

# STATISTICAL REPORTS

*Ecology*, 85(4), 2004, pp. 923–929  
© 2004 by the Ecological Society of America

## ESTIMATING POPULATION TREND AND PROCESS VARIATION FOR PVA IN THE PRESENCE OF SAMPLING ERROR

DAVID F. STAPLES,<sup>1,3</sup> MARK L. TAPER,<sup>1</sup> AND BRIAN DENNIS<sup>2</sup>

<sup>1</sup>*Department of Ecology, 310 Lewis Hall, Montana State University, Bozeman, Montana 59717 USA*

<sup>2</sup>*CNR Room 105, University of Idaho, Moscow, Idaho 83844 USA*

**Abstract.** Time series of population abundance estimates often are the only data available for evaluating the prospects for persistence of a species of concern. With such data, it is possible to perform a population viability analysis (PVA) with diffusion approximation methods using estimates of the mean population trend and the variance of the trend, the so-called process variation. Sampling error in the data, however, may bias estimators of process variation derived by simple methods. We develop a restricted maximum likelihood (REML)-based method for estimating trend, process variation, and sampling error from a single time series based on a discrete-time model of density-independent growth coupled with a model of the sampling process. Transformation of the data yields a conventional linear mixed model, in which the variance components are functions of the process variation and sampling error. Simulation results show essentially unbiased estimators of trend, process variation, and sampling error over a range of process variation/sampling error combinations. Example data analyses are provided for the Whooping Crane (*Grus americana*), grizzly bear (*Ursus arctos horribilis*), California Condor (*Gymnogyps californianus*), and Puerto Rican Parrot (*Amazona vittata*). This REML-based method is useful for PVA methods that depend on accurate estimation of process variation from time-series data.

**Key words:** measurement error; mixed linear model; modeling; population growth rate; population viability analysis; process variation; PVA; sampling error; time series; trend estimation.

### INTRODUCTION

Conservation or management policies generally require assessment of a population's prospects for persistence. Unfortunately, these assessments often must be made with little data. Commonly, the most extensive data available are count-based data in which the entire population or a subset of the population is sampled over multiple years. Count-based data may include time series of population abundance estimates, catch-per-unit effort, or samples of a portion of the life cycle, e.g., spawning redds, nesting adults, or mother-cub pairs. Dennis et al. (1991) presented methods for fitting an exponential growth model to count-based data and techniques to perform a population viability analysis (PVA) with the fitted model parameter estimates using diffusion approximations (DA). These PVA techniques include the calculation of risk metrics such as the mean time to extinction, distribution of extinction times, probability of hitting a lower threshold (not necessarily extinction), or the distribution of the population density

in the future. The methods presented in Dennis et al. (1991) have been used to calculate risks for a variety of species (Dennis et al. 1991, Nicholls et al. 1996, Gerber et al. 1999, Morris and Doak 2002).

Count-based PVAs using diffusion approximation (DA) methods are useful for analyzing risk to a population. However, it has been noted that these methods are vulnerable to sampling or measurement error in the data. Ludwig (1999) found that errors in estimated abundance could increase confidence intervals for quasi-extinction to the point at which estimates become meaningless. Meir and Fagan (2000) found that sampling error had its largest effect on estimates of extinction risk for cases in which it was uncertain whether the population would go extinct, presumably a situation applicable to many species of conservation interest.

PVA techniques with the Dennis et al. (1991) model depend on the estimates of two population parameters in the exponential growth model: the long-term geometric mean growth rate and the variance of the realized growth due to environmental effects. Estimation of the latter parameter, termed process variation, is confounded by sampling error in the data. In this paper, the term "sampling error" is used to indicate any deviation of the observed count from the true population

Manuscript received 9 July 2003; revised 23 September 2003; accepted 29 September 2003. Corresponding Editor: A. M. Ellison.

<sup>3</sup> E-mail: staples@montana.edu

size. This includes error from sampling only a portion of the population (the traditional definition of sampling error) in addition to error from incorrect measurements within a chosen sampling unit (conventionally called observation, or measurement, error). Sampling error adds to the variability in the data leading to positively biased estimators of process variation in the model of Dennis et al. (1991). This bias leads to poor predictive precision and pessimistic PVA projections. An ability to partition variation in the data due to sampling error from variation due to environmental effects would greatly improve the quality of count-based PVA.

Holmes (2001) described a regression method for estimating process variation and sampling error from a single time series of population estimates. Holmes' method, however, does not necessarily produce parameter estimates with desirable statistical properties. A particularly worrisome feature of the Holmes method is that estimators of process variation can be biased high or low and the magnitude and direction of bias is unknown in actual applications. For more detailed discussions of the Holmes method, see Morris and Doak (2002) and Holmes and Fagan (2002).

We present a new restricted maximum likelihood (REML)-based method for estimating trend, process variation, and sampling error variance from a single time series of count-based data that can be performed in statistical computing packages such as SAS PROC MIXED (Littell et al. 1996). These parameter estimates are obtained by embedding a model of the sampling process within a model of population growth. This allows the estimation of process variation while accounting for variation in the data due to sampling error. The method presented, however, has data requirements more strict than those of Dennis et al. (1991) in that the REML-based method requires uniformly spaced population estimates in time, with no missing data points.

#### METHODS

The underlying change in population density over time is modeled with a stochastic, discrete-time model of exponential growth. This model represents an assumption of density-independent growth for the population, a reasonable assumption for depleted populations. The population density at time  $(t + 1)$  is thus

$$\mathbf{N}(t + 1) = \mathbf{N}(t) \exp(\mu) \mathbf{E}(t) \quad (1)$$

for  $t = 1, 2, \dots, T$ ; where  $\mathbf{N}(t)$  represents the actual population density at time  $(t)$  and  $\exp(\mu)$  is the deterministic per-unit-abundance growth rate in population density. Natural population growth is not constant from year to year, so a multiplicative term,  $\mathbf{E}(t)$ , is included to represent the process variation or deviation from the long-term trend at time  $(t)$  due to environmental effects. The random variables  $\mathbf{E}(1), \mathbf{E}(2), \dots$  are assumed to be independent and identically distributed (iid) as lognormal(0,  $\sigma_p^2$ ), that is,  $\ln(\mathbf{E}(t))$  has a normal distribution

with a mean of 0 and variance  $\sigma_p^2$ . We write this as  $\mathbf{E}(t) \sim \text{lognormal}(0, \sigma_p^2)$  and  $\ln(\mathbf{E}(t)) \sim \text{normal}(0, \sigma_p^2)$ . Dennis et al. (1991) used this model as an approximation of a more complicated demographic population model under the assumption of constant trend ( $\mu$ ) and process variation ( $\sigma_p^2$ ) in the absence of density dependence. The parameters  $\mu$  and  $\sigma_p^2$  are the main quantities of interest when using DA methods for PVA. If this model is used for abundance estimates, a discrete random variable, as opposed to density, a continuous random variable, then caution may be required should the population abundance not be large enough to justify the continuous approximation with the exponential model.

If process error is the sole source of error, the estimates of the parameters  $\mu$  and  $\sigma_p^2$  given in Dennis et al. (1991) are unbiased. However, estimation is complicated when observed variability in the data reflects sampling error as well as process variation. Holmes (2001) proposed a model in which the population growth model (Eq. 1) is embedded in a model that includes sampling error. The observed data are modeled as

$$\mathbf{O}(t) = \mathbf{N}(t) \mathbf{Z}(t) \quad (2)$$

where  $\mathbf{O}(t)$  is the observed estimate of population density at time  $(t)$  and  $\mathbf{Z}(t)$  represents sampling error. The random variables  $\mathbf{Z}(1), \mathbf{Z}(2), \dots$  are assumed to be iid lognormal( $\mu_s, \sigma_s^2$ ) and independent of  $\mathbf{E}(t)$ . Bias in measurement is represented by  $\mu_s$ . By combining the population growth and sampling models, the observation at time  $(t + 1)$  is represented as

$$\mathbf{O}(t + 1) = \mathbf{N}(t) \exp(\mu) \mathbf{E}(t) \mathbf{Z}(t + 1). \quad (3)$$

Define  $\mathbf{M}(t) = \ln(\mathbf{N}(t))$ ,  $\varepsilon(t) = \ln(\mathbf{E}(t))$ , and  $\varphi(t) = \ln(\mathbf{Z}(t))$ . Note that  $\varepsilon \sim \text{normal}(0, \sigma_p^2)$  and  $\varphi \sim \text{normal}(\mu_s, \sigma_s^2)$ . Log transformation of the observed values yields

$$\mathbf{Y}(t) = \ln[\mathbf{O}(t)] = \mathbf{M}(t) + \varphi(t) \quad (4)$$

$$\mathbf{Y}(t + 1) = \mathbf{M}(t) + \mu + \varepsilon(t) + \varphi(t + 1). \quad (5)$$

Note that the entries of  $\mathbf{Y}$  are autocorrelated because the effect of each realization of the process variation,  $\varepsilon(t)$ , is passed on to later population densities through the exponential growth of the population, e.g.,  $\mathbf{Y}(t + 2) = \mathbf{M}(t) + 2\mu + \varepsilon(t) + \varepsilon(t + 1) + \varphi(t + 1)$  and  $\mathbf{Y}(t + 3) = \mathbf{M}(t) + 3\mu + \varepsilon(t) + \varepsilon(t + 1) + \varepsilon(t + 2) + \varphi(t + 2)$  and so on. The differences  $\mathbf{W}(t) = \mathbf{Y}(t + 1) - \mathbf{Y}(t)$  for  $t = 1, 2, \dots, T - 1$  give the growth rate plus error at each individual time step due to process variation at time  $t$  and sampling error at times  $t$  and  $t + 1$ . The mixed model representation of  $\mathbf{W}(t)$  and  $\mathbf{W}(t + 1)$  are

$$\mathbf{W}(t) = \mu + \varepsilon(t) + \varphi(t + 1) - \varphi(t) \quad (6)$$

$$\mathbf{W}(t + 1) = \mu + \varepsilon(t + 1) + \varphi(t + 2) - \varphi(t + 1). \quad (7)$$

The vector  $\mathbf{W}$  is a series of the observed log population growth rate at each of the  $(T - 1)$  one-step time in-

tervals in the time series of observations. The distribution of  $\mathbf{W}$  is multivariate-normal with common mean  $\mu$  and variance-covariance matrix  $\Sigma$ . The empirical growth observations in  $\mathbf{W}$  can be expressed as a conventional linear mixed model:

$$\mathbf{W} = \mathbf{X}\mu + \nu \tag{8}$$

where  $\mathbf{X}$  is a column of ones with length  $(T - 1)$  and  $\nu \sim (0, \Sigma)$ . Autocorrelation in  $\mathbf{Y}$  due to exponential population growth is removed in the creation of  $\mathbf{W}$  by the differencing step. The entries of  $\mathbf{W}$  are still correlated, however, because a single realization of the sampling error will be shared by two successive entries of  $\mathbf{W}$ , e.g.  $\varphi(t + 1)$  appears in both  $\mathbf{W}(t)$  and  $\mathbf{W}(t + 1)$ . This creates a one-step covariance in  $\mathbf{W}$ . Accordingly, the variance-covariance matrix for  $\mathbf{W}$  has a structure composed of the variance of individual entries of  $\mathbf{W}$  on the main diagonal and the one-step covariance on the diagonals above and below the main with zeroes elsewhere (this is known as a banded Toeplitz(2) structure (Graybill 1983)). Let  $\sigma_1$  and  $\sigma_2$  represent the variance and one-step covariance of  $\mathbf{W}$ , respectively. Then,

$$\Sigma_{(T-1)(T-1)} = \begin{bmatrix} \sigma_1 & \sigma_2 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \sigma_2 & \sigma_1 & \sigma_2 & 0 & & & & \cdot \\ 0 & \sigma_2 & \sigma_1 & \sigma_2 & 0 & & & \cdot \\ 0 & 0 & \sigma_2 & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & \cdot & \cdot & \cdot & \sigma_2 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & \sigma_2 & \sigma_1 \end{bmatrix} \tag{9}$$

The parameters in  $\Sigma$  are functions of the process variation and sampling error. The variance of an individual entry of  $\mathbf{W}$  is  $\sigma_1 = 2\sigma_s^2 + \sigma_p^2$ . The one-step covariance of the entries of  $\mathbf{W}$  is  $\sigma_2 = -\sigma_s^2$ .

In simulations, maximum likelihood estimators of  $\mu$ ,  $\sigma_s^2$ , and  $\sigma_p^2$  were strongly biased. Biased estimators are a common problem in variance component problems when using maximum likelihood. That is, the maximum likelihood estimate of a sample variance is biased as it ignores the difference of the sample mean from the true mean. This bias is overcome by using restricted maximum likelihood (REML) estimation (Searle et al. 1992) in which a transformation of the data is used to eliminate the fixed effect ( $\mu$ ) before estimating  $\sigma_1$  and  $\sigma_2$ . The data transformation used in this case is again a one-step difference of the  $\mathbf{W}$  series:

$$\mathbf{J}(t) = \mathbf{W}(t + 1) - \mathbf{W}(t) \tag{10}$$

where  $\mathbf{J}(t)$  is the transformed data vector of length  $(T - 2)$ . This transformation can be constructed as  $\mathbf{J} = \mathbf{U}'\mathbf{W}$ , where  $\mathbf{U}$  is a  $(T - 1) \times (T - 2)$  contrast matrix with negative 1's on the main diagonal and 1's on the diagonal below the main. The parameters  $\sigma_1$  and  $\sigma_2$  are then estimated with maximum likelihood from the

transformed data. Because the covariance matrix of  $\mathbf{J}$  is related to that of  $\mathbf{W}$  by the equation  $\Theta = \mathbf{U}'\Sigma\mathbf{U}$ , the covariance matrix  $\Theta$  is also a function of  $\sigma_1$  and  $\sigma_2$ . The values  $\hat{\sigma}_1$  and  $\hat{\sigma}_2$  that minimize the negative log-likelihood function

$$L(\sigma_1, \sigma_2 | \mathbf{J}) = \frac{1}{2}(T - 2)\ln(2\pi) + \frac{1}{2}\ln(|\Theta|) + \frac{1}{2}\mathbf{J}'\Theta^{-1}\mathbf{J} \tag{11}$$

are then the REML estimates of  $\sigma_1$  and  $\sigma_2$ . Good initial estimates of  $\sigma_1$  and  $\sigma_2$  for the minimization can be obtained by the method-of-moments, i.e., the initial estimates of  $\sigma_1$  and  $\sigma_2$  are the sample variance and the sample one-step covariance of  $\mathbf{W}$ , respectively. The estimate,  $\mathbf{V}$ , of the covariance matrix,  $\Sigma$ , is obtained by inserting  $\hat{\sigma}_1$  and  $\hat{\sigma}_2$  into Eq. 9.

To estimate  $\mu$ , a generalized least squares estimator is used. This is an extension of ordinary least squares, i.e., a straight average of the entries in  $\mathbf{W}$ , and takes advantage of nonzero covariance in the data. The generalized least squares estimator of the trend is

$$\hat{\mu} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{W} \tag{12}$$

Kackar and Harville (1984) showed that generalized least squares estimators with estimated weights are unbiased. Inference on  $\hat{\mu}$  requires an estimate of its variance. If the data have no sampling error, this variance is  $\mathbf{V} = \text{var}(\mathbf{W}) \mathbf{I}_{(T-1)}$ , where  $\mathbf{I}_{(T-1)}$  is an  $(T - 1)$  identity matrix and the estimator for trend is the same as that presented in Dennis et al. (1991). If  $\Sigma$  is known, then  $\hat{\mu} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{W}$  and the variance of  $\hat{\mu}$  is exactly  $\Phi = \text{var}(\hat{\mu}) = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ . Because  $\Sigma$  is unknown, the estimated covariance matrix must be used as in Eq. 12, resulting in a biased estimator of  $\text{var}(\hat{\mu})$ , namely,

$$\text{var}(\hat{\mu}) = \hat{\Phi} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \tag{13}$$

Kenward and Roger (1997) presented a technique to remove the bias in  $\hat{\Phi}$ , but application of this technique is not easily implemented in SAS and simulation results suggest that the bias in  $\hat{\Phi}$  is small, at least for the larger data sets examined. Therefore, the Kenward and Roger adjustment  $\hat{\Phi}$  to was not used for this analysis.

The behavior of estimators for the REML-based procedure was explored with analysis of simulated time-series data at nine combinations of process variation and sampling error. For each simulation, 2000 time series of length 50 were generated by first simulating the actual population growth with process variation and then "observing" the population densities with sampling error. Process variance and sampling error variance were each set at three levels: large, L = 0.09; medium, M = 0.0225; and small, S = 0.0025. For each of the nine variation combinations, simulations with and without a true nonzero underlying trend were analyzed. The simulations with zero trend, random walks,

consisted of population growth that had random fluctuations due to process variation only. Simulations with a nonzero trend had  $\mu = -0.05$ , a decline in total population density of  $\sim 5\%$  per year.

The performance of the REML-based procedure is compared with that of the Dennis et al. (1991) model in the simulations just described and in example analyses of data on the Whooping Crane (*Grus americana*), grizzly bear (*Ursus arctos horribilis*) in Yellowstone National Park, California Condor (*Gymnogyps californianus*), and Puerto Rican Parrot (*Amazona vittata*). The Whooping Crane data are winter population counts conducted 1938–1993 in the Aransas National Wildlife Refuge (USFWS 1994). The grizzly bear data are yearly estimated numbers of adult females in the Greater Yellowstone population for 1959–1997 (Morris and Doak 2002). The California Condor data are the October Surveys conducted 1965–1980 that have been described as problematic due to inconsistent sampling efforts (Synder and Johnson 1985). The Puerto Rican Parrot data are 1969–1989 censuses for an intensively managed population in the Luquillo Forest of Puerto Rico (Dennis et al. 1991).

As an example of the behavior of extinction metrics using estimators from the Dennis et al. (1991) and REML-based methods, predictions of the probability of reaching a threshold of 75% of the current population size (so-called quasi-extinction) are compared for simulated data with a slight positive trend. For these simulations, the true underlying trend is a 2% per year increase in population size, with the process variation and sampling variance taking on the same nine combinations as in the previously mentioned simulations.

## RESULTS

In simulations, the REML-based method provided an unbiased estimator of trend and only a slight bias in estimators for variance components. Performance of the REML-based method was similar for simulations with and without a nonzero trend. Increases in sampling error or process variation resulted in more variable estimates of sampling error. The level of sampling error affected the distribution of process variation estimates similarly for the three levels of process variation examined and for simulations with and without a nonzero trend. The distributions of process variation estimates for the REML-based method and the model of Dennis et al. (1991) are compared in Fig. 1 over the range of sampling error for simulations with medium-level process variation and  $\sim 5\%$  per year decline. In addition to decreasing the bias in process variation estimates, the REML-based method has less variability in process variation estimates. In data sets with large sampling error, however, estimates of process variation occasionally are extremely high with the REML-based method. These high estimates are typically higher than the total variation observed in the time series, but occur  $<1\%$  of the time in the simulated data. The sampling

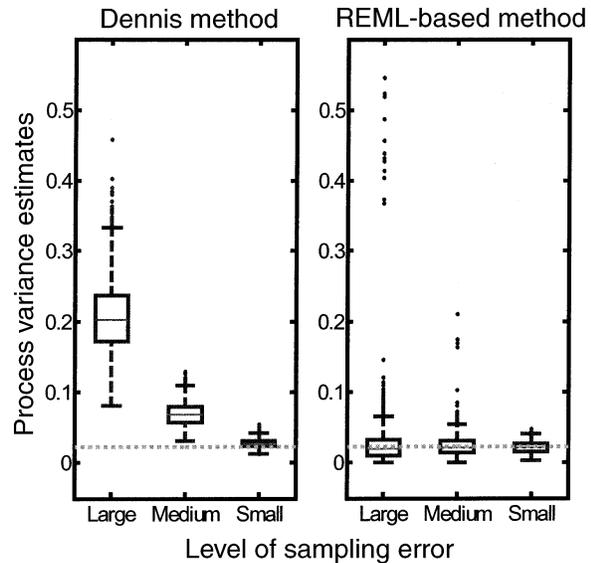


FIG. 1. Comparison of the Dennis et al. (1991) model and REML-based estimates of process variation over a range of sampling error. There were 2000 simulations at a medium level of process variation with 5% decline per year. The horizontal dotted line represents the true process variation value. Boxes represent the interquartile range, and error bars extend a maximum distance of 1.5 times the interquartile range. Dots represent extreme values.

error estimates are zero or very close to zero for the time series with high process variation estimates. In practice, a near-zero sampling error estimate accompanied by a process variation estimate that is larger than the variance in the time series (specifically the variance of the  $\mathbf{W}$  vector) should warn the user of unreliable variance component estimates. In such cases, we would recommend using the Dennis et al. (1991) model process variation estimator while acknowledging that it is likely to be biased high.

Bias and variance in predictions of quasi-extinction depend on the true probability of reaching quasi-extinction and the level of sampling error in the data. For high process variation, when quasi-extinction was highly likely, the higher variance of predictions for the REML-based method made it a poorer choice than the Dennis et al. method despite its bias. When the true probability of quasi-extinction was near zero (when process variation was small), the REML-based method was better due to the large bias in the method of Dennis et al. (1991) attributable to sampling error in the data. The distributions of predictions for both estimation methods over the nine variation combinations are shown in Fig. 2.

Parameter estimates and the probability of quasi-extinction for the example data analyses are given in Table 1. The trend estimates for the grizzly bear and Whooping Crane are similar for the Dennis et al. (1991) and REML-based approaches, but the REML-based model gives a smaller standard error for the trend es-

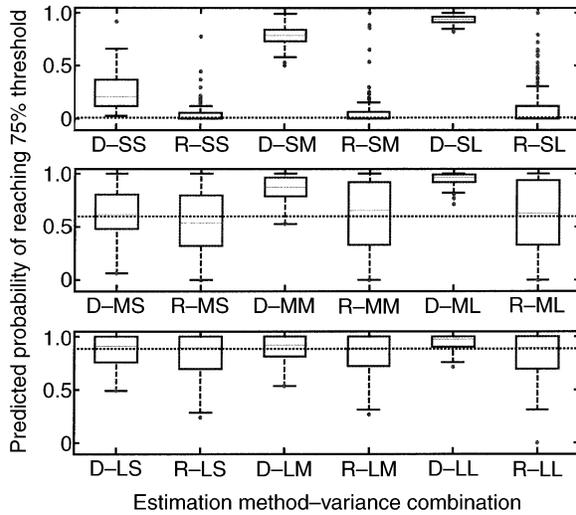


FIG. 2. Predicted probability of reaching a quasi-extinction level of 75% of the current population size using the Dennis et al. (1991) model (D) and the REML-based model (R). Population growth is ~2% increase per year. SS, SM, SL, . . . , LL represent, respectively, the level of process variance and sampling error variance: S, small; M, medium; L, large). The true probability of quasi-extinction is given by the horizontal dotted line. Boxes, error bars, and dots are as defined in Fig. 1.

imate in both cases. For both analysis methods, the estimates of trend, standard error of the trend estimate, and the estimated process variation are the same for the Puerto Rican Parrot data because the estimated sampling error is less than zero. In contrast, the California Condor data analysis resulted in an infinite likelihood in the SAS analysis (code available as a Supplement in *Ecological Archives*) due to the estimate of process variation being less than zero (see *Discussion*). Therefore the results given for the California Condor were calculated with the MATLAB code available in *Ecological Archives*.

DISCUSSION

The REML-based method may be performed in standard statistical packages such as SAS PROC MIXED. Because SAS places no boundaries on variance component estimates, there are occasional instances when the sampling error estimate is negative or the procedure fails to converge on a solution. The former is due to SAS allowing  $\hat{\sigma}_2$  to be larger than zero (recall that the sampling error is the negative of  $\hat{\sigma}_2$ ). In this case, we recommend setting the sampling error estimate to the boundary value of zero, thereby recovering the Dennis et al. (1991) method; see the following discussion of the Puerto Rican Parrot. The latter occurs when the likelihood function is infinite in the unbounded optimization. An ad hoc solution is to set the process variation estimate to the minimum boundary value of zero; see the following discussion of the California Condor.

The MATLAB code provided in *Ecological Archives* automatically makes these adjustments.

A *t*-test for significant trend may be performed in SAS if non-boundary parameter estimates exist, i.e., a minimum is found for the likelihood function and the sampling error estimate is greater than zero. Correct degrees of freedom for the test, however, are not clear for all circumstances. We suggest using the Kenward and Roger (1997) adjustment option in PROC MIXED for the degrees of freedom. However, this test is problematic if the estimated degrees of freedom are less than four (D. F. Staples and M. L. Taper, *personal observation*). In such cases, we recommend using the same degrees of freedom as the Dennis et al. (1991) method minus one for the additional parameter being estimated ( $T - 4$ ), while being aware this is likely to be a liberal test (D. F. Staples, *unpublished data*). Improved inferential techniques will be presented in a subsequent paper (M. L. Taper, D. F. Staples, R. Boik, M. Ferrari, and B. Dennis, *unpublished manuscript*).

Assumptions and diagnostics

The unbiased estimator of process variation rectifies a significant liability when using the Dennis et al. (1991) model for PVA, i.e., the vulnerability of conclusions to sampling error in the data. It does, however, retain other assumptions for PVA from the Dennis et al. model. Both methods assume that there are no large environmental events such as catastrophes or bonanzas that dramatically change the population size. Dennis et al. (1991) presented model diagnostic procedures that can be used to detect extreme events in the data. We would like to note that the detection of outliers does not mean that the point should be dropped for the REML-based method, because the “extreme” event

TABLE 1. Parameter estimates and risk metric comparisons for Dennis et al. (1991) and REML-based methods for four species.

Species and analysis method	Trend (1 SE)	Process variance	Sampling variance	<i>P</i> (lower threshold) <sup>†</sup>
Grizzly bear, GB (39 years of data)				
Dennis	0.0213 (0.0185)	0.0131		0.40
REML	0.0211 (0.0148)	0.0082	0.0023	0.24
Whooping Crane, WC (56 years of data)				
Dennis	0.0377 (0.0187)	0.0194		0.32
REML	0.0372 (0.0159)	0.0137	0.0028	0.21
California Condor, CC (16 years of data)				
Dennis	-0.0768 (0.0885)	0.1176		1.00
REML	-0.0948 (0.0131)	0 <sup>‡</sup>	0.0579	1.00
Puerto Rican Parrot, PP (21 years of data)				
Dennis	0.0273 (.0275)	0.0151		0.33
REML	0.0273 (.0275)	0.0151	0 <sup>‡</sup>	0.33

<sup>†</sup> Probability of population size reaching a lower threshold equal to 0.75 of the last population size.

<sup>‡</sup> Estimates at boundary.

may be due to sampling error. If it is due to sampling error, the REML-based method can use the information for better estimates of process variation, thereby avoiding the inflation of the process variation estimate by the extreme event.

The process variation (or environmental error) is also assumed to be temporally uncorrelated for both methods. A two-tailed Durbin-Watson test (Sen and Srivastava 1990) on the residuals of the Dennis et al. (1991) model can be used to evaluate whether errors are correlated. If there is a positive correlation, neither model may be appropriate. A negative correlation supports the assumption of the REML-based method that one-step correlation in the observed growth rates (the **W** series) is due to sampling error. Preliminary simulation results suggest that REML-based sampling error estimators are biased low and process variation estimators are biased high in the presence of positively correlated environmental error (D. F. Staples, *unpublished data*). The magnitude of the bias depends on the magnitude of the correlation. Development of improved tests for correlated environmental error and methods to incorporate it into the REML-based method are the focus of ongoing research.

Another assumption is a constant trend and process variation throughout the time series. Violations of this may occur due to demographic stochasticity resulting from small population size or other factors that change the growth rate of the population, e.g., density dependence or the Allee effect. Demographic stochasticity can change the variance in growth rate, especially when population estimates vary from low to high densities. Density dependence may be tested using methods presented in Dennis and Taper (1994) or, alternatively, model identification techniques may be used as in Taper and Gogan (2002).

#### *Example data analyses*

In addition to the smaller standard error in the trend estimates, another benefit of the REML-based method for the grizzly bear and Whooping Crane data sets is the removal of sampling error from the observed variation in the data, resulting in smaller estimates of process variation. This decreases bias in the process variation and results in a less biased analysis of the risk to the populations. For the Puerto Rican Parrot data, the estimate of sampling error is at the boundary value of zero, indicating that sampling error contributes little of the variability observed in the data. This is reasonable, considering that the data come from a small, intensively managed population and represent full population censuses. With the sampling error set to zero, the REML-based method is identical to the Dennis et al. (1991) method.

The California Condor data represent the opposite end of the parameter spectrum compared to the Puerto Rican Parrot. In contrast to a zero estimate of sampling error, the REML-based method gives an estimate of

process variation that is at the boundary value of zero for the condor data, indicating that the variation in the data is mainly due to sampling error. This again is a reasonable conclusion, considering that the data come from the "October Surveys" that have been described as problematic due to sampling irregularities (Wilbur 1980), and the population was undergoing a severe crash at the time and probably had little variability in the true negative growth rate. If the process variation is set to zero and a significance test is performed, the REML-based method does give a significantly negative growth rate ( $P < 0.001$ ) in contrast to a nonsignificant trend in the Dennis et al. (1991) method ( $P = 0.40$ ). This case is, however, an instance in which the degrees of freedom estimate from the Kenward and Roger (1997) method is extremely small (0.07); therefore,  $T - 4$  degrees of freedom was used for the test. Although inference in this case is questionable due to uncertainty about degrees of freedom, it is reassuring to get a significant negative trend in a population known to have experienced such a severe decline.

#### *Conclusion*

The restricted maximum likelihood (REML)-based method is a useful addition to PVA techniques using diffusion approximation (DA) methods. It allows partitioning out variability in the data due to sampling error for more precise estimates of trend and less bias in the process variation estimator. When sampling error is small in the data, as for the Puerto Rican Parrot, the REML-based method recovers the estimators in the Dennis et al. (1991) model. The REML-based method does, however, have more strict data requirements than the Dennis et al. model at present. Data must come from uniform sampling intervals with no missing points and the REML-based method is likely to require more data points because of the additional parameter to be estimated. In cases with too few data points, an estimate of the trend, which is relatively straightforward and unbiased, may be the best metric available for assessing risk to the population. We caution that these models or any other PVA models are only tools for a manager and results must be conditional on the validity of assumptions in the model and the nature of the population itself.

#### ACKNOWLEDGMENTS

We would like to thank Robert J. Boik, John J. Borkowski, and James F. Robison-Cox for their help in developing the method presented. We would also like to thank reviewers Ben Bolker, Aaron M. Ellison, Lev Ginzburg, and Jay Ver Hoef for their constructive comments. Funding for this work was provided by Montana Fish, Wildlife, and Parks grant number Z0862, EPA STAR grant number G9A10014, and Northwest Power and Conservation Council Grant number PTMM11.

#### LITERATURE CITED

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

- Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. *Ecological Monographs* **64**:205–224.
- Gerber, L. R., D. P. DeMaster, and P. M. Kareiva. 1999. Gray whales and the value of monitoring data in implementing the U.S. Endangered Species Act. *Conservation Biology* **13**:1215–1219.
- Graybill, F. A. 1983. *Matrices with applications in statistics*. Second edition. Wadsworth International Group, Belmont, California, USA.
- Holmes, E. E. 2001. Estimating risks in declining populations with poor data. *Proceedings of the National Academy of Sciences (USA)* **98**:5072–5077.
- Holmes, E. E., and W. E. Fagan. 2002. Validating population viability analysis for corrupted data sets. *Ecology* **83**:2379–2386.
- Kackar, R. N., and D. A. Harville. 1984. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association* **79**:853–862.
- Kenward, M. G., and J. H. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**:983–997.
- Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. *SAS system for mixed models*. SAS Institute, Cary, North Carolina, USA.
- Ludwig, D. 1999. Is it meaningful to estimate a probability of extinction? *Ecology* **80**:298–310.
- Meir, E., and W. F. Fagan. 2000. Will observation error and biases ruin the use of simple extinction models? *Conservation Biology* **14**:148–154.
- Morris, W. F., and D. F. Doak. 2002. *Quantitative conservation biology: theory and practice of population viability analysis*. Sinauer Associates, Sunderland, Massachusetts, USA.
- Nicholls, A. O., P. C. Viljoen, M. H. Knight, and A. S. vanJaarsveld. 1996. Evaluating population persistence of censused and unmanaged herbivore populations from the Kruger National Park, South Africa. *Biological Conservation* **76**:57–67.
- Searle, S. R., G. Casella, and C. E. McCulloch. 1992. *Variance components*. John Wiley, New York, New York, USA.
- Sen, A., and M. Srivastava. 1990. *Regression analysis: theory, methods, and applications*. Springer-Verlag, New York, New York, USA.
- Snyder, N. F. R., and E. V. Johnson. 1985. Photographic censusing of the 1982–1983 California Condor population. *Condor* **87**:1–13.
- Taper, M. L., and P. J. P. Gogan. 2002. The Northern Yellowstone elk: density dependence and climatic conditions. *Journal of Wildlife Management* **66**:106–122.
- USFWS (U.S. Fish and Wildlife Service). 1994. *Whooping Crane recovery plan*. U.S. Fish and Wildlife Service, Albuquerque, New Mexico, USA.
- Wilbur, S. R. 1980. Estimating the size and trend of the California condor population. *California Fish and Game* **66**:40–48.

#### SUPPLEMENT

A supplement providing data files and code is available in ESA's Electronic Data Archive: *Ecological Archives* E085-025-S1.