Research

Stage, age and individual stochasticity in demography

Hal Caswell

Hal Caswell (hcaswell@whoi.edu), Biology Dept MS-34, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA, and: Max Planck Inst. for Demographic Research, Rostock, Germany.

Demography is the study of the population consequences of the fates of individuals. Individuals are differentiated on the basis of age or, in general, life cycle stages. The movement of an individual through its life cycle is a random process, and although the eventual destination (death) is certain, the pathways taken to that destination are stochastic and will differ even between identical individuals; this is individual stochasticity. A stage-classified demographic model contains implicit age-specific information, which can be analyzed using Markov chain methods. The living stages in the life cycles are transient states in an absorbing Markov chain; death is an absorbing state. This paper presents Markov chain methods for computing the mean and variance of the lifetime number of visits to any transient state, the mean and variance of longevity, the net reproductive rate R_0, and the cohort generation time. It presents the matrix calculus methods needed to calculate the sensitivity and elasticity of all these indices to any life history parameters. These sensitivities have many uses, including calculation of selection gradients. It is shown that the use of R_0 as a measure of fitness or an invasion exponent gives erroneous results except when $R_0 = \lambda = 1$. The Markov chain approach is then generalized to variable environments (deterministic environmental sequences, periodic environments, iid random environments, Markovian environments). Variable environments are analyzed using the vec-permutation method to create a model that classifies individuals jointly by the stage and environmental condition. Throughout, examples are presented using the North Atlantic right whale (*Eubaleana glacialis*) and an endangered prairie plant (*Lomatium bradshawii*) in a stochastic fire environment.

The essence of demography is the connection between the fates of individual organisms and the dynamics of populations. There exist diverse mathematical frameworks in which this connection can be studied (e.g. Keyfitz 1967, Nisbet and Gurney 1982, Metz and Diekmann 1986, Caswell 1989, DeAngelis and Gross 1992, Caswell et al. 1997, Tuljapurkar and Caswell 1997). Regardless of the type of equations used, demographic analysis must account for differences among individuals, and the ways in which those differences affect the vital rates. Among the many ways that individuals may differ, age has long had a kind of conceptual priority. Age is universal in the sense that every organism becomes one minute older with the passage of one minute of time. Age is also often associated with predictable changes in the vital rates. However, in some organisms characteristics other than age provide more and better information about an individual. Ecologists recognized this long ago, and have developed demographic theory based on size, maturity, physiological condition, instar, spatial location, etc. – what we refer to in general as 'stage-classified' demography. Human demographers, who were responsible for the classical age-classified theory, by no means deny the importance of other properties (employment, parity, health, etc.); see Land and Rogers (1982), Goldman (1994), Robine et al. (2003), or Jasney et al. (2006) for just a sample of the kinds of issues that arise.

There now exists a comprehensive stage-classified demographic theory, applicable regardless of taxon or habitat (Caswell 2001). It includes linear and nonlinear, time-invariant and time-varying, deterministic and stochastic models. It describes growth rate, population structure, stability and bifurcations, transient dynamics and oscillations, extinctions and quasi-extinctions. It includes sensitivity analyses that make it possible to explore how these properties respond to changes in the parameters of the population and its environment.

Even when the demographic model is entirely stage-classified, however, age is still implicitly present. Individuals in a given stage may differ in age, and individuals of a given age may be found in many different stages, but each individual still becomes one unit of age older with the passage of each unit of time. Extracting this implicit age-dependent information makes it possible to calculate interesting age-specific properties, such as survivorship, longevity, life expectancy, generation time, and net reproductive rate (Cochran and Ellner 1992, Caswell 2001, 2006, Tuljapurkar and Horvitz 2006, Horvitz and Tuljapurkar 2008).

Consider a newborn individual. As it develops through the stages of its life cycle, it may grow, shrink, mature, move, reproduce, and allocate resources among its biological processes. At each moment, it is exposed to various mortality risks. Because these processes are stochastic, the lives of any two individuals may differ. These random

differences (which I propose to call individual stochasticity) imply that the age-specific properties of an individual (say, longevity) are random variables – there is a distribution among individuals that should be characterized by its mean, moments, etc.

In this paper, I show how to calculate some of these implicit age-specific properties from any stage-classified model. The trick is to formulate the life cycle as a Markov chain, and to generalize the 'life' cycle to include death as a stage. Because death is permanent, it is called an absorbing state, and the theory of absorbing Markov chains provides the starting point for our analysis (Feichtinger 1971, Caswell 2001).

A Markov chain is a stochastic model for the movement of a particle among a set of states (e.g. Kemeny and Snell 1976, Iosifescu 1980). The probability distribution of the next state of the particle may depend on the current state, but not on earlier states. In our context, a 'particle' is an individual organism. The states correspond to the stages of the life cycle, plus death (or perhaps multiple types of death, for example deaths due to different causes). This structure is ideally suited to asking questions about individual stochasticity, because it accounts for all the possible pathways, and their probabilities, that an individual can follow through its life. I will focus on discrete-time models, but much of the theory can no doubt be generalized to continuous-time models.

Perturbation analysis (sensitivity and elasticity analysis) plays a particularly important role in demography. The sensitivity of population properties (e.g. λ) to changes in the underlying parameters is useful in conservation, management, pest control, ecotoxicology, epidemiology and evolution. In this paper, I will present perturbation analyses for almost all of the properties I consider. These rely on new mathematical results (Caswell 2006) obtained using matrix calculus. Although the approach produces amusingly complicated-looking formulas, with a little practice it is surprisingly easy to implement (Caswell 2006, 2007, 2008, 2009, Verdy and Caswell 2008). The best way to get that practice is, naturally, to practice. So, in this paper, I will pose some of the derivations as challenges. The answers are given, in detail, in Supplementary material Appendix A, but I urge you to try some of them before turning to the solutions. Equations in the Supplementary Appendices are numbered (A—1), (B—1), etc.

I will present examples of the calculations, but cannot say whether the results are specific to these studies, or are general. If they are general, I cannot say what aspects of the life history influence them. Those questions can be answered only by a collection of comparative studies; one goal of this paper is to present the techniques necessary to begin such a collection. These Markov chain models are intimately connected with empirical data on identified individuals followed over time (by marking, banding, or photographic identification; see Caswell and Fujiwara 2004, Lebreton et al. 2009). As a result, powerful statistical methods (e.g. parameter estimation, model selection, confidence intervals, statistical significance tests) are available for the results I will present here.

The use of Markov chains in demographic analysis is not new. As far as I know, Feichtinger (1971, 1973) was the first to use discrete-time absorbing Markov chains in demography, paying particular attention to competing risks and multiple causes of death. At around the same time, Hoem (1969) applied continuous-time Markov chains in the analysis of insurance systems (with states such as 'active,' 'disabled,' and 'dead'). Later, Cochran and Ellner (1992) independently proposed the use of Markov chains to generate age-classified statistics from stage-classified models, but minimized the use of matrix notation in their presentation. Influenced by Feichtinger's work, and relying heavily on Iosifescu's (1980) treatment of absorbing Markov chains, I extended the calculations using matrix notation (Caswell 2001, Keyfitz and Caswell 2005). Later (Caswell 2006) I introduced sensitivity analysis, and presented results for both time-invariant and time-varying models. At the same time, Tuljapurkar and Horvitz (2006) presented a more extensive investigation of time-variation.

Examples

The calculations will be demonstrated by means of two case studies. The first is a stage-classified model for the North Atlantic right whale *Eubalaena glacialis*. Later, a stochastic matrix model for the threatened prairie plant *Lomatium bradshawii* will appear as part of a study of variable environments.

The North Atlantic right whale is a large, highly endangered baleen whale (Kraus and Rolland 2007). Once abundant in the north Atlantic, it was decimated by whaling, beginning as much as a thousand years ago (Reeves et al. 2007). By 1900 the western North Atlantic stock had been effectively eliminated, and the eastern North Atlantic stock hunted to near extinction. The population has recovered only slowly since receiving at least nominal protection in 1935, and now numbers only about 300 individuals. Right whales migrate along the Atlantic coast of North America, from summer feeding grounds in the Gulf of Maine and Bay of Fundy to winter calving grounds off the southeastern US. They are killed by ship collisions and entanglement in fishing gear (Moore et al. 2005), and may also be affected by pollution of coastal waters (Rolland et al. 2007).

Individual right whales are photographically identifiable by scars and callosity patterns. Since 1980, the New England Aquarium has surveyed the population, accumulating a database of over 10 000 sightings (Crone and Kraus 1990). Treating the first year of identification of an individual as marking, and each year of resighting as a recapture, permits the use of mark–recapture statistics to estimate demographic parameters of this endangered population (Caswell et al. 1999, Fujiwara and Caswell 2001, 2002, Fujiwara 2002, Caswell and Fujiwara 2004).

Figure 1 shows a life cycle graph used by Caswell and Fujiwara (2004) as the basis of a stage-structured matrix population model for the right whale. The stages are calves, immature females, mature but non-reproductive females, mothers, and 'resting' mothers (because of the long period of parental care and gestation, right whales do not reproduce in the year after giving birth). This life cycle is typical of large, long-lived monovular mammals and birds.

The model is parameterized in terms of survival probabilities $\sigma_1, \ldots, \sigma_5$, the probability of maturation γ_2, and the birth probability γ_3. The projection matrix is
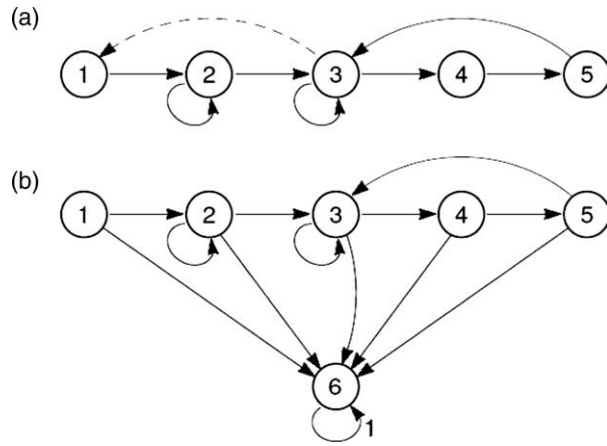
Figure 1. Life cycle graphs for females of the North Atlantic right whale (*Eubalaena glacialis*). Projection interval is one year. Stages: 1 = calf, 2 = immature, 3 = mature, 4 = mother, 5 = post-breeding female. (a) the graph corresponding to the population projection matrix $\mathbf{A}$. Solid arcs indicate transitions, dashed arcs indicate reproduction. Each time a mature female enters stage 4, she produces a calf. See Caswell and Fujiwara (2004) for explanation and parameter estimates. (b) the graph corresponding to the absorbing Markov chain, showing transitions of extant individuals only. Stage 6 = death is an absorbing state.

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & F & 0 & 0 \\ \sigma_1 & \sigma_2(1-\gamma_2) & 0 & 0 & 0 \\ 0 & \sigma_2\gamma_2 & \sigma_3(1-\gamma_3) & 0 & \sigma_5 \\ 0 & 0 & \sigma_3\gamma_3 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4 & 0 \end{pmatrix} \quad (1)$$

The fertility term in the (1, 3) position is $F = 0.5\sigma_3\gamma_3\sqrt{\sigma_4}$, accounting for the sex ratio, the survival of mature females, their probability of giving birth if they survive, and the effect of survival of the mother on survival of the calf. For reasons related to parameter estimation, σ_5 is constrained to equal σ_3.

A note on notation

In this paper, matrices are denoted by upper case bold symbols (e.g. $\mathbf{A}$), vectors by lower case bold symbols ($\mathbf{n}$); a_{ij} is the (i, j) entry of the matrix $\mathbf{A}$, n_i is the ith entry of the vector $\mathbf{n}$. Vectors are columns, and $\mathbf{n}^T$ is the row vector obtained as the transpose of $\mathbf{n}$. Logarithms are natural. In addition to the ordinary matrix product, the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ and the Hadamard product $\mathbf{A} \circ \mathbf{B}$ will appear (Supplementary material Appendix A). The symbol diag ($\mathbf{x}$) denotes the square matrix with $\mathbf{x}$ on the diagonal and zeros elsewhere. The symbol $\mathbf{e}$ denotes a vector of 1s; the vector $\mathbf{e}_i$ is a vector with 1 in the ith entry and zeros elsewhere. The identity matrix is $\mathbf{I}$. Sometimes dimensions are indicated by subscripts, as in $\mathbf{I}_s$ for the $s \times s$ identity matrix, or $\mathbf{e}_{sp \times 1}$ for a vector of ones of length sp.

Markov chains

The familiar life cycle graph (Fig. 1a) corresponds to a projection matrix $\mathbf{A}$, in which a_{ij} gives the per-capita production of stage i individuals at $t+1$ by a stage j individual at t. This production may occur by the transition of an individual from stage j to stage i, or by the production of one or more new individuals (by reproduction, fragmentation, etc.). So, we partition $\mathbf{A}$ into a matrix $\mathbf{U}$ describing transition probabilities of extant individuals and a matrix $\mathbf{F}$ describing the production of new individuals

$$\mathbf{A} = \mathbf{U} + \mathbf{F} \quad (2)$$

The column sums of $\mathbf{U}$ are all less than or equal to 1. Because individuals eventually die and pass out of the stages contained in $\mathbf{U}$, those stages are called transient states.

An absorbing Markov chain

If we include death explicitly (Fig. 1b) and remove the arcs representing reproduction, we obtain the graph corresponding to the transition matrix for an absorbing Markov chain

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & 0 \\ \hline \mathbf{m} & 1 \end{array} \right) \quad (3)$$

The element m_j of the vector $\mathbf{m}$ is the probability of mortality of an individual in stage j. Death is an absorbing state. I will assume that at least one absorbing state is accessible from any transient state in $\mathbf{U}$, and that the spectral radius of $\mathbf{U}$ is strictly less than 1. This guarantees that, with probability 1, every individual ends up in the absorbing state.

The right whale

Fujiwara (2002, Caswell and Fujiwara 2004) estimated $\mathbf{U}$ by applying multi-stage mark-recapture methods to the photographic identification catalog. Although the best model, out of a large number evaluated, included significant time variation in survival and birth rates, here I will analyze a single matrix obtained from a time-invariant model. The complete transient matrix $\mathbf{U}$ and the fertility matrix $\mathbf{F}$ are

$$\mathbf{U} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.90 & 0.85 & 0 & 0 & 0 \\ 0 & 0.12 & 0.71 & 0 & 1.00 \\ 0 & 0 & 0.29 & 0 & 0 \\ 0 & 0 & 0 & 0.85 & 0 \end{pmatrix} \quad (4)$$

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0.13 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5)$$

Expected visits to transient states: the fundamental matrix

As the syllogism asserts, all men are mortal; absorbtion is certain. Our question is, how long does absorbtion take and what happens en route? From a demographic perspective, this is asking about the lifespan of an individual and the events that happen during that lifetime. The key to these questions is the fundamental matrix of the absorbing Markov chain. Consider an individual that is presently in

transient state j. As time passes, it will visit other transient states, repeating some, skipping others, until it eventually dies. Let v_{ij} denote the number of visits to transient state i that our individual, starting in transient state j, makes before being absorbed. The v_{ij} are random variables, reflecting individual stochasticity.

The entries of the matrix $\mathbf{U}$ give the probabilities of visiting each of the transient states after one time step. The entries of $\mathbf{U}^2$ give the probabilities of visiting each of the transient states after two time steps. Adding the powers of $\mathbf{U}$ gives the expected number of visits to each transient state, over a lifetime, in a matrix $\mathbf{N}$; i.e.

$$
\begin{aligned}
\mathbf{N} &= \left(E(v_{ij})\right) \\
&= \sum_{t=0}^{\infty} \mathbf{U}^t \\
&= (\mathbf{I} - \mathbf{U})^{-1}
\end{aligned}
\tag{6}
$$

The right whale

The fundamental matrix for the right whale is calculated from (6) to be

$$
\mathbf{N} = \begin{pmatrix}
1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\
5.88 & 6.52 & 0.00 & 0.00 & 0.00 \\
16.35 & 18.11 & 22.94 & 19.49 & 22.94 \\
\mathbf{4.74} & 5.25 & 6.65 & 6.65 & 6.65 \\
4.02 & 4.46 & 5.65 & 5.65 & 6.65
\end{pmatrix}
\tag{7}
$$

The first column corresponds to calves. On average, a calf will spend 1 year as a calf, 5.9 years as a juvenile, 16.3 years as a mature but non-breeding female, etc. Row 4 of $\mathbf{N}$ is of particular interest. Stage 4 represents mothers, so n_{4j} is the expected number of reproductive events that a female in stage j will experience during her remaining lifetime. Based on this model, a newborn calf could expect to give birth $n_{41} = 4.74$ times. A mature female could expect to give birth $n_{43} = 6.65$ times; the difference reflects the likelihood of mortality between birth and maturity. (Note that $n_{43} = n_{44} = n_{45} = n_{55}$ and $n_{53} = n_{54}$. This seems to be due to the fact, specific to these data, that the survival probability of stages 3 and 5 is indistinguishable from 1.0, and influences the results below.)

We would like to know how the entries of $\mathbf{N}$ vary in response to changes in the vital rates. To accomplish this, we need matrix calculus, which is the topic of the next section.

Sensitivity via matrix calculus

The output of a demographic calculation may be a scalar (e.g. λ), a vector (e.g. the stable stage distribution), or a matrix (e.g. $\mathbf{N}$). The parameters of interest may also be scalars, vectors (e.g. a vector of survival probabilities), or matrices (e.g. the entries of $\mathbf{U}$). Thus, sensitivity analysis requires a way to differentiate scalar-, vector- and matrix-valued functions of scalar, vector or matrix arguments. This is provided by matrix calculus. The approach I describe here is due to Magnus and Neudecker (1985, 1988); it is called the vector-rearrangement method in the review paper of

Nel (1980). Because matrix calculus has only recently been applied in ecology (Caswell 2006, 2007, 2008, 2009, Verdy and Caswell 2008), I summarize some of the basic methods. For more details see Caswell (2007), the mathematical introduction in Abadir and Magnus (2005), or the complete treatment in Magnus and Neudecker (1988).

Differentiation

If x and y are scalars, differentiating y with respect to x gives the familiar derivative dy/dx. If $\mathbf{y}$ is a $n \times 1$ vector and x is a scalar, differentiating $\mathbf{y}$ with respect to x gives the $n \times 1$ vector

$$
\frac{d\mathbf{y}}{dx} = \begin{pmatrix}
\dfrac{dy_1}{dx} \\
\vdots \\
\dfrac{dy_n}{dx}
\end{pmatrix}
\tag{8}
$$

If y is a scalar and x is a $m \times 1$ vector, differentiating y with respect to $\mathbf{x}$ gives the $1 \times m$ vector (called the gradient vector)

$$
\frac{dy}{d\mathbf{x}^T} = \left(\frac{\partial y}{\partial x_1} \cdots \frac{\partial y}{\partial x_m} \right)
\tag{9}
$$

Note the orientation of $d\mathbf{y}/dx$ as a column vector and $dy/d\mathbf{x}^T$ as a row vector.

If $\mathbf{y}$ is a $n \times 1$ vector and $\mathbf{x}$ a $m \times 1$ vector, differentiating $\mathbf{y}$ with respect to $\mathbf{x}$ gives the $n \times m$ matrix (called the Jacobian matrix)

$$
\frac{d\mathbf{y}}{d\mathbf{x}^T} = \left(\frac{dy_i}{dx_j} \right)
\tag{10}
$$

Derivatives involving matrices are written by transforming the matrices into vectors using the vec operator (which stacks the columns of the matrix into a column vector), and then applying the rules for vector differentiation. That is, the derivative of the $m \times n$ matrix $\mathbf{Y}$ with respect to the $p \times q$ matrix $\mathbf{X}$ is the $mn \times pq$ matrix

$$
\frac{d\text{vec } \mathbf{Y}}{d(\text{vec } \mathbf{X})^T}
\tag{11}
$$

For notational convenience, I will write $\text{vec}^T\mathbf{X}$ for $(\text{vec }\mathbf{X})^T$.

These definitions (unlike some alternatives; see Magnus and Neudecker 1985) lead to the familiar chain rule of calculus. If $\mathbf{Y}$ is a function of $\mathbf{X}$, and $\mathbf{X}$ is a function of $\mathbf{Z}$, then

$$
\frac{d\text{vec } \mathbf{Y}}{d\text{vec}^T \mathbf{Z}} = \frac{d\text{vec } \mathbf{Y}}{d\text{vec}^T \mathbf{X}} \frac{d\text{vec } \mathbf{X}}{d\text{vec}^T \mathbf{Z}}
\tag{12}
$$

In practice, matrix derivatives are constructed by forming differentials, where the differential of a matrix (or vector) is the matrix (or vector) of differentials of the elements; i.e.

$$
d\mathbf{X} = (dx_{ij})
\tag{13}
$$

The key result was proven by Magnus and Neudecker (1985, 1988) who showed that if, for some matrix $\mathbf{Q}$, it can be shown that

$$d\mathbf{y} = \mathbf{Q}\,d\mathbf{x} \qquad (14)$$

then

$$\frac{d\mathbf{y}}{d\mathbf{x}^{\mathsf{T}}} = \mathbf{Q} \qquad (15)$$

This is the "first identification theorem" of Magnus and Neudecker (1985).

The combination of the chain rule and the identification theorem permits more complicated expressions involving differentials to be turned into derivatives with respect to an arbitrary vector, say $\mathbf{u}$. If it can be shown that

$$d\mathbf{y} = \mathbf{Q}\,d\mathbf{x} + \mathbf{R}\,d\mathbf{z} \qquad (16)$$

then

$$\frac{d\mathbf{y}}{d\mathbf{u}^{\mathsf{T}}} = \mathbf{Q}\frac{d\mathbf{x}}{d\mathbf{u}^{\mathsf{T}}} + \mathbf{R}\frac{d\mathbf{z}}{d\mathbf{u}^{\mathsf{T}}} \qquad (17)$$

for any $\mathbf{u}$.

To arrive at expressions where the identification theorem can be applied, we will make extensive use the Kronecker product, defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \qquad (18)$$

The vec operator and the Kronecker product are related by a theorem due to Roth (1934); if

$$\mathbf{Y} = \mathbf{ABC} \qquad (19)$$

then

$$\operatorname{vec}\mathbf{Y} = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A})\operatorname{vec}\mathbf{B} \qquad (20)$$

Finally, we will want to explore elasticities, or proportional sensitivities. To convert a matrix of sensitivities of $\mathbf{y}$ to $\mathbf{x}$ to elasticities requires multiplying the jth column by x_j and dividing the ith row by y_i. This can be accomplished by matrix multiplication as

$$\operatorname{diag}(\mathbf{x})^{-1}\left(\frac{d\mathbf{y}}{d\mathbf{x}^{\mathsf{T}}}\right)\operatorname{diag}(\mathbf{y}) \qquad (21)$$

Sensitivity of the fundamental matrix

Let us apply matrix calculus to find the sensitivity of the fundamental matrix $\mathbf{N}$ (Caswell 2006). This result will appear in the sensitivity analysis of most other demographic quantities. Let $\boldsymbol{\theta}$ be a vector of parameters (of dimension $p \times 1$) on which the entries of the transition matrix $\mathbf{U}$ depend. The fundamental matrix satisfies

$$\mathbf{I} = \mathbf{N}\mathbf{N}^{-1} \qquad (22)$$

Differentiating both sides gives

$$\mathbf{0} = (d\mathbf{N})\mathbf{N}^{-1} + \mathbf{N}(d\mathbf{N}^{-1}) \qquad (23)$$

Applying the vec operator and Roth's theorem to both sides gives

$$\operatorname{vec}\mathbf{0} = [(\mathbf{N}^{-1})^{\mathsf{T}} \otimes \mathbf{I}_s]\,\operatorname{dvec}\mathbf{N} + (\mathbf{I}_s \otimes \mathbf{N})\,\operatorname{dvec}\mathbf{N}^{-1} \qquad (24)$$

Solving for $\operatorname{dvec}\mathbf{N}$ gives

$$\operatorname{dvec}\mathbf{N} = [(\mathbf{N}^{-1})^{\mathsf{T}} \otimes \mathbf{I}_s]^{-1}(\mathbf{I}_s \otimes \mathbf{N})\,\operatorname{dvec}\mathbf{U} \qquad (25)$$

To simplify this, it helps to know two facts about the Kronecker product:

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \qquad (26)$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD}) \qquad (27)$$

provided that the sizes of the matrices permit the indicated operations. Thus $\operatorname{dvec}\mathbf{N}$ in (25) simplifies to

$$\operatorname{dvec}\mathbf{N} = (\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N})\operatorname{dvec}\mathbf{U} \qquad (28)$$

The identification theorem (15) implies

$$\frac{\operatorname{dvec}\mathbf{N}}{\operatorname{dvec}^{\mathsf{T}}\mathbf{U}} = \mathbf{N}^{\mathsf{T}} \otimes \mathbf{N} \qquad (29)$$

and the chain rule permits us to write

$$\frac{\operatorname{dvec}\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} = (\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N})\frac{\operatorname{dvec}\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \qquad (30)$$

The left-hand side of (30) is a matrix, of dimension $s^2 \times p$, containing the sensitivity of every entry of $\mathbf{N}$ to every parameter in $\boldsymbol{\theta}$. The matrix $\operatorname{dvec}\mathbf{U}/d\boldsymbol{\theta}^{\mathsf{T}}$ is an $s^2 \times p$ matrix containing the sensitivities of all the elements of $\mathbf{U}$ to all the elements of $\boldsymbol{\theta}$. From (21), the elasticity of the fundamental matrix is given by

$$\operatorname{diag}(\operatorname{vec}\mathbf{N})^{-1}\frac{\operatorname{dvec}\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}}\operatorname{diag}(\boldsymbol{\theta}) \qquad (31)$$

The right whale

As an example, we use (30) and (31) to calculate the elasticity of the expected lifetime number of reproductive events, $E(\nu_{41}) = n_{41}$, with respect to the survival probabilities $\sigma_1, \ldots, \sigma_4$, the maturation probability γ_2, and the breeding probability γ_3. Figure 2 shows that the number of breeding events is most elastic to mature female survival (σ_3), and less so to the survival of immature females or mothers (σ_2 and σ_4). Changes in the probability of giving birth, γ_3, have, remarkably enough, no impact on the expected number of reproductive events.

The elasticity of n_{41} to σ_3 (survival of mature females) is approximately 30. This implies that a 1% increase in σ_3 would produce about a 30% increase in the expected number of reproductive events.

From stage to age

The fundamental matrix summarizes the age-specific information implicit in the transient matrix $\mathbf{U}$, even if the model is stage-classified and age does not appear explicitly. We now extend this, to explore a series of age-specific demographic indices and their sensitivity analyses. Some are well known (R_0, generation time), others little explored (variance in longevity, for example). They can, however, all be easily calculated from any stage-classified model.

Variance in visits to transient states

The number of visits to any transient state is a random variable; the fundamental matrix $\mathbf{N}$ gives its mean. Some individuals will visit that state more often, some less often,

(a)

Lifetime reproductive events
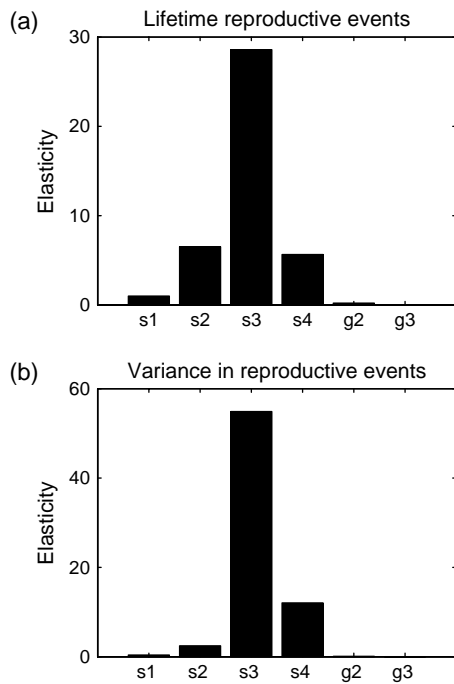

(b)

Variance in reproductive events

Figure 2. (a) the elasticity, to each of the vital rates, of the expected lifetime number of reproductive events ($E(v_{41})$) for the right whale. (b) the elasticity of the variance in the lifetime number of reproductive events, $V(v_{41})$ for the right whale. Vital rates: s_1–s_4 are survival probabilities ($s_5 = s_3$ by assumption in this model); g_2 is the probability of maturation, and g_3 is the probability of reproduction.

Net reproductive rate

Figure 4. The elasticity, to each of the vital rates, of the net reproductive rate (R_0) for the right whale. Parameters as in Fig. 2.

some not at all. This basic property of individual stochasticity can be described by the variance of v_{ij}. Iosifescu (1980, theorem 3.1) gives a formula for all the moments of the v_{ij}; from this we can calculate the matrix of variances

$$\mathbf{V} = (V(v_{ij})) = (2\mathbf{N}_{\text{diag}} - \mathbf{I})\mathbf{N} - \mathbf{N} \circ \mathbf{N} \qquad (32)$$

(Caswell 2006) where $\circ$ denotes the Hadamard, or element-by-element, product and $\mathbf{N}_{\text{diag}}$ is a matrix with the diagonal elements of $\mathbf{N}$ on its diagonal and zeros elsewhere. The standard deviations of the number of visits to transient states are the square roots of the elements of $\mathbf{V}$.

(a)

Life expectancy
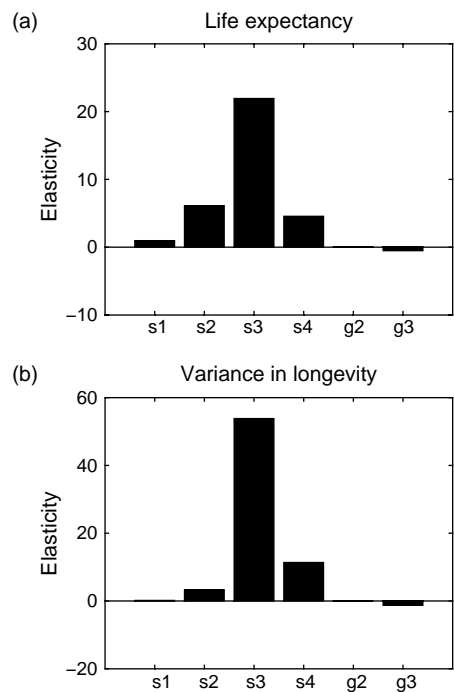

(b)

Variance in longevity

Figure 3. (a) the elasticity, to each of the vital rates, of life expectancy for a female right whale calf. (b) elasticity of the variance in longevity for a female right whale calf. Parameters as in Fig. 2.
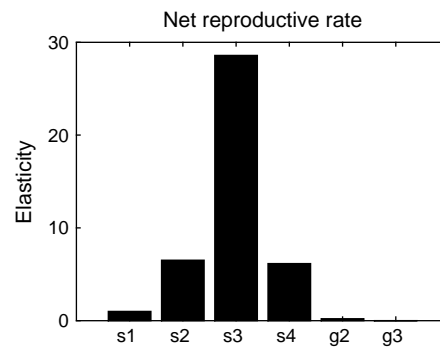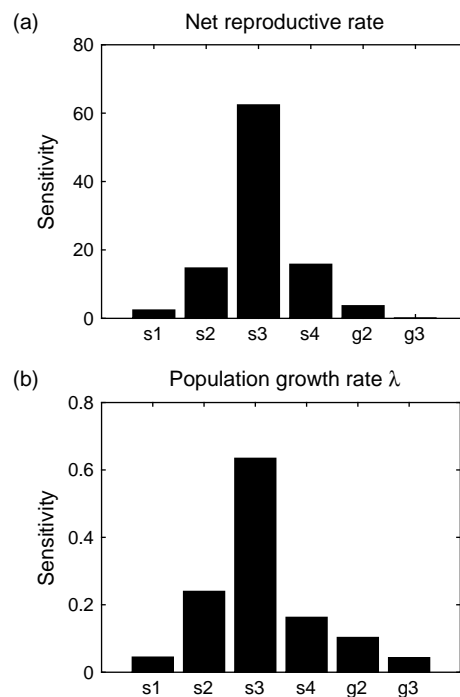

(a)

Net reproductive rate

(b)

Population growth rate λ

Figure 5. (a) the sensitivity, to each of the vital rates, of the net reproductive rate R_0 for the right whale. (b) the sensitivity of population growth rate λ. The derivative of λ is the selection gradient; use of the derivative of R_0 leads to erroneous predictions unless the population is at equilibrium. Parameters as in Fig. 2.
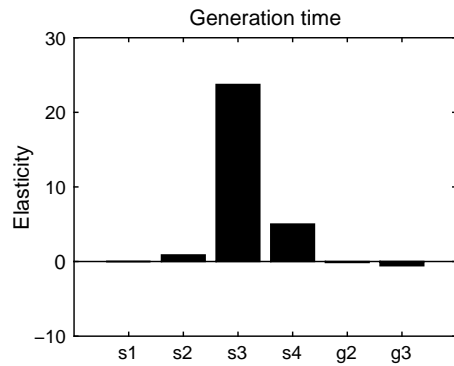
Figure 6. The elasticity, to each of the vital rates, of the cohort generation time for a newborn calf right whale. Parameters as in Fig. 2.

The right whale

For the right whale, the matrix of variances calculated from (32) is

$$
\mathbf{V} = \begin{pmatrix}
0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\
36.18 & 35.95 & 0.00 & 0.00 & 0.00 \\
466.44 & 484.80 & 503.32 & 494.86 & 503.32 \\
35.80 & 36.98 & 37.54 & 37.54 & 37.54 \\
33.28 & 34.94 & 37.54 & 37.54 & 37.54
\end{pmatrix}
$$

(33)

and the corresponding standard deviations are

$$
\left(\mathrm{SD}(v_{ij})\right) = \begin{pmatrix}
0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\
6.02 & 6.00 & 0.00 & 0.00 & 0.00 \\
21.60 & 22.02 & 22.43 & 22.25 & 22.43 \\
5.98 & 6.08 & 6.13 & 6.13 & 6.13 \\
5.77 & 5.91 & 6.13 & 6.13 & 6.13
\end{pmatrix}
$$

(34)

The variance in the v_{ij} is the result of luck, not heterogeneity. That is, it is the variance among a group of individuals all experiencing exactly the same stage-specific
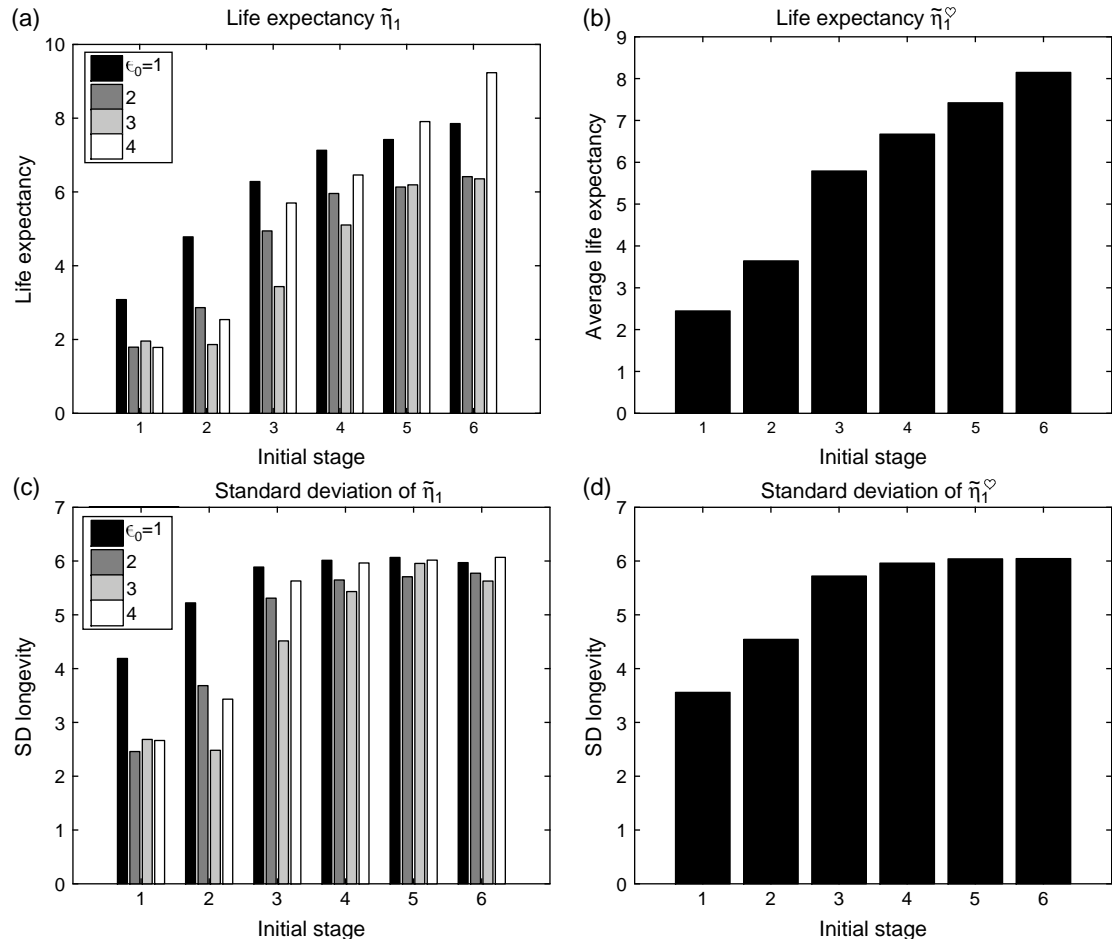

Figure 7. The expectation and standard deviation of longevity for *Lomatium bradshawii* in a stochastic fire environment. (a) the expected longevity conditional on the initial environment (ϵ_0). Environments: 1 = year of a fire, 2 = 1 year post-fire, 3 = 2 years post-fire, 4 = three or more years post-fire. (b) the expected longevity averaged over the stationary distribution of initial environments. (c) the standard deviation of longevity conditional on the initial environment. (d) the standard deviation of longevity over the stationary distribution of initial environments. The frequency of fire is 0.5 and the temporal autocorrelation $\rho = 0.7$. Based on matrices from Caswell and Kaye (2001).
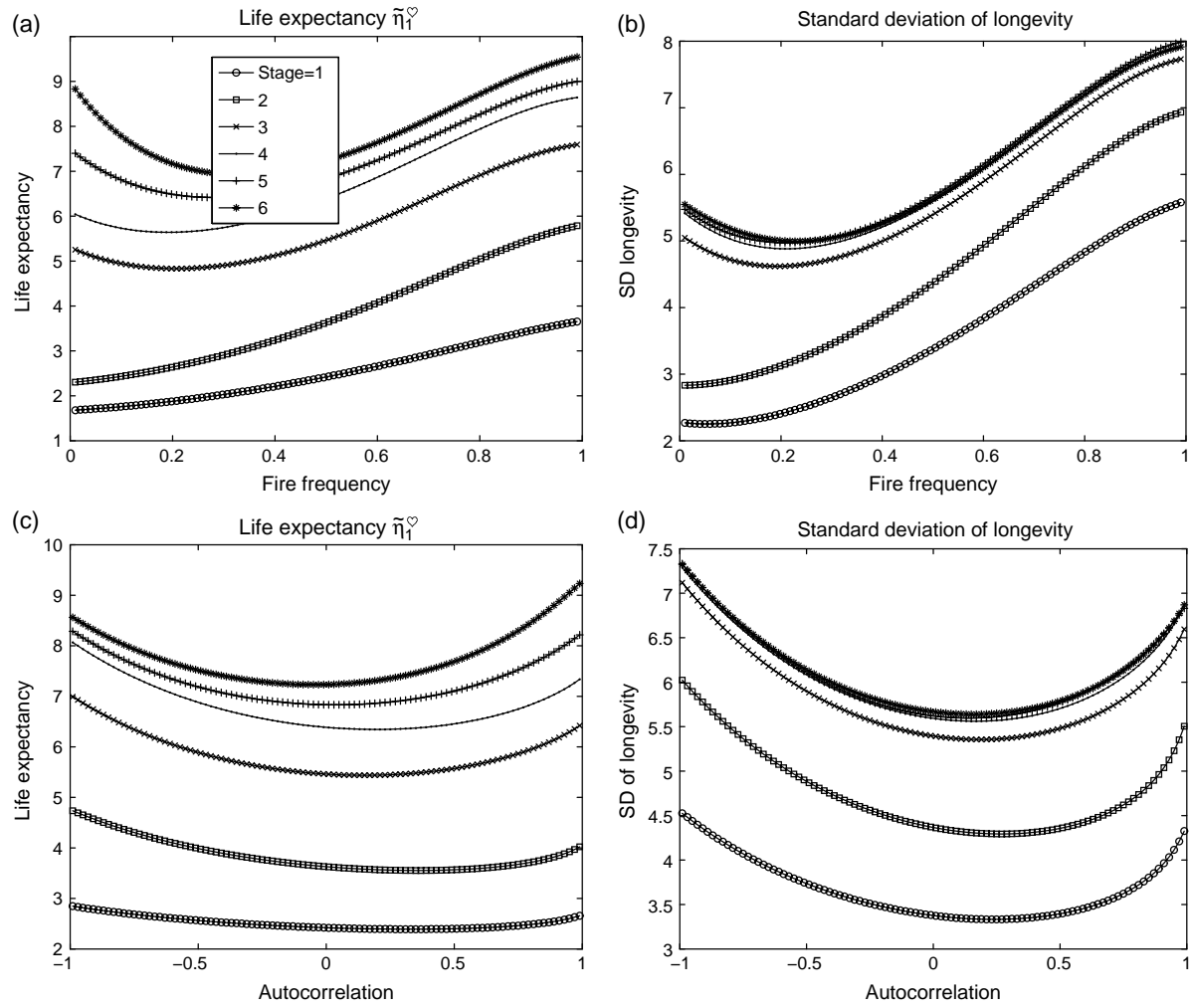
1769

Figure 8. The expectation and standard deviation of longevity, averaged over the stationary distribution of initial environments, for *Lomatium bradshawii*, as a function of the initial stage, the fire frequency f, and and the temporal autocorrelation ρ. Parameters as in Fig. 7.

transition and mortality probabilities in $\mathbf{U}$. As such, it can provide a null model for studies of heterogeneity in quantities such as the number of reproductive events. This idea has been explored independently, and in more detail, by Tuljapurkar et al. (2009).

The sensitivity of the variance is derived in Supplementary material Appendix A.1 as

$$\frac{d\text{vec}\,\mathbf{V}}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[2(\mathbf{N}^{\mathsf{T}}\otimes \mathbf{I}_s)\ \text{diag}\ (\text{vec}\ \mathbf{I}_s) + 2(\mathbf{I}_s \otimes \mathbf{N}_{\text{diag}}) - \mathbf{I}_{s^2} \right.$$

$$\left. - 2\text{diag}(\text{vec}\,\mathbf{N}) \right] \frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}}$$

(35)

Elasticities of $\mathbf{V}$ are calculated using (21).

Hint – before looking at Supplementary material Appendix A.1, to derive (35), write $\mathbf{N}_{\text{diag}} = \mathbf{I} \circ \mathbf{N}$, differentiate

(32), and use the fact that vec $(\mathbf{A} \circ \mathbf{B}) = \text{diag}\ (\text{vec}\,\mathbf{A})$ vec $\mathbf{B} = \text{diag}\ (\text{vec}\,\mathbf{B})\text{vec}\,\mathbf{A}$.

The right whale

The elasticities of $V(v_{41})$, calculated from (35) and (31), are shown in Figure 2b. They are roughly proportional to the elasticities of $E(v_{41})$; that is, the vital rates that have large effects on the expected number of reproductive events also have large effects on the variance.

Longevity and life expectancy

Longevity is an important demographic characteristic (Carey 2003). Mean longevity, or life expectancy, is one of the most widely reported demographic statistics, used to compare populations, species, countries, regions, historical periods, etc. and to examine the effects of evolutionary, management, medical, and social processes. The longevity

of an individual is the sum of the time spent in all of the transient states before final absorption. Let the random η_{ij} denote the longevity of an individual currently in stage j. Then

$$\eta_j = \sum_i \nu_{ij} \qquad (36)$$

A vector $E(\boldsymbol{\eta})$ of expected longevities, or life expectancies, is obtained by summing the columns of $\mathbf{N}$:

$$E(\boldsymbol{\eta}^T) = \mathbf{e}^T\mathbf{N} \qquad (37)$$

where $\mathbf{e}$ is a vector of ones. Often, life expectancy at birth is of primary interest. If stages are numbered so that birth corresponds to stage 1, then life expectancy at birth is

$$E(\eta_1) = \mathbf{e}^T\mathbf{N}\mathbf{e}_1 \qquad (38)$$

where $\mathbf{e}_1$ is a vector with 1 in the first entry and zeros elsewhere.

The sensitivity of life expectancy in age-classified models has been studied by Pollard (1982), Keyfitz (1977; see Keyfitz and Caswell 2005, section 4.3), and Vaupel (1986, Vaupel and Canudas Romo 2003). For stage-classified models, the sensitivity of $E(\boldsymbol{\eta})$ is (Caswell 2006)

$$\frac{dE(\boldsymbol{\eta})}{d\boldsymbol{\theta}^T} = (\mathbf{I}_s \otimes \mathbf{e}^T)(\mathbf{N}^T \otimes \mathbf{N})\frac{d\text{vec }\mathbf{U}}{d\boldsymbol{\theta}^T} \qquad (39)$$

Hint – to obtain (39), differentiate both sides of (37), apply the vec operator, and use (30) for the derivative of $\mathbf{N}$. See Supplementary material Appendix A.2 for the derivation.

The right whale

For the right whale, the vector of life expectancies is

$$E(\boldsymbol{\eta}^T) = \begin{pmatrix} 32.0 & 34.3 & 35.2 & 31.8 & 36.2 \end{pmatrix} \qquad (40)$$

Because mortality rates vary relatively little among stages, the life expectancies of the stages differ by only about 15%. Thus life expectancy for a calf implied by these data was 32 years. The elasticities of life expectancy to the vital rates are shown in Fig. 3. Life expectancy is most elastic to mature female survival σ_3, and less so to σ_2 and σ_4. This partly reflects the longer amount of time spent as a mature female, compared to an immature female or mother; see (7). The elasticity to the birth rate γ_3 is negative, because of the reduced survival of mothers. A 1% increase in γ_3 will lead to a 0.51% decrease in life expectancy. This is one possible measure of the cost of reproduction.

Variance in longevity

Like the number of visits to a transient state, longevity is a random variable, the variability of which is a measure of individual stochasticity. Individuals differ in longevity depending on the pathways taken from birth to death. This variance has been explored by human demographers, using life table methods, as one way of studying the inequality in life span generated by a given mortality schedule, and how that inequality has changed over time (e.g. Wilmoth and Horiuchi 1999, Shkolnikov et al. 2003, Edwards and Tuljapurkar 2005).

The variance of the time to absorbtion, calculated by Caswell (2006) from results in Iosifescu (1980, theorem 3.2), is

$$V(\boldsymbol{\eta}^T) = \mathbf{e}^T\mathbf{N}(2\mathbf{N} - \mathbf{I}) - E(\boldsymbol{\eta}^T) \circ E(\boldsymbol{\eta}^T) \qquad (41)$$

The sensitivity of the variance in longevity is

$$\frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^T} = \Big[2(\mathbf{N}^T \otimes \mathbf{e}^T) + 2(\mathbf{I}_s \otimes \mathbf{e}^T\mathbf{N}) - (\mathbf{I}_s \otimes \mathbf{e}^T)$$
$$- 2\text{diag}(E(\boldsymbol{\eta}))(\mathbf{I}_s \otimes \mathbf{e}^T) \Big](\mathbf{N}^T \otimes \mathbf{N})\frac{d\mathbf{U}}{d\boldsymbol{\theta}^T} \qquad (42)$$

The first entry of (42) gives the sensitivity of the variance in longevity starting in stage 1.

Hint – to derive (42), differentiate (41) and apply the vec operator and Roth's theorem to each term, using (39) for the derivative of $E(\eta)$. See section A.3 for details.

The right whale

For the right whale, the variance and standard deviation of longevity are given by

$$V(\boldsymbol{\eta})^T = \begin{pmatrix} 1157 & 1167 & 1172 & 1163 & 1172 \end{pmatrix} \qquad (43)$$

$$SD(\boldsymbol{\eta})^T = \begin{pmatrix} 34.0 & 34.2 & 34.2 & 34.1 & 34.2 \end{pmatrix} \qquad (44)$$

The life expectancy at birth of 32 years has a standard deviation of about 34 years. Note that this result implies a very long positive tail of longevity. The interpretation of this result is tricky; I will return to it in the Discussion.

The elasticities of the variance of longevity of a calf are shown in Fig. 3b. The variance in longevity is increased by increases in σ_3, less so by increases in σ_2 and σ_4. The pattern of the elasticities is strikingly similar to that of the elasticities of $E(\eta_1)$.

The net reproductive rate

In age-classified demography, the net reproductive rate R_0 measures lifetime reproductive output. It also appears in epidemiology, where it measures the potential of a disease to spread (e.g. Diekmann et al. 1990, van den Driessche and Watmough 2002). The classical net reproductive rate satisfies three conditions:

C_1: R_0 measures the expected lifetime production of offspring.

C_2: R_0 measures the rate of increase per generation (in contrast to the rate of increase per unit of time, which is given by λ or r).

C_3: R_0 is an indicator function for population persistence. If $R_0 > 1$ then an individidual will, on average, produce more than enough offspring to replace itself, the next generation will be larger than the present generation, and the population will grow. If $R_0 < 1$, each generation is smaller than the one before, and the population will decline to extinction.

In classical demography (Lotka 1939, Rhodes 1940),

$$R_0 = \int_0^\infty l(x)m(x)dx \qquad (45)$$

where $l(x)$ is survivorship to age x and $m(x)$ is the maternity function. It is not difficult to show that R_0 defined in this way satisfies conditions C_1, C_2, and C_3.

In stage-classified models, however, the calculation of R_0 must account for the multiple pathways that an individual may follow through the life cycle, and the production of multiple kinds of offspring along each of these pathways. Rogers (1974, Lebreton 1996) considered R_0 in the context of an age-classified population distributed across a set of spatial regions. However, these calculations assume that age-specific survival and fertility schedules are available for each region. A more general solution was provided by Cushing and Yicang (1994) for stage-classified populations with no age-specific information. Their analysis produces an index that satisfies as many as possible of the conditions C_1, C_2 and C_3. De Camino-Beck and Lewis (2007, 2008) have derived graph-theoretic ways to calculate R_0.

Consider an initial cohort at $t = 0$ with structure $\mathbf{x}_0$, and call this the first generation. This cohort will produce offspring according to $\mathbf{Fx}_0$. The survivors of the cohort at $t = 1$ will produce offspring according to $\mathbf{FUx}_0$. The survivors at $t = 2$ will produce offspring $\mathbf{FU}^2\mathbf{x}_0$, and so on. The second generation is composed of all the offspring of the first generation, obtained by summing over the lifetime of the cohort

$$\mathbf{x}(1) = \left(\mathbf{F} \sum_{i=0}^\infty \mathbf{U}^i \right) \mathbf{x}_0 = (\mathbf{FN})\mathbf{x}_0 \qquad (46)$$

Iterating this process leads to a model for the growth from one generation to the next

$$\mathbf{x}(k+1) = \mathbf{FNx}(k) \qquad (47)$$

Cushing and Yicang (1994) define R_0 as the per-generation growth rate, given by the dominant eigenvalue ρ of $\mathbf{FN}$,

$$R_0 = \rho[\mathbf{FN}] \qquad (48)$$

Thus the Cushing-Yicang measure of R_0 clearly satisfies condition C_2. Cushing and Yicang (1994) also prove (theorem 3) that R_0 defined in this way is less than, equal to, or greater than 1 if and only if λ is less than, equal to, or greater than one, respectively, thus satisfying condition C_3.

The relation between lifetime offspring production and R_0 (condition C_1) is more complicated when the life cycle contains multiple types of offspring. If only a single type of offspring is produced (call it stage 1), then $\mathbf{F}$ will have nonzero entries only in its first row, and $\mathbf{FN}$ will be upper triangular, with its dominant eigenvalue appearing in the (1, 1) position, i.e. the sum of the fertilities of each stage weighted by the expected time spent in that stage. This is precisely the expected lifetime offspring production, so for the case of a single type of offspring, the Cushing-Yicang R_0 also satisfies C_1.

However, if the life cycle contains multiple types of offspring (say stages 1, ..., h), the upper left $h \times h$ corner of $\mathbf{FN}$ will contain the expected lifetime production of offspring of types 1, ..., h, by individuals starting life as types 1, ..., h. Since such a life cycle contains more than

one kind of expected lifetime production of offspring, R_0 cannot satisfy C_1 in the sense of being *the* expected lifetime reproduction. Instead, R_0 is calculated from all these expectations (as the dominant eigenvalue of this $h \times h$ submatrix). It determines per-generation growth and population persistence as a function of the expected lifetime production of all types of offspring in a way that satisfies C_2 and C_3.

The right whale
The right whale produces only a single type of offspring. The fundamental matrix $\mathbf{N}$ is given by (7), the fertility matrix is given by (5), and the generation growth matrix is

$$\mathbf{FN} = \begin{pmatrix} 2.18 & 2.42 & 3.06 & 2.60 & 3.06 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (49)$$

The dominant eigenvalue of $\mathbf{FN}$ is its (1, 1) entry

$$R_0 = \sum_j f_{1j} E(\nu_{j1}) = 2.18 \qquad (50)$$

It is interesting to compare $R_0 = 2.18$ with $E(\nu_{14}) = 4.74$. Only female offspring are counted in R_0, whereas $E(\nu_{14})$ counts reproductive events regardless of the sex of the offspring produced. Still, R_0 is less than half of $E(\nu_{14})$, because of the less than perfect survival of calves from t to $t + 1$.

Net reproductive rate in periodic models
Periodic time-varying models (Caswell 2001, chapter 13) are an interesting special case of the multiple offspring type problem. In a periodic model, apparently identical offspring (e.g. seeds) produced at different phases of the cycle (e.g. seasons) are, in effect, of different types. To the extent that they face different environments, they will differ in their expected offspring production, and R_0 will differ depending on the phase of the cycle in which it is calculated.

The net reproductive rate in a periodic environment was calculated by Hunter and Caswell (2005a) in a study of the sooty shearwater, a pelagic seabird nesting on offshore islands in New Zealand. The year was divided into two short phases, in which breeding and harvest of chicks occur, and a longer phase encompassing the rest of the year. Let $\mathbf{B}_i = \mathbf{U}_i + \mathbf{F}_i$ be the projection matrix in phase i of the cycle. Without loss of generality, consider an environment with a period of 2 (e.g., winter and summer). The population is projected over a year, starting in phase 1, by

$$\mathbf{A}_1 = \mathbf{B}_2 \mathbf{B}_1 \qquad (51)$$

which is decomposed as

$$\begin{aligned} \mathbf{A}_1 &= (\mathbf{U}_2 + \mathbf{F}_2)(\mathbf{U}_1 + \mathbf{F}_1) \\ &= \mathbf{U}_2\mathbf{U}_1 + \mathbf{U}_2\mathbf{F}_1 + \mathbf{F}_2\mathbf{U}_1 + \mathbf{F}_2\mathbf{F}_1 \end{aligned} \qquad (52)$$

The first term includes only transitions, whereas the last three terms all describe some aspect of reproduction. Thus the annual matrix is $\mathbf{A}_1 = \hat{\mathbf{U}} + \hat{\mathbf{F}}$, where

$$\hat{\mathbf{U}}_1 = \mathbf{U}_2\mathbf{U}_1 \qquad (53)$$

$$\hat{\mathbf{F}}_1 = \mathbf{U}_2\mathbf{F}_1 + \mathbf{F}_2\mathbf{U}_1 + \mathbf{F}_2\mathbf{F}_1 \qquad (54)$$

and

$$R_0^{(1)} = \rho\left[\hat{\mathbf{F}}_1(\mathbf{I} - \hat{\mathbf{U}}_1)^{-1}\right] \qquad (55)$$

where the superscript 1 indicates that this is the net reproductive rate of a generation beginning in season 1. The corresponding matrices for a generation starting in season 2 are obtained from

$$\mathbf{A}_2 = \mathbf{B}_1\mathbf{B}_2 \qquad (56)$$

and lead to a net reproductive rate $R_0^{(2)}$. It is easily verified that $R_0^{(1)} \neq R_0^{(2)}$ in general. This contrasts with the population growth rate λ, which is independent of cyclic permutation of the seasons. However, since λ is the same for $\mathbf{A}_1$ and $\mathbf{A}_2$, it must be the case that $R_0^{(1)}$ and $R_0^{(2)}$ are both greater than or less than 1 together.

Sensitivity of the net reproductive rate

Since R_0 is obtained as an eigenvalue, its sensitivity to parameter changes is easy to derive. Let $\mathbf{x}$ and $\mathbf{y}$ be the right and left eigenvectors of $\mathbf{FN}$ corresponding to R_0. Then (Caswell 2006) the sensitivity of R_0 is

$$\frac{dR_0}{d\boldsymbol{\theta}^{\mathsf{T}}} = (\mathbf{y}^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes \mathbf{x}^{\mathsf{T}})\frac{d\text{vec}\,\mathbf{F}}{d\boldsymbol{\theta}^{\mathsf{T}}} + (\mathbf{y}^{\mathsf{T}} \otimes \mathbf{x}^{\mathsf{T}}\mathbf{F})(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N})\frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}}$$

$$(57)$$

The first term captures the effects of changing fertility, the second term captures effects of changes in survival and transitions. The derivation of (57) is given in Supplementary material Appendix A.4.

Hint – to derive (57), write $R_0 = \rho[\mathbf{FN}]$ and write dR_0 in terms of the right and left eigenvectors of $\mathbf{FN}$ and the differential of $\mathbf{FN}$. Then expand $d(\mathbf{FN}) = (d\mathbf{F})\mathbf{N} + \mathbf{F}d(\mathbf{N})$ and apply the vec operator and the chain rule. See (A-22) and (A-23).

The right whale

The elasticity of R_0 is shown in Fig. 4; R_0 is most elastic to σ_3, less so to σ_2 and σ_4. Remarkably, the elasticity of R_0 to the birth probability γ_3 is zero (actually, $\sim 10^{-9}$). This is a case where lifetime reproductive output is affected strongly by survival, slightly by maturation, but not at all by the probability of breeding given survival. This seems to be a consequence of the lower survival probability of mothers; an increase in γ_3 increases the probability of reproduction, but reduces the lifetime over which that reproduction will be realized.

Invasion exponents, selection gradients and R_0

Selection on life history traits can be studied in terms of the invasion exponent, which measures the rate at which a mutation, introduced at low densities, will increase in the environment created by a resident phenotype (Metz et al. 1992, Ferriére and Gatto 1993; for a recent introduction see Otto and Day 2007). The selection gradient on a trait is the derivative of the invasion exponent with respect to the

value of the trait. If the derivative is positive, selection favors an increase in the trait, and vice-versa. The invasion exponent in a density-independent model is given by log λ. In a density-dependent model, the invasion exponent is given by the growth rate at equilibrium, $\lambda[\hat{\mathbf{n}}]$. The net reproductive rate R_0 is not, strictly speaking, an invasion exponent, but because it measures expected lifetime reproduction, it is attractive as a measure of fitness (e.g. discussion in Kozlowski 1999). Using R_0 as a measure of fitness will lead to erroneous conclusions unless the selection gradients, measured in terms of λ and of R_0, give the same answers, i.e. unless $dR_0/d\theta \propto d\log\lambda/d\theta$.

For an age-classified model, we write R_0 in terms of the net maternity function $\phi(x,\theta) = l(x,\theta)m(x,\theta)$ where both survival and reproduction depend on some parameter θ. Then

$$R_0(\theta) = \int_0^\infty \phi(x,\theta)dx \qquad (58)$$

The growth rate $r = \log \lambda$ is the solution to

$$1 = \int_0^\infty \phi(x,\theta)e^{-r(\theta)x}dx \qquad (59)$$

Differentiating (58) and (59) gives

$$\frac{dR_0}{d\theta} = \int_0^\infty \frac{d\phi(x,\theta)}{d\theta}dx \qquad (60)$$

$$\frac{dr}{d\theta} = \frac{\int_0^\infty e^{-rx}\frac{d\phi(x,\theta)}{d\theta}dx}{\int_0^\infty x\phi(x,\theta)e^{-rx}dx} \qquad (61)$$

Equation (61) is Hamilton's (1966) famous result; the denominator is the generation time measured as the average age of reproduction in the stable age distribution.

When $R_0 = 1$ and $r = 0$, it follows from (60) and (61) that the gradients $dr/d\theta$ and $dR_0/d\theta$ are proportional. Use of either will lead to the same conclusions about selection. But when $r \neq 0$, this is not the case. If $r > 0$, then $dr/d\theta$ is reduced for traits that operate at later ages, because $d\phi/dx$ is weighted by e^{-rx}. It is an open problem to generalize this result to stage-classified models, and prove that

$$\frac{d\log\lambda}{d\boldsymbol{\theta}^{\mathsf{T}}} \propto \frac{dR_0}{d\boldsymbol{\theta}^{\mathsf{T}}} \qquad (62)$$

when $\lambda = R_0 = 1$. In a few cases I have examined, it appears to be true numerically. As the following example shows, it is certainly the case that when $\lambda \neq 1$, the derivatives are not generally proportional.

The right whale

The lack of proportionality between the selection gradients in terms of λ and of R_0 means that predictions can differ dramatically depending on which is used as the invasion exponent. This is particularly true when tradeoffs exist between two or more traits. For the right whale, $\lambda = 1.025$ and $R_0 = 2.183$. Figure 5 shows the sensitivity of λ and of R_0; while the patterns are similar, they are not proportional,

and the use of R_0 as an invasion exponent would result in erroneous predictions. Suppose a trait existed that would increase the birth probability γ_3 at the cost of a reduction in calf survival σ_1, with the cost measured by $c = -d\sigma_1/d\gamma_3$. An increase in this trait would be favored by selection provided that

$$c < \frac{\partial\lambda/\partial\gamma_3}{\partial\lambda/\partial\sigma_1} = 0.96 \qquad (63)$$

But if expected lifetime reproduction was used as an invasion exponent, the analysis would conclude that selection would favor an increase in the trait only if

$$c < \frac{\partial R_0/\partial\gamma_3}{\partial R_0/\partial\sigma_1} = 0.0 \qquad (64)$$

That is, according to R_0, any cost whatsoever of increased birth rate would prevent selection from favoring it. According to λ (and correctly, in this case), selection would favor increased birth rate provided that the cost was not too great. In spite of the superficial similarity of the patterns in Fig. 5, the evolutionary implications are quite different, reflecting the impact of timing of life history events on λ. The sensitivities of λ to σ_2 and γ_2, which influence early survival and the age at maturity, are larger than the sensitivities of R_0 to the same parameters.

Cohort generation time

Generation time measures the typical age at which offspring are produced, or the age at which the typical offspring is produced. It appears in the IUCN criteria for classifying threatened species (IUCN 2001) as well as in various evolutionary considerations. There are several definitions of generation time (Coale 1972); here we will examine the cohort generation time, defined as the mean age of production of offspring in a cohort of newborn individuals. From the definition it is clear why calculation of generation time is a challenge in stage-classified models, in which the age of parents does not appear. Moreover, in stage-classified models, individuals may be born into several stages (e.g. cleisthogamous vs chasmogamous seeds; LeCorff and Horvitz 2005), each with a different subsequent pattern of development, survival, and fertility. There could be a different generation time for each type of offspring, and if individuals may produce more than one type of offspring, the average age at which they are produced could differ from one kind of offspring to another.

Thus, we expect to have a generation time that measures the mean age of production of offspring of type i by an individual born in stage j. Write this as a vector $\mu^{(j)}$. Then it can be shown (Supplementary material Appendix A.5) that

$$\mu^{(j)} = \text{diag}\,(\mathbf{FNe}_j)^{-1}\,\mathbf{FNUNe}_j \qquad (65)$$

The sensitivity of $\mu^{(j)}$ is obtained by a methodical application of matrix calculus to (65). To simplify notation, define (for this calculation only)

$$\mathbf{X} = \text{diag}\,(\mathbf{FNe}_j) \qquad (66)$$
$$\mathbf{r} = \mathbf{FNUNe}_j \qquad (67)$$

The resulting sensitivity of $\mu^{(j)}$ is

$$\frac{d\mu^{(j)}}{d\theta^\top} = -(\mathbf{r}^\top \otimes \mathbf{I})(\mathbf{X}^{-1} \otimes \mathbf{X}^{-1})\,\text{diag}\,(\text{vec}\,\mathbf{I})$$

$$\times \left[(\mathbf{ee}_j^\top \mathbf{N}^\top \otimes \mathbf{I})\frac{\text{dvec}\,\mathbf{F}}{d\theta^\top} + (\mathbf{ee}_j \otimes \mathbf{F})\frac{\text{dvec}\,\mathbf{N}}{d\theta^\top} \right]$$

$$+ \left\{ [(\mathbf{NUNe}_j)^\top \otimes \mathbf{I}]\frac{\text{dvec}\,\mathbf{F}}{d\theta^\top} + [(\mathbf{UNe}_j)^\top \otimes \mathbf{F}] \right.$$

$$\times \frac{\text{dvec}\,\mathbf{N}}{d\theta^\top} + [(\mathbf{Ne}_j)^\top \otimes \mathbf{FN}]\frac{\text{dvec}\,\mathbf{U}}{d\theta^\top}$$

$$\left. + [\mathbf{e}_j^\top \otimes \mathbf{FNU}]\frac{\text{dvec}\,\mathbf{N}}{d\theta^\top} \right\} \qquad (68)$$

Hint – to derive (68), it helps to note that, for any vector z, one can write $\text{diag}(\mathbf{z}) = \mathbf{I} \circ \mathbf{ze}^\top$. Apply this to $\mathbf{X}$, differentiate all the terms in $\mu^{(j)}$, and apply the vec operator. With any luck, you will come out to this answer. See Supplementary material Appendix A.5.1 for derivation.

The right whale
The elasticities of the generation time $\mu^{(1)}$ of a calf are shown in Fig. 6. Changes in early survival (σ_1 and σ_2) have little effect. Adult survival σ_3 and, to a lesser extent, σ_4 increase the generation time by extending the reproductive lifespan. The maturation probability γ_2 and the birth probability γ_3 have negative effects on generation time, because they speed up reproduction.

Variable environments

All the results presented so far have pertained to constant environments. The extension to variable environments is obviously interesting, and there are exciting recent developments in this area (Caswell 2006, Tuljapurkar and Horvitz 2006). Several cases can be considered:

- Deterministic aperiodic environments. These usually appear as specific historical sequences; e.g. the specific sequence of vital rates exhibited by the right whale between 1980 and 1998 (Caswell 2006). That sequence is fixed, and is neither random nor periodic.
- Periodic environments. A periodic model may describe seasonal variation within a year, or may approximate inter-annual variability in events such as floods, fires, or hurricanes.
- Stochastic iid environments. In such environments, successive states are drawn independently from a fixed probability distribution; hence the identifier iid, short for 'independent and identically distributed.'
- Markovian stochastic environments. In a Markovian environment the probability distribution of the next environmental state may depend on the current state. This permits study of the effects of environmental

autocorrelation. Markovian environments include periodic and iid environments as special cases.

See Tuljapurkar (1990) for a thorough discussion of types of stochastic environments.

When studying variable environments, it is important to distinguish period and cohort calculations. Period calculations are based on the vital rates in a given year. They describe the results of the hypothetical situation where the conditions of year t are maintained indefinitely, and compare those to the results for conditions in year $t+1$, etc. Period calculations are a way to summarize the effects of changing environment. But an individual born in year t does not live its life under the conditions of year t. It spends its first year of life under the conditions in year t, its second year under the conditions of year $t+1$, and so on. Results calculated in this way are called cohort calculations, because they describe a cohort born in year t and living through the environmental sequence starting then. Period-specific calculations are easy; simply apply the time-invariant calculation to the vital rates of each year and tabulate the results. Cohort calculations, however, must account for all the possible environmental sequences through which a cohort may pass. Caswell (2006) and Tuljapurkar and Horvitz (2006) independently introduced two different, complementary approaches to doing so. I will present the former approach here.

After presenting the analysis, I will present an example using a model for a prairie plant in a stochastic fire environment. The matrices for this example are large, and are presented in Supplementary material Appendix B. I will refer to those matrices at appropriate places throughout the derivations.

A model for variable environments

In a variable environment, the transient matrix $\mathbf{U}$ is a time-varying matrix $\mathbf{U}(t)$. We can define a fundamental matrix by

$$\mathbf{N} = \mathbf{I} + \mathbf{U}(0) + \mathbf{U}(1)\mathbf{U}(0) + \mathbf{U}(2)\mathbf{U}(1)\mathbf{U}(0) + \cdots \quad (69)$$

The (i, j) element of $\mathbf{N}$ is the expected number of visits to transient state i by an individual starting in transient state j at time 0, and experiencing the specific sequence of environments $\mathbf{U}(0)$, $\mathbf{U}(1),\ldots$. Thus there will be a different matrix $\mathbf{N}$ for each possible environmental sequence.

Tuljapurkar and Horvitz (2006), whose paper I highly recommend, work directly from (69) to develop the means and variances of $\mathbf{N}$, $\boldsymbol{\eta}$, and survivorship, in periodic, iid, and Markovian environments. My previous paper (Caswell 2006) presented a different approach, in which an individual is jointly classified by its stage and its environment. This approach is an extension of the vec-permutation model developed by Hunter and Caswell (2005b) for spatial models. The relations between these two approaches, formal proofs of their equivalence, and extensions of both to additional demographic variables are intresting open problems.

Suppose that there are q environmental states $\varepsilon = 1,\ldots,$ q and s stages, $g = 1,\ldots,$ s. Corresponding to environment i is a $s \times s$ transient matrix $\mathbf{U}_i$. Assemble the matrices $\mathbf{U}_i$ into a block-diagonal matrix

$$\mathbb{U} = \begin{pmatrix} \mathbf{U}_1 & & \\ & \ddots & \\ & & \mathbf{U}_q \end{pmatrix} \quad (70)$$

of dimension $sq \times sq$ (see equation (B-1)).

The transitions among environmental states are defined by a $q \times q$ column-stochastic matrix $\mathbf{D}$. Use the matrix $\mathbf{D}$ to construct a block-diagonal environmental transition matrix

$$\mathbb{D} = \begin{pmatrix} \mathbf{D} & 0 & \cdots & 0 \\ 0 & \mathbf{D} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \mathbf{D} \end{pmatrix} \quad (71)$$

of dimension $sq \times sq$ (B-2).

Suppose that there are 4 environmental states. In an aperiodic deterministic environment

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad (72)$$

That is, the environment moves deterministically from state 1 to state 2 to state 3 to state 4. Setting $d_{44} = 1$ solves the problem of what to do at the end of the sequence, by the (possibly satisfactory) trick of letting the final state repeat indefinitely. In a periodic environment,

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (73)$$

In an iid environment in which environment i occurs with probability π_i,

$$\mathbf{D} = \begin{pmatrix} \pi_1 & \pi_1 & \pi_1 & \pi_1 \\ \pi_2 & \pi_2 & \pi_2 & \pi_2 \\ \pi_3 & \pi_3 & \pi_3 & \pi_3 \\ \pi_4 & \pi_4 & \pi_4 & \pi_4 \end{pmatrix} \quad (74)$$

In a Markovian environment, $\mathbf{D}$ is a column stochastic transition matrix describing the transition probabilities. I will assume that the environmental Markov chain is ergodic, with a stationary probability distribution denoted by $\boldsymbol{\pi}$. This gives the long-term frequency of occurrence of each environmental state.

The state of the cohort is described by a matrix $\mathbf{X}$, of dimension $s \times q$, with rows corresponding to stages and columns to environments, and where $x_{ij}(t)$ is the expected number of individuals in stage i and environmental state j at time t.

$$\mathbf{X}(t) = \begin{pmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & & \vdots \\ x_{s1} & \cdots & x_{sq} \end{pmatrix} \quad (75)$$

We rearrange $\mathbf{X}$ into a vector by applying the vec operator to $\mathbf{X}^\mathsf{T}$,

$$\text{vec } \mathbf{X}^T = \begin{pmatrix} x_{11} & \cdots & x_{1q} | \cdots & \cdots & \cdots | x_{s1} & \cdots & x_{sq} \end{pmatrix}^T \tag{76}$$

The first block of entries gives stage 1 individuals in environments 1 through q. The second block gives stage 2 individuals in environments 1 through q, and so on.

To describe the dynamics of the cohort, suppose that individuals first move among stages, according to the vital rates determined by the current environment, and then the environment changes to a new state according to $\mathbf{D}$. Then

$$\text{vec }^T\mathbf{X}(t+1) = \mathbb{D}\,\mathbf{K}_{s,q} \cup \mathbf{K}_{s,q}^T \text{vec}^T\mathbf{X}(t) \tag{77}$$

The matrix $\mathbf{K}_{s,q}$ is the vec-permutation matrix (Henderson and Searle 1981, Hunter and Caswell 2005b), also called the commutation matrix (Magnus and Neudecker 1979), which permutes the entries of a vector so that

$$\text{vec }^T\mathbf{X} = \mathbf{K}_{s,q}\text{vec } \mathbf{X} \tag{78}$$

Like all permutation matrices, its transpose is equal to its inverse (see Henderson and Searle 1981 or Hunter and Caswell 2005b for its calculation). In the model its role is to rearrange the population vector into a form appropriate for multiplication by the block-diagonal matrices $\mathbb{B}$ and $\mathbb{D}$.

Working from right to left, (77) first rearranges the vector, then applies the block-transition matrix $\mathbb{U}$, then reverses the rearrangement of the vector, and finally applies the environmental transition block matrix $\mathbb{D}$ to obtain the expected cohort at $t+1$. This gives a transition matrix for the joint process,

$$\tilde{\mathbf{U}} = \mathbb{D}\,\mathbf{K}_{s,q} \cup \mathbf{K}_{s,q}^T \tag{79}$$

that incorporates the demographic transitions within each environment and the patterns of time variation among environments (see (B-3) for an example). Here and in what follows, the tilde distinguishes the matrix from the environment-specific matrices. Note that (77) computes the expected population at $t+1$ from the expected population at t. If this approach was taken using projection matrices instead of transient transition matrices, the eigenvalues of the resulting product matrix would give the growth rate of the mean population, but not the stochastic growth rate (which is always less than or equal to the growth rate of the mean population). For calculations such as life expectancy, which are explicitly properties of the expected population, the difference does not arise.

Matrices of similar form, but not using this formalism, were introduced by Horvitz to study populations in habitat patches where the habitat patches change state over time, for example in recovering from disturbance (Horvitz and Schemske 1986, Pascarella and Horvitz 1998). Horvitz introduced the term "megamatrix" to describe these models. It can be shown (Caswell unpubl.) that a megamatrix arises as a special case of (79) when the population is classified by stages within environmental states, the demographic matrices are applied first, and the environmental transition matrices $\mathbf{D}_i$ are identical for all stages, as is the case in (71).

The fundamental matrix

Since $\tilde{\mathbf{U}}$ is the transient matrix of an absorbing Markov chain, the fundamental matrix in the time-varying environment is

$$\tilde{\mathbf{N}} = (\mathbf{I}_{sp} - \tilde{\mathbf{U}})^{-1} \tag{80}$$

The elements of $\tilde{\mathbf{N}}$ give the expected number of visits to each stage, in each environment, as a function of the starting stage and starting environment.

Notation alert

Any attempt to develop a complete system of notation for $\tilde{\mathbf{N}}$ would obscure more than it would clarify. However, pictures can help a great deal. As I present the fundamental matrix and some of the properties calculated from it, I will use diagrams for a simple case with three stages and two environments. I will often indicate the dimension of matrices and vectors with subscripts. I will use g to denote stages ($g = 1, 2, \ldots, s$) and ϵ to denote environments ($\epsilon = 1, \ldots, q$). I will use superscripts on $\tilde{\mathbf{N}}$ and quantities derived from it, to distinguish different ways of combining information across environmental states (Table 1).

Recall that in a constant environment, v_{ij} was the number of visits to stage i, starting in stage j. Now we must consider the visits to stage i in environment ϵ, starting in stage j and environment ϵ_0, so we write

$$\tilde{\mathbf{N}} = E(v_{ij,\epsilon}|\epsilon_0) \tag{81}$$

The structure of $\tilde{\mathbf{N}}$ when $s = 3$ and $q = 2$ is

Table 1. Superscript notation for time-varying models. The tilde on the variables indicates that they are calculated from the time-varying transient matrix $\tilde{\mathbf{U}}$ in (79). Visits to transient states and times to absorbtion depend on the initial and final demographic and environmental states. The superscripts (‡, §, ♡) indicate various choices of summing and averaging over the environmental states. The superscripts are shown here for the fundamental matrix $\tilde{\mathbf{N}}$.

Symbol	Definition	Description	Equation
$\tilde{\mathbf{N}}$	$E(v_{ij,\epsilon}\|\epsilon_0)$	Expected visits to state i in environment ϵ, starting from state j in environment ϵ_0	(81)
$\tilde{\mathbf{N}}^\ddagger$	$E(v_{ij}\|\epsilon_0)$	Expected visits to state i, summed over environments, starting from state j in environment ϵ_0	(82)
$\tilde{\mathbf{N}}^{\ddagger\ddagger}$		Rearrangement of the rows and columns of $\tilde{\mathbf{N}}^\ddagger$	(83)
$\tilde{\mathbf{N}}^\S$	$E(v_{ij,\epsilon})$	Expected visits to state i and environmental state ϵ, averaged over initial environmental states.	(84)
$\tilde{\mathbf{N}}^{\S\S}$		Rearrangement of the rows and columns of $\tilde{\mathbf{N}}^\S$	(85)
$\tilde{\mathbf{N}}^\heartsuit$	$E(v_{ij})$	Expected visits to state i summed over environments, starting from state j and averaged over initial environmental states.	(86)

$\widetilde{\mathbf{N}}_{sq \times sq}$:

		$g = 1$		$g = 2$		$g = 3$	
		$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$
$g = 1$	$\epsilon = 1$						
	$\epsilon = 2$						
$g = 2$	$\epsilon = 1$						
	$\epsilon = 2$						
$g = 3$	$\epsilon = 1$						
	$\epsilon = 2$						

(see (B-4) for an example). From $\tilde{\mathbf{N}}$ we can obtain the expected number of visits to each stage, regardless of the environment in which those visits occur, by aggregating rows. The resulting matrix $\mathbf{N}^{\ddagger}$ is

$$\tilde{\mathbf{N}}^{\dagger} = E(\nu_{ij} | \epsilon_0) = (\mathbf{I}_s \otimes \mathbf{e}_{q \times 1}^{\mathsf{T}}) \tilde{\mathbf{N}} \tag{82}$$

where $\mathbf{e}_{q \times 1}$ is a vector of ones. The structure of $\mathbf{N}^{\ddagger}$ is:

$\widetilde{\mathbf{N}}_{s \times sq}^{\ddagger}$:

	$g = 1$		$g = 2$		$g = 3$	
	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$
$g = 1$						
$g = 2$						
$g = 3$						

(see (B-5) for an example). If it is useful to group stages within initial environments, rather than grouping environments within stages, $\mathbf{N}^{\ddagger}$ can be rearranged as

$$\tilde{\mathbf{N}}^{\ddagger\ddagger} = \tilde{\mathbf{N}}^{\ddagger} \mathbf{K}_{s,q} \tag{83}$$

with the structure

$\widetilde{\mathbf{N}}_{s \times sq}^{\ddagger\ddagger}$:

	$\epsilon_0 = 1$			$\epsilon_0 = 2$		
	$g = 1$	$g = 2$	$g = 3$	$g = 1$	$g = 2$	$g = 3$
$g = 1$						
$g = 2$						
$g = 3$						

The matrices $\tilde{\mathbf{N}}^{\ddagger}$ and $\tilde{\mathbf{N}}^{\ddagger\ddagger}$ both display expected visits to each stage as a function of initial environment. To describe the fates of individuals without specifying their initial environment, we take an expectation over the stationary distribution $\boldsymbol{\pi}$ of initial environments. This gives

$$\tilde{\mathbf{N}}^{\S} = E[\nu_{ij,\epsilon}] = \tilde{\mathbf{N}}(\mathbf{I}_s \otimes \boldsymbol{\pi}) \tag{84}$$

The structure of $\tilde{\mathbf{N}}^{\S}$ is

$\widetilde{\mathbf{N}}_{sq \times s}^{\S}$:

		$\epsilon_0 = \bar{\epsilon}$		
		$g = 1$	$g = 2$	$g = 3$
$g = 1$	$\epsilon = 1$			
	$\epsilon = 2$			
$g = 2$	$\epsilon = 1$			
	$\epsilon = 2$			
$g = 3$	$\epsilon = 1$			
	$\epsilon = 2$			

(see (B-6) for an example). The rows of $\tilde{\mathbf{N}}^{\S}$ can be rearranged to display stages within environments, giving

$$\tilde{\mathbf{N}}^{\S\S} = \mathbf{K}_{s,q}^{\mathsf{T}} \tilde{\mathbf{N}}^{\S} \tag{85}$$

with the structure

$\widetilde{\mathbf{N}}_{sq \times s}^{\S\S}$:

		$\epsilon_0 = \bar{\epsilon}$		
		$g = 1$	$g = 2$	$g = 3$
$\epsilon = 1$	$g = 1$			
	$g = 2$			
	$g = 3$			
$\epsilon = 2$	$g = 1$			
	$g = 2$			
	$g = 3$			

Finally, aggregating over destination environments and averaging over initial environments gives a matrix containing the expected visits to stages as a function of initial stage, averaged over environments

$$\tilde{\mathbf{N}}^{\heartsuit} = E[\nu_{ij}] = (\mathbf{I}_s \otimes \mathbf{e}_{q \times 1}^{\mathsf{T}}) \tilde{\mathbf{N}}(\mathbf{I}_s \otimes \boldsymbol{\pi}) \tag{86}$$

The structure of $\tilde{\mathbf{N}}^{\heartsuit}$ is

$\widetilde{\mathbf{N}}_{s \times s}^{\heartsuit}$:

	$\epsilon_0 = \bar{\epsilon}$		
	$g = 1$	$g = 2$	$g = 3$
$g = 1$			
$g = 2$			
$g = 3$			

(see (B-7) for an example). The matrix $\bar{\mathbf{N}}^{\heartsuit}$, obtained by the simple calculation (86), is 'the' fundamental matrix for the variable environment. It could be compared directly to the fundamental matrix in a constant environment (e.g. the environment defined by one of the environmental states).

Longevity in a variable environment

Life expectancy, as a function of initial stage and initial environment, is obtained by summing the columns of $\tilde{\mathbf{N}}$,

$$E(\tilde{\boldsymbol{\eta}}^{\mathsf{T}}) = E(\boldsymbol{\eta}^{\mathsf{T}} | \epsilon_0) = \mathbf{e}_{sq \times 1}^{\mathsf{T}} \tilde{\mathbf{N}} \tag{87}$$

The structure of $E(\tilde{\boldsymbol{\eta}}^{\mathsf{T}})$ is:

$E\left(\tilde{\boldsymbol{\eta}}^{\mathsf{T}}\right)$:

$g = 1$		$g = 2$		$g = 3$	
$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$	$\epsilon_0 = 1$	$\epsilon_0 = 2$

(see (B-8) for an example). Averaging this conditional life expectancy over the stationary distribution $\boldsymbol{\pi}$ of initial environments gives

$$E(\tilde{\boldsymbol{\eta}}^{\heartsuit}) = E(\tilde{\boldsymbol{\eta}})(\mathbf{I}_s \otimes \boldsymbol{\pi}) \tag{88}$$

(see (B-9) for an example). This measure of life expectancy in a variable environment is directly comparable to $E(\boldsymbol{\eta})$ calculated from the same life history in a constant environment.

Variance in longevity

In a constant environment, the variance among individuals in longevity is due to individual stochasticity. In a time-varying environment, the variance contains an additional component due to differences among individuals as a function of their environment at birth. Applying (41) to $\tilde{\mathbf{N}}$ we obtain the variances conditional on the initial environment:

$$V[\tilde{\boldsymbol{\eta}}^\mathsf{T}|\epsilon_0] = E(\tilde{\boldsymbol{\eta}}^\mathsf{T})(2\tilde{\mathbf{N}} - \mathbf{I}_{sq}) - E(\tilde{\boldsymbol{\eta}}^\mathsf{T}) \circ E(\tilde{\boldsymbol{\eta}}^\mathsf{T}) \qquad (89)$$

As indicated by the notation, $V[\tilde{\boldsymbol{\eta}}^\mathsf{T}|\epsilon_0]$ is a conditional variance of $\tilde{\boldsymbol{\eta}}$, given the initial environment ϵ_0. The initial environment is distributed according to the stationary distribution $\boldsymbol{\pi}$, so the unconditional longevity $\boldsymbol{\eta}$ follows a finite mixture distribution with mixing distribution $\boldsymbol{\pi}$. The unconditional variance of $\boldsymbol{\eta}$, taking account of both sources of variability, is

$$V[\tilde{\boldsymbol{\eta}}^\mathsf{T}] = V[E(\tilde{\boldsymbol{\eta}}^\mathsf{T}|\epsilon_0)] + E_\pi[V(\tilde{\boldsymbol{\eta}}^\mathsf{T}|\epsilon_0)] \qquad (90)$$

where E_π denotes the expectation over the stationary distribution $\boldsymbol{\pi}$ of initial environments (Rényi 1970, p. 275, theorem 1). This can be rearranged as

$$V[\tilde{\boldsymbol{\eta}}^\mathsf{T}] = E_\pi[\tilde{\boldsymbol{\eta}}^\mathsf{T} \circ \tilde{\boldsymbol{\eta}}^\mathsf{T}] - E_\pi[\tilde{\boldsymbol{\eta}}^\mathsf{T}] \circ E_\pi[\tilde{\boldsymbol{\eta}}^\mathsf{T}] + E_\pi[V(\boldsymbol{\eta}^\mathsf{T}|\epsilon_0)]$$

$$= [E(\tilde{\boldsymbol{\eta}}^\mathsf{T}) \circ E(\tilde{\boldsymbol{\eta}}^\mathsf{T})](\mathbf{I}_s \otimes \boldsymbol{\pi}) - [E(\tilde{\boldsymbol{\eta}}^\heartsuit) \circ E(\tilde{\boldsymbol{\eta}}^\heartsuit)]^\mathsf{T}$$

$$+ V[\tilde{\boldsymbol{\eta}}^\mathsf{T}|\epsilon_0](\mathbf{I}_s \otimes \boldsymbol{\pi}) \qquad (91)$$

(Frühwirth-Schnatter 2006, p. 10)

A time-varying example: *Lomatium bradshawii*

Lomatium bradshawii is an endangered herbaceous perennial plant, found in only a few isolated populations in prairies of Oregon and Washington. These habitats were, until recent times, subject to natural and anthropogenic fires, to which *L. bradshawii* seems to have adapted. Fall-season fires increase plant size and seedling recruitment, but the effect fades within a few years. Populations in burned areas have higher growth rates and lower probabilities of extinction than unburned populations (Caswell and Kaye 2001).

A stochastic demographic model for *L. bradshawii* was developed by Caswell and Kaye (2001, Kaye et al. 2001, Kaye and Pyke 2003), based on data from an experimental study using controlled burning. Individuals were classified into six stages based on size and reproductive status: yearlings, small and large vegetative plants, and small, medium, and large reproductive plants. The environment was classified into four states defined by fire history: the year of a fire and 1, 2, and 3+ years post-fire. Projection matrices were estimated in each environment; the example here is based on one of the two sites (Rose Prairie) in the original study. The matrices are given in Caswell and Kaye (2001).

L. bradshawii performs well under recently burned conditions, but less well in sites that have not been recently burned. For example, the values of λ are

Years since fire:	0	1	2	≥ 3
Growth rate λ :	1.18	1.12	0.48	0.88

Caswell and Kaye (2001) found a minimum frequency of fire (0.4–0.5) below which the stochastic growth rate was negative and the population would be unable to persist. Effects of autocorrelation were small, but positive autocorrelation reduced the stochastic growth rate.

As an example of a time-varying analysis, let us examine *L. bradshawii* in a Markovian environment. Let f be the long-term frequency of fire, and ρ the temporal autocorrelation. Then the transition matrix for environmental states is

$$\mathbf{D} = \begin{pmatrix} p & q & q & q \\ 1-p & 0 & 0 & 0 \\ 0 & 1-q & 0 & 0 \\ 0 & 0 & 1-q & 1-q \end{pmatrix} \qquad (92)$$

where $q = f(1-\rho)$ and $p = \rho + q$.

The matrices involved in the analysis are given in Supplementary material Appendix B. The first step is to assemble the block diagonal matrices $\mathbb{U}$ and $\mathbb{D}$, which are shown in (B-1) and (B-2), respectively. Combining these with the vec-permutation matrix leads to the time-varying transient matrix given in $\acute{\mathbf{U}}$ (B-3). The time-varying fundamental matrix $\tilde{\mathbf{N}} = (\mathbf{I}_{24} - \acute{\mathbf{U}})^{-1}$ is given in (B-4).

Figure 7a shows the life expectancy $E(\tilde{\eta}|\epsilon_0)$ of *L. bradshawii* as a function of initial stage and initial environmental state, from (87). Life expectancy increases with the stage (size) of a plant. A seedling has its greatest life expectancy in the year of a fire, less in an environment three or more years post-fire. A large flowering plant, in contrast, has its greatest life expectancy in an environment three or more years post-fire. When the environment-dependence is averaged over the stationary distribution of environmental states, there is a smooth increase in life expectancy from ~ 2.5 years for a seedling to 8 years for a large flowering plant (Fig. 7b). The standard deviation of longevity also increases with stage, in a pattern very similar to that of the expectation.

These patterns in the mean and variance of longevity (Fig. 7) depend on the stochastic properties of the environment — in this case, the frequency f and auto-correlation ρ of fires. Even with an environmental model this simple, the effects of f and ρ can be complicated. I know of no previous attempts to examine their effects on longevity. To do so, I calculated life expectancy with $f = 0.5$ for autocorrelation $-1 < \rho < 1$, and with $\rho = 0$ for fire frequency $0 < f < 1$.

The life expectancy of early life cycle stages increases monotonically with fire frequency (Fig. 8a), but the life expectancy of large reproductive plants is greatest at either low or high fire frequencies. The standard deviation of longevity increases with f (Fig. 8b). As $f \to 1$, the standard deviation of longevity is approximately twice the mean.

The autocorrelation of fires has little effect on the life expectancy of seedlings, but a larger effect on that of large plants. For the latter, life expectancy is maximized as $\rho \to -1$ (alternating fire and non-fire years) or as $\rho \to 1$ (long periods of fires alternating with long periods without fire). The standard deviation of longevity also shows a strong U-shaped response to ρ for all stages. The generality of this pattern is unknown.

Sensitivity analysis and extensions

It will be an enjoyable task to extend the sensitivity analysis to accommodate variable environments, both in this approach and in the Tuljapurkar–Horvitz approach, but it is an open problem at this time. These two approaches are complementary. The Tuljapurkar–Horvitz approach lends itself to connections with the theory of random matrix products (e.g. in deriving asymptotic growth rates of variances in survivorship; see Horvitz and Tuljapurkar 2008), but my approach does not. On the other hand, the Tuljapurkar-Horvitz approach requires separate calculations for each environment. My approach obtains results for all initial environments in a single matrix multiplication, albeit using the complicated matrix $\tilde{U}$ constructed using the vec-permutation matrix. Because $\tilde{U}$ is a transient matrix for a single Markov chain, it can be extended to calculate other properties of absorbing chains, including multiple absorbing states, conditional chains, and passage times (Caswell 2001, 2006). The matrix calculus approach makes for easy computations, but usually does not yield readily interpretable formulas. There are many interesting and challenging open problems.

Discussion

Taking advantage of the Markov chain embedded within a population projection matrix opens up a wealth of demographic information. The age-classified information extracted from a stage-classified model can form a valuable component of behavioral studies, especially if the model (like the right whale example) includes reproductive behavior as part of the life cycle structure. Longevity provides a powerful way to compare mortality schedules among species, populations, or environmental conditions, but it has been inaccessible to stage-classified analysis prior to the development of Markov chain methods. The generation time characterizes an important population time scale, with implications in conservation (IUCN 2001), but there has been no way to compute it from stage-classified models. Absorbing Markov chains form the basis of multi-state mark-recapture statistical methods (Arnason 1973, Brownie et al. 1993, Fujiwara and Caswell 2002, Lebreton and Pradel 2002, Caswell and Fujiwara 2004, Lebreton et al. 2009), and thus provide natural links to parameter estimation. Thus, although I have not discussed estimation here, all the results for the right whale are maximum likelihood estimates. With a little extra effort, it would be possible to associate confidence intervals with, say, the elasticities of the variance in longevity to changes in survival. The Markov chain approach makes many demographic properties empirically estimable from data.

A Markov chain naturally treats individual trajectories (i.e. individual lives) as stochastic realizations of an underlying stochastic process. This individual stochasticity is a neglected component of demography. It complements environmental stochasticity (externally imposed random changes in vital rates) and demographic stochasticity (randomness in the number of survivors when applying a rate to a population). Individual stochasticity reflects randomness in the pathways that individuals take through the life cycle. It expresses itself in inter-individual variation in such quantities as the number of visits to states or longevity.

Individual stochasticity must be distinguished from individual heterogeneity, which results from differences among individuals in their inherent properties (e.g. frailty). Heterogeneity has important consequences for demography (Vaupel et al. 1979, Vaupel and Yashin 1985, Vaupel and Carey 1993) and evolution (where it determines the opportunity for selection). However, simply documenting inter-individual variance does not demonstrate the existence of heterogeneity. A standard of comparison is needed, and the calculation of individual stochasticity provides such a standard (Tuljapurkar et al. 2009).

Stage-classified life cycles may have consequences that are not yet appreciated, but must be considered when interpreting the results. For example, any stage-classified model eventually leads to an age-independent mortality rate (Horvitz and Tuljapurkar 2008), and so is of limited use in the study of senescence. This fact has consequences for life expectancy and variance in longevity that are not well understood (at least by me). For the right whale, expected longevity at birth is 32 years with a standard deviation of 34 years. It is unlikely that there are appreciable numbers of whales alive at even one standard deviation above this mean. The high survival probability and the assumption of age-independence lead to the high standard deviation. Those of us who work with stage-classified models are accustomed to this. We discount its importance because it (often) has little effect on λ, but it will be important to determine the consequences of the simplifying assumptions represented by the self-loops on the life cycle graph.

The elasticity results shown here reveal an interesting pattern. For the right whale, the elasticities of the mean and of the variance of longevity and the number of reproductive events show a surprisingly similar pattern (Fig. 2, 3). A possible explanation can be found by examining the elasticity of absorbtion time in a simple unstructured model, with one transient state (alive) and one absorbing state (dead). Let σ be the survival probability. The time to absorbtion in this model has a geometric distribution with mean E and variance V given by

$$E = \frac{1}{1 - \sigma} \tag{93}$$

$$V = \frac{\sigma}{(1 - \sigma)^2} \tag{94}$$

Differentiating E and V, and a bit of algebra, suffices to show that, for any parameter θ,

$$\frac{E}{V}\left(\frac{dV/d\theta}{dE/d\theta}\right) = \frac{1 + \sigma}{\sigma} \tag{95}$$

That is, the ratio of the elasticities of the mean and of the variance is a constant over all parameters. Perhaps the near-proportionality of the elasticities in Fig. 2 and 3 is because, as far as survival is concerned, the right whale is similar to the simple one-stage life cycle. The differences in survival among the stages are small ($0.85 \leq \sigma_i \leq 1$).

This paper does not begin to exhaust the information that can be extracted from the Markov chain formulation of

a stage-classified model. Two examples of particular interest are competing risks and passage times. If more than one absorbing state exists (e.g. death at different stages, or from different causes), then the risks of absorbtion compete, because an individual can only be absorbed (i.e. die) once. It is possible to calculate the probability of absorbtion in each state, and to explore the effects of changing one risk on the probability of experiencing another (Feichtinger 1971, Caswell 2001, 2006). Passage times refer to the time required to get from one stage to another in the life cycle. An important passage time is the birth interval: the time from one birth to the next. This can only be calculated for individuals that do reproduce a second time (otherwise the interval is infinite), and so it requires developing a chain that is conditional on successfully reaching the reproductive state (Caswell 2001 unpubl.). In species that produce only one or a few offspring, reproduction cannot be adjusted in response to the environment by changing offspring number, and so changes in the birth interval are particularly important in such species.

There has been an explosion of new demographic analyses in the last two decades. Both theoretical and empirical research has benefited, and our ability to explore the environmental and evolutionary determinants of the 'distribution and abundance of animals' (and plants!) has never been greater. The theoretical part needs no extra justification: these are interesting questions in themselves (by the most basic of criteria: they lead to more, interesting, questions). But it is always even more exciting when theoretical research connects with, and enlightens, empirical investigation. I have a personal fondness for that kind of research, and I hope that this paper contributes to it.

References

Abadir, K. M. and Magnus, J. R. 2005. Matrix algebra. Econometric exercises 1. – Cambridge Univ. Press.

Arnason, A. N. 1973. The estimation of population size, migration rates, and survival in stratified populations. – Res. Popul. Ecol. 15: 1–8.

Brownie, C. et al. 1993. Capture–recapture studies for multiple strata including non-Markovian transitions. – Biometrics 49: 1173–1187.

Carey, J. R. 2003. Longevity: the biology and demography of life span. – Princeton Univ. Press.

Caswell, H. 1989. Matrix population models (1st ed.). – Sinauer.

Caswell, H. 2001. Matrix population models: construction, analysis and interpretation (2nd ed.). – Sinauer.

Caswell, H. 2006. Applications of Markov chains in demography. – In: Langville, A. N. and Stewart, W. J. (eds), MAM2006: Markov anniversary meeting. Boson Books, Raleigh, North Carolina, USA, pp. 319–334.

Caswell, H. 2007. Sensitivity analysis of transient population dynamics. – Ecol. Lett. 10: 1–15.

Caswell, H. 2008. Perturbation analysis of nonlinear matrix population models. – Demogr. Res. 18: 59–116.

Caswell, H. 2009. Sensitivity and elasticity of density-dependent population models. – J. Difference Eq. Appl. 15: 349–369.

Caswell, H. and Fujiwara, M. 2004. Beyond survival estimation: mark–recapture, matrix population models, and population dynamics. – Anim. Biodiv. Conserv. 27: 471–488.

Caswell, H. and Kaye, T. N. 2001. Stochastic demography and conservation of an endangered perennial plant (*Lomatium bradshawii*) in a dynamic fire regime. – Adv. Ecol. Res. 32: 1–51.

Caswell, H. et al. 1997. Structured population models: many methods, a few general principles. – In: Tuljapurkar, S. and Caswell, H. (eds), Structured population models in marine, terrestrial and freshwater systems. Chapman and Hall, pp. 3–18.

Caswell, H. et al. 1999. Declining survival probability threatens the North Atlantic right whale. – Proc. Natl Acad. Sci. USA 96: 3308–3313.

Coale, A. J. 1972. The growth and structure of human populations. – Princeton Univ. Press.

Cochran, M. E. and Ellner, S. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. – Ecol. Monogr. 62: 345–364.

Crone, M. J. and Kraus, S. D. 1990. Right whale (*Eubalaena glacialis*) in the western North Atlantic: a catalog of identified individuals. – New England Aquarium, Boston, MA.

Cushing, J. M. and Yicang, Z. 1994. The net reproductive value and stability in matrix population models. – Nat. Resour. Model. 8: 297–333.

DeAngelis, D. L. and Gross, L. J. 1992. Individual based models and approaches in ecology. – Chapman and Hall.

de Camino-Beck, T. and Lewis, M. A. 2007. A new method for calculating net reproductive rate from graph reduction with applications to the control of invasive species. – Bull. Math. Biol. 69: 1341–1354.

de Camino-Beck, T. and Lewis, M. A. 2008. On net reproductive rate and the timing of reproductive output. – Am. Nat. 172: 128–139.

Diekmann, O. et al. 1990. On the definition and the computation of the basic reproduction number R_0 in models for infectious diseases in heterogeneous population. – J. Math. Biol. 28: 365–382.

Edwards, R. D. and Tuljapurkar, S. 2005. Inequality in life spans and a new perspective on mortality convergence across industrialized countries. – Popul. Dev. Rev. 31: 645–674.

Feichtinger, G. 1971. Stochastische modelle demographischer prozesse. Lecture notes in operations research and mathematical systems. – Springer.

Feichtinger, G. 1973. Markovian models for some demographic processes. – Stat. Hefte 14: 310–334.

Ferriére, R. and Gatto, M. 1993. Chaotic population dynamics can result from natural selection. – Proc. R. Soc. Lond. B 251: 33–38.

Fujiwara, M. 2002. Mark–recapture statistics and demographic analysis. PhD thesis. – MIT/WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering.

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

Früuhwirth-Schnatter, S. 2006. Finite mixture and Markov switching models. – Springer.

Fujiwara, M. and Caswell, H. 2002. Estimating population projection matrices from multi-stage mark–recapture data. – Ecology 83: 3257–3265.

Goldman, N. 1994. Social factors and health: the causation-selection issue revisited. – Proc. Natl Acad. Sci. USA 91: 1251–1255.

Hamilton, W. D. 1966. The moulding of senescence by natural selection. – J. Theor. Biol. 12: 12–45.

Henderson, H. V. and Searle, S. R. 1981. The vec-permutation matrix, the vec operator and Kronecker products: a review. – Linear Multilinear Algebra 9: 271–288.

Hoem, J. M. 1969. Markov chain models in life insurance. – Sonderdruck aus Blatter der Deutschen Gelellschaft für Versicherungsmathematik 9: 91–107.

Horvitz, C. C. and Schemske, D. W. 1986. Seed dispersal and environmental heterogeneity in a neotropical herb: a model of population and patch dynamics. – In: Estrada, A. and Fleming, T. H. (eds), Frugivores and seed dispersal. Dr W. Junk Publishers, pp. 169–186.

Horvitz, C. C. and Tuljapurkar, S. 2008. Stage dynamics, period survival, and mortality plateaus. – Am. Nat. 172: 203–215.

Hunter, C. M. and Caswell, H. 2005a. Selective harvest of sooty shearwater chicks: effects on population dynamics and sustainability. – J. Anim. Ecol. 74: 589–600.

Hunter, C. M. and Caswell, H. 2005b. The use of the vec-permutation matrix in spatial matrix population models. – Ecol. Modell. 188: 15–21.

Iosifescu, M. 1980. Finite Markov processes and their applications. – Wiley.

IUCN 2001. IUCN Red List categories and criteria: ver. 3.1. IUCN Species Survival Commission.

Jasny, B. et al. 2006. Science looks at life: special section. – Science 312: 1893.

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

Kaye, T. N. et al. 2001. The effect of fire on the population viability of an endangered prairie plant. – Ecol. Appl. 11: 1366–1380.

Kemeny, J. G. and Snell, J. L. 1976. Finite Markov chains. – Springer.

Keyfitz, N. 1967. Reconciliation of population models: matrix, integral equation and partial fraction. – J. R. Stat. Soc. 130: 61–83.

Keyfitz, N. 1977. Applied mathematical demography. – Wiley.

Keyfitz, N. and Caswell, H. 2005. Applied mathematical demography (3rd ed.). – Springer.

Kozlowski, J. 1999. Adaptation: a life history perspective. – Oikos 86: 185–194.

Kraus, S. D. and Rolland, R. M. 2007. The urban whale: North Atlantic right whales at the crossroads. – Harvard Univ. Press.

Land, K. C. and Rogers, A. 1982. Multidimensional mathematical demography. – Academic Press.

Lebreton, J.-D. 1996. Demographic models for subdivided equations: the renewal equation approach. – Theor. Popul. Biol. 49: 291–313.

Lebreton, J-D. and Pradel, R. 2002. Multistate recapture models: modelling incomplete individual histories. – J. Appl. Stat. 29: 353–369.

Lebreton, J.-D. et al. 2009. Modeling individual animal histories with multistate capture–recapture models. – Adv. Ecol. Res. 41: 85–166.

LeCorff, J. and Horvitz, C. C. 2005. The contrasting roles of chasmogamous and cleistogamous reproduction in population growth and population spread in an ant-dispersed neotropical herb, Calathea micans (Marantaceae). – Ecol. Modell. 188: 41–51.

Lotka, A. J. 1939. Théorie analytique des associations biologiques. Part II. Analyse démographique avec application particulière à l'espèce humaine. – Actualités Scientifiques et Industrielles, no. 780. Hermann et Cie, Paris, France. (published in translation as: Analytical Theory of Biological Populations, translated by D. P. Smith and H. Rossert. Plenum Press, 1998)

Magnus, J. R. and Neudecker, H. 1979. The commutation matrix: some properties and applications. – Ann. Stat. 7: 381–394.

Magnus, J. R. and Neudecker, H. 1985. Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. – J. Math. Psychol. 29: 474–492.

Magnus, J. R. and Neudecker, H. 1988. Matrix differential calculus with applications in statistics and econometrics. – Wiley.

Metz, J. A. J. and Diekmann, O. 1986. The dynamics of physiologically structured populations. – Springer.

Metz, J. A. J. et al. 1992. How should we define 'fitness' for general ecological scenarios? – Trends Ecol. Evol. 7: 198–202.

Moore, M. J. et al. 2005. Morphometry, gross morphology and available histopathology in North Atlantic right whale (Eubalaena glacialis) mortalities (1970–2002). – J. Cetacean Res. Manage. 6: 199–214.

Nel, D. G. 1980. On matrix differentiation in statistics. – S. Afr. Stat. J. 14: 137–193.

Nisbet, R. M. and Gurney, W. S. C. 1982. Modelling fluctuating populations. – Wiley.

Otto, S. P. and Day, T. 2007. A biologist's guide to mathematical modeling in ecology and evolution. – Princeton Univ. Press.

Pascarella, J. B. and Horvitz, C. C. 1998. Hurricane disturbance and the population dynamics of a tropical understory shrub: megamatrix elasticity analysis. – Ecology 79: 547–563.

Pollard, J. H. 1982. The expectation of life and its relationship to mortality. – J. Inst. Actuaries 109: 225–240.

Rényi, A. 1970. Probability theory. – Elsevier.

Reeves, R. R. et al. 2007. Near-annihilation of a species: right whaling in the North Atlantic. – In: Kraus, S. D. and Rolland, R. M. (eds), The urban whale: north Atlantic right whales at the crossroads. Harvard Univ. Press, pp. 39–74.

Rhodes, E. C. 1940. Population mathematics. I. – J. R. Stat. Soc. 103: 61–89.

Robine, J.-M. et al. 2003. Determining health expectancies. – Wiley.

Rogers, A. 1974. The multiregional net maternity function and multiregional stable growth. – Demography 11: 473–481.

Rolland, R. M. et al. 2007. The inner whale: hormones, biotoxins and parasites. – In: Kraus, S. D. and Rolland, R. M. (eds), The urban whale: north Atlantic right whales at the crossroads. Harvard Univ. Press, pp. 232–272.

Roth, W. E. 1934. On direct product matrices. – Bull. Am. Math. Soc. 40: 461–468.

Shkolnikov, V. M. et al. 2003. Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples. – Demogr. Res. 8: 305–358.

Tuljapurkar, S. 1990. Population dynamics in variable environments. – Springer.

Tuljapurkar, S. and Caswell, H. 1997. Structured population models in marine, terrestrial and freshwater systems. – Chapman and Hall.

Tuljapurkar, S. and Horvitz, C. C. 2006. From stage to age in variable environments: life expectancy and survivorship. – Ecology. 87: 1497–1509.

Tuljapurkar, S. et al. 2009. Dynamic heterogeneity in life histories. – Ecol. Lett. 12: 93–106.

van den Driessche, P. and Watmough, J. 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. – Math. Biosci. 180: 29–48.

Vaupel, J. W. 1986. How change in age-specific mortality affects life expectancy. – Popul. Stud. 40: 147–157.

Vaupel, J. W. and Yashin, A. I. 1985. Heterogeneitys ruses: some surprising effects of selection on population dynamics. – Am. Stat. 39: 176–185.

Vaupel, J. W. and Carey, J. R. 1993. Compositional interpretations of medfly mortality. – Science 260: 1666–1667.

Vaupel, J. W. and Canudas Romo, V. 2003. Decomposing change in life expectancy: a bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. – Demography 40: 201–216.

Vaupel, J. W. et al. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. – Demography 16: 439–454.

Verdy, A. and Caswell, H. 2008. Sensitivity analysis of reactive ecological dynamics. – Bull. Math. Biol. 70: 1634–1659.

Wilmoth, J. R. and Horiuchi, S. 1999. Rectangularization revisited: variability of age at death within human populations. – Demography 36: 475–495.

Supplementary material (available online as Appendix O17620 at <www.oikos.ekol.lu.se/appendix>). Appendix A. Appendix B.