

UNCERTAINTY AND VARIABILITY IN DEMOGRAPHY AND POPULATION GROWTH: A HIERARCHICAL APPROACH

JAMES S. CLARK¹

Department of Biology and Nicholas School of the Environment, Duke University, Durham, North Carolina 27708 USA

Abstract. Estimates of uncertainty are the basis for inference of population risk. Uncertainty is estimated from models fitted to data that typically include a deterministic model (e.g., population growth) and stochastic elements, which should accommodate errors in sampling and any sources of variability that affect observations. Prediction from fitted models (of, say, demography) to new variables (say, population growth) requires propagation of these stochastic elements. Ecological models ignore most forms of variability, because they make statistical models complex, and they pose computational challenges. Variability associated with space, time, and among individuals that is not accommodated by demographic models can make parameter estimates and growth rate predictions unrealistic.

I adapt a hierarchical approach to the problem of estimating population growth rates and their uncertainties when individuals vary and that variability cannot be assigned to specific causes. In contrast to an overfitted model that would assign a different parameter value to each individual, hierarchical models accommodate individual differences, but assume that those differences derive from an underlying distribution—they belong to a “population.” The hierarchical model can be implemented in classical (frequentist) and Bayesian frameworks (I demonstrate both) and analyzed using Markov chain Monte Carlo simulation. Results show that population growth models that rely on standard propagation of estimation error but ignore variability among individuals can misrepresent uncertainties in ways that erode credibility.

Key words: Bayesian; demography; hierarchical Bayes; matrix models; population growth; uncertainty.

INTRODUCTION

The need to understand population risk challenges ecologists to accurately estimate variability (Andrewartha and Birch 1954, Strong 1986, Turchin 1995). Time series of population abundance can provide rather direct evidence of growth rate (e.g., Ogden 1993) and of factors that affect it (Elton 1927, Sinclair et al. 1993, Bjornstad et al. 1999). They require reliable, repeated censuses, which are often unavailable for species having broad, diffuse distributions. Observation errors are difficult to incorporate in the classical time series models (Carpenter et al. 1994, Hilborn and Mangel 1997, Calder et al. 2003) and are often ignored. There are few time series of sufficient duration to confidently evaluate influences (Turchin 1995).

Demographic models represent a second approach that provides insight into population growth (Caswell 1989, Easterling et al. 2000). Populations too sparse to census comprehensively can often be studied from demography of sample individuals. Even where time series of total abundance exist, demographic data may

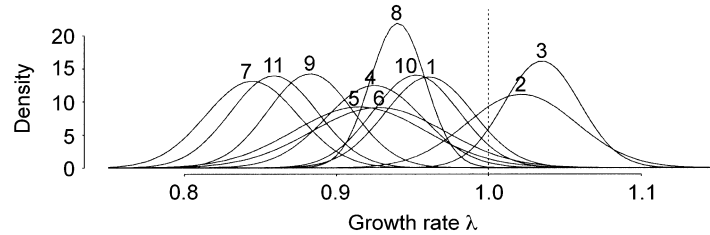
still be needed to understand the basis for population change (Kieth and Windberg 1978, Krebs et al. 2001).

The demographic approach entails statistical challenges that concern how individual- and group-level variability affects growth rate. By contrast with time series of density, population growth rate typically is not estimated directly from data. Rather, the demographic rates are estimated, and the growth rate is calculated using Lotka's equation or a transition matrix. Probability statements thus require not only a means for estimating uncertainty at different scales, but also for propagating demographic uncertainty to growth rate uncertainty. Fig. 1 shows propagated error distributions using standard methods (e.g., Lande 1988) for the growth rates of Northern Spotted Owl (NSO) estimated from 11 data sets. In cases where there is information on the variables that influence demography, they can be used as covariates. In the case of NSO, a recent study made use of 10 years of demographic and environmental data to assess how variability in fecundity and survivorship for three life history stages affect population growth (Franklin et al. 2000). More commonly, the factors that might impact demographic rates are unknown, data are unavailable, or both. In the typical case, demographic rates are fixed constants that apply to all individuals within an age or stage class, and “uncertainty” is limited to estimation error.

Manuscript received 3 December 2001; revised 16 April 2002; accepted 1 May 2002; final version received 21 June 2002. Corresponding Editor: O. N. Bjornstad. For reprints of this Special Feature, see footnote 1, p. 1349.

¹ E-mail: jimclark@duke.edu

FIG. 1. Eleven normal densities defined by means and standard errors (Eqs. 2, 3) from studies of the Northern Spotted Owl. The densities are: (1) Lande (1988); (2) Franklin (1992) males and (3) females; (4) McKelvey et al. (1993); (5) Anderson and Burnham (1990) northwest California; (6) H. J. Andrews, (7, 8) southwest Oregon, and (9) northwest Washington; (10) U.S. Fish and Wildlife (1990) northwest California, and (11) southwest Oregon.



In view of the long-standing attention to “scaling” in ecology, the paucity of formal multilevel analysis is surprising. In public health sciences, the practice of inferring individual risk solely on the basis of group-level (“environmental”) variables is interestingly termed the “ecological fallacy” (Schwartz 1994, Diez-Roux 1998), highlighting the fact that group-level attributes may poorly characterize, mask, or even misrepresent the factors that affect individuals. The term “fallacy” concerns inference at one level based on data collected at another. On the one hand, individual risk is constrained by environmental context, such as local availability of resources. On the other hand, risk is often linked to individual behavior. “Contextual” or “multilevel” analyses address how individual-level responses can change with context. Within classical statistics, multilevel analyses are difficult. There is no standard framework that can accommodate the diversity of model structures that vary from one study to the next.

Here, I demonstrate that the estimates and uncertainties in population growth calculated from demographic data (e.g., Fig. 1) miss the inherent variability associated with individuals, and I describe a straightforward approach to the problem. The statistical models that treat individuals within a class as identical and independent of those in other classes and data sets can produce inflated confidence in growth estimates. They often assume (and, thus, propagate) only estimation error, which declines with sample size, regardless of how well a model “fits” the data.

The method for quantifying uncertainty in population growth entails a hierarchical structure for demographic modeling and propagation. The method is general. It is not restricted to particular types of life histories. It can be applied to models based on Lotka’s equation or on a transition matrix. It does not depend on the extensive data sets that will rarely be available for a few species and never for most. It explicitly quantifies both estimation errors and variability that cannot be assigned to specific sources. It can be implemented within either a classical (empirical Bayes) or fully Bayesian framework (I demonstrate both), although Bayesian methods will typically be most flexible. Because I rely on simulated data, this demonstration does not include observation error, although mark-recapture probabilities (Jolly 1965, Lebreton et al. 1992, Lavine et al. 2002) and multistage sampling (MacGibbon and

Tomberlin 1989, Stroud 1994) readily submit to this framework (e.g., Clark et al. 2003). The approach is not a substitute for extensive knowledge of factors that affect population growth (Clark et al. 1999, Turchin et al. 1999, Franklin et al. 2000, Krebs et al. 2001). It provides a simple framework to allow for stochasticity that is especially valuable when the sources can be precisely known.

THE MODEL FRAMEWORK

To illustrate the methodology of the hierarchical approach, I provide context by starting with the traditional framework.

A traditional approach to demographic data modeling

Ecologists model variability in demographic data by subdividing a population into classes or by parameterizing functions that might explain demographic rates. I use the example of survival, but the same arguments apply to fecundity. The binomial is used to model survival, with likelihood

$$L(y; \theta) \propto \theta^y (1 - \theta)^{n-y} \quad (1)$$

where θ is the survival probability, having maximum likelihood estimate (MLE) $\hat{\theta} = y/n$. This model assumes that n individuals are subject to risk $1 - \theta$, and a random number y survive. An alternative representation $f(y|\theta)$ emphasizes that the likelihood is a joint distribution of data.

If information is available on risk factors, the binomial likelihood is extended in one of two standard ways. First, the population might be (further) subdivided on the basis of age, stage, size, location, sex, and so forth, and each new class might be assigned its own survival parameter θ_j . In other words, “build a bigger matrix.” This subdivision cannot proceed indefinitely; errors swell with shrinking sample sizes, leaving a hodge-podge of overfitted, uninformative estimates (Vandermeer 1978, Moloney 1986). Alternatively, if risk factors \mathbf{x} are quantified, then survival can be related to risk, e.g., $\theta(\mathbf{x}; \beta)$, where β are estimated parameters (e.g., logistic regression). Inference is made about the collection of θ_j estimates (for discrete groups) or about parameters β . Eq. 1 is the default approach when knowledge or data are limited.

Parameter error estimates come from the notion of a parameter sampling distribution. A classical analysis

generally relies on the central limit theorem to assign probabilities to a statistic $\hat{\theta}$, because sampling distributions are usually asymptotically normal:

$$\theta|\mathbf{y} \sim N(\hat{\theta}, [I(\mathbf{y})]^{-1}) \tag{2}$$

where $\hat{\theta}$ is the MLE and where

$$I(\mathbf{y}) = - \left[\frac{\partial^2}{\partial \theta^2} \ln L(\theta; \mathbf{y}) \right]_{\hat{\theta}} \tag{3}$$

is the observed Fisher information, \mathbf{y} represents a vector of the number of survivors, and L is the likelihood (confidence intervals actually come from distributions centered on the confidence limits themselves, but, for normal distributions, the two are equivalent, e.g., Cousins 1995). Fig. 2b, c, and d show sampling distributions for survival and fecundity parameters for Northern Spotted Owl based on the model used by Lande (1988, McKelvey et al. 1993). For a Bayesian approach we have the posterior

$$\theta|\mathbf{y} \sim N(\theta_{\pi}, [I_{\pi}(\mathbf{y})]^{-1}) \tag{4}$$

where θ_{π} is the posterior mode and where

$$I_{\pi}(\mathbf{y}) = - \left\{ \frac{\partial^2}{\partial \theta^2} \ln [L(\theta; \mathbf{y}) \pi(\theta)] \right\}_{\theta_{\pi}} \tag{5}$$

is the ‘‘generalized’’ observed Fisher Information and $\pi(\theta)$ is the prior distribution of θ (e.g., Carlin and Louis 2000).

The variances of these normal distributions (Eqs. 2 and 4) are inversely proportional to sample size n , so they tend to zero variance as sample size increases. For the binomial model (Eq. 1), we have classical and (for a noninformative prior) Bayesian ‘‘standard errors’’:

$$SE_{\theta} = [I(\mathbf{y})]^{-1/2} = \sqrt{\frac{\hat{\theta}(1 - \hat{\theta})}{n}} \tag{6}$$

$$SE_{\theta} = [I_{\pi}(\mathbf{y})]^{-1/2} = \sqrt{\frac{\theta_{\pi}(1 - \theta_{\pi})}{n}} \tag{7}$$

respectively (for the noninformative prior, likelihood and posterior contain the same information). The asymptotic 95% confidence interval is $\hat{\theta} \pm 1.96 SE_{\theta}$. For discrete stages, standard errors can be calculated for each $\hat{\theta}_j$. If θ is assumed to vary continuously with variable x , then there are asymptotic standard errors for parameters β .

By contrast with confidence intervals on demographic parameters, confidence intervals on growth rate λ are calculated. The typical error propagation method comes from the linear term of a Taylor series:

$$\text{var}[\lambda] \approx \sum_{j=1}^m \left(\frac{\partial \lambda}{\partial \theta_j} \right)^2 SE_{\theta_j}^2 \tag{8}$$

(e.g., Stuart and Ord 1994), with asymptotic confidence interval $\lambda \hat{\theta} \pm 1.96 SE_{\lambda}$ and predictive distributions in Fig. 1.

Fig. 2 compares an analysis using standard errors reported from the literature with one that assumes the

same error structure, but having five times as many samples. Rather than the normal approximation (Eq. 8), I used a resampling routine (bootstrap) to generate the predictive distributions for λ in Fig. 2a, but they are similar for large sample size. The figure illustrates the assumption of a ‘‘true’’ parameter value that is approached as sampling effort increases. The predictive distribution on λ likewise collapses with large n .

The problem

Neither of the alternatives for modeling parameter variability (Eqs. 2 and 4) is ideal for most demographic data: differences among individuals or subpopulations are large, the causes are poorly understood, and the explanatory variables are not quantified. The Bernoulli (coin-flipping) process that gives rise to Eq. 1 does not accommodate such differences, and nothing in the estimation error of θ compensates for it. Eq. 1 only applies when all individuals are flipping identical coins. The implications of Fig. 2 are not appreciated by most ecologists or by the managers who interpret their analyses. Fig. 2 could be realistic if the underlying demographic rates were fixed and identical for all individuals. The problem results from the fact that the model allows no place for variation among individuals or groups. It does not depend on technical details; it exists regardless of whether the approach is classical (Eqs. 2, 6) or Bayesian (Eqs. 4, 7). While adding more stages to a matrix model allows for a greater range of rates, there is still within-stage variability. The problem is not solved by independently fitting more parameters to smaller data sets (e.g., highly subdivided populations). Overfitting results, because the individual parameters have no predictive capacity. The assumption that each is independent is unrealistic. Independent estimates do not ‘‘inform’’ one another, as would be expected if individuals belong to a ‘‘population.’’

Finally, variability among individuals is not the ‘‘demographic stochasticity’’ of discrete-state, difference-equation models (May 1973, Hensen et al. 2001). Demographic stochasticity is associated with the realization of discrete events (birth and death), it exists regardless of whether underlying parameters vary, and it becomes important when densities are low. Individual variation matters at all densities, because it affects the distribution of rates, whether a process is modeled as continuous or discrete.

A hierarchical solution

Here, I adapt a hierarchical strategy for modeling population growth rate that accommodates variability that cannot be assigned to a specific cause. It explicitly models random effects on parameters and estimation error of ‘‘hyperparameters’’ that summarize parameter densities. Rather than overfitting with many independent parameters, it admits parameter variability, but assumes that individuals derive from a population. Hierarchical models are a recent development designed

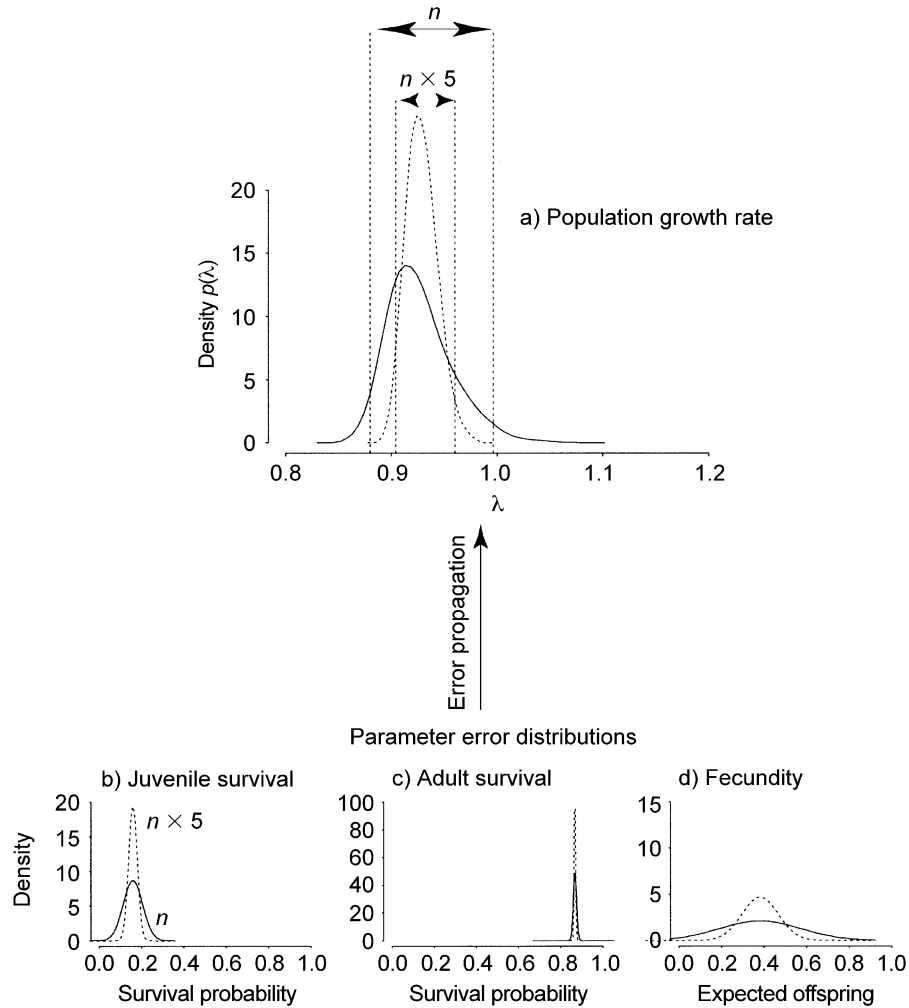


FIG. 2. Parameter sampling distributions and their effect on uncertainty in growth rate of the population (a). The density of growth rates is obtained by resampling from parameter sample distributions. Dashed lines assume the same data structure, but sample sizes five times as large.

to allow for processes that are more variable than admitted by traditional models (Gelman et al. 1995, Carlin and Louis 2000).

A hierarchical model.—Consider the sampling distribution for a variable y represented by the density $f(\mathbf{y}|\theta)$. The likelihood is the joint probability conditioned on a parameter θ :

$$L(\theta; \mathbf{y}) = \prod_{j=1}^n f(y_j|\theta). \tag{9}$$

For the coin-flipping process in Eq. 1, y_j assumes values of 1 or 0, with probability θ and $1 - \theta$, respectively. A hierarchical model assumes that parameters have distributions of their own. For example, survival probability θ may vary among individuals. Let $\pi(\theta|\alpha, \beta)$ represent the density of parameter θ ; the “hyperparameters” α and β describe this density. The hierarchical version of Eq. 9 is as follows:

$$g(\mathbf{y}|\alpha, \beta) = \int_{-\infty}^{\infty} \prod_{i=1}^n f(y_i|\theta_i)\pi(\theta_i|\alpha, \beta) d\theta$$

$$= \prod_{i=1}^n f(y_i|\alpha, \beta). \tag{10}$$

Within a classical framework, Eq. 10 could be substituted for Eq. 9:

$$L(\alpha, \beta; \mathbf{y}) \equiv g(\mathbf{y}|\alpha, \beta). \tag{11}$$

Rather than fitting θ , we fit α and β . Likelihood Eq. 11 has a broader spread than likelihood Eq. 9, because it accommodates the additional variability conferred by the distribution of θ . This approach is termed parametric empirical Bayes (PEB; e.g., Morris 1983, Ver Hoef 1996).

The hierarchical approach that led to Eq. 10 differs from a Bayesian analysis, which might begin with the

same two distributions, but treat them in a different way. A nonhierarchical Bayesian model writes a posterior density for θ using Bayes' theorem:

$$p(\theta | \mathbf{y}) = \frac{\prod_{i=1}^n f(y_i | \theta) \pi(\theta)}{\int_{-\infty}^{\infty} \prod_{i=1}^n f(y_i | \theta) \pi(\theta) d\theta} = \frac{f(\mathbf{y} | \theta)}{g(\mathbf{y})} \times \pi(\theta). \quad (12)$$

(The parameters of $\pi(\theta)$ are suppressed for brevity.) The posterior distribution contains the marginal distribution of the data as a normalizing constant (the denominator of Eq. 12), but the likelihood itself (the assumed distribution of data conditioned on θ) remains nonhierarchical (it is Eq. 9). As data dominate the prior in Eq. 12, the posterior assumes the shape of a normalized likelihood, described by Eq. 4. Unlike the hierarchical (PEB) model, the nonhierarchical Bayesian posterior converges to the classical result (Eq. 4 looks like Eq. 2). Like the classical nonhierarchical model, Eq. 12 does not admit variability in θ . It simply expresses our posterior belief in values of θ , which collapses to a point estimate with large sample size.

The hierarchical Bayesian method adopted here makes use of the likelihood Eq. 9, but assigns priors $u(\alpha)$ and $v(\beta)$ to yield the joint posterior:

$$p(\theta, \alpha, \beta | \mathbf{y}) = \frac{\prod_{i=1}^n f(y_i | \theta_i) \pi(\theta_i | \alpha, \beta) u(\alpha) v(\beta)}{g(\theta | \alpha, \beta)} \quad (13)$$

where θ_i indicates that each individual can have a different demographic rate, but all are linked through hyperparameters. (In theory, priors need not be independent of one another.)

The hierarchical model (Eq. 11 or 13) admits variability. Hyperparameters $[\alpha, \beta]$ have the estimation error, while the individual parameters θ_i have distributions that depend on data for the i th individual and for the entire population. Two parameters define a conditional distribution for the θ_i ($\pi(\theta_i | \alpha, \beta)$ in Eq. 10), and this distribution does not collapse with increasing sample size. The posterior is obtained using Monte Carlo simulation.

A predictive distribution for λ .—Because λ is calculated, not estimated directly, we require methods for constructing a distribution for λ . We shall see below that the assumption of normality can be unreasonable for the hierarchical model, so Eq. 8 is inappropriate. Draws are taken from the posterior (Bayesian) or sampling (classical) distributions. For each resampled parameter set, the growth rate λ is calculated (using Lotka's equation or matrix methods) and used to assemble the predictive distribution. Because estimates are typically in the form of a list constructed by Markov chain Monte Carlo methods (MCMC; Bayesian) or bootstrap

TABLE 1. Structure used for data simulation.

Parameter	Parameter estimate	Variance among individuals	Parameter distributions [†]	n
s_0	0.159	0.01	Beta(1.97, 10.4)	63
s	0.868	0.01	Beta(9.08, 1.38)	114
b	0.382	0.09	Gam(1.62, 4.24)	197

[†] Distributions used to draw random data sets; n parameter values are drawn from each distribution, followed by n data values from Binomial (survival) or Poisson (fecundity) distributions.

(classical), the distribution of λ is straightforward. Confidence intervals incorporate estimation error and individual variability.

APPLICATION

To demonstrate the approach, I use the example of the Northern Spotted Owl (NSO). Many analyses of this familiar example employ simple demographic models, including several excellent studies of parameter variability (Franklin et al. 2000). Because NSO has a broad distribution, low density, long life, and large home range, they exemplify a population that is difficult to census comprehensively. I adopt the simplest model, because key points are not model specific. The analysis consists of (1) generating random data having sample sizes and parameter means and standard errors of literature values (Table 1), (2) development of the hierarchical models, and (3) comparison with the traditional analysis. I focus on effects of sample size and the substructure assumed in statistical models.

Data structure

Parameter sampling distributions from the literature are drawn as solid lines in Fig. 2b, c, and d. I used the standard assumption of a homogeneous population of size n and compared the analysis with that for two levels of substructure. The first assumes 10 subpopulations of size n_i approximately equal to $n/10$ (rounded to the nearest integer). The second assumes that variability exists at the individual level.

I used standard distributions to describe variability among subgroups or individuals. Fecundity is a Poisson process, with offspring production from the i th subgroup $y_i \sim \text{Pois}(b_i)$. Fecundity is assumed to vary among subgroups according to the gamma density $b_i \sim \text{Gam}(\alpha_b, \beta_b)$. This mixture incorporates the common observation, including from NSO (Barrowclough and Coats 1985), that variability in offspring production is greater than Poisson. I assumed the mean value \hat{b} used by McKelvey et al. (1993; Table 1) and a conservative variance of $V_b = 0.09$. Gamma parameters were selected by moment matching with $\alpha_b = \hat{b}/V_b$ and $\beta_b = \hat{b}/V_b$ (Table 1). Simulation consisted of drawing n_i Poisson parameters from $\text{Gam}(\alpha_b, \beta_b)$ followed by n_i Poisson variates.

Identical methods were used for survivorship s and s_0 . Survivorship is $y_i \sim \text{Bin}(n_p, s_i)$ where s_i is survivorship of the i th subgroup. For individual variation, $y_i \sim \text{Bin}(1, s_i) = \text{Bernoulli}(s_i)$. Variability among groups follows a beta density $s_i \sim \text{Beta}(\alpha_s, \beta_s)$. I use the mean value \hat{s} of McKelvey et al. (1993) and variance $V_s = 0.01$. Beta parameters were selected by moment matching, $\alpha_s = \hat{s}[\hat{s}(1 - \hat{s})/V_s - 1]$ and $\beta_s = \alpha_s(1 - \hat{s})/\hat{s}$. I drew n_i binomial parameters from $\text{Beta}(\alpha_s, \beta_s)$ followed by n_i binomial variates.

A standard population model

The simple model of Lande (1988, McKelvey et al. 1993) involves three survival rates, juvenile, subadult, and adult, a fecundity parameter, and a maturation age. Subadult and adult survival rates are similar, and McKelvey et al. (1993) use the same value. The transition matrix is

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & b \\ s_0 & 0 & 0 \\ 0 & s & s \end{bmatrix}. \tag{14}$$

Roots were obtained directly from Lotka's equation, $\lambda = (s \pm \sqrt{s^2 + 4bs_0s})/2$ (Lande 1988). McKelvey et al. (1993) calculated $\hat{\lambda} = 0.925$ and determined that a value of $\lambda = 1$ was outside the 95% CI. The growth rate can also be estimated by eigenanalysis of \mathbf{A} (Eq. 14). An effect of sample size on propagated error is shown in Fig. 2a.

A hierarchical analysis

The hierarchical structure was applied in both classical (PEB) and Bayesian contexts.

PEB.—For the PEB analysis, the likelihood function is a mixture, obtained by marginalizing over parameter distributions. For fecundity, I use Eq. 10, together with the gamma distributed b to obtain the likelihood that is negative binomial (NB):

$$\begin{aligned} L(\alpha_b, \beta_b; \mathbf{y}) &= \int_{-\infty}^{\infty} \prod_{i=1}^n \text{Poisson}(y_i | b) \text{Gam}(b | \alpha_b, \beta_b) db \\ &= \prod_{i=1}^n \text{NB}(y_i | \alpha_b, \beta_b). \end{aligned}$$

For survival I use the marginal beta-binomial (BetaBin):

$$\begin{aligned} L(\alpha, \beta; \mathbf{y}) &= \int_{-\infty}^{\infty} \prod_{i=1}^n \text{Bin}(y_i | n, s) \text{Beta}(s | \alpha, \beta) ds \\ &= \prod_{i=1}^n \text{BetaBin}(y_i | n, \alpha, \beta). \end{aligned}$$

Parameter sampling errors were estimated by nonparametric bootstrap. Simulated data were sampled with replacement and fitted to negative binomial or beta-binomial likelihoods.

Hierarchical Bayes.—The Bayesian analysis makes use of Eq. 13 and a Gibbs sampler, a type of MCMC.

The algorithm described in the Appendix results in posterior distributions of parameters and hyperparameters. A kernel estimator with a narrow bandwidth (Silverman 1986) was used to smooth marginal posteriors produced by MCMC simulation. All CIs are 95%.

Predictive distribution for λ .—Distributions of λ were produced from bootstrapped parameter estimates (PEB) or MCMC (Bayes).

RESULTS

To illustrate how a nonhierarchical analysis might compare with a hierarchical model that admits modest structure, I initially assume 10 subgroups simulated using parameters from Table 1. Fig. 3 shows bootstrapped hyperparameter distributions and “parameter” distributions for the PEB model. Parameter sample distributions (Fig. 3b) are determined by hyperparameter distributions (Fig. 3c). The PEB model for fecundity is unstable, because bootstrapped resamples contained too many zeros to consistently yield parameter estimates—to produce Fig. 3, I bounded the search algorithm below $\alpha_b = 20$ (Table 2). Although asymptotics provide approximate CIs (Eq. 6), they are not useful for this parameter having mean near zero. George et al. (1993) report similar broad dispersion in MLEs for a comparable PEB model when fully Bayesian inference yielded more consistent results. Parameter distributions for survivorship are homogeneous (Fig. 3c). The spread in parameter distributions (Fig. 3b) describes the variability that could have been missed by a traditional model (compare Fig. 2).

The Bayesian analysis (Fig. 4, Table 3) yields stable hyperparameter estimates, despite many parameters. For each of the three demographic rates, there are two hyperparameters plus 10 parameters, one for each subgroup. The parameter distributions in Fig. 4b are integrated over the n individual posteriors, which, collectively, are summarized by two hyperparameter posteriors (Fig. 4c). The predictive distribution for λ is broad, reflecting the variability in parameters, which the nonhierarchical model does not permit (Fig. 4a).

Despite contrasting approaches of PEB and hierarchical Bayes, at the levels of parameters and population growth rate the results are similar to each other and different from the traditional model. Subgroup variability results in similar parameter distributions for the two hierarchical models (Fig. 5b), whereas the nonhierarchical model overestimates parameter confidence (Fig. 5c). The predictive distributions for λ that come from PEB and hierarchical Bayes are also similar and contrast with the unrealistically narrow distribution predicted by the traditional approach (Fig. 5a). Neither of the broad, hierarchical models excludes $\lambda = 1$ at 95% confidence.

The hierarchical model does not collapse simply because sample size is large. The comparison of sample size effects from Fig. 2 was repeated using hierarchical Bayes. The effects on hyperparameters are noticeable

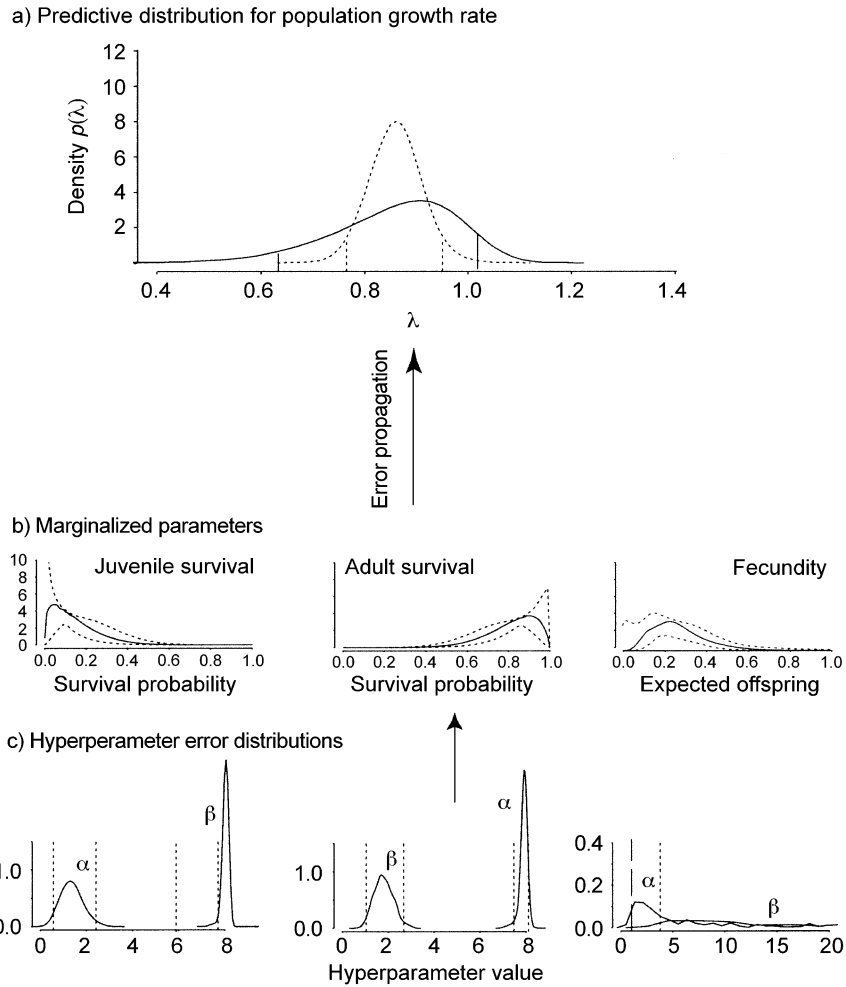


FIG. 3. Hierarchical model analysis using a bootstrap and PEB: (a) comparison of hierarchical (dashed) and nonhierarchical (solid) models; (b) parameter distributions and 95% envelopes calculated by resampling from hyperparameter distributions in (c). Vertical lines in (a) and (c) are 95% confidence intervals.

TABLE 2. Parameters and quantiles for a PEB analysis of a data set that assumed 10 subgroups.

Para- meter†	Hyper para- meter	MLE	1 SE	Quantiles	
				0.025	0.975
s_0	α_s	0.143	0.0441		
	β_s	1.34	0.459	0.557	2.41
		8.05	0.260	7.71	8.22
s		0.816	0.363		
	α_s	92.2	0.139	7.56	8.10
	β_s	20.8	0.397	1.04	2.58
b		0.270	0.190		
	α_b	2.02	1.69	1.04	6.79
	β_b	7.48	5.80	3.87	20.0

† Parameter distributions are marginalized over hyperparameter distributions.

(Fig. 6c), but parameter posteriors (Fig. 6b) and the shape of the predictive distribution for λ (Fig. 6a) are little affected.

If random effects exist at the individual level, parameters and growth rates become more dispersed (Fig. 7a). To isolate the effects of substructure in Table 1, the examples in Fig. 7 are all Bayesian. As previously, there is individual variation, and total population sizes and parameters come from Table 1. Models assume homogeneity (7a, bottom: nonhierarchical Bayes), 10 subgroups (7a, middle), and individual variation (7a, top). Ten simulated data sets for each level of structure show that the hierarchical model consistently reproduced the underlying mean growth rate, with broad spread reflecting variability. The nonstructured model predicts variable mean values but narrow credible intervals. Fig. 7b shows that the hierarchical model shows only slightly narrower confidence intervals when sample size is increased fivefold.

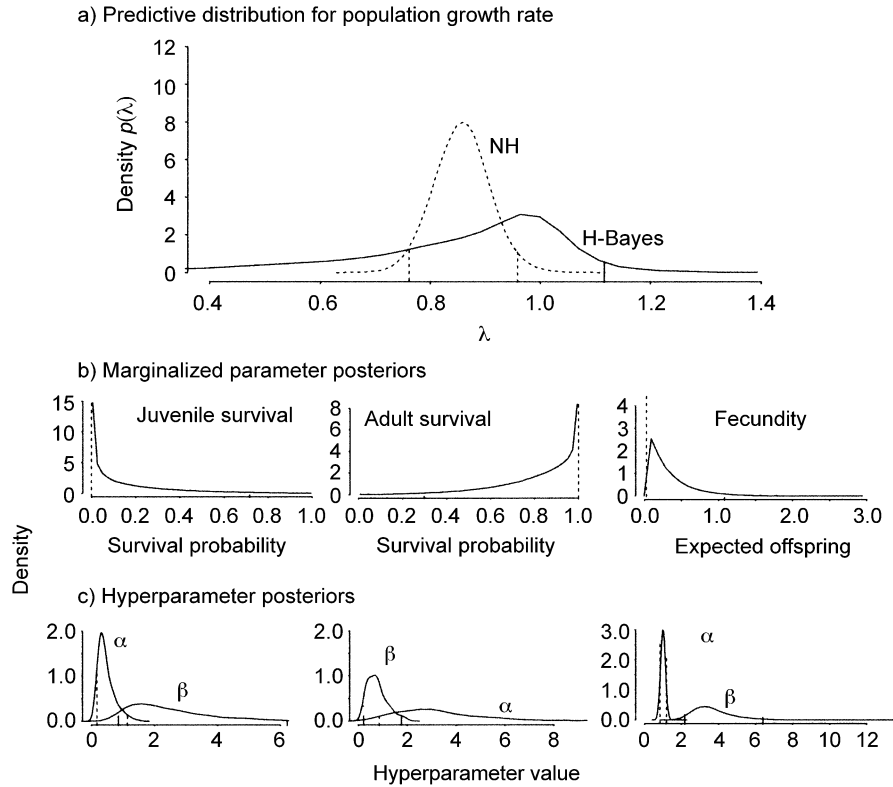


FIG. 4. Same as Fig. 3 for the hierarchical Bayesian analysis. Abbreviations are: NH, nonhierarchical; H-Bayes, hierarchical Bayes.

DISCUSSION

Demographic varies among individuals

Change in a demographic rate like survival might come to pass in at least three different ways:

- A) All individuals are subject to the same rate, and all experience a shift from one rate to another rate.
- B) Individuals might be subject to different rates, but all rates change by the same magnitude.
- C) Individuals are subject to different rates, and the average rate changes with the changing fraction of the population exposed to different risks.

Ecologists could agree that scenario C is most common, and scenario A is usually unrealistic. Individuals die from many causes, and they vary in their exposure to different risks. The population survival rate averages over individuals having differential susceptibility to disease, predation, pathogens, accidents, and so forth. Individuals affect the population mean as they enter and leave high-risk categories; the more individuals with high exposure to one risk or another, the lower the mean survival. A change in the mean usually does not occur because the risk from all mortality sources change in concert.

Traditional models do not accommodate scenario C; they do not even accommodate scenario B. Hierarchical models are designed for scenarios B and C. Observed

variation in fecundity and mortality is typically larger than admitted by the standard distributions used to describe them (Poisson and binomial, respectively; Barrowclough and Coats 1985, Franklin et al. 2000, Wyckoff and Clark 2002). These overdispersed distributions are the motivation for hierarchical modeling.

Hierarchical modeling: where, when, and how much?

Ecologists recognize that different factors operate at different scales (e.g., Carpenter 1996), and that the re-

TABLE 3. Bayesian analysis of a data set that assumed 10 subgroups, showing parameters with posterior moments and quantiles.

Parameter†	Hyperprior value	Posterior mean	Posterior quantiles	
			0.025	0.975
s_0		0.155	0.00000361	0.716
α_s	0.5	0.387	0.176	1.14
β_s	0.5	2.14	0.859	6.21
s		0.812	0.297	0.999
α_s	0.5	2.95	0.844	8.69
β_s	0.5	0.708	0.215	1.76
b		0.286	0.00637	1.08
α_b	1	0.999	0.8323	1.19
β_b	1	3.49	2.19	6.39

† Parameters are marginalized over parameter distributions for n individuals.

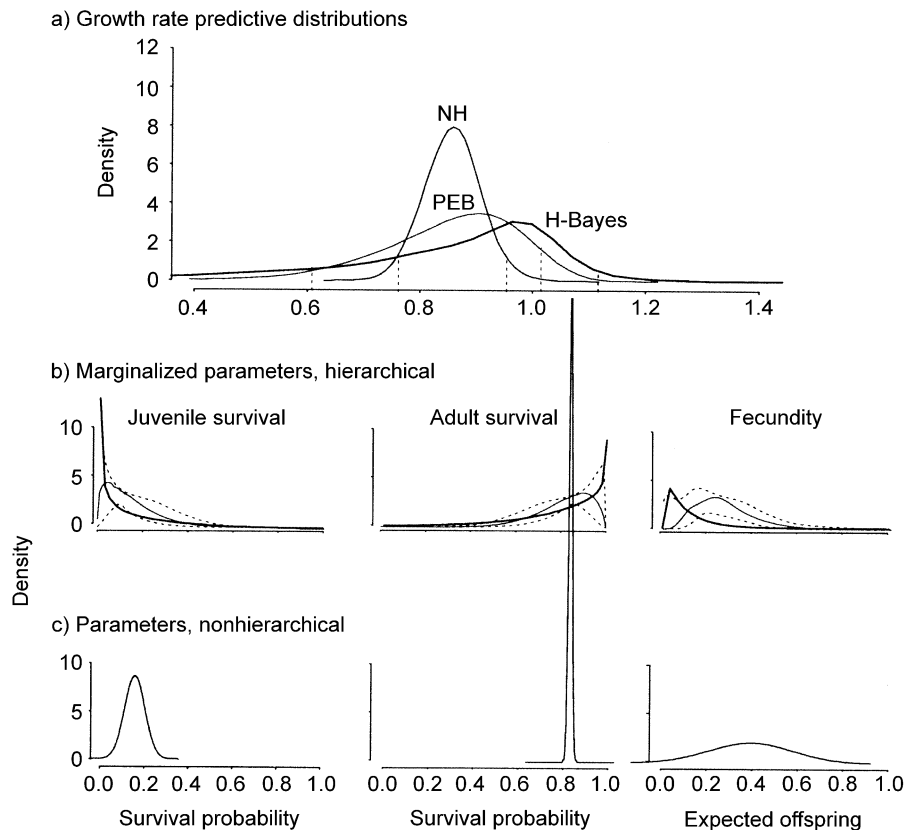


FIG. 5. Comparison of the standard nonhierarchical (NH) and two hierarchical approaches.

relationships among the same variables can vary from one scale to the next (e.g., Chase and Leibold 2002). To illustrate fundamental impacts on inference, I have focused on a simple model structure. The approach is readily extended to admit contextual influences that can be used to explore covariates that operate at different scales or, when influences are unknown, random effects structured in appropriate ways. There are few data sets for which comprehensive information permits incorporation of the many factors that control variability in demographic rates (Clark et al. 1999). Franklin et al. (2000) provide an excellent, albeit atypical, example where extensive data permit a wide-ranging analysis of effects. Hierarchical models provide a way forward by allowing for the variability that cannot be assigned to specific causes. The advantages of a hierarchical approach are not limited to population growth estimates that are calculated from demographic data. For example, hierarchical structures can be applied to individual growth responses (Gelfand et al. 1990, Clark et al. 2003), to mortality (Stroud 1994, Lavine et al. 2002), to population time series (Bjornstad et al. 1999), and to population spread (Wikle 2003). They are not restricted to individual effects, but also apply when variability is structured in space, in time, and among

groups (e.g., MacGibbon and Tomberlin 1989, Zeger and Karim 1991).

Hierarchical models represent a middle ground between traditional models that do not admit individual variation (solid bar in Fig. 8a) and overfitted models that, taken to the extreme, could assign an estimate to each observation (a histogram of estimates in Fig. 8b). The marginalized posterior density of fecundity taken across all observations (Fig. 8a) integrates n posteriors; each observation has its own posterior. But, unlike n independent estimates, the individual estimates are each informed by the full data set (Fig. 8b). The collective posterior includes the variability among observations, without assuming that they simply represent n independent populations, each with a sample size of 1, and it is summarized by hyperparameters.

Hierarchical Bayes is more straightforward (and simple) than classical methods (e.g., Zeger and Karim 1991, Gelfand and Sahu 1999). For example, a classical analysis that allows for dependence between adjacent cohorts of demersal marine fish (Myers and Cadigan 1993) could accommodate only parts of the data, parameters did not consistently converge or "make sense" (e.g., correlations >1), and the complex and highly specialized analysis is not flexible to alternative

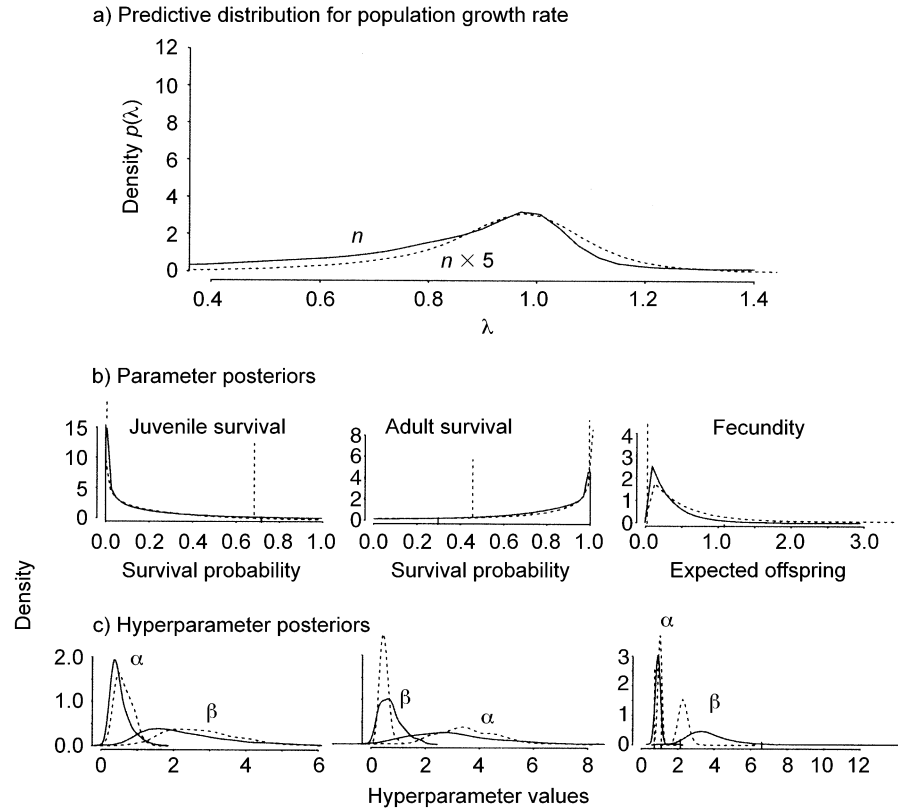


FIG. 6. Increased sample size (from n to $n \times 5$) for the hierarchical model reduces spread on hyperparameter posteriors (c), but does not have much effect on parameters posteriors (b) or on the growth rate uncertainty (a).

assumptions. The hierarchical Bayes model, analyzed with a sampling based approach, readily admits diverse assumptions without resort to diverse structures (Gelfand and Smith 1990).

The cost of ignoring variability

Confidence intervals can be unrealistically narrow, because standard analysis does not admit that demo-

graphic rates vary among individuals. Classical models treat demographic rates as fixed constants, whereas individuals display dramatic differences in mortality risk and reproductive potential. These differences result from age, size, sex, habitat, behavior, and their interactions with environmental variability. In the classical context, it is difficult to allow for the variability that all ecologists recognize as typical of natural populations.

Forecast failure can be one of the most important consequences of analyses that ignore individual variability. Managers require realistic estimates of uncertainty. Fig. 5a contrasts uncertainty estimated from the standard nonhierarchical model, predicting low probability of a sustainable population, with the hierarchical forecast acknowledging variability. Hierarchical models admitting modest individual variation might include $\lambda = 1$ within a 95% CI. This result does not increase optimism, because it also raises the possibility that growth rates are perilously low. Rather than optimism, the message is one of more realistic evaluation of the possibilities. Narrow confidence intervals have the short-term advantage of fostering a sense of precision. The mid- to long-term consequence can be loss of credibility.

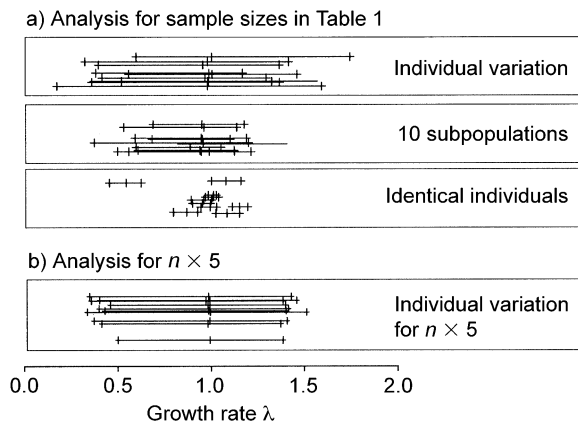


FIG. 7. (a) Comparison of growth-rate 95% confidence intervals for three levels of structure for 10 simulated data sets. (b) Confidence intervals for 10 simulated data sets assuming five times as many samples.

ACKNOWLEDGMENTS

For discussion and comments on the manuscript, I thank O. Bjørnstad, K. Calder, M. Dietze, P. de Valpine, J.

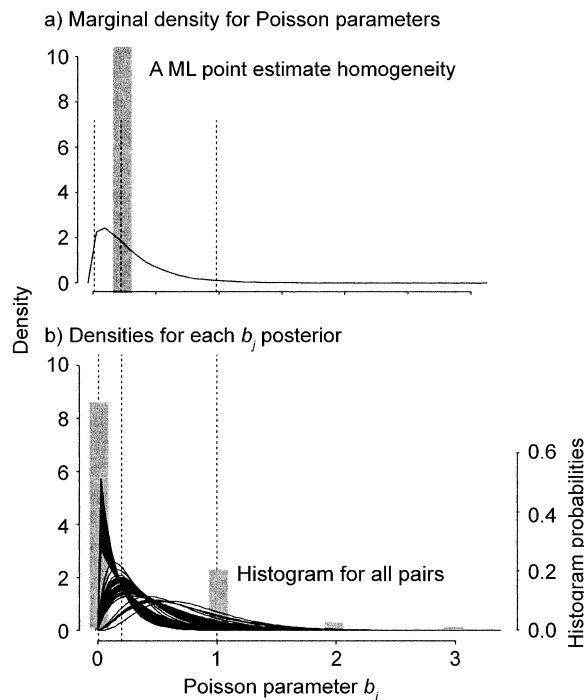


FIG. 8. Three assumptions for fecundity parameter distributions. A nonhierarchical model assumes that all individuals have the same fixed parameter value (shaded value at the MLE in [a]). An overfitted model assumes each individual has its own parameter value (shaded histogram bars representing the fraction of the population having a given value in [b]). The hierarchical model yields a conditional distribution for each individual that is informed by all other individuals (individual densities in [b]).

HilleRisLambers, I. Ibanez, S. LaDeau, J. McLachlan, J. Mohan, M. Wolosin, and an anonymous reviewer.

LITERATURE CITED

- Anderson, D. R., and K. P. Burnham. 1990. Demographic analysis of Northern Spotted Owl populations. Appendix to 1990 status review: Northern Spotted Owl (*Strix occidentalis caurina*). U.S. Fish and Wildlife Service, Washington, D.C., USA.
- Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, Illinois, USA.
- Barrowclough, G. F., and S. L. Coats. 1985. The demography and population genetics of owls, with special reference to the conservation of the spotted owl (*Strix occidentalis*). Pages 74–85 in R. J. Gutierrez and B. Carey, editors. Ecology and management of the spotted owl in the Pacific Northwest. USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, USA.
- Bjørnstad, O. N., J.-M. Fromentin, J. C. Stenseth, and J. Gjøsæter. 1999. Cycles and trends in cod populations. Proceedings of the National Academy of Sciences (USA) **96**: 5066–5071.
- Calder, C., M. Lavine, P. Møller, and J. S. Clark. 2003. Incorporating multiple sources of stochasticity into dynamic population models. Ecology **84**:1395–1402.
- Carlin, B. P., and T. A. Louis. 2000. Bayes and empirical Bayes methods for data analysis. Chapman and Hall, Boca Raton, Florida, USA.
- Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology **77**:677–680.
- Carpenter, S. R., K. L. Cottingham, and C. A. Stow. 1994. Fitting predator–prey models to time series with observation errors. Ecology **75**:1254–1264.
- Caswell, H. 1989. Matrix population models. Sinauer, Sunderland, Massachusetts, USA.
- Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity–biodiversity relationship. Nature **416**: 427–430.
- Clark, J. S., B. Beckage, P. Camill, B. Cleveland, J. HilleRisLambers, J. Lichter, J. MacLachlan, J. Mohan, and P. Wyckoff. 1999. Interpreting recruitment limitation in forests. American Journal of Botany **86**:1–16.
- Clark, J. S., M. Dietze, I. Ibanez, and J. Mohan. 2003. Coexistence: how to identify trophic trade-offs. Ecology **84**: 17–31.
- Cousins, R. D. 1995. Why isn't every physicist a Bayesian? American Journal of Physics **63**:398–410.
- Diez-Roux, A. V. 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. American Journal of Public Health **88**:216–222.
- Easterling, M. R., S. P. Ellner, and P. M. Dixon. 2000. Size-specific sensitivity: applying a new structured population model. Ecology **81**:694–708.
- Elton, C. 1927. Animal ecology. Sedgwick and Jackson, London, UK.
- Franklin, A. B. 1992. Population regulation in Northern Spotted Owls: theoretical implications for management. Pages 815–827 in D. R. McCullough and R. H. Barrett, editors. Wildlife 2001: populations. Elsevier Applied Science, London, UK.
- Franklin, A. B., D. R. Anderson, R. J. Gutierrez, and K. P. Burnham. 2000. Climate, habitat quality, and fitness in Northern Spotted Owl populations in northwestern California. Ecological Monographs **70**:539–590.
- Gelfand, A. E., S. E. Hills, A. Racine-Poon, and A. F. M. Smith. 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. Journal of the American Statistical Association **85**:972–985.
- Gelfand, A. E., and S. K. Sahu. 1999. Identifiability, improper priors, and Gibbs sampling for generalized linear models. Journal of the American Statistical Association **94**:247–253.
- Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association **85**:398–409.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman Hall, London, UK.
- George, E. I., U. E. Makov, and A. F. M. Smith. 1993. Conjugate likelihood distributions. Scandinavian Journal of Statistics **20**:147–156.
- Henson, S. M., R. F. Costantino, J. M. Cushing, R. A. Desharnais, B. Dennis, and A. A. King. 2001. Lattice effects observed in chaotic dynamics of experimental populations. Science **294**:602–605.
- Hilborn, R., and M. Mangel. 1997. The ecological detective. Princeton University Press, Princeton, New Jersey, USA.
- Jolly, G. M. 1965. Explicit estimates from capture–recapture data with both death and dilution—stochastic model. Biometrika **52**:225–247.
- Kieth, L. B., and L. A. Windberg. 1978. A demographic analysis of the snowshoe hare cycle. Wildlife Monographs **58**:1–70.
- Krebs, C. J., R. Boonstra, S. Boutin, and A. R. E. Sinclair. 2001. What drives the 10-year cycle of snowshoe hares? BioScience **51**:25–35.
- Lande, R. 1988. Demographic models of the Northern Spotted Owl (*Strix occidentalis caurina*). Oecologia **75**:601–607.

- Lavine, M., B. Beckage, and J. S. Clark. 2002. Statistical modeling of seedling mortality. *Journal of Agricultural, Biological, and Environmental Statistics* **7**:21–41.
- Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* **62**:67–118.
- MacGibbon, B., and T. J. Tomberlin. 1989. Small area estimates of proportions via empirical Bayes techniques. *Survey Methodology* **15**:237–252.
- May, R. M. 1973. *Stability and complexity in model ecosystems*. Princeton University Press, Princeton, New Jersey, USA.
- McKelvey, K., B. R. Noon, and R. H. Lamberson. 1993. Conservation planning for species occupying fragmented landscapes: the case of the Northern Spotted Owl. Pages 424–450 *in* P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, editors. *Biotic interactions and global change*. Sinauer, Sunderland, Massachusetts, USA.
- Moloney, K. A. 1996. A generalized algorithm for determining category size. *Oecologia* **69**:179–180.
- Morris, C. 1983. Parametric empirical Bayes inference: theory and applications. *Journal of the American Statistical Association* **78**:47–65.
- Myers, R. A., and N. G. Cadigan. 1993. Density-dependent juvenile mortality in marine demersal fish. *Canadian Journal of Fisheries and Aquatic Sciences* **50**:1576–1590.
- Ogden, J. 1993. Population increase and nesting patterns of the black noddy *Anous minutus* in Pisonia forest on Heron Island: Observations from 1978, 1979, and 1992. *Australian Journal of Ecology* **18**:395–403.
- Schwartz, S. 1994. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *American Journal of Public Health* **84**:819–824.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Chapman and Hall, London, UK.
- Sinclair, A. R. E., J. M. Gosline, G. Holdsworth, C. J. Krebs, S. Boutin, J. N. M. Smith, R. Boostra, and M. Dale. 1993. Can the solar cycle and climate synchronize the snowshoe hare cycle in Canada? Evidence from tree rings and ice cores. *American Naturalist* **141**:173–198.
- Strong, D. R. 1986. Density vagueness: abiding the variance in the demography of real populations. Pages 257–268 *in* J. Diamond and T. J. Case, editors. *Community ecology*. Harper and Row, Cambridge, UK.
- Stroud, T. 1994. Bayesian analysis of binary survey data. *Canadian Journal of Statistics* **22**:33–45.
- Stuart, A., and J. K. Ord. 1994. *Kendall's advanced theory of statistics*. Edward Arnold, London, UK.
- Turchin, P. 1995. Population regulation: old arguments and a new synthesis. Pages 19–40 *in* N. Cappuccino and P. W. Price, editors. *Population dynamics: new approaches and synthesis*. Academic, San Diego, California, USA.
- Turchin, P., A. D. Taylor, and J. D. Reeve. 1999. Dynamical role of predators in population cycles of a forest insect: an experimental test. *Science* **285**:1068–1071.
- U.S. Fish and Wildlife Service. 1990. Endangered and threatened wildlife and plants: determine of threatened status for the Northern Spotted Owl. *Federal Register* **55**:26114–26196.
- Vandermeer, J. 1978. Choosing category size in a stage projection matrix. *Oecologia* **32**:79–84.
- Ver Hoef, J. M. 1996. Parametric empirical Bayes methods for ecological applications. *Ecological Applications* **6**:1047–1055.
- Wikle, C. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* **84**:1382–1394.
- Wyckoff, P. H., and J. S. Clark. 2002. Growth and mortality for seven co-occurring tree species in the southern Appalachian Mountains: implications for future forest composition. *Journal of Ecology* **90**:604–615.
- Zeger, S. L., and M. R. Karim. 1991. Generalized linear models with random effects: a Gibbs sampling approach. *Journal of the American Statistical Association* **86**:79–86.

APPENDIX

A Gibbs sampler for the hierarchical model is available in ESA's Electronic Data Archive: *Ecological Archives* E084-031-A1.