidly fluctuating populations. Only the slowly fluctuating populations were limited to shorter forecasts. This agrees with our previous simulation studies (Holmes 2001, 2004; Sabo *et al.* 2004; Sabo & Gerber 2007) and is supported by cross-validations using time series from species of conservation concern (this study; Holmes & Fagan 2002; Holmes *et al.* 2005).

One reason that our simulation results are more optimistic is that we use a much smaller value of σ_b^2 in the CSEG. Recall that σ_b^2 represents how variance increases in a time series; it is a stochastic average over all the variations buffeting a population. Based on CSEG estimates taken from hundreds of vertebrate time series where we separated σ_b^2 and σ_w^2, we consider the range 0.01–0.05 to be realistic. In contrast, studies exhibiting extreme (0,1) problems typically use models with such high process error variance that they produce σ_b^2 values in the range of 0.1–0.5. Such values produce time series that resemble insect trajectories more than vertebrate trajectories. It could be argued that our optimist outlook represents a taxonomic bias, however, this bias is also present in the monitoring of species of conservation concern.

When and where can statistical models for extinction be useful?

CSEGs can be used to estimate quasi-extinction probabilities, which are in turn useful for establishing the status of threatened and endangered species. Quasi-extinction probabilities are also useful in 'conservation triage' – ranking the risks faced by various species to prioritize attention and aid (McCarthy *et al.* 2003). However, simple statistical models (like the CSEG) will not be useful for all types of analyses performed in a PVA. Ignoring the mechanisms responsible for dynamics means that we cannot find mechanistic solutions to population decline. For example, only stage- or age-structured models can aid practitioners in identifying the life stages or vital rates most critical to a declining population. Many previous studies have applied PVA in this way (Crouse *et al.* 1987; Doak *et al.* 1994; Heppell *et al.* 1996; Caswell 2000; Wisdom *et al.* 2000). In general, statistical models of quasi-extinction risk will not prove useful for answering such 'What should we do and how should we do it?' questions.

Although we have argued in this paper that with 20 years of data the (0,1) problem is neither insurmountable nor ubiquitous, quasi-extinction estimates are estimates and they definitely have uncertainty and sometimes high uncertainty. This paper has been focused on accurately specifying that uncertainty. To then make decisions based on these estimates, it is critical that they be used in a decision-

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

making framework that allows one to account for the uncertainty surrounding the estimates and that does not rely on point estimates. This is an active research area and there are a diversity of options, both frequentist and Bayesian, for approaching this problem (Ludwig 1996; Wade 2000; Sæther & Engen 2002; Taylor *et al.* 2002; Staples *et al.* 2005; Nicholson & Possingham 2007). In addition, research is ongoing into formal decision theory frameworks for PVA that allow policy-makers to weigh the uncertainty concerning the true risk level against the costs of those different levels being true (Goodman 2002; Dorazio & Johnson 2003; Drechsler & Burgman 2004).

CONCLUSION

We do not dispute the importance of mechanistic models, which provide testable hypotheses about underlying mechanisms and can be used to simulate the effects of specific management actions. When the only goal is forecasting quasi-extinction, however, we argue for a statistical model with few free parameters. Appropriate statistical models can be derived from the theory of stochastic population processes and can be easily estimated from the most common type of data for species of concern – time series of population counts. It is true that in some cases the risk estimates will have high variance, but we argue that statistical models both reduce this variance and allow us to quantify the uncertainty in more a traditional manner. In contrast, for all but the most well-studied species, any mechanistic population model will be replete with poorly estimated parameters. In such cases, meaningful uncertainty estimates are impossible. Critics of PVA tend to argue that the devil is in the details. We counter that the urge to account for intricate biological details ignores the fact that these details average out thanks to the stochastic and multiplicative nature of population growth. If we focus only on the details, we lose sight of the common patterns that govern the ensemble behaviour of population trajectories.

ACKNOWLEDGEMENTS

We thank B. Dennis, S. Beissinger, K. Zikan and three anonymous referees for their constructive comments and suggestions on the manuscript. J. Rango, A. Folarin, J. Sorensen, J. Lippe and T. Mongillo helped assemble the time series data. SV was supported by a fellowship from the National Research Council. WFF was supported by Dept. of Defense–SERDP contract SI-1475. This work was stimulated by a two-day workshop conducted at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant #DEB-0553768), the University of California, Santa Barbara, and the State of California.

REFERENCES

Akçakaya, H.R., Halley, J.M. & Inchausti, P. (2003) Population-level mechanisms for reddened spectra in ecological time series. *J. Anim. Ecol.*, 72, 698–702.

Beissinger, S.R. & McCullough, D.R. (2002) *Population Viability Analysis*. University of Chicago Press, Chicago.

Beissinger, S.R. & Westphal, M.I. (1998) On the use of demographic models of population viability in endangered species management. *J. Wildl. Manage.*, 62, 821–841.

den Boer, P.J. (1991) Seeing the trees for the wood: random walks or bounded fluctuations of population size? *Oecologia*, 86, 484–491.

den Boer, P.J. & Reddingius, J. (1989) On the stabilization of animal numbers. Problems of testing. 2. Confrontation with data from the field. *Oecologia*, 79, 484–491.

Boyce, M.S. (1992) Population viability analysis. *Annu. Rev. Ecol. Syst.*, 23, 481–506.

Brook, B.W., O'Grady, J.J., Chapman, A.P., Burgman, M.A., Akçakaya, H.R. & Frankham, R. (2000) Predictive accuracy of population viability analysis in conservation biology. *Nature*, 404, 385–387.

Caswell, H. (2000) *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer, Sunderland, MA, USA.

Cohen, J. (1977) Ergodicity of age structure in populations with Markovian vital rates, III: finite-state moments and growth rate; an illustration. *Adv. Appl. Prob.*, 9, 462–475.

Coulson, T., Mace, G.M., Hudson, E. & Possingham, H. (2001) The use and abuse of population viability analysis. *Trends Ecol. Evol.*, 16, 219–221.

Crouse, D.T., Crowder, L.B. & Caswell, H. (1987) A stage-based population-model for Loggerhead sea-turtles and implications for conservation. *Ecology*, 68, 1412–1423.

DeMaster, D., Angliss, R., Cochrane, J., Mace, P., Merrick, R., Miller, M. *et al.* (2004) *Recommendations to NOAA Fisheries: ESA Listing Criteria by the Quantitative Working Group 10 June 2004*. U.S. Department of Commerce, NOAA Tech. Memo. NMFS-F/SPO-67, Washington DC, 85 pp.

Dennis, B. & Costantino, R.F. (1988) Analysis of steady-state populations with the gamma abundance model: application to Tribolium. *Ecology*, 69, 1200–1213.

Dennis, B. & Otten, M.R.M. (2000) Joint effects of density dependence and rainfall on abundance of San Joaquin Kit Fox. *J. Wildl. Manage.*, 64, 388–400.

Dennis, B. & Taper, M.L. (1994) Density dependence in time series observations of natural populations: estimation and testing. *Ecol. Monogr.*, 64, 205–224.

Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L. & Staples, D.F. (2006) Estimating density dependence, process noise, and observation error. *Ecol. Monogr.*, 76, 323–341.

Diserud, O.H. & Engen, S. (2000) A general and dynamic species abundance model, embracing the lognormal and the gamma models. *Am. Nat.*, 155, 498–511.

Doak, D., Kareiva, P. & Kleptetka, B. (1994) Modeling population viability for the desert tortoise in the Western Mojave desert. *Ecol. Appl.*, 4, 446–460.

Dorazio, R.M. & Johnson, F.A. (2003) Bayesian inference and decision theory – a framework for decision making in natural resource management. *Ecol. Appl.*, 13, 556–563.

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works

Drechsler, M. & Burgman, M.A. (2004) Combining population viability analysis with decision analysis. *Biodivers. Conserv.*, 13, 115–139.

Ellner, S.P., Fieberg, J., Ludwig, D. & Wilcox, C. (2002) Precision of population viability analysis. *Conserv. Biol.*, 16, 258–261.

Engen, S., Sæther, B.-E. & Møller, A.P. (2001) Stochastic population dynamics and time to extinction of a declining population of barn swallows. *J. Anim. Ecol.*, 70, 789–797.

Engen, S., Lande, R. & Sæther, B.-E. (2003) Demographic stochasticity and allee effects in populations' with two sexes. *Ecology*, 84, 2378–2386.

Engen, S., Lande, R., Sæther, B.-E. & Weimerskirch, H. (2005) Extinction in relation to demographic and environmental stochasticity in age-structured models. *Math. Biosci.*, 195, 210–227.

Engen, S., Lande, R. & Sæther, B.-E. (2007) Using reproductive value to estimate key parameters in density-independent age-structured populations. *J. Theor. Biol.*, 244, 308–317.

Fagan, W.F. & Holmes, E.E. (2006) Quantifying the extinction vortex. *Ecol. Lett.*, 9, 51–60.

Fieberg, J. & Ellner, S.P. (2000) When is it meaningful to estimate an extinction probability? *Ecology*, 81, 2040–2047.

Foley, P. (1994) Predicting extinction times from environmental stochasticity and carrying capacity. *Conserv. Biol.*, 8, 124–137.

Fox, G.A. & Kendall, B.E. (2002) Demographic stochasticity and the variance reduction effect. *Ecology*, 83, 1928–1934.

Furstenberg, H. & Kesten, H. (1960) Products of random matrices. *Ann. Math. Stat.*, 31, 457–469.

Gerber, L.R., McCallum, H., Lafferty, K.D., Sabo, J.L. & Dobson, A. (2005) Exposing extinction risk analysis to pathogens: Is disease just another form of density dependence? *Ecol. Appl.*, 15, 1402–1414.

Gilpin, M.E. & Soule', M.E. (1986) Minimum viable populations: processes of extinction. In: *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soule', M.E.). Sinauer Associates, Sunderland, MA, pp. 19–34.

Ginzburg, L., Slobodkin, L.B., Johnson, K. & Bindman, A.G. (1982) Quasi-extinction probabilities as a measure of impact on population growth. *Risk. Anal.*, 21, 171–181.

Goodman, D. (2002) Predictive Bayesian population viability analysis: a logic for listing criteria, delisting criteria, and recovery plans. In: *Population Viability Analysis* (eds Beissinger, S.R. & McCullough, D.R.). University of Chicago Press, Chicago, IL, pp. 447–469.

Halley, J. & Kunin, W. (1999) Extinction risk and the 1/*f* family of noise models. *Theor. Popul. Biol.*, 56, 215–239.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, UK.

Heppell, S.S., Crowder, L.B. & Crouse, D.T. (1996) Models to evaluate headstarting as a management tool for long-lived turtles. *Ecol. Appl.*, 6, 556–565.

Hinrichsen, R. (2002) The accuracy of alternative stochastic growth rate estimates for salmon populations. *Can. J. Fish. Aquat. Sci.*, 59, 1014–1023.

Holmes, E.E. (2001) Estimating risks in declining populations with poor data. *Proc. Natl Acad. Sci. USA*, 98, 5072–5077.

Holmes, E.E. (2004) Beyond theory to application and evaluation: diffusion approximations for population viability analysis. *Ecol. Appl.*, 14, 1272–1293.

Holmes, E.E. & Fagan, W.E. (2002) Validating population viability analysis for corrupted data sets. *Ecology*, 83, 2379–2386.

Holmes, E.E. & Semmens, B. (2004) Population viability analysis for metapopulations: a diffusion approximation approach. In: *Ecology, Genetics, and Evolution of Metapopulations* (eds Hanski, I. & Gaggiotti, O.E.), Elsevier Press, Burlington, MA, pp. 565–598.

Holmes, E.E., Fagan, W.F., Rango, J.J., Folarin, A.J.A.S., Lippe, J.E. & McIntyre, N.E. (2005) *Cross Validation of Quasi-Extinction Risks From Real Time Series: An Examination of Diffusion Approximation Methods*. U.S. Department of Commerce, NOAA Tech. Memo. NMFS-NWFSC-67, Washington, DC. 37 pp.

Inchausti, P. & Halley, J. (2001) Investigating long-term ecological variability using the Global Population Dynamics Database. *Science*, 293, 655–657.

Inchausti, P. & Halley, J. (2002) The long-term temporal variability and spectral colour of animal populations. *Evol. Ecol. Res.*, 4, 1033–1048.

Ives, A.R., Dennis, B., Cottingham, K.L. & Carpenter, S.R. (2003) Estimating community stability and ecological interactions from time-series data. *Ecol. Monogr.*, 73, 301–330.

Lande, R. (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.*, 142, 911–927.

Lande, R. & Orzack, S.H. (1988) Extinction dynamics of age-structured populations in a fluctuating environment. *Proc. Natl Acad. Sci. USA*, 85, 7418–7421.

Lande, R., Engen, S. & Saether, B. (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, Oxford.

Lewontin, R.C. & Cohen, D. (1969) On population growth in a randomly varying environment. *Proc. Natl Acad. Sci. USA*, 62, 1056–1060.

Lindley, S.T. (2003) Estimation of population growth and extinction parameters from noisy data. *Ecol. Appl.*, 13, 806–813.

Ludwig, D. (1996) Uncertainty and the assessment of extinction probabilities. *Ecol. Appl.*, 6, 1067–1076.

Ludwig, D. (1999) Is it meaningful to estimate a probability of extinction? *Ecology*, 80, 298–310.

Mace, G.M. & Lande, R. (1991) Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. *Conserv. Biol.*, 5, 148–157.

McCarthy, M.A., Andelman, S.J. & Possingham, H.P. (2003) Reliability of relative predictions in population viability analysis. *Conserv. Biol.*, 17, 982–989.

Morris, W.F. & Doak, D.F. (2003) *Quantitative Conservation Biology: Theory and Practice of Population Viability Analysis*. Sinauer Associates, Sunderland, MA USA.

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

Nicholson, E. & Possingham, H.P. (2007) Making conservation decisions under uncertainty for the persistence of multiple species. *Ecol. Appl.*, 17, 251–265.

Pimm, S. (1991) *The Balance of Nature? Ecological Issues in the Conservation of Species and Communities*. Chicago University Press, Chicago, IL.

Pimm, S. & Redfearn, A. (1988) The variability of natural populations. *Nature*, 334, 613–614.

Polya, G. (1968) *Mathematics and Plausible Reasoning*, Vol. II. Princeton University Press, Princeton, NJ.

Sæther, B.-E. & Engen, S. (2002) Including uncertainty in population viability analysis using population prediction intervals. In: *Population Viability Analysis* (eds Beissinger, S.R. & McCullough, D.R.), University of Chicago Press, Chicago, IL, pp. 191–212.

Sabo, J.L. & Gerber, L.R. (2007) Predicting extinction risk in spite of predator-prey oscillations. *Ecol. Appl.*, 17, 1543–1554.

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

Shumway, R.H. & Stoffer, D.S. (1982) An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Ser. Anal.*, 3, 253–264.

Smitalova, K. & Sujan, S. (1992) *A Mathematical Treatment of Dynamical Models in Biological Science*. Ellis Horwood Ltd, New York, NY, USA.

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

Staples, D.F., Taper, M.L. & Shepard, B.B. (2005) Risk-based viable population monitoring. *Conserv. Biol.*, 19, 1908–1916.

Taylor, B.L. (1995) The reliability of using population viability analysis for risk classification of species. *Conserv. Biol.*, 9, 551–558.

Taylor, B.L., Wade, P.R., Ramakrishnan, U., Gilpin, M. & Akcakaya, H.R. (2002) Incorporating uncertainty in population viability analyses for the purpose of classifying species by risk. In: *Population Viability Analysis* (eds Beissinger, S.R. & Mccullough, D.R.), University of Chicago Press, Chicago, IL, pp. 239–252.

Tuljapurkar, S.D. (1989) An uncertain life: demography in random environments. *Theor. Popul. Biol.*, 35, 227–294.

Tuljapurkar, S.D. & Orzack, S.H. (1980) Population dynamics in variable environments. I. long-run growth rates and extinction. *Theor. Popul. Biol.*, 18, 314–342.

de Valpine, P. (2002) Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-Gaussian state-space models. *Bull. Mar. Sci.*, 70, 455–471.

de Valpine, P. & Hastings, A. (2002) Fitting population models incorporating process noise and observation error. *Ecol. Monogr.*, 72, 57–76.

Wade, P. (2000) Bayesian methods in conservation biology. *Conserv. Biol.*, 14, 1308–1316.

Wilcox, C. & Possingham, H. (2002) Do life history traits affect the accuracy of diffusion approximations for mean time to extinction? *Ecol. Appl.*, 12, 1163–1179.

Wisdom, M.J., Mills, L.S. & Doak, D.F. (2000) Life stage simulation analysis: estimating vital-rate effects on population growth for conservation. *Ecology*, 81, 628–641.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Summary of population time series used in cross-validation analyses.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/full/10.1111/j.1461-0248.2007.01105.x.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Editor, Steve Beissinger
Manuscript received 27 March 2007
First decision made 11 May 2007
Manuscript accepted 30 July 2007

© 2007 Blackwell Publishing Ltd/CNRS. No claim to original US government works