Contributed Papers

Reliability of Absolute and Relative Predictions of Population Persistence Based on Time Series

KELLY C. LOTTS,* THOMAS A. WAITE,*‡ AND JOHN A. VUCETICH†

*Department of Evolution, Ecology, and Organismal Biology, Ohio State University, 300 Aronoff Laboratory, 318 W. 12th Avenue, Columbus, OH 43210, U.S.A.
†School of Forest Resources and Environmental Sciences, Michigan Technological University, Houghton, MI 49931, U.S.A.

Abstract: *Conventional population viability analysis (PVA) is often impractical because data are scarce for many threatened species. For this reason, simple count-based models are being advocated. The simplest of these models requires nothing more than a time series of abundance estimates, from which variance and autocorrelation in growth rate are estimated and predictions of population persistence are generated. What remains unclear, however, is how many years of data are needed to generate reliable estimates of these parameters and hence reliable predictions of persistence. By analyzing published and simulated time series, we show that several decades of data are needed. Predictions based on short time series were very unreliable mainly because limited data yielded biased, unreliable estimates of variance in growth rate, especially when growth rate was strongly autocorrelated. More optimistically, our results suggest that count-based PVA is sometimes useful for relative risk assessment (i.e., for ranking populations by extinction risk), even when time series are only a decade long. However, some conditions consistently lead to backward rankings. We explored the limited conditions under which simple count-based PVA may be useful for relative risk assessment.*

Key Words: diffusion approximation, endangered species, extinction, population viability analysis, risk assessment, time series

Confiabilidad de Predicciones Absolutas y Relativas de la Persistencia de Poblaciones Basadas en Series de Tiempo

Resumen: *El análisis de viabilidad poblacional (AVP) a menudo es impráctico porque los datos para muchas especies amenazadas son escasos. Por esta razón, los modelos simples basados en conteos están siendo apoyados. El más simple de estos modelos requiere de nada más que una serie de tiempo de estimaciones de abundancia, a partir de las cual se estiman la varianza y autocorrelación de la tasa de crecimiento y se generan predicciones de la persistencia de la población. Sin embargo, aun no esta claro cuantos años de datos son necesarios para generar estimaciones confiables de estos parámetros y por tanto predicciones confiables de la persistencia. Analizando series de tiempo publicadas y simuladas, mostramos que se requieren datos de varias décadas. Predicciones basadas en series de tiempo cortas fueron muy poco confiables principalmente porque los datos limitados produjeron estimaciones sesgadas, poco confiables de la varianza de la tasa de crecimiento, especialmente cuando la tasa de crecimiento estaba fuertemente autocorrelacionada. Más optimistamente, nuestros resultados sugieren que el AVP basado en conteos a veces es útil para la evaluación de riesgo relativo (i.e. para clasificar poblaciones por riesgo de extinción), aun cuando la serie de tiempo sólo es de una década. Sin embargo, algunas condiciones conducen consistentemente a clasificaciones regresivas. Exploramos las condiciones limitadas bajos las que los AVP basados en conteos pueden ser útiles para la evaluación de riesgo relativo.*

Palabras Clave: análisis de viabilidad poblacional, aproximación difusión, especies en peligro, evaluación de riesgo, extinción, series de tiempo

1

Introduction

Population viability analysis (PVA) is no longer the exclusive domain of conservation biologists concerned with promoting the persistence of a particular species. With increasing frequency, it is being advocated as a tool in management and policy (Boyce 1992; Beissinger & McCullough 2002; Morris & Doak 2002; Shaffer et al. 2002). For example, three U.S. federal acts (the Federal Endangered Species Act [ESA] of 1973, the National Forest Management Act, and the Marine Mammal Protection Act) all require some unspecified form of risk analysis. Meanwhile, the World Conservation Union (IUCN) has developed risk assessment criteria that include risk of extinction (IUCN 1994). Some conservation biologists have recommended a formal role for PVA in listing and delisting decisions under ESA (Rohlf 1991; Thompson 1991; Scott et al. 1995; Gerber & DeMaster 1999; Shelden et al. 2001; Shaffer et al. 2002) and in IUCN assessment (Mace & Lande 1991; Mace et al. 1992; but see also Burgman et al. 1999; Keith et al. 2000), whereas others have criticized such recommendations because PVA can be inaccurate, expensive, and difficult to perform (e.g., Taylor 1995; Ralls et al. 1996; Beissinger & Westphal 1998; Coulson et al. 2001; Ellner et al. 2002; but see also Brook et al. 2000, 2002). These difficulties are pronounced for complex PVAs, which require estimates of many parameters (Ludwig & Walters 1985; Harcourt 1995; Beissinger & Westphal 1998; Ludwig 1999; Meir & Fagan 2000) and hence more data than are usually available (Mace & Lande 1991; Harcourt 1995; Morris et al. 2002). Because simple PVAs have minimal data requirements, it may be appropriate to increase our use of these models (Holmes 2001; Belovsky et al. 2002; Morris & Doak 2002; Morris et al. 2002).

The simplest of the count-based PVAs requires nothing more than count data. This model predicts mean time to extinction (MTE) from a time series of abundance estimates (Dennis et al. 1991; Foley 1994). Various limitations of this model have been evaluated (Sæther et al. 1998; Hakoyama & Iwasa 2000; Wilcox & Possingham 2002), but no study to date has asked how many count data are needed for adequate performance (see, however, Taylor **Q1** 1995; Moilanen et al. 1998; Thomas et al. 2002). We evaluated the number of years of abundance estimates required to generate reliable estimates of variance in growth rate, a key parameter in the model. We also evaluated how serial autocorrelation influences the number of years of data required. Finally, we evaluated whether the model is useful for relative risk assessment.

Diffusion Approximation

For all analyses, we used Foley's (1994) model, which is based on diffusion analysis (Dennis et al. 1991) and incorporates both density dependence and carrying capacity. We simulated time series with the population model (Eq. 2 in Foley 1994)

$$n_{t+1} = r_t + n_t, \tag{1}$$

where n_t is the natural logarithm of population size (N) at time t and r_t is the ln-transformed growth rate. Population size (n) is constrained to the interval $(0, k)$, and the population is considered extinct when $n = 0$. For the simple case of no autocorrelation in growth rate, r_t is an independently distributed normal random variable with mean μ and variance v_r. We assumed $\mu = 0$ throughout. The model uses abundance data to estimate MTE by the relationship (Eq. 6 in Foley 1994)

$$MTE = \frac{2n_0}{v_r}\left(k - \frac{n_0}{2}\right), \tag{2}$$

where n_0 is initial population size, v_r is variance in growth rate, and k is the population ceiling. Equation 2 applies when the long-term population growth rate is approximately zero. All parameters were calculated following ln-transformation of abundance estimates. Variance in growth rate was calculated directly from the measurements of r. To correct for errors in v_r resulting from any correlation between r_t and r_{t+1}, autocorrelation ρ was estimated for each series and used to calculate effective variance (Eq. 10 in Foley 1994):

$$v_{re} \approx \frac{1+\rho}{1-\rho}v_r, \tag{3}$$

which is substituted for v_r in Eq. 2.

Traditional diffusion approximation assumes that (1) population trajectories follow Brownian motion; (2) variance in population size (N) and v_r are stationary (Dennis et al. 1991); (3) successive population growth rates are uncorrelated; (4) growth rate is unaffected by population density; and (5) growth rate exhibits no long-term trend (for a complete list of assumptions see Holmes 2001; Morris & Doak 2002). Foley's (1994) method accounts for violations in two of these assumptions. First, any assumption about variance in N is extraneous because a population's behavior is modeled from its growth rate r, which is calculated from the first difference of n. Second, serial autocorrelation in r is corrected with Eq. 3. Here, we evaluated the performance of this correction. Violations of other assumptions were left uncorrected.

Methods

Global Population Dynamics Database

We selected published time series of abundance from the Global Population Dynamics Database (GPDD; Natural Environment Research Council [NERC] 2002), **Q2** which is based on taxonomic group (Mammalia), length

(minimum 20 years), and measurement type (census or population abundance estimate). We excluded time series generated from records of fur harvest or mortality. Altogether, we chose 34 continuous time series representing 21 species.

To explore how length of time series affects estimates of MTE, we first used a moving window to generate as many segments as possible from each original time series. This method is analogous to a jackknife procedure. For example, a 20-year time series was divided into 16 5-year segments (using years 1–5, 2–6, and so on), 15 6-year segments, and so on. Then, a prediction of MTE was calculated for each segment with Eq. 2. We did not intend to generate realistic predictions for any particular time series. Therefore, we arbitrarily set carrying capacity at $k = 15$ ($\sim$3 million), a value far higher than the largest observed population size, and we scaled from 0 to 1 each MTE estimate originating from a particular time series (by dividing it by the maximum predicted MTE for that series). These steps guaranteed that extinction dynamics were governed entirely by variation in r and not by k or average r.

Simulated Time Series

To simulate autocorrelated growth rates, r_t (in Eq. 1) is replaced by (Eq. A26 in Foley 1994):

$$r_t = \rho r_{t-1} + \varepsilon_t, \tag{4}$$

where ρ is the correlation between r_t and r_{t-1}, and ε_t is an independently distributed normal random variable with mean zero and variance $(1 - \rho^2)v_r$. Initial population size and carrying capacity were set at 50 ($n_0 = 3.9$) and 100 ($k = 4.6$), respectively. Values for v_r and ρ were selected based on the 5th, 50th, and 95th percentiles of the GPDD time series (Fig. 1). Using these biologically realistic values, we simulated 1000 50-year time series for each of six combinations of v_r and ρ, in each case holding one parameter constant to evaluate the role of the other.

Estimating Variance in Growth Rate

We used these simulations to assess how estimates of variance in r and effective variance in r are affected by length of time series. We also evaluated how ρ affects estimates of both v_r and v_{re}. In addition, we identified how estimates of v_{re} are affected by uncertainty in estimates of autocorrelation ($\hat{\rho}$). This final analysis evaluated the performance of the autocorrelation correction (Eq. 2).

Estimating Population Persistence

We also used the simulated time series to evaluate how many years of count data are needed to generate reliable predictions of population persistence. To do so, we calculated two types of MTE predictions for each of six combinations of parameter values (Fig. 1). First, the asymptotic
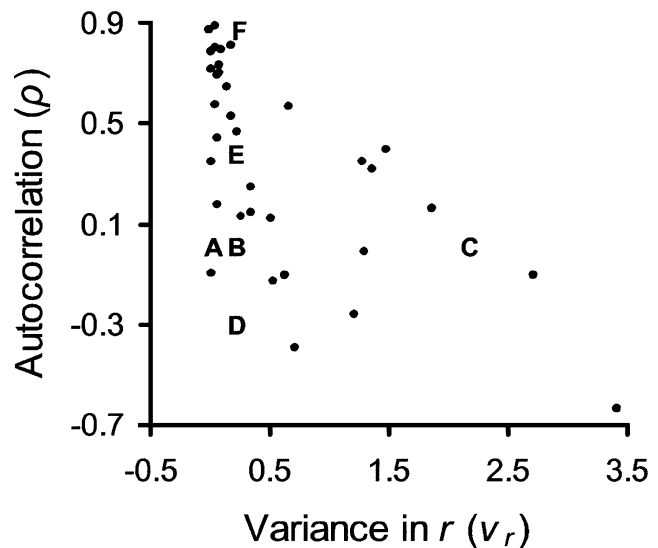

Figure 1. Relationship between ρ and v_r (see Eq. 3 for definition of v_r) for time series from the Global Population Dynamics Database. Letters correspond to parameter-value combinations used in simulations and relative-risk assessments.

MTE of the underlying process (MTE$_{asym}$) was calculated using these specified values of ρ, k, n_0, and v_r. Second, an MTE prediction (MTE$_t$) was generated for each length of the time series (10–50 years) with specified values of ρ, k, and n_0 and estimated variance in growth rates ($\hat{v}_r$). Then, isolating the effect of uncertainty in v_r on predictions of MTE, we quantified comparisons between MTE$_t$ and MTE$_{asym}$ predictions as percent prediction error:

$$\left(\frac{MTE_t - MTE_{asym}}{MTE_{asym}}\right) \times 100. \tag{5}$$

Finally, we used the simulated time series to evaluate how uncertainty in both v_r and ρ affect estimation of MTE. To do so, we used k, n_0, $\hat{v}_r$, and estimated autocorrelation ($\hat{\rho}$) to estimate MTE (MTE$_{est}$). We also used this measure of MTE in the analysis of relative risk assessment.

Recognizing that the probability of extinction within some time frame is the preferred measure of population persistence for some conservationists, we extended our analysis to include this metric. We converted predictions of MTE$_t$ to estimates of probability of extinction within X ($=50$) years with Vucetich & Waite's (1998a, 1999) expression

$$P\langle E \mid X \rangle = -\exp\left(\frac{-X}{MTE_t}\right) + 1, \tag{6}$$

which assumes that times to extinction are exponentially distributed.
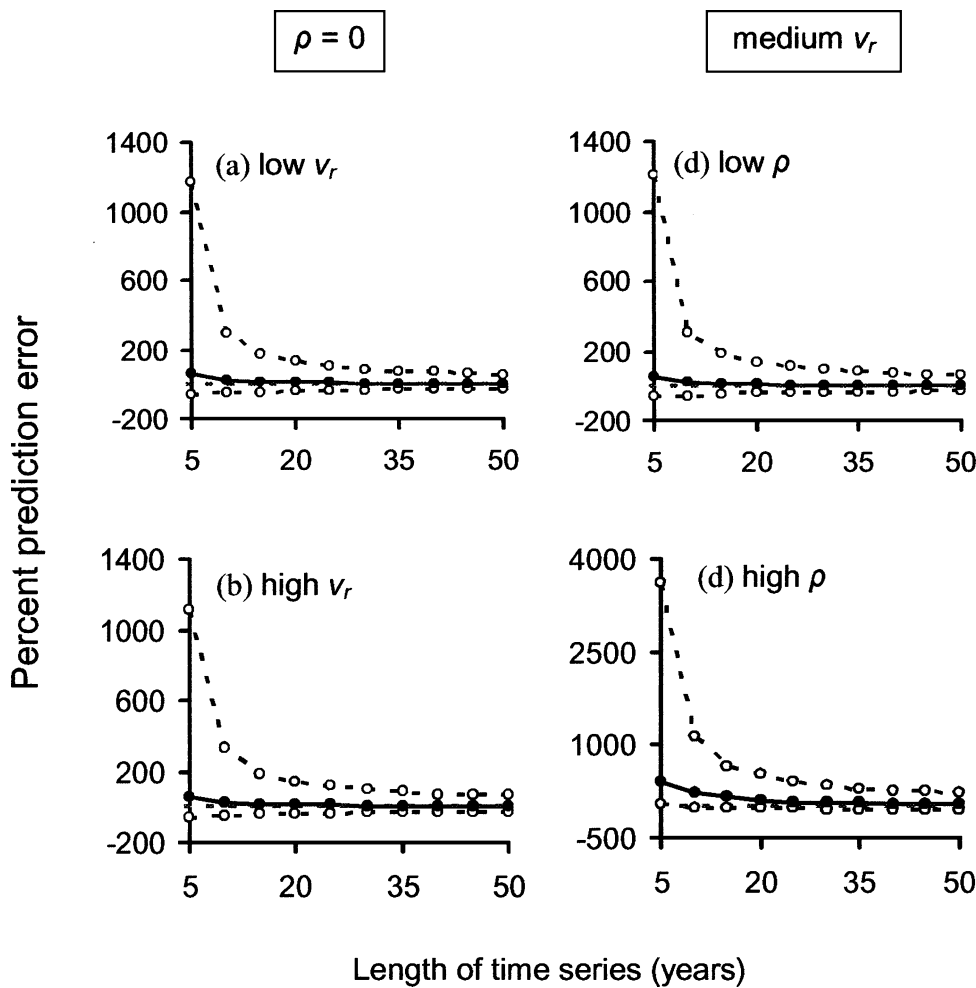
$\rho = 0$ medium v_r

Figure 2. Relationship between length of time series (years) and the percent difference between asymptotic mean time to extinction (MTE_{asym}) predicted by known parameters of the simulated time series and mean time to extinction estimated by time series of each shorter segment (MTE_t, an MTE prediction) as calculated by Eq. 5. The 5th, 50th, and 95th percentiles are shown. Panels (a) and (b) show prediction error for time series simulated at zero autocorrelation between growth rates ($\rho = 0$) and (a) low (0.01) and (b) high (2.16) v_r. Panels (c) and (d) show prediction error for time series simulated with medium (0.21) v_r and (c) low (−0.31) and (d) high (0.83) levels of ρ. Note the y-axis scale in (d) differs from the scale in a–c. Asymptotic prediction error occurs at y = 0.

Relative Risk Assessment

To examine whether Foley's model could be used to make reliable assessments of relative extinction risk based on short (10-year) time series, we initially tested the model's ability to correctly rank populations by MTE for 15 realistic combinations of v_r and ρ (Fig. 1; see also Ariño & Pimm 1995). We used MTE_{est} here because it requires estimates of both v_r and ρ and is thus the most sensible metric for real data. For each pair of populations considered, we determined the correct ranking by comparing the true underlying risks as estimated by MTE_{asym}. Pairs of MTE_{est} predictions based on 10-year simulated time series were compared in all possible combinations, and we quantified the proportion of 1000 simulations that were correctly ranked for each combination. In this preliminary analysis, we found that under some conditions the ranking of populations could be consistently backward (i.e., worse than chance). These preliminary results are not reported here. Instead, we present the results from a follow-up analysis, designed to explore more fully the conditions favoring both highly accurate as well as backward ranking of populations by risk.
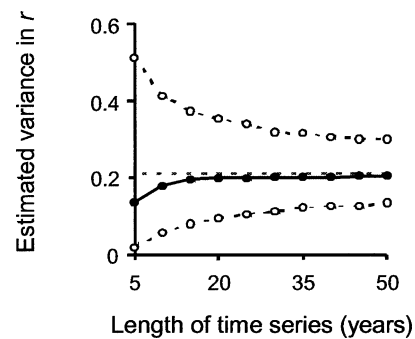

Figure 3. Relationship between length of time series (years) and estimated variance in growth rate ($\hat{v}_r$).Comp: no italics. The 5th, 50th, and 95th percentiles are shown. Graph shows $\hat{v}_r$ for time series simulated at medium (0.21) variance in growth rate v_r and zero autocorrelation between growth rates ($\rho = 0$). Dotted line without symbols represents true underlying v_r..
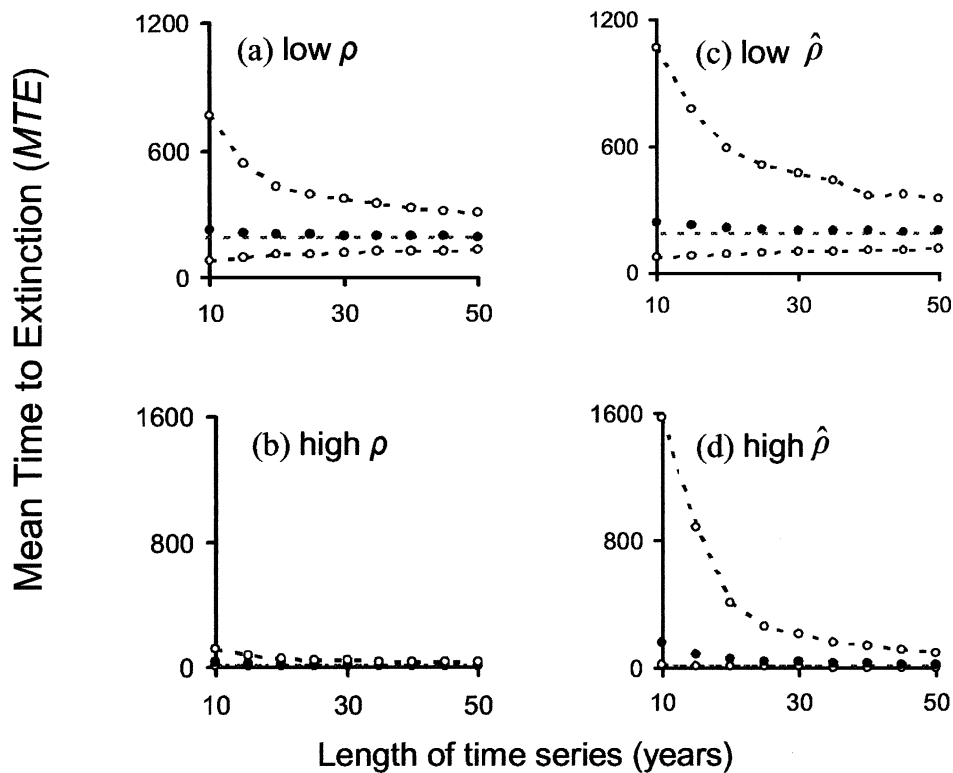
Figure 4. Relationship between length of time series (years) and MTE (mean time to extinction). The 5th, 50th, and 95th percentiles are shown. Panels show MTE_t (an MTE prediction) for time series simulated at medium (0.21) v_r and (a,b) low (−0.31) and (c,d) high (0.83) autocorrelation ρ. The MTE_t predictions in panels (a), (b), (c), and (d) were calculated using ρ and $\hat{v}_{re}$ and $\hat{\rho}$ and $\hat{v}_{re}$, respectively. Dotted line without symbols represents MTE_{asym} (asymptotic MTE), calculated using parameter values v_r and ρ.

Results and Interpretation

Estimating Variance in Growth Rate and Population Persistence

Analysis of both GPDD time series (not shown) and simulated time series (Fig. 2) revealed that predictions of MTE based on short segments were extremely unreliable because of bias and sampling error in estimation of v_r (Fig. 3). Most MTE predictions based on short time series exceeded predictions based on longer time series, but for any given short time series, MTE could be very large or very small. Short time series produced highly unreliable predictions of long-term MTE, particularly when population growth rate was strongly positively autocorrelated (Fig. 4). Overall, predictions of extinction risk (including $P(E \mid 50)$ [Fig. 5]) tended to be overly optimistic, although any individual prediction could either overestimate or vastly underestimate extinction risk.

Relative Risk Assessment

Relative risk assessment was most accurate when two populations differed substantially in MTE_{est} (Fig. 6) and in underlying v_r (Fig. 7). Indeed, even when their MTEs differed only moderately, pairs of simulated populations were ranked correctly nearly every time, provided they had different underlying levels of v_r (i.e., provided the two populations differed by at least one level of the five assumed values [0.01, 0.06, 0.21, 0.70, 2.16]). However, in some cases, two populations with different underlying

v_r were ranked incorrectly most of the time (e.g., see bars corresponding to $v_r = 0.06$ in Fig. 7c). These incorrect rankings occurred both when populations' MTEs were similar and when they were vastly different (Fig. 7). When populations were the same or similar in underlying v_r but differed by even one level of ρ, they were correctly ranked <75% (Fig. 7) and sometimes < 50% (Fig. 7c,d) of the time. These results highlight some serious limitations to using this model for relative risk assessment.

Discussion

Effects of Length of Time Series on Predictions of Persistence

The reliability of persistence predictions increases with length of time series because the reliability of parameter estimates increases with sample size. Short time series do not adequately estimate population parameters (e.g., Vucetich & Waite 1998b) such as variance in r, leading instead to downwardly biased and extremely uncertain estimates of variance in r (Fig. 3). Short time series also lead to biased estimates of autocorrelation (Foley 1994). Together, these biases lead to overly optimistic predictions of population persistence (Figs. 4 & 5). All count-based methods of PVA are likely to be biased in this way because they all rely on an estimate of variance in r (Morris & Doak 2002). Beyond this bias, great uncertainty in parameter estimates translates into great uncertainty in extinction risk predictions for any given short time series.
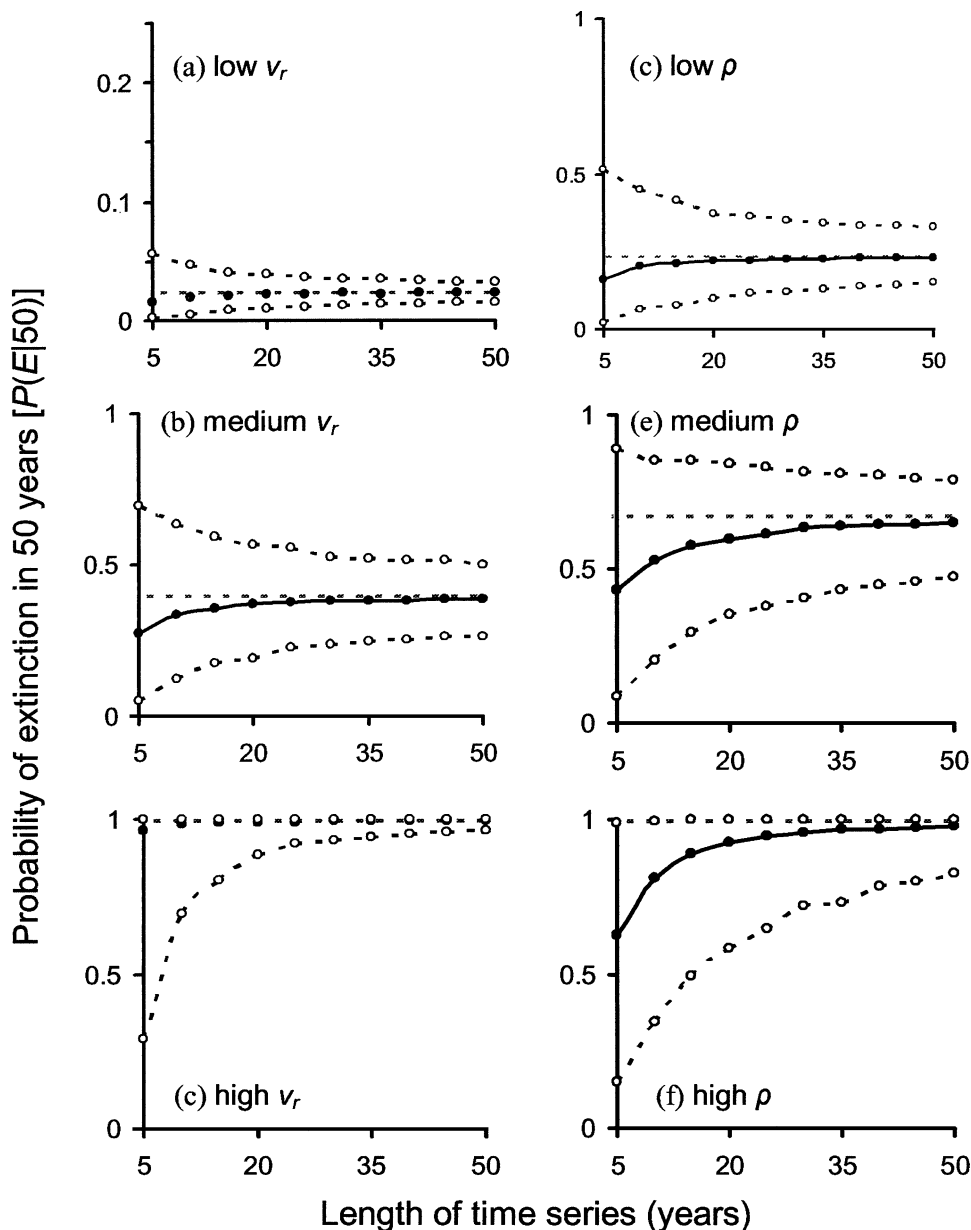
Figure 5. Relationship between length of time series (years) and probability of extinction P(E | 50)$_{asym}$ *converted from* MTE$_{asym}$ *(asymptotic mean time to extinction) and* P(E | 50)$_t$ *converted from* MTE$_t$ *(an MTE prediction). The 5th, 50th, and 95th percentiles are shown. Panels (a)–(c) show* P(E | 50)$_t$ *for time series simulated at* $\rho = 0$ *and (a) low (0.01), (b) medium (0.21), and (c) high (2.16)* v$_r$*. Note scale of y-axis in panel (a) differs from scale in (b)–(c). Panels (d)–(f) show* P(E | 50)$_t$ *for time series simulated with medium (0.21)* v$_r$ *and (d) low (−0.31), (e) medium (0.37), and (f) high (0.83) levels of* ρ*.*

Our study builds on earlier work showing that estimates of variance in *r* are inflated at small population sizes and deflated around the population ceiling (Wilcox & Possingham 2002). Despite these and other considerations (Middleton & Nisbet 1997; Hakoyama & Iwasa 2000; Sæther et al. 2000), it remains common practice to use the present method for estimating *r* and variance in *r* (Wilcox & Possingham 2002, but see also Holmes 2001). Count-based PVA should be performed and interpreted with these limitations in mind.

Implications for Absolute and Relative Threat Assessment

Our findings present a challenge to conservation biologists who aim to err on the side of caution. Although simple count-based PVAs have been advocated for use in threat assessment because they require so few data (Allen et al. 1992; Belovsky et al. 2002; Morris & Doak 2002), the usefulness of these models depends strongly on the amount of count data available. Predictions of population persistence tend to be overly optimistic, not cautious, and any given prediction is highly dubious (see also Ludwig 1999; Taylor et al. 2002). Our analysis suggests that prediction error may be acceptably small only if several decades of count data are available. Thus, absolute risk assessment based on short time series may be worse than useless, especially because the unwary may place too much confidence in predictions generated by a complex computational procedure.
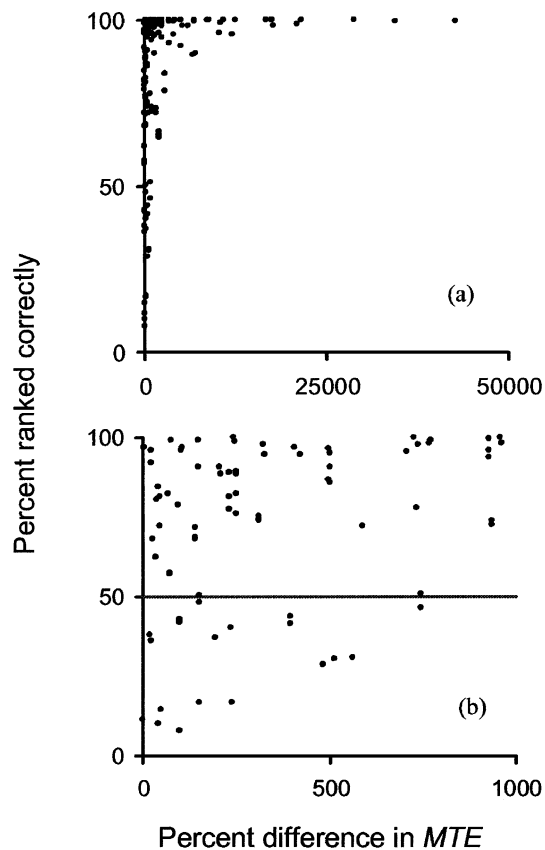
Figure 6. Percent of rankings by MTE$_{est}$ (estimated mean time to extinction) that were correct relative to percent difference between MTE$_{asym}$ (asymptotic MTE) for the two 10-year time series. Note x-axis scale differs in two panels.

Because predictions improve with length of time series, it is tempting for us to recommend massive data-collection efforts. With such data, conservation managers would have the option of using count-based models for risk assessment. However, more data are no panacea. Even if the parameters could be estimated perfectly, confidence intervals around MTE estimates are extremely wide (95% CI: 2.5–370% of MTE; Vucetich & Waite 1998*a*, 1999). Reflecting this limitation, some workers have cautioned that simple PVAs should be used only to weigh the relative usefulness of management options, assess relative extinction risk, or make qualitative risk assessments (Beissinger & Westphal 1998; Fagan et al. 1999; Fieberg and Sæther et al. 2000; Belovsky et al. 2002; Reed et al. 2002; Wilcox & Possingham 2002; McCarthy et al. 2003).

Our study (see also McCarthy et al. 2003) reinforces the general recommendation that count-based PVA should be used for predicting relative, rather than absolute, extinction risk (for specific recommendations see Table 2 in Beissinger & Westphal 1998). Our results show that relative extinction risk of pairs of populations can be assessed effectively, even with only 10 years of count data, pro-

vided the populations differ sufficiently in their levels of risk. If two populations' MTEs differ by an order of magnitude, they are ranked correctly virtually every time (Fig. 6), suggesting that the method may work best when it is not even needed. Pairs of populations can also be consistently ranked correctly if they differ in underlying variance in *r* (Fig. 7). We emphasize, however, that in some cases pairs of populations are correctly ranked less than half the time. Incorrect rankings occur when two populations differ in both variance in *r* and autocorrelation. Typically, when time series are short and when autocorrelation in *r* is strongly positive, variance in *r* is underestimated and MTE is overestimated. Thus, when one of two populations under comparison has strongly autocorrelated *r*, that population's MTE tends to be vastly overestimated. The consequence is a consistently reversed ranking of the two populations' true levels of extinction risk (Fig. 7). Therefore, under some conditions, the relative risk faced by two populations would be more accurately assessed by a coin toss than by ranking based on MTE predictions. Additional analysis is needed before offering any prescription as to how much count data are needed or how pairs of populations must compare in their parameter values.

Future Directions

Our study shows that, under some conditions, a simple count-based model can be used to assess relative extinction risk. This guardedly optimistic conclusion may prove too optimistic. After all, we purposely chose to evaluate the simplest of all count-based PVA models. More realistic models require estimation of additional parameters and thus may yield even less reliable predictions. Future studies should include a detailed evaluation of the utility of both simple count-based models and more complex models for relative risk assessment and should evaluate how many years of data are required for adequate assessment. Such studies should investigate whether these models yield reliable rankings when there are more than two populations and whether they do so for other taxa and life histories (Wilcox & Possingham 2002). Ultimately, it may be prudent to use crude methods of relative risk assessment, which would have the advantage of inspiring only limited confidence. More optimistically, extensive simulation studies may lead to the development of clear guidelines for relative risk assessment based on simple count-based models.
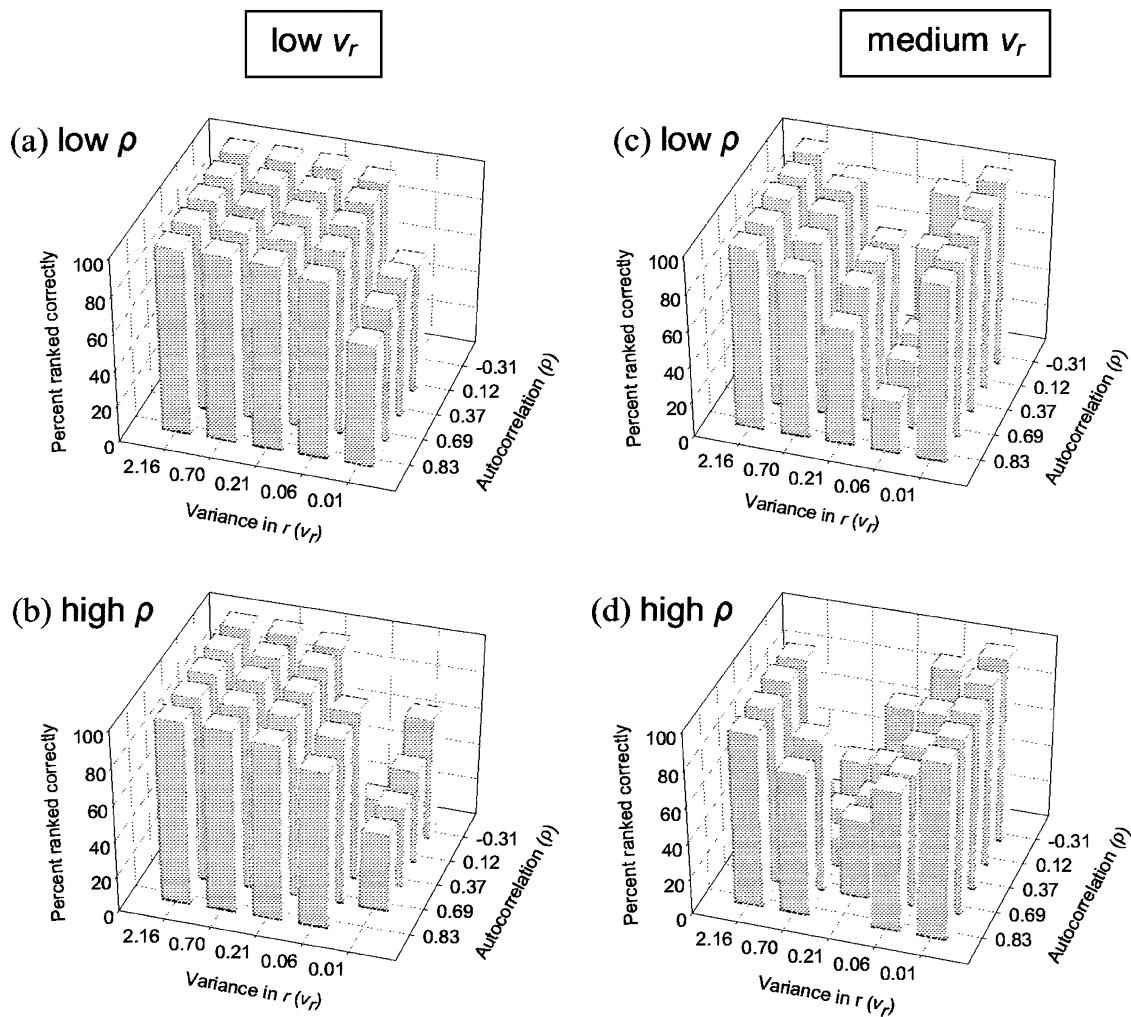
Acknowledgments

Figure 7. Percent of rankings by MTE$_{est}$ (estimated mean time to extinction) that were correct. Panels show accuracy of relative risk assessment when simulated time series with one set of parameters ("standard") were compared with simulated 10-year time series with different underlying parameters. Standard series were simulated with (a)–(b) low (0.01) v$_r$ and (c)–(d) medium (0.21) v$_r$, and with (a, c) low (−0.31) and (b, d) high (0.83) ρ. Each panel has one missing bar where the standard series was not compared with itself.

Biology and their collaborators for making available the Global Population Dynamics Database.

Literature Cited

Allen, E. J., J. M. Harris, and L. J. S. Allen. 1992. Persistence-time models for use in viability analyses of vanishing species. Journal of Theoretical Biology **155:**33–53.

Ariño, A., and S. L. Pimm. 1995. On the nature of population extremes. Evolutionary Ecology **9:**429–433.

Beissinger, S. R. 2002. Population viability analysis: past, present, future. Pages 5–17 in S. R. Beissinger and D. R. McCullough, editors. 2002. Population viability analysis. University of Chicago Press, Chicago.

Beissinger, S. R., and D. R. McCullough, editors. 2002. Population viability analysis. University of Chicago Press, Chicago.

Beissinger, S. R., and M. I. Westphal. 1998. On the use of demographic models of population viability in endangered species management. Journal of Wildlife Management **62:**821–841.

Belovsky, G. E., C. Mellison, C. Larson, and P. A. Van Zandt. 2002. How good are PVA models? Testing their predictions with experimental data on the brine shrimp. Pages 257–283 in S. R. Beissinger and D. R. McCullough, editors. Population viability analysis. University of Chicago Press, Chicago.

Boyce, M. S. 1992. Population viability analysis. Annual Review of Ecology and Systematics **23:**481–506.

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

Brook, B. W., M. A. Burgman, H. R. Akçakaya, J. J. O'Grady, and R. Frankham. 2002. Critiques of PVA ask the wrong questions: throwing the heuristic baby out with the bath water. Conservation Biology **16:**262–263.

Burgman, M. A., D. A. Keith, F. J. Rohlf, and C. R. Todd. 1999. Probabilistic classification rules for setting conservation priorities. Biological Conservation **89:**227–231.

Coulson, T., G. M. Mace, E. Hudson, and H. Possingham. 2001. The use and abuse of population viability analysis. Trends in Ecology & Evolution **16:**219–221.

Dennis, B., P. L. Munholland, and J. A. Scott. 1991. Estimation of growth

Q4

and extinction parameters for endangered species. Ecological Monographs **6:**115–143.

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

Fagan, W. F., E. Meir, and J. Moore. 1999. Variation thresholds for extinction and their implications for conservation strategies. The American Naturalist **154:**510–520.

Fieberg, J., and S. P. Ellner. 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. Conservation Biology **8:**124–137.

Gerber, L. R., and D. P. DeMaster. 1999. A quantitative approach to endangered species act classification of long-lived vertebrates: application to the North Pacific humpback whale. Conservation Biology **13:**1203–1214.

Hakoyama, H., and Y. Iwasa. 2000. Extinction risk of a density-dependent population estimated from a time series of population size. Journal of Theoretical Biology **204:**337–359.

Harcourt, A. H. 1995. Population viability estimates: theory and practice for a wild gorilla population. Conservation Biology **9:**134–142.

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

IUCN (World Conservation Union). 1994. IUCN Red List categories. IUCN Species Survival Commission, Gland, Switzerland.

Keith, D. A., T. D. Auld, M. K. J. Ooi, and B. D. E. Mackenzie. 2000. Sensitivity analysis of decision rules in World Conservation Union (IUCN) Red List criteria using Australian plants. Biological Conservation **94:**311–319.

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

Ludwig, D., and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences **42:**1066–1072.

Mace, G., and R. Lande. 1991. Assessing extinction threats: towards a reevaluation of IUCN threatened species categories. Conservation Biology **5:**148–157.

Mace, G., N. Collar, J. Cooke, K. Gaston, J. Ginsberg, N. Leader Williams, M. Maunder, and E. J. Milner-Gulland. 1992. The development of new criteria for listing species on the IUCN Red List. Species **19:**16–22.

McCarthy, M. A., S. J. Andelman, and H. P. Possingham. 2003. Reliability of relative predictions in population viability analysis. Conservation Biology **17:**982–989.

McCarthy, M. A., H. P. Possingham, J. R. Day, and A. J. Tyre. 2001. Testing the accuracy of population viability analysis. Conservation Biology **15:**1030–1038.

McCarthy, M. A., and C. Thompson. 2001. Expected minimum population size as a measure of threat. Animal Conservation **4:**351–355.

Meir E., and W. F. Fagan. 2000. Will observation error and biases ruin the use of simple extinction models? Conservation Biology **14:**148–154.

Middleton, D. A. J., and R. M. Nisbet. 1997. Population persistence time: estimates, models, and mechanisms. Ecological Applications **7:**107–117.

Moilanen, A., A. T. Smith, and I. Hanski. 1998. Long-term dynamics in a metapopulation of the American pika. The American Naturalist **152:**530–542.

Morris, W. F., P. L. Bloch, B. R. Hudgens, L. C. Moyle, and J. R. Stinchcombe. 2002. Population viability analysis in endangered species

recovery plans: past use and future improvements. Ecological Applications **12:**708–712.

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

NERC Centre for Population Biology, Imperial College. 1999. The Global Population Dynamics Database. NERC Centre for Population Biology, Imperial College, London. Available from http://cpbnts1.bio.ic.ac.uk/gpdd/(accessed February 2004).

Ralls, K., D.P. Demaster, and J. A. Estes. 1996. Developing a criterion for delisting the southern sea otter under the US endangered species act. Conservation Biology **10:**1528–1537.

Reed, J. M., L. S. Mills, J. B. Dunning Jr., E. S. Menges, K. S. McKelvey, R. Frye, S. R. Beissinger, M. Anstett, and P. Miller. 2002. Emerging issues in population viability analysis. Conservation Biology **16:**7–19.

Rohlf, D. J. 1991. Six biological reasons why the endangered species act doesn't work—and what to do about it. Conservation Biology **5:**273–282.

Sæther, B. E., S. Engen, A. Islam, R. McCleery, and C. Perrins. 1998. Environmental stochasticity and extinction risk in a population of a small songbird, the great tit. The American Naturalist **151:**441–450.

Sæther, B. E., S. Engen, R. Lande, P. Arcese, and J. N. M. Smith. 2000. Estimating the time to extinction in an island population of song sparrows. Proceedings of the Royal Society of London Series B **267:**621–626.

Scott, J. M., T. H. Tear, and L. S. Mills. 1995. Socioeconomics and the recovery of endangered species: biological assessment in a political world. Conservation Biology **9:**214–216.

Shaffer, M. L., L. H. Watchman, W. J. Snape III, and I. K. Latchis. 2002. Population viability analysis and conservation policy. Pages 123–142 in S. R. Beissinger and D. R. McCullough, editors. Population viability analysis. University of Chicago Press, Chicago.

Shelden, K. E. W., D. P. DeMaster, D. J. Rugh, and A. M. Olson. 2001. Developing classification criteria under the US Endangered Species Act: Bowhead whales as a case study. Conservation Biology **15:**1300–1307.

Taylor, B. L. 1995. The reliability of using population viability analysis for risk classification of species. Conservation Biology **9:**551–558.

Taylor, B. L., P. R. Wade, U. Ramakrishnan, M. Gilpin, and H. R. Akçakaya. 2002. Incorporating uncertainty in population viability analyses for the purpose of classifying species by risk. Pages 239–252 in S. R. Beissinger and D. R. McCullough, editors. Population viability analysis. University of Chicago Press, Chicago.

Thompson, G. G. 1991. Determining minimum viable populations under the Endangered Species Act. Technical memorandum NMFS F/NWC-198. National Oceanic and Atmospheric Administration, Seattle, Washington.

Vucetich, J. A., and T. A. Waite. 1998*a*. On the interpretation and application of mean times to extinction. Biodiversity and Conservation **7:**1539–1547.

Vucetich, J. A., and T. A. Waite. 1998*b*. Number of censuses required for demographic estimation of effective population size. Conservation Biology **12:**1023–1030.

Vucetich, J. A., and T. A. Waite. 1999. On the interpretation and application of mean extinction times (Volume 7, pp. 1539–1547 [Erratum]). Biodiversity and Conservation **8:**1005.

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

Queries

Q1 Author: Thomas et al. 2002 not listed in Literature Cited.

Q2 Author: NERC 2002 year differs in Literature Cited.

Q3 Author: A two-person reference wouldn't take "et al." See also comment in Literature Cited; should this be "Fieberg and Ellner 2000"?

Q4 Author: Beissinger, 2002 not cited in text.

Q5 Author: Fieberg and Ellner. 2000 not cited in text; there is a cite for Fieber and Saether et al. 2000, pg. 6 of ms.

Q6 Author: McCarthy et al. 2001 not cited in text.

Q7 Author: McCarthy and Thompson. 2001 not cited in text.