Ecological Monographs, 79(4), 2009, pp. 575–594

Integral projection models for populations in temporally varying environments

Mark Rees[1] and Stephen P. Ellner[2,3]

[1]*Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN United Kingdom*
[2]*Department of Ecology and Evolutionary Biology, E145 Corson Hall, Cornell University, Ithaca, New York 24853-2701 USA*

Abstract. Most plant and animal populations have substantial interannual variability in survival, growth rate, and fecundity. They also exhibit substantial variability among individuals in traits such as size, age, condition, and disease status that have large impacts on individual fates and consequently on the future of the population. We present here methods for constructing and analyzing a stochastic integral projection model (IPM) incorporating both of these forms of variability, illustrated through a case study of the monocarpic thistle *Carlina vulgaris*. We show how model construction can exploit the close correspondence between stochastic IPMs and statistical analysis of trait–fate relationships in a "mixed" or "hierarchical" models framework. This correspondence means that IPMs can be parameterized straightforwardly from data using established statistical techniques and software (vs. the largely ad hoc methods for stochastic matrix models), properly accounting for sampling error and between-year sample size variation and with vastly fewer parameters than a conventional stochastic matrix model. We show that the many tools available for analyzing stochastic matrix models (such as stochastic growth rate, λ_S, small variance approximations, elasticity/sensitivity analysis, and life table response experiment [LTRE] analysis) can be used for IPMs, and we give computational formulas for elasticity/sensitivity analyses. We develop evolutionary analyses based on the connection between growth rate sensitivity and selection gradients and present a new method using techniques from functional data analysis to study the evolution of function-valued traits such as size-dependent flowering probability. For *Carlina* we found consistent selection against variability in both state-specific transition rates and the fitted functions describing state dependence in demographic rates. For most of the regression parameters defining the IPM there was also selection against temporal variance; however, in some cases the effects of nonlinear averaging were big enough to favor increased temporal variation. The LTRE analysis identified year-to-year variation in survival as the dominant factor in population growth variability. Evolutionary analysis of flowering strategy showed that the entire functional relationship between plant size and flowering probability is at or near an evolutionarily stable strategy (ESS) shaped by the size-specific trade-off between the benefit (fecundity) and cost (mortality) of flowering in a temporally varying environment.

Key words: Carlina vulgaris; *evolutionarily stable strategies; function-valued traits; integral projection model; life table response experiment; mixed models; population forecasting; random environment; sensitivity analysis; stochastic growth rate; structured population; temporal variability*.

Introduction

The integral projection model (IPM) is a discrete-time structured population model that allows population structure to be defined by a mix of continuous and discrete individual attributes such as size, age, disease state, or any other attributes that affect demographic rates (Easterling et al. 2000, Ellner and Rees 2006). Integral projection models retain the most useful properties of matrix projection models (Ellner and Rees 2006) while eliminating the sometimes undesirable feature that all structuring variables must be broken up into discrete categories, which are often arbitrary for continuous attributes such as size. The IPM framework is also particularly useful when demographic performance is affected by multiple attributes, because the continuous-response structure of the model facilitates parsimonious modeling of trait–fate relationships. For example, demographic analyses of some thistles have found that although flowering probability depends on both size and age, it can be described by a logistic regression model with only three fitted parameters (Childs et al. 2003). Extending a size-structured IPM to include age-dependent flowering then requires the estimation of one extra parameter, rather than the many age- and size-class-specific flowering probabilities that would appear in a conventional matrix projection model.

Most plant and animal populations show substantial interannual variability in survival, growth, and fecundi-

Manuscript received 8 August 2008; revised 18 March 2009; accepted 30 March 2009; final version received 29 April 2009. Corresponding Editor: B. E. Kendall.
[3] Corresponding author. E-mail spe2@cornell.edu

ty, often spanning many orders of magnitude (Hairston et al. 1996, Metcalf et al. 2003, Rees et al. 2006). Stochastic matrix models are probably the most widely used tool for modeling and predicting the population fluctuations that result from variability in demographic rates and the associated risk of population extinction and responses to interventions (Caswell 2001, Morris and Doak 2002). However, stochastic IPMs can be constructed and parameterized from the same type of field data that would be required for a stochastic matrix model (e.g., Childs et al. 2004, Rees et al. 2006), and these models retain the basic mathematical properties of stochastic matrix models such as the existence of a long-term stochastic growth rate, λ_S (Ellner and Rees 2007). In addition, these models provide a framework to study selection in stochastic environments by extending analytic methods originally developed for matrix models, rather than having to resort to individual-based simulations of competition and natural selection.

Our goal in this paper is to outline a set of general tools that make stochastic IPMs a practical alternative to stochastic matrix models, and in many cases a preferable alternative, for demographic analyses of structured populations in temporally varying environments. Our previous paper (Ellner and Rees 2007), which derived the basic properties of stochastic IPMs, corresponds to Chapter 4 in Tuljapurkar (1990) or equivalently the first three sections of Chapter 14 in Caswell (2001). This paper aims to provide the rest: how the standard ways of getting information from a stochastic matrix model, sensitivity/elasticity analysis, small variance approximations, life table response experiments (LTRE), evolutionary demography, and so on, can all be done for stochastic IPMs. This might seem to require a book rather than a paper. Fortunately, existing books cover most of it (Tuljapurkar 1990, Caswell 2001) because the computational analysis and interpretation of IPMs and matrix models are very similar. Indeed, the same computer code can often be used when an IPM is implemented using the methods in Appendix A. However, the underlying theory is not just a matter of approximating IPMs by a large matrix and invoking matrix model theory. The key perturbation formulas (Eqs. 10 and 12 below) and their applications are derived here for stochastic IPMs, and approximation by matrices is only a technique for numerical calculations.

We present the ideas through a tutorial-style example using a field study of the monocarpic thistle *Carlina vulgaris*, although the ideas and methods are applicable to a wide range of plant and animal life cycles, including for example: (1) dormant and active life stages, (2) cross-classification by several attributes, such as size and a measure of condition or quality, and (3) changes between discrete and continuous structure over the life cycle. The *Carlina* example illustrates the close correspondence between stochastic IPMs and regression models with random effects, usually called "mixed models" by frequentists and "hierarchical models" by Bayesians (e.g., Clark et al. 2007). This correspondence, first exploited by Williams and Crone (2006), means that stochastic IPMs can be parameterized from data using well-grounded statistical techniques and widely available software, in contrast to the largely ad hoc methods that are available for stochastic matrix models. Another important benefit is a drastic reduction in the number of parameters that must be estimated from data, such as eliminating the many within- and between-year pairwise correlations among matrix entries that must be tested for significance and then estimated.

The smooth trait–fate models that are used to build an IPM could be used instead to parameterize a conventional matrix model, as some practitioners have suggested (Gross et al. 2006), in order to retain the familiar "look and feel" of matrix models. But the costs for sticking with the familiar are the conceptual and numerical inaccuracies that result from using discrete classes to model continuous variation. The theory we present here shows that these inaccuracies can be eliminated (in an appropriately fitted model) just by changing one line of your computer code: the one that specifies the size of the matrix.

Having built your model, what can you do with it? Continuing the *Carlina* example we show how prospective demographic analysis (i.e., sensitivity/elasticity analysis) can be applied to the kernel, to the regression parameters and functions defining the kernel in a stochastic IPM, and perform a retrospective analysis (i.e., LTRE) exploiting the mixed-effects modeling framework to test the assumptions of the standard first-order approximation. For cases in which the first-order analysis is found wanting, we present a more general analysis based on generalized additive models. We then develop evolutionary demographic analyses based on the connection between growth rate sensitivity and selection gradients and develop a new method using techniques from functional data analysis to study the evolution of function-valued traits such as size-dependent flowering probability in a temporally varying environment.

In terms of theory, our aims here are to summarize the basic properties and present formulas for computing the main quantities of interest and to facilitate application of IPMs based on ecological field data. The three application areas that we highlight (prospective sensitivity analysis, retrospective sensitivity analysis, and evolutionary demography) are largely independent of one another, but they all depend on the basic theory in *Basic properties: stochastic growth rates*. . . .

Building a Stochastic Integral Projection Model for *Carlina vulgaris*

In this section we illustrate the manner in which a stochastic IPM can be built from field data, paying particular attention to the statistical models for varia-

tion in demographic rates and how these are used to construct the IPM.

Field study

Carlina vulgaris, a monocarpic thistle of base-rich soils (mainly on limestone or calcareous sand), is native in Western, Central and Eastern Europe and has been introduced to North America and New Zealand. Under very favorable conditions individuals can flower in their second year (Klinkhamer et al. 1991, 1996, Rees et al. 2000) but more commonly reproduction is delayed until the third year or later. In the United Kingdom, the probability of flowering is related to both plant size and age (Rose et al. 2002), but for the sake of presentation here we ignore age dependence; this has little effect on model predictions. Flowering occurs between June and August. Seeds are retained in the flower heads until they are dispersed during dry sunny days in late autumn, winter, or spring (P. J. Grubb, *unpublished data*). Seeds germinate from April to June, and there is little evidence of a persistent seed bank (Eriksson and Eriksson 1997, de Jong et al. 2000).

The study, described in detail by Rose et al. (2002), spanned 16 years and followed the fates of over 1400 individuals. The ln-transformed length of the longest leaf was used as the measure of plant size. Rose et al. (2002) found that local conspecific density had no statistically significant effect on mortality and explained $<1\%$ of the variance in growth. We therefore omit from our models any effects of local density, though this could be accommodated in principle using a spatially structured IPM (E. Jongejans, K. Shea, and S. P. Ellner, *unpublished manuscript*).

Model structure

The state of a size-structured population is described by $n(x, t)$, the distribution of size x at time t, meaning that $\int_a^b n(x, t)\,dx$ is the number of individuals at time t whose size is between a and b. The general structure of a time-varying size-structured IPM is

$$n(y, t+1) = \int_L^U K(y, x; \boldsymbol{\theta}(t)) n(x, t)\,dx = \int_L^U [P(y, x; \boldsymbol{\theta}(t)) + F(y, x; \boldsymbol{\theta}(t))] n(x, t)\,dx. \quad (1)$$

Here $P(y, x; \boldsymbol{\theta})$ is the survival–growth kernel, describing individuals that are size x at time t and survive to reach size y at time $t + 1$. $F(y, x; \boldsymbol{\theta})$ is the fecundity kernel, describing the production of size y offspring by size x parents. $\boldsymbol{\theta}(t)$ is a vector of year-specific parameters, assumed to vary randomly over time, and L and U are lower and upper bounds on the range of possible sizes x. In this paper we model the kernel's time dependence as resulting from environmentally driven variation in parameters defining the kernel, such as the intercept and slope parameters of a logistic regression for the effect of size on survival.

The structure of the kernels reflects the life cycle of the species. *Carlina* individuals were censused in August. Flowering is fatal, and it is believed that mortality unrelated to flowering acts primarily over winter. We therefore assume that mortality unrelated to flowering acts before growth and flowering, so the survival–growth kernel can be written as

$$P(y, x; \boldsymbol{\theta}) = p_s(x; \boldsymbol{\theta})[1 - p_f(x; \boldsymbol{\theta})] g(y, x; \boldsymbol{\theta}) \quad (2)$$

where $p_s(x; \boldsymbol{\theta})$ and $p_f(x; \boldsymbol{\theta})$ are the probabilities of survival and of flowering conditional on survival, respectively, and $g(y, x; \boldsymbol{\theta})$ is the conditional probability density of size y next year, given current size x. The fecundity kernel $F(y, x; \boldsymbol{\theta})$ is given by

$$F(y, x; \boldsymbol{\theta}) = p_s(x; \boldsymbol{\theta}) p_f(x; \boldsymbol{\theta}) f_n(x; \boldsymbol{\theta}) f_d(y; \boldsymbol{\theta}) p_e \quad (3)$$

where $f_n(x; \boldsymbol{\theta})$ is the number of seeds produced by a size x plant, $f_d(y; \boldsymbol{\theta})$ is the probability distribution of recruit size, and p_e is the probability a recruit establishes.

Modeling temporal variation: fixed-effects models

Temporal variation in demographic rates can be modeled statistically as either fixed or random effects (Pinheiro and Bates 2000). "Fixed effects" means estimating a separate set of demographic parameters for each year when data on current state and future fate were collected. A kernel is then constructed for each year, and the population is projected forward by randomly selecting from the set of fitted year-specific kernels (Childs et al. 2004). This corresponds to the "matrix selection" procedure for constructing a stochastic matrix model.

Constructing a fixed-effects IPM for *Carlina* is largely an exercise in conventional linear and logistic regression, with "year" as an unordered categorical covariate. Rose et al. (2002) fitted models of this sort for *Carlina* with both size and age dependence. Through similar steps we obtained the size-dependent models for survival, growth, flowering probability, and the distribution of recruit sizes summarized in Table 1. Because "plants sit still and wait to be counted" we assume that measurement errors in plant size and fate are negligible. Other situations may require more complex models allowing random errors in census data or incomplete information (e.g., Metcalf et al. 2009), but the overall modeling strategy would be the same.

No data are available on size dependence of seed production, so following Rose et al. (2002) we assume, based on studies of similar species (Metcalf et al. 2006), that seed production is proportional to the square of untransformed linear size, giving $f_n(x, \boldsymbol{\theta}) = \exp(A + 2x)$ for some constant A. Because of the way that seedling establishment probability is estimated, A can be set arbitrarily and assumed not to vary over time, so we set $A = 1$.

TABLE 1. Structure of the fixed-effects demographic models for *Carlina vulgaris*.

Demographic process	Model
Size dynamics	
Rosette growth	$y = \mathcal{N}(a_g(t) + b_g(t)x, \sigma_\varepsilon)$
Recruit size	$R = \mathcal{N}(a_R(t), \sigma_\omega)$
Probability of survival	$\text{logit}(p_s(x)) = m_0(t) + m_s(t)x$
Probability of flowering	$\text{logit}(p_f(x)) = \beta_0(t) + \beta_s x$
Seed production	$f_n = \exp(A + Bx)$

Notes: In all cases x is size in year t (log-transformed) and y is the size the following year, R is recruit size, and $\mathcal{N}(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ. ε is growth variability, ω is recruit size variability, and m_0, m_s, a_g, b_g, β_0, β_s, A, B are fitted regression coefficients of the statistical models shown in this table. Parameters indicated as functions of t mean that a statistically significant effect of year was found, so the fitted model includes a different value of that parameter for each year of the study. A time-varying intercept parameter, such as $a_g(t)$, is present whenever the main effect of year was statistically significant, and a time-varying slope, such as $b_g(t)$, is present whenever the interaction between year and size was statistically significant. Models were fitted using the lm and glm functions in R.

Modeling temporal variation: fixed vs. random effects

The alternative to matrix selection for constructing a stochastic matrix model is element selection, in which matrix elements are drawn from a fitted multivariate statistical distribution (Morris and Doak 2002). The corresponding approach for IPMs is to fit mixed-effects demographic models, meaning that some parameters in the statistical models for state–fate relationships are modeled as random effects drawn each year from a fitted probability distribution. The models are summarized in Table 2; here we describe how the models were fitted.

The mixed-effects IPM approach has several important advantages over a fixed-effects IPM or conventional stochastic matrix model. Firstly, there is a great reduction in the number of parameters to estimate. The fixed-effects *Carlina* model in Table 1 requires 90 parameters to describe temporal variation (6 coefficients $\times$ 15 years), while mixed models for exactly the same processes require 15 parameters to specify the means, variances, and covariances among time-varying regression coefficients. Element selection for a conventional stochastic matrix model would also require estimating far more than 15 parameters, because of the need to specify at least the mean and variance for all potentially nonzero matrix entries and all of the pairwise covariances. Within-year covariances are common, in part because of inherent constraints (e.g., if individuals are likely to grow this year they must be unlikely to stay the same size or to shrink) and in part because a good year for some individuals is likely to be good for all (see Morris and Doak [2002:253–290] for a discussion of these issues). In a mixed-effects IPM these correlations result automatically from the model structure, without parameter covariance. For example, a higher intercept in the regression of future size on current size simultaneously increases the probability of growth and decreases the probability of shrinkage, for individuals of all sizes.

Secondly, parameters estimated from years when there are few data are given equal weight in simulating a fixed-effects or matrix selection model. Years are appropriately weighted to reflect unequal sample sizes when a mixed-effects model is fitted.

Finally, the between-year variability in fixed effects or matrix selection approaches is biased upwards, because some of the apparent interannual parameter variability is actually due to sampling variability (e.g., Kendall 1998). Current methods for fitting mixed-effects models account for sampling variability and estimate (for example) the actual variance in survival probability that underlies the observed year-to-year variation in survival of sampled individuals.

These advantages come at a cost: the need to assume a parametric distribution for year effects. Modern developments in mixed models make it possible to explore a range of different distributions and see how this influences the model's goodness of fit and predictions

TABLE 2. Estimated demographic functions in the mixed-models analysis for *Carlina vulgaris*.

Demographic process	Model	Parameter estimates
Size dynamics: rosette growth and recruit size	$y = a_g + b_g x + \varepsilon$ $R = a_R + \omega$	$b_g \sim \mathcal{N}(0.74, 0.13)$ $\varepsilon \sim \mathcal{N}(0, 0.29)$ $\omega \sim \mathcal{N}(0, 0.50)$ $a_g, a_R \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ $\boldsymbol{\mu} = (1.14, 3.16)$ $\boldsymbol{\Sigma} = \begin{pmatrix} 0.037 & 0.041 \\ 0.041 & 0.075 \end{pmatrix}$
Probability of survival	$\text{logit}(p_s(x)) = m_0 + m_s x$	$m_0 \sim \mathcal{N}(-2.28, 1.16)$ $m_s \sim \mathcal{N}(0.90, 0.41)$
Probability of flowering	$\text{logit}(p_f(x)) = \beta_0 + \beta_s x$	$\beta_0 \sim \mathcal{N}(-16.19, 1.03)$ $\beta_s = 3.88$
Seed production	$f_n = \exp(A + Bx)$	$A = 1$, $B = 2$

Notes: In all cases x is size in year t (log-transformed) and y is the size the following year, R is recruit size, $\mathcal{N}(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ, and MVN($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$) is a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and variance–covariance matrix $\boldsymbol{\Sigma}$.

or even to fit year distributions nonparametrically if sufficient data are available (Zhang and Davidian 2001, Ghidey et al. 2004). But if a study spans only a few years, it is hard to draw reliable conclusions about parameter distributions, and fixed-effects modeling might then be preferred, with some shrinkage of year-specific parameters to correct for sampling variability (see for example McCullough and Searle [2001: section 2.4b]). All of our theoretical results (such as the formulas in Table 3) also apply to fixed-effects stochastic IPMs in which future years are assumed to be a random draw from the set of years used to construct the model and also to completely deterministic IPMs if the assumptions in Ellner and Rees (2006) are satisfied.

Modeling temporal variation: fitting demographic mixed models

Our approach to fitting mixed models follows Pinheiro and Bates (2000). The fixed-effects analysis identifies which parameters should be modeled as being time-varying. We then used normal probability plots of the year-specific parameter estimates to confirm that it is reasonable to assume (as all our fitted mixed models do) that the parameter variability is Gaussian.

To determine which parameters should be modeled as covarying, we calculated the correlation coefficients between the different year-specific parameter estimates from the fixed-effects models. Very few significant correlations were found. The rosette growth intercept, a_g, and the mean new recruit size, a_R, were positively correlated ($r = 0.77$, $n = 15$, $P < 0.001$). The growth intercept was also highly negatively correlated with the growth slope, b_g ($r = -0.97$), because sizes are in the range 1.8–4.6, whereas the intercepts are extrapolated back to size 0. The correlation was reduced to -0.19 when current size values were centered to zero mean, suggesting that slope and intercept vary independently. This conclusion was confirmed by fitting a linear mixed model using the lme function in R (R Development Core Team 2007), which indicated that treating slope and intercept as random effects improved the fit of the model ($\chi^2_1 = 8.15$, $P < 0.004$), whereas allowing the slope and intercept to covary did not ($\chi^2_1 = 2.34$, $P > 0.1$). Even with centering, the slopes and intercepts for survival were correlated ($r = 0.67$), but this was due to a single year in which there were few individuals and only two died, resulting in the largest estimated intercept and slope, although neither was significantly different from zero. Removing that year yielded $r = -0.09$, suggesting that slope and intercept are approximately independent.

We consequently modeled growth and recruitment with a mixed model in which the growth intercept and the mean recruit size in the following year are drawn from a bivariate normal distribution and used WinBUGS (Lunn et al. 2000) to fit the model using a Bayesian approach. We used standard noninformative priors for the variances (inverse γ; Spiegelhalter et al. [2003]) and a uniform prior on [−1, 1] for the correlation coefficient; see Cam et al. (2002) for a more detailed description of the approach. The estimated correlation between the random effects was 0.77 with the interval [0.48, 0.95] containing 95% of the mass of the posterior distribution. The models for survival and flowering are standard generalized linear mixed models that we fitted using WinBUGS and glmmML (Bröström 2008) in R.

Modeling establishment probability

There was no relationship between estimated total seed production one year and the total number of recruits the following year, most likely because establishment is microsite-limited and dependent upon disturbance (Rose et al. 2002). The decoupling of recruitment from seed production implies a density-dependent model, in which higher seed production gives lower recruitment probability. Such a model is easily simulated, but general theory for density-dependent stochastic IPMs is lacking. We consider here both density-dependent and density-independent IPMs for *Carlina*. The resident model is density dependent and describes a numerically dominant population with microsite-limited recruitment. The resident's recruitment probability is given by the total number of recruits the following year divided by current seed production:

$$p_e(t) = \frac{R(t+1)}{\int_L^U p_s\Big(x;\boldsymbol{\theta}(t)\Big)p_f\Big(x;\boldsymbol{\theta}(t)\Big)f_n(x)n(x,t)\,dx}. \quad (4)$$

For evolutionary analyses, we use the invader model, which is density independent and describes a rare subpopulation differing from the resident in some way that does not affect how seedlings compete for microsites. The invader's seed production does not influence its probability of establishment because it is assumed to be rare, and so the invader's $p_e(t)$ is also given by Eq. 4. So for the invader, $p_e(t)$ is just one additional component of the vector $\boldsymbol{\theta}(t)$ describing year-to-year variability (Ellner and Rees 2007: Appendix A). We obtained $p_e(t)$ values by simulating the resident population (with total recruit numbers drawn at random from the 16 observed values), computing $p_e(t)$ using Eq. 4, and appending the resulting $p_e(t)$ values to the generated sequence of year-specific parameters $\boldsymbol{\theta}(t)$.

Putting the pieces together

The statistical models summarized in Table 2, combined with Eqs. 2, 3, and 4, completely specify the resident and invader IPMs for *Carlina vulgaris*. The key point of this section is that once the demography has been analyzed in a mixed-models framework, the stochastic IPM follows automatically.

Let l denote the inverse logit (i.e., logistic) function $l(u) = \exp(u)/(1 + \exp(u))$, and $\varphi(y \mid \mu, \sigma)$ the probability density for a normal distribution with mean μ and standard deviation σ. For simulating the resident model, the vector of year-specific parameters is drawn from the distributions specified in Table 2:

TABLE 3. Summary of formulas for stochastic sensitivities and elasticities of the long-run growth rate of the population, λ_S, with respect to kernel values and underlying parameters of the kernel.

Sensitivity measure	Notation and formula
Sensitivity of log λ_S to perturbing K_t by arbitrary kernel εC_t	$\frac{\partial \log \lambda_S}{\partial \varepsilon} = E\left[\frac{\langle v_{t+1}, C_t w_t\rangle}{\langle v_{t+1}, K_t w_t\rangle}\right]$
Sensitivity of log λ_S to perturbing K_t by independent kernel εH_t with zero mean	$\frac{\partial \log \lambda_S}{\partial \varepsilon^2} = -\frac{1}{2}\mathrm{Var}\left[\frac{\langle v_{t+1}, H_t w_t\rangle}{\langle v_{t+1}, K_t w_t\rangle}\right] = -\frac{1}{2}E\left[\frac{\langle v_{t+1}, H_t w_t\rangle^2}{\langle v_{t+1}, K_t w_t\rangle^2}\right]$
Sensitivity of λ_S to kernel value $K_t(y, x)$ = sensitivity of λ_S to mean of $K_t(y, x)$	$s_S(y, x) = s_S^{\mu}(y, x) = \lambda_S E\left[\frac{v_{t+1}(y) w_t(x)}{\langle v_{t+1}, K_t w_t\rangle}\right]$
Elasticity of λ_S to kernel value $K_t(y, x)$	$e_S(y, x) = E\left[\frac{v_{t+1}(y) w_t(x) K_t(y, x)}{\langle v_{t+1}, K_t w_t\rangle}\right]$
Elasticity of λ_S to the mean of kernel value $K_t(y, x)$	$e_S^{\mu}(y, x) = \bar{K}(y, x) E\left[\frac{v_{t+1}(y) w_t(x)}{\langle v_{t+1}, K_t w_t\rangle}\right]$
Elasticity to standard deviation of kernel value $K_t(y, x)$	$e_S^{\sigma}(y, x) = e_S(y, x) - e_S^{\mu}(y, x) = 0 \quad$ if $\mathrm{Var}[K_t(y, x)] = 0$
Sensitivity of λ_S to the standard deviation of time-varying kernel value $K_t(y, x)$	$s_S^{\sigma}(y, x) = \lambda_S e_S^{\sigma}(y, x) \Big/ \sqrt{\mathrm{Var}[K_t(y, x)]}$
Sensitivity of λ_S to the variance of time-varying kernel value $K_t(y, x)$	$s_S^{\sigma^2}(y, x) = 0.5 s_S^{\sigma}(y, x) \Big/ \sqrt{\mathrm{Var}[K_t(y, x)]}$
Sensitivity to adding independent variability to a time-invariant kernel value $K(y, x)$	$s_S^{\sigma^2,0}(y, x) = -\frac{\lambda_S}{2} E\left[\frac{[v_{i+1}(y) w_i(x)]^2}{\langle v_{i+1}, K_i w_i\rangle^2}\right]$
Sensitivity of λ_S to parameter θ_i = sensitivity of λ_S to mean of θ_i	$s_{S,i} = s_{S,i}^{\mu} = \lambda_S E\left[\left\langle v_{t+1}, \frac{\partial K_t}{\partial \theta_i} w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right]$
Elasticity of λ_S to parameter θ_i	$e_{S,i} = E\left[\theta_i(t)\left\langle v_{t+1}, \frac{\partial K_t}{\partial \theta_i} w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right]$
Elasticity of λ_S to the mean of θ_i	$e_{S,i}^{\mu} = \bar{\theta}_i E\left[\left\langle v_{t+1}, \frac{\partial K_t}{\partial \theta_i} w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right]$
Elasticity of λ_S to the standard deviation of θ_i	$e_{S,i}^{\sigma} = e_{S,i} - e_{S,i}^{\mu} = 0 \quad$ if $\mathrm{Var}(\theta_i) = 0$
Sensitivity of λ_S to the standard deviation of time-varying parameter $\theta_i(t)$	$s_{S,i}^{\sigma} = \lambda_S e_{S,i}^{\sigma} \Big/ \sqrt{\mathrm{Var}(\theta_i)}$
Sensitivity of λ_S to the variance of time-varying parameter $\theta_i(t)$	$s_{S,i}^{\sigma^2} = 0.5 s_{S,i}^{\sigma} \Big/ \sqrt{\mathrm{Var}(\theta_i)}$
Sensitivity of λ_S to added variance in time-invariant parameter θ_i	$s_{S,i}^{\sigma^2,0} = \frac{\lambda_S}{2}\left(E\left[\left\langle v_{t+1}, \frac{\partial^2 K_t}{\partial \theta_i^2} w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right] - E\left[\left\langle v_{t+1}, \frac{\partial K_t}{\partial \theta_i} w_t\right\rangle^2 \Big/ \langle v_{t+1}, K_t w_t\rangle^2\right]\right)$

Note: K_t is the base kernel, C_t and H_t are perturbation kernels, E denotes expectation, v and w are the left and right eigenfunctions of the unperturbed kernel, and θ_i is the ith component of $\boldsymbol{\theta}$.

$$\boldsymbol{\theta}(t) = [m_0(t),\, m_s(t),\, \beta_0(t),\, a_g(t),\, b_g(t),\, a_R(t)] \quad (5)$$

is drawn from the distributions specified in Table 2. Given these, the survival–growth kernel (Eq. 2) is

$$P(y,x;\boldsymbol{\theta}) = l(m_0 + m_s x)[1 - l(\beta_0 + \beta_s x)] \times \varphi(y \mid \mu = a_g + b_g x, \sigma = \sigma_\varepsilon) \quad (6)$$

and using Eq. 3 the fecundity kernel is

$$F(y,x;\boldsymbol{\theta}) = l(m_0 + m_s x) l(\beta_0 + \beta_s x) e^{A+Bx} \times \varphi(y \mid \mu = a_R, \sigma = \sigma_\omega) p_e(t). \quad (7)$$

The resulting kernel $K(y, x, \boldsymbol{\theta}(t))$ is used to iterate the population from its current size and structure to the following year. Numerical iteration of the model is straightforward using standard numerical methods (Ellner and Rees 2006), but for the sake of completeness we summarize suggested methods in Appendix A. All results presented here are based on using 50 or more mesh points for evaluation of integrals with respect to size and 10 000 iterations for computing time averages, discarding the first 500 iterations.

Basic Properties: Stochastic Growth Rates and Small Variance Approximations

We now aim to convince you that just about any analysis that could be done for a density-independent, stochastic matrix model can be done for a density-independent, stochastic IPM such as the *Carlina* invader IPM. The technical assumptions underpinning our results for stochastic IPMs are described in Appendix B; see also Ellner and Rees (2006, 2007). The main practical consequences of the technical assumptions are: (1) that the kernel $K(y, x, \boldsymbol{\theta})$ must be a continuous function of $(y, x, \boldsymbol{\theta})$ and (2) that the distributions of all model parameters and individual-level state variables must be bounded. Using an unbounded statistical distribution or state variable to model a finite data set is harmless for short-term forecasting but can cause unrealistic long-term behavior (Ellner and Rees 2006). The fitted mixed-effects demographic models therefore must be modified in two ways: (1) by limiting the possible range of parameter variation, e.g., by truncation at the upper and lower 99.9% percentiles of the fitted distribution and (2) by limiting the possible range of size distributions for survivors and offspring, by truncation of those distributions at values far outside the range of the empirically observed values.

We begin by reviewing the basic population dynamic properties of stochastic IPMs derived in our previous paper (Ellner and Rees 2007). We then develop a small variance approximation for stochastic growth rate, present new methods and formulas for prospective (elasticity/sensitivity) and retrospective (life table response experiment) demographic analyses, and develop some new techniques useful for evolutionary demography analyses using IPMS. We illustrate all the new methods using the *Carlina* IPM.

Stochastic growth rate

The most basic property of a stochastic IPM is the long-run growth rate of the population, usually denoted $\log \lambda_S$ (Caswell 2001, Tuljapurkar and Haridas 2006), where λ_S corresponds to the dominant eigenvalue of a time-invariant projection matrix or kernel. Under the assumptions listed in Appendix B, Ellner and Rees (2007) prove that: (1) The logarithm of total population size $N(t)$, or of any part of the population, has a long-term growth rate $\log \lambda_S$, which is constant with probability 1, i.e., $[\log N(t)]/t \to \log \lambda_S$. (2) The initial population structure has no effect on the long-term behavior. (3) The long-term growth rate equals the average one-step growth rate, $\log \lambda_S = E \log(N(t+1)/N(t))$. If future states of the environment process become unpredictable quickly enough, the long-term distribution of total population size is lognormal (Ellner and Rees 2007). Any irreducible aperiodic finite-state Markov Chain forgets its past fast enough for this property, but some environmental variables have a longer-lasting "memory" of the past (Halley and Inchausti 2004), and population fluctuations driven by such variables would not necessarily tend to a lognormal distribution.

Small variance approximation

Let $v(x)$, $w(x)$ denote the dominant left and right eigenfunctions of the mean kernel (the reproductive value and stable state distribution, respectively). We scale these so that $\langle v, w\rangle = 1$, where angle brackets denote the inner product $\langle f, g\rangle = \int_{\mathbf{x}} f(x)g(x)\,dx$. The mean kernel $\bar{K}(y,x) = E[K(y,x;\boldsymbol{\theta})]$ is assumed to satisfy the assumptions of Ellner and Rees (2006). The dominant eigenvalue for $\bar{K}$ will be denoted λ_1, following Caswell (2001).

In Appendix C we show (following Tuljapurkar [1990] and Appendix B of Tuljapurkar and Haridas [2006]) that if year-to-year variability is small then

$$\log \lambda_S \approx \log \lambda_1 - \frac{\mathrm{Var}\langle v, K_t w\rangle}{2\lambda_1^2} + \sum_{j=1}^{\infty} c_j \quad (8)$$

with the three terms on the right-hand side of Eq. 8 corresponding to the mean kernel, the kernel variance, and effects of temporal autocorrelations at time-lags $j \geq 1$. Here K_t is shorthand for $K(y, x; \boldsymbol{\theta}(t))$, Var is short for "variance," $K_t w = \int_{\mathbf{x}} K_t(y, x) w(x)\,dx$, and the serial correlation effects c_j are defined in Appendix C.

Eq. 8 is the generalization to IPMs of Eq. 12.1.16 in Tuljapurkar (1990). The variance term can be written as

$$-\frac{1}{2\lambda_1^2} \iiiint s(y_2,x_2)s(y_1,x_1) \times \mathrm{Cov}[K_t(y_2,x_2), K_t(y_1,x_1)]\; dx_1\, dx_2\, dy_1\, dy_2 \quad (9)$$

analogous to Eq. 14.71 in Caswell (2001), where s is the sensitivity function for the mean kernel, $s(y, x) = v(y)w(x)/\langle v, w\rangle$. We have not found a pretty expression for the autocorrelation terms c_j. However, the effect of environmental autocorrelation is limited by the damping ratio ρ of the mean kernel, as in matrix models (Tuljapurkar and Haridas 2006). Specifically, $|c_j| \leq C\rho^j$ for a constant C, depending on the mean kernel (see Appendix C: Eq. C.6). So if the mean model converges rapidly onto its stable distribution w, the effects of environmental autocorrelation will be weak.

Prospective Demographic Analyses

Methods: sensitivity and elasticity general theory

Perturbation analysis of a stochastic IPM can be carried out at three levels: at the level of kernel entries $K(y, x)$, at the level of parameters in the demographic functions making up the kernel (e.g., perturbing the mean slope in a regression model), and at the level of demographic functions (e.g., perturbing survival at a specific size or at all sizes simultaneously). The first of these is identical in practice to sensitivity analysis of a matrix model with respect to perturbations of individual matrix entries, because it can be done by using a large matrix to approximate the IPM and then applying matrix model methods. The second level is analogous to perturbation analysis of matrix models with respect to "lower-level parameters" that affect several matrix entries simultaneously (Caswell 2001: sections 9.2.3 and 10.4.3; Haridas and Tuljapurkar 2005). The third level, however, reflects how an IPM kernel is constructed from the demographic models that use continuous individual-level state variables to predict individual fates. At each of these levels, perturbation analyses may be performed with respect to changes in the mean or standard deviation/variance. Elasticities with respect to standard deviation are necessarily zero whenever there is no temporal variation. So to allow comparisons between constant and time-varying elements, we develop sensitivities to variance, which are nonzero for both constant and time-varying parameters or kernel elements.

All of these analyses are based on two general perturbation formulas that involve the time-varying population structure w_t and reproductive value v_t. These are computed as follows:

$$\tilde{w}_{t+1} = K_t w_t, \qquad w_{t+1} = \tilde{w}_{t+1} \Big/ \int \tilde{w}_{t+1}(x)\, dx$$

and

$$\tilde{v}_{t-1} = v_t K_{t-1} = \int v_t(y) K_{t-1}(y, x)\, dy,$$

$$v_{t-1} = \tilde{v}_{t-1} \Big/ \int v_{t-1}(x)\, dx.$$

Under our assumptions the population structure and reproductive value converge to stationary processes that are independent of their initial values (Ellner and Rees 2007: section 4).

The first formula concerns perturbing from $K_t(y, x)$ to $K_t(y, x) + \varepsilon C_t(y, x)$ where ε is a small constant and the perturbation kernels C_t preserve the model assumptions stated above for small ε and otherwise are completely arbitrary (for example, they could be time-varying or constant, zero or nonzero mean, independent of K or correlated with it, etc.). The sensitivity of λ_S to the perturbation is defined as

$$\left.\frac{\partial \lambda_S(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0}$$

where $\lambda_S(\varepsilon)$ is the stochastic growth rate of the perturbed model. For stochastic IPMs there is no mathematical proof that this derivative exists, but the formulas below apply whenever $\lambda_S(\varepsilon)$ is a smooth function of ε.

Following Tuljapurkar (1990:89) or equivalently section 14.4.1 in Caswell (2001), we have

$$\log \lambda_S(\varepsilon) = \log \lambda_S(0) + \varepsilon E\left[\frac{\langle v_{t+1}, C_t w_t\rangle}{\langle v_{t+1}, K_t w_t\rangle}\right] + O(\varepsilon^2). \quad (10)$$

where E denotes expectation and v_t and w_t are the stationary reproductive value and population structure sequences. The expectation on the right-hand side of Eq. 10 is therefore the sensitivity of $\log \lambda_S$:

$$\frac{\partial \log \lambda_S}{\partial \varepsilon} = \frac{1}{\lambda_S}\frac{\partial \lambda_S}{\partial \varepsilon} = E\left[\frac{\langle v_{t+1}, C_t w_t\rangle}{\langle v_{t+1}, K_t w_t\rangle}\right]. \quad (11)$$

Eqs. 10 and 11 are used in practice as follows. A long sequence of population structures $\{w_t\}_{t=0}^{T}$ and reproductive values $\{v_t\}_{t=0}^{T}$ for the unperturbed model is computed once, with arbitrary choices of w_0 and v_T. These sequences can then be used to evaluate the sensitivity for any perturbation, by using a time average from $t = k$ to $t = T - k$ (for some $1 \ll k \ll T$) to approximate the expectation.

The second perturbation formula concerns perturbing $K_t(y, x)$ to $K_t(y, x) + \varepsilon H_t(y, x)$, where the perturbation kernels H_t have mean zero and are independent of one another and of the "base" kernels K_t. The most important examples arise when considering the effect of variability on time-invariant parts of the kernel or time-invariant parameters. Eq. 10 applies to any perturbation, but when the perturbation kernel is zero-mean and independent of everything else in the model (as we assume about H), the sensitivity Eq. 11 is always zero. So to compare impacts of adding variability to different time-invariant parts of the kernel, or to different time-invariant parameters, we need to compute the order ε^2 term in Eq. 10. The result (derived in Appendix D) is as follows:

$$\log \lambda_S(\varepsilon) = \log \lambda_S(0) - \frac{\varepsilon^2}{2}\mathrm{Var}\left[\frac{\langle v_{t+1}, H_t w_t\rangle}{\langle v_{t+1}, K_t w_t\rangle}\right] + O(\varepsilon^3)$$
$$= \log \lambda_S(0) - \frac{\varepsilon^2}{2}E\left[\frac{\langle v_{t+1}, H_t w_t\rangle^2}{\langle v_{t+1}, K_t w_t\rangle^2}\right] + O(\varepsilon^3). \quad (12)$$

This is (not coincidentally) very similar to the small variance approximation (Eq. 8).

Sensitivity and elasticity to kernel values $K_t(y, x)$

Using Eqs. 10 and 12 it is straightforward to derive the IPM analogs of the various sensitivities and elasticities that have been proposed for stochastic matrix models (Tuljapurkar 1990, Tuljapurkar et al. 2003) or any others you want to consider. The results are collected in Table 3, and derivations are in Appendix E. All of the expressions in Table 3 can be evaluated using one long simulation of the model to approximate the expectations by time averages.

An important complication is the need for separate treatment of time-invariant kernel entries (i.e., a transition (y, x) where $K_t(y, x) \equiv \bar{K}(y, x)$), when considering effects of increased variability. For time-varying kernel entries, the sensitivity and elasticity of λ_S to increased variability are computed from Eq. 10; in our *Carlina* model all kernel entries are time varying. But when some kernel entries are time-invariant, the relevant measure for comparing these to one another is the sensitivity of λ_S to increasing the variance from 0, computed using Eq. 12, which we denote $s_S^{\sigma^2,0}(y, x)$. Time-varying entries can be compared with time-invariant entries by computing the sensitivity of λ_S to the variance of time-varying entries, $s_S^{\sigma^2}(y, x)$, using the formula in Table 3.

Most properties of the stochastic elasticities of matrix models carry over to kernel value elasticities of stochastic IPMs. The stochastic elasticity function $e_S(y, x)$ and elasticity to the mean $e_S^{\mu}(y, x)$ are both nonnegative everywhere. If the overall variability of the kernel is small, Eq. 9 shows that (in the absence of temporal correlations) any increase in variability causes λ_S to decrease, so $e_S^{\sigma}(y, x)$ is negative for all time-varying kernel entries and $e_S^{\sigma^2,0}(y, x)$ is negative for all time-invariant kernel entries, and the response to a unit increase in variability is largest for entries at which the sensitivity function of the mean kernel is large. The double integral of $e_S(y, x)$ over all size-transitions equals 1 (for the simple reason that a 5% increase in the entire kernel every year leads to a 5% increase in λ_S). By definition $e_S^{\mu}(y, x) + e_S^{\sigma}(y, x) = e_S(y, x)$. Define $e_S(A)$ to be the integral of $e_S(y, x)$ over the set A and analogously for other elasticities. Then $0 < e_S^{\mu}(A) + e_S^{\sigma}(A) = e_S(A) \leq 1$, implying that $-e_S^{\sigma}(A) < e_S^{\mu}(A) \leq (1 - e_S^{\sigma}(A))$.

It is difficult to provide general advice on how long a simulation is needed for the expectations in the perturbation formulas in Table 3 to be approximated well by time averages. In practice it should be sufficient to divide the simulation into two halves and compute any quantity of interest twice. A substantial difference between these would imply that a longer simulation is needed.

Sensitivity and elasticity to underlying parameters

For a parametric kernel $K_t(y, x) = K(y, x; \boldsymbol{\theta}(t))$, we can also examine the effects of perturbing underlying parameters, e.g., perturbing the slope coefficient in a regression model describing size-dependent growth or survival. As with kernel values, time-varying and constant parameters need to be considered separately (see Table 3). Because the elasticity and sensitivity of λ_S to the standard deviation of time-invariant parameters is necessarily zero, it becomes informative to compute the sensitivities with respect to variance ($s_{S,i}^{\sigma^2}$ and $s_{S,i}^{\sigma^2,0}$ for time-varying and time-invariant parameters, respectively).

Elasticities with respect to underlying parameters are less clearly interpretable than the corresponding sensitivities, because parameters may be negative. For example, a linear regression model for changes in log-transformed size may have a negative intercept, meaning that an individual of size 1 (in the model's units) will shrink on average. A proportional change in a negative parameter makes it go further negative, so the elasticity and sensitivity have opposite signs. Stochastic sensitivities play an important role in evolutionary demography because they can be interpreted as a fitness gradient for evolutionary stability analysis, i.e., for predicting whether a homogeneous population will be open to invasion by a rare alternative type (though not for analysis of short-term selection response in a genetically heterogeneous population, as Lande [2007] has recently pointed out). So for negative underlying parameters, the sensitivity rather than the elasticity will correctly indicate the predicted direction of long-term adaptive dynamics. This issue doesn't arise in conventional matrix models because the underlying parameters are intrinsically positive, either matrix entries themselves or underlying probabilities of events such as survival, growth, or breeding.

Sensitivity and elasticity to kernel functions

Sensitivities and elasticities to the demographic functions that define the kernel can be computed by deriving the perturbation kernel and applying either Eq. 10 or Eq. 12, as appropriate. The specifics of this process depend upon how the kernel is built up from demographic rate functions and upon which function is perturbed. In most cases Eq. 10 will be appropriate for perturbations to an entire function such as $p_s(x)$, because the perturbation is unlikely to be independent of the unperturbed kernel. However, for adding variability to a constant rate function or to a constant part of a rate function, the sensitivity to the standard deviation will be zero and then Eq. 12 should be used to calculate the sensitivity to the variance. In Appendix F we work through two examples: the elasticity to the size-

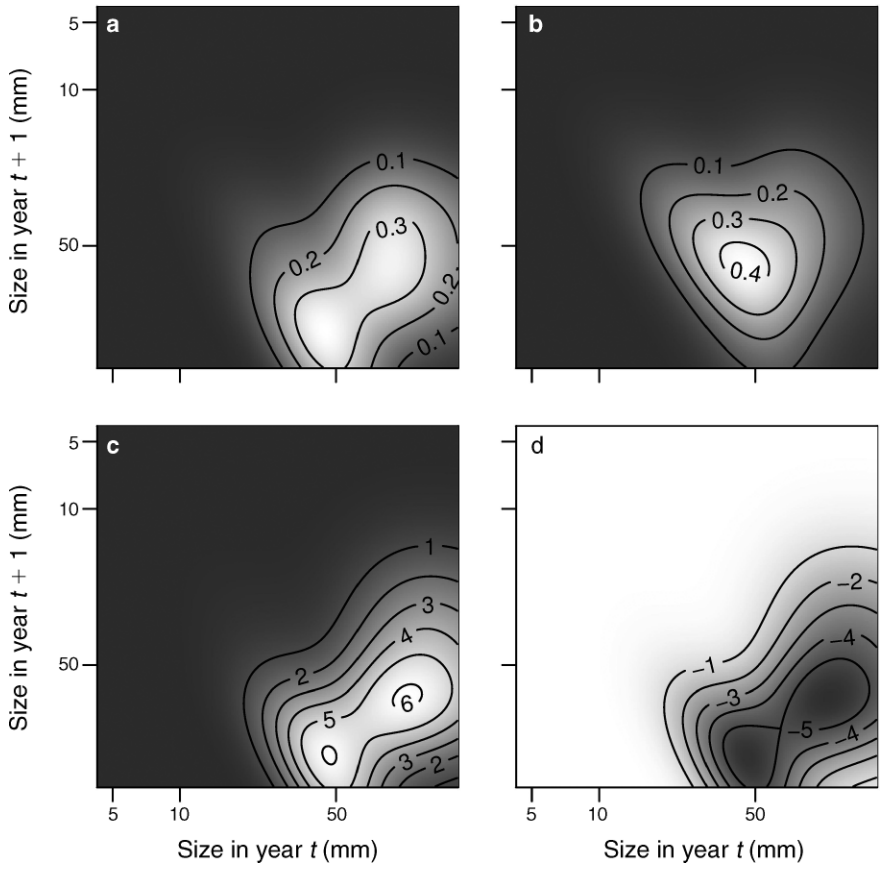

FIG. 1. Shaded contour plots of elasticity and sensitivity functions (see Table 3). The values along the contours are (a) the elasticities of the mean kernel $e(y, x)$, (b) stochastic elasticities to kernel values $e_S(y, x)$, (c) stochastic elasticities to the mean kernel value $e_S^{\mu}(y, x)$, and (d) stochastic elasticities to the standard deviation of kernel values $e_S^{\sigma}(y, x)$.

dependent survival function $p_s(x)$ or survival at a particular size $p_s(x_0)$ in a simple model for a size-structured iteroparous species and the elasticity to flowering probability in our *Carlina* model.

Results: prospective analysis for the density-independent Carlina *invader model*

The elasticities to kernel values for the mean kernel $e(y, x)$, the stochastic elasticities $e_S(y, x)$, the stochastic elasticities to the mean $e_S^{\mu}(y, x)$, and to the standard deviation $e_S^{\sigma}(y, x)$ all indicate the importance of transitions into the size range at which individuals have a high probability of flowering (Fig. 1): flowering probability increases with size, and individuals of size 50 mm have ~50% chance of flowering. The various elasticities to kernel values are correlated (Table 4), so knowing the elasticity function $e(y, x)$ of the mean kernel provides a great deal of information on the various stochastic elasticities. The stochastic elasticities to the mean and standard deviation are almost perfectly negatively correlated. Transitions where λ_S is highly sensitive to a change in the mean also have a large response to changes in standard deviation, but the former is positive (a higher mean always helps) while the latter is negative (variance hurts).

The sensitivities to the parameters defining the kernel reveal that all the parameter means are under positive selection (Table 5). For the parameters defining the survival (m_0, m_s), growth (a_g, b_g), and fecundity (A, B) relationships this is not surprising. What is perhaps more surprising is that there is also selection for increased variance in growth (σ_ε^2) and in the size distribution of recruits (σ_ω^2). This is a consequence of working on a log scale for size; because actual size is

TABLE 4. Correlation matrix for the elasticities of the mean kernel value $e(y, x)$, the stochastic elasticities $e_S(y, x)$, the stochastic elasticities to the mean $e_S^{\mu}(y, x)$, and the stochastic elasticities to the standard deviation $e_S^{\sigma}(y, x)$.

	e_S	e_S^{μ}	e_S^{σ}
e	0.75	0.91	−0.89
e_S		0.50	−0.49
e_S^{μ}			−0.99

Note: To compute the values we evaluated each elasticity function at a 100 × 100 grid of (y, x) values, converted these matrices to vectors, and computed the pairwise correlation coefficients for those vectors.

TABLE 5. Stochastic sensitivities for the parameters defining the kernel.

Parameter	$s^{\mu}_{S,i}$	$s^{\sigma^2}_{S,i}$
m_0	0.32	−0.06
m_s	1.23	−0.99
β_0	0.07	−0.04
β_s	0.29	**−0.18**
a_g	1.03	−0.55
b_g	3.79	−1.83
σ^2_ε	1.32	**−0.10**
A	0.44	**0.08**
B	1.94	**1.46**
a_R	0.99	−0.12
σ^2_ω	1.34	**−0.04**

Notes: See Table 2 for descriptions of the parameters and Table 3 for definitions of the sensitivities; σ^2_ε is the variance of growth increments, σ^2_ω is the variance of recruit size, and a_R is mean recruit size. The $s^{\mu}_{S,i}$ are the stochastic sensitivities to the mean, and $s^{\sigma^2}_{S,i}$ are the sensitivities to the variance calculated for both time-varying parameters and time-invariant parameters (boldface).

$\exp(x)$, variance in x increases the average plant or recruit size on an arithmetic scale (due to Jensen's inequality). In all cases there is much stronger selection on the slopes compared to the intercepts (ratios in the range 3.7–4.4). The weakest selection is on β_0, which suggests that β_0 is close to an evolutionarily stable strategy (ESS). There is selection for an increase in the slope of the flowering function, β_s. As β_s increases, the relationship between the probability of flowering and plant size approaches a step function, with no flowering below a size threshold and 100% flowering above the threshold.

The stochastic sensitivities to parameter variance (Table 5) suggest there is selection against temporal variation in most parameters apart from those defining the seed production function (A and B). To clarify the factors influencing the sensitivity to parameter variance we can express $s^{\sigma^2,0}_{S,i}$ as

$$s^{\sigma^2,0}_{S,i} = \frac{\lambda_S}{2}\left\{E\left[\left\langle v_{t+1}, \frac{\partial^2 K_t}{\partial\theta_i^2} w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right]\right.$$

(nonlinear averaging),

$$-\mathrm{Var}\left[\frac{\partial \log \lambda(t)}{\partial\theta_i(t)}\right]$$

(variance in sensitivity of $\log\lambda(t)$ to $\theta_i(t)$), and

$$\left.-\left[\frac{\partial \log \lambda_S}{\partial\theta_i}\right]^2\right\} \qquad (13)$$

(sensitivity of $\log\lambda_S$ to θ_i, squared) (see Appendix E for the derivation). So there is an exact partitioning of $s^{\sigma^2,0}_{S,i}$ in terms of nonlinear averaging (which changes mean demographic rates) and two aspects of sensitivity to the parameter mean. The same partitioning applies to $s^{\sigma^2}_{S,i}$ as a small-fluctuations approximation, rather than being exact. Increasing variance can therefore have opposing effects on population growth, possibly increasing the mean through nonlinear averaging but increasing the variance of annual growth rates. Depending on the relative magnitudes of these increases the net effect on λ_S may be either positive or negative. For the *Carlina* model, variance effects dominate for most parameters, resulting in selection against temporal variation.

There is no clear relationship between the sensitivity of λ_S to the mean of a parameter ($s^{\mu}_{S,i}$) and the sensitivity to the variance ($s^{\sigma^2}_{S,i}$; Fig. 2a). Given that multiple factors contribute to these sensitivities, this should not be

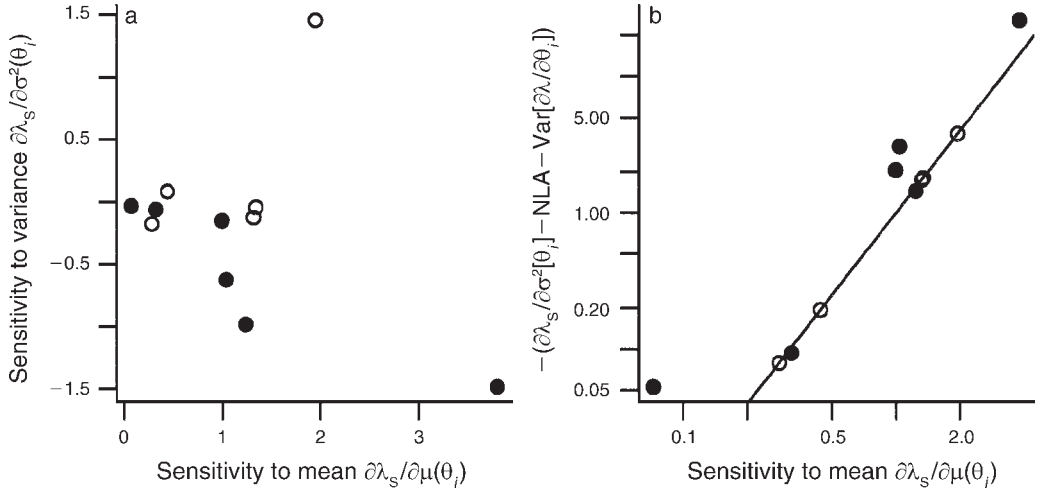

FIG. 2. Relationship between (a) the sensitivity to the mean of θ_i (the ith component of $\boldsymbol{\theta}$) ($s^{\mu}_{S,i}$) and the sensitivity to the variance of θ_i and (b) the sensitivity to the mean of θ_i ($s^{\mu}_{S,i}$) and the sensitivity to the variance of θ_i having removed the effects of nonlinear averaging (NLA) and variance in sensitivities of $\lambda(t)$ to the mean of θ_i. The line in (b) is the predicted relationship based on the partitioning of $s^{\sigma^2,0}_{S,i}$ given in Eq. 13. Open circles are for time-invariant parameters, solid circles for time-varying ones. See Table 3 for a summary of formulas for sensitivities.

TABLE 6. Stochastic elasticities for the parameters defining the kernel.

Parameter	$e_{S,i}$	$e^{\mu}_{S,i}$	$e^{\sigma}_{S,i}$
m_0	−0.89	−0.73	−0.17
m_s	0.77	1.11	−0.33
β_0	−1.23	−1.16	−0.08
β_s	1.08	...	...
a_g	1.13	1.17	−0.04
b_g	2.72	2.78	−0.06
σ^2_ε	0.11	...	...
A	0.44	...	...
B	3.85	...	...
a_R	3.06	2.08	−0.02
σ^2_ω	0.33	...	...

Notes: See Table 2 for descriptions of the parameters. The $e_{S,i}$ are the stochastic elasticities, $e^{\mu}_{S,i}$ are the stochastic elasticities to the mean, and $e^{\sigma}_{S,i}$ are the elasticities to the standard deviation. Ellipses indicate a parameter that is not time varying in the unperturbed model, so necessarily $e^{\mu}_{S,i} = e_{S,i}$ and $e^{\sigma}_{S,i} = 0$.

surprising. It is therefore more interesting to see how well the partitioning of Eq. 13 works for time-varying parameters. To do this we subtracted the effects of nonlinear averaging and variance in the sensitivities of $\lambda(t)$ from the estimated stochastic sensitivities, giving

$$-\left\{\frac{2}{\lambda_S}s^{\sigma^2}_{S,i} - E\left[\left\langle v_{t+1}, \frac{\partial^2 K_t}{\partial\theta_i^2}w_t\right\rangle \Big/ \langle v_{t+1}, K_t w_t\rangle\right] + \mathrm{Var}\left[\frac{\partial \log \lambda(t)}{\partial\theta_i(t)}\right]\right\} \approx \left[\frac{\partial \log \lambda_S}{\partial\theta_i}\right]^2.$$

The resulting relationship is exact for constant parameters and reasonable also for time-varying ones (Fig. 2b), suggesting that partitioning provides useful information for time-varying parameters. In general, elasticities of the parameters defining the kernel suggest that λ_S is more sensitive to proportional changes in a parameter's mean than its standard deviation (Table 6).

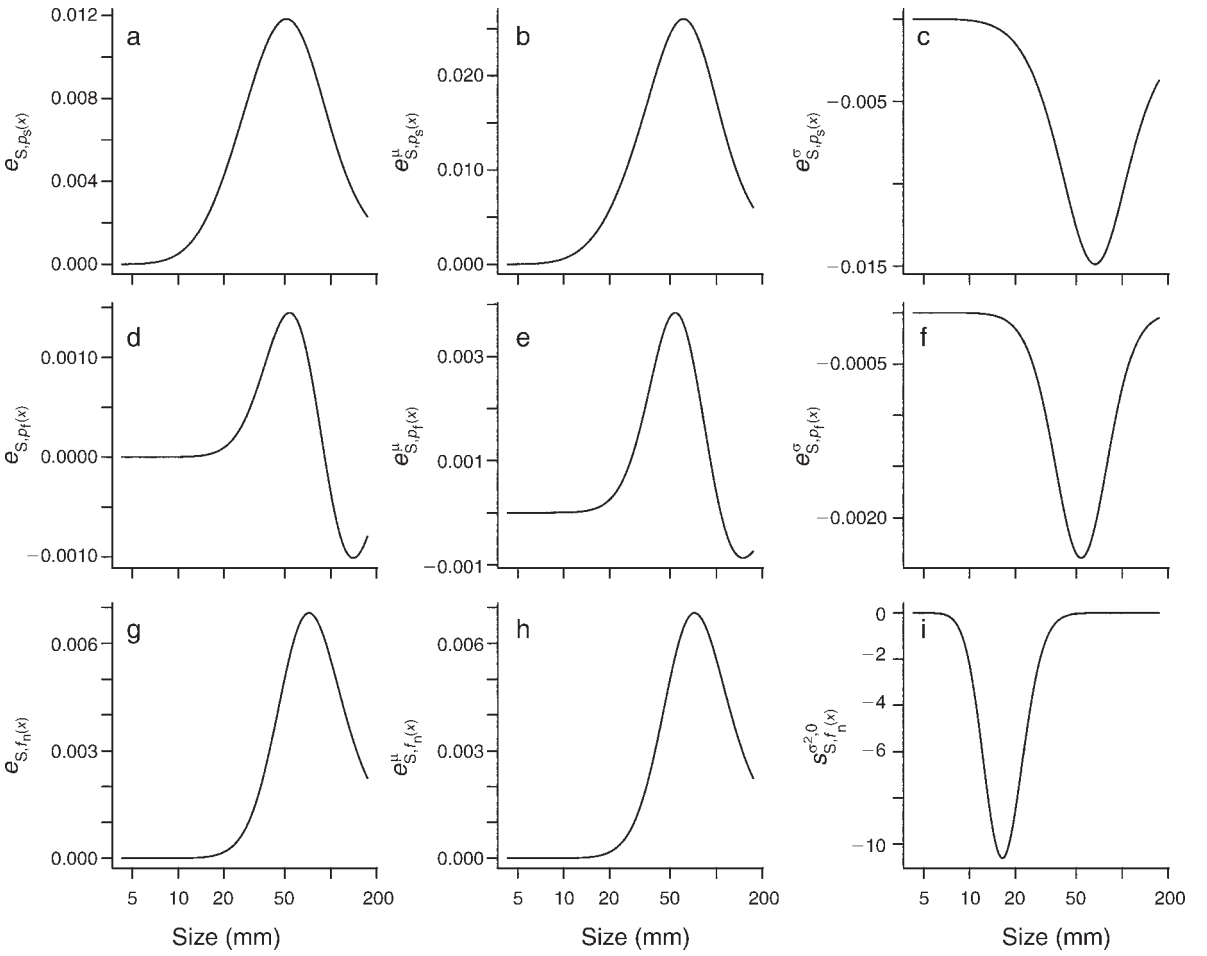

FIG. 3. Elasticity analysis of (a–c) the survival, $p_s(x)$, (d–f) the probability of flowering, $p_f(x)$, and (g–i) the seed production, $f_n(x)$, functions; see Table 2 for descriptions of the functions. Panels (a), (d), and (g) contain the stochastic elasticities; panels (b), (e), and (h) contain the stochastic elasticities to the mean; panels (c) and (f) contain the stochastic elasticities to the standard deviation; and panel (i) contains the stochastic sensitivity to the variance. Size is plotted on a log scale.

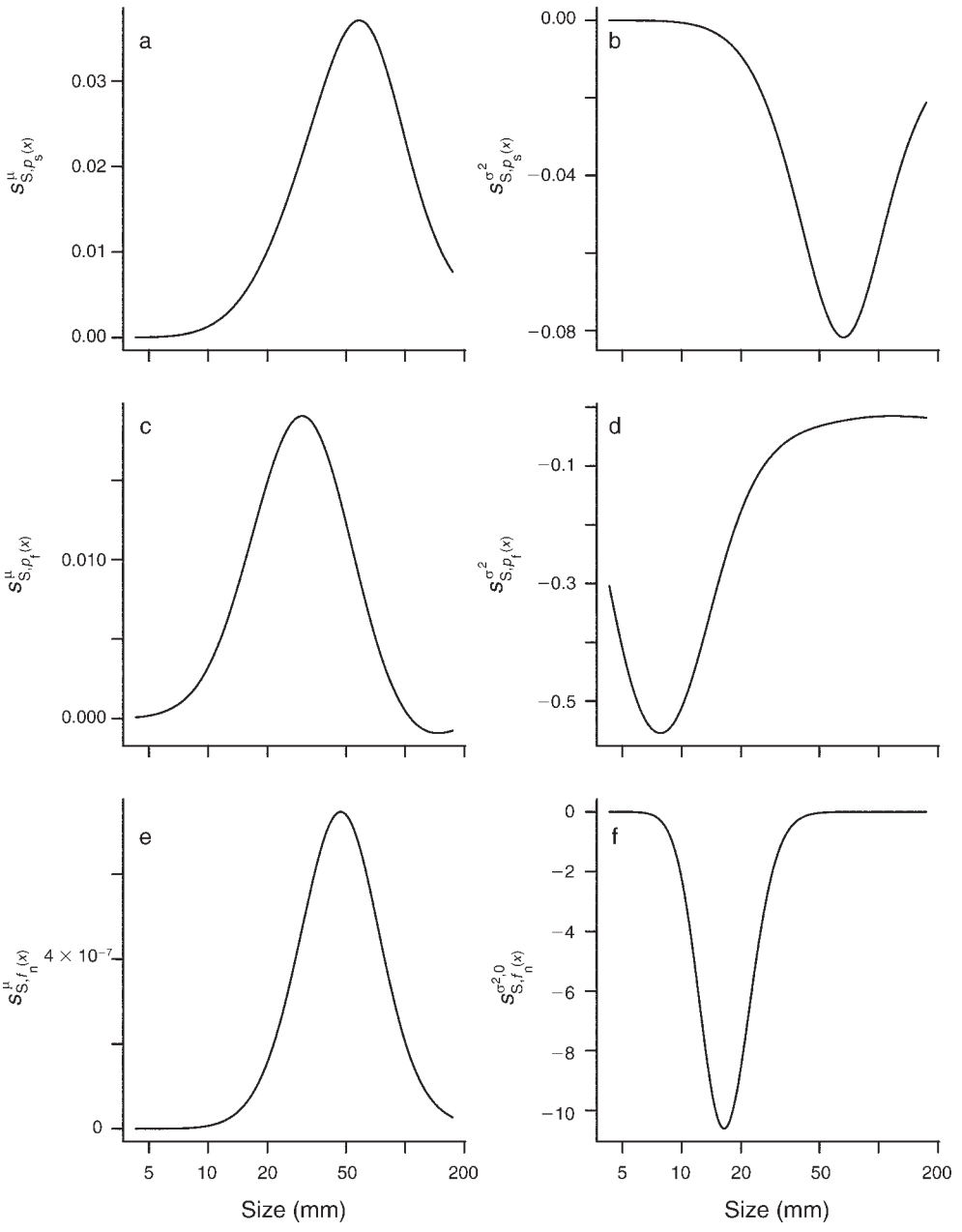

FIG. 4. Sensitivity analysis of (a, b) the survival, $p_s(x)$, (c, d) probability of flowering, $p_f(x)$, and (e, f) seed production, $f_n(x)$, functions; see Table 2 for descriptions of the functions. Panels (a), (c), and (e) contain the stochastic sensitivities to the mean, and panels (b), (d), and (f) contain the stochastic sensitivities to the variance. Size is plotted on a log scale.

Elasticity analysis for the functions used to construct the kernel clearly demonstrates the importance of the probabilities of survival and flowering and seed production functions of flowering individuals (average flowering size is 52 mm) in determining the stochastic growth rate (Fig. 3a, d, g). Increasing the mean of these functions increases the stochastic growth rate as expected (Fig. 3b, e, h); increasing variability has the opposite effect (Fig. 3c, f, i). The seed production function, which is not time varying, is most sensitive to temporal variation of relatively small individuals (~20 mm; Fig. 3i), whereas the elasticities of the probability of survival and flowering functions suggest that temporal variation in these vital rates for flowering individuals has the greatest impact on the stochastic growth rate (Fig. 3c, f). The stochastic sensitivities for all functions show a similar pattern (Fig. 4), but with the maximum sensitivity to the variance in the probability of flowering occurring at smaller sizes than in the elasticity analysis (~10 mm).

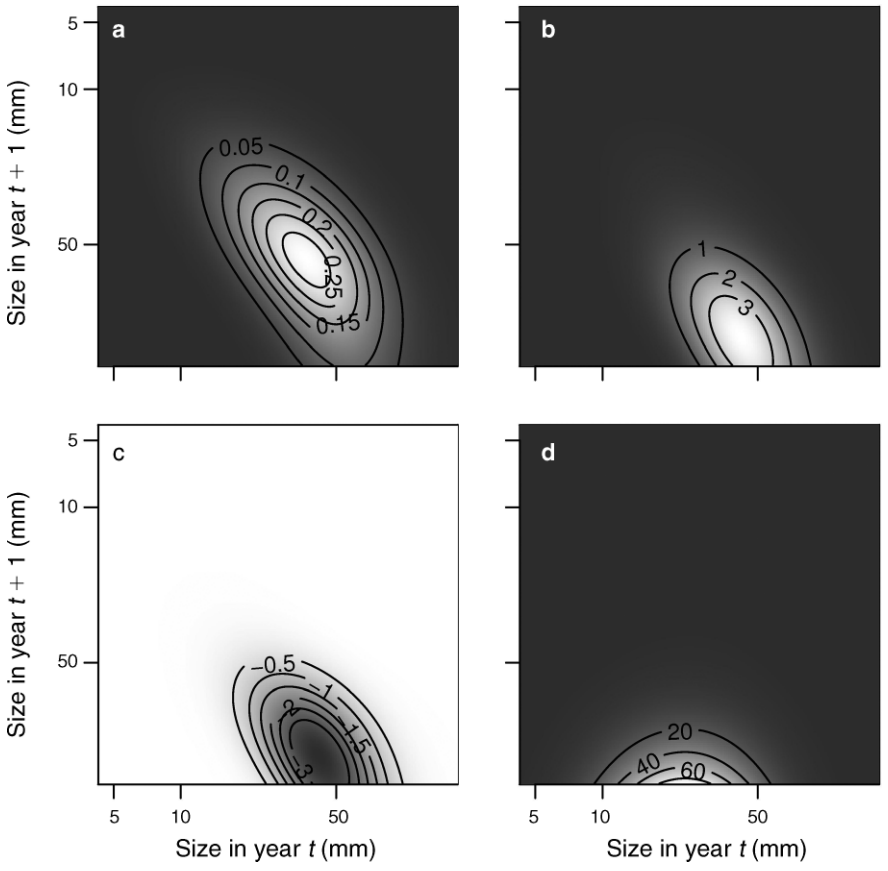

FIG. 5. Shaded contour plots of stochastic elasticities and sensitivities for the growth function, $g(y, x)$. The values along the contours are for (a) the stochastic elasticities, (b) the stochastic elasticities to the mean, (c) stochastic elasticities to the standard deviation, and (d) the stochastic sensitivity.

Elasticity analysis for the growth function, $g(y, x)$, is shown in Fig. 5. Interpreting the elasticities to the growth function is complicated by the fact that $\int g(y, x)\,dy = 1$, and so an increase in one growth transition has to be compensated by decreases somewhere else; the same issue arises in the perturbation analysis of Markov chains (see Caswell 2001: section 9.6). For locations at which the elasticity is small, increases there that were compensated by decreases elsewhere would generally decrease λ_S. Increases where the elasticity is large, compensated by decreases elsewhere, would generally increase λ_S. With this interpretation, the elasticity analysis of $g(y, x)$ demonstrates the importance of growth transitions into the flowering size classes (Fig. 5a). Increases in the mean of the growth function increase the stochastic growth rates as expected (Fig. 5b), while increases in the standard deviation have the opposite effect (Fig. 5c). In contrast to the elasticity analysis, the sensitivity of the stochastic growth rate is extremely large for rare transitions from small to large sizes. The difference between the elasticities and sensitivities reflects the fact that a rare transition generally will have a low elasticity, because even a large proportional change will be small in absolute terms.

RETROSPECTIVE DEMOGRAPHIC ANALYSES

Methods: life table response experiments (LTRE)

The general goal of LTRE analyses is to decompose the year-to-year variation in the value of λ for the year-specific model into contributions from the variances and covariances of underlying parameters. In keeping with our emphasis on the mixed-models framework we consider only the random designs situation (in the terminology of Caswell 2001: chapter 10) but fixed-design and regression design analyses are equally possible (Williams and Crone 2006). The first-order expression for the variance in λ is

$$\mathrm{Var}(\lambda) \approx \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(\theta_i, \theta_j) s_i s_j \qquad (14)$$

where s_i is the sensitivity to θ_i in the mean kernel given by

$$s_i = \left\langle v, \frac{\partial \bar{K}}{\partial \theta_i} w \right\rangle \Big/ \langle v, w \rangle.$$

Because the kernel is a nonlinear function of the parameters θ_i, we must estimate the mean kernel and

$\partial\bar{K}/\partial\theta_i$ by simulating from θ and averaging the year-specific K's and $\partial K/\partial\theta_i$'s. The covariances in Eq. 14 are all mixing distribution parameters in the fitted demographic models. For example from Table 2 we have $\mathrm{Cov}(b_g, b_g) = \mathrm{Var}(b_g) = 0.13^2$. Dividing through by the sum of the terms on the right-hand size of Eq. 14 allows us to quantify the relative contribution of each temporally varying parameter to the year-to-year variability in λ, within the accuracy of the first-order approximation. Note that relative contributions to $\mathrm{Var}(\lambda)$ and to $\mathrm{Var}(\log\lambda)$ are equal to first order, because the first-order expansions for λ and $\log\lambda$ only differ by a constant factor.

To explore the adequacy of the first-order approximation, we used a Monte Carlo approach. Specifically we drew parameter sets from the fitted distribution of $\boldsymbol{\theta}(t)$, computed the corresponding values of λ for the year-specific model, and used statistical analyses to estimate the impact of each parameter on λ using a linear model, $\lambda = \beta_0 + \Sigma_i\ \beta_i\theta_i + \varepsilon$. The variance decomposition for λ under the regression model is

$$\mathrm{Var}(\lambda) \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(\theta_i, \theta_j)\beta_i\beta_j + \mathrm{Var}(\varepsilon). \qquad (15)$$

So the fitted slopes are analogous to the sensitivities in the first-order approximation. The regression approach has the advantages that the β's are estimated over the full range of variation of the θ_i's and that regression diagnostics can be used to explore the adequacy of the linear model. If a nonlinear model is required, this could be fitted as a generalized additive model (GAM) $\lambda = \beta_0 + \Sigma_j\ \eta(\theta_j) + e$ where $\eta(\theta_j)$ is a smooth function, with the identifiability constraints that $E[\eta(\theta_j)] = 0$ for all j. The variance decomposition under the GAM model is

$$\mathrm{Var}(\lambda) \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(\eta(\theta_i), \eta(\theta_j)) + \mathrm{Var}(\varepsilon). \qquad (16)$$

Alternatively, if the number of between-parameter correlations is relatively small, parameters can be separated into uncorrelated parameter groups and the relative contribution of each group to $\mathrm{Var}(\lambda)$ can be found using nonparametric variance decomposition (see Ellner and Guckenheimer [2006] section 8.5.2 for a general discussion and Ellner and Fieberg [2003] for an application to a stochastic matrix model).

For the LTRE analyses we used the density-independent *Carlina* model setting the probability of establishment, $p_e(t)$, equal to 0.00095 so that λ_S was approximately that observed in the field.

Results: life table response experiments (LTRE)

The first-order analysis (Table 7) suggests that yearly variation in growth and survival each account for ~35% of the variation in λ and that variation in recruit size and the interaction between recruit size and growth intercept account for a further ~15% each. The only term that is unimportant is the variation in the flowering intercept, which accounts for ~1% of the variation in λ. We obtain essentially the same results from the linear-regression-based analysis, Eq. 15.

TABLE 7. Life table response experiment (LTRE) analyses using the first-order approximation $\mathrm{Cov}(\theta_i, \theta_j)s_is_j$, the linear regression analysis $\mathrm{Cov}(\theta_i, \theta_j)\beta_i\beta_j$, and the generalized additive model (GAM)-based analysis of contributions to variability in $\log(\lambda)$.

Parameter	$\mathrm{Cov}(\theta_i, \theta_j)s_is_j$	$\mathrm{Cov}(\theta_i, \theta_j)\beta_i\beta_j$	GAM (log analysis)
m_0	0.12	0.11	0.13
m_s	0.23	0.20	0.28
m_0:m_s			0.16
β_0	0.01	0.02	0.01
β_s			
a_g	0.05	0.05	0.04
b_g	0.30	0.28	0.21
σ^2_ε			
A			
B			
a_R	0.16	0.19	0.09
σ^2_ω			
a_g:a_R	0.14	0.15	0.08
r^2		0.77	0.97

Note: In all cases the contributions to variation have been normalized to sum to 1, and empty cells indicate negligibly small contributions. For the growth : recruit size interaction we multiply the covariance term by 2, as there are two covariance terms in the expansion (Eq. 14).

However, the linear model only accounts for 77% of the variation in λ and regression diagnostic plots suggest substantial systematic variation from the fitted model and variance increasing with the fitted values. In order to improve the statistical properties of the model we therefore fitted a GAM model to $\log(\lambda)$ that included a bivariate spline term for interactions between survival intercept and slope. This model accounted for 97% of the variation in $\log(\lambda)$ and had much better regression diagnostics. For the GAM model yearly variation in survival accounted for nearly 60% of the variation in $\log(\lambda)$, with growth and recruit size variation accounting for the rest, and as in the previous analysis yearly variation in flowering intercept was relatively unimportant (~1%).

EVOLUTIONARY DEMOGRAPHY

Methods: evolutionarily stable flowering strategies and function-valued trait evolution

Previously we have explored the evolution of flowering strategies using stochastic fixed-effects IPMs (Childs et al. 2004, Rees et al. 2004, 2006, Ellner and Rees 2007, Metcalf et al. 2008). The flowering strategy was defined by the fitted parameters of a logistic regression describing how the probability of flowering varies with plant size. In all these studies temporal variation in the flowering strategy could not be simply incorporated into the analysis and so it was ignored. In contrast in the mixed-model framework temporal variation in the

flowering strategy is easily incorporated and we can explore how the mean or standard deviation of the intercept or slope of the logistic regression evolves. In *Carlina* three parameters define the regression model for the probability of flowering, namely, the mean and standard deviation of the intercept, $\beta_0 \sim N(-16.19, 1.03)$, and the size-dependent slope, β_s (Table 2). The *Carlina* mixed-effects IPM also differs from our previous studies in that temporal variation in recruit size and its covariance with the growth intercept was explicitly modeled.

Evolutionarily stable flowering strategies were identified using an iterative sequence of invasions. Specifically we generated $\mathbf{\theta}(t)$ using the resident model and then maximized λ_S for the invader model as a function of the invader's flowering strategy (e.g., the mean intercept). If the maximum value of λ_S was close to 1 within a specified tolerance then the resident flowering strategy is the ESS. If there are strategies with $\lambda_S > 1$ then the resident flowering strategy was set to that of the invader with the maximum λ_S and the process repeated (see Childs et al. [2004] for details).

The size-dependent flowering probability is a function-valued trait, meaning that the entire shape of the $p_f(x)$ function is in principle shaped by natural selection. Our preceding evolutionary analyses took a parametric approach, assuming that logit p_f is a linear function of x specified by slope and intercept parameters, as we estimated from our field data. Dieckmann et al. (2006) have pointed out that presupposing a specific parametric form for the ESS of a function-valued trait is risky, because the optimum within the parametric family may be highly suboptimal relative to more general functional forms. Parvinen et al. (2006) have shown how the calculus of variations can be used to find ESS function-valued traits in a wide class of deterministic models with continuous population structure. Our results make it possible to do the same for stochastic population dynamics modeled by an IPM.

Techniques from functional data analysis (Ramsay and Silverman 2005) make nonparametric optimization of function-valued traits computationally tractable. Specifically, we used a B-spline expansion:

$$\mathrm{E}[\mathrm{logit}\, p_f(x)] = \sum_{j=1}^{d} c_j \varphi_j(x) \qquad (17)$$

where the functions $\{\varphi_j(x)\}_{j=1}^{d}$ are the cubic B-spline basis on the size range $[L, U]$ of the *Carlina* model with evenly spaced interior knots. The value of d determines the potential complexity of the functions that can be represented, so to be sure that d is large enough to accommodate the ESS functional form, the results should be checked by verifying that increasing d has no effect. For the calculations presented here we used $d = 6$ initially and then repeated the calculations with $d = 7$.

The question we posed is whether the parametric ESS is really optimal, or could it be invaded by a more general flowering strategy represented by Eq. 17? We subjected the invader to the same constraints as the parametric ESS: an upper bound on the slope of logit $p_f(x)$ and stochastic variation in the intercept. The optimization problem is then to maximize λ_S as a function of the coefficients c_j in Eq. 17, for an invader with

$$\mathrm{logit}\, p_f(x) = \sum_{j=1}^{d} c_j \varphi_j(x) + N(0, 1.03)$$

facing a resident with the estimated (Table 2) or ESS parametric flowering strategies, subject to the constraint that $d[\mathrm{logit}\, p_f]/dx \leq \beta_s = 3.883$.

Results: evolutionarily stable flowering strategies and evolution of function-valued traits

When the means of both the mean of β_0 and β_s are allowed to evolve, the predicted strategy (ESS) is a step function, where the probability of flowering is either 0 or 1, depending upon whether plants are smaller or larger than the threshold. The predicted mean size at flowering is ~58 mm, slightly larger than observed in the field, but the distribution of flowering sizes is very different with fewer small flowering plants and more in the range 35–55 mm than observed in the field. The observed size-dependent flowering is gradual, however, as a result of variation in the threshold size for flowering, variable growth between the time the decision to flower is made and when the size at flowering was recorded, or an imperfect correlation between size and the threshold condition for flowering (Rees and Rose 2002, Childs et al. 2003). To allow for this we treated the slope of the flowering function, β_s, as fixed and determined the ESS mean intercept (see Childs et al. 2003). The ESS mean intercept was −14.14, which was not significantly different from the estimated value (−16.19 ± 1.66; parameter ± SE), and the mean flowering size at the ESS was 51.0 mm, again not significantly different from that observed in the field (52 mm); the predicted stable size distribution is close to that observed in the field (results not presented).

The mean flowering intercept β_0 estimated from field data is slightly lower than the ESS, which explains why there is selection for an increase in the mean of the probability of flowering function (Fig. 3e); the ESS flowering probability is slightly greater (at all sizes) than that estimated from the field data, so increasing the mean of $p_f(x)$ increases fitness. In contrast, if the mean of β_0 is set to a value higher than the ESS, say $\beta_0 = -10$, then increases in the mean of $p_f(x)$ would move the population even further from the ESS and the elasticities for the mean of $p_f(x)$ are all negative (results not presented).

The perturbation analysis of flowering probability (Figs. 3e and 4c) shows that when the resident has the flowering strategy estimated from field data ($\beta_0 = -16.19$, $\beta_s = 3.883$), there is selection for more flowering by mid-size plants and less flowering by large plants.

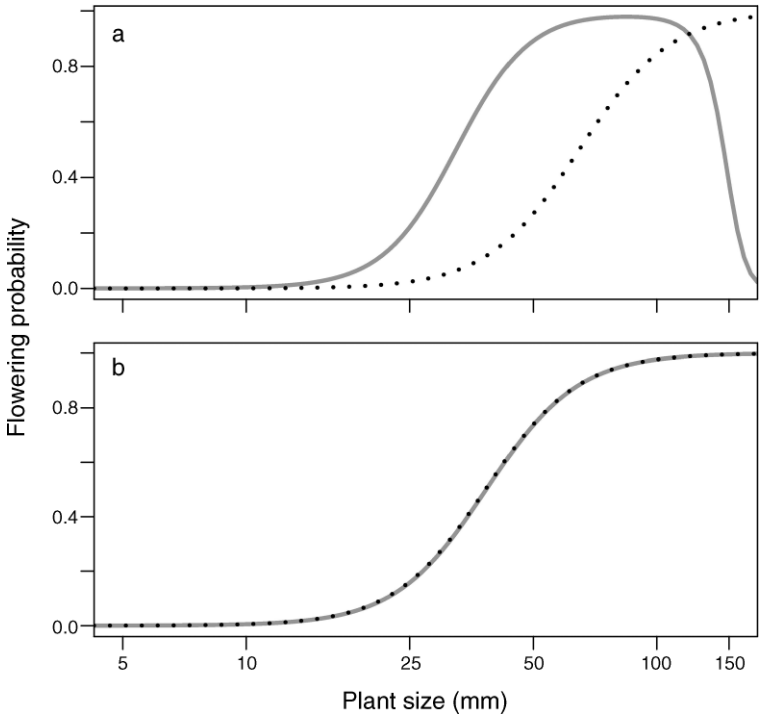

FIG. 6. A nonparametric approach to evolution of flowering strategy as a function-valued trait. In both panels, the resident type has flowering strategy logit $p_f(x) = \beta_0 + \beta_s x$ as in Table 2, with random variability in β_0. The dotted curve shows the parametric flowering strategy, plotted for β_0 held constant at its expected value. The solid gray curve shows the optimal invader specified by a B-spline expansion for the expectation of logit $p_f(x)$; the plotted values are the right-hand side of Eq. 17. In panel (a) the resident has $E[\beta_0] = -16.19$, the value estimated from the field data. In panel (b) the resident has $E[\beta_0] = -14.15$, the estimated parametric evolutionarily stable strategy (ESS).

This results in the optimal nonparametric flowering strategy for the invader being very different from that of the resident, with flowering probability not even monotonically increasing in size (Fig. 6a). In contrast, when the resident has the parametric ESS, the optimal nonparametric invader is identical to the resident within the accuracy of our numerical calculations (Fig. 6b). This shows that our parametric ESS described by a logistic regression model is in fact able to resist invasion by any alternative strategy and is therefore the ESS flowering strategy.

DISCUSSION

For readers who care most about the how-to, our message about stochastic IPMS can be summarized as an aphorism: fit smoothly and use a big matrix. "Smooth" applies in several senses: (1) if a trait varies continuously in the study population, then model it that way using an IPM rather than a conventional matrix model; (2) model demographic rate functions as continuous functions of underlying continuous traits; (3) use a mixed-models framework to model temporal variation by fitting continuous probability distributions for the parameters defining the demographic rate functions, rather than by estimating a discrete set of values corresponding to the years for which you happen to have data. Our *Carlina* case study illustrates that these recommendations have significant practical advantages on top of the conceptual advantage of better representing biological reality.

If you do fit smoothly, the mathematical theory guarantees that you can safely use the biggest matrix your computer permits, because bigger matrices give more and more accurate numerical approximations to the limit of truly continuous trait distributions. The situation for conventional matrix models is the opposite: growing the matrix eventually leads to nonsense, because the increase in the number of parameters (matrix entries) entails a decrease in statistical precision. With a big matrix implementing an IPM you need computationally efficient ways to do sensitivity analyses, LTRE, or whatever else you like, and we have provided those as summarized in Table 3. Recognizing that the matrix operations are really numerical approximations to integrals also makes it possible to use more efficient numerical integration methods (e.g., Easterling et al. 2000), which may be important when there are many individual-level state variables or when many model runs are needed, such as for computing population prediction intervals (Sæther et al. 2000).

In addition to its practical advantages for model fitting, the mixed-models framework is advantageous for

LTRE analyses because the standard first-order analysis can be tested by generating many yearly kernels and fitting statistical models for the relationship between parameter values and population growth rate. In our example there was excellent agreement between the first-order expansion and linear regression analysis. But neither of these was accurate, because regression diagnostics revealed large systematic deviations from the assumed linear relationship. This problem was corrected by using a generalized additive model, which revealed that the importance of survival variability was nearly twice as important as the linear analyses had suggested.

Elasticities and sensitivities have close links to evolutionary theory. For example Lande (2007, 2008) has demonstrated that the expected change in (1) the allele frequency (p) in a single-locus haploid (asexual) or randomly mating diploid population or (2) the mean phenotype of quantitative characters ($\bar{z}$) in a sexual population is governed by the gradient of the long-run growth rate of the population, $\log(\lambda_S)$, with respect to p or $\bar{z}$, which is calculated by our Eq. 10. The stochastic sensitivities computed for our "invader" model in which the resident is at equilibrium ($\lambda_S = 1$) have a direct interpretation as expected selection gradients for the resident because

$$\frac{\partial \log \lambda_S}{\partial \varepsilon} = \frac{1}{\lambda_S}\frac{\partial \lambda_S}{\partial \varepsilon} = \frac{\partial \lambda_S}{\partial \varepsilon}.$$

Our results indicate directional selection for increases in the means of all the model parameters (Table 5). The weak selection on β_0 is consistent with it being close to an ESS at which there is no directional selection. The ESS occurs because both the costs (in terms of survival) and benefits (seed production) of higher size-specific flowering are included in the model. For other parameters we lack cost estimates, so the predictions of directional selection must be interpreted with caution. For example, if there is a trade-off between growth and survival, as observed in many monocarpic perennials (Metcalf et al. 2006), the parameters defining the growth and survival curves may in fact be at a joint ESS. This problem of incorrectly specifying or ignoring costs is common to many forms of selection analysis. For example Merila et al. (2001) show that in 13 studies in which there are estimates of the strength of directional selection on heritable traits using regression-based approaches (Lande and Arnold 1983), none of the traits changed in the direction predicted.

Regression-based approaches to characterizing selection in structured populations also ignore how the perturbation formulas that define the selection gradients (Eqs. 10 and 12) depend on the numbers (w_t) and reproductive values (v_{t+1}) of individuals. The relatively small impact of large individuals on λ_S (Figs. 1, 3, 4, and 5) is a consequence of there being very few of them, even though they have highest survivorship, flowering probability, and seed production. Standard selection analyses based on regression of relative fitness on trait values analyses (Lande and Arnold 1983) ignore population structure, whereas the approaches presented here allow for population structure.

The growth sensitivities suggest strong selection for rapid growth (Fig. 5d), and this might in part explain the enormous variance in growth rate even for plants of the same size seen in most natural plant populations (Metcalf et al. 2003, 2006). Being able to respond to environmental variation in growth brings enormous fitness benefits, which is why plant growth responses are so plastic.

We found a strong negative correlation between the elasticities of λ_S to the mean and to the variance for kernel entries (Fig. 1, Table 4). This is exactly what would be expected based on the small variance approximation, Eq. 9. Temporal variability in parameters often reduces λ_S, but this need not always be true. For example, in the density-independent *Carlina* model there is selection for temporal variation in the slope of the seed production function, B. In this case the effect of nonlinear averaging is sufficient to offset the negative effects of temporal variance.

The sensitivities of growth rate to temporal variance in kernel entries are always negative, whereas sensitivities to variance in underlying parameters can be positive when nonlinear averaging causes a large increase in the mean kernel. This might seem to complicate the interpretation of sensitivity analysis at the level of the underlying parameters. However, the simple results for kernel entries result from considering perturbations that affect one very small segment of the population independent of all others (Doak et al. 2005), which is usually difficult to justify biologically. In contrast, most underlying parameters in the demographic models really do vary independently, so the results are straightforward to interpret.

The models in this paper can be viewed as the "mean field" description for a population large enough that demographic stochasticity can be ignored. Exact representation of demographic stochasticity would require individual-level stochastic simulations, sacrificing the entire analytic tool kit that is available for integral or matrix models. But there is a very useful middle ground for stochastic matrix models, in which the mean field model is perturbed by random variables whose mean and variance approximate the net effect of many-individuals-level random events; a well-known example is the "LPA model" for experimental *Tribolium* populations with demographic stochasticity (e.g., Dennis et al. 2001: Eqs. 10–12). We are not aware of anything similar for integrodifference equations. Given the other advantages of IPMs, an approach to modeling "medium-size" populations in which demographic stochasticity is neither negligible nor dominant would be a valuable contribution. This is especially true for models with spatial location as one of the individual-level continuous traits: "very many" individuals of about

the same size is far easier to rationalize than "very many" individuals at about the same location. Another important direction is to develop the theory of density-dependent stochastic IPMs, for which a lot of the hard work has already been done (Hardin et al. 1988, Benaïm and Schreiber 2009) in the setting of spatial or matrix models.

Acknowledgments

Research support was provided by NSF grant OCE 0326705 in the NSF/NIH Ecology of Infectious Diseases program (to S. P. Ellner) and Natural Environment Research Council grant NER/A/S/2002/00940 (to M. Rees). We thank Rob Freckleton, Jessica Metcalf, and two anonymous referees for comments on the manuscript.

Literature Cited

Benaïm, M., and S. J. Schreiber. 2009. Persistence of structured populations in random environments. Theoretical Population Biology 76:19–34.

Bröström, G. 2008. glmmML: generalized linear models with clustering. R package version 0.72-0. R Foundation for Statistical Computing, Vienna, Austria.

Cam, E., W. A. Link, E. G. Cooch, J. Y. Monnat, and E. Danchin. 2002. Individual covariation in life-history traits: seeing the trees despite the forest. American Naturalist 159: 96–105.

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

Childs, D. Z., M. Rees, K. E. Rose, P. J. Grubb, and S. P. Ellner. 2003. Evolution of complex flowering strategies: an age- and size-structured integral projection model. Proceedings of the Royal Society B 270:1829–1838.

Childs, D. Z., M. Rees, K. E. Rose, P. J. Grubb, and S. P. Ellner. 2004. Evolution of size-dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. Proceedings of the Royal Society B 271:425–434.

Clark, J. S., M. Dietze, S. Chakraborty, P. K. Agarwal, I. Ibanez, S. LaDeau, and M. Wolosin. 2007. Resolving the biodiversity paradox. Ecology Letters 10:647–659.

de Jong, T. J., P. G. L. Klinkhamer, and J. L. H. de Heiden. 2000. The evolution of generation time in metapopulations of monocarpic perennial plants: some theoretical considerations and the example of the rare thistle *Carlina vulgaris*. Evolutionary Ecology 14:213–231.

Dennis, B., R. A. Desharnais, J. M. Cushing, S. M. Henson, and R. F. Costantino. 2001. Estimating chaos and complex dynamics in an insect population. Ecological Monographs 71:277–303.

Dieckmann, U., M. Heino, and K. Parvinen. 2006. The adaptive dynamics of function-valued traits. Journal of Theoretical Biology 241:370–389.

Doak, D. F., W. F. Morris, C. Pfister, B. E. Kendall, and E. M. Bruna. 2005. Correctly estimating how environmental stochasticity influences fitness and population growth. American Naturalist 166:E14–E21.

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., and J. Fieberg. 2003. Using PVA for management despite uncertainty: effects of habitat, hatcheries, and harvest on salmon. Ecology 84:1359–1369.

Ellner, S. P., and J. Guckenheimer. 2006. Dynamics models in biology. Princeton University Press, Princeton, New Jersey, USA.

Ellner, S. P., and M. Rees. 2006. Integral projection models for species with complex demography. American Naturalist 167: 410–428.

Ellner, S. P., and M. Rees. 2007. Stochastic stable population growth in integral projection models: theory and application. Journal of Mathematical Biology 54:227–256.

Eriksson, A., and O. Eriksson. 1997. Seedling recruitment in semi-natural pastures: the effects of disturbance, seed size, phenology and seed bank. Nordic Journal of Botany 17:469–482.

Ghidey, W., E. Lesaffre, and P. Eilers. 2004. Smooth random effects distribution in a linear mixed model. Biometrics 60: 945–953.

Gross, K., W. F. Morris, M. S. Wolosin, and D. F. Doak. 2006. Modeling vital rates improves estimation of population projection matrices. Population Ecology 48:79–89.

Hairston, N. G., Jr., S. Ellner, and C. M. Kearns. 1996. Overlapping generations: the storage effect and the maintenance of biotic diversity. Pages 109–145 *in* O. E. Rhodes, R. K. Chesser, and M. H. Smith, editors. Population dynamics in ecological space and time. University of Chicago Press, Chicago, Illinois, USA.

Halley, J. M., and P. Inchausti. 2004. The increasing importance of 1/f-noises as models of ecological variability. Fluctuation and Noise Letters 4:R1–R26.

Hardin, D. P., P. Takáč, and G. F. Webb. 1988. Asymptotic properties of a continuous-space discrete-time population model in a random environment. Journal of Mathematical Biology 26:361–374.

Haridas, C. V., and S. Tuljapurkar. 2005. Elasticities in variable environments: properties and implications. American Naturalist 166:481–495.

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

Klinkhamer, P. G. L., T. J. de Jong, and J. L. H. de Heiden. 1996. An 8-year study of population-dynamics and life-history variation of the biennial *Carlina vulgaris*. Oikos 75:259–268.

Klinkhamer, P. G. L., T. J. de Jong, and E. Meelis. 1991. The control of flowering in the monocarpic perennial *Carlina vulgaris*. Oikos 61:88–95.

Lande, R. 2007. Expected relative fitness and the adaptive topography of fluctuating selection. Evolution 61:1835–1846.

Lande, R. 2008. Adaptive topography of fluctuating selection in a Mendelian population. Journal of Evolutionary Biology 21: 1096–1105.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10: 325–337.

McCullough, C. E., and S. R. Searle. 2001. Generalized, linear, and mixed models. John Wiley and Sons, New York, New York, USA.

Merila, J., B. C. Sheldon, and L. E. B. Kruuk. 2001. Explaining stasis: microevolutionary studies in natural populations. Genetica 112:199–222.

Metcalf, C. J. E., M. Rees, J. M. Alexander, and K. Rose. 2006. Growth–survival trade-offs and allometries in rosette-forming perennials. Functional Ecology 20:217–225.

Metcalf, C. J. E., K. E. Rose, D. Z. Childs, A. W. Sheppard, P. J. Grubb, and M. Rees. 2008. Evolution of flowering decisions in a stochastic, density-dependent environment. Proceedings of the National Academy of Sciences (USA) 105: 10466–10470.

Metcalf, J. C., K. E. Rose, and M. Rees. 2003. Evolutionary demography of monocarpic perennials. Trends in Ecology and Evolution 18:471–480.

Metcalf, C. J. E., D. A. Stevens, M. Rees, S. M. Louda, and K. H. Keeler. 2009. Using Bayesian inference to understand the allocation of resources between sexual and asexual reproduction. Journal of the Royal Statistical Society—Series C 58:143–170.

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

Parvinen, K., U. Dieckmann, and M. Heino. 2006. Function-valued adaptive dynamics and the calculus of variations. Journal of Mathematical Biology 52:1–26.

Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-Plus. Springer-Verlag, London, UK.

R Development Core Team. 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramsay, J. O., and B. W. Silverman. 2005. Functional data analysis. Second edition. Springer-Verlag, New York, New York, USA.

Rees, M., D. Z. Childs, J. C. Metcalf, K. E. Rose, A. W. Sheppard, and P. J. Grubb. 2006. Seed dormancy and delayed flowering in monocarpic plants: selective interactions in a stochastic environment. American Naturalist 168:E53–E71.

Rees, M., D. Z. Childs, K. E. Rose, and P. J. Grubb. 2004. Evolution of size dependent flowering in a variable environment: partitioning the effects of fluctuating selection. Proceedings of the Royal Society B 271:471–475.

Rees, M., M. Mangel, L. A. Turnbull, A. Sheppard, and D. Briese. 2000. The effects of heterogeneity on dispersal and colonisation in plants. Pages 237–265 *in* M. J. Hutchings, E. A. John, and A. J. A. Stewart, editors. Ecological consequences of environmental heterogeneity. Blackwell, Oxford, UK.

Rees, M., and K. E. Rose. 2002. Evolution of flowering strategies in *Oenothera glazioviana*: an integral projection model approach. Proceedings of the Royal Society B 269: 1509–1515.

Rose, K. E., M. Rees, and P. J. Grubb. 2002. Evolution in the real world: stochastic variation and the determinants of fitness in *Carlina vulgaris*. Evolution 56:1416–1430.

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

Spiegelhalter, D., A. Thomas, N. Best, W. Gilks, and D. Lunn. 2003. BUGS: Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, UK.

Tuljapurkar, S. 1990. Population dynamics in variable environments. Springer-Verlag, London, UK.

Tuljapurkar, S., and C. V. Haridas. 2006. Temporal autocorrelation and stochastic population growth. Ecology Letters 9: 324–334.

Tuljapurkar, S., C. C. Horvitz, and J. B. Pascarella. 2003. The many growth rates and elasticities of populations in random environments. American Naturalist 162:489–502.

Williams, J. L., and E. E. Crone. 2006. The impact of invasive grasses on the population growth of *Anemone patens*, a long-lived native forb. Ecology 87:3200–3208.

Zhang, D. W., and M. Davidian. 2001. Linear mixed models with flexible distributions of random effects for longitudinal data. Biometrics 57:795–802.

APPENDIX A

Numerical solution methods (*Ecological Archives* M079-020-A1).

APPENDIX B

Mathematical assumptions (*Ecological Archives* M079-020-A2).

APPENDIX C

Derivation of the small variance approximation (*Ecological Archives* M079-020-A3).

APPENDIX D

Effect of adding small independent variability with zero mean (*Ecological Archives* M079-020-A4).

APPENDIX E

Derivations of stochastic elasticities and sensitivities (*Ecological Archives* M079-020-A5).

APPENDIX F

Perturbations to vital rate functions (*Ecological Archives* M079-020-A6).