UNDERSTANDING AND PREDICTING THE EFFECTS OF SPARSE DATA ON DEMOGRAPHIC ANALYSES

Daniel F. Doak,[1,4] Kevin Gross,[2] and William F. Morris[3]

[1]*Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95064 USA*
[2]*Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695 USA*
[3]*Biology Department, Duke University, Box 90338, Durham, North Carolina 27708-0338 USA*

Abstract. Demographic models are an increasingly important tool in population biology. However, these models, especially stochastic matrix models, are based upon a multitude of parameters that must usually be estimated with only a few years of data and limited sample sizes within each year, calling into question the accuracy of the results of these models. We first discuss how these data limitations create sampling uncertainty and bias in the estimated parameters for a stochastic demography model. Next, we ask whether limited data can favor the construction of deterministic models that ignore variation and correlation of rates. With less than five years of data, the mean squared error of deterministic models will sometimes be smaller than that of stochastic models, favoring the use of simple models, even when their predictions are known to be biased. Finally, we introduce a procedure to estimate the sampling variation around population growth rate estimates made from demographic models that are based on specified sampling durations and intensities.

Key words: covariance; demography; matrix; rarity; stochasticity; variability; viability.

INTRODUCTION

Demographic models are among the most widely used analytical tools in life history studies, population ecology, and, especially, conservation biology, where they form the backbone of population viability analysis (PVA). However, demographic analyses—including both life tables and matrix models—are also among the most data hungry of methods, requiring estimates of birth, death, and other vital rates for each step of an individual's life. In spite of these needs, there has been relatively little discussion in the demographic literature of what data limitations mean for the construction and interpretation of these models (but see Ludwig 1999, Fieberg and Ellner 2001, Ellner et al. 2002), and in particular, little work that would help a field biologist understand how much demographic data must be collected to have confidence in modeling results, and how the complexity of different demographic models and the scarcity of information with which to parameterize them will interact to influence this confidence.

Two distinct kinds of data scarcity can effect the accuracy of vital rate estimates and hence the accuracy of demographic predictions: the sampling of small numbers of individuals within each year, so that vital rate estimates made for each transition period are uncertain; and the collection of data over only a small number of years, so estimates of the mean, variance, and covariance of vital rates over time are unreliable.

We will refer to these as sampling intensity and sampling duration. One or both kinds of data limitation afflict virtually anyone who tries to develop a demographic model. Here, we discuss some of the ramifications of this data scarcity. In particular, we focus on the problems of parameterizing stochastic demographic models, which are increasingly advocated, but which require even more information than do simple deterministic models.

We have three goals in this paper. First, to explain how limited data not only increase the uncertainty in all parameter estimates needed for a demographic analysis, but also create biased estimates of the variance and correlation parameters needed for stochastic models. While this bias is to some extent correctable, few ecologists seem to realize that this is a key complication in the construction of stochastic models. Second, to analyze the trade-off of building more realistic, but complex, stochastic models vs. simpler, but less realistic models that do not include variance and correlation. While variance and correlation in vital rates can strongly affect demography, there is also a great deal of uncertainty around estimates of variance and correlation with limited data. Consequently, we ask: when is the cost of including estimates of variance and correlation in a demographic model worth the benefit? The basic idea behind this question is now well-publicized in ecology: there is a trade-off between model realism (i.e., complexity) and the "noise" created by building predictions on the back of poorly estimated parameters (Ludwig et al. 1988, Hilborn and Mangel 1997, Burnham and Anderson 1998). However, this trade-off has not been directly assessed for the incorporation of sto-

chasticity into demography models. Third, we address what may be the most widely posed question among empirical ecologists who are initiating demographic studies: how many data are enough to get reasonably accurate predictions? While this is obviously impossible to answer with certainty, we develop a way to get a better answer than a simple guess or a shrug of the shoulders—the usual responses to this question.

DEMOGRAPHIC MODELS IN A NUTSHELL

To make the following discussion of parameter estimation and model construction comprehendible, we briefly review the basic form and analysis of demographic models. Demographic models generally use age, size, or stage (e.g., juveniles and adults) to classify individuals. For the sake of clarity, we will couch our discussion in terms of the rates needed to construct a size-based model, but the issues and methods are the same when using other classifying variables. We consider a simple model with four size classes and a set structure, in which an individual can either survive and remain in the same size class from one year to the next, or survive and grow enough to enter the next largest size class. We can characterize such a life history with three types of vital rates: fecundities (f_a), growth probabilities (g_a), and survival rates (s_a), each of which are indexed by the size class (a) to which they apply. We will assume that, once in the largest size class, there is no senescence, and also that the fecundities, and hence the model as a whole, only apply to females. We also assume that reproduction is synchronous and that the youngest class in the model consists of newborns, with the census of the population occurring just after reproduction. These assumptions yield the following transition matrix $\mathbf{A}$ for four size classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & e_{13} & e_{14} \\ e_{21} & e_{22} & 0 & 0 \\ 0 & e_{32} & e_{33} & 0 \\ 0 & 0 & e_{43} & e_{44} \end{pmatrix}$$

$$= \begin{bmatrix} 0 & 0 & f_3 s_3 & f_4 s_4 \\ s_1 & s_2(1 - g_2) & 0 & 0 \\ 0 & s_2 g_2 & s_3(1 - g_3) & 0 \\ 0 & 0 & s_3 g_3 & s_4 \end{bmatrix}. \quad (1)$$

Here, the matrix element e_{ab} is the number of individuals in class a that we expect to arise one year in the future from each individual of size b now. For the top row of the matrix, these elements are functions of survival and fecundity, and for the remaining rows, the elements are functions of survival and growth probabilities. The second matrix in the equation shows how each matrix element is defined in terms of the vital rates.

With an estimated projection matrix in hand, there are several key predictions that are usually made. First is the annual population growth rate, λ, which summarizes the long-term rate of change for a population governed by $\mathbf{A}$. If population size is shown by a column vector, $\mathbf{N}_t$, that contains the numbers of individuals in each size class (n_1, n_2, n_3, n_4 for our example), then $\mathbf{N}_{t+1} = \mathbf{A}\mathbf{N}_t = \mathbf{A}^{t+1}\mathbf{N}_0$, and as t becomes large, the annual rate of change of the population will converge to that predicted by λ. At the same time, the population will converge to the stable stage distribution (SSD), in which the fraction of the whole population that is in each stage remains constant. Besides λ and the SSD, the most commonly used outputs of a simple projection matrix are the sets of sensitivity and elasticity values. Sensitivities estimate the rate with which λ will change with changes in each matrix element, and elasticities are rescalings of these values that estimate the proportional change in λ resulting from proportional changes in each matrix element (Caswell 2001).

In reality, a single matrix constructed from a set of mean vital rates will often do a poor job of representing population dynamics, because one or more of these rates will vary substantially between years. One way to include this environmental stochasticity is to estimate a separate matrix for each year of data collected, and then randomly choose among them in simulating the future (Fieberg and Ellner 2001). While this approach is straightforward, it does not allow any ready analysis of, or correction for, the problems created by limited data. Alternatively, we can make explicit estimates of the mean and variance of each vital rate and the covariance, or correlation, between all pairs of rates across years. With these estimates, both analytical approximations or simulations of the model can provide results analogous to those of a deterministic model, as well as estimates of extinction risks through time (Caswell 2001, Morris and Doak 2002). The stochastic analogue to λ, λ_s, represents the average expected rate of population change, accounting for the effects of environmental stochasticity. Because it is more or less normally distributed and hence easy to deal with statistically, the natural logarithm of the of λ_s, log λ_s, forms the basis of most stochastic demography, including our analyses here.

UNCERTAINTY AND BIAS CREATED BY LIMITED DATA

There are three interrelated issues to consider in using limited data to estimate the parameters of a stochastic matrix model. The first of these is perhaps the most important, and is also the most widely understood. The second two are much less appreciated, but are still critical to understand when building estimates of demographic rates.

First is the simple problem that estimates of vital rate means are highly uncertain when data are limited. While there are various tactics to reduce this problem in the estimation of vital rate parameters (e.g., Easterling et al. 2002, Morris and Doak 2002, Dixon et al. 2005), it will never go away. Quantifying this uncer-

tainty in estimates of means is the grist of most basic statistics classes.

The second issue to appreciate in making parameter estimates for demographic models is that sampling error not only creates uncertainty but also bias in estimates of variances and correlations. The emphasis in standard statistics in estimating and testing effects on *means* is undoubtedly why this basic aspect of sampling theory is so underappreciated by ecologists. Estimated temporal variance in a demographic rate is generally biased high; observed variation across years is due in part to real changes in rates across time (environmental stochasticity) but also in part to variation caused by estimation errors. Conversely, if vital rates that are probabilities, such as survival and growth, have very high variation from year to year, sampling error can bias estimates of this variation low (Kendall 1998). In either case, to arrive at an unbiased estimate of the environmental variation in a demographic rate, one must estimate and remove the portion of observed variation due to sampling errors. This biasing of temporal variance estimates has recently received considerable attention in the technical demographic literature, with at least three suggestions of how to make such a correction (Engen et al. 1998 [whose method is presented for estimation of demographic stochasticity], Kendall 1998, White 2000; Morris and Doak 2002 include programs to implement these corrections). However, these methods are all relatively new, and the vast majority of stochastic demography studies published to date did not use them in estimating the environmental stochasticity in vital rates.

Covariances between two rates are not biased by sampling errors, at least if we can assume that sampling errors for the two rates do not themselves covary. However, correlation coefficients, equal to the covariance of two rates divided by their standard deviations, are biased, with increasing amounts of sampling error leading to estimated correlation coefficients that approach zero even when the true correlation is substantial. This would seem to suggest that we should use the unbiased covariance estimates together with the corrected variance estimates to arrive at unbiased estimates of correlations. However, for any limited sample of data, this procedure will yield many correlation estimates that are greater than one or less than -1, simply due to random (but unbiased) variation in the estimated covariance values. Until some solution is found to this problem, we are left estimating correlations directly, realizing that they will on average be biased toward zero.

The third parameter estimation problem involves the entire set of correlation coefficients between all the vital rates estimated for a population; because of sampling errors or missing data, the full set of correlations you estimate may not be able to exist. For example, three vital rates can't all be highly negatively correlated with each of the others. While this is an extreme ex-

ample, more subtle forms of this problem commonly arise. This may seem like an esoteric issue, but it becomes important if your goal in estimating demographic rates is to make stochastic predictions; if the set of estimated correlations is impossible, then you can't use them in simulation models to predict population growth or extinction. We describe this issue and how to deal with it in more detail in Appendix A.

THE COSTS AND BENEFITS OF INCLUDING STOCHASTICITY IN DEMOGRAPHIC MODELS

Even when properly accounting for bias in estimates of vital rate variances, the uncertainty in estimated means, variances and correlations arising from low sampling duration and intensity mean that a PVA or other demographic model based on limited data can give highly misleading predictions. The extent to which PVA results can be trusted without truly extensive sampling within and across years has been vigorously debated in the literature, but no clear conclusion exists to date (Ludwig 1999, Brook et al. 2000, Ellner et al. 2002). While it is clear that even moderate amounts of environmental stochasticity in combination with limited data can yield inaccurate results, we would still like to know how to build the best model with the data at hand, and in particular, whether it is worth the hassle to estimate and simulate the variance and correlation of vital rates in order to arrive at stochastic predictions.

Three classes of demographic models are commonly parameterized and used: deterministic models, based only on estimated mean rates; stochastic variance models, which ignore correlations between rates, but do use estimated variances; and full stochastic models, which use estimated means, variances, and correlations for all or most vital rates. The advantage of using a deterministic model is that its predictions rely on a smaller number of parameters, and with sparse data, this means a smaller number of badly estimated parameters. However, this benefit may be more than balanced by the greater realism of stochastic models. Although these use more parameter estimates, they do so to include real features of the population—temporal variability—and thus may give more accurate predictions. If environmental variability in vital rates is moderate or strong, deterministic matrix models will generally give predictions of population growth that are biased high, predicting more rapid growth rates than will actually occur (Tuljapurkar 1990, Caswell 2001). It is this bias that has led to the frequent calls to include variation and correlation in demographic PVA models (e.g., Burgman et al. 1993, Beissinger and Westphal 1998, Morris and Doak 2002). Nonetheless, we usually have only one short stretch of years with which to parameterize a model, and the increased bias of a simplified model may be outweighed by less random variation in its predictions, relative to an unbiased, but potentially more inaccurate model built with many poor estimates of variation and correlation. In statistical terms, there

TABLE 1.   Sets of parameter values used in the simulation models.

Parameter	Vital rate							
	s_1	s_2	s_3	s_4	g_2	g_3	f_3	f_4
Mean	0.1	0.3	0.8	0.9	0.5	0.75	5	10
Variances								
High	0.0675	0.1575	0.1200	0.0675	0.1875	0.1400	20	20
Medium	0.0500	0.0500	0.0200	0.0200	0.0100	0.0100	7	10
Low	0.0125	0.0125	0.0050	0.0050	0.0025	0.0025	1.75	2.50

Notes: A total of nine sets of vital rate parameters were used, corresponding to combinations of the three sets of variance values with the three sets of correlation coefficients. In the high-correlation models, all correlation coefficients were +0.80; in the low-correlation models, all correlation coefficients were +0.20. In the mixed correlation models, the correlation between each fecundity (*f*) and any survival (*s*) or growth (*g*) rates was −0.80 (corr(f_i,s_j) = corr(f_i,g_j) = −0.80), while all other correlations were +0.80. Mean values of vital rates were the same in all models.

is a trade-off between the bias and sampling variation of the estimated population growth rate.

To explore this issue, we developed a set of simulation models to test the accuracy of different demographic models parameterized with differing durations and intensities of data sampling. In all, we simulated nine demographic scenarios, each with a different set of vital rate means, variances, and correlations (Table 1), but all based on the life history shown in Eq. 1. These rates generate expected stochastic growth rates ranging from $\lambda_s = 0.81$ to 1.01. Each simulation consisted of three basic steps (see Appendix B for details):

1) Generation of a set of ''real,'' but unobserved, vital rate values across 20 years. For each demographic scenario, we generated 500 sets of these real annual rates.

2) Based on each set of real rates, simulation of an observed data set, taken with limited sampling duration and intensity. We varied sampling intensity from 10 to 40 individuals in each size class, and sampling duration from three to 10 years. For any single simulation, we assumed that the same number of individuals were followed in each size class and each year. Using these simulated data, we estimated means, variances, and covariances of vital rates, employing Kendall's (1998) method of variance correction for survival and growth rates and White's (2000) method for fecundities.

3) For each simulated data set, we predicted log λ_s using the deterministic growth rate predicted from the mean matrix and using the stochastic variance and full stochastic versions of Tuljapurkar's (1990) approximation, each parameterized with the estimated vital rates. For each sampling pattern and model type, we compared the estimated log λ_s values with the growth rate predicted from the real vital rate parameters by calculating mean squared error (MSE) and bias.

As expected, the average predictions of the deterministic models were optimistically biased, with positive differences between the mean deterministic predictions and true log λ_s for almost all model forms and sampling regimes (Fig. 1a–c). Differences between the full and variance model predictions and expected log

λ_s values lie closer to zero, although they too are for the most part positively biased for models with moderate to high vital rate variances. This trend is largely due to optimistic estimates of growth for models using the high or mixed correlation structures, which both feature strong and mostly positive correlations between vital rates. In these situations, the tendency of estimated correlations to be biased low is expected to result in just such a bias in the predictions of full models. Predictions of variance models, which implicitly assume that all correlations are zero, are only somewhat more biased than those of full models, but also show a tendency towards more variable results.

The greater average bias in deterministic model predictions would suggest that the stochastic models should provide better estimates of future numbers. However, while the average bias of deterministic models is larger, the variance in the deterministic predictions is substantially lower than that of either stochastic model for some sets of vital rates and sampling regimes (Fig. 1d). In particular, deterministic models showed lower sampling variance for all simulations using high vital rate variances and for some sampling regimes with medium variances. This lower variance in predictions means that even with their biases, deterministic models can sometimes have lower MSE than do full or variance-only models.

To show how model form and sampling regimes influence the predictive power of these models, we separately plot MSE against sampling duration and intensity for results from different vital rate scenarios. Here and below, we only present results for the full and deterministic models, since the full and variance models show nearly identical results for virtually all parameter combinations. Unsurprisingly, with higher sampling intensity and higher sampling duration, the MSE of all model predictions declines (Figs. 2 and 3; results for low-variance models are qualitatively identical). For the most part, there is little interaction between intensity and duration of sampling. Duration has by far the larger effect on MSE, and if duration is low, intensive within-year sampling usually has little effect
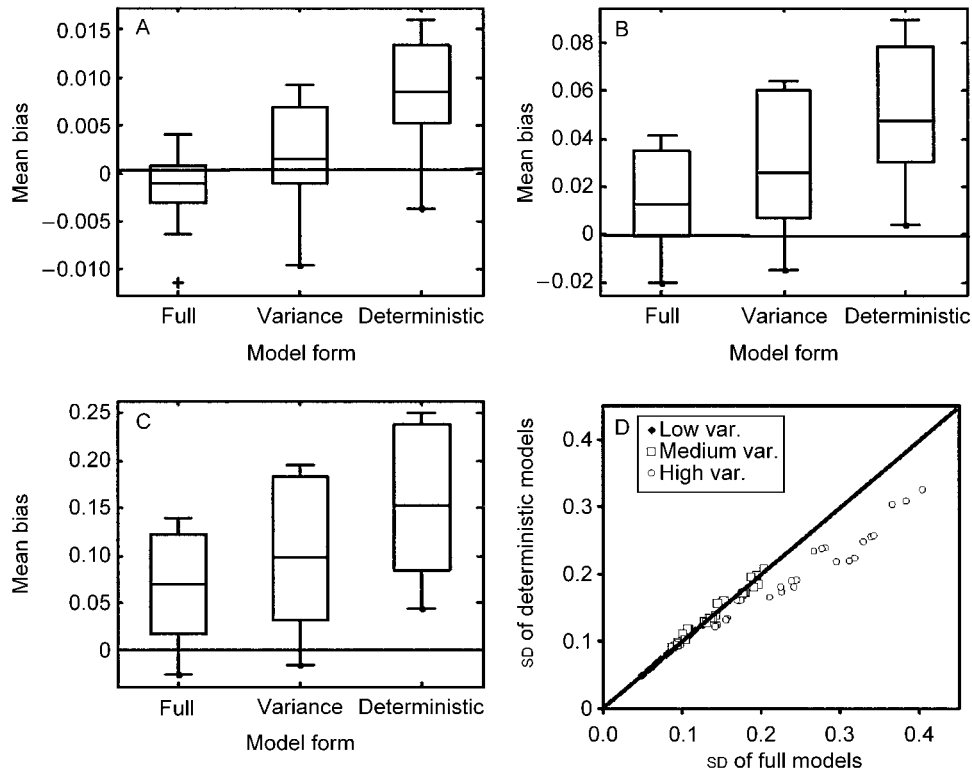
SPECIAL FEATURE

FIG. 1. Bias and variance in predictions of deterministic, variance-only, and full stochastic models. (A–C) Box plots showing the distribution of mean differences between predicted and expected log λ_s values for 27 means, one for each combination of correlation parameters, sampling intensity, and sampling duration: (A) low-variance models, (B) medium-variance models, and (C) high-variance models. (D) The standard deviation in predicted log λ_s (stochastic lambda) values for each combination of vital rate parameters and sampling regimes for full and deterministic models. Symbols indicate results from models with high, medium, and low variance in vital rates (Table 1). Points falling below the diagonal line indicate situations in which the deterministic predictions were less variable than those of the full models for the same vital rate parameters and sampling regimes.

on MSE. The exception to this comes at the lowest sampling durations, when higher intensity can substantially reduce MSE for some vital rate combinations (e.g., Fig. 2b).

For parameter sets with low vital rate variances, the difference in quality of different model predictions are quite small, amounting to <10% of MSE of the full models in all cases, and usually <5% (results not shown). For medium and high variance models, differences between the predictions of deterministic and full models are more substantial and more complex. With high correlations in vital rates (Figs. 2a and 3a), Full models have substantially higher accuracy at all but the lowest sampling durations. In contrast, with low correlation and Medium variance in vital rates, the two model forms show very similar MSEs, and with low correlation and high variances, the deterministic models have much lower MSEs for all but the longest sampling durations (Figs. 2b and 3b). Finally, with mixed correlations in vital rates, full models are more accurate with high duration and less so for short sampling durations (Figs. 2c and 3c). Within these details, the general pattern is that full model MSE values are quite

sensitive to data scarcity, rising sharply with decreasing sampling duration and also increasing with lower within-year sample sizes. In contrast, deterministic model MSEs are generally less responsive, such that there is some sampling duration at which the deterministic MSE becomes less than the full MSE, with the exact value depending critically on the underlying vital rate parameters.

Altogether, these results provide a rather mixed picture of the problems of predictive power of stochastic models. For some life histories and sampling regimes deterministic models and full stochastic models essentially tie one another in accuracy, as judged by their MSE values. In these cases, the lower bias of stochastic models will make them preferable to the overly optimistic deterministic predictions. However, when variances in vital rates are high, full stochastic models are sometimes more accurate than simple deterministic predictions, but are also sometimes spectacularly less so. With real data, we can't know the real underlying rates, so we could also ask, averaging across all the vital rate combinations we ran, which model form has the lower MSE values? The answer is deterministic
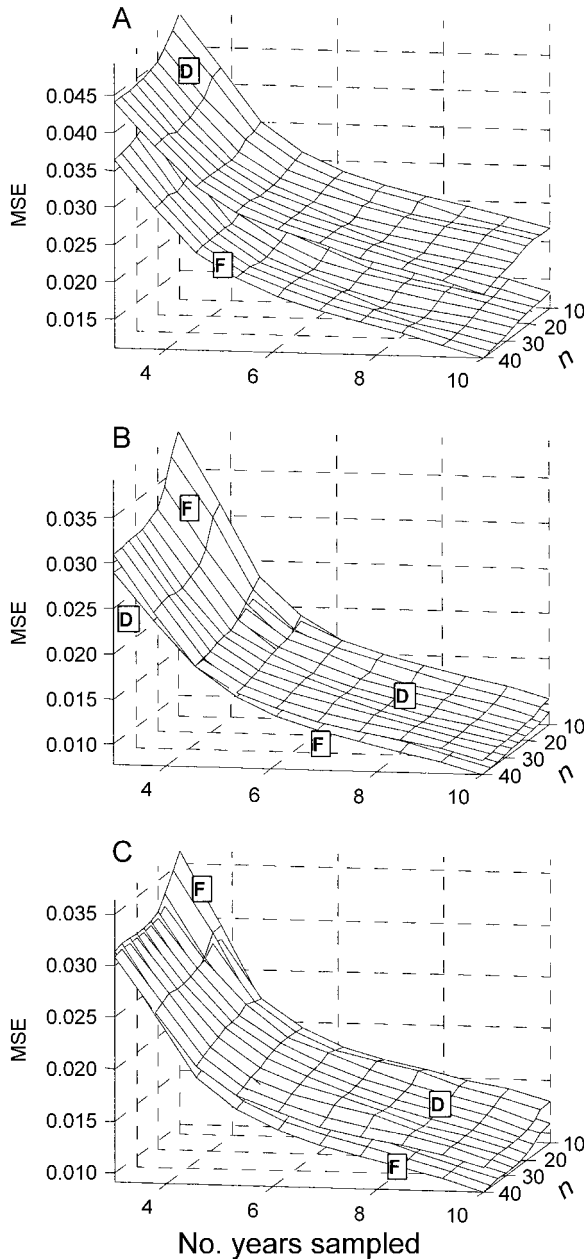
FIG. 2. Mean-squared error (MSE) of predicted population sizes for medium-variance models for sample sizes $n = 10$–40 and 4–10 sampling years. Two surfaces are shown (drawn with cubic interpolation), one for deterministic and one for full stochastic models. The labels "D" and "F" indicate which model corresponds to each surface (deterministic and full, respectively). The model with the lower MSE provides more accurate predictions of population size. Results are from life histories with (A) high positive correlations, (B) low positive correlations, and (C) mixed correlations.

models for sampling durations of 3–5 yr and full models for durations of 5–10 yr (Fig. 3d), implying that if we have few data, simpler models may provide better predictions than will stochastic model forms. To put this in perspective, only nine of the 29 studies listed

in Fieberg and Ellner's (2001) review of stochastic demography analyses were based on five or more years of vital rate estimates. Depending on your love of stochastic analysis, this may seem a disheartening or a comforting conclusion. However, we reiterate that we have only explored results for one life history pattern, and it is conceivable that other, very different life histories will show quite different patterns.

HOW MANY DATA ARE ENOUGH?

The final question we consider is how high sampling intensity and duration must be to achieve reasonable accuracy in demographic predictions. In particular, how do the duration and intensity of sampling translate into uncertainty of population growth predictions? Obviously, you can't really answer this question without perfect knowledge of all the rates and processes you are trying to estimate. However, by making a few guesses and assumptions, you can use the derivation we give below to arrive at a good estimate of the accuracy you can hope to achieve (see Caswell [2001] for a review of other approaches, mostly for assessing uncertainty in deterministic λ after data collection is completed).

To begin with, you must guess the form of the matrix model and the values of the vital rate means, variances and correlations governing your species. While you can't know the real values, usually natural history knowledge or data from related species can be used to make rough estimates of the form of the life history of your species and the values of its vital rates. As Dixon et al. (2005) emphasize, data on related species are a great untapped source of information for such guesses. We will call the guessed mean and environmental variance for a vital rate i, ν_i and $\sigma_{i,E}^2$, and its correlation with another rate j, $\rho_{i,j}$.

Second, you must have some guess as to how the within-year sampling intensity for a vital rate, the M_i number of individuals measured, influences the sampling variance of the estimated annual values of the rate. For survival and growth probabilities of individuals with identical mean values, this within-year sampling variance is $\sigma_{i,W}^2 = \nu_{i,t}(1 - \nu_{i,t})/M_i$, where $\nu_{i,t}$ is the mean probability in year t. For fecundities, in which each individual must give birth to an integer number of offspring, we will assume that the variance between individuals is Poisson, so that the variance of a mean annual estimate is $\sigma_{i,W}^2 = \nu_{i,t}/M_i$. There are many other forms of individual variation in fecundities that will give more or less variation in annual mean values and which can be used in place of the Poisson-distributed variation we assume.

With guesses as to the true demography of your species and the form of sampling variance, there are two ways to proceed. First, simulations like those described in the last section can directly yield estimates of the sampling variance in estimated growth rates. However, constructing such a simulation is not a task that most
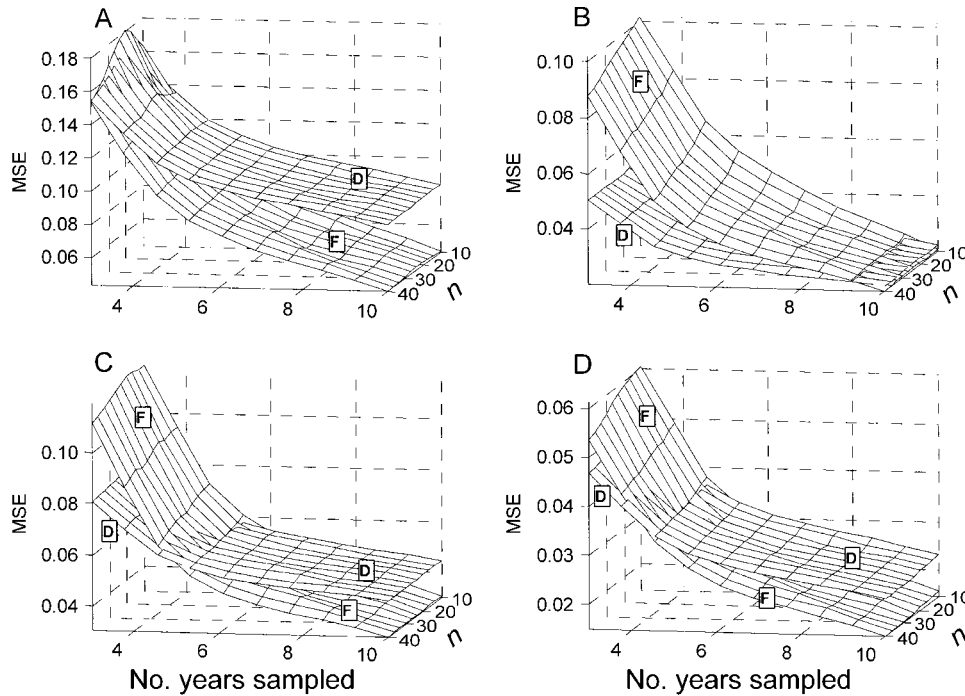
FIG. 3.   Mean-squared error (MSE) of predicted population sizes for high-variance models and all models combined for sample sizes n = 10–40 and 4–10 sampling years. Two surfaces are shown (drawn with cubic interpolation), one for deterministic models (labeled "D") and one for full stochastic models (labeled "F"). The model with the lower MSE provides more accurate predictions of population size. Results are from life histories with (A) high positive correlations, (B) low positive correlations, and (C) mixed correlations. Results for MSE over all vital rate parameter sets are shown in (D).

empirical ecologists are likely to tackle at the onset of their own field study. Therefore, we derived an analytical approximation for the sampling variance of $\log\lambda_s$, based on Tuljapurkars's (1990) "small noise" approximation for $\log\lambda_s$. Like the derivation of this approximation, our further use of it relies on assumptions of small, normally-distributed variation in vital rates (see Appendix C for details).

The basic result is an equation for the sampling variance in estimated population growth rate as a function of the vital rate parameters, the number of years sampled, and the number of individuals sampled in each year to estimate each rate:

$$\mathrm{Var}(\log \hat{\lambda}_s) \approx \sum_i \mathrm{Var}(\hat{v}_i)\left(\frac{\partial \log \lambda_s}{\partial v_i}\right)^2$$

$$+ \frac{1}{\lambda_1^4} \sum_i \mathrm{Var}(\hat{\sigma}_{i,E})\left(\sum_j S_i S_j \rho_{i,j}\sigma_{j,E}\right)^2$$

$$+ \frac{1}{\lambda_1^4} \sum_{i,j} \mathrm{Var}(\hat{\rho}_{i,j})(S_i S_j \sigma_{i,E}\sigma_{j,E})^2. \quad (2)$$

Here, S_i is the sensitivity of the growth rate predicted by the mean matrix, λ_1, to changes in v_i, and $\partial \log \lambda_s/\partial v_i$ is the sensitivity of the log-stochastic growth rate to the mean vital rate v_i (see Appendix C for the formula for this sensitivity). $\mathrm{Var}(\hat{v}_i)$, $\mathrm{Var}(\hat{\sigma}_{i,E})$, and $\mathrm{Var}(\hat{\rho}_{i,j})$ are approximations for the sampling variation in estimated

means, variances and correlations of the vital rates, respectively, and are where limited sampling has its influence. Assuming normally distributed variation and uncorrelated errors among rates, these quantities are:

$$\mathrm{Var}(\hat{v}_i) \approx \frac{1}{N}\left(\sigma_{i,E}^2 + \frac{\sigma_{i,W}^2}{M_i}\right)$$

$$\mathrm{Var}(\hat{\sigma}_{i,E}) \approx \frac{N}{2(N-1)^2}\left(\sigma_{i,E}^2 + \frac{2\sigma_{i,W}^2}{M_i}\right) \quad (3)$$

$$\mathrm{Var}(\hat{\rho}_{i,j}) \approx \frac{N(\rho_{i,j}^2 - 1)^2}{(N-1)^2}.$$

In addition to these basic contributions to Var(log $\hat{\lambda}_s$), a further source of variability comes from the correlations between the estimated means and variances of vital rates that are probabilities. Because the mean of a set of probabilities sets a cap on the possible values of their variance (Morris and Doak 2004), there can be substantial correlations between estimated means and variances across sets of samples, with the sign and the magnitude of the correlation depending on the true mean and variance of the vital rate in question. While Eq. 2 does a good job of approximating Var(log $\hat{\lambda}_s$), these additional terms have nontrivial effects on the predictions, and we discuss how to estimate them in Appendix C. With the addition of these terms, the approximation does an excellent job of predicting the

sampling variation we could expect for a given life history and sampling regime, with the caveat that it will tend to give optimistic estimates of accuracy for populations with very high real environmental variability in vital rates (Appendix C). A MATLAB program that automates all the calculations needed to arrive at Var(log $\hat{\lambda}_s$) estimates with a range of sampling regimes is available in a Supplement to this paper.

To illustrate the use of this approximation, we first ask how sampling intensity and duration will affect estimates of growth rates for a winter annual with highly variable demographic rates, *Collinsia verna*. We use the single stochastic matrix model formulated by Doak et al. (2002) from the much more extensive results of Kalisz (1991). For this species, we first plotted Var(log $\hat{\lambda}_s$) for sampling intensities ranging from 10 to 200 plants followed to estimate each of the four vital rates for this matrix, and for sampling durations ranging from three to 20 years (Fig. 4a). The results indicate that sampling of 50 individuals or less is sufficient to achieve good accuracy, but that uncertainty of estimates will only gradually decline with increased sampling duration, with no clear cut-off beyond which further gains in accuracy are minimal. To further examine these results, and to present them in terms of annual growth rates, which can be easier to interpret, we estimated the 95% confidence limits on $\hat{\lambda}_s$ as exp(log λ_s $\pm$ $2\sqrt{\text{Var}(\log \hat{\lambda}_s)}$) and plotted these bounds against sampling duration for intensities of 20, 100, and 200 individuals (Fig. 4b). This plot emphasizes the minimal gains that result even from the ten-fold increase in sampling intensity from 20 to 200 individuals. At the same time, environmental variation is high enough for this species that even 200 years of data would not be enough to clearly predict whether λ_s is above or below one. Luckily, this species is not endangered, so the exact value of λ_s is not of critical importance.

To illustrate a contrasting situation, and one where estimating population growth rate is of real importance, we analyzed the duration of sampling needed to reach narrow confidence limits on $\hat{\lambda}_s$ for desert tortoise populations in the western Mojave desert. For this threatened and sparse population, we assumed that sample sizes of 30 animals for each rate in each year were the best that are ever likely to be achieved, and used this sampling intensity in conjunction with vital rates presented in Morris and Doak (2002) from data in Doak et al. (1994). For the tortoise, only 10 years of data would be sufficient to reach fairly narrow confidence limits on population growth estimates (Fig. 4c), and even less would be needed to have confidence that decline was occurring, given the vital rates we use. Given that considerably more years of data than this formed the basis of these models (Doak et al. 1994), we can thus have considerable confidence that the population truly was declining.

As these examples show, clear and simple results are easily generated using this approximation. Three more
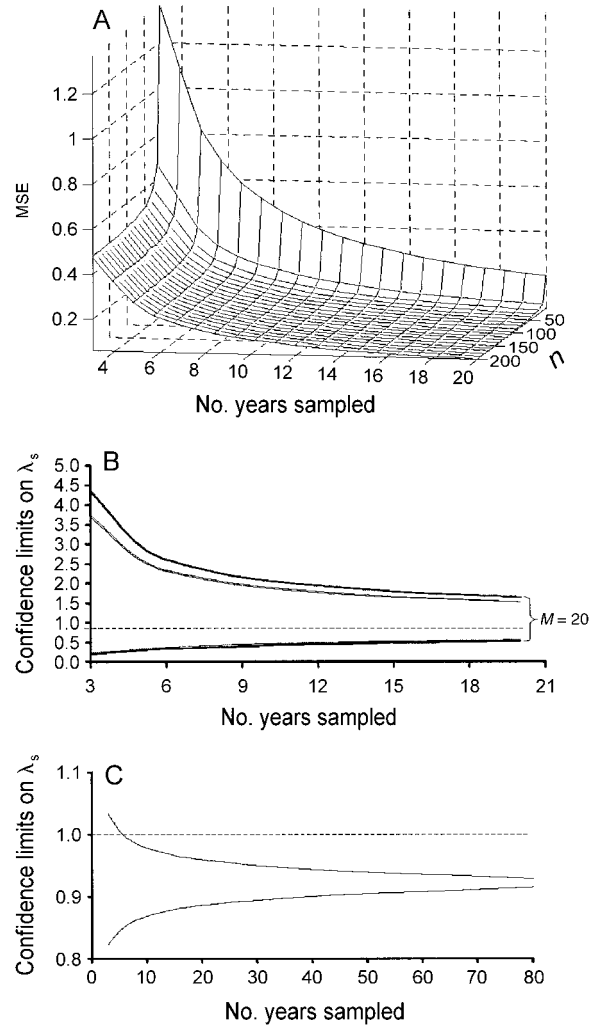

FIG. 4. Effects of sampling regimes on uncertainty in growth rates for *Collinsia verna* and the desert tortoise. (A) Predicted sampling variance in estimates of log λ_s (stochastic lambda) for different sampling intensities and durations for *C. verna*. (B) Upper and lower 95% confidence limits around estimates of λ_s for sampling intensities of $M = 200$, 100, and 20 individuals for *C. verna*. The outermost two lines are for 20 individuals sampled for each vital rate in each year, while the limits for intensities of 100 and 200 individuals are inward and nearly identical. (C) 95% confidence limits around estimates of λ_s for the desert tortoise. In (B) and (C), dashed lines indicate $\lambda_s = 1$, a stable population.

points about its use should be mentioned. First, while our examples all use the same sampling intensity for all rates, there is no difficulty in analyzing more realistic patterns of sampling intensity. Given that larger, older, and less common individuals are frequently the most important members of populations and are also less intensively sampled, these inequities in sampling intensity can have important effects on Var(log $\hat{\lambda}_s$). Second, we usually really will be guessing about the "true" vital rates used to make these approximations, and it therefore behooves us to make several different

guesses, especially about patterns of environmental variability, about which we are likely to be least certain. Finally, simply by leaving out the terms in Eq. 2 for the effects of sampling variation in vital rate variances and correlations, the approximation can also be used to assess the variance in predictions of deterministic models (see also Alvarez-Buylla and Slatkin 1994), allowing a contrast between the best model form to use in generating predictions with limited data.

CONCLUSIONS

In the ecological literature, a great deal of effort has gone into arguing the merits of different demographic outputs and developing elaborate forms of population analysis. Ironically, the problems of how to use these models with real data have received far less attention. Our simulations of limited sampling show that inaccuracies in estimated growth rates can be large with even reasonable amount of sampling effort (see also Fieberg and Ellner 2001). Still, we find that stochastic models will generally perform better than will deterministic ones with moderate to long sampling durations. However, with less than five years of data stochastic models may give uncomfortably variable answers. If sampling duration is low, and the variation in vital rates is high, improved estimates of population behavior may come from deterministic models—a counterintuitive result. In this case, the inaccuracies involved in estimating the effects of environmental stochasticity are so large that they outweigh their own strong effects on population dynamics. More encouragingly, our results suggest that unfeasibly large sampling duration or intensity is not necessary in order for stochastic models to perform reasonably well.

While these simulations provide some insights into the interplay of model complexity and data scarcity, this is only a first stab at the problem of when and how to judge the right degree of complexity to build into a demographic model, whether for use in a PVA or more basic ecological work. Furthermore, we have addressed only one of the many decisions that must be made in setting up a demographic study or analyzing its results (Caswell 2001, Easterling et al. 2000, Gross 2002, Morris and Doak 2002, Kaye and Pyke 2003). In particular, we have emphasized only a single measure of model performance, the variance in predicted population growth rate. Many other outputs of demographic models are of keen interest, including elasticity values and extinction risk, which are likely to be even more sensitive to the problems of data scarcity. In addition, any use of demographic models to forecast growth or extinction relies on measured rates remaining relatively stable into the future, an increasingly worrisome assumption.

It could be assumed that it is always better to construct less biased predictions of population growth, regardless of accuracy. However, we would argue against this view. A more precise estimate of population growth, even if biased, is far more useful than a less reliable one. This is especially true if we know the direction of this bias, as is the case with deterministic growth rate predictions. Especially for PVA models, it can be much more useful to have a prediction of the most optimistic likely outcome than it is to have such an uncertain, but unbiased, prediction that it can support few or no conclusions at all. This said, anyone who is actually doing a demographic analysis should probably perform both stochastic and deterministic analyses, and carefully draw inferences using information from both.

While there have been numerous calls for the presentation of demographic results to include analyzes of their uncertainty, the practical difficulties of doing so mean that most demographic studies still do not include any such quality assessment. The approximation for the sampling variance of population growth rate that we provide should allow such an assessment to be made far more easily. This estimator can also be used to plan future demographic studies to better fulfill the goals of investigators. A great deal of heartache can be expended in the decisions about how long and how intensively to pursue a demographic study. To date, these decisions have been almost always made from intuitions about the trade-off between crippled knees (at least for plant demographers) and data quantity, but with almost no analysis of the power of the study. By providing a formula for sampling uncertainty in growth rates, and software to easily implement it, our hope is that these decisions can be made more carefully, improving the collection and interpretation of demographic information.

LITERATURE CITED

Alvarez-Buylla, E. R., and M. Slatkin. 1994. Finding confidence limits on population growth rates: three real examples revised. Ecology **75**:255–260.

Beissinger, S., 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.

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

Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in conservation biology. Chapman and Hall, London, UK.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer, New York, New York, USA.

Caswell, H. 2001. Matrix population models: construction, analysis and interpretation. Second edition. Sinauer, Sunderland, Massachusetts, USA.

Dixon, P. M., A. M. Ellison, and N. J. Gotelli. 2005. Improving the precision of estimates of the frequency of rare events. Ecology **86**:1114–1123.

Doak, D. F., P. Kareiva, and B. Klepetka. 1994. Modeling population viability for the desert tortoise in the Mojave desert. Ecological Applications **4**:446–460.

Doak, D. F., D. M. Thomson, and E. S. Jules. 2002. PVA for plants: understanding the demographic consequences of seed banks for population health. Pages 312–337 *in* S. R. Beissinger and D. R. McCullough, editors. Population viability analysis. University of Chicago Press, Chicago, Illinois, USA.

Easterling, M. R., S. P. Ellner, and P. M. Dixon. 2000. Size-specific sensitivity: applying a new structured population model. Ecology **81**:694–708.

Ellner, S. P., J. Fieberg, D. Ludwig, and C. Wilcox. 2002. Precision of population viability analysis. Conservation Biology **16**:1–5.

Engen, S., O. Bakke, and A. Islam. 1998. Demographic and environmental stochasticity: concepts and definitions. Biometrics **54**:840–846.

Fieberg, J., and S. P. Ellner. 2001. Stochastic matrix models for conservation and management: a comparative review of methods. Ecology Letters **4**:244–266.

Gross, K. 2002. Efficient data collection for estimating growth rates of structured populations. Ecology **83**:1762–1767.

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

Kalisz, S. 1991. Experimental determination of seed bank age structure in the winter annual, *Collinsia verna*. Ecology **72**:575–585.

Kaye, T. N., and D. A. Pyke. 2003. The effect of stochastic technique on estimates of population viability from transition matrix models. Ecology **84**:1464–1476.

Kendall, B. E. 1998. Estimating the magnitude of environmental stochasticity in survivorship data. Ecological Applications **8**:184–193.

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

Ludwig, D., C. J. Walters, and J. Cooke. 1988. Comparison of two models and two estimation methods for catch and effort data. Natural Resource Modeling **2**:457–498.

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

Morris, W. F., and D. F. Doak. 2004. Buffering of life histories against environmental stochasticity: accounting for a spurious correlation between the variabilities of vital rates. American Naturalist **163**:579–590.

Tuljapurkar, S. 1990. Population dynamics in variable environments. Spring-Verlag, New York, New York, USA.

White, G. C. 2000. Population viability analysis: data requirements and essential analyses. Pages 288–331 *in* L. Boitani and T. K. Fuller, editors. Research techniques in animal ecology: controversies and consequences. Columbia University Press, New York, New York, USA.

APPENDIX A

A description of problems with estimating and using correlations in vital rates is presented in ESA's Electronic Data Archive: *Ecological Archives* E086-062-A1.

APPENDIX B

A description of the simulation models used is presented in ESA's Electronic Data Archive: *Ecological Archives* E086-062-A2.

APPENDIX C

An approximation for sampling variance in stochastic growth rate estimates is presented in ESA's Electronic Data Archive: *Ecological Archives* E086-062-A3.

SUPPLEMENT

A MATLAB program to calculate the sampling variance in estimated log(stochastic lambda) values given different sampling intensities and durations is available in ESA's Electronic Data Archive: *Ecological Archives* E086-062-S1.

SPECIAL FEATURE