

Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach

Hal Caswell

Received: 20 February 2011 / Accepted: 16 May 2011 / Published online: 14 June 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Matrix population models in which individuals are classified by both age and stage can be constructed using the vec-permutation matrix. The resulting age-stage models can be used to derive the age-specific consequences of a stage-specific life history or to describe populations in which the vital rates respond to both age and stage. I derive a general formula for the sensitivity of any output (scalar, vector, or matrix-valued) of the model, to any vector of parameters, using matrix calculus. The matrices describing age-stage dynamics are almost always reducible; I present results giving conditions under which population growth is ergodic from any initial condition. As an example, I analyze a published stage-specific model of Scotch broom (*Cytisus scoparius*), an invasive perennial shrub. Sensitivity analysis of the population growth rate finds that the selection gradients on adult survival do not always decrease with age but may increase over a range of ages. This may have implications for the evolution of senescence in stage-classified populations. I also derive and analyze the joint distribution of age and stage at death and present a sensitivity analysis of

this distribution and of the marginal distribution of age at death.

Keywords Age-stage classified · Reducible matrices · Matrix calculus · Senescence · Age at death · Ergodicity

Introduction

The first step in developing any kind of structured population model is choosing one or more variables in terms of which to describe the population structure. The job of these *i*-state variables is to encapsulate all the information about the past experience of an individual that is relevant to its future behavior (Metz and Diekmann 1986, Caswell 2001, Chapter 3). Classical demography (for both humans and for nonhuman animals and plants) uses age as a *i*-state, but other, more biologically relevant criteria (e.g., size, developmental stage, parity, physiological condition, etc.) are now widely used in ecology, with age-classified models viewed as a special case.

However, it has long been recognized that cases exist where it is important to classify individuals by *both* age and stage.

1. Even in a stage-classified model, age still exists; every individual becomes older, by one unit of age, with the passage of each unit of time. There is increasing interest in extracting the age-specific demographic consequences of stage-classified models (e.g., Feichtinger 1971; Caswell 2001, 2006, 2009b; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008). Models that include both age

Electronic supplementary material The online version of this article (doi:10.1007/s12080-011-0132-2) contains supplementary material, which is available to authorized users.

H. Caswell (✉)
Biology Department MS-34, Woods Hole Oceanographic
Institution, Woods Hole, MA 02543, USA
e-mail: hcaswell@whoi.edu

H. Caswell
Max Planck Institute for Demographic Research,
Rostock, Germany

and stage provide information on those consequences that goes beyond current methods based on the fundamental matrix of the stage-classified model.

2. If the vital rates depend on both age and stage, a model that includes both is necessary to reveal the joint action of age- and stage-specific processes (e.g., Goodman 1969; Logofet 2002). Such models, of course, require information on the joint age dependence and stage dependence of the vital rates and thus are challenging to construct (see Law 1983 and van Groenendael and Slim 1988 for examples). A special case that has been extensively explored is the multiregional case, in which the stage variable describes spatial location (e.g., Rogers 1966, 1995; Lebreton 1996).

Here, I present a model framework in which individuals are classified by age and stage, using the vec-permutation matrix approach (so-called for the role that the vec-permutation matrix plays in rearranging age and stage categories in the population vector). This formalism was introduced by Hunter and Caswell (2005) for populations classified by stage and location (see applications by Ozgul et al. 2009; Goldberg et al. 2010; Strasser et al. 2010) and has been applied to time-varying models classified by stage and environmental state (Caswell 2006, 2009b, 2011a) and to stage-structured epidemic models (Klepac and Caswell 2010).

Matrix models can describe both population dynamics and cohort dynamics. Population dynamics (population growth, age and stage structure, reproductive value) depend on both the transitions of extant individuals and the production of new individuals by reproduction. In contrast, cohort dynamics (survivorship, life expectancy, age at death, generation time) depend only on the fates of already existing individuals. The framework I introduce here permits both kinds of analysis.

Perturbation analysis calculates the response of model outputs to changes in the parameters. Demographic studies are almost always concerned with change: over time, in response to external factors (e.g., experimental treatments, environmental influences, policy interventions, or historical events), or as differences among populations (e.g., among regions, related species, or among populations differentiated by social factors). Evolutionary demography focuses differences caused by genetic variation; the fate of a new phenotype depends on the changes it produces in fitness. Hence, perturbation analysis is an important tool in ecology, management, human demography, and evolutionary biology (Caswell 2001). I will develop a general formula for the sensitivity of any dependent

variable on changes in any parameter[s] influencing either age- or stage-dependent dynamics, using matrix calculus methods (Caswell 2007, 2008, 2009a, b).

The perturbation analysis of fitness provides selection gradients, which are particularly relevant to the evolution of senescence (increases in mortality rates with age; e.g., Medawar 1952, Hamilton 1966, Charlesworth 1994, Rose 1991, Baudisch 2005, Vaupel 2010). Hamilton (1966) showed that the magnitude of the selection gradient on mortality is nonincreasing with age and is strictly decreasing with age after maturity. This means that relatively large increases in mortality at late ages can be compensated for by smaller—often much smaller—reductions in mortality at early ages. Any trait that creates such changes will be favored, and the accumulation of such traits in the population leads to senescence, but see Baudisch (2005, 2008) for a discussion of the care required in the definition of “such traits.”

As is well-known, however, Hamilton’s conclusions are specific to age-classified life cycles. Selection gradients may behave quite differently in stage-classified models, leading some to suggest that stage-classified species, particularly those where demography is strongly size-dependent, might exhibit “negative senescence” (Caswell 1982; Vaupel et al. 2004). However, in order to evaluate this argument, it is essential to see how selection gradients change with both age *and* stage. The model framework developed to be presented here makes this possible, and an example is presented in section “Population growth rate and selection gradients”). These results greatly expand the range of ecological data that can be applied to questions about the evolution of senescence.

Model construction

The construction and analysis of these models require a number of different matrices and operators; some of the notation is collected in Table 1.

Individuals are classified into stages $1, \dots, s$ and age classes $1, \dots, \omega$. The model treats the processes of moving among stages and moving among age classes as periodic, or alternating. First, stage-specific demography operates to move individuals among stages and to produce new offspring. Then aging acts to move individuals to the next older age, and the process repeats.

Define a stage-classified projection matrix \mathbf{A}_i , of dimension $s \times s$, for each age class, $i = 1, \dots, \omega$. Decompose \mathbf{A}_i into

$$\mathbf{A}_i = \mathbf{U}_i + \mathbf{F}_i \quad (1)$$

Table 1 Mathematical notation used in this study

Quantity	Description	Dimension
$\mathbf{A}_i, \mathbf{F}_i, \mathbf{U}_i$	Stage-classified projection, fertility, and transition matrices for age class i	$s \times s$
$\mathbf{D}_U, \mathbf{D}_F$	Age transition matrices for individuals already present in the population and for new individuals produced by reproduction	$\omega \times \omega$
$\mathbb{A}, \mathbb{F}, \mathbb{U}, \mathbb{D}$	Block-diagonal matrices.	$s\omega \times s\omega$
$\tilde{\mathbf{A}}, \tilde{\mathbf{U}}$, etc.	Age-stage matrices constructed from block-diagonal matrices using the vec-permutation matrix	$s\omega \times s\omega$
$\mathbf{K}_{s,\omega}, \mathbf{K}$	Vec-permutation matrix	$s\omega \times s\omega$
\mathbf{I}_s	Identity matrix	$s \times s$
$\mathbf{1}_s$	Vector of ones	$s \times 1$
\mathbf{e}_i	The i th unit vector, with a 1 in the i th entry and zeros elsewhere	Various
\mathbf{E}_{ij}	A matrix with a 1 in the (i, j) position, and zeros elsewhere	Various
\otimes	Kronecker product	
\circ	Hadamard, or element-by-element, product	
vec \mathbf{X}	The vec operator, which stacks the columns of a $m \times n$ matrix \mathbf{X} into a $mn \times 1$ vector	
$\mathcal{D}(\mathbf{x})$	A diagonal matrix with \mathbf{x} on the diagonal and zeros elsewhere	

Dimensions are shown, where relevant, for matrices and vectors; s denotes the number of stages and ω the number of age classes

where \mathbf{U}_i contains the transition probabilities of extant individuals and \mathbf{F}_i describes the generation of new individuals by reproduction.

Aging is described by two matrices, each of dimension $\omega \times \omega$ (shown here for 3×3 , but easily generalized),

$$\mathbf{D}_U = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{dimension } \omega \times \omega \quad (2)$$

$$\mathbf{D}_F = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \omega \times \omega \quad (3)$$

The matrix \mathbf{D}_U applies to extant individuals; such an individual advances to the next age class. I have set the (ω, ω) entry of \mathbf{D}_U to 1, so that the last age class contains individuals of age ω and older. If this entry were set to 0, all individuals in the last age class would die. The matrix \mathbf{D}_F applies to individuals newly created by reproduction; such newborn individuals are placed in the first age class, regardless of the age of their parents.

Using the matrices $\mathbf{A}_i, \mathbf{U}_i, \mathbf{F}_i, \mathbf{D}_U$, and \mathbf{D}_F , construct block-diagonal matrices, each of dimension $s\omega \times s\omega$. For example,

$$\mathbb{A} = \begin{pmatrix} \mathbf{A}_1 & & \\ & \ddots & \\ & & \mathbf{A}_\omega \end{pmatrix}. \quad (4)$$

The complete set of block-diagonal matrices is given in Online Resource A. These block-diagonal matrices can be written

$$\mathbb{A} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{A}_i) \quad (5)$$

$$\mathbb{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{U}_i) \quad (6)$$

$$\mathbb{F} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{F}_i) \quad (7)$$

$$\mathbb{D}_U = \mathbf{I}_s \otimes \mathbf{D}_U \quad (8)$$

$$\mathbb{D}_F = \mathbf{I}_s \otimes \mathbf{D}_F \quad (9)$$

where \mathbf{E}_{ii} is of dimension $\omega \times \omega$.

If the demography is truly stage-dependent, so that $\mathbf{A}_i = \mathbf{A}$, for $i = 1, \dots, \omega$, then the block-diagonal matrices \mathbb{A}, \mathbb{F} , and \mathbb{U} reduce to, e.g.,

$$\mathbb{A} = \mathbf{I}_\omega \otimes \mathbf{A} \quad (10)$$

with corresponding expressions for \mathbb{F} and \mathbb{U} .

The state of the population at time t could be described by a two-dimensional array

$$\mathcal{N}(t) = \begin{pmatrix} n_{11} & \cdots & n_{1\omega} \\ \vdots & & \vdots \\ n_{s1} & \cdots & n_{s\omega} \end{pmatrix} (t) \quad s \times \omega \quad (11)$$

where rows correspond to stages and columns to age classes. However, such a two-dimensional array cannot be projected directly; instead, it is transformed to a vector,

$$\mathbf{n}(t) = \text{vec } \mathcal{N}(t) = \begin{pmatrix} n_{11} \\ \vdots \\ n_{s1} \\ \vdots \\ n_{s\omega} \end{pmatrix} (t) \quad s\omega \times 1 \quad (12)$$

using the *vec* operator, which stacks the columns of the matrix one above the next. The vector $\mathbf{n}(t)$ created in this way contains the stages arranged within age classes. An alternative configuration, with ages arranged within stages, is obtained by applying the *vec* operator to \mathcal{N}^T :

$$\text{vec } \mathcal{N}^T(t) = \begin{pmatrix} n_{11} \\ \vdots \\ n_{1\omega} \\ \vdots \\ n_{s1} \\ \vdots \\ n_{s\omega} \end{pmatrix} (t) \quad s\omega \times 1. \quad (13)$$

The two vectors $\text{vec } \mathcal{N}$ and $\text{vec } \mathcal{N}^T$ are related by the *vec*-permutation matrix \mathbf{K} (Henderson and Searle 1981), also known as the commutation matrix (Magnus and Neudecker 1979),

$$\text{vec } \mathcal{N}^T = \mathbf{K}_{s,\omega} \text{vec } \mathcal{N} \quad (14)$$

where

$$\mathbf{K}_{s,\omega} = \sum_{i=1}^s \sum_{j=1}^{\omega} \mathbf{E}_{ij} \otimes \mathbf{E}_{ij}^T \quad (15)$$

where \mathbf{E}_{ij} is of dimension $s \times \omega$. Where no confusion seems likely to arise, I will suppress the subscripts and write $\mathbf{K}_{s,\omega}$ as \mathbf{K} . As with any permutation matrix, $\mathbf{K}^T = \mathbf{K}^{-1}$.

The goal of the model is to project the age-stage vector $\mathbf{n} = \text{vec } \mathcal{N}$ from t to $t + 1$. The complete projection is given by

$$\mathbf{n}(t + 1) = (\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U} + \mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F}) \mathbf{n}(t). \quad (16)$$

This deserves some explanation. Consider the first term on the right-hand side, $\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U}$. Reading from right to left, it first operates on the vector $\mathbf{n}(t)$ with the block diagonal matrix \mathbf{U} , which moves surviving extant individuals among stages without changing their age. Then the resulting vector is rearranged by the *vec*-permutation matrix \mathbf{K} to group individuals by age classes within each stage. The block diagonal matrix \mathbb{D}_U then moves each surviving individual to the next older age class. Finally, \mathbf{K}^T rearranges the vector back to the stage-within-age arrangement of $\mathbf{n}(t)$.

The second term in Eq. (16), $\mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F}$, carries out a similar sequence of transformations for the generation of new individuals. First, newborn individuals are produced according to the block-diagonal fertility matrix \mathbf{F} . The resulting vector is rearranged by the *vec*-permutation matrix, and then the matrix \mathbb{D}_F places all the newborn individuals into the first age class. Finally, \mathbf{K}^T rearranges the vector to the stage-within-age arrangement.

I will write the age-stage projection matrix in Eq. (16) as

$$\tilde{\mathbf{A}} = (\mathbf{K}^T \mathbb{D}_U \mathbf{K} \mathbf{U} + \mathbf{K}^T \mathbb{D}_F \mathbf{K} \mathbf{F}) \quad (17)$$

$$= (\tilde{\mathbf{U}} + \tilde{\mathbf{F}}). \quad (18)$$

The matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{U}}$, and $\tilde{\mathbf{F}}$ that operate on the age-stage vector \mathbf{n} are denoted with a tilde ($\tilde{\mathbf{A}}$, $\tilde{\mathbf{U}}$, $\tilde{\mathbf{F}}$); these matrices define the age-stage classified model and can be subjected to all the usual demographic analyses.

Perturbation analysis

Age-stage models pose particular challenges for perturbation analysis, because interest naturally focuses on changes in the matrices \mathbf{F}_i and \mathbf{U}_i ($i = 1, \dots, \omega$), which are deeply embedded within $\tilde{\mathbf{F}}$, $\tilde{\mathbf{U}}$, and $\tilde{\mathbf{A}}$. However, the computations are possible using matrix calculus, a formalism that permits consistent calculation of the derivatives of scalar-, vector- or matrix-valued functions of scalar-, vector-, or matrix-valued arguments (Magnus and Neudecker 1985, 1988). See Abadir and Magnus (2005) for an introduction and Caswell (2006, 2007, 2008, 2009a, b, 2010, 2011a, b), Jenouvrier et al. (2010), Klepac and Caswell (2010), and Strasser et al. (2010) for ecological applications. A brief outline of the methods is provided in “[Online Resource B](#)”.

I will present the perturbation analysis in terms of a generic dependent variable ξ , which is a scalar- or vector-valued function of $\tilde{\mathbf{A}}$. In the examples to follow, ξ will be either the population growth rate λ or the joint

distribution of age and stage at death in a cohort, but it could be any variable calculated from $\tilde{\mathbf{A}}$. Let θ be a vector of parameters; these could be entries of the matrices, or lower-level parameters determining those entries. The goal of perturbation analysis is to obtain the derivative of ξ with respect to θ . This derivative is a matrix whose (i, j) entry is the derivative of ξ_i with respect to θ_j :

$$\frac{d\xi}{d\theta^T} = \left(\frac{d\xi_i}{d\theta_j} \right). \tag{19}$$

By the chain rule,

$$\frac{d\xi}{d\theta^T} = \frac{d\xi}{d\text{vec}^T \tilde{\mathbf{A}}} \frac{d\text{vec} \tilde{\mathbf{A}}}{d\theta^T}. \tag{20}$$

The first term in Eq. (20) is the derivative of ξ with respect to the matrix $\tilde{\mathbf{A}}$. If, for example, ξ was the dominant eigenvalue λ , then this term would be the matrix calculus version of the well-known eigenvalue sensitivity equation.

The second term in Eq. (20) requires differentiating $\tilde{\mathbf{A}}$ with respect to the parameters that determine it. From Eq. (17), write

$$\tilde{\mathbf{A}} = \mathbf{Q}_U \mathbf{U} + \mathbf{Q}_F \mathbb{F} \tag{21}$$

where $\mathbf{Q}_U = \mathbf{K}^T \mathbb{D}_U \mathbf{K}$ and $\mathbf{Q}_F = \mathbf{K}^T \mathbb{D}_F \mathbf{K}$ are the (constant) matrix products appearing in the definition of $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{F}}$ in Eq. (17).

Differentiating $\tilde{\mathbf{A}}$ in Eq. (21) gives

$$d\text{vec} \tilde{\mathbf{A}} = (\mathbf{I}_{sw} \otimes \mathbf{Q}_U) d\text{vec} \mathbf{U} + (\mathbf{I}_{sw} \otimes \mathbf{Q}_F) d\text{vec} \mathbb{F}. \tag{22}$$

This requires the differentials of \mathbf{U} and \mathbb{F} . Differentiating \mathbf{U} in Eq. (6) gives

$$d\mathbf{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes d\mathbf{U}_i). \tag{23}$$

Applying the vec operator to $d\mathbf{U}$ gives

$$d\text{vec} \mathbf{U} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec} \mathbf{U}_i \tag{24}$$

using the results of Magnus and Neudecker (1985, Theorem 11; see Appendix B of Klepac and Caswell 2010) on the derivative of the Kronecker product. Differentiation of \mathbb{F} proceeds in the same fashion, yielding

$$d\text{vec} \mathbb{F} = \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec} \mathbf{F}_i. \tag{25}$$

In the special case where \mathbf{U} and \mathbb{F} are constructed from single stage-classified matrices \mathbf{U} and \mathbf{F} , as in Eqs. (10), (24) and (25) simplify to

$$d\text{vec} \mathbf{U} = (\mathbf{I}_{\omega} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec} \mathbf{U} \tag{26}$$

$$d\text{vec} \mathbb{F} = (\mathbf{I}_{\omega} \otimes \mathbf{K} \otimes \mathbf{I}_s) (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) d\text{vec} \mathbf{F}. \tag{27}$$

Substituting Eqs. (24) and (25) into Eq. (22) and then substituting Eq. (22) into Eq. (20) yields the general result for the derivative

$$\begin{aligned} \frac{d\xi}{d\theta^T} = \frac{d\xi}{d\text{vec}^T \tilde{\mathbf{A}}} & \left[(\mathbf{I}_{sw} \otimes \mathbf{Q}_U) \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) \right. \\ & \times (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) \frac{d\text{vec} \mathbf{U}_i}{d\theta^T} + (\mathbf{I}_{sw} \otimes \mathbf{Q}_F) \\ & \times \sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) \\ & \left. \times (\text{vec} \mathbf{I}_{\omega} \otimes \mathbf{I}_{s^2}) \frac{d\text{vec} \mathbf{F}_i}{d\theta^T} \right]. \tag{28} \end{aligned}$$

Notice that Eq. (28) requires only three pieces of demographic information: the derivatives of \mathbf{U}_i and \mathbf{F}_i with respect to the parameters (whatever those may be in the case at hand) and the sensitivity of the dependent variable ξ (whatever that may be) to the elements of the matrix $\tilde{\mathbf{A}}$ from which it is calculated. All the other pieces of Eq. (28) are constants.¹

Case studies

Here I present two examples of the use of the age-stage model to extract age-classified information from a stage-classified model. The first example is a sensitivity analysis of the population growth rate λ , obtaining the sensitivity of λ to both age- and stage-specific survival, and permitting examination of how selection pressures on senescence-inducing traits would vary from stage to stage. The second example is an analysis of the joint distribution of age and stage at death. This has, to my knowledge, never been calculated before.

These examples are based on a stage-classified model (Parker 2000) for Scotch broom (*Cytisus scoparius*). Scotch broom is a large (up to 4 m tall) leguminous shrub, introduced into North America from Europe in the late 19th century. It is an invasive plant, considered a pest in the northwestern parts of North America.

¹Some of these constant matrices may be large, depending on s and ω , but they are very sparse; the sparse matrix technology available in MATLAB can be extremely useful in implementation.

Stage-classified demographic models have been used to evaluate potential management policies for the plant (Parker 2000) and to investigate its potential for spatial spread (Neubert and Parker 2004).

The model contains seven stages (stage 1 = seeds, 2 = seedlings, 3 = juveniles, 4 = small adults, 5 = medium adults, 6 = large adults, 7 = extra-large adults), and parameters were estimated at a number of locations in Washington State. As is typical with many perennial plant species, survival is low for seeds and seedlings but increases dramatically in larger stages. Parker's (2000) study presented estimated projection matrices for plants at the edge, at intermediate locations, and at the center of an invading stand. Plants near the center experience more crowding, with resulting reduced rates of survival, growth, and fertility.

Population growth rate and selection gradients

The population growth rate λ , the stable age or stage distribution \mathbf{w} , and age or stage-specific reproductive value vector \mathbf{v} are given by the dominant eigenvalue and corresponding right and left eigenvectors of the population projection matrix, respectively. In evolutionary demography, λ measures the fitness of a phenotype, in that it gives the eventual rate at which descendants of an individual with that phenotype will increase. The selection gradient on a vector of traits $\boldsymbol{\theta}$ is given by

$$\frac{d\lambda}{d\boldsymbol{\theta}^\top} \quad (29)$$

These gradients play a fundamental role in evolutionary biodemography, whether evolution is conceived of in terms of population genetics, quantitative genetics, adaptive dynamics, or mutation accumulation (e.g., Metz et al. 1992; Dercole and Rinaldi 2008; Caswell 2001; Rice 2004; Barfield et al. 2011). If the gradient is positive, selection favors an increase in the trait, and vice-versa.

In this application, ξ in Eq. (20) is the dominant eigenvalue λ . Let \mathbf{w} and \mathbf{v} be the right and left eigenvectors corresponding to λ , scaled so that $\mathbf{v}^\top \mathbf{w} = 1$. Then, in Eq. (28),

$$\frac{d\lambda}{d\text{vec}^\top \tilde{\mathbf{A}}} = \mathbf{w}^\top \otimes \mathbf{v}^\top. \quad (30)$$

(Caswell 2010).

In this model, the vital rates are functions only of stage; the phenotype is blind to the age of the individual. However, the terms in the summations in Eq. (28)

give the selection gradients on traits operating at a different age. That is,

$$\begin{aligned} \frac{d\lambda}{d\boldsymbol{\theta}^\top} \Big|_{\text{age}=i} &= \frac{d\lambda}{d\text{vec}^\top \tilde{\mathbf{A}}} \left[(\mathbf{I}_{sw} \otimes \mathbf{Q}_U) (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) \right. \\ &\quad \times (\text{vec } \mathbf{I}_w \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{U}_i}{d\boldsymbol{\theta}^\top} \\ &\quad + (\mathbf{I}_{sw} \otimes \mathbf{Q}_F) (\mathbf{E}_{ii} \otimes \mathbf{K} \otimes \mathbf{I}_s) \\ &\quad \left. \times (\text{vec } \mathbf{I}_w \otimes \mathbf{I}_{s^2}) \frac{d\text{vec } \mathbf{F}_i}{d\boldsymbol{\theta}^\top} \right]. \quad (31) \end{aligned}$$

Thus, these terms reveal the selection patterns that would operate on a mutation that was able to detect the age of an individual within a given stage, or that affected age differentially depending on the stage of the individual.

To examine the selection gradients on survival, it is necessary to separate survival from inter-stage transitions in \mathbf{U} . Let $\boldsymbol{\sigma}$ be the vector of stage-specific survival probabilities. The matrix \mathbf{U} can be written as the product of a diagonal matrix containing the survival probabilities and a matrix $\boldsymbol{\Gamma}$ is a matrix of transition probabilities, conditional on survival;

$$\mathbf{U} = \boldsymbol{\Gamma} \mathcal{D}(\boldsymbol{\sigma}). \quad (32)$$

I assume that \mathbf{F} is independent² of $\boldsymbol{\sigma}$. Thus

$$d\mathbf{U} = \boldsymbol{\Gamma} \mathcal{D}(d\boldsymbol{\sigma}). \quad (33)$$

Applying the vec operator gives

$$\begin{aligned} d\text{vec } \mathbf{U} &= (\mathbf{I}_s \otimes \boldsymbol{\Gamma}) \text{vec } \mathcal{D}(d\boldsymbol{\sigma}) \\ &= (\mathbf{I}_s \otimes \boldsymbol{\Gamma}) \mathcal{D}(\text{vec } \mathbf{I}_s) (\mathbf{I}_s \otimes \mathbf{I}_s) d\boldsymbol{\sigma} \end{aligned} \quad (34)$$

which implies that

$$\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\sigma}^\top} = (\mathbf{I}_s \otimes \boldsymbol{\Gamma}) \mathcal{D}(\text{vec } \mathbf{I}_s) (\mathbf{I}_s \otimes \mathbf{I}_s). \quad (35)$$

Setting $\boldsymbol{\theta} = \boldsymbol{\sigma}$ and substituting Eqs. (35) and (30) into Eq. (20) gives the selection gradient on $\boldsymbol{\sigma}$. Substituting Eq. (35) and Eq. (30) into Eq. (31), with $d\text{vec } \mathbf{F}/d\boldsymbol{\theta}^\top = \mathbf{0}$, gives the selection gradient on $\boldsymbol{\sigma}$ as a function of age and stage.

²By assuming that \mathbf{F} does not depend on $\boldsymbol{\sigma}$, I am in effect choosing a prebreeding census and excluding neonatal mortality from $\boldsymbol{\sigma}$. It is not difficult to include \mathbf{F} in the analysis; the implications of doing so will be explored elsewhere.

Table 2 The projection matrix **A** for Scotch broom for the Discovery Park population, 1993–1994, edge conditions; taken from the Appendix of Parker (2000)

$$\mathbf{A} = \begin{pmatrix} 0.740 & 0 & 3.400 & 47.1 & 108.700 & 1120.0 & 3339.0 \\ 0.001 & 0.310 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.350 & 0.310 & 0 & 0 & 0 & 0 \\ 0 & 0.038 & 0.290 & 0.024 & 0 & 0 & 0 \\ 0 & 0 & 0.069 & 0.390 & 0.320 & 0 & 0.091 \\ 0 & 0 & 0 & 0.440 & 0.440 & 0.530 & 0.091 \\ 0 & 0 & 0 & 0 & 0.029 & 0.400 & 0.730 \end{pmatrix} \quad (36)$$

Results

The projection matrix **A** for Scotch broom is given in Table 2. The matrix **U** is obtained from **A** by setting all elements in the first row except for a_{11} to zero. The matrix **F** is a 7×7 matrix with the elements of row 1, columns 2–7 of **A** in the corresponding positions, and zeros elsewhere. The maximum age was set to $\omega = 30$. The aging matrices **D_U** and **D_F** are given by Eqs. (2) and (3) with $\omega = 30$. Because the vital rates do not depend on age, the dominant eigenvalues of **A** and **Ã** should be identical, and they are; $\lambda = 1.268$.

The selection gradients on stage-specific survival (i.e., sensitivities of λ to σ) are shown in Fig. 1. There is a steady decline with increasing stage, from seeds to medium-sized adults, but then an increase for large and extra-large adults. A quite different pattern emerges when the selection gradients are calculated as functions of both age and stage, using Eq. (31). These results are

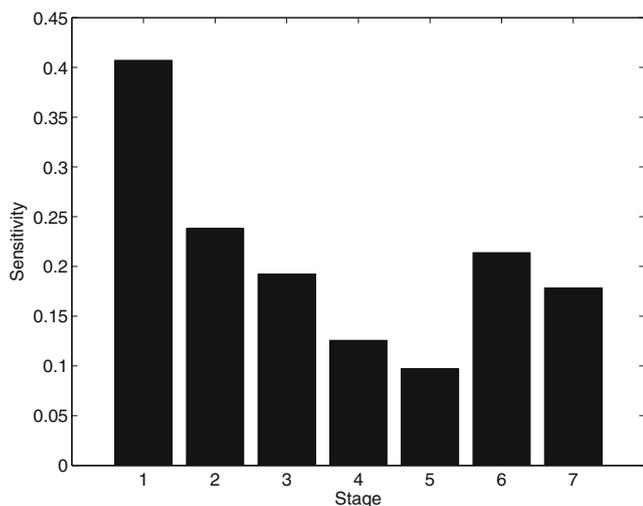


Fig. 1 Sensitivity of population growth rate λ to stage-specific survival probabilities. Calculated for the stage-classified model of Scotch broom (*C. scoparius* using data from Parker (2000). Stages: 1 = seeds, 2 = seedlings, 3 = juveniles, 4 = small adults, 5 = medium adults, 6 = large adults, 7 = extra-large adults

shown in Fig. 2. The age-specific selection gradients on survival in stages 1–3 are strictly decreasing with age. But the age-specific selection gradients on survival in the adult stages 4–7 increase with age, level off, and then decline. The increase is longer and more pronounced in the larger adult stages.

To my knowledge, this pattern has never been documented before. Carrying out the same analysis for a set of eight different populations of Scotch broom (Parker 2000), in different locations and different years, shows that they all exhibit this pattern to one degree or another (see Online Resource C). The consequences of these selection gradients for the evolution of senescence are still unknown. However, any conclusions that follow from the general decline in selection gradients with age would not apply to traits that affect age-specific survival differentially depending on developmental stage. Traits that affect survival in adult stages should postpone senescence for at least some time.

The elasticities of λ to σ show a similar pattern (Online Resource C). These elasticities are the sensitivities of $\log \lambda$ to changes in age-stage-specific mortality (with opposite sign),

$$\frac{\sigma_i}{\lambda} \frac{d\lambda}{d\sigma_i} = - \frac{d \log \lambda}{d\mu_i} \quad (37)$$

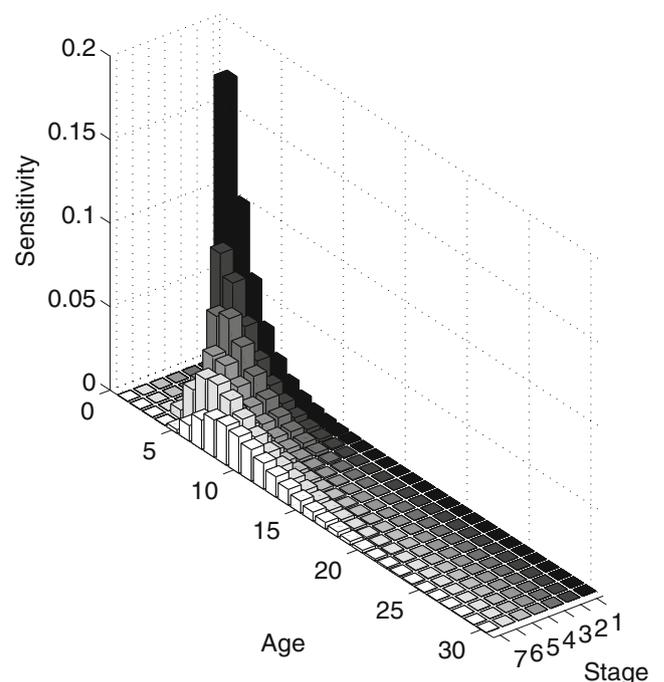


Fig. 2 Sensitivity of population growth rate λ to stage-specific survival as a function of age, for Scotch broom. Stages defined as in Fig. 1

where μ_i is the mortality rate of stage i . Thus, conclusions about senescence also hold for traits that cause perturbations to mortality.

Cohort dynamics: the distribution of age and stage at death

The pattern of longevity within a population is captured by the probability distribution of the age at death, one of the standard results of age-classified life table analysis. The moments of the age at death and their sensitivity can also be calculated directly from stage-classified models using Markov chain methods (Feichtinger 1971; Caswell 2001, 2006, 2009b; Tuljapurkar and Horvitz 2006; Horvitz and Tuljapurkar 2008). My goal here, however, is to go beyond that, to the full joint distribution of stage and age at death, along with the marginal distributions of age at death and stage at death, implied by an age-stage classified model.

To do this, note that the cohort projection matrix $\tilde{\mathbf{U}}$ describes movement of individuals among transient states of an absorbing Markov chain, where the absorbing state is death, or death classified by stage or age at death. The transition matrix of the chain is

$$\tilde{\mathbf{P}} = \begin{pmatrix} \tilde{\mathbf{U}} & \mathbf{0} \\ \tilde{\mathbf{M}} & \mathbf{I} \end{pmatrix}. \tag{38}$$

The matrix $\tilde{\mathbf{P}}$ is column stochastic, written in column-to-row orientation.

Each row of $\tilde{\mathbf{M}}$ corresponds to an absorbing state, and \tilde{m}_{ij} is the probability of a transition from transient state j to absorbing state i . To compute the distribution of age and stage at death, we define the absorbing states to correspond to the age-stage combination at death. Thus, $\tilde{\mathbf{M}}$ contains probabilities of death on the diagonal and zeros elsewhere,

$$\tilde{\mathbf{M}} = \mathbf{I}_{s\omega} - \mathcal{D} \left(\mathbf{1}_{s\omega}^T \tilde{\mathbf{U}} \right). \tag{39}$$

The fundamental matrix of the Markov chain in Eq. (38) is

$$\tilde{\mathbf{N}} = \left(\mathbf{I} - \tilde{\mathbf{U}} \right)^{-1}. \tag{40}$$

The (i, j) element of $\tilde{\mathbf{N}}$ is the expected number of visits that an individual in state j will make to transient state i before death.

Consider the eventual fate of an individual starting in transient state j . Let

$$\tilde{b}_{ij} = P[\text{eventual absorption in } i \mid \text{starting in } j]. \tag{41}$$

The \tilde{b}_{ij} are the elements of the matrix $\tilde{\mathbf{B}}$ ($s\omega \times s\omega$) given by

$$\tilde{\mathbf{B}} = \tilde{\mathbf{M}}\tilde{\mathbf{N}}. \tag{42}$$

(Iosifescu 1980, Theorem 3.3; see Caswell 2001, Section 5.1). Since the absorbing states (the rows of $\tilde{\mathbf{M}}$) correspond to combinations of age and stage at death, column j of $\tilde{\mathbf{B}}$ gives the joint distribution of age and stage at death, starting from state j :

$$\mathbf{b}_{\cdot j} = \tilde{\mathbf{B}}\mathbf{e}_j. \tag{43}$$

The rows of $\tilde{\mathbf{B}}$ correspond to combinations of stage and age at death. Summing the rows over stages gives the marginal distribution of age at death, starting in column j of $\tilde{\mathbf{B}}$, as

$$\mathbf{g}_j = (\mathbf{I}_\omega \otimes \mathbf{1}_s^T) \mathbf{b}_{\cdot j}. \tag{44}$$

Similarly, summing over ages gives the marginal distribution of stage at death:

$$\mathbf{h}_j = (\mathbf{1}_s \otimes \mathbf{1}_\omega^T) \mathbf{K} \mathbf{b}_{\cdot j}. \tag{45}$$

Perturbation analysis

In the general sensitivity Eq. (20), the dependent variable $\xi = \mathbf{b}_{\cdot j}$. This depends only on $\tilde{\mathbf{U}}$, so the first term in Eq. (20) can be shown to be

$$\frac{d\xi}{\text{dvec } \tilde{\mathbf{A}}} = \frac{d\mathbf{b}_{\cdot j}}{\text{dvec } \tilde{\mathbf{U}}} \tag{46}$$

$$= - \left(\mathbf{e}_j^T \tilde{\mathbf{N}}^T \otimes \mathbf{I}_{s\omega} \right) \mathcal{D} \left(\text{vec } \mathbf{I}_{s\omega} \right) \left(\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega} \mathbf{1}_{s\omega}^T \right) + \left(\mathbf{e}_j^T \tilde{\mathbf{N}}^T \otimes \tilde{\mathbf{B}} \right). \tag{47}$$

The desired derivative $d\mathbf{b}_{\cdot j}/d\theta^T$ is obtained by substituting Eq. (47) for $d\xi/d\text{vec } \tilde{\mathbf{A}}$ in Eq. (28), setting $\text{dvec } \mathbf{F}_i/d\theta^T = 0$.

The sensitivities of the marginal distributions of age and stage at death are then given by

$$\frac{d\mathbf{g}_j}{d\theta^T} = (\mathbf{I}_\omega \otimes \mathbf{1}_s^T) \frac{d\mathbf{b}_{\cdot j}}{d\theta^T} \tag{48}$$

$$\frac{d\mathbf{h}_j}{d\theta^T} = (\mathbf{1}_s \otimes \mathbf{1}_\omega^T) \mathbf{K} \frac{d\mathbf{b}_{\cdot j}}{d\theta^T}. \tag{49}$$

The derivation of these perturbation results, a particularly nice application of matrix calculus, is given in Appendix A.

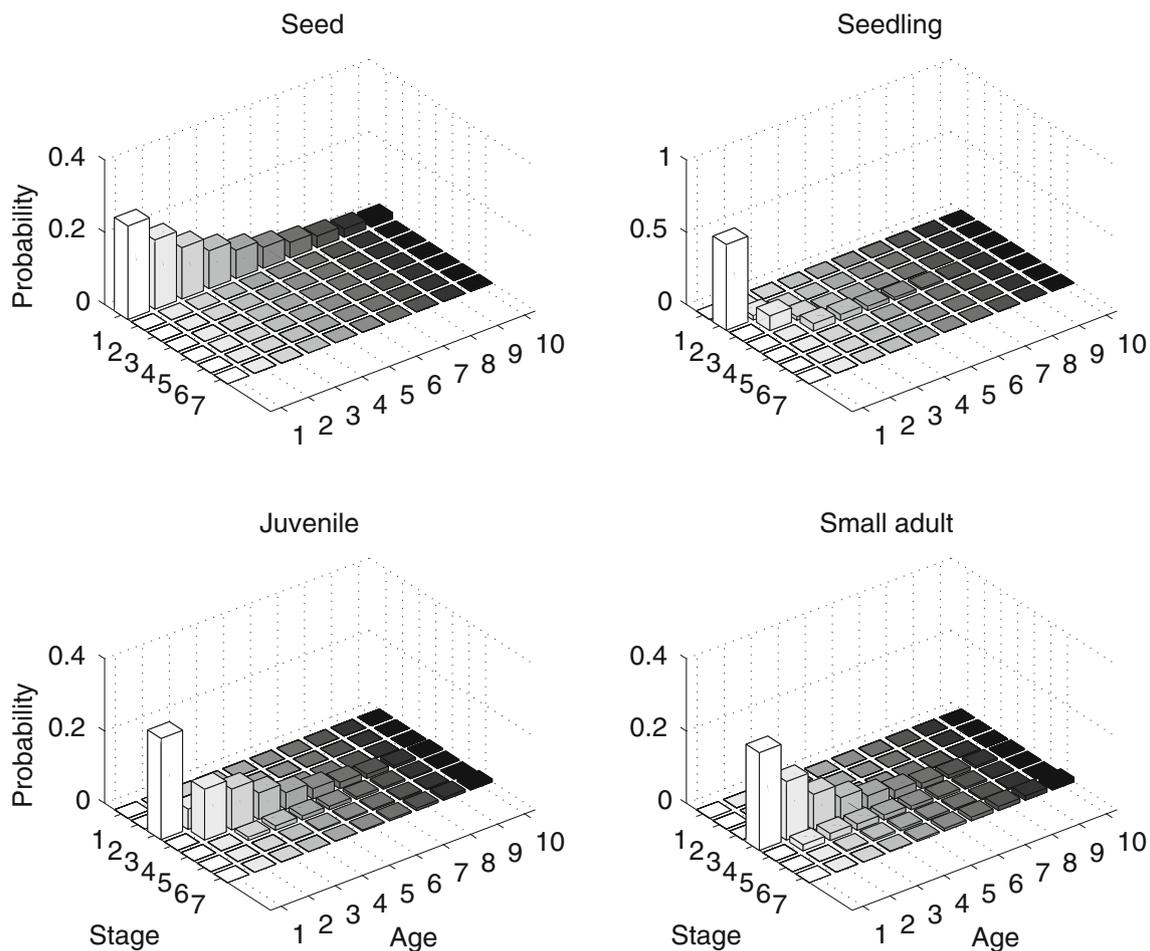


Fig. 3 The joint probability distribution of age and stage at death for an individual seed, seedling, juvenile, or small adult of Scotch broom. Stages as in Fig. 1

Results

Figure 3 shows the joint distribution of age and stage at death for a seed of age 1 (one definition of “newborn” in this life cycle), with $\omega = 40$. Almost all seeds will die as seeds, because the germination probability is low, $a_{21} = 0.001$; see Eq. (29). The fates of seedlings (another possible choice for newborn status) are more diverse, and those of juveniles and small adults even moreso; the distributions show what proportion will die as seedlings, juveniles, etc., and at what ages (Fig. 3).

The marginal distribution of age at death, for individuals in each initial stage, is given in Fig. 4. Not surprisingly, larger stages have an age distribution of death shifted to later ages, including some probability of survival to age class ω (≥ 40 years in this calculation).

The sensitivity of \mathbf{g}_2 (the marginal distribution of age at death for a seedling) is shown in Fig. 5. Changes in the survival of seeds (σ_1) have no effect on this distribution, because seedlings have already left the seed stage. Changes in σ_2 – σ_7 shift the distribution to

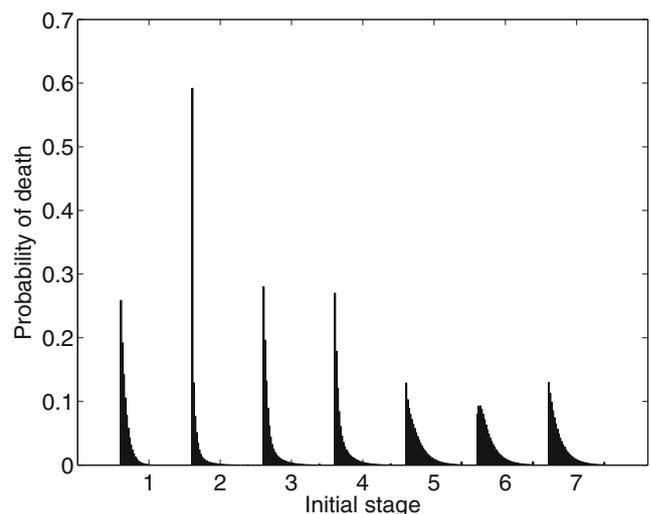


Fig. 4 The marginal distributions of age at death for individuals of Scotch broom in each stage. The maximum age in the model is $\omega = 40$. Stages as in Fig. 1

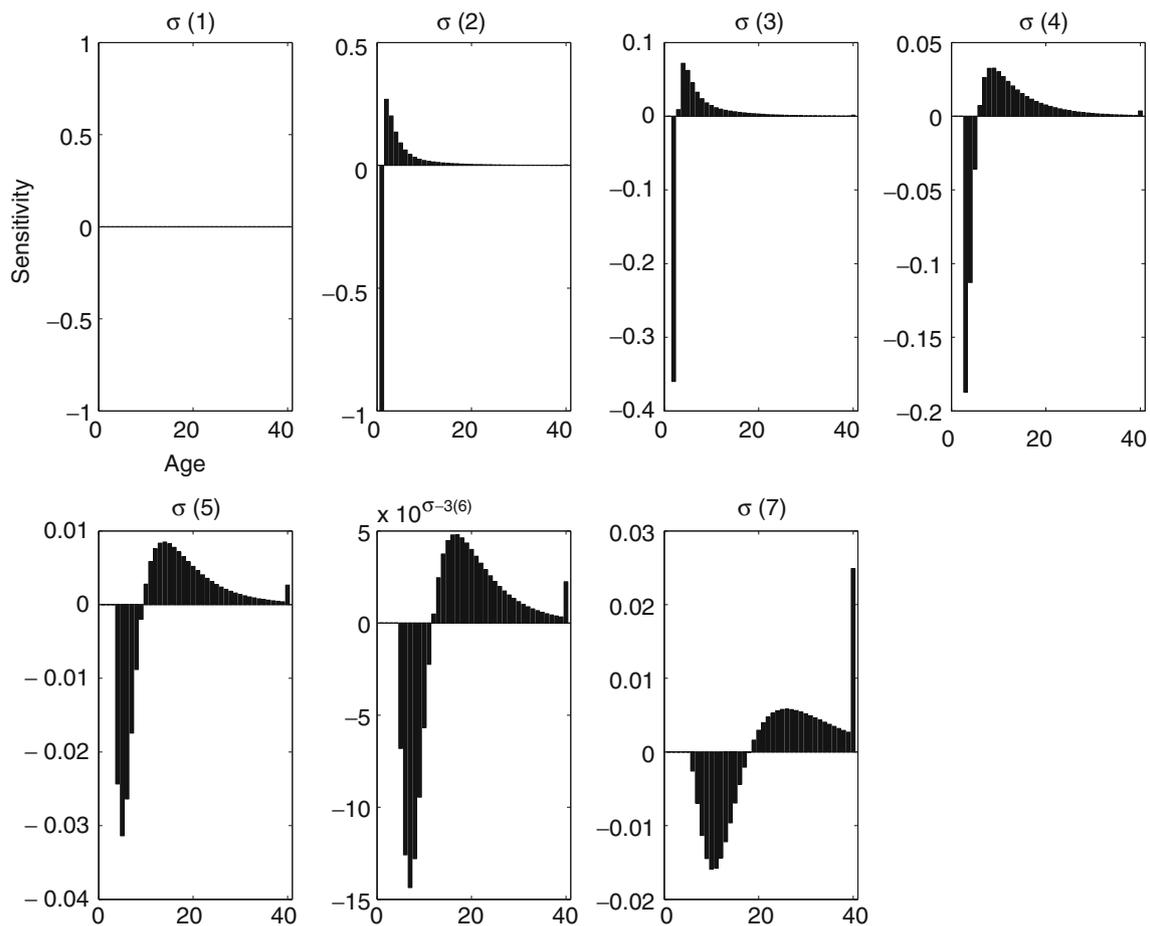


Fig. 5 Sensitivity of the marginal distribution of age at death, g_2 , to the survival probabilities of each stage, for an individual starting in stage 2 (seedlings). Stages as in Fig. 1

progressively older ages, by reducing the probability of death at young ages and increasing it at older ages.

Discussion

Models in which individuals are classified by both age and stage can extend demographic analyses in several directions. They permit biodemographic analyses of aging to take advantage of the many stage-classified demographic analyses accumulated by ecologists (cf. Caswell 2001). They also permit human demographers to take account of factors other than age in determining mortality, longevity, fertility, and population dynamics.

Age- and stage-specific demographic processes are regularly combined in demography using multistate life table (MSLT) methods (e.g., Rogers 1975; Willikens 2002; Hougaard 2000). These are usually focused on cohort dynamics and associated survival statistics (but see Rogers 1975, Chapter 5 for an explicit consideration of population projection). MSLT models are written

as continuous-parameter, discrete-state Markov chains, where the parameter represents age and the states represent stages. In order to solve the resulting equations, the dynamics must be approximated over a (usually short) finite age interval; this would correspond to the sequence of matrices \mathbf{A}_i in the model here. The age-stage model described by $\tilde{\mathbf{A}}$ is a way to solve the discretized equations in one step and makes possible a variety of analyses that are difficult or impossible in the usual MSLT formulation. Further investigation of the relation between the continuous MSLT methods and the age-stage models will be interesting.

Age-stage models require stage-classified projection matrices \mathbf{A}_i (or their components \mathbf{U}_i or \mathbf{F}_i) for ages $i = 1, \dots, \omega$. These can be obtained in several ways. The simplest is to use a single age-invariant matrix \mathbf{A} , as in the *Cytisus* example, and infer the age-specific properties it produces. Or, one could start with an age-schedule of mortality and modify it by stage-specific mortality differentials. For example, Honeycutt et al. (2003) developed a Markov chain model for diabetes

prevalence in the United States, in which the relative risk of death for individuals with diabetes was used to modify a general age-specific mortality rate.

Given sufficient longitudinal data on both age and stage, it is possible to estimate the stage-specific matrices \mathbf{A}_i as explicit functions of age; see Peeters et al. (2002), for an example of a study of human heart disease, and Lebreton et al. (2009) for a review of methods used in multistate capture–mark–recapture analysis in ecology. Needless to say, the data requirements for a full age-stage parameterization are challenging. I suspect that the development of estimation methods at intermediate levels of detail will be an important step; this study will help in that development.

Reducibility of age-stage matrices

The properties of $\tilde{\mathbf{A}}$ raise an important theoretical and technical issue regarding population growth, fitness, and selection gradients. The use of λ as a measure of fitness is usually justified by the strong ergodic theorem (Cohen 1979; Caswell 2001, Section 4.5.2), which guarantees the eventual convergence to the stable population structure and growth at a rate given by the dominant eigenvalue λ . A sufficient condition for this convergence is that the projection matrix be irreducible; i.e., that there exist a pathway connecting any two stages (Caswell 2001, Section 4.5).

General results about the irreducibility of block-structured matrices are difficult; see Csetenyi and Logofet Csetenyi and Logofet (1989), Logofet (1993, Chap. 3), and Logofet and Belova (2007) for some important graph-theoretical results. However, the age-stage matrices $\tilde{\mathbf{A}}$ developed here are unusual among population models in that they are (almost) always reducible, because they contain categories to which there are no possible pathways. This arises because age 1 individuals are produced only by reproduction. Hence, there can never be age 1 individuals in any stage that is not produced by reproduction. For example, Scotch broom reproduces only by seeds, so age 1 seeds appear in the model. However, the matrix $\tilde{\mathbf{A}}$ also contains entries corresponding to age 1 seedlings, age 1 juveniles, age 1 adults, etc. These do not exist, and because there are no pathways to these stages from any other stages, the matrix $\tilde{\mathbf{A}}$ is reducible.

The Perron–Frobenius theorem guarantees that a reducible nonnegative matrix will have a real, nonnegative, dominant eigenvalue that is at least as large as any of the others. However, the asymptotic population growth rate and structure may depend on initial conditions (Caswell 2001, Section 4.5.4). This means that one must ascertain that the eigenvalues and eigenvec-

tors under analysis correspond to initial conditions of interest.

Appendix B shows that a necessary and sufficient condition for population growth to be described by the dominant eigenvalue λ of $\tilde{\mathbf{A}}$, regardless of the (nonnegative and nonzero) initial population vector, is that the left eigenvector \mathbf{v} be strictly positive and that this corresponds to a particular form a block-triangular form of $\tilde{\mathbf{A}}$. This provides a simple check for the ergodicity of population growth and justifies the use of λ as a population growth rate and measure of fitness.

Primitivity may be difficult to evaluate for an age-stage matrix (but see Logofet 1993), but as with any projection matrix model, the long-term average growth rate of a primitive matrix is still given by the dominant real eigenvalue.

The matrix $\tilde{\mathbf{A}}$ for Scotch broom in Eq. (29) is reducible, as shown by calculating $(\mathbf{I}_{s\omega} + \tilde{\mathbf{A}})^{s\omega}$ and finding that this matrix contains zeros (Caswell 2001). However, the left eigenvector \mathbf{v} is strictly positive, so we know that the population eventually grows at the rate λ regardless of initial conditions.

A protocol for age-stage models

The approach outlined here gives a step-by-step procedure for constructing and analyzing age-stage matrix population models.

1. Choose a question. Are you interested in population dynamics (growth, structure, transients)? Or in cohort dynamics (survival, longevity)? Or in some combination of the two?
2. Obtain the stage-classified projection matrices \mathbf{A}_i for ages $i = 1, \dots, \omega$.
3. Decompose $\mathbf{A}_i = \mathbf{U}_i + \mathbf{F}_i$.
4. Construct the block-diagonal matrices \mathbb{A} , \mathbb{F} , \mathbb{U} , and \mathbb{D} , according to Eqs. (5)–(10).
5. Construct the age-stage matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{F}}$, $\tilde{\mathbf{U}}$ using Eq. (17) and, if appropriate for the question at hand, $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{P}}$ using Eqs. (38) and (39).
6. Analyze the model, e.g., by computing eigenvalues, eigenvectors, the fundamental matrix, etc., as appropriate. If necessary, check for reducibility and ergodicity using the methods in “[Reducibility of age-stage matrices](#)”.
7. For sensitivity analysis,
 - a. choose a set of dependent variables ξ and a vector of parameters θ ,
 - b. compute the sensitivity matrix $d\xi/d\text{vec}^T \tilde{\mathbf{A}}$,

c. compute the matrices:

$$\frac{d\text{vec } \mathbf{A}_i}{d\theta^\top}, \frac{d\text{vec } \mathbf{U}_i}{d\theta^\top}, \text{ and } \frac{d\text{vec } \mathbf{F}_i}{d\theta^\top} \quad (50)$$

d. compute $d\xi/d\theta^\top$ according to (20).

The explicit connection between matrix population models and absorbing Markov chain theory makes it possible to analyze both population dynamics and cohort dynamics in a unified framework (cf. Feichtinger 1971; Caswell 2001, 2006, 2009b). Cohort dynamics are, in essence, the demography of individuals. It may seem paradoxical to speak of the demography of individuals, but that is what it is, because the *statistical* properties of a cohort (e.g., average lifespan) are *probabilistic* properties of an individual (e.g., life expectancy). Demography in general, and matrix population models in particular, provides the link between the individual and the population.

Acknowledgements This research was supported by National Science Foundation Grant DEB-0816514 and by a Research Award from the Alexander von Humboldt Foundation. I am grateful to the Max Planck Institute for Demographic Research for hospitality while these ideas were developed, and I thank A. Baudisch, S. Jenouvrier, J. Kellner, D. Logofet, M. Rebke, E. Shyu, J. Vaupel, and several anonymous reviewers for comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix A: Derivation of the sensitivity of the distribution of deaths

This appendix contains the derivation, using matrix calculus, of the sensitivity of the distributions of age and stage at death in Eqs. 48 and 49. For a mathematical introduction to matrix calculus, see Abadir and Magnus (2005), for introductions in the context of demography, see Caswell (2007, 2008).

The columns of the matrix $\tilde{\mathbf{B}}$ are the joint distributions of age and stage at death. Consider column j of $\tilde{\mathbf{B}}$,

$$\mathbf{b}_{.j} = \tilde{\mathbf{B}}\mathbf{e}_j. \quad (51)$$

Differentiate both sides of Eq. (51),

$$d\mathbf{b}_{.j} = (d\tilde{\mathbf{B}})\mathbf{e}_j \quad (52)$$

and then apply the vec operator. Since $d\mathbf{b}_{.j}$ is a column vector, the vec operator has no effect on the left hand side, so

$$d\mathbf{b}_{.j} = (\mathbf{e}_j^\top \otimes \mathbf{I}_{s\omega})d\text{vec } \tilde{\mathbf{B}}. \quad (53)$$

However, from Eq. (42), $\tilde{\mathbf{B}} = \tilde{\mathbf{M}}\tilde{\mathbf{N}}$, so

$$d\tilde{\mathbf{B}} = (d\tilde{\mathbf{M}})\tilde{\mathbf{N}} + \tilde{\mathbf{M}}(d\tilde{\mathbf{N}}). \quad (54)$$

Apply the vec operator to obtain

$$d\text{vec } \tilde{\mathbf{B}} = (\tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega})d\text{vec } \tilde{\mathbf{M}} + (\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}})d\text{vec } \tilde{\mathbf{N}}. \quad (55)$$

Caswell (2006) showed that the differential of the fundamental matrix is

$$d\text{vec } \tilde{\mathbf{N}} = (\tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{N}})d\text{vec } \tilde{\mathbf{U}}. \quad (56)$$

The differential of $\tilde{\mathbf{M}}$ is obtained as follows. Note that

$$\tilde{\mathbf{M}} = \mathbf{I}_{s\omega} - \mathcal{D}(\mathbf{1}^\top \tilde{\mathbf{U}}) \quad (57)$$

$$= \mathbf{I}_{s\omega} - \mathbf{I}_{s\omega} \circ (\mathbf{1}\mathbf{1}^\top \tilde{\mathbf{U}}). \quad (58)$$

Differentiating gives

$$d\tilde{\mathbf{M}} = -\mathbf{I}_{s\omega} \circ [\mathbf{1}_{s\omega}\mathbf{1}_{s\omega}^\top (d\tilde{\mathbf{U}})]. \quad (59)$$

Applying the vec operator gives

$$d\text{vec } \tilde{\mathbf{M}} = -\mathcal{D}(\text{vec } \mathbf{I}_{s\omega}) (\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega}\mathbf{1}_{s\omega}^\top) d\text{vec } \tilde{\mathbf{U}}. \quad (60)$$

Substituting Eqs. (56) and (60) into Eq. (55) gives

$$d\text{vec } \tilde{\mathbf{B}} = \left[-(\tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega})\mathcal{D}(\text{vec } \mathbf{I}_{s\omega}) (\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega}\mathbf{1}_{s\omega}^\top) + (\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}})(\tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{N}}) \right] d\text{vec } \tilde{\mathbf{U}}. \quad (61)$$

Substituting this into Eq. (53) gives

$$d\mathbf{b}_{.j} = \left[-(\mathbf{e}_j^\top \otimes \mathbf{I}_{s\omega})(\tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega}) \times \mathcal{D}(\text{vec } \mathbf{I}_{s\omega}) (\mathbf{I}_{s\omega} \otimes \mathbf{1}_{s\omega}\mathbf{1}_{s\omega}^\top) + (\mathbf{e}_j^\top \otimes \mathbf{I}_{s\omega})(\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{M}})(\tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{N}}) \right] d\text{vec } \tilde{\mathbf{U}}. \quad (62)$$

Equation (62) can be simplified to obtain Eq. (47), using the fact that

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD}), \quad (63)$$

provided the products exist.

Appendix B: Population growth and reducible matrices

Some ergodic properties of population growth under the action of reducible matrices are described by Caswell (2001, Section 4.5.4). Here we can extend the analysis.

Let \mathbf{A} be a reducible nonnegative projection matrix. By permutation of its rows and columns (i.e., renumbering the stages in the life cycle), \mathbf{A} can be transformed to a block lower-triangular form. Here is an example:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_{11} & 0 & 0 & 0 \\ \mathbf{B}_{21} & \mathbf{B}_{22} & 0 & 0 \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & 0 \\ \mathbf{B}_{41} & \mathbf{B}_{42} & \mathbf{B}_{43} & \mathbf{B}_{44} \end{pmatrix}. \tag{64}$$

In this form, all the diagonal blocks \mathbf{B}_{ii} are either irreducible matrices or 1×1 (i.e., scalar) zero matrices. The block triangular form is unique, up to a renumbering of the blocks (Gantmacher 1959) and permutation of indices within blocks. It corresponds to a decomposition of the state space into a set of subspaces; let R_i be the subspace corresponding to the block \mathbf{B}_{ii} .

Some or all of the subdiagonal blocks in (64) may be zero. For reasons that will become apparent, consider an example where $\mathbf{B}_{21} = \mathbf{B}_{43} = \mathbf{0}$; i.e.,

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_{11} & 0 & 0 & 0 \\ 0 & \mathbf{B}_{22} & 0 & 0 \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & 0 \\ \mathbf{B}_{41} & \mathbf{B}_{42} & 0 & \mathbf{B}_{44} \end{pmatrix}. \tag{65}$$

Gantmacher (1959, Section 13.4) calls a block \mathbf{B}_{ii} *isolated* if there are no other nonzero blocks on its row, that is, if $\mathbf{B}_{ij} = 0$ for $j < i$. I will call such a block *row-isolated* and introduce the term *column-isolated* to describe any block \mathbf{B}_{ii} with no other nonzero blocks in its column, that is, $B_{ji} = 0$ for $j > i$. In Eq. (65), \mathbf{B}_{11} and \mathbf{B}_{22} are row-isolated and \mathbf{B}_{33} and \mathbf{B}_{44} are column-isolated.

If \mathbf{B}_{ii} is row-isolated, then the life cycle graph contains no pathways from any state outside of the subspace R_i to any state inside R_i , and R_i is a source. If \mathbf{B}_{ii} is column-isolated, then the life cycle graph contains no pathways from any state in R_i to any state outside R_i , and R_i is a sink.

The eigenvalues of \mathbf{A} are the eigenvalues of the diagonal blocks \mathbf{B}_{ii} . Let λ_1 be the dominant eigenvalue of \mathbf{A} , with right and left eigenvectors \mathbf{w}_1 and \mathbf{v}_1 . The Perron–Frobenius theorem guarantees that λ_1 , \mathbf{w}_1 , and \mathbf{v}_1 are real and nonnegative. Gantmacher (1959, Chapter 13, Theorem 6) proves that the eigenvector \mathbf{w}_1 is strictly positive if and only if λ_1 is an eigenvalue of

every row-isolated block and is not an eigenvalue of any of the nonrow-isolated blocks. This makes it easy to demonstrate the following corollary.

Corollary (Positivity of \mathbf{v}_1) *Let \mathbf{v}_1 be the left eigenvector corresponding to $\lambda_1[\mathbf{A}]$. Then \mathbf{v}_1 is strictly positive if and only if $\lambda_1[\mathbf{A}]$ is an eigenvalue of every column-isolated block, and is not an eigenvalue of any noncolumn-isolated block.*

To see this, note that \mathbf{v}_1 is the right eigenvector of \mathbf{A}^T . The column-isolated blocks of \mathbf{A} become row-isolated blocks of the block lower-triangular form of \mathbf{A}^T , and application of Gantmacher’s Theorem 6 proves the Corollary.

For example, transposing Eq. (65) gives

$$\mathbf{A}^T = \begin{pmatrix} \mathbf{B}_{11}^T & 0 & \mathbf{B}_{31}^T & \mathbf{B}_{41}^T \\ 0 & \mathbf{B}_{22}^T & \mathbf{B}_{32}^T & \mathbf{B}_{42}^T \\ 0 & 0 & \mathbf{B}_{33}^T & 0 \\ 0 & 0 & 0 & \mathbf{B}_{44}^T \end{pmatrix}. \tag{66}$$

Reversing the order of the rows and columns gives the block lower-triangular form

$$\begin{pmatrix} \mathbf{B}_{44}^T & 0 & 0 & 0 \\ 0 & \mathbf{B}_{33}^T & 0 & 0 \\ \mathbf{B}_{42}^T & \mathbf{B}_{32}^T & \mathbf{B}_{22}^T & 0 \\ \mathbf{B}_{41}^T & \mathbf{B}_{31}^T & 0 & \mathbf{B}_{11}^T \end{pmatrix}. \tag{67}$$

The column-isolated blocks in \mathbf{A} (\mathbf{B}_{33} and \mathbf{B}_{44}) now appear as row-isolated blocks in \mathbf{A}^T .

The usefulness of the Corollary follows from the population projection model

$$\mathbf{n}(t + 1) = \mathbf{A}\mathbf{n}(t) \quad \mathbf{n}(0) = \mathbf{n}_0 \tag{68}$$

and its solution³

$$\mathbf{n}(t) = \sum_{i=1}^s c_i \lambda_i^t \mathbf{w}_i \tag{69}$$

$$= \sum_{i=1}^s (\mathbf{v}_i^T \mathbf{n}_0) \lambda_i^t \mathbf{w}_i \tag{70}$$

(Caswell 2001). If \mathbf{n}_0 is such that $c_1 = \mathbf{v}_1^T \mathbf{n}_0$ is positive, then λ_1^t will eventually dominate all other terms in the solution and the population will grow at the rate λ_1 with stable structure \mathbf{w}_1 . We know the following about c_1 :

1. If \mathbf{A} is irreducible, then by the Perron–Frobenius theorem \mathbf{v}_1 is strictly positive, so any nonnegative, nonzero initial population \mathbf{n}_0 leads to a positive value of c_1 and eventual growth at the rate λ_1 .

³This holds provided that \mathbf{A} is diagonalizable, which is a generic property for linear operators (Hirsch and Smale 1974, p. 157).

2. If \mathbf{A} is reducible and \mathbf{v}_1 is strictly positive, any nonnegative, nonzero \mathbf{n}_0 leads to a positive value of c_1 and growth at the rate λ_1 .
3. If \mathbf{A} is reducible and \mathbf{v}_1 contains zero entries corresponding to a subspace R_i , then initial conditions with positive support only in R_i will lead to $c_1 = 0$, and λ_1 will make no contribution to population growth from those initial vectors.

In the first two cases, population growth is ergodic from any nonzero initial population. In the third case, there exists a basin of attraction leading to growth according to λ_1 and a basin (or basins) of attraction for growth according to the dominant eigenvalues of the diagonal blocks \mathbf{B}_{ii} corresponding to the zero entries of \mathbf{v}_1 .

References

- Abadir KM, Magnus JR (2005) Matrix algebra. Econometric exercises 1. Cambridge University Press, Cambridge
- Barfield M, Holt RD, Gomulkiewicz R (2011) Evolution in stage-structured populations. *Am Nat* 177:397–409
- Baudisch A (2005) Hamilton's indicators of the force of selection. *Proc Natl Acad Sci* 101:8263–8268
- Baudisch A (2008) Inevitable aging? Contributions to evolutionary–demographic theory. Springer, Berlin
- Caswell H (1982) Stable population structure and reproductive value for populations with complex life cycles. *Ecology* 63:1223–1231
- Caswell H (2001) Matrix population models: construction, analysis, and interpretation, 2nd edn. Sinauer Associates, Sunderland, Massachusetts
- Caswell H (2006) Applications of Markov chains in demography. In: Langville AN, Stewart WJ (eds) MAM2006: Markov anniversary meeting. Bosen Books, Raleigh, pp 319–334
- Caswell H (2007) Sensitivity analysis of transient population dynamics. *Ecol Res* 10:1–15
- Caswell H (2008) Perturbation analysis of nonlinear matrix population models. *Demogr Res* 18:59–116
- Caswell H (2009a) Sensitivity and elasticity of density-dependent population models. *J Differ Equ Appl* 15:349–369
- Caswell H (2009b) Stage, age, and individual stochasticity in demography. *Oikos* 118:1763–1782
- Caswell H (2010) Reproductive value, the stable stage distribution, and the sensitivity of the population growth rate to changes in vital rates. *Demogr Res* 23:531–548
- Caswell H (2011a) Beyond R_0 : demographic models for variability in lifetime reproductive output. *PLoS ONE* (in press)
- Caswell H (2011b) Perturbation analysis of continuous-time absorbing Markov chains. *Numer Linear Algebr Appl* (in press)
- Charlesworth B (1994) Evolution in age-structured populations, second edn. Cambridge University Press, Cambridge
- Cohen JE (1979) Ergodic theorems in demography. *Bull Am Math Soc* 1:275–295
- Csetenyi AI, Logofet DO (1989) Leslie model revisited: some generalizations to block structures. *Ecol Model* 48:277–290
- Dercole F, Rinaldi S (2008) Analysis of evolutionary processes: the adaptive dynamics approach and its applications. Princeton University Press, Princeton
- Feichtinger G (1971) Stochastische modelle demographischer prozesse. Lecture notes in operations research and mathematical systems. Springer, Berlin
- Gantmacher FR (1959) The theory of matrices. Chelsea Publishing, New York
- Goldberg EE, Lynch HJ, Neubert MG, Fagan WF (2010) Effects of branching spatial structure and life history on the asymptotic growth rate of a population. *Theor Ecol* 3:137–152
- Goodman LA (1969) The analysis of population growth when the birth and death rates depend upon several factors. *Biometrics* 25:659–681
- Hamilton WD (1966) The moulding of senescence by natural selection. *J Theor Biol* 12:12–45
- Henderson HV, Searle SR (1981) The vec-permutation, the vec operator and Kronecker products: a review. *Linear Multilinear Algebra* 9:271–288
- Hirsch MW, Smale S (1974) Differential equations, dynamical systems, and linear algebra. Academic, New York
- Honeycutt AA, Boyle JP, Broglio KR, Thompson TJ, Hoerger TJ, Geiss LS, Venkat Narayan KM (2003) A dynamic Markov model for forecasting diabetes prevalence in the United States through 2050. *Health Care Manage Sci* 6:155–164
- Horvitz CC, Tuljapurkar S (2008) Stage dynamics, period survival, and mortality plateaus. *Am Nat* 172:203–215
- Hougaard P (2000) Analysis of multivariate survival data. Springer, New York
- Hunter CM, Caswell H (2005) The use of the vec-permutation matrix in spatial matrix population models. *Ecol Model* 188:15–21
- Iosifescu M (1980) Finite Markov processes and their applications. Wiley, New York
- Jenouvrier S, Caswell H, Barbraud C, Weimerskirch H (2010) Mating behavior, population growth and the operational sex ratio: a periodic two-sex model approach. *Am Nat* 175:739–752
- Klepac P, Caswell H (2010) The stage-structured epidemic: a multi-state matrix population model approach. *Theor Ecol*. doi:10.1007/s12080-010-0079-8
- Law R (1983) A model for the dynamics of a plant population containing individuals classified by age and size. *Ecology* 64:224–230
- Lebreton J-D (1996) Demographic models for subdivided populations: the renewal equation approach. *Theor Popul Biol* 49:291–313
- Lebreton J-D, Nichols JD, Barker RJ, Pradel R, Spendelov JA (2009) Modeling individual animal histories with multistate capture-recapture models. *Adv Ecol Res* 41:87–173
- Logofet DO (1993) Matrices and graphs: stability problems in mathematical ecology. CRC, Boca Raton
- Logofet DO (2002) Three sources and three constituents of the formalism for a population with discrete age and stage structures. *Math Model* 14:11–22 (in Russian, with English summary)
- Logofet DO, Belova IN (2007) Nonnegative matrices as a tool to model population dynamics: classical models and contemporary expansions. *J Math Sci* 155:894–907 (in Russian with English summary)
- Magnus JR, Neudecker H (1979) The commutation matrix: some properties and applications. *Ann Stat* 7:381–394
- Magnus JR, Neudecker H (1985) Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *J Math Psychol* 29:474–492
- Magnus JR, Neudecker H (1988) Matrix differential calculus with applications in statistics and econometrics. Wiley, New York

- Medawar PB (1952) An unsolved problem in biology. H K Lewis, London
- Metz JAJ, Diekmann O (1986) The dynamics of physiologically structured populations. Springer, Berlin
- Metz JAJ, Nisbet RM, Geritz SAH (1992) How should we define ‘fitness’ for general ecological scenarios? *Trends Ecol Evol* 7:198–202
- Neubert MG, Parker IM (2004) Projecting rates of spread for invasive species. *Risk Anal* 24:817–831
- Ozgul A, Oli MK, Armitage KB, Blumstein DT, van Vuren DH (2009) Influence of local demography on asymptotic and transient dynamics of a yellow-bellied marmot metapopulation. *Am Nat* 173:517–530
- Parker IM (2000) Invasion dynamics of *Cytisus scoparius*: a matrix model approach. *Ecol Appl* 10:726–743
- Peeters A, Mamun AA, Willekens F, Bonneux L (2002) A cardiovascular life history: a life course analysis of the original Framingham Heart Study cohort. *Eur Heart J* 23:458–466
- Rice SH (2004) Evolutionary theory: mathematical and conceptual foundations. Sinauer, Sunderland, MA
- Rogers A (1966) The multiregional matrix growth operator and the stable interregional age structure. *Demography* 3:537–544
- Rogers A (1975) Introduction to multiregional mathematical demography. Wiley, New York
- Rogers A (1995) Multiregional demography: principles, methods and extensions. Wiley, New York
- Rose MR (1991) Evolutionary biology of aging. Oxford University Press, Oxford
- Strasser CA, Neubert MG, Caswell H, Hunter CM (2010) Contributions of high- and low-quality patches to a metapopulation with stochastic disturbance. *Theor Ecol*. doi:10.1007/s12080-010-0106-9
- Tuljapurkar S, Horvitz CC (2006) From stage to age in variable environments: life expectancy and survivorship. *Ecology* 87:1497–1509
- van Groenendael JM, Slim P (1988) The contrasting dynamics of two populations of *Plantago lanceolata* classified by age and size. *J Ecol* 76:585–599
- Vaupel JW, Baudisch A, Dölling M, Roach DA, Gampe JJW (2010) Biodemography of human ageing. *Nature* 464:536–542
- Vaupel JW, Baudisch A, Dölling M, Roach DA, Gampe J (2004) The case for negative senescence. *Theor Popul Biol* 65:339–351
- Willekens FJ (2002) Multistate demography. *Encyclopaedia of Population*, Macmillan Reference, New York