Popul Ecol (2006) 48:79–89
DOI 10.1007/s10144-005-0238-8

ORIGINAL ARTICLE

Kevin Gross · William F. Morris · Michael S. Wolosin
Daniel F. Doak

Modeling vital rates improves estimation of population projection matrices

Received: 12 June 2005 / Accepted: 15 September 2005 / Published online: 10 November 2005

Abstract Population projection matrices are commonly used by ecologists and managers to analyze the dynamics of stage-structured populations. Building projection matrices from data requires estimating transition rates among stages, a task that often entails estimating many parameters with few data. Consequently, large sampling variability in the estimated transition rates increases the uncertainty in the estimated matrix and quantities derived from it, such as the population multiplication rate and sensitivities of matrix elements. Here, we propose a strategy to avoid overparameterized matrix models. This strategy involves fitting models to the vital rates that determine matrix elements, evaluating both these models and ones that estimate matrix elements individually with model selection via information criteria, and averaging competing models with multimodel averaging. We illustrate this idea with data from a population of *Silene acaulis* (Caryophyllaceae), and conduct a simulation to investigate the statistical properties of the matrices estimated in this way. The simulation shows that compared with estimating matrix elements individually, building population projection matrices by fitting and averaging models of vital-rate estimates can reduce the statistical error in the population projection matrix and quantities derived from it.

Keywords Information criteria · Model selection · Multimodel averaging · Vital rates

K. Gross (✉)
Biomathematics Program, North Carolina State University,
CB 8203, Raleigh, NC 27695, USA
E-mail: gross@stat.ncsu.edu
Tel.: +1-919-5134690
Fax: +1-919-5151909

W. F. Morris · M. S. Wolosin
Department of Biology and Program in Ecology,
Duke University, Durham, NC, USA

D. F. Doak
Department of Ecology and Evolutionary Biology,
University of California at Santa Cruz, Santa Cruz, CA, USA

Introduction

Population projection matrices (PPMs) are popular tools among ecologists and wildlife managers for modeling and analyzing population dynamics (Tuljapurkar and Caswell 1997; Caswell 2001; Morris and Doak 2002; and references therein). Among many other applications, PPMs have been used to assess conservation strategies for northern spotted owl (Lande 1988), loggerhead sea turtles (Crowder et al. 1994), and grizzly bears (Wielgus 2002); to explain population trends of an imperiled whale population (Fujiwara and Caswell 2001); to detect the signal of environmental change in the population dynamics of long-lived plants (Doak and Morris 1999); to study weed-control strategies in agricultural systems (Mertens et al. 2002); to characterize the effect of environmental disasters (Monson et al. 2000); and to analyze the demography of unusual species (Gotelli 1991).

In a nutshell, PPMs are discrete-time, stage-structured models of population dynamics based on the demographic rates of individuals. PPMs divide a species' life history into distinct stages. Stages can be either components of an individual's life history (e.g., age classes or developmental stages) or partitions of a continuous variable, such as size. The element in row i and column j of a PPM is the average number of individuals in stage i contributed by an individual in stage j over the model's time step (Caswell 2001). Contributions are typically either transition rates between life stages of a single individual or reproduction rates (contributions to the youngest life-history stages). If initial stage-specific densities of a population are known, population dynamics can be projected by repeatedly multiplying the PPM by a vector of stage-specific densities. The eigenstructure of the PPM provides an abundance of useful information, such as the long-term population growth

rate (λ, the dominant eigenvalue of the matrix), the damping ratio (ρ, a measure of the rate at which transient dynamics decay, equal to the ratio of the magnitudes of the two largest eigenvalues), the long-term relative abundance and reproductive value of different stages (the right and left dominant eigenvectors of the matrix, respectively), and sensitivities of population growth to particular demographic transitions (Caswell 2001).

A drawback of PPMs is that they often encourage estimating too many parameters from too few data, particularly if each matrix element is considered a separate parameter to be estimated. This problem can be pronounced when PPMs are used for rare or endangered species where few individuals lead to few data. Overparameterization compromises PPMs by increasing the statistical variability in estimates of matrix elements. To counter overparameterization, matrix elements can be modeled with statistical models that have fewer fitted parameters. Indeed, some statistical models for matrix elements have already been proposed for PPMs based on partitions of a continuous state variable (e.g., Batista et al. 1998; Morris and Doak 2002). Importantly, however, estimating matrix elements with statistical models is not always preferable to estimating each matrix element separately. Inappropriate statistical models can produce bad estimates of matrix elements, despite fitting fewer parameters. Consequently, attempts to estimate matrix elements with statistical models must be coupled with a data-driven evaluation of those models. Model selection techniques provide the machinery for these evaluations. As a final step, multimodel averaging allows competing model outputs to be combined as a weighted average, with weights based on the outcome of the model selection (Burnham and Anderson 2002).

The goal of this paper is to outline, illustrate, and investigate through simulation a strategy for improving the estimation of PPMs by estimating matrix elements with one or more statistical models, arbitrating among these models with model selection, and constructing multimodel averages of the competing models (Burnham and Anderson 2002). Our goal is not to provide a one-size-fits-all recipe but to describe a framework that is flexible enough to be adapted to many situations. We chose information criteria as our basis for model selection (Akaike 1973) although other model selection methods, such as cross-validation, could be used instead. In short, information criteria measure the empirical support for competing models by estimating the average discrepancy between a fitted model and reality (Burnham and Anderson 2002). There are important subtleties in using information criteria to arbitrate among models for matrix elements, and we elaborate on these below. We include a simulation because the theoretical properties of information criteria have been established only for large samples (Akaike 1973; Hurvich and Tsai 1989; Burnham and Anderson 2002), and simulations are needed to examine their performance with smaller data sets.

The rest of this paper is structured as follows. We first introduce a motivating example with the cushion plant *Silene acaulis* (Caryophyllaceae). We then describe our general scheme for estimating, evaluating, and averaging models for matrix elements and apply it to the *Silene* data. The *Silene* example is followed by a simulation study that compares the statistical properties of PPMs with separately estimated matrix elements to PPMs based on reduced-parameter models and multimodel averages. A discussion closes the paper.

Motivating example

S. acaulis is a long-lived alpine cushion plant that is common in both alpine and arctic tundra habitats throughout the circumboreal zone (Hultén 1974). Plants are composed of tight aggregations of short-leaved branch tips (termed "rosettes") that hug the ground, making either rosette number or two-dimensional area good measures of plant size. Each cushion has a single taproot, and there is no clonal reproduction. *S. acaulis* is gynodioecious, but there are no differences in growth or survival rates of females and hermaphrodites in this population, so we follow Morris and Doak (1998) in averaging reproductive rates across the sexes and classifying individuals by size alone. More details about the biology of *Silene* and the data set we use here can be found in Morris and Doak (1998, 2005).

The data we consider here come from a survey of 549 plants at one location in the Wrangell Mountains in southeastern Alaska, USA (the Ridge site in Morris and Doak 1998, 2005). In 1998, plants were tagged and measured, and fruits were counted if present. One year later, surviving tagged plants were measured again. Plant size was measured in one of two ways. For plants with fewer than 20 rosettes, plant size was measured by counting the number of rosettes. Plants with more than 20 rosettes were photographed, and the two-dimensional cushion area was quantified with image analysis software. Survival, growth, and fruit production for these data are shown in Fig. 1. We divided the *Silene* into 12 stages—seeds in the seed bank, seedlings, four stages for plants classified by rosette number (1, 2–5, 6–10, and 11–20 rosettes), and six stages for the larger plants classified by two-dimensional area (<12.5, 12.5–25, 25–50, 50–100, 100–200, and >200 cm^2). Henceforth, we call plants classified by rosette number "small plants" and plants classified by two-dimensional area "large plants." We do not consider the issue of choosing stage boundaries here (see Vandermeer 1978; Moloney 1986).

Approach

In the most general terms, this approach entails proposing and fitting one or more statistical models for demographic transitions, using model selection to arbi-

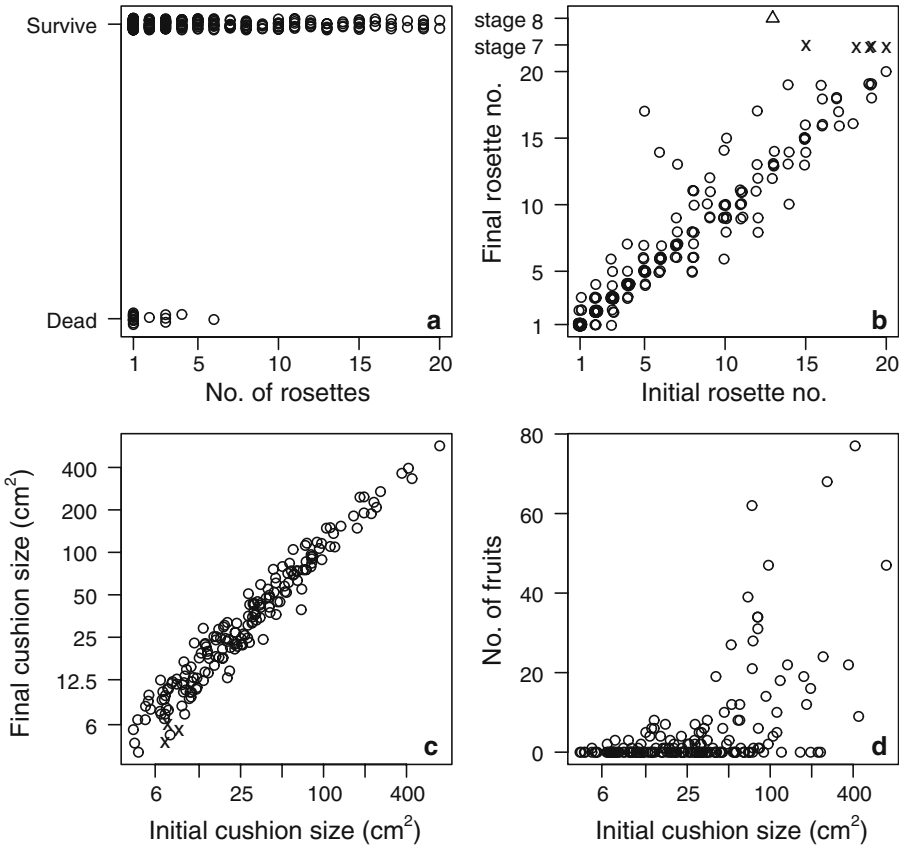

Fig. 1 Demographic data for *Silene acaulis*. **a** Survival of small plants. Points are jittered for clarity. **b** Final size of small plants (in number of rosettes) versus size 1 year earlier. Plants that grew to stages categorized by cushion area are shown with ×'s and Δ's. **c** Final size of large plants versus size 1 year earlier. Plants that shrank to stages categorized by rosette number are shown with ×'s. **d** Number of fruits produced versus size for large plants

trate among the various models, and then calculating multimodel-average estimates. The particular choice of statistical models will depend on both biological context and the type of data available, but separately estimated matrix elements should be included as one of the candidate models. We have found it convenient to decompose matrix elements into stage-specific vital rates—fecundities, survival rates, and growth rates conditional on survival—and to fit models to those vital rates instead of the matrix elements themselves. For example, a matrix element for transition to a larger size class is the product of two vital rates, the survival rate and the growth rate to the larger size class, given survival. We use the small-sample version of Akaike's information criteria, AIC_c (Hurvich and Tsai 1989; Burnham and Anderson 2002) for model selection, and we use Akaike weights based on AIC_c for model averaging (Burnham and Anderson 2002).

This approach can be applied to any PPM, regardless of how stages are defined. However, in the common case when stages are partitions of a continuous state variable, it is natural to estimate vital rates or matrix elements by fitting a model of smooth changes in a vital rate with state and then discretizing the fitted model to generate vital-rate estimates (e.g., Batista et al. 1998; Morris and Doak 2002). Because this situation is relatively common, and because using information criteria to arbitrate among models constructed in this way requires some care, we discuss this scenario in detail using the survival of small *Silene* plants to illustrate. The remainder of the *Silene* example and the simulation provide additional illustrations.

Consider estimating stage-specific survival rates for small *Silene* plants. To develop some notation, write the survival rate of stage i as s_i, $i=3,...,6$, write the number of individuals in stage i as n_i, and write the number that survive as y_i. Write the initial size (rosette number) of individual k in stage i as x_{ik}, $k=1,...,n_i$. The probability model implied by the projection matrix is that the y_i's are independent realizations of binomial random variables with parameters n_i and s_i, respectively, so that the kernel of the likelihood function for the s_i's is

$$\mathcal{L}(s_3,\ldots,s_6|n_3,\ldots,n_6,y_3,\ldots,y_6) \propto \prod_{i=3}^{6} s_i^{y_i}(1-s_i)^{n_i-y_i}. \quad (1)$$

To estimate the s_i's separately, we maximize Eq. 1, which gives the standard maximum likelihood estimates (MLEs) $\hat{s}_i = y_i/n_i$.

To estimate the s_i's by fitting fewer parameters, and to do so by leveraging the (presumably) smooth relationship between survival and size, we first posit a smooth survival versus size function. A sensible choice here is to model the logit of survival probability as a linear function of size

$$s(x;a_0,a_1)=\frac{\exp\{a_0+a_1x\}}{1+\exp\{a_0+a_1x\}}, \quad (2)$$

where $s(x;\ a_0,\ a_1)$ is the survival probability of a size x individual, and a_0 and a_1 are parameters to be estimated. To estimate a_0 and a_1, we need to average $s\ (x;\ a_0,\ a_1)$ over the distribution of sizes in each stage. That is to say, suppose $\phi_i(x)$ is the distribution of sizes in stage i and L_i and U_i are the lower and upper size limits of stage i. Then the average survival rate for individuals in stage i is

$$s_i(a_0,a_1)=\int_{L_i}^{U_i} s(x;a_0,a_1)\phi_i(x)\mathrm{d}x. \quad (3)$$

We choose to estimate $\phi_i(x)$ by the empirical size distribution observed in the data, and so plugging this into Eq. 3 gives

$$s_i(a_0,a_1)=\frac{1}{n_i}\sum_{k=1}^{n_i} s(x_{ik};a_0,a_1). \quad (4)$$

Plugging Eq. 4 into Eq. 1 gives a likelihood function for a_0 and a_1

$$\mathcal{L}(a_0,a_1|n_3,\ldots,n_6,y_3,\ldots,y_6) \propto \prod_{i=3}^{6}[s_i(a_0,a_1)]^{y_i}[1-s_i(a_0,a_1)]^{n_i-y_i} \quad (5)$$

and maximizing this likelihood gives MLEs for a_0 and a_1. Plugging these MLEs into Eq. 4 gives estimates for s_3,..., s_6 that are functions of two fitted parameters.

By evaluating the likelihoods in Eqs. 1 and 5 at their MLEs, the two parameters sets can be compared by $\mathrm{AIC_c}$. $\mathrm{AIC_c}$ is defined as

$$\mathrm{AIC_c}=-2\ \log\ \mathcal{L}\left(\hat{\theta}\right)+2K\left(\frac{n}{n-K-1}\right), \quad (6)$$

where $\mathcal{L}\left(\hat{\theta}\right)$ is the full likelihood (not just the kernel) evaluated at the MLE of the parameters θ, n is the number of data points, and K is the number of estimated parameters. The model with the smallest (most negative) value of $\mathrm{AIC_c}$ is deemed $\mathrm{AIC_c}$-best, and $\mathrm{AIC_c}$ differences ($\Delta\mathrm{AIC_c}$) measure the lack of empirical support for inferior models. Although there are no hard and fast rules for interpreting $\Delta\mathrm{AIC_c}$, Burnham and Anderson (2002, p. 70) suggest that models with $\Delta\mathrm{AIC_c}>10$ have negligible empirical support. $\Delta\mathrm{AIC_c}$ can also be used to calculate Akaike weights, w_i,

$$w_i=\frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_r \exp\left(-\frac{1}{2}\Delta_r\right)}, \quad (7)$$

where the sum is over the models under consideration, Δ_i is shorthand for $\Delta\mathrm{AIC_c}$ for model i, and $\Delta_i=0$ for the best fitting model (Burnham and Anderson 2002). The estimates from the competing models can then be averaged (using the w_i's as weights) to generate multimodel-average estimates. In the case of small *Silene* survival, the reduced-parameter estimates are $\mathrm{AIC_c}$-best, and $\Delta\mathrm{AIC_c}$ between the two models is 5.6. Consequently, the Akaike weights are 0.909 for the reduced-parameter model and 0.091 for the individually estimated vital rates.

The important point here is that in order for $\mathrm{AIC_c}$ to be used as a basis for model comparison, the parameter estimates need to be the MLEs with respect to the likelihood implied by the PPM. In this case, that means that a_0 and a_1 need to be estimated by maximizing Eq. 5, *not* by a simple logistic regression of the data in Fig. 1a. The logistic model in Eq. 2 is not the model that is being fit to the data because the PPM implies that all individuals in a stage have the same survival probability. Instead, Eq. 2 together with Eq. 4 is a statistical model that connects the stage-specific vital rates in neighboring stages, allowing those vital rates to be estimated using fewer fitted parameters.

Application to *Silene* data

Matrix elements for *Silene* seeds and seedlings were estimated from a separate data set, and we do not consider their estimation here. For small and large plants, we reparameterize nonreproduction matrix elements as the product of stage-specific survival rates and growth rates given survival. Thus, there are six sets of rates to consider: survival, growth, and reproduction rates for small and large plants.

Survival

We have already estimated survival rates for small plants. Mortality of plants with greater than 20 rosettes is extremely rare. From a larger study of *Silene* demography in the Wrangell Mountains (Morris and Doak 2005), only ten deaths were observed among 3,584 plant-years of observations. Using these data, we set survival for all large plant classes to 0.9972 in this study. For consistency across all stages, survival rates of small plants were also multiplied by 0.9972.

Growth

Among the small plants that survived, transitions to stages 3–8 were observed. The final stages of small plants that survived are independent realizations of multinomial random variables. Specifically, if n_i individuals begin in stage i and survive ($i=3,\ldots,6$), y_{ij} of these end in stage j ($j=3,\ldots,8$), and the conditional growth rates from stage i to stage j are g_{ij}, then the kernel of the likelihood for the g_{ij}'s is

$$\mathcal{L}(g_{33},\ldots,g_{68}|n_3,\ldots,n_6,y_{33},\ldots,y_{68}) \propto \prod_{i=3}^{6}\prod_{j=3}^{8} g_{ij}^{y_{ij}}. \quad (8)$$

Estimating each rate separately yields the MLEs $\hat{g}_{ij} = y_{ij}/n_i$.

We based our reduced-parameter model for the g_{ij}'s on a model of final stage versus initial size using a proportional odds cumulative logistic model (Agresti 2002). In notation, let $\gamma_j(x)$ be the probability that an individual with initial size x ends up in stage j, $j = 3,...,8$. The $\gamma_j(x)$'s determine the cumulative transition probabilities $\Gamma_j(x)$

$$\Gamma_j(x) = \sum_{k=3}^{j} \gamma_k(x). \tag{9}$$

The logits of the $\Gamma_j(x)$'s are linear functions of initial size, x:

$$\Gamma_j(x; b, z_3, \ldots, z_7) = \frac{\exp\{z_j + bx\}}{1 + \exp\{z_j + bx\}} \tag{10}$$

for $j = 3,...,7$ ($\Gamma_8(x) = 1$). Eqs. 9 and 10 are converted to average conditional growth rates by

$$g_{ij}(b, z_3, \ldots, z_7) = \frac{1}{n_i} \sum_{k=1}^{n_i} \gamma_j(x_{ik}; b, z_3, \ldots, z_7). \tag{11}$$

Plugging Eq. 11 into Eq. 8, maximizing the resulting likelihood, and plugging the MLEs into Eq. 11 gives a set of estimates for the g_{ij}'s as a function of six parameters.

To calculate the AIC_c of the separately estimated growth rates, we set the number of estimated parameters equal to the number of nonzero estimates, 13, although one could argue that the true number of estimated growth rates is 24. In either case, AIC_c decisively chooses the individually estimated growth rates. ΔAIC_c for the reduced-parameter model was 59.26 ($w \approx 0$), and so the multimodel-average growth rates were nearly identical to the individually estimated growth rates. The individually estimated growth rates were preferable here because data were abundant ($n = 348$), and the growth rates based on the proportional odds cumulative logistic model did not accommodate the relatively high rates of growth and shrinkage observed from stages 5–6.

For large plants, we illustrate a different method here that is based on modeling final size (instead of final stage) as a function of initial size. For the sake of modeling large plant growth rates, we converted the final sizes of the three (of 51) stage 7 plants that shrunk to stage 6 to a size in cushion area (Fig. 1c) and calculated growth rates as if their final stage had been stage 7. Afterward, we adjusted g_{76} and g_{77} accordingly.

The likelihood for the g_{ij}'s among large plants has the same form as Eq. 8, with $i, j = 7,...,12$. We consider two reduced-parameter models for the g_{ij}'s. In both models, log final size is modeled as a normal random variable, with mean and variance determined by functions of log initial size. Thus, if L_j and U_j are the lower and upper size limits to stage j ($U_{12} = \infty$), then the probability that an individual with initial size x ends up in stage j, given that it survives, is

$$\gamma_j(x) = \int_{\ln L_j}^{\ln U_j} f\left(y; \mu(x), \sigma^2(x)\right) \mathrm{d}y \tag{12}$$

where $f(\cdot\,|\mu, \sigma^2)$ is the probability density function of a normal random variable with mean μ and variance σ^2. In both models, the variance is a (piecewise) linear function of log initial size: $\sigma^2(x) = \max(d_0 + d_1 \ln x, 0)$. In the first model, $\mu(x)$ is a linear function of log initial size: $\mu(x) = c_0 + c_1 \ln x$, and in the second model, $\mu(x)$ is a quadratic function of log initial size: $\mu(x) = c_0 + c_1 \ln x + c_2(\ln x)^2$. To produce a likelihood function for these models, we plug the appropriate expressions for $\mu(x)$ and $\sigma^2(x)$ into Eq. 12, average Eq. 12 by an expression similar to Eq. 11, and plug the resulting g_{ij}'s into a likelihood similar to Eq. 8. This yields estimates for the g_{ij}'s based on four or five fitted parameters.

As with small plants, we calculated AIC_c for the separately estimated growth rates by counting the number of estimated parameters as the number of nonzero estimates (15 instead of 36). AIC_c favored the model based on a linear relationship between mean log final size and log initial size. ΔAIC_c's for the quadratic model and the separately estimated g_{ij}'s were 1.69 and 17.24, respectively, giving Akaike weights of 0.700, 0.300, and ≈ 0 for the linear, quadratic, and separately estimated models. The reduced-parameter models outperformed the separately estimated growth rates here because few plants were observed in the largest stages ($n = 11$, 8 for stages 11 and 12, respectively), and the reduced parameter models fit the data well.

Fecundity

Because *Silene* surveys were performed before seeds were dropped, matrix elements for reproduction combine fruit production, seed survival, and germination. The latter two quantities were estimated from a separate study (3.75% of fruits survive to the next survey without germinating, and an additional 0.78% both survive and germinate), and so the task here is to estimate average fruit production for each stage, $f_3,...,f_{12}$. Among small plants, only a few plants in stage 6 set fruit, so we set $\hat{f}_3 = \hat{f}_4 = \hat{f}_5 = 0$ and estimate f_6 by the average number of fruits per stage 6 plant. For large plants, we assume that the number of fruits per plant is Poisson distributed, so if there are a total of y_i fruits produced by the n_i plants in stage i, $i = 7,...,12$, the kernel of the likelihood function is

$$\mathcal{L}(f_7, \ldots, f_{12} | n_7, \ldots, n_{12}, y_7, \ldots, y_{12}) \propto \prod_{i=7}^{12} \mathrm{e}^{-n_i f_i} f_i^{y_i}. \tag{13}$$

Estimating each average fecundity separately, the MLEs are simply $\hat{f}_i = y_i/n_i$.

We fit two reduced-parameter models for f_i's: one that models the log of mean fruit production as a linear function of the log initial size, and a second that models the log of mean fruit production as a quadratic function of the log initial size. For the first model, if $f(x)$ is the average fruit production for a plant of size x, then

$$f(x; h_0, h_1) = \exp\{h_0 + h_1 \ln x\}. \quad (14)$$

The quadratic model is simply $f(x; h_0, h_1, h_2) = \exp\left\{h_0 + h_1 \ln x + h_2(\ln x)^2\right\}$. Eq. 14 is averaged across stages by an averaging similar to Eq. 4 or Eq. 11, the resulting averages are plugged into Eq. 13, and the likelihood is maximized.

For the *Silene* data, AIC_c favors the separately estimated f_i's decisively. ΔAIC_c's for the linear and quadratic models were 165 and 110, respectively. Thus, the multimodel estimates of the f_i's equal the separately estimated f_i's. Separately estimated fecundity estimates were favored because both reduced parameter models were unable to accommodate the small observed fecundity in stage 11 relative to stages 10 and 12 (Fig. 1d).

Silene summary

The population projection matrix produced by estimating each vital rate separately is shown in Table 1, and the matrix produced by multimodel averaging is shown in Table 2. The main difference between the two matrices involves the shrinkage rates for large plants (matrix elements immediately above the main diagonal for stages 7–12). Estimating matrix elements separately produces shrinkage rates that vary markedly between neighboring stages while the multimodel-average PPM yields more gradual changes in shrinkage rates. In this case, the data were not informative enough to support separate estimates of shrinkage rates for each large plant stage, and the multimodel average PPM adjusted accordingly by smoothing the shrinkage rates across stages.

For the remaining matrix elements, the multimodel average estimates were similar to the separately estimated elements for several reasons. For small plant survival, the reduced-parameter and separately estimated estimates were similar. For small plant growth, there were enough data to justify estimating each rate separately. For large plant fecundity, the proposed reduced-parameter models did not fit the data well.

The differences in shrinkage rates did not impact the estimate of λ, the long-term population multiplication rate—the PPM in Table 1 yields $\lambda = 1.0047$, and the PPM in Table 2 gives $\lambda = 1.0048$. Instead, the main impact of the smoothed shrinkage rates appeared in the stable stage distribution (SSD, Fig. 2). Most notably, the PPM with separately estimated elements predicts a jump in the SSD between stages 10 and 11 while the

Table 1 Population projection matrix for *Silene acaulis* with each matrix element estimated separately

Stage	1	2	3	4	5	6	7	8	9	10	11	12
1	0.320	0	0	0	0	0.003	0.014	0.050	0.078	0.460	0.362	1.158
2	0.067	0	0	0	0	0.001	0.003	0.011	0.016	0.096	0.075	0.241
3	0	0.874	0.880	0.026	0	0	0	0	0	0	0	0
4	0	0	0.026	0.875	0.098	0	0	0	0	0	0	0
5	0	0	0	0.044	0.724	0.140	0	0	0	0	0	0
6	0	0	0	0.009	0.156	0.738	0.059	0	0	0	0	0
7	0	0	0	0	0	0.100	0.763	0.026	0	0	0	0
8	0	0	0	0	0	0.020	0.176	0.709	0.146	0	0	0
9	0	0	0	0	0	0	0	0.262	0.681	0.031	0	0
10	0	0	0	0	0	0	0	0	0.170	0.779	0	0
11	0	0	0	0	0	0	0	0	0	0.187	0.816	0.125
12	0	0	0	0	0	0	0	0	0	0	0.181	0.873

Table 2 Projection matrix for *Silene acaulis* estimated by multimodel averaging

Stage	1	2	3	4	5	6	7	8	9	10	11	12
1	0.320	0	0	0	0	0.003	0.014	0.050	0.078	0.460	0.362	1.158
2	0.067	0	0	0	0	0.001	0.003	0.011	0.016	0.096	0.075	0.241
3	0	0.874	0.883	0.026	0	0	0	0	0	0	0	0
4	0	0	0.026	0.869	0.099	0	0	0	0	0	0	0
5	0	0	0	0.043	0.729	0.139	0	0	0	0	0	0
6	0	0	0	0.009	0.158	0.737	0.057	0	0	0	0	0
7	0	0	0	0	0	0.100	0.743	0.063	0	0	0	0
8	0	0	0	0	0	0.020	0.196	0.708	0.076	0	0	0
9	0	0	0	0	0	0	0.001	0.225	0.749	0.069	0	0
10	0	0	0	0	0	0	0	0.001	0.172	0.750	0.078	0
11	0	0	0	0	0	0	0	0	0	0.177	0.700	0.066
12	0	0	0	0	0	0	0	0	0	0	0.219	0.932

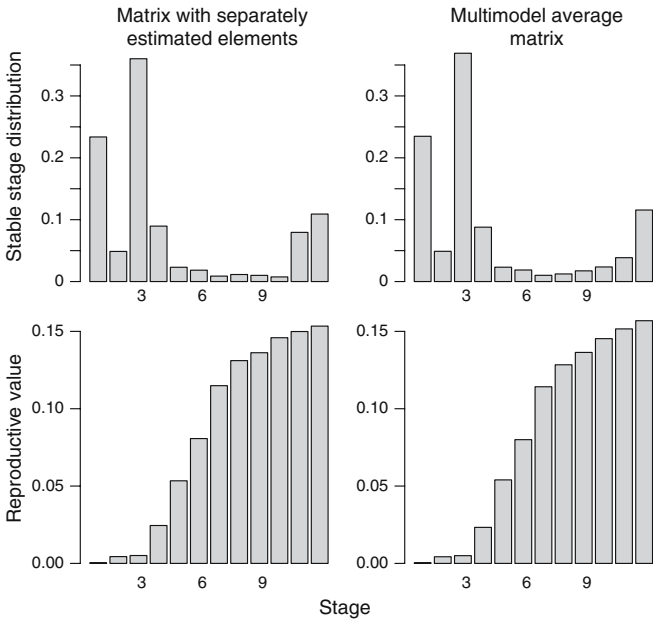

Fig. 2 Stable-stage distributions (*top*) and reproductive values (*bottom*) for population projection matrices for *Silene acaulis*. *Panels on the left* are from the projection matrix with separately estimated elements, and *panels on the right* are from the projection matrix estimated by multimodel averaging

multimodel-average matrix predicts a smoother increase in the SSD from stages 7–12.

These differences in SSD have substantial impacts on an elasticity analysis of the PPM. The elasticity of a matrix element or vital rate is the partial derivative of log λ with respect to the log of that matrix element or vital rate, and is a scaled measure of how a change in the matrix element or vital rate will affect population growth (Caswell 2001). Elasticity analyses are used to identify targets for conservation management strategies (Caswell 2000). Although *Silene* is not a species of conservation concern, we calculate elasticities here as an illustration of how they can differ between PPMs estimated from the same data set. For the matrix with separately estimated elements, the elasticities of large plant survival rates (s_7–s_{12}) are 0.02, 0.04, 0.03, 0.03, 0.29, and 0.41. For the multimodel-average matrix, these elasticities are 0.03, 0.04, 0.06, 0.09, 0.15, and 0.45. Thus, if we assume that the multimodel-average matrix is superior (see the simulation below), a matrix with separately estimated rates overestimates the importance of the survival of plants in the second largest stage and underestimates the importance of other large plants' survival.

Simulation study

The *Silene* example illustrates that a PPM can be estimated by averaging models for matrix elements, and it suggests that multimodel-averaged PPMs may differ from PPMs with separately estimated elements in subtle ways that may be important for management. However, it is difficult to draw any conclusive comparisons between estimation methods on the basis of a single data set where the "true" vital rates are unknown. Thus, we conducted a simulation study to compare the statistical properties of multimodel-average PPMs to PPMs with separately estimated elements. This simulation is based on the *Silene* example but does not match it exactly. In the simulation, we have made some of the true vital rates piecewise constant functions that change at stage boundaries. We have done this not for biological reality, but instead to "stack the deck" in favor of separately estimated vital rates in some components of the matrix. If multimodel-average matrices outperform matrices with rates estimated separately here, we can be confident that they will perform at least as well for real species where it is unlikely that vital rates will be piecewise constant functions that change exactly at size–class boundaries.

The simulation envisions a population of plants that are grouped into seed and seedling stages plus five size classes: 0–10, 10–20, 20–40, 40–80, and >80 (units are arbitrary). Forty percent of seeds in the seed bank survive each year but do not germinate, 20% germinate to seedlings, and 75% of seedlings survive and grow to the smallest size class. Seeds do not germinate in the year that they are produced. Survival probabilities are 75%, 80%, 85%, 90%, or 95% for plants in size classes 1–5, respectively. For those plants that survive, the log of plant size at the next survey is a normally distributed random variable with mean $0.5 + x - 0.025x^2$, where x is the log size at the initial survey. The standard deviation of log size at the next survey is 0.40, 0.35, 0.30, 0.25, and 0.20 for plants in size classes 1–5, respectively. Finally, the number of viable seeds produced by a plant (contributions to the following survey's seed bank) is a Poisson random variable with mean 2, 4, 6, or 8 for plants in size classes 2–5.

When this population is at its SSD, the average transition rates among stages are

$$A = \begin{bmatrix} 0.4 & 0 & 0 & 2 & 4 & 6 & 8 \\ 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.75 & 0.713 & 0.036 & 0 & 0 & 0 \\ 0 & 0 & 0.033 & 0.390 & 0.041 & 0 & 0 \\ 0 & 0 & 0.004 & 0.349 & 0.499 & 0.053 & 0 \\ 0 & 0 & 0 & 0.025 & 0.303 & 0.643 & 0.175 \\ 0 & 0 & 0 & 0 & 0.007 & 0.204 & 0.755 \end{bmatrix}. \quad (15)$$

In our simulation, we assumed that individuals were sampled from a population at the SSD, and so the matrix in Eq. 15 is the matrix that our simulations seek to estimate. λ for this matrix is 1.021.

We consider the seed and seedling transition rates (the leftmost two columns of A) to be estimated from other data and only estimate the transitions from the five size classes. We simulated 500 data sets each for $n = 50$, $n = 100$, and $n = 200$ individuals from these size classes, with each data set containing at least ten individuals

from each class and at least one individual growing from each size class. For each data set, we estimated stage-specific survival rates (s_i, $i = 3,...,7$), growth rates from stage i to stage j conditional on survival (g_{ij}, $i, j = 3,...,7$), and average fecundities (f_i, $i = 3,...,7$). All three series of vital rates were estimated both individually and by reduced-parameter models similar to those used in the *Silene* example. Survival rates were estimated by the logistic-based model in Eqs. 2 and 4, and growth rates were estimated by the model in Eq. 12 with log mean final size modeled as a linear function of log initial size. We did not model log mean final size as a quadratic function of log initial size because it would be unlikely in practice to choose a model that matched the data-generating mechanism so well. Our model for fecundity rates was based on a (piecewise) linear relationship between average fruit production and initial size, $f(x;\ h_0,\ h_1) = \max(h_0 + h_1\ x,\ 0)$, instead of the loglinear relationship in Eq. 14.

For each data set, we constructed PPMs from the separately estimated vital rates and from the multimodel-average vital rates. As we did with *Silene*, we counted the number of estimated parameters for separately estimated growth rates as the number of nonzero estimates. For illustration, we also constructed a PPM using only the reduced-parameter vital-rate models. We calculated $\hat{\lambda}$, the damping ratio ($\hat{\rho}$), the SSD ($\hat{\mathbf{w}}$), and reproductive values ($\hat{\mathbf{v}}$) for each PPM and calculated mean squared errors (MSEs) of all quantities of interest for each sample size.

Results

In nearly all the simulations with $n = 100$ and $n = 200$ and in most of the simulations with $n = 50$, the AIC_c-best estimates were the separately estimated fecundity rates and the survival and growth rates estimated with reduced-parameter models. ΔAIC_c's were usually large enough that the multimodel estimates were almost equal to the AIC_c-best estimates. For nonreproduction matrix elements, MSEs were generally smaller for multimodel-average matrices than for matrices with separately estimated vital rates (Fig. 3). In contrast, the reduced parameter model chosen for fecundity estimates was

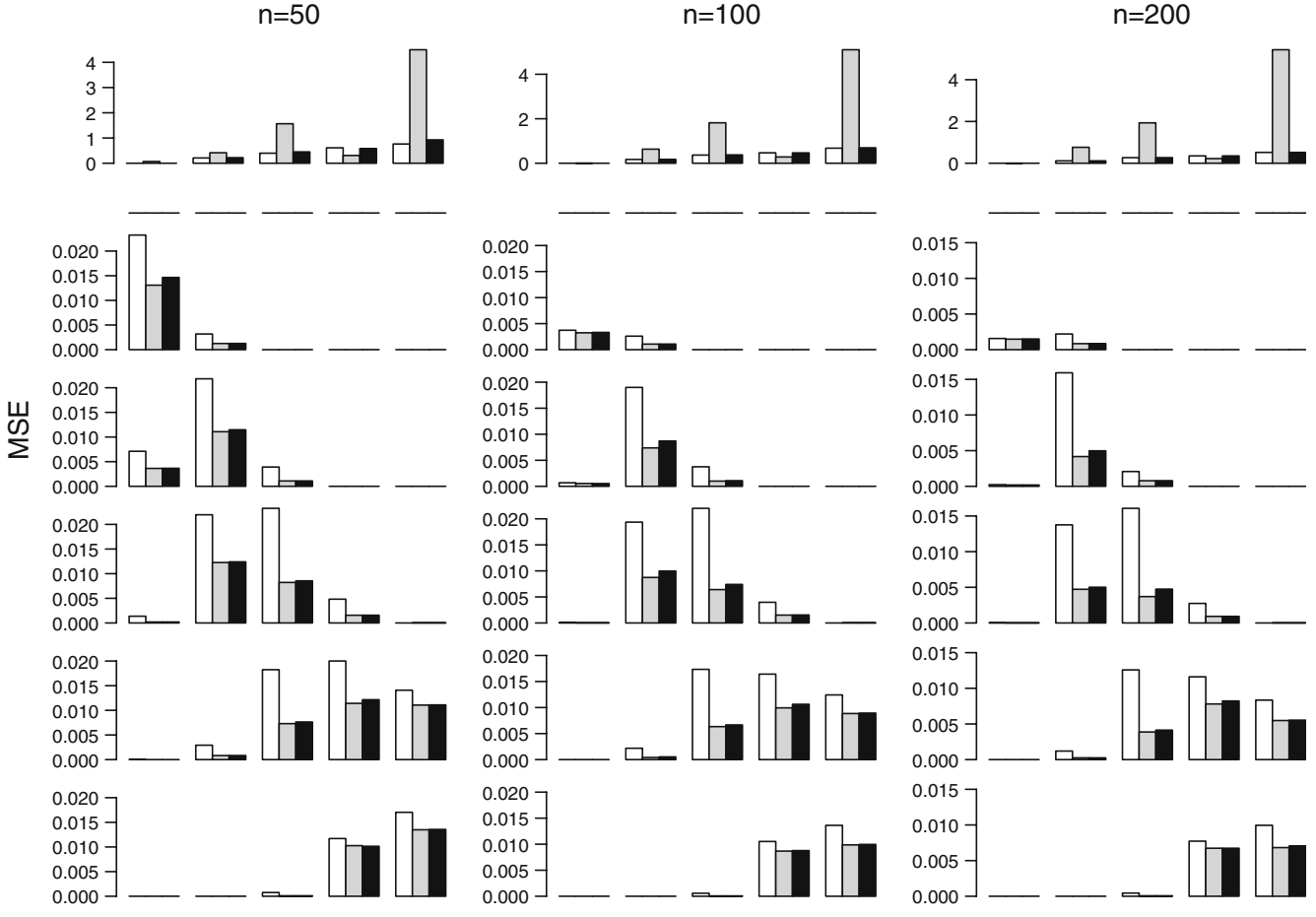

Fig. 3 Mean squared errors for estimates of matrix elements from the simulation study. *Bar graphs* are positioned as they would be in the matrix: *top row* is for reproduction estimates, etc. Only columns 3–7 of the projection matrix are shown. Within each bar graph, the bars are mean squared errors (MSEs) for (*left to right*): separately estimated vital rates (*white*), reduced-parameter vital-rate models (*gray*), and multimodel averages (*black*)

poor, and so multimodel averaging did not reduce the MSE of matrix elements for reproduction.

When $n = 50$, multimodel-average PPMs reduced the MSE of $\hat{\lambda}$ by 25% compared with PPMs with individually estimated elements, but the differences in MSEs were negligible when $n = 100$ and $n = 200$ (Fig. 4). For all three sample sizes, the damping ratio was estimated more precisely by the multimodel-averaged matrix than the matrix with individually estimated rates. For the SSD and reproductive values, improvements in MSEs from multimodel-average matrices varied, ranging from large (e.g., $\hat{v}_4 - \hat{v}_7$) to slight or none (e.g., $\hat{v}_1 - \hat{v}_3$), and in some cases, MSEs from multimodel-average matrices were larger (e.g., $\hat{w}_7$; Fig. 4). In the large majority of cases, however, multimodel-average PPMs produced more precise estimates than matrices with separately estimated rates. MSEs of sensitivities and elasticities for matrix elements were also generally smaller for multimodel-average matrices (data not shown). The specific results observed here—multimodel-average PPMs reduced MSEs for nonreproduction matrix elements but not for reproduction-based elements—are probably not general but instead are attributable to the data-generating models and reduced-parameter models chosen. However, we suspect that the broad pattern—multimodel-average matrices do not reduce MSE everywhere but on the whole provide better estimates than matrices built from separately estimated vital rates or from reduced-parameter models—is general. These results underscore the benefit of combining models that reduce the number of fitted parameters in PPMs with model selection techniques that evaluate the models chosen.

Discussion

Population projection matrices are widely used by ecologists and wildlife managers to analyze or project population dynamics on the basis of transition rates among life history stages. From a statistical perspective, however, PPMs invite overparameterization, especially when the species being analyzed is rare or endangered and consequently data are scarce. Here, we have out-

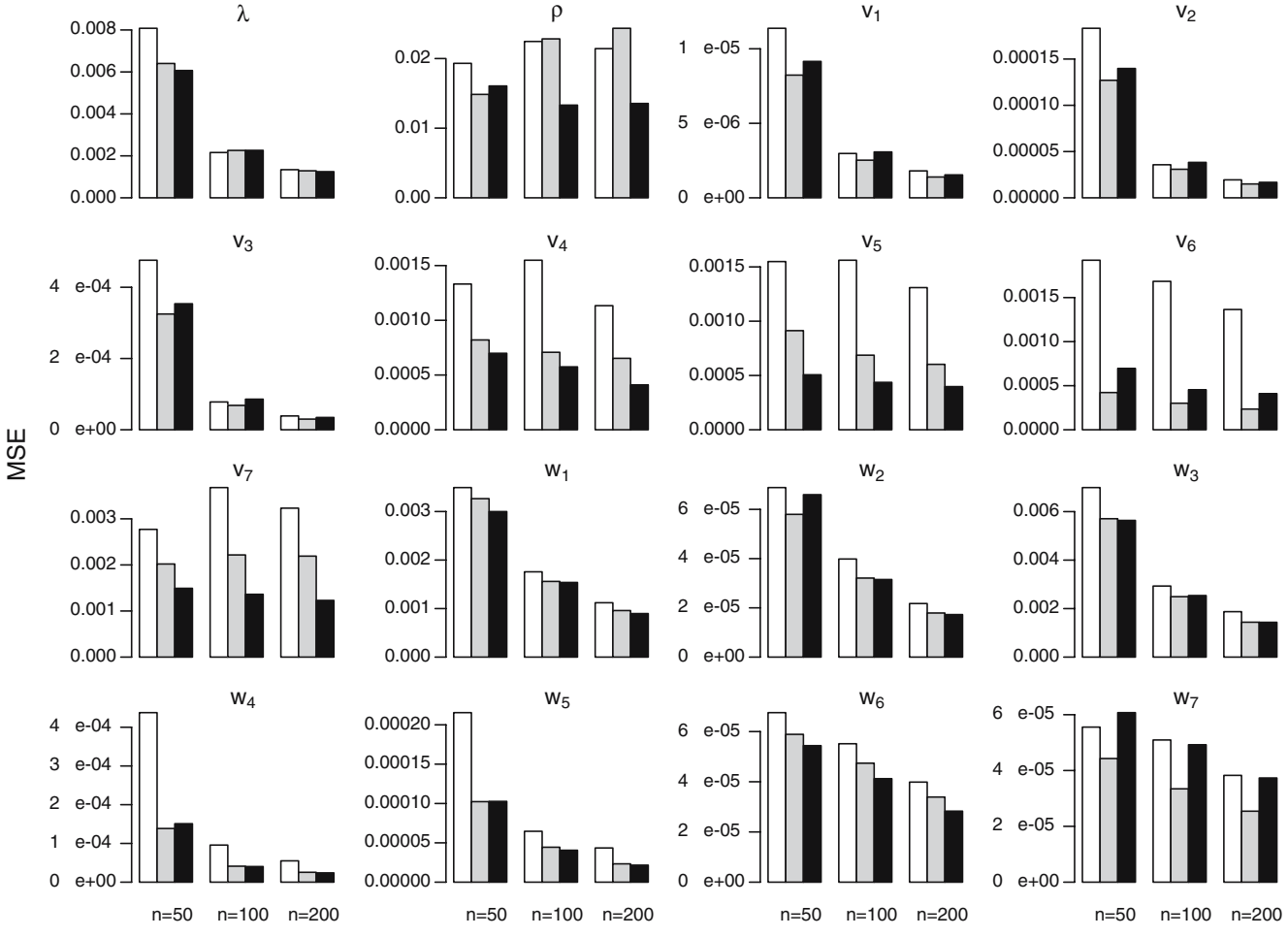

Fig. 4 Mean squared errors (MSEs) of $\hat{\lambda}$, damping ratios ($\hat{\rho}$), reproductive values ($\hat{\mathbf{v}}$) and the stable-stage distribution ($\hat{\mathbf{w}}$) from the simulation study. Within each bar graph, the *bars* are MSEs for (*left to right*): separately estimated vital rates (*white*), reduced-parameter vital-rate models (*gray*), and multimodel averages (*black*)

lined, applied, and investigated a scheme for avoiding overparameterization in PPMs. This scheme involves proposing and fitting models for matrix elements (or the stage-specific vital rates that determine matrix elements), evaluating these models and separately estimated elements by model selection, and averaging the competing estimates with multimodel averaging. We have used an information criterion, AIC_c, for model selection although other model selection methods could be used. When an information criterion is used, care must be taken to ensure that the likelihood on which the criterion is based (and the one maximized to estimate parameters) is the likelihood implied by the PPM.

Our simulation results suggest that multimodel-average matrices will only affect the estimate of λ when the total sample size is small. On the other hand, estimates of most other demographic measures such as sensitivities, elasticities, and damping ratios will be more precise with multimodel-average PPMs, even with relatively large data sets. We speculate that this is so because λ is an aggregate measure of the data (and thus is somewhat robust to possible matrix overparameterization) while other demographic measures depend more heavily on individual matrix elements. When matrix elements are estimated separately, the quality of individual estimates depends not just on total sample size but on the distribution of data among stages. If sample sizes differ among stages, then this data imbalance can compromise the quality of separately estimated matrix elements (and hence sensitivity and elasticity analyses) even when total sample size is large. The method we propose here provides a safeguard against estimating more matrix parameters than the data support. Thus, we conclude that multimodel averaging will have the most benefit either when PPMs are used to conduct more detailed demographic analyses than simply calculating λ or when the availability of individuals makes data more scarce for some matrix stages than for others.

In the examples we have considered here, some stages are partitions of a continuous state variable, and reduced-parameter vital-rate models are based on smooth rate versus size relationships. In this setting, an alternative modeling option is an integral projection model (IPMs) that does not divide individuals into stages (Easterling et al. 2000; Rees and Rose 2002). Our point here is not to argue against IPMs. However, we anticipate that situations will continue to arise in which ecologists choose to use matrix models, and if matrix models are to be used, their elements should be estimated well. We think the procedures described here contribute to that end.

Lastly, we emphasize that the general scheme here is not limited to cases where stages are partitions of a continuous state variable or where individuals are marked and observed in consecutive surveys. The general idea is flexible enough to accommodate a variety of settings. For example, when stages are levels of a categorical state variable (e.g., egg, larva, pupa, adult), alternative reduced-parameter models that equate vital rates in neighboring stages may be appropriate [cf. §5.2 of Burnham and Anderson (2002)]. We have conducted a second simulation study for this scenario and obtained similar results (data not shown)—multimodel-average matrices only slightly reduced the MSE of $\hat{\lambda}$, but they substantially reduced the MSEs of some matrix elements, the stable stage distribution, reproductive values, sensitivities, and elasticities. Also, we have only considered cases in which the probability of being unable to relocate a tagged individual is negligible. In other settings, such as capture–recapture studies of animal populations, this may not be the case, and methods for estimating PPMs need to account for capture histories (Nichols et al. 1992; Fujiwara and Caswell 2001). Although we do not pursue this here, we speculate that consideration of reduced-parameter models for elements of PPMs could be fruitfully combined with methods for capture–recapture data as well (cf. Fujiwara and Caswell 2001).

Acknowledgements We thank Hal Caswell, Massa Nakaoka, and an anonymous reviewer for helpful comments and discussion. This work was supported by NSF grant DEB-0087096 to WFM and DFD.

References

Agresti A (2002) Categorical data analysis, 2nd edn. Wiley-Interscience, New York

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petran B and Csáki F (eds) International symposium on information theory. Akadémiai Kiadi, Budapest, pp 267–281

Batista WB, Platt WJ, Macchiavelli RE (1998) Demography of a shade-tolerant tree (*Fagus grandifolia*) in a hurricane-disturbed forest. Ecology 79:38–53

Burnham KP, Anderson DR (2002) Model selection and multimodel inference, 2nd edn. Springer, Berlin Heidelberg New York

Caswell H (2000) Prospective and retrospective perturbation analyses: their roles in conservation biology. Ecology 81:619–627

Caswell H (2001) Matrix population models, 2nd edn. Sinauer, Sunderland

Crowder LB, Crouse DT, Heppell SS, Martin TH (1994) Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. Ecol Appl 4:437–445

Doak DF, Morris WF (1999) Detecting population-level consequences of ongoing environmental change without long-term monitoring. Ecology 80:1537–1551

Easterling MR, Ellner SP, Dixon PM (2000) Size-specific sensitivity: applying a new structured population model. Ecology 81:694–708

Fujiwara M, Caswell H (2001) Demography of the endangered North Atlantic right whale. Nature 414:537–541

Gotelli NJ (1991) Demographic models for *Leptogorgia virgulata*, a shallow-water gorgonian. Ecology 72:457–467

Hultén E (1974) Flora of Alaska and neighboring territories. Stanford University Press, Stanford

Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. Biometrika 76:297–307

Lande R (1988) Demographic models of the northern spotted owl (*Strix occidentalis caurina*). Oecologia 75:601–607

Mertens SK, van den Bosch F, Heesterbeek JAP (2002) Weed populations and crop rotations: Exploring dynamics of a structured periodic system. Ecol Appl 12:1125–1141

Moloney KA (1986) A generalized algorithm for determining category size. Oecologia 69:176–180

Monson DH, Doak DF, Ballachey BE, Johnson A, Bodkin JL (2000) Long-term impacts of the Exxon Valdez oil spill on sea otters, assessed through age-dependent mortality patterns. Proc Nat Acad Sci USA 97:6562–6567

Morris WF, Doak DF (1998) Life history of the long-lived gynodioecious cushion plant *Silene acaulis* (Caryophyllaceae), inferred from size-based population projection matrices. Am J Bot 85:784–793

Morris WF, Doak DF (2002) Quantitative conservation biology. Sinauer, Sunderland

Morris WF, Doak DF (2005) How general are the determinants of the stochastic population growth rate across nearby sites? Ecol Monogr 75:119–137

Nichols JD, Sauer JR, Pollock KH, Hestbeck JB (1992) Estimating transition probabilities for stage-based population projection matrices using capture-recapture data. Ecology 73:306–312

Rees M, Rose KE (2002) Evolution of flowering strategies in *Oenothera glazioviana*: an integral projection model approach. Proc Roy Soc Lond B 269:1509–1515

Tuljapurkar S, Caswell H (1997) Structured-population models in marine, terrestrial, and freshwater systems. Chapman and Hall, New York

Vandermeer J (1978) Choosing category size in a stage projection matrix. Oecologia 32:79–84

Wielgus RB (2002) Minimum viable population and reserve sizes for naturally regulated grizzly bears in British Columbia. Biol Cons 106:381–388