The diffusion approximation overestimates the extinction risk for count-based PVA

Bruce E. Kendall
The Ecology Centre
University of Queensland
St Lucia, QLD 4072 AUSTRALIA†
and
Donald Bren School of Environmental Science and Management
University of California, Santa Barbara
Santa Barbara, CA 93106-5131 USA‡

Email: kendall at bren.ucsb.edu


Abstract

Simple population models are increasingly being used to predict extinction risk using historical abundance estimates. A very simple model, the stochastic exponential growth (SEG) model, is surprisingly robust. Extinction risk is commonly computed for this model using a mathematical approximation (the ‘diffusion approximation’) that assumes continuous breeding throughout the year, an assumption that is violated by many species. Here I show that, for an organism with seasonal breeding, the diffusion approximation systematically overestimates the extinction risk. I demonstrate the conditions generating large bias (high environmental variance, intermediate extinction risk), and reanalyze 100

†Address through 31 August 2009. Tel: +61 (7) 3365 2450; Fax: +61 (7) 3365 1655
‡Permanent address. Tel: +1 (805) 893-7539; Fax: +1 (805) 893-7612.
populations of conservation concern. Analyzing several policy applications, I find that the bias may be most important when classifying the risk status of species. The SEG model is still sound, but associated risk estimates should be calculated by performing stochastic simulations (as with all other population viability models) rather than by evaluating the diffusion approximation.

1 Introduction

Over the three decades since its formal introduction, population viability analysis (PVA) has been used to confirm the threatened status of individual species or populations (e.g., Shaffer, 1981; Lande, 1988), prioritize among populations based on relative risk (e.g., Menges, 1990; Forsman et al., 1996), develop quantitative recovery criteria (e.g., Gerber et al., 1999; Schultz & Hammond, 2003), and evaluate the prospects for reintroduction and restoration programs (e.g., South et al., 2000; Bell et al., 2003). Analyses that incorporate demography and other ecological processes allow the effects of various potential management strategies to be evaluated (e.g., Crouse et al., 1987; Crowder et al., 1994). Modern practice recognizes that the probability of absolute extinction is impossible to estimate accurately (there are too many unpredictable factors influencing the fate of the species when only a few individuals are left), and instead focuses on evaluating the risk of decline to a ‘quasi-extinction threshold’ below which extinction risks start to rapidly increase (e.g., Akçakaya & Sjögren-Gulve, 2000; Morris & Doak, 2002).

Any PVA starts with a population model; usually the model is stochastic. Having selected a starting population size ($N_0$) and a quasi-extinction threshold ($N_X$), one can use model simulations to estimate a wide variety of quasi-extinction risk statistics. Some of the more common include the probability that population size drops below $N_X$ on or before time $t$ ($G(t, d)$, where $d = \ln(N_0/N_X$)), the mean ($T_e$), median ($T_{\text{med}}$), or most likely ($T_{\text{mode}}$) time to first reach $N_X$, or the expected minimum population size (EMP; McCarthy & Thompson, 2001) over some time horizon.

The simplest stochastic population model is the stochastic exponential growth (SEG) model, in which the annual growth rate fluctuates randomly (reflecting environmental stochasticity) around a fixed, density-independent mean. Despite lacking biological detail (no density feedback, age structure, demographic stochasticity, or spatial dynamics), it is a robust approximating model for more complex population processes, especially in the face of limited data (Holmes, 2004; Sabo et al., 2004; Holmes et al., 2007). This simplicity presents two advantages. First, the model has only two parameters (the mean and variance of growth rate on a log scale, which are conventionally denoted $\mu$ and $\sigma^2$), which can be estimated from a simple time series of abundances, even if those abundance estimates are imprecise (Dennis et al., 1991; Holmes, 2001; Morris & Doak, 2002; Lindley, 2003; Staples et al., 2004; Buonaccorsi et al., 2006; Staudenmayer & Buonaccorsi, 2006; Buonaccorsi & Staudenmayer, 2009). Second, analytical formulas can be derived for all of the extinction risk metrics listed
above. These can be evaluated much more quickly than simulations, a great benefit when computers were slow. These formulas come from a mathematical technique known as a ‘diffusion approximation,’ and within the conservation community, the latter term is often used to refer to the SEG model (an unfortunate semantic custom, as the same mathematical technique can be applied to other models — e.g., Wilcox & Possingham (2002)).

After being introduced to ecologists by Dennis et al. (1991) and subsequently featured in a major textbook (Morris & Doak, 2002), the SEG model and its associated analytical approximations of quasi-extinction risk have been widely used for empirical PVAs. For example, it has been used in mammals (Nicholls et al., 1996; Lima et al., 1998; Gerber et al., 1999; Knight et al., 2001; Buenau & Gerber, 2004; Watson et al., 2005; Gonzalez-Suarez et al., 2006; Watson & Chadwick, 2007), birds (Gaston & Nicholls, 1995; Knight et al., 2001), insects (Schultz & Hammond, 2003), and plants (Thomson & Schwartz, 2006; Jacqueymyn et al., 2007).

The diffusion approximation (in the mathematical sense) generally works well when modeling continuous time processes, even in cases where one might expect it to break down (e.g., few individuals). However, many species of conservation concern (including most of those in the list of studies above), especially those that live in strongly seasonal environments, exhibit discrete breeding seasons. The discrete-time and continuous-time versions of the SEG model are not interchangeable, and the diffusion approximation is not guaranteed to apply in discrete time. In contrast, the discrete-time SEG model can be simulated, and quasi-extinction risk statistics generated from the simulations. In this paper, I ask, “Does the diffusion approximation provide an unbiased estimate of quasi-extinction risk in the discrete-time SEG model?” The answer, in brief, is “No.” After demonstrating the bias and its dependence on model parameters, I reanalyze the published studies to quantify the magnitude of bias of each of these. Finally, I examine the extent to which the bias might change three types of management decisions.

2 Methods

2.1 Patterns of bias

I used the diffusion approximation and simulations to calculate $G(t, d)$, $T_e$, $T_{med}$, and EMP, spanning a broad range of parameter values: $\mu$ ranged from $-0.25$ to $0.25$ and $\sigma^2$ ranged from $0.01$ to $1.8$, reflecting the range of parameter estimates in the empirical studies described below (for $T_{med}$, I focused on $\mu \leq 0$). For $G(t, d)$, I evaluated time horizons ranging from $1$ to $100$ years, and quasi-extinction thresholds corresponding to $50\%$ ($d = \ln 2$) to $90\%$ ($d = \ln 10$) declines. For EMP I examined time horizons of $10$, $50$, and $100$ years. For $T_e$ and $T_{mode}$, with much higher computational demands, I only evaluated $d = \ln 2$.

The simulation analyses of $G(t, d)$ and EMP were based on $10,000$ replicates, although this was increased to $50,000$ for some of the illustrative figures. The
analysis of $T_e$ and $T_{med}$ required 100,000 replicates of simulations that ran for up to 50,000 years.

I took the estimates from the simulations to represent ‘truth,’ and the deviations from these values by the diffusion approximation to be the bias. There are up to four parameters ($\mu$, $\sigma^2$, $d$, and $t$), so visualization of some patterns can be difficult. For $G(t,d)$ I used regression to understand the general patterns of bias across these parameters. Using a generalized linear model with Poisson-distributed residuals and a log link function, I modeled the bias as a function of linear and quadratic terms of the four parameters, as well as the predicted quasi-extinction risk itself. I did not include interaction terms because there was no evidence of nonlinearity in the residuals. I evaluated the bias in the remaining extinction risk metrics graphically.

2.2 Application to real populations

I identified 12 peer-reviewed journal articles that reported parameter estimates for the SEG model in real populations of conservation concern and used the diffusion approximation to estimate one or more quasi-extinction metrics. These studies encompassed 100 populations of 71 species; the estimates of $\mu$, $\sigma^2$, and $d$ for each population/species are tabulated in the Supporting Information.

I performed two analyses to evaluate potential biases in PVAs of real populations. The first was a standardized comparison: for each population, I evaluated the probability of a 50% (or more) decline within 10 years ($G(10, \ln 2)$). I also repeated the analyses reported in the original paper ($G(t,d)$, $T_e$, $T_{med}$, or $T_{mode}$), using the quasi-extinction thresholds and time horizons reported for each population. I calculated bias in the same way as in the previous section.

A useful scale with which to evaluate the magnitude of estimation bias is the precision of the estimate. Estimates of extinction risk are notoriously imprecise (McCarthy et al., 1996; Ludwig, 1999; Fieberg & Ellner, 2000), and in an analysis of an endangered plant, Thomson & Schwartz (2006) found that the 95% confidence interval for $G(50)$ spanned the entire [0,1] interval in most populations. With such low precision, bias might seem unimportant. However, at shorter time horizons, such as 10 years, the confidence interval of extinction risk can be much narrower, and the bias may be a substantial fraction of that interval. Confidence intervals for $\mu$ and $\sigma^2$ are reported for two thirds of the populations. Using these values, I performed a parametric bootstrap (assuming normally distributed uncertainty in $\mu$ and $\chi^2$ distributed uncertainty in $\sigma^2$), and using the simulation approach, calculated 95% confidence intervals for $G(10, \ln 2)$. I then compared the bias in quasi-extinction risk with the width of the confidence interval.

2.3 Projected impacts on policy

I examined the potential impact of the extinction risk bias across three applications PVA: ranking species or populations by their level of relative risk; setting
recovery goals; and classifying species into qualitative risk categories under the IUCN red list.

I examined risk ranking in two ways. First, using $G(10, \ln 2)$, I compared the ranking of all pairs of real populations under the diffusion approximation and simulations. I then used logistic regression to quantify how the probability of misranking depends on the difference in quasi-extinction risk between the two populations. Second, for those studies that quantified extinction risk for multiple populations or species, I examined how often the ranking changed when using simulations instead of the diffusion approximation, using the risk measure from the original publication.

Schultz & Hammond (2003) used the diffusion approximation to estimate the mean growth rate that would produce an acceptably low probability of extinction for the Fender’s blue butterfly, and recommended using these as recovery goals (although not formally incorporated into the draft recovery plan (U.S. Fish and Wildlife Service, 2008), the recommendations appear to have been influential). I repeated these analyses using simulations to explore the extent to which the recovery targets may have been overly conservative.

Criterion E of the IUCN red list allows classification of species via a quantitative analysis of extinction risk. A PVA based on the SEG is not appropriate for this, as the model is a poor approximation of dynamics at low population size. However, it is an appropriate tool for projecting likely future declines (Criterion A4). For example, a species is classified as Vulnerable if its abundance is projected to decline by at least 30% over the next ten years or three generations. This could be assessed by setting the quasi-extinction threshold to match a 30% decline and asking whether $G(10, \ln 1.43) > 0.5$. It is unclear whether this has ever been done (most species accounts on the red list do not specify the analyses used to project decline), but if it were, we would want the projections to be unbiased. Using the values of $\mu$ and $\sigma^2$ from the empirical populations, I calculated $G(10, d)$ for the appropriate threshold levels of $d$ and assigned an IUCN risk category accordingly, using both simulations and the diffusion approximation (I did not use species-specific generation times because, in most cases, the data are for populations rather than whole species; I am simply using them to provide typical parameter values for species of conservation concern).

3 Results

3.1 Patterns of bias

The diffusion approximation consistently over-estimated $G(t, d)$. The bias is close to zero when $\sigma^2 < 0.01$, and is highest (taking on values of 0.2 or greater) when $\mu$ is positive and $\sigma^2$ is large, although the details differ somewhat depending on $t$ and $d$ (Fig. 1). The bias tends to be higher for short time horizons and small values of $d$. The regression explained 99.3% of the deviance, and all coefficients were significantly different from zero except for the quadratic term in $\mu$; there was no evidence of nonlinearity in the residuals. The coefficients
Figure 1: Absolute bias in quasi-extinction probability estimated by the diffusion approximation, as a function of $\mu$ and $\sigma^2$, for two values of $d$ and time horizon ($t$). Positive values indicate that the diffusion approximation overestimates extinction risk. Points indicate the values of $\mu$ and $\sigma^2$ from the empirical literature.
Figure 2: Mean (red) and median (black) times to quasi-extinction, based on the diffusion approximation (lines) and simulations (points). $\mu = -0.05; d = \ln 2$.

Figure 2: Mean (red) and median (black) times to quasi-extinction, based on the diffusion approximation (lines) and simulations (points). $\mu = -0.05; d = \ln 2$.

(tabulated in the Supporting Information) reveal that, across the full span of parameters, the bias in quasi-extinction probability increases linearly with $\mu$, increases at a decelerating rate with $\sigma^2$, and decreases at decelerating rates with $d$ and $t$. The bias also depends on $G$, but in a humped fashion, with the bias highest when $G \approx 0.5$. This evidently explains the nonlinearities apparent in Fig. 1: the sharp bends in the contours correspond to parameter combinations in which $G(t, d) = 0.5$.

The diffusion approximation consistently underestimates both $T_e$ and $T_{med}$ (Fig. 2; Supporting Information). The difference is particularly striking for $T_e$: in contrast to the diffusion approximation, estimates from the simulations rise steadily as $\sigma^2$ increases, and at high levels of $\sigma^2$, the bias is 50% of simulation-based estimate. Estimates of $T_{med}$ decrease with $\sigma^2$ for both approaches, such that both estimates are small in absolute magnitude, but the relative bias increases steadily with increasing $\sigma^2$ (Fig. S1).

The diffusion approximation consistently underestimates EMP. The relative bias increases approximately linearly with both $\mu$ and $\sigma^2$, although there are some curious nonlinearities with negative values of $\mu$ and longer time horizons (Fig. 3). The relative bias reaches 0.8 at the highest levels of $\mu$ and $\sigma^2$, where the EMP predicted by the diffusion approximation is only 20% of the simulation value.
3.2 Application to real populations

The distribution of biases in $G(10, \ln 2)$ was bimodal, with peaks near zero and 0.1; 42% of the populations had biases greater than 0.1 (Fig. 4a). Small biases were mostly associated with very low quasi-extinction probabilities, whereas the highest biases were associated with $G \approx 0.5$ (Fig. 4b). When using the time horizons and quasi-extinction thresholds from the original studies (typically, larger values of both $t$ and $d$), most populations clustered at a high quasi-extinction probability, with biases smaller than 0.05 (Fig. 4c). For those studies with some estimate of quasi-extinction time, the bias was mostly less than 10 years (83%), but could be as large as 100 years (Fig. 4d).

31% of the populations had a 95% confidence interval for extinction risk with width less than 0.5. As predicted by recent theory (Ellner & Holmes, 2008), most of these were associated with populations for which the risk of decline was close to zero ($G(10, \ln 2) < 0.02$); bias is a minimal factor in these cases. In the remainder of the populations, the width of the confidence interval ranged from 0.63 to 0.95, with a mean of 0.78, and was uncorrelated with the bias. For 53% of these populations, the bias was between 10% and 20% of the confidence interval width.

3.3 Projected impacts on policy

Across all the empirical populations/species, using $G(10, \ln 2)$ as the risk metric, the pairwise ranking of quasi-extinction risk was almost always congruent if the difference in quasi-extinction risk between the two populations was greater than
Figure 4: Patterns of bias in real populations. (a) Histogram of overestimates of quasi-extinction probability ($G(10, \ln 2)$). (b) As in (a), but plotted vs. actual quasi-extinction probability, as evaluated by simulations. (c) Bias in quasi-extinction probability vs. actual quasi-extinction probability, using original authors’ values for $t$ and $d$, for those studies in which quasi-extinction probability was evaluated. (d) Bias in time to quasi-extinction vs. actual quasi-extinction time, using authors’ values for $d$ and their choice(s) of time statistics (mean = points, median = triangles, mode = pluses), for those studies that calculated quasi-extinction times. In panels (a) through (c), the bias represents overestimates of quasi-extinction probability by the diffusion approximation; in (d) the bias represents underestimates of quasi-extinction time by the diffusion approximation (sign was changed to allow plotting on a log scale).
Figure 5: The frequency with which the relative extinction risk is mis-classified by the diffusion approximation, as a function of the difference in extinction risk between the two populations. Based on all pairwise comparisons of the populations represented in the empirical data. Points indicate the fraction of comparisons mis-ranked within each interval on the horizontal axis; the line is the fitted probability from a logistic regression.
Table 1: Classification of species into IUCN threat categories, using $G(10) > 0.5$ and $d$ appropriate for 30% decline (Vulnerable), 50% decline (Endangered) or 80% decline (Critically Endangered), using $\mu$ and $\sigma^2$ from the empirical populations. Columns represent classification via simulations, and rows represent classification via the diffusion approximation.

<table>
<thead>
<tr>
<th>DA classification</th>
<th>simulation classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT/LC</td>
<td>25 0 0 0</td>
</tr>
<tr>
<td>V</td>
<td>8 7 0 0</td>
</tr>
<tr>
<td>E</td>
<td>1 11 23 0</td>
</tr>
<tr>
<td>CE</td>
<td>0 0 9 16</td>
</tr>
</tbody>
</table>

0.1 (Fig. 5). However, the error rate increases as the difference in risk gets small, to a maximum of nearly 30%.

Within each of the seven studies that analyzed more than two species or populations, the diffusion approximation misranked the relative risk using at least one of the risk measures (see Supporting Information). For example, the population of Fender’s blue butterfly (Schultz & Hammond, 2003) that had the sixth highest risk according to simulations had the eighth highest risk according to the diffusion approximation, and so would incorrectly have been excluded from a list of the top six or top seven (Fig. S2). Other ranking errors were of a similar magnitude.

The values of $\lambda$ required to ensure the desired low quasi-extinction probability in Fender’s blue butterfly are higher when estimated by the diffusion approximation than when estimated by simulations (Fig. 6). The bias ranges from 0.011 to 0.047.

Using the quasi-extinction probability as a measure of projected future decline under the IUCN red list, the diffusion approximation assigned 26% of Least Concern/Near Threatened species, 61% of Vulnerable species, and 41% of Endangered species to a higher risk category; none were assigned to a lower risk category (Table 1).

4 Discussion

Using four different metrics (probability of quasi-extinction by a certain time ($G$), mean ($T_c$) and median ($T_{med}$) times to quasi-extinction, and expected minimum population size (EMP) over a certain time), the diffusion approximation consistently overestimates the quasi-extinction risk in the stochastic exponential growth model. The bias was generally highest when the environmental variance ($\sigma^2$) was large. For some risk metrics, the bias was largest at short time horizons and intermediate quasi-extinction probabilities (the situations in which quantitative PVA is most likely to be useful). The bias approached zero
Figure 6: The value of $\lambda = \exp(\mu + \sigma^2/2)$ required to ensure that at least one population will persist for 100 years with a probability of 0.95, as a function of the number of independent populations and initial population size. $N_x = 1$ and $\sigma^2 = 0.79$. Open points are based on diffusion approximation, and filled points are based on simulations. Compare with Figure 4 of Schultz & Hammond (2003).
as \( \sigma^2 \) approached zero, or when the quasi-extinction probability was nearly zero or one. The existence of a bias was hinted at by Sabo et al. (2004), but the current paper is the first comprehensive documentation of the phenomenon and provides the first quantitative estimates of the bias magnitude.

Across a range of empirical applications to species of conservation concern, the bias in \( G \) was as much as 0.2 (which is a substantial fraction of the [0,1] probability scale). \( T_c \) and \( T_{med} \) could be overestimated by 20 years or more. The bias in EMP was as much as 70% of the true value. However, for many populations/species the bias was quite small, especially when the overall risk was high. In those populations for which I could calculate confidence intervals of extinction risk, the bias was never more than 20% of the width of the 95% confidence interval, suggesting that bias is a relatively minor component of the estimation error. However, whereas the confidence interval width can be reduced by using more data, the bias will always remain.

In two of the policy applications that I examined, the impact of the bias was modest. The relative risk of populations was sometimes misranked, but only if the risk difference between them was small. In such cases the choice of which population to protect will likely be driven by factors such as feasibility and cost. The biases in the estimates of \( \lambda \) suggested by Schultz & Hammond (2003) as recovery criteria for the Fender’s blue butterfly are small relative to the large growth rates required. However, the effect of the bias is equivalent to about a 70% overestimate of the initial population size required to achieve viability at a given level of \( \lambda \). Furthermore, in a slower-growing species, a bias in \( \lambda \) of 0.05 (corresponding to a 5% annual growth rate) might be considered large.

The largest potential policy impact of the bias was on IUCN redlist classifications when using the PVA to project future declines: overall, 29% of the species were classified to a higher-than-appropriate level by the diffusion approximation, and none were underclassified.

The primary source of the bias probably comes from the fact that with continuous breeding, the population is constantly moving up and down: it can drop below the quasi-extinction threshold during the year but then recover by the end of the year. This would be counted as an quasi-extinction by the stochastic differential equation (to which the diffusion approximation gives a good estimate). In discrete time, there is a pulse of recruitment each year, followed by gradual decline due to mortality; only if this endpoint is below the quasi-extinction threshold would the discrete-time model count an extinction event. The stochastic differential equation to which the diffusion approximation applies is a poor description of this sawtoothed trajectory within a year. Of course, the bias does not apply to species that breed continuously throughout the year.

The bias described here differs from the inaccuracies associated with a ‘diffusion approximation’ reported by Wilcox & Possingham (2002). In that model, incorporating negative density feedback and demographic stochasticity, the direction of the errors varied across parameter space. These biases were caused by deliberate discrepancies between the biological assumptions of the diffusion approximation and simulation models. They modeled absolute extinction, and
not encounter the phenomenon of populations recovering from quasi-extinction between censuses.

Many of the empirical values reported for $\sigma^2$ are likely to be overestimates — none were corrected for observation error in the abundance data, and some were based on time series of low-abundance populations for which demographic stochasticity may be inflating the variance. Observation error can inflate estimates of $\sigma^2$ by as much as 100% (H. Wilson unpublished results; also see examples in Lindley, 2003; Staples et al., 2004; Staudenmayer & Buonaccorsi, 2006). However, even if the true variances are half the reported estimates, many populations are still subject to high bias.

In summary, the diffusion approximation systematically overestimates the extinction risk for populations characterized by discrete breeding seasons. The only benefit of the diffusion approximation for these species is computational speed; and that is only relevant in large-scale simulation studies. Thus, I recommend that when using the stochastic exponential growth model to assess the extinction risk of real populations, the simulation approach should be used.

Acknowledgements

I thank Howard Wilson and other members of the Spatial Ecology Lab, as well as Corey Bradshaw and three anonymous reviewers, for comments that improved this paper. This research was supported in part by the US National Science Foundation under Grant No. 615024 and by the Spatial Ecology Laboratory at the University of Queensland.

References


