EURING PROCEEDINGS

Incorporating 'recruitment' in matrix projection models: estimation, parameters, and the influence of model structure

Evan G. Cooch · Emmanuelle Cam · Hal Caswell

Received: 6 September 2009/Revised: 9 July 2010/Accepted: 2 August 2010/Published online: 19 August 2010 © Dt. Ornithologen-Gesellschaft e.V. 2010

Abstract Advances in the estimation of population parameters using encounter data from marked individuals have made it possible to include estimates of the probability of recruitment in population projection models. However, the projected growth rate of the population, and the sensitivity of projected growth to changes in recruitment, can vary significantly depending upon both the structural form of the model and how recruitment is parameterized. We show that the common practices of (1) collapsing some age classes into a single, terminal 'aggregated' age-class, and (2) parameterizing recruitment using the proportion of recruited individuals (breeders) in a given age-class may confound analysis of age-based (Leslie) matrix projection models in some instances, relative to state-based projection models where recruited and pre-recruited individuals are treated as separate states. Failing to account for these differences can lead to misinterpretation of the relative role of recruitment in the dynamics of an age-structured population. We show that such problems can be avoided, either by structural changes to the terminal aggregated age-class in age-based models, or by using using a state-based model instead. Since all the

Communicated by M. Schaub.

E. G. Cooch (⊠) Department of Natural Resources, Cornell University, Ithaca, NY, USA e-mail: evan.cooch@cornell.edu

E. Cam Evolution et Diversité Biologique, Université Paul Sabatier, Toulouse, France

H. Caswell Biology, Woods Hole Oceanographic Institution, Woods Hole, MA, USA metrics of general interest from a classical age-based matrix models are readily derived from a state-based model equivalent, this suggests there may be little reason to use the classical age-based approach in situations where recruitment is a parameter of interest.

Keywords Matrix projection models · Perturbation analysis · Recruitment

Introduction

It is well known that the projected growth rate of an agestructured population reflects the timing (schedule) of reproduction across various age classes, except in stationary populations where the net reproductive rate adequately characterizes growth (Caswell 2001). Thus, the probability of making a permanent life-history transition from a prerecruit (an individual that has not bred) to a recruit (an individual that has bred at least once) at a given age (referred to hereafter as 'recruitment') is of fundamental importance to life history theory (Stearns 1992; Caswell 2001, and references therein), and population dynamics in general (Caswell 2001).

Many early attempts to estimate recruitment in the field relied on ad hoc estimation methods, which may suffer significant bias in most cases. For example, the use of return rate (proportion of a number of individuals marked as offspring which are observed breeding for the first time at some later age) to assess recruitment implicitly assumes that individuals of all ages have the same encounter probability, and that the first encounter of an individual provides an unbiased estimate of the age of recruitment.

However, an individual observed breeding at age *i* could have bred previously at an earlier age, and simply not have

been observed. Recent advances in the analysis of capture– mark–encounter data have provided a robust framework for estimation of and analysis of variation in recruitment while controlling for possible differences in encounter probability and imperfect detection of breeding individuals (Clobert et al. 1994; Nichols and Kendall 1995; Pradel and Lebreton 1999; Schwarz and Arnason 2000; Williams et al. 2002; Crespin et al. 2006; Jenouvrier et al. 2008; Nevoux et al. 2010; see Lebreton et al. 2009 for a recent review). Recent work by Cam et al. (2005) considered the influence of assumptions concerning equivalence of survival probability of pre-breeders and breeders if pre-breeders are unobservable (Clobert et al. 1994; Pradel and Lebreton 1999).

Despite advances in our ability to estimate recruitment, questions remain concerning the most appropriate measure of recruitment to use in population models, reflecting differences in how recruitment is defined. While earlier discussions by Frederiksen and Pradel (2001) and Schwarz and Arnason (2001) clarified some issues, we believe several questions remain. In particular, they did not address the possible interaction between model structure and how recruitment is defined and entered into population models.

Here, we briefly review common measures of recruitment. We show that analysis of the influence of recruitment on population growth can be significantly affected by the choice of recruitment parameter, and the manner in which recruitment is entered into the population projection model. We show that, in general, recruitment should be defined as the probability that an as yet inexperienced (pre-breeding) individual of a given age starts to breed at that age, and that recruitment is most conveniently, and robustly, analyzed using a multi-stage projection matrix including separate stages for pre-recruit and recruited individuals, rather than the more commonly used age-based (Leslie) matrix models, which assume that all individuals of a given age are identical and have a common reproductive rate.

Measures of recruitment

We define the 'probability of recruitment' as the probability (age-specific) of making a one-time, permanent state transition from pre-recruit (an individual that has not bred) to a recruited or 'breeding' individual. Recruited individuals are those which have bred at least once. We distinguish 'recruitment' from subsequent 'breeding propensity', which we define as the probability that, given that you have bred before, that you breed in a given year. We do not consider the parametrization and modeling of breeding propensity here.

Pradel and Lebreton (1999) and Frederiksen and Pradel (2001) considered 3 related parameters (a, α and π) which together can be used to describe recruitment. Each of these parameters can be estimated using encounter data from a sample of individuals marked as offspring.

 a_i , the probability that an as yet inexperienced individual of age *i* starts to breed at that age, given in general by

$$a_{i} = \frac{(1 - \gamma_{i}) \sum_{j=1}^{i-1} a_{j} \prod_{k=0}^{j-1} (1 - a_{k})}{\gamma_{i} \prod_{j=1}^{i-1} (1 - a_{j})}$$
(1)

where γ_i is the probability that an individual alive and breeding in the population at time (age) *i* was also alive and in the population at time *i* – 1 (Pradel 1996). In addition, under some assumptions (Pradel and Lebreton 1999), this parameter a_i is strictly equivalent to the parameter ψ_i^{xy} derived using multi-stage approaches (where ψ_i^{xy} is the probability of moving from stage *x* to stage *y* over the interval from *i* to *i* + 1; sensu Brownie et al. 1993). a_i is the ratio of the number of first time breeders to the sum of first time breeders and prebreeders at age *i* (Pradel and Lebreton 1999).

 α_i , the probability that an individual of age *i* is a first-time breeder, where

$$\alpha_i = a_i \prod_{y \le j < i} (1 - a_j), \ i \le f$$
(2)

 α_i is the proportion of first-time breeders among all individuals alive at age *i* (Pradel and Lebreton 1999), where *y* is the youngest breeding age, and *f* is the age of full-breeding (i.e. where $\alpha_f = 1$), and assuming equal survival of recruits and pre-recruits at the same age.

 π_i , the probability that an individual of age *i* is a breeder (i.e. the proportion of all animals of age *i* that are breeders) (Frederiksen and Pradel 2001), where

$$\pi_i = \sum_{i=y} \alpha_i \tag{3}$$

where *y* is the youngest breeding age.

Clearly, a_i is equivalent to recruitment as we have defined it. However, age-based (Leslie) projection models are parameterized using net (average) contributions from one age class to another. In such models, recruitment is a component of the net fertility contributions to the offspring age class. In such cases, the appropriate recruitment parameter to include in the model is the age-specific 'breeding proportion'; i.e. the proportion of individuals of a given age class which breed, and contribute to the offspring class.

Schwarz and Arnason (2000) suggested that the probability that an animal that (1) survives until it starts breeding will (2) do so (i.e. start breeding) at age j + 1 (their parameter b_{ij}) be referred to as 'age-specific breeding proportion' (or probability). However, Frederiksen and Pradel (2001) showed that b_{ij} is in fact a separate population parameter, and is not appropriately defined as 'age-specific breeding proportion'. They suggested that the term 'agespecific breeding proportion' should be applied to π_i (Eq. 3), and that it is π_i that is the relevant lower-level element of the fertility term in age-based projection matrix models (Frederiksen and Pradel 2001; Schwarz and Arnason 2001). We note that this implies that recruited individuals breed every year (i.e. breeding propensity is 1).

Recruitment and population growth

We introduce some potential issues in parameterizing recruitment in population projection models by means of a series of simple numerical examples. We assume throughout that recruitment is in situ, and that the number of breeding individuals is unaffected by asymmetric immigration and emigration. While it is possible to accommodate such movement in matrix projection models (e.g., Cooch et al. 2001), we do not address this issue here. We also assume equal survival among pre-recruits and recruits of the same age (cf. Cam et al. 2005). We begin by considering the classical age-based projection model assumed by Frederiksen and Pradel (2001) and Schwarz and Arnason (2001).

Age-based matrix model

Consider a population where the earliest age of breeding y is 2 years of age (i.e. no recruitment prior to 2 years of age). Fecundity B increases monotonically from age 2 to age 4, and is independent of age for all individuals 4 years and older (age of full-breeding, f = 4 years). We make two assumptions concerning recruitment. First, we assume that all individuals aged $\geq f$ have been recruited to the breeding population, and that once recruited, individuals breed every year. We also assume that both recruitment and breeding occur immediately after survival of an individual from time *i* to i + 1 (such that the recruitment and fecundity parameters are indexed based on the age of the individual at time i + 1). Survival of newborns over their first year S_{α} differs from subsequent 'adult' survival S_a , which is independent of age. Following Frederiksen and Pradel (2001) and Schwarz and Arnason (2001), let π_i (Eq. 3) be the proportion of breeders among all animals aged *i*.

Assuming a post-breeding census, we represent this population model using a life-cycle graph (Fig. 1; Caswell 2001). The projection matrix model corresponding to the life-cycle graph in Fig. 1 is

$$\begin{pmatrix} 0 & S_a \pi_2 B_2 & S_a \pi_3 B_3 & S_a \pi_4 B_4 \\ S_o & 0 & 0 & 0 \\ 0 & S_a & 0 & 0 \\ 0 & 0 & S_a & S_a \end{pmatrix}$$
(4)

Note that the life-cycle graph includes a 'self-loop' on node 4. Such a self-loop generates a 'terminal node' which contains all surviving individuals greater than or equal to

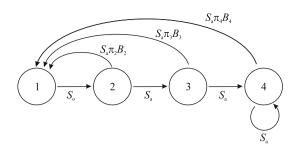


Fig. 1 Life-cycle graph and structure of the hypothetical age-based population model. Transitions between age classes (nodes) are indicated by *arrows*. Transition *labels* indicate the probability of individuals at one stage (*start of arrow*) moving or contributing to the node at the *end of the arrow* over the projection interval. Node 1 refers to offspring, 2 to yearlings, 3 to 2 year individuals, and 4 to individuals \geq 3 years of age. We assume transitions occur over the time scale of 1 year. Parameters *S* and *B* refer to age-specific survival and fertility, respectively. The parameter π_i refers to the proportion of animals of age *i* that are recruits (Eq. 3), which is assumed to represent the proportion of breeding individuals in each age-class. The parameterization and indexing of the fertility arcs reflects the assumption that recruitment and fecundity occur immediately following survival

the minimum age on entry into that node (3 years of age, in this example). Such truncation of the life-cycle graph (by means of a self-loop on a terminal node) is appropriate only if there is no further change with age in any of the parameters contributing to either survival or fertility arcs emanating from the terminal node. When this not the case, truncation results in a model which is only an approximation to the true underlying infinite matrix (see Gosselin and Lebreton 2009); metrics derived from such approximations will be biased (here, we use the term 'bias' to indicate a departure from the metric derived from the 'true' (correct) projection model). In the present example, we have assumed that there was an age-class f at which $a_{i>f} = \pi_{i>f} = 1.0$ (i.e. an age class where all individuals are recruits). In our example, this age class f corresponds to the terminal node in the life-cycle graph. Since all lower-level parameters (S, B, π) for individuals in the terminal node are independent of age, metrics calculated from the projection matrix corresponding to this life-cycle graph (e.g., projected growth rate λ) will be unbiased.

However, suppose that $a_i < 1, \forall i$ (i.e. there is no age at which the probability of recruiting is 1), but where there is an age f beyond which $a_i \ge f$ is constant. Given this age invariance in a_i for $i \ge f$, it might seem reasonable to truncate the life-cycle at age f. However, following Eqs. 2 and 3, age invariance in a is not equivalent to age invariance in π unless a = 1 at some age. Thus, truncating the life-cycle graph at age f would result in a terminal node comprised of exponentially decreasing proportions of newly recruited f year individuals, f + 1 year individuals which may either be (1) new recruits or (2) individuals that

recruited at age f, and so forth. Such heterogeneity would bias projected growth rate λ ; if there is no age class f for which $a_{i>f} = 1.0$, then if the dimension of the age-based projection matrix is less than the age class at which the limit for $\pi_i = 1.0$, the projected growth rate λ will be relatively negatively biased.

There are number of ways to avoid this problem. First, we could reparameterize the current model by recalculating the value of π included in the fertility arc emanating from the terminal node as an average of the π_i values, weighted by the expected frequency of each age class *i* in terminal node. Alternatively, we could increase the number of nodes in the life-cycle graph, such that $\pi \approx 1$ in the terminal node. The number of additional nodes (i.e. age classes) would be determined by the age class (f) at which $\pi_f \approx 1.0$, which can be found either by (1) finding the approximate limit for π_i (Eq. 3), or (2) by using a Markov chain decomposition (discussed below) to determine the expected proportion of breeders in each age class (i.e. π_i), and thus identifying the age class for which this proportion is ≈ 1 . Since the limit of π_i (Eq. 3) is ≈ 1 if $a_i < 1, \forall i$, either approach will only approximate the underlying infinite matrix (Gosselin and Lebreton 2009); the accuracy of the approximation will be a function of the number of age classes included in either (1) or (2).

Consider the following simple numerical example. Let $S_o = 0.5$, $S_a = 0.8$, $B_2 = 0.6$, $B_3 = 0.7$, $B_{4+} = 0.8$. Let $a_2 = 0.35$, $a_3 = 0.7$, and $a_{4+} = 0.9$. From Eq. 2, $\alpha_2 = a_2 = 0.35$, $\alpha_3 = a_3(1 - a_2)(1 - a_1) = 0.455$, and $\alpha_4 = a_4(1 - a_3)(1 - a_2)(1 - a_1) = 0.1755$. From Eq. 3, the proportion of breeders π_i at each age *i* are calculated as: $\pi_2 = 0.35$, $\pi_3 = (0.350 + 0.455) = 0.805$, and $\pi_{4+} = (0.350 + 0.455 + 0.1755) = 0.9805$. If we ignore for the moment that our estimate of $\pi_{4+} < 1$, and use the truncated life-cycle shown in Fig. 1, then from Eq. 4

10	S_a	$\pi_2 B_2$	$S_a \pi_3 B_3$	$S_a \pi_{4+} B_4 $	
S	$S_o = 0$		0	0	
	$0 S_a$		0	0	
10)	0	S_a	S_a)	
	$\begin{pmatrix} 0 \end{pmatrix}$	0.168	0.451	0.628	
_	0.5	0	0	0	
_	0	0.8	0	0	
	0 /	0	0.8	0.8 /	

The projected growth rate from this projection matrix is $\lambda = 1.0369$.

However, since $\pi_{4+} < 1$, we expect that this projected growth rate is negatively biased. Given $a_2 = 0.35$, $a_3 = 0.7$, and $a_{4+} = 0.9$, $\pi_i \approx 1.0$ (to within three significant digits) for i = 6. Extending our age-based model

(Eq. 4) to include at least 6 age classes results in the following projection matrix

(0	0.168 0 0.8 0 0 0 0	0.451	0.628	0.639	0.640	
0.5	0	0	0	0	0	
0	0.8	0	0	0	0	(5)
0	0	0.8	0	0	0	(5)
0	0	0	0.8	0	0	
(0	0	0	0	0.8	0.8 /	

which yields a projected growth rate of $\lambda \approx 1.0389$, which as expected is higher than the projected growth when only 4 age classes were used.

State-based model

While extending the age-based matrix to include more age classes is straightforward, a more direct approach is to use a model based on a state-based projection matrix where age-specific transitions between stages are explicitly included (sensu Lebreton 2005). In such a state-based matrix model, we treat recruits (i.e. first-time breeders) and pre-recruits as discrete 'states', with age-specific transition probabilities between the two states. To do so, we must consider the probability that if alive and in state N (non-breeder) at age *i*, that the individual will be alive and in state B (breeder) at time i + 1. This is simply the product of the probability of surviving from *i* to i + 1, and the probability of recruiting at time *i*, which is a_i (Eq. 1; Pradel and Lebreton 1999).

The life-cycle graph for the state-based equivalent for our original truncated age-based model (Fig. 1) is shown in Fig. 2. The projection matrix corresponding to this statebased life-cycle graph is:

Again, we consider the case where there is no age class f for which $a_{i \ge f} = 1.0$. We showed previously ("Age-based matrix model") that under this assumption, estimates of population growth from an age-based projection matrix will be negatively biased if the dimension of the matrix is less than the age class at which the limit for $\pi_i \approx 1.0$ (i.e. if the model is truncated at an age class where $\pi < 1$). As in our age-based model example, let $a_2 = 0.35$, $a_3 = 0.7$, and $a_{4+} = 0.9$. Using an extended age-based matrix approach with 6 age-classes (Eq. 5), we calculated projected growth rate as $\lambda \approx 1.0388$. Using the state-based matrix (Eq. 6), we obtain an identical value for projected

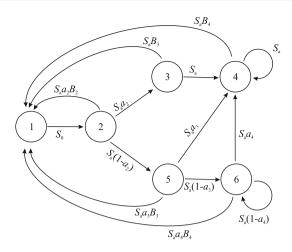


Fig. 2 Life-cycle graph and structure of the hypothetical state-based population model. Transitions between stage classes (nodes) are indicated by *arrows* on the graph. Transition *labels* indicate the probability of individuals at one stage (*start of arrow*) moving or contributing to the node at the *end of the arrow*. Transitions occur over the time scale of 1 year. Node *1* refers to offspring, 2 to yearlings (which do not breed), 3 to breeding 2 year individuals, and 4 to breeding individuals 3+ years of age. Nodes 5 and 6 refer to non-breeder (pre-recruit) individual's age 2 and 3+ years, respectively

growth rate, $\lambda = 1.0388$, without increasing the dimension of the state-based model, even though $a_i < 1.0, \forall i$.

From the state-based model, we can calculate the expected proportions of breeders, π_i , for all age classes, including $i \ge 3$ years. If we decompose the projection matrix A into a fecundity matrix F and a transition matrix T (Cochran and Ellner 1992; Caswell 2001; Tuljapurkar and Horvitz 2006), then the individuals in age class i + 1have a stage distribution proportional to $\lambda^{-i} \mathbf{T}^i \mathbf{F} \mathbf{w}$. Thus, the stage distribution of newborn individuals i = 0 is given as **Fw**, of 1 year-olds is λ^{-1} **TFw**, of 2 year olds is $\lambda^{-2} \mathbf{T}^2 \mathbf{F} \mathbf{w}$, and so on. The proportions of individuals of age *i* in each stage can be determined by normalizing the totals in each age to sum to 1. For the present example, π_i is simply the normalized proportion of breeders, calculated for each age class *i*. If $a_2 = 0.35$, $a_3 = 0.7$, and $a_{4+} = 0.9$, then for age i = 2-6, the equilibrium proportion of individuals in each state (nodes $1 \rightarrow 6$) are

Stage	i = 2	<i>i</i> = 3	i = 4	<i>i</i> = 5	<i>i</i> = 6
1	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000
3	0.350	0.000	0.000	0.000	0.000
4	0.000	0.805	0.980	0.998	1.000
5	0.650	0.000	0.000	0.000	0.000
6	0.000	0.195	0.020	0.002	0.000

For age i = 2,35% of the individuals are expected to be breeders (corresponding to node 3). For age 3, 80.5% are expected to be breeders (node 4). At age 4, 98.05% are expected to be breeders (again, node 4), and so forth. These correspond to the values of $\pi_2 = 0.35$, $\pi_3 = 0.805$, and $\pi_4 = 0.9805$ calculated for the age-based model ("Age-based matrix model"). We note that all surviving 6 year-old (and older) individuals are in stage 4 (recruited breeders; Fig. 2). This is consistent with the limit $\pi_6 \approx 1.0$ which we used to determine the number of age classes needed in the extended age-based projection matrix model to minimize bias in projected growth rate ("Age-based matrix model", Eq. 5).

Recruitment parameterization and perturbation analysis

The preceding suggests that state-based matrix projection models, parameterized directly using estimates of a_i , may have several advantages when compared to more traditional age-based matrix model approaches. In particular, estimates of asymptotic population growth rate are unbiased, even under circumstances where $a_i < 1, \forall i$. Moreover, exact methods for estimation of age-specific parameters from state-based models have been developed (Cochran and Ellner 1992; Caswell 2001; Tuljapurkar and Horvitz 2006).

Two important metrics commonly derived from matrix models are the relative sensitivity and elasticity of the projected growth rate λ to variation in a particular element x_{ij} of the matrix **A** (sensitivity and elasticity analysis are collectively referred to as perturbation analysis, since they assess the projected numerical response of growth rate to perturbation of particular matrix elements; Benton and Grant 1999; de Kroon et al. 2000; Caswell 2001).

Our interest here concerns the role model structure may play in analysis of the influence of variation in recruitment on population growth. We assume recruitment is an agespecific probability, such that not all individuals recruit at the same age. This differs from many analyses of comparative life histories that have historically used integer age of first breeding as the parameter of interest, rather than recruitment, although clearly age of first breeding is directly related to the underlying age-specific recruitment probabilities. Such an approach is perhaps convenient, especially when only partial life-table data are available, such that direct estimates of recruitment are unavailable. Here, we assume that the mean age of first-breeding is noninteger, with non-zero variance, reflecting age-specific differences in recruitment, which we can estimate in some fashion using data from marked individuals. Oli and Zinner (2001) proposed assessing the sensitivity of projected growth λ to age of first breeding using implicit differentiation of the characteristic equation for λ with respect to the mean age of first breeding, estimated using a weighted average based on proportions of individuals recruiting at a given age (which is related, but not equivalent to π ; Eq. 3). Here, however, we focus on recruitment (*a*; Eq. 1) directly, rather than parameter(s) (like mean age of first breeding) which are functions of the underlying parameter *a*.

Perturbation analysis: a numerical example

Consider the following example. We assume a population censused after breeding where individuals can breed for the first time at 1 year (with probability a_1), with full breeding by age 2 years (i.e. $a_i \ge 2$ is a constant, with $a_{i\ge 2} > a_1$). Age-specific fecundity is given as B_i . Survival probabilities are also age-specific (S_o and S_a for offspring and adults, respectively). We assume that once recruited, individuals breed every year. Using a state-based modeling approach ("State-based model"), we represent this population by means of the life-cycle graph shown in Fig. 3.

The projection matrix corresponding to this life-cycle graph is:

$$\begin{pmatrix} S_o a_1 B_1 & S_a B_2 & S_a a_2 B_2 \\ S_o a_1 & S_a & S_a a_2 \\ S_o (1-a_1) & 0 & S_a (1-a_2) \end{pmatrix}$$
(7)

We wish to determine if the structure of the projection model (state- or age-based) affects our analysis of the sensitivity of projected growth rate λ to changes in recruitment. Such a comparison requires construction of an age-based model which is 'equivalent' to the state-based model (Fig. 3). Does an 'equivalent' age-based model exist? As we demonstrate, the answer depends on the assumptions you make, and on what constitutes model equivalency.

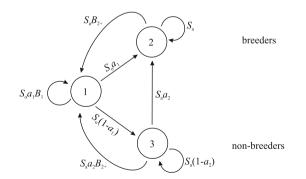


Fig. 3 Life-cycle graph and structure of the hypothetical state-based population model, assuming a post-breeding census. Transitions between stage classes (nodes) are indicated by *arrows* on the graph. Transition *labels* indicate the probability of individuals at one node (*start of arrow*) moving or contributing to the node at the *end of the arrow*. Transitions occur over the time scale of 1 year. Node *1* refers to newborns, 2 breeding individuals, and 3 non-breeding individuals

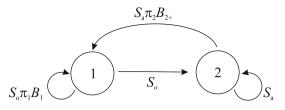


Fig. 4 Life-cycle graph of an age-based population model equivalent to the state-based model shown in Fig. 3, assuming $a_2 = 1$. Node *1* refers to newborns, and 2 to individuals ≥ 1 year of age. Transitions between age classes (nodes) are indicated by *arrows* on the graph. Transition labels indicate the probability of individuals from one node (*start of arrow*) moving or contributing to the node at the *end of the arrow*, assuming a post-breeding census. Transitions occur over the time scale of 1 year

Provided $a_f = 1$ for some age f (such that π_f is guaranteed to equal 1), we have shown ("Age-based matrix model") that we can construct an age-based projection matrix of dimension f nodes (age classes) which is yields the identical dominant eigenvalue as the state-based model. If $a_i < 1, \forall i$, that we can approximate the underlying infinite matrix by increasing the dimension of the age-based model such that $\pi \approx 1$ in the terminal node. To simplify interpretation of any differences between the analysis of state- and age-based matrix models, we begin by assuming $a_2 = 1$. Under this assumption, we construct an age-based life-cycle graph (Fig. 4) which we expect to yield a projected growth rate identical to that calculated from the state-based model (Fig. 3).

The corresponding projection matrix is

$$\begin{pmatrix} S_o \pi_1 B_1 & S_a \pi_2 B_2 \\ S_o & S_a \end{pmatrix}$$
(8)

which when re-parameterized in terms of a_i yields:

$$\begin{pmatrix} S_o a_1 B_1 & S_a [a_1 + a_2(1 - a_1)] B_2 \\ S_o & S_a \end{pmatrix}$$
(9)

It can be shown that the non-zero roots of the characteristic polynomial for both models (Eqs. 7 and 9) are, as expected, identical. (Although the characteristic polynomial for Eq. 7 is order 3, the matrix is singular, such that one of the roots is 0; the remaining roots are square, and identical to the roots for Eq. 9.) This equality occurs if and only if there is an age i = f for which $a_i \ge f = 1$ (in the present example, $a_2 = 1$).

However, does such an equality in non-zero eigenvalues between two models indicate their strict equivalence? For the moment, we proceed with our perturbation analysis of the state- and age-based models under the (naïve) assumption that it *does* imply equivalence. In the process, we demonstrate that, in fact, the assumption of strict equivalence is incorrect, and why (see "Influence of model structure on perturbation analysis – a conundrum?"). We adopt this approach to demonstrate the ease with which a simple but critical error might be made.

To compare sensitivities for the state- and age-based matrix models numerically, we used the following parameter values: $S_o = 0.5$, $S_a = 0.8$, $a_1 = 0.5$, $a_2 = 1.0$, $B_1 = 0.7$, $B_2 = 0.9$. (Since a_1 and a_2 are measured on the same scale, for purposes of convenience we restrict further discussion to comparisons of sensitivities only.)

The state- (Eq. 7) and age-based (Eq. 9) matrices given these parameter values are, respectively:

$$\mathbf{A}_{\text{stage}} = \begin{pmatrix} 0.175 & 0.720 & 0.720 \\ 0.250 & 0.800 & 0.800 \\ 0.250 & 0 & 0 \end{pmatrix}$$
$$\mathbf{A}_{\text{age}} = \begin{pmatrix} 0.175 & 0.720 \\ 0.500 & 0.800 \end{pmatrix}$$

As expected, the projected growth rate is identical for both models ($\lambda = 1.164$).

The sensitivity matrices for the state- and age-based model are, respectively:

$$\mathbf{S}_{\text{stage}} = \begin{pmatrix} 0.269 & 0.312 & 0.058 \\ 0.532 & 0.617 & 0.114 \\ 0.532 & 0.617 & 0.114 \end{pmatrix}$$
$$\mathbf{S}_{\text{age}} = \begin{pmatrix} 0.269 & 0.370 \\ 0.532 & 0.731 \end{pmatrix}$$

For the state-based matrix (Eq. 7), the sensitivity of λ with respect to the a_1 is 0.094, and with respect to a_2 is 0.042. Using the age-based matrix parametrization (Eq. 9), the sensitivity of λ with respect to a_1 is 0.094, which is precisely the same value calculated for the state-based matrix. However, the sensitivity of λ with respect to a_2 calculated from the age-based matrix is 0.133, which is markedly different from the value calculated for a_2 from the state-based matrix (0.042).

Influence of model structure on perturbation analysis – a conundrum?

In the preceding example, the projected growth rate λ from the state-based model was found to be almost twice as sensitive to variation in probability of recruitment at age 1 (0.094) versus age 2 (0.042). However, if we had used an age-based matrix approach, the relative sensitivities were reversed, and of different numerical value for a_2 (0.0942 for a_1 vs. 0.1330 for a_2).

Taken at face value, this would imply that in the statebased matrix, there is a greater influence of a change in the probability of breeding as a 1 year old on population growth, while in the age-based matrix, the probability of breeding as a 2 year old has the greatest absolute influence on population growth. Since sensitivities are strictly analogous to selection differentials (Caswell 2001), and since a_1 and a_2 are measured on the same scale (as [0,1]bounded probabilities), this would imply that in the statebased matrix, there is stronger selection on probability of breeding as a 1 year old, while in the age-based matrix, selection acts most strongly on probability of breeding as a 2 year old.

Clearly, this is problematic since we have assumed that both models are strictly equivalent (based on equality of projected growth rate λ between the two models). As such, we would not anticipate that sensitivity analysis for a particular parameter would depend on the particular model formulation used. What is the explanation for this difference?

There are at least two important issues here which relate what would appear to be a conundrum: the first relates to why the sensitivity of λ to $a_f(a_2$ in the preceding example) differs between state- and age-based models, even when $a_f = 1$. The second concerns the situation when $a_i < 1, \forall i$.

When there is full breeding $(a_f = 1)$

The apparent difference in sensitivity of λ to a_f even when $a_f = 1$ reflects an inconsistency between the aggregation of all recruited individuals in the terminal self-loop and the definition of the parameter a_i . Recall from Eq. (1) that a_i is the probability that a pre-recruit makes a permanent single time-step transition to breeder. Thus, in order to evaluate the sensitivity of projected growth to variation in a_i , the terminal node cannot represent an aggregation of both (1) individuals which make the permanent state-transition from pre-recruit to recruit in a given time step i with (2) individuals which have made the permanent state-transition at a previous time step $\langle i$. Thus, the solution is to modify the structure of the age-based model by creating separate 'classes' (nodes in the life-cycle diagram) for individuals breeding for the first time at the age of full recruitment, followed by a terminal node representing surviving individuals older than the age of full breeding. Specifically, if f is the age at which all individuals are recruits (i.e. $\pi_f \approx 1$, where π_i is the probability that an individual of age *i* is a breeder; Eq. 3), we extend the life-cycle graph to consist of f + 1 nodes, where the new terminal node aggregates surviving individuals which recruited at an earlier age (and this does not contain any new recruits). The sensitivity analysis would be based on the projection model corresponding to this modified life-cycle graph.

We demonstrate this as follows. In our example, $a_1 = 0.5$, and $a_2 = 1.0$. Thus, $\pi_1 = 0.5$, and $\pi_2 = 1.0$ - meaning, 100% of individuals age 2 or older are recruits (i.e. f = 2). We modify the original age-structured lifecycle graph (Fig. 4) by adding a third node (Fig. 5) to temporally separate first-time recruits at age 2 (node 2)

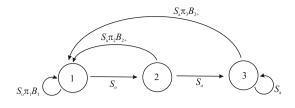


Fig. 5 Modified life-cycle graph and structure of the age-based population model presented in Fig. 4, where node 1 refers to newborns, 2 to individuals 1 year of age, and new node 3 to individuals ≥ 2 years of age

from surviving individuals (age >2 years) that recruited at a younger age. The important distinction between the two life-cycle graphs is that in Fig. 4 the fertility arc for node 2 represents contributions from *both* individuals recruited at age 2 and older, whereas in Fig. 5, the fertility arc for node 2 represents contributions from individuals recruiting at age 2 only, while the fertility arc for node 3 represents contributions from individuals which have previously recruited (since $\pi_2 = 1.0$ for this example; meaning that all individuals in node 2 will recruit over the next time step).

The projection matrix model corresponding to Fig. 5 is (parameterized in terms of π to preserve space) is

$$\begin{pmatrix} S_o \pi_1 B_1 & S_a \pi_2 B_2 & S_a \pi_3 B_2 \\ S_o & 0 & 0 \\ 0 & S_a & S_a \end{pmatrix} = \begin{pmatrix} 0.175 & 0.720 & 0.720 \\ 0.500 & 0 & 0 \\ 0 & 0.800 & 0.800 \end{pmatrix}$$
(10)

The fertility elements for nodes 2 and 3 are identical (0.720) because for this example, $\pi_2 = \pi_3 = 1.0$, and thus the product terms on the fertility arcs for both nodes are the same. We discuss the importance of this later.

The sensitivity matrix corresponding to this modified age-based matrix is

$$\mathbf{S}_{\text{age,modified}} = \begin{pmatrix} 0.269 & 0.116 & 0.254 \\ 0.532 & 0.229 & 0.502 \\ 0.532 & 0.229 & 0.502 \end{pmatrix}$$

Using this modified model, the sensitivity of λ with respect to a_2 is

$$\frac{\partial \lambda}{\partial a_2} = \frac{\partial \lambda}{\partial x_{1,2}} \cdot \frac{\partial x_{1,2}}{\partial a_2} + \frac{\partial \lambda}{\partial x_{1,3}} \cdot \frac{\partial x_{1,3}}{\partial a_2} = 0.042$$

which is *identical* to the value calculated earlier for the state-based matrix.

Why does this modification appear to work? In the original formulation of the model (Fig. 4) the terminal node consisted of all individuals of age $\geq f$. By partitioning this single node into two nodes (Fig. 5), one consisting of new recruits at age f (node 2), and another consisting of

previously recruited individuals age $> f \pmod{3}$, we have effectively partitioned the fertility contributions of the two classes to population growth. Since $\pi_{i>2} = 1$, then the fertility contributions from nodes 2 and 3 are identical (0.720). What differs between them is the relative proportion of individuals in each node at equilibrium. In the original model (Fig. 4), the stable age vector at equilibrium is $\mathbf{w}^T = (0.421, 0.579)$, such that 57.9% of the population at equilibrium would be 2+ years of age. For the modified model (Fig. 5), the stable age vector is $\mathbf{w}^T = (0.421,$ 0.181, 0.398), such that 18.1% of the population at equilibrium would be 2 years of age, and 39.8% would be 3+ years of age. Note that (0.181 + 0.398) = 0.579; in other words, in extending the life-cycle graph by expanding the aggregated terminal node, we are simply partitioning the proportions of individuals aggregated in the terminal node. Since the sensitivity of growth rate λ to variation in a matrix element $x_{i,j}$ is in part a function of the stable age distribution (Caswell 2001), then partitioning the stable age distribution by adding the additional terminal node also partitions the sensitivity contribution between the two nodes. Comparison of the sensitivity matrices for the original and modified age-based projection models make this clear:

$$\begin{split} \mathbf{S}_{age} &= \begin{pmatrix} 0.269 & 0.370 \\ 0.532 & 0.731 \end{pmatrix} \\ \mathbf{S}_{age,modified} &= \begin{pmatrix} 0.269 & 0.116 & 0.254 \\ 0.532 & 0.229 & 0.502 \\ 0.532 & 0.229 & 0.502 \end{pmatrix} \end{split}$$

We see that the sensitivity of population growth to fertility contributions for the terminal node in the original model (0.370) has been partitioned in the modified model (0.116 + 0.254 = 0.370), where the partitions simply reflect the proportions of 2 and 3+ year individuals in the modified model (0.181 and 0.398, respectively) relative to the proportions of 2+ year individuals in the original model (0.579). For example, the proportion of 2 year individuals in the modified model relative to the proportion of 2+ year individuals in the original model (0.181/ 0.579 = 0.313) represents the proportional partition of the sensitivity for 2 year individuals in the modified model (0.313 × 0.370 = 0.116).

Evaluating the sensitivity of population growth λ with respect to some lower-level parameter (like a_i) involves summing the products of (1) the partial derivative of λ with respect to each matrix element x_{ij} that contains the parameter (i.e. the sensitivity of λ with respect to a given matrix element $x_{i,j}$) and (2) the partial derivative of the x_{ij} with respect to the parameter (Caswell 2001). For this example,

$$\frac{\partial \lambda}{\partial a_2} = \frac{\partial \lambda}{\partial x_{1,2}} \cdot \frac{\partial x_{1,2}}{\partial a_2} + \frac{\partial \lambda}{\partial x_{1,3}} \cdot \frac{\partial x_{1,3}}{\partial a_2}$$

We note that the product

$$\frac{\partial \lambda}{\partial x_{1,2}} \cdot \frac{\partial x_{1,2}}{\partial a_2} = 0.116 \cdot \frac{\partial (S_a[a_1 + a_2(1 - a_1)]B_2)}{\partial a_2}$$
$$= 0.116 \cdot (S_a[1 - a_1]B_2)$$
$$= 0.116(0.36)$$
$$= 0.042$$

which is the value reported earlier for the sensitivity of λ to a_2 .

However, since a_2 still occurs as a parameter in the fertility contributions for both node 2 and node 3 in our example, then if $\partial \lambda / \partial a_2 = \partial \lambda / \partial x_{1,2}$, then $\partial \lambda / \partial x_{1,3}$ must equal 0. In fact, this will always be true given that $a_f = a_{f+1} = 1$, because (1) a_{f+1} and it's complement $(1 - a_{f+1})$ occur as a product in $x_{i,f+1}$ (since $a \ge f$ is a constant, then $\alpha_{f+1} = a_{f+1}(1 - a_f)$ is equivalent to $\alpha_f = a_f(1 - a_f)$), and (2) the partial derivative of this function with respect to a_{f+1} evaluated at a_{f+1} will always be 0.

For our example,

$$\begin{aligned} \frac{\partial \lambda}{\partial x_{1,3}} &\cdot \frac{\partial x_{1,3}}{\partial a_2} \\ &= 0.254 \cdot \frac{\partial (S_a[a_1 + a_2(1 - a_1) + a_2(1 - a_2)(1 - a_1)]B_2)}{\partial a_2} \\ &= 0.254 \cdot (S_a[1 - a_1 + (1 - a_2)(1 - a_1) - a_2(1 - a_1)]B_2) \\ &= 0.254(0) \\ &= 0 \end{aligned}$$

When there is less than full breeding $(a_f < 1)$

The approach presented in the preceding is only approximate when $a_i < 1 \ \forall i$. In such cases, the calculated sensitivity will differ from the correct value by an amount reflecting the choice of the number of additional nodes required for $\pi \approx 1$. If we let a_2 in our example equal 0.9, then the sensitivity of population growth λ (where $\lambda = 1.160$, assuming all the other parameter values remain the same) to a_2 calculated using the state-based model (Fig. 3) is 0.0475. Using a Markov chain decomposition of the state-based model, we find that $\pi \approx 1$ (to within 3) decimal places) for 4 age classes (f = 4). If we modify the age-based model to include 5 nodes (f + 1 = 5), the sensitivity of population growth λ to a_2 is 0.0476, which is very close but not identical to the value calculated for the state-based model for this example. The degree of difference gets smaller with each additional node added to the model, but does not reach zero since the limit of π_i (Eq. 3) is ≈ 1 if $a_i < 1, \forall i$ (such that the sum of product terms in the sensitivity is a decreasing infinite series in a_2 . This contrasts with the preceding case where $a_f = 1$, where the series is finite, terminating at node f + 1).

Thus, if an age-based matrix model with a terminal self-loop is used to derive sensitivities of population growth to perturbation of age-specific recruitment probability (a_i) , then the contribution of this node to the estimated sensitivity of λ to a_i will be the sum of the sensitivities for all age classes contained in this node, which will lead to positive bias in the calculated sensitivity overall. This is true even if there exists an age-class *i* for which $a_i = 1$. A straightforward solution to eliminate this bias involves estimating the sensitivity separately for individuals recruiting at the age of full breeding by partitioning the terminal node into new and previously recruited individuals.

Discussion

While the effects of model structure (in particular, the effect of truncating the life-cycle graph) on the calculations of some metrics derived from age-based matrices have been noted elsewhere (McDonald and Caswell 1993; Doak and Morris 1999; Mollet and Cailliet 2002, 2003; Yearsley and Fletcher 2002; Grear and Elderd 2008; Gosselin and Lebreton 2009), the potential difficulties in incorporating recruitment into population projection models presented in this paper have not, to our knowledge, been previously described. We have shown how projected population growth rate, and relative sensitivity of growth to changes in age-specific recruitment, can be significantly affected not only by the choice of how the probability of recruitment is parameterized, but also by the choice of how the population is modeled. In particular, for age-based models where (1) adult age classes are aggregated into a single terminal stage (characterized by the presence of a self-loop on the terminal node of the life-cycle graph), and (2) parameterizing recruitment using the proportion of recruited individuals (breeders) in a given age-class, both the projected population growth rate (sensu Gosselin and Lebreton 2009) and the relative sensitivity of population growth to agespecific changes in recruitment will be biased.

It is worth noting that the inclusion of the terminal-self loop in an age-based matrix does not necessarily bias the sensitivity of population growth to changes in all lowerlevel parameters. In general, the sensitivity to any lowerlevel parameter in the terminal node is the sum of the sensitivities for that parameter for all age classes accumulated in that node. For lower-level parameters that do not change with age beyond the minimal age class represented by the terminal node (i.e. which are constant), the calculated sensitivity is not affected by the number of age classes contained in that node, so long as any other lowerlevel parameters in that node are also constant. What makes sensitivity analysis of a_i different from another parameter (say, S_a), even when it is constant over age (say, for all age classes including and above the age of full breeding)? The difference is in the nature of the parameter being considered—a parameter such as adult survival S_a occurs as a scalar constant, whereas the probability that an individual of a given age is recruited is determined by a limiting series, which unless care is taken, confounds newly and previously recruited individuals in the terminal node. Such a confounding for (say) adult survival S_a is not a problem since individuals which do not survive are not included with surviving individuals in the terminal node.

State-based models: a more robust approach?

While the addition of one extra 'age' (node) (potentially >1 additional node for situations where $a_i < 1, \forall i$) separating newly and previously recruited individuals accounts for the problem, such that the calculated sensitivity of population growth to changes in recruitment is unbiased, this approach is somewhat cumbersome and not particularly intuitive. As such, there appear to be clear advantages to using statebased models to assess the role of recruitment on population dynamics. Multi-state models are very general, and can be structured to account for an arbitrary number of pathways of varying complexity (Caswell 2001; Lebreton 2005; Lebreton et al. 2009). State transitions are more easily and directly parameterized in the state-based approach, and have more direct correspondence with parameters estimated from analysis of data from individuals marked as offspring (Pradel and Lebreton 1999; Lebreton and Pradel 2002). In the context of this study, analysis of multi-state population models appears to be robust to possible violations of assumptions that are often involved in the estimation of recruitment. The most notable example of this involves the common assumption that there exists an age for which $a_i = 1$ (Pradel 1996; Pradel and Lebreton 1999). Such an assumption may not, in fact, be justified by the biology of the species, and relying on this assumption to derive estimates of recruitment may complicate analysis of classical age-based projection models. In fact, even if we had robust estimates of breeding proportions available from a methodology (unspecified) which did not rely on the assumption of full breeding, there may still be difficulties in using these estimates in a classical age-based model.

Model complexity and parameterization: a trade-off

Yearsley and Fletcher (2002) considered the general problem of the trade-off between model complexity and model parameterization. They showed that, under a specific

set of conditions, it was possible to construct a simplified version of a general ('baseline') model which retained several key properties (e.g., projected asymptotic growth rate λ). They noted, however, two important considerations. First, they acknowledged that the motivation for constructing a simplified model is generally the desire to minimize the number of parameters that need to be estimated (and the amount of data needed to estimate those parameters). In some instances, the estimation of some parameters, especially those in aggregated stages, may not be straightforward, or even possible, with current methods.

Second, and perhaps more critically, model construction should be strongly conditioned by the ultimate use of the model. In particular, they noted that not all properties of a complex (general) model can be conserved when creating a simplified model: "the undesirable effects of covariation between a model's structure and quantities of interest have to be weighed against the benefits of model simplification". Doak and Morris (1999), Mollet and Cailliet (2003) and Lebreton (2005) make similar points. Here, we have demonstrated such a situation where the sensitivity of projected growth λ to variation in recruitment can be strongly influenced by model structure, and model parametrization. The common use of a truncated age-based matrix for comparative analysis of life-histories (e.g., Oli and Dobson 2003) or prospective analysis can potentially complicate assessment of the relative importance of key demographic parameters (McDonald and Caswell 1993; Doak and Morris 1999; Mollet and Cailliet 2003; Grear and Elderd 2008; Gosselin and Lebreton 2009). In general, many of these potential difficulties can be mitigated by using a multi-state model approach. Using such a state-structured matrix model, estimates of population growth and the sensitivity of growth to variation in recruitment appear to be unbiased.

Acknowledgments We wish to thank David Thomson and Jean-Dominique Lebreton for inviting this contribution. Res Altwegg, David Fletcher, Jay Rotella, Michael Schaub and Brett Sandercock made a number of helpful comments on a preliminary draft. We also thank Jim Nichols for early discussions concerning recruitment estimation and modeling, and especially Jean-Dominique Lebreton for encouraging us to consider some issues in more depth than we might have otherwise.

References

- Benton TG, Grant A (1999) Elasticity analysis as an important tool in evolutionary and population ecology. Trends Ecol Evol 14:467–471
- Brownie C, Hines JE, Nichols JD, Pollock KH, Hestbeck JB (1993) Capture-recapture studies for multiple strata including non-Markovian transitions. Biometrics 49:1173–1187
- Cam E, Cooch EG, Monnat JY (2005) Earlier recruitment or earlier death? On the assumption of homogenous survival rates in recruitment studies. Ecol Monogr 75:419–434

- Caswell H (2001) Matrix population models—construction, analysis, and interpretation, 2nd edn. Sinauer, USA
- Clobert J, Lebreton JD, Allaine D, Gaillard JM (1994) The estimation of age-specific breeding probabilities from recaptures or resightings in vertebrate populations: II longitudinal models. Biometrics 51:429–438
- Cochran ME, Ellner SP (1992) Simple methods for calculating agebased life history parameters for stage-structured populations. Ecol Monogr 62:345–364
- Cooch EG, Rockwell RF, Brault S (2001) Retrospective analysis of demographic response to environmental change: a lesser snow goose example. Ecol Monogr 71:377–400
- Crespin L, Harris MP, Lebreton JD, Frederiksen M, Wanless S (2006) Recruitment to a seabird population depends on environmental factors and on population size. J Anim Ecol 75:228–238
- Doak DF, Morris WF (1999) Detecting population-level consequences of ongoing environmental change without long-term monitoring. Ecology 80:1537–1551
- Frederiksen M, Pradel R (2001) Comment on Schwarz and Arnason: estimation of age-specific breeding probabilities from capturerecapture data. Biometrics 57:975
- Gosselin F, Lebreton JD (2009) Asymptotic properties of infinite Leslie matrices. J Theor Biol 256:157–163
- Grear J, Elderd B (2008) Population growth rate estimation: sparse data, partial life cycle analysis and Jensen's inequality. Oikos 117:1587–1593
- Jenouvrier S, Tavecchia G, Thibault JC, Choquet R, Bretagnolle V (2008) Recruitment processes in long-lived species with delayed maturity: estimating key demographic parameters. Oikos 117:620–628
- de Kroon H, van Groenendael J, Ehrlen J (2000) Elasticities: a review of methods and model limitations. Ecology 81:607–618
- Lebreton JD (2005) Age, stages, and the role of generation time in matrix models. Ecol Model 188:22–29
- Lebreton JD, Pradel R (2002) Multistate recapture models: modeling incomplete individual histories. J Appl Stat 29:353–369
- Lebreton JD, Nichols JD, Barker RJ, Pradel R, Spendelow JA (2009) Modeling individual animal histories with multistate capturerecapture models. Adv Ecol Res 41:87–173

- McDonald DB, Caswell H (1993) Matrix methods for avian demography. Current ornithology, vol 10. Plenum, New York, pp 139–185
- Mollet HF, Cailliet GM (2002) Comparative population demography of elasmobranchs using life history tables, Leslie matrices and stage-based matrix models. Mar Freshw Res 53:503–516
- Mollet HF, Cailliet GM (2003) Reply to comments by Miller et al. (2003) on Mollet and Cailliet (2002): confronting models with data. Mar Freshw Res 54:739–744
- Nevoux M, Weimerskirch H, Barbraud C (2010) Long- and shortterm influence of environment on recruitment in a species with highly delayed maturity. Oecologia 162:383–392
- Nichols JD, Kendall WL (1995) The use of multistate capturerecapture models to address questions in evolutionary ecology. J Appl Stat 22:835–846
- Oli MK, Dobson FS (2003) The relative importance of life-history variables to population growth rate in mammals: Cole's prediction revisited. Am Nat 161:422–440
- Oli MK, Zinner B (2001) Partial life-cycle analysis: a model for birthpulse populations. Ecology 82:1180–1190
- Pradel R (1996) Utilization of capture-mark-recapture for the study of recruitment and population growth rates. Biometrics 52:703–709
- Pradel R, Lebreton JD (1999) Comparison of different approaches to the study of local recruitment of breeders. Bird Study 46:74–81
- Schwarz CJ, Arnason AN (2000) Estimation of age-specific breeding probabilities from capture-recapture data. Biometrics 56:59–64
- Schwarz CJ, Arnason AN (2001) Comment on Schwarz and Arnason: estimation of age-specific breeding probabilities from capturerecapture data - authors reply. Biometrics 57:976
- Stearns S (1992) The evolution of life histories. Oxford University Press, Oxford
- Tuljapurkar S, Horvitz CC (2006) From stage to age in variable environments: life expectancy and survivorship. Ecology 87:1497–1509
- Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic, London
- Yearsley JM, Fletcher D (2002) Equivalence relationships between stage-structured population models. Math Biosci 179:131–143