



**UvA-DARE (Digital Academic Repository)**

**Age × stage-classified demographic analysis: a comprehensive approach**

Caswell, H.; de Vries, C.; Hartemink, N.; Roth, G.; van Daalen, S.F.

*Published in:*  
Ecological Monographs

*DOI:*  
[10.1002/ecm.1306](https://doi.org/10.1002/ecm.1306)

[Link to publication](#)

*Citation for published version (APA):*  
Caswell, H., de Vries, C., Hartemink, N., Roth, G., & van Daalen, S. F. (2018). Age × stage-classified demographic analysis: a comprehensive approach. *Ecological Monographs*, 88(4), 560-584.  
<https://doi.org/10.1002/ecm.1306>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



# CONCEPTS & SYNTHESIS

EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

*Ecological Monographs*, 0(0), 2018, pp. 1–25

© 2018 The Authors. *Ecological Monographs* published by Wiley Periodicals, Inc. on behalf of Ecological Society of America

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Age $\times$ stage-classified demographic analysis: a comprehensive approach

HAL CASWELL,<sup>1</sup> CHARLOTTE DE VRIES, NIENKE HARTEMINK, GREGORY ROTH, AND SILKE F. VAN DAALEN

*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands*

**Abstract.** This paper presents a comprehensive theory for the demographic analysis of populations in which individuals are classified by both age and stage. The earliest demographic models were age classified. Ecologists adopted methods developed by human demographers and used life tables to quantify survivorship and fertility of cohorts and the growth rates and structures of populations. Later, motivated by studies of plants and insects, matrix population models structured by size or stage were developed. The theory of these models has been extended to cover all the aspects of age-classified demography and more. It is a natural development to consider populations classified by both age and stage. A steady trickle of results has appeared since the 1960s, analyzing one or another aspect of age  $\times$  stage-classified populations, in both ecology and human demography. Here, we use the vec-permutation formulation of multistate matrix population models to incorporate age- and stage-specific vital rates into demographic analysis. We present cohort results for the life table functions (survivorship, mortality, and fertility), the dynamics of intra-cohort selection, the statistics of longevity, the joint distribution of age and stage at death, and the statistics of life disparity. Combining transitions and fertility yields a complete set of population dynamic results, including population growth rates and structures, net reproductive rate, the statistics of lifetime reproduction, and measures of generation time. We present a complete analysis of a hypothetical model species, inspired by poecilogonous marine invertebrates that produce two kinds of larval offspring. Given the joint effects of age and stage, many familiar demographic results become multidimensional, so calculations of marginal and mixture distributions are an important tool. From an age-classified point of view, stage structure is a form of unobserved heterogeneity. From a stage-classified point of view, age structure is unobserved heterogeneity. In an age  $\times$  stage-classified model, variance in demographic outcomes can be partitioned into contributions from both sources. Because these models are formulated as matrices, they are amenable to a complete sensitivity analysis. As more detailed and longer longitudinal studies are developed, age  $\times$  stage-classified demography will become more common and more important.

**Key words:** age-stage classification; elasticity; generation time; heterogeneity; Markov chain models; matrix population models; mortality; net reproductive rate; sensitivity; survivorship; vec-permutation matrix.

### INTRODUCTION

As demographic data become more detailed, they reveal more, and more diverse, differences among individuals. These types of heterogeneity led to the development of increasingly complicated demographic models. Demographic analysis in ecology can be divided into three periods. The first, building on the work of Lotka and many others, used age-classified models and life table functions to compute survival, fertility, life expectancy, intrinsic rates of increase, stable age distributions, and reproductive values. It became a fully developed population theory by the mid-20th century (e.g., Lotka 1939, Keyfitz 1968, Coale 1972). It was quickly adopted by population biologists and ecologists, from the mortality studies of Pearl and Miner (1935) and the comparative life table studies of Deevey (1947) to the use of the Euler-Lotka equation to calculate population growth rates (e.g., Birch 1948, Leslie and Park 1949). Fisher (1930), Norton (1928), and

Charlesworth (1994), among others, applied age-classified theory to evolutionary questions, and it became the basis of evolutionary demography and life history theory (e.g., Cole 1954, Lewontin 1965, Hamilton 1966, Stearns 1992).

The second period saw the widespread adoption of stage-classified methods based largely on matrix population models classified by stage (Lefkovich 1965, Werner and Caswell 1977) where stage might refer to size, instar, developmental stage, physiological condition, behavioral status, etc. Plant ecologists were at the forefront of this development because the modular construction and plastic growth of plants tend to make age less suitable as a state variable in many cases. Today, age-classified models are still more common among animal models and stage-classified models more common among plant models (Salguero-Gómez et al. 2015, 2016).

Stage-classified matrix models made it possible to compute population growth rates and structures, sensitivity, and elasticity analyses, and their extensions to periodic and stochastic environments, as well as to develop sophisticated treatments of nonlinear dynamics. This led eventually to a fully developed theory of stage-classified demography in a variety of mathematical frameworks (Nisbet and Gurney

Manuscript received 12 March 2017; revised 23 February 2018; accepted 21 March 2018. Corresponding Editor: Aimée T. Classen.

<sup>1</sup>E-mail: h.caswell@uva.nl

1982, Metz and Diekmann 1986, Tuljapurkar 1990, Tuljapurkar and Caswell 1997, Easterling et al. 2000, Caswell 2001, Ellner et al. 2016).

Now, increasingly extensive and detailed individual data are being collected and it is becoming apparent that combining age- and stage-classification is sometimes useful. This is accompanied, in both ecology and human demography, by an interest in incorporating more biological or social detail into the analyses.

Our goal here is to present a comprehensive methodological treatment of age  $\times$  stage-classified matrix population models. The methods will be presented in a way that can be applied to any species, with any life cycle, described by any pattern of transitions among any set of stages, with any type of reproduction of any kind(s) of offspring. We will provide a hypothetical example to suggest how a particular species might be analyzed.

#### *Explicit and implicit age dependence*

As a first step, we recognize that age dependence may be either implicit or explicit. Any stage-classified model produces implicit age-classified results because, even though age has no effect on the vital rates, an individual still becomes one unit older with the passage of one unit of time. Age-specific results (e.g., life expectancy) computed from such a model are implicit in the stage structure of the life cycle. Methods based on absorbing Markov chains are widely used to extract implicit age dependence from stage-classified models (e.g., Feichtinger 1971, Cochran and Ellner 1992, Caswell 2001, 2006, 2009, Steinsaltz and Evans 2004, Tuljapurkar and Horvitz 2006, Horvitz and Tuljapurkar 2008). These are not, however, age  $\times$  stage-classified models and the results do not reveal anything about the interaction of age-specific and stage-specific rates.

This paper addresses *explicit* age dependence, in which individuals are classified by their age and stage, and the vital rates depend on both variables. The literature on explicit age  $\times$  stage demography is scattered, with fragmentary results appearing in the ecological and demographic literatures. Goodman (1969) first introduced the matrix formulation of age-stage models. Even before that, a continuous-time age  $\times$  size model written as a partial differential equation was presented by Sinko and Streifer (1967). Csetenyi and Logofet (1989) and Logofet (2002, 2013) explored the graph-theoretic implications of the block structure appearing in age  $\times$  stage models, in relation to the irreducibility and primitivity of the resulting model. Recent papers (Caswell 2012, Caswell and Shyu 2012) have developed matrix models and their sensitivity analysis, for both cohort properties and population growth rate, for age  $\times$  stage-classified models. These were used in an evolutionary analysis of the selection gradients on senescence, as a function of age and of stage, by Caswell and Salguero-Gómez (2013). Steiner et al. (2012) explored some aspects of longevity and later (Steiner et al. 2014) developed a powerful age  $\times$  stage-classified analysis of the relation between generation time and population growth rate, with applications to life history theory (generation time will be considered here in the section entitled *Population dynamics: generation time*).

Multiregional models are an important special case of age  $\times$  stage-classified models. They classify individuals by age and spatial location, and were one of the earliest applications of matrix population models (Rogers 1968) and have been extensively developed in human demography (e.g., Rogers 1975) and ecology (Lebreton 1996, 2005, Lebreton et al. 2000). There is a rich literature in health demography of multistate models (e.g., Willekens 2014) in which individuals are classified by age and health status (see, e.g., Wu et al. [2006] for colorectal cancer, Zhou et al. (2016) for dementia, or Honeycutt et al. (2003) for diabetes). These models focus on longevity and occupancy times in various age-stage combinations, and include only the survival and transition portion of the life cycle. Some epidemiological models have combined age and infection status in nonlinear matrix models that include the infection process (e.g., Klepac and Caswell 2011, Metcalf et al. 2012).

Perhaps the human demographic usage closest in conceptual structure to ecological practice are multistate population projections, which are used to forecast the short-term, transient dynamics of populations under scenarios of changing mortality, fertility, and migration rates. The basic age-specific projections are sometimes augmented by a stage classification to examine the joint age  $\times$  stage dynamics. The flavor of these efforts is shown by studies on, e.g., the interaction of age with citizenship status (Sánchez Gassen 2014), disability (Van der Gaag et al. 2015), or education level (Loichinger 2015).

Ecological examples of explicit age  $\times$  stage models are not common. Law (1983) considered the theoretical problem of age  $\times$  stage models but presented only a hypothetical example. De Roos (2008) analyzes a continuous-time model in which age and stage can be combined. More empirical studies can be found in, e.g., van Groenendael and Slim (1988) and Zuidema et al. (2009).

#### *Heterogeneity*

The more closely one looks at individual organisms, the more ways in which they appear to differ. The problem of heterogeneity is to figure out how those differences affect population dynamics. This is the essence of demography: accounting for differences due to age, development, sex, physiological condition, breeding status, etc. The analyses to be presented here are a way to incorporate diverse types of heterogeneity into demographic analysis.

#### *Outline*

The rest of the paper is organized as follows:

1. A systematic method for the construction of age  $\times$  stage-classified models.
2. The life table functions that define survivorship, mortality, and the age and stage at death of individuals, and quantify intra-cohort selection.
3. The fertility functions that describe reproduction of individuals across the life cycle.
4. The statistics of longevity for cohorts, including life expectancy and also measures of variability, state occupancy times, and life disparity.

5. Characteristics of population dynamics, including the population growth rate, stable structure, reproductive value, net reproductive rate, and measures of generation time.
6. Sensitivity and elasticity analysis, showing a completely general formula for the effects of changes in any parameter[s] affecting any of the age-specific or stage-specific rates on any quantity calculated from the model.

TABLE 1. Age × stage-classified demographic analysis.

	Notation	Equation
<b>Model construction</b>		
Population vector	$\tilde{\mathbf{N}}$	3
Forbidden age-stage combinations	$\mathbf{U}_i, \mathbf{F}_i$	8, 9
Block-diagonal matrices	$\mathbf{U}, \mathbf{F}, \mathbf{D}, \mathbf{H}$	10
Age × stage-classified projection matrices	$\tilde{\mathbf{U}}, \tilde{\mathbf{F}}, \tilde{\mathbf{A}}$	11–13
<b>Life table functions</b>		
Survivorship vector[s]	$\ell(x), \mathbf{L}$	17, 46
Survivorship of mixed cohort	$\ell(x   \boldsymbol{\pi})$	19
Mortality rate vector	$\boldsymbol{\mu}(x)$	22
Mortality rate of mixed cohort	$\mu(x   \boldsymbol{\pi})$	23
<b>Intra-cohort selection</b>		
Joint age × stage	$\tilde{\mathbf{m}}(x   \tilde{\mathbf{n}}_0)$	25
Marginal stage	$\tilde{\mathbf{m}}_{\text{stage}}(x   \tilde{\mathbf{n}}_0)$	26
Marginal age	$\tilde{\mathbf{m}}_{\text{age}}(x   \tilde{\mathbf{n}}_0)$	27
<b>Fertility</b>		
Fertility matrix	$\mathbf{F}_x$	
Weighted fertility vector	$\mathbf{f}_{\text{weighted}}(x)$	29
Mixed fertility vector	$\mathbf{f}_{\text{mixed}}(x)$	30
Weighted and mixed fertility vector	$f(x)$	32
<b>Longevity statistics</b>		
Fundamental matrix	$\tilde{\mathbf{N}}$	34
Moments and variance of longevity	$\tilde{\eta}_i, V(\tilde{\eta})$	35–37
Decomposition of variance in longevity	$V_{\text{within}}, V_{\text{between}}$	43, 45
<b>Distribution of age and stage at death</b>		
Joint age × stage distribution	$\tilde{\mathbf{B}}$	50
Marginal age distribution	$\mathbf{B}_{\text{age}}$	51
Marginal stage distribution	$\mathbf{B}_{\text{stage}}$	52
Life disparity, years of life lost	$\tilde{\eta}^\dagger$	57
<b>Population dynamics</b>		
Population projection	$\tilde{\mathbf{n}}(t+1) = \tilde{\mathbf{A}}\tilde{\mathbf{n}}(t)$	60
<b>Stable population structure</b>		
Marginal age structure	$\tilde{\mathbf{w}}_{\text{age}}$	61
Marginal stage structure	$\tilde{\mathbf{w}}_{\text{stage}}$	62
Reproductive value	$\mathbf{v}$	63
Net reproductive rate	$R_0$	67
<b>Mean lifetime reproduction</b>		
Next generation matrix	$\mathbf{R}_{11}$	73
Weighted lifetime reproduction	$\mathbf{r}_{\text{weighted}}$	75
Mixed lifetime reproduction	$\mathbf{r}_{\text{mixed}}$	76
Total lifetime reproduction	$\rho$	77
Cohort generation time	$\boldsymbol{\Gamma}$	88
<b>Sensitivity analysis</b>		
Sensitivity of output $\xi$ to parameters $\boldsymbol{\theta}$	$d\xi/d\boldsymbol{\theta}^\top$	89

Note: Results are shown for model construction, life table functions, longevity statistics, and population dynamics, with equation numbers in which the results are presented.

Table 1 is a list of all the demographic results that will be presented, keyed to the equations in which they are derived.

*Notation.*—Matrices are denoted by uppercase boldface letters (e.g.,  $\mathbf{U}$ ), and vectors by lowercase boldface letters (e.g.,  $\mathbf{n}$ ). Block-diagonal matrices are denoted by blackboard font (e.g.,  $\mathbb{U}$ ). Matrices and vectors associated with the full age × stage model are denoted, e.g.,  $\tilde{\mathbf{U}}, \tilde{\mathbf{n}}$ ; these matrices are block structured and contain entries for all combinations of age classes and stages. The number of age classes is  $\omega$  and the number of stages is  $s$ . The notation for matrices used in model construction is tabulated for easy reference in Table 2.

The unit vector  $\mathbf{e}_i$  is a vector with a 1 in the  $i$ th entry and zeros elsewhere. The unit matrix  $\mathbf{E}_{ij}$  is a matrix with a 1 in the  $(i,j)$  entry and zeros elsewhere; the dimensions will be indicated if not clear from the context. The dimensions of certain matrices and vectors are denoted by subscripts;  $\mathbf{I}_s$  is an identity matrix of order  $s$  and  $\mathbf{1}_s$  is a  $s \times 1$  vector of ones. When convenient, MATLAB (MathWorks, Natick, MA, USA) notation will be used to refer to rows and columns of matrices; thus  $\mathbf{X}(i,:)$  is the  $i$ th row and  $\mathbf{X}(:,j)$  the  $j$ th column of  $\mathbf{X}$ . The diagonal matrix with  $\mathbf{x}$  on the diagonal and zeros elsewhere is denoted  $\mathcal{D}(\mathbf{x})$ . The symbol  $\circ$  denotes the Hadamard, or element-by-element product; the symbol  $\otimes$  denotes the Kronecker product. The vec operator transforms a matrix to a vector by stacking the columns on top of each other. The symbol  $\|\mathbf{x}\|$  denotes the 1-norm of the vector  $\mathbf{x}$ . The transpose of the matrix  $\mathbf{X}$  is  $\mathbf{X}^\top$ . The matrix  $\mathbf{K}$  is the vec-permutation matrix (Henderson and Searle 1981); see Box 2. The mean and variance are denoted by  $E(\cdot)$  and  $V(\cdot)$  respectively.

CONSTRUCTING AGE × STAGE MODELS

*The population vector*

Each individual is jointly classified by its age and its stage; the population composition at any time can be written as

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

where rows correspond to stages (1, ...,  $s$ ) and columns to age classes (1, ...,  $\omega$ ). The population vector  $\tilde{\mathbf{n}}$  is obtained from  $\mathcal{N}$  as

$$\tilde{\mathbf{n}} = \text{vec}\mathcal{N} \quad (2)$$

$$= \begin{pmatrix} n_{11} \\ \vdots \\ n_{s1} \\ \vdots \\ \overline{n_{1\omega}} \\ \vdots \\ n_{s\omega} \end{pmatrix} \quad (3)$$

In  $\tilde{\mathbf{n}}$ , stages are grouped within age classes. The vector can be transformed so that age classes are grouped within stages,

TABLE 2. Matrices used in constructing the age  $\times$  stage-classified projection matrix, where  $\omega$  denotes the number of age classes and  $s$  the number of stages.

Symbol	Expression	Size	Description
Matrix describing the full population			
$\tilde{\mathbf{A}}$	$\tilde{\mathbf{U}} + \tilde{\mathbf{F}}$	$s\omega \times s\omega$	Age $\times$ stage population projection matrix
Matrices describing transitions and survival of existing individuals			
$\tilde{\mathbf{U}}$	$\mathbf{K}^T \mathbb{D} \mathbf{K} \mathbf{U}$	$s\omega \times s\omega$	Age $\times$ stage transition and survival matrix
$\mathbf{U}$	$\sum_{j=1}^{\omega} (\mathbf{E}_{jj} \otimes \mathbf{U}_j)$	$s\omega \times s\omega$	Block diagonal age $\times$ stage transition matrix
$\mathbb{D}$	$\sum_{i=1}^{\omega} (\mathbf{E}_{ii} \otimes \mathbf{D}_i)$	$s\omega \times s\omega$	Block diagonal age transition matrix
$\mathbf{U}_j$	$\mathbf{U}_j$	$s \times s$	Stage transition matrix for age class $j$
$\mathbf{D}_i$	$\mathbf{D}_i$	$\omega \times \omega$	Age transition matrix for stage $i$
Matrices describing reproduction			
$\tilde{\mathbf{F}}$	$\mathbf{K}^T \mathbb{H} \mathbf{K} \mathbf{F}$	$s\omega \times s\omega$	Age $\times$ stage fertility matrix
$\mathbf{F}$	$\sum_{j=1}^{\omega} (\mathbf{E}_{jj} \otimes \mathbf{F}_j)$	$s\omega \times s\omega$	Block diagonal fertility matrix
$\mathbb{H}$	$\sum_{i=1}^s (\mathbf{E}_{ii} \otimes \mathbf{H}_i)$	$s\omega \times s\omega$	Block diagonal age assignment matrix
$\mathbf{F}_j$	$\mathbf{F}_j$	$s \times s$	Fertility matrix for age class $j$
$\mathbf{H}_i$	$\mathbf{H}_i$	$\omega \times \omega$	Age assignment matrix for offspring of stage $i$

Note: The index  $i$  denotes stages,  $i = 1, \dots, s$ ; the index  $j$  denotes age classes,  $j = 1, \dots, \omega$ .

$$\text{vec}(\mathcal{N}^T) = \mathbf{K}_{s,\omega} \text{vec} \mathcal{N} \quad (4)$$

$$\mathbf{A} = \mathbf{U} + \mathbf{F}. \quad (5)$$

where  $\mathbf{K}_{s,\omega}$  is the vec permutation matrix (Magnus and Neudecker 1979, Henderson and Searle 1981). This transformation is essential to the analysis of the model, as will be seen below. An explicit formula for  $\mathbf{K}_{s,\omega}$  is given in Box 2. All results can be obtained using this arrangement, by properly reformulating the relevant matrices.

#### Age- and stage-specific demography

The influence of age and stage on the vital rates is captured in four sets of matrices:

1. Matrices  $\mathbf{U}_j$  for  $j = 1, \dots, \omega$  (dimension  $s \times s$ ). These matrices contain stage transitions, including mortality, of living individuals in each age class. Because the entries of the  $\mathbf{U}_j$  refer to stages, and each matrix corresponds to an age class, the set includes arbitrarily complicated interactions of age and stage in determining mortality and transitions.
2. Matrices  $\mathbf{F}_j$  for  $j = 1, \dots, \omega$  (dimension  $s \times s$ ). These matrices contain fertilities, describing the stage-specific per capita production of new individuals by reproduction, for individuals in each age class. Because the entries of the  $\mathbf{F}_j$  refer to stages, and each matrix corresponds to an age class, the set includes arbitrarily complicated interactions of age and stage in determining fertility.
3. Matrices  $\mathbf{D}_i$  for  $i = 1, \dots, s$  (dimension  $\omega \times \omega$ ). These matrices contain age transitions for individuals in each stage.
4. Matrices  $\mathbf{H}_i$  for  $i = 1, \dots, s$  (dimension  $\omega \times \omega$ ). These matrices assign newly produced offspring, of parents in stage  $i$ , to an initial age class (usually the first).

The matrices  $\mathbf{U}_j$  and  $\mathbf{F}_j$  correspond to the familiar decomposition of a projection matrix  $\mathbf{A}$  into components due to transitions and survival of extant individuals ( $\mathbf{U}$ ) and due to fertility ( $\mathbf{F}$ )

Now, however, the matrices are defined for each of the  $\omega$  age classes. This flexible formulation permits age to influence any of the stage-specific vital rates, in any way, described in either parametric or nonparametric terms. Similarly, it permits stages to influence age-specific rates in any way.

The matrix  $\mathbf{F}_j$  describes reproduction by all stages in age class  $j$ . An age  $\times$  stage model must account for the possibility that multiple types of offspring may be produced (e.g. a plant may produce seedlings of different sizes). In the special case where all offspring are of the same stage,  $\mathbf{F}_j$  will have positive entries in only one row, usually chosen to be the first.

The matrices  $\mathbf{U}_j$  and  $\mathbf{F}_j$  move individuals among stages, account for survival, and create new offspring of various stages, at rates that depend on the current age. Following these processes, individuals must be allocated to their next age class. The matrices  $\mathbf{D}_i$  advance surviving individuals from one age class to the next; the simplest such model contains ones on the subdiagonal and the  $\omega, \omega$  corner; e.g., (for  $\omega = 3$ )

$$\mathbf{D}_i = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad i = 1, \dots, s. \quad (6)$$

The  $\omega, \omega$  entry converts the final age class into an open-ended category. Setting it to 0 would kill all individuals at age  $\omega$ .

As written in Eq. 6, advancement in age involves no deaths; all mortality is included in the  $\mathbf{U}_j$ . This need not be the case. If desired, an additional age-specific mortality hazard, affecting stage  $i$  could be incorporated by replacing the ones in  $\mathbf{D}_i$  by survival probabilities  $< 1$ . Such mortality can always be incorporated into the  $\mathbf{U}_j$ , so we will not explore its incorporation into the  $\mathbf{D}_i$  here.

### Box 1. Creating marginals and mixtures

The Kronecker product expressions for computing marginals (e.g., Eqs. 26 and 27) and mixtures [e.g., equation (19)] may appear confusing at first sight. There is a simple trick to deriving them, which we will reveal here. It relies on the properties of the  $\text{vec}$  operator and on Roth's theorem that, for any matrices  $\mathbf{X}$ ,  $\mathbf{Y}$  and  $\mathbf{Z}$ ,

$$\text{vec}(\mathbf{XYZ}) = (\mathbf{Z}^\top \otimes \mathbf{X})\text{vec } \mathbf{Y}$$

(Roth 1934). In this paper, we have agreed to organize  $\tilde{\mathbf{n}}$  by grouping stages within age classes; that is, by applying the  $\text{vec}$  operator to the array  $\mathcal{N}$  in Eq. 1, in which stages appear as rows and ages as columns

$$\mathcal{N} = \begin{pmatrix} n_{11} & \cdots & n_{1\omega} \\ \vdots & & \vdots \\ n_{s1} & \cdots & n_{s,\omega} \end{pmatrix}.$$

Matrices (e.g.,  $\tilde{\mathbf{U}}$ ,  $\tilde{\mathbf{N}}$ ) and vectors (e.g.,  $\tilde{\mathbf{w}}$ ,  $\tilde{\mathbf{v}}$ ) inherit the same block structure.

Any linear combination of rows of  $\mathcal{N}$  (i.e., combinations of stages) can be written as a matrix  $\mathbf{R}$ , and any linear combination columns of  $\mathcal{N}$  (i.e., of ages) can be written as a matrix  $\mathbf{C}$ , and applied to  $\mathcal{N}$  as

$$\mathbf{R}\mathcal{N}\mathbf{C}.$$

Applying the  $\text{vec}$  operator to this gives

$$(\mathbf{C}^\top \otimes \mathbf{R})\text{vec}\mathcal{N}.$$

Thus  $(\mathbf{C}^\top \otimes \mathbf{R})$  operating on the rows of any age × stage block-structured matrix or vector captures the operations implied by  $\mathbf{R}$  and  $\mathbf{C}$  operating on  $\mathcal{N}$ , e.g., adding all stages to get a marginal age structure

$$\left. \begin{array}{l} \mathbf{R} = \mathbf{1}_s^\top \\ \mathbf{C} = \mathbf{I}_\omega \end{array} \right\} \Rightarrow (\mathbf{I}_\omega \otimes \mathbf{1}_s^\top) \text{vec}\mathcal{N} \quad \omega \times 1$$

e.g., adding all ages to get a marginal stage distribution

$$\left. \begin{array}{l} \mathbf{R} = \mathbf{I}_s \\ \mathbf{C} = \mathbf{1}_\omega^\top \end{array} \right\} \Rightarrow (\mathbf{1}_\omega^\top \otimes \mathbf{I}_s) \text{vec}\mathcal{N} \quad s \times 1$$

e.g., a mixture, defined by a vector  $\boldsymbol{\pi}_s$ , all of age class 1

$$\left. \begin{array}{l} \mathbf{R} = \boldsymbol{\pi}_s^\top \\ \mathbf{C} = \mathbf{e}_1 \end{array} \right\} \Rightarrow (\mathbf{e}_1^\top \otimes \boldsymbol{\pi}_s^\top) \text{vec}\mathcal{N} \quad 1 \times 1.$$

These matrices operate on arrays with the structure of  $\text{vec}\mathcal{N}$ ; i.e., on the rows of an object in which columns have stages arranged within ages. To operate on columns, for example on the columns of the fundamental matrix  $\tilde{\mathbf{N}}$ , the arrays must be transposed (the transpose of a Kronecker product is the product of the transposes). Thus the mixture in Eq. A.7 applied to the columns of  $\tilde{\mathbf{N}}$  would be

$$(\text{vec}\mathcal{N})^\top (\mathbf{e}_1 \otimes \boldsymbol{\pi}_s).$$

The matrix  $\mathbf{H}_i$  allocates the newborn individuals, regardless of the age of their parent, into age class 1. Thus (e.g., for  $\omega = 3$ )

$$\mathbf{H}_i = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad i = 1, \dots, s. \quad (7)$$

*Impossible age × stage combinations.*—An operational decision is required to deal with “impossible” age-stage combinations. For example, if stages are size classes, perhaps large individuals never occur in young age classes, and small individuals never occur in old age classes. In such cases, we set the rows and columns of the  $\mathbf{U}_j$  corresponding to these impossible combinations to 0

$$\text{stage } i \text{ and age } j \text{ is impossible} \implies \begin{cases} \mathbf{U}_j(:, i) = 0 \\ \mathbf{U}_j(i, :) = 0 \end{cases} \quad (8)$$

Because no offspring can be produced by a parent of an impossible age-stage combination

$$\text{stage } i \text{ and age } j \text{ is impossible} \implies \begin{cases} \mathbf{F}_j(:, i) = 0 \\ \mathbf{F}_j(i, :) = 0 \end{cases} \quad (9)$$

#### The age $\times$ stage-classified projection matrices

To construct the age  $\times$  stage model using the vec-permutation matrix formulation, we first construct block diagonal matrices  $\mathbf{U}$ ,  $\mathbf{F}$ ,  $\mathbf{D}$ , and  $\mathbf{H}$ , which contain the matrices  $\mathbf{U}_j$ ,  $\mathbf{F}_j$ ,  $\mathbf{D}_j$ , and  $\mathbf{H}_j$ , respectively, on the diagonal. That is

$$\mathbf{U} = \sum_{j=1}^{\omega} (\mathbf{E}_{jj} \otimes \mathbf{U}_j) = \begin{pmatrix} \mathbf{U}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{U}_{\omega} \end{pmatrix} \quad (10)$$

with similar construction for the others. The block diagonal matrices are all of dimension  $s\omega \times s\omega$ .

The age  $\times$  stage-classified projection matrices are defined in terms of the block-diagonal matrices as

$$\tilde{\mathbf{U}} = \mathbf{K}^T \mathbf{D} \mathbf{K} \mathbf{U} \quad (11)$$

$$\tilde{\mathbf{F}} = \mathbf{K}^T \mathbf{H} \mathbf{K} \mathbf{F} \quad (12)$$

$$\tilde{\mathbf{A}} = \tilde{\mathbf{U}} + \tilde{\mathbf{F}} \quad (13)$$

where  $\mathbf{K} = \mathbf{K}_{s,\omega}$  is the vec-permutation matrix. From right to left in Eq. 11, the matrix  $\mathbf{U}$  first moves individuals among stages within their age classes. The vec-permutation matrix  $\mathbf{K}$  rearranges age classes within stages, the matrix  $\mathbf{D}$  advances individuals to the next age class, and the matrix  $\mathbf{K}^T$  returns the vector to its original arrangement. In Eq. 12, the matrix  $\mathbf{F}$  produces new offspring, then  $\mathbf{K}$  rearranges age classes within stages, and the matrix  $\mathbf{H}$  assigns all newborn individuals to the first age class.

A little manipulation of these matrices reveals that  $\tilde{\mathbf{A}}$  has a block-Leslie form (e.g., for  $\omega = 3$ )

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 \\ \mathbf{U}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 & \mathbf{U}_3 \end{pmatrix} \quad (14)$$

where the  $\mathbf{0}$  matrices are of dimension  $s \times s$ . The formulation of  $\tilde{\mathbf{n}}$  in Eq. 3, in which stages are grouped within age classes, leads to this familiar structure (e.g., Goodman 1969, Feeney 1970, Lebreton 1996). The formulation of  $\tilde{\mathbf{n}}$  with age classes grouped within stages, where  $\tilde{\mathbf{n}} = \text{vec}(\mathcal{N}^T)$ , rearranges the blocks in Eq. 14 (e.g., Rogers 1968, Cohen 1982; see Caswell 2001: Section 4.3).

The systematic construction of  $\tilde{\mathbf{A}}$  by the vec-permutation matrix algorithm in Eqs. 11 and 12 is enormously beneficial. The biological content of the model, and hence the data collection and data analysis effort, comes from assembling the demographic information in the  $\mathbf{U}_j$  and  $\mathbf{F}_j$ . Of the  $s^2\omega^2$  entries of  $\tilde{\mathbf{A}}$ , at most only  $2s^2\omega$  entries contain demographic

information. The expressions for  $\tilde{\mathbf{U}}$  and  $\tilde{\mathbf{F}}$  isolate these components and make it more efficient to conduct analyses, especially sensitivity analyses.

In a later section (A model species example), we develop a hypothetical example to demonstrate the calculations, for a life cycle including two types of larvae and two types of adults.

#### LIFE TABLE FUNCTIONS

The classical life table functions are the survivorship  $\ell(x)$ , the mortality rate  $\mu(x)$ , and the distribution of age at death  $b(x)$ , where  $x$  represents age. Together, these functions are a time-honored way to characterize age-specific demography, longevity, and life histories (e.g., Pearl and Miner 1935, Dublin and Lotka 1937, Deevey 1947, Slobodkin 1961, Hutchinson 1978). As is well known, any one of them suffices to calculate any of the others.

In this section, we derive the life table functions for age  $\times$  stage-classified models. The extra dimension of stage classification adds a rich set of additional life table perspectives. In an age-classified model, for example, survival to some age  $x$  either happens or it does not; if it does, it follows a defined pathway of ages  $1, \dots, x$ . In an age  $\times$  stage-classified model, survival to a given age can take place via a potentially infinite set of developmental pathways through the combinations of ages and life history stages. Each pathway has its own probability of occurrence, and the life table functions must integrate over all those pathways and probabilities. In addition, if there are multiple types of offspring, a cohort may start off with a mixture of different stages at birth, which will affect the survival and mortality results. Fortunately, the matrix formulation makes these calculations possible.

#### Survivorship and mortality

Survivorship  $\ell(x)$  is the probability that an individual survives from age 0 to age  $x$ ,

$$\ell(x) = P(\text{survival from age 0 to age } x) \quad 1 \times 1. \quad (15)$$

In an age-classified model,  $\ell(x)$  is calculated by projecting an initial age 0 cohort of size 1; the resulting values give the proportion of the cohort surviving.

In an age  $\times$  stage-classified model, a survivorship vector is obtained by projecting a cohort using  $\tilde{\mathbf{U}}$

$$\tilde{\ell}(x) = (\mathbf{1}_{s\omega}^T \tilde{\mathbf{U}}^x)^T \quad x = 0, 1, \dots \quad (16)$$

The survivorship vector  $\tilde{\ell}(x)$ , of dimension  $s\omega \times 1$ , inherits the block age-stage structure of  $\tilde{\mathbf{U}}$ ; its entries give the probability of survival to age  $x$  of an individual starting in every age-stage combination. We extract the vector of survivorship starting from the first age class by

$$\ell(x) = (\mathbf{e}_1^T \otimes \mathbf{1}_s) \tilde{\ell}(x) \quad s \times 1 \quad (17)$$

where  $\mathbf{e}_1$  is a unit vector of length  $\omega$  with 1 in the first position and zeros elsewhere. The  $i$ th entry of  $\ell(x)$  gives the survivorship to age  $x$  of an individual starting life in stage  $i$ , accounting for the joint effects of age and stage on survival,

### Box 2. Computing the vec-permutation matrix

The vec-permutation matrix (Magnus and Neudecker 1979, Henderson and Searle 1981) connects the vec operator and the matrix transpose. If  $\mathbf{X}$  is a  $m \times n$  matrix, then

$$\text{vec } \mathbf{X}^\top = \mathbf{K}_{m,n} \text{vec } \mathbf{X}.$$

The matrix can be calculated as

$$\mathbf{K}_{m,n} = \sum_{i=1}^m \sum_{j=1}^n (\mathbf{E}_{ij} \otimes \mathbf{E}_{ij}^\top)$$

where  $\mathbf{E}_{ij}$  is a matrix, of dimension  $m \times n$ , with a 1 in the  $(i,j)$  entry and zeros elsewhere.

integrating over all age-stage developmental pathways between birth and age  $x$ .

If the cohort begins as a mixture of stages specified by a mixing distribution  $\boldsymbol{\pi}$ , the resulting survivorship function is

$$\ell(x|\boldsymbol{\pi}) = \boldsymbol{\pi}_s^\top \boldsymbol{\ell}(x) \quad (18)$$

$$= (\mathbf{e}_1^\top(\omega) \otimes \boldsymbol{\pi}^\top) \tilde{\boldsymbol{\ell}}(x) \quad \mathbf{1} \times \mathbf{1}. \quad (19)$$

Note that our convention (Eq. 8) that transitions involving impossible age-stage combinations are set to 0 implies that  $\ell_i(x) = 0$  if stage  $i$  does not occur among the types of offspring.

In age-classified models, the mortality rate, or hazard function,  $\mu(x)$  is defined by the relationship

$$\ell(x+1) = \ell(x)e^{-\mu(x)}, \quad (20)$$

so that

$$\mu(x) = -\log \frac{\ell(x+1)}{\ell(x)}. \quad (21)$$

That is, mortality rates are given by the slope (with the sign reversed) of the log of the survivorship function.

In the age × stage-classified model, using the vector  $\boldsymbol{\ell}(x)$  in Eq. 17, we obtain a vector of mortality rates as a function of birth stage as

$$\boldsymbol{\mu}(x) = -\log \left[ \mathcal{D}[\boldsymbol{\ell}(x)]^{-1} \boldsymbol{\ell}(x+1) \right] \quad s \times \mathbf{1} \quad (22)$$

where  $\boldsymbol{\ell}(x)$  is given by Eq. 17 and the log function is applied elementwise. The mortality schedule for birth stage  $i$ ,  $\mu_i(x)$  is undefined if stage  $i$  and age class 1 is an impossible combination.

The apparent mortality schedule (i.e., what would appear to an observer ignorant of the mixed composition of the cohort) for a mixed cohort with mixing distribution  $\boldsymbol{\pi}$ , is

$$\mu(x|\boldsymbol{\pi}_s) = -\log \left( \frac{\ell(x+1|\boldsymbol{\pi})}{\ell(x|\boldsymbol{\pi})} \right) \quad (23)$$

where  $\ell(x|\boldsymbol{\pi})$  is given by (19). This schedule is “apparent” because it results from the mixture of different initial stages; no individual actually experiences  $\mu(x|\boldsymbol{\pi})$ . Eqs. 22 and 23 provide mortality schedules resulting from the full age-stage dynamics;  $\boldsymbol{\mu}(x)$  and  $\mu(x|\boldsymbol{\pi})$  capture the full complexity of age- and stage-specific development. We will return briefly to the survivorship function, and present an alternative method of computation, in the section entitled *The fundamental matrix and longevity*.

*Intra-cohort selection.*—In an age-classified model, as a cohort ages, it shrinks as individuals die. In an age × stage-classified model, a cohort shrinks and its stage composition changes. Stages with higher mortality rates and/or shorter stage durations decrease in relative frequency due to intra-cohort selection.

To analyze intra-cohort selection, define an initial cohort  $\tilde{\mathbf{n}}_0$  and project it forward by

$$\tilde{\mathbf{n}}(x) = \tilde{\mathbf{U}}^x \tilde{\mathbf{n}}_0. \quad (24)$$

The projected cohort vector  $\tilde{\mathbf{n}}(x)$  gives the complete joint distribution of abundance by age and stage at age  $x$ , dependent on the initial cohort  $\tilde{\mathbf{n}}_0$ . Normalizing  $\tilde{\mathbf{n}}(x)$  to sum to 1 gives the joint age-stage frequency distribution,

$$\tilde{\mathbf{m}}(x|\tilde{\mathbf{n}}_0) = \frac{\tilde{\mathbf{n}}(x)}{\|\tilde{\mathbf{n}}(x)\|} \quad (25)$$

from which the marginal age and stage distributions can be calculated as:

$$\mathbf{m}_{\text{stage}}(x|\tilde{\mathbf{n}}_0) = (\mathbf{1}_\omega^\top \otimes \mathbf{I}_s) \tilde{\mathbf{m}}(x|\tilde{\mathbf{n}}_0) \quad (26)$$

$$\mathbf{m}_{\text{age}}(x|\tilde{\mathbf{n}}_0) = (\mathbf{I}_\omega \otimes \mathbf{1}_s^\top) \tilde{\mathbf{m}}(x|\tilde{\mathbf{n}}_0) \quad (27)$$

The general approach to the computation of such mixture and marginal quantities is outlined in Box 1. In general,  $\tilde{\mathbf{m}}(x)$  will converge, as  $x$  increases, to the right eigenvector corresponding to the dominant eigenvalue of  $\tilde{\mathbf{U}}$  (Horvitz and Tuljapurkar 2008).



The choice of  $\tilde{\mathbf{n}}_0$  depends on the question of interest. By definition, a cohort consists of individuals of the same age, which we can take as age class 1. Define  $\boldsymbol{\pi}_0$  as the distribution of stages within this age class. Then

$$\tilde{\mathbf{n}}_0 = (\mathbf{e}_1^\top \otimes \mathbf{I}_s) \boldsymbol{\pi}_0 \quad (28)$$

where  $\mathbf{e}_1$  is a vector of length  $\omega$  with a 1 in the first entry and zeros elsewhere. If the goal is to analyze a cohort all members of which start in stage  $j$ , then  $\boldsymbol{\pi}_0$  is a unit vector of length  $s$  with a 1 in the  $j$ th location and zeros elsewhere.

*The fertility function.*—In an age-classified model, the (scalar) fertility function  $f(x)$  gives the mean number of offspring produced, per unit time, by a parent aged  $x$ . In an age  $\times$  stage-classified model the fertility matrices  $\mathbf{F}_x$ , for  $x = 1, \dots, \omega$ , are the multivariate analogue of  $f(x)$ . The  $(i, j)$  entry of  $\mathbf{F}_x$  is the expected number of offspring of type  $i$  produced, per unit time, by a parent of type  $j$  and age class  $x$ .

The fertility matrix  $\mathbf{F}_x$  can be simplified to obtain three different age-specific fertility measures.

1. **Weighted offspring production.** Multiple offspring types, when they exist, are combined into a weighted sum, with weights defined by, e.g., body size, parental investment, reproductive value, etc. Let  $\mathbf{c}$  be such a vector of weights ( $s \times 1$ ). The vector of weighted fertility at age  $x$  is

$$\mathbf{f}_{\text{weighted}}(x) = (\mathbf{c}^\top \mathbf{F}_x)^\top \quad s \times 1. \quad (29)$$

The  $j$ th entry of  $\mathbf{f}_{\text{weighted}}(x)$  is the weighted mean number of offspring produced by a parent of stage  $j$  at age  $x$ .

2. **Mixed offspring production.** The vector giving the offspring, of all types, produced at age  $x$  by a mixture of stages given by a mixing distribution  $\boldsymbol{\pi}(x)$ , is

$$\mathbf{f}_{\text{mixed}}(x) = \mathbf{F}_x \boldsymbol{\pi}(x) \quad s \times 1. \quad (30)$$

One source for  $\boldsymbol{\pi}(x)$  is the stage structure of a cohort of individuals of age  $x$ . Given an initial cohort composition  $\tilde{\mathbf{n}}_0$  containing only individuals in age class 1, the appropriate mixture is

$$\boldsymbol{\pi}(x) = \mathbf{m}_{\text{stage}}(x | \tilde{\mathbf{n}}_0) \quad (31)$$

given by Eq. 26. The  $j$ th entry of  $\mathbf{f}_{\text{mixed}}(x)$  is the number of type  $j$  offspring produced per individual in a mixed cohort at age  $x$ , with a specified composition.

3. **Mixed and weighted offspring production.** The scalar fertility function, giving the weighted number of offspring produced by a mixed cohort at age  $x$ , is

$$f(x) = \mathbf{c}^\top \mathbf{F}_x \boldsymbol{\pi}(x) \quad 1 \times 1. \quad (32)$$

If only one type of offspring exists, and if stages do not differ in their fertility,  $f(x)$  reduces to the familiar age-specific fertility function.

The functions  $\mathbf{F}_x$ ,  $\mathbf{f}_{\text{weighted}}(x)$ ,  $\mathbf{f}_{\text{mixed}}(x)$ , and  $f(x)$  capture different aspects of the (possibly complicated) age- and stage-dependence of reproduction.

#### COHORT DYNAMICS: LONGEVITY STATISTICS

Longevity is the age at death of an individual. Statistics of longevity are frequently used to summarize mortality schedules. The most familiar such statistic is the life expectancy, or mean age at death. Measures of variation (the variance, standard deviation, coefficient of variation, and life discrepancy) provide information on stochastic variation in longevity among individuals. For example, many human populations have recently exhibited declines in variance in longevity, accompanying increases in life expectancy (e.g., Edwards and Tuljapurkar 2005, Tuljapurkar and Edwards 2011, Vaupel et al. 2011, Van Raalte and Caswell 2013, Engelman et al. 2014).

In age  $\times$  stage-classified models, longevity depends on how mortality varies over age and among stages, and we now turn to the calculation of these statistics.

#### The fundamental matrix and longevity

Longevity statistics are obtained by treating  $\tilde{\mathbf{U}}$  as the transient matrix of an age  $\times$  stage-classified absorbing Markov chain (Feichtinger 1971, Caswell 2001, 2006, 2009, 2013). If the states are numbered so that the transient (alive) states precede the absorbing (dead) states, the transition matrix for such a chain can be written

$$\tilde{\mathbf{P}} = \begin{pmatrix} \tilde{\mathbf{U}} & \mathbf{0} \\ \tilde{\mathbf{M}} & \mathbf{I} \end{pmatrix} \quad (33)$$

where  $\tilde{\mathbf{M}}$  is a matrix of transition probabilities from transient to absorbing states. We will consider  $\tilde{\mathbf{M}}$  in detail in the section entitled *Age and stage at death*.

The fundamental matrix for the chain is

$$\tilde{\mathbf{N}} = (\mathbf{I}_{s\omega} - \tilde{\mathbf{U}})^{-1}. \quad (34)$$

The matrix  $\tilde{\mathbf{N}}$  inherits the block structure of  $\tilde{\mathbf{U}}$ , with stages arranged within age classes. The  $(i, j)$  element of  $\tilde{\mathbf{N}}$  is the expected time spent in transient state  $i$  by an individual starting in transient state  $j$  (see Caswell 2006, 2009, 2013 for the higher moments and variances), where  $i$  and  $j$  range over all combinations of age and stage.

Longevity is measured by the time to eventual absorption in one of the states representing death. Well known results from Markov chain theory (Iosifescu 1980: Theorem 3.2) give the vectors of the mean and the second moments of time to absorption, which satisfy

$$\tilde{\boldsymbol{\eta}}_1^\top = \mathbf{1}_{s\omega}^\top \tilde{\mathbf{N}} \quad (35)$$

$$\tilde{\boldsymbol{\eta}}_2^\top = \tilde{\boldsymbol{\eta}}_1^\top (2\tilde{\mathbf{N}} - \mathbf{I}_{s\omega}). \quad (36)$$

The entries of  $\tilde{\boldsymbol{\eta}}_1$ , of dimension  $s\omega \times 1$ , give the life expectancy of individuals of each age-stage combination. The entries are arranged as in  $\tilde{\mathbf{n}}$ , with stages within age classes.

The vector  $\tilde{\eta}_2$  contains the second moments of longevity, so the vector of variances of longevity is

$$V(\tilde{\eta}) = \tilde{\eta}_2 - (\tilde{\eta}_1 \circ \tilde{\eta}_1). \quad (37)$$

This vector contains the variances in remaining longevity among individuals of every possible age-stage combination, reflecting stochasticity in advancement through age classes and transitions among stages. This variance can be decomposed into contributions within and between stages (cf. Hartemink et al. 2017), as we show in the next section.

*Partitioning the variance in longevity.*—In a cohort at age  $x$ , individuals will vary in longevity for two reasons: the stochasticity inherent in aging and transitions, and the heterogeneity inherent in the stage distribution at age  $x$ . To analyze the variance in remaining longevity at age  $x$ , define the vector  $\eta(x)$ , of dimension  $s \times 1$ , whose  $i$ th entry gives the remaining longevity of individuals of stage  $i$  at age  $x$ , and the scalar  $\eta(x)$  that gives the remaining longevity of an individual of age  $x$  in a mixed cohort. Extract the mean and variance of remaining longevity in age class  $x$  as

$$E[\eta(x)] = (\mathbf{e}_x^\top(\omega) \otimes \mathbf{I}_s)E(\tilde{\eta}) \quad (38)$$

$$V[\eta(x)] = (\mathbf{e}_x^\top(\omega) \otimes \mathbf{I}_s)V(\tilde{\eta}) \quad (39)$$

where  $\mathbf{e}_x(\omega)$  is a unit vector of length  $\omega$  with 1 in the  $x$ th position and zeros elsewhere,  $E(\cdot)$  is the mean and  $V(\cdot)$  is the variance.

Let  $\pi(x)$ , of dimension  $s \times 1$  be the mixing distribution of stages at age  $x$ . This could be obtained as the vector  $\mathbf{m}_{\text{stage}(x)}$  in Eq. 26. Then the mean longevity of age class  $x$ , treated as a mixture of stages, is

$$E_{\pi(x)}(\eta(x)) = \pi^\top(x)E[\eta(x)] \quad (40)$$

and the variance is partitioned into a component due to stochasticity within a given stage and a component due to heterogeneity among the stages,

$$V[\eta(x)] = V_{\text{within}}[\eta(x)] + V_{\text{between}}[\eta(x)] \quad (41)$$

where the within-group variance is the expectation over  $\pi(x)$  of the variances,

$$V_{\text{within}}[\eta(x)] = E_{\pi(x)}(V[\eta(x)]) \quad (42)$$

$$= \pi^\top(x)V[\eta(x)] \quad (43)$$

The between-group variance is the variance over  $\pi(x)$  of the means,

$$V_{\text{between}}[\eta(x)] = V_{\pi(x)}(E[\eta(x)]) \quad (44)$$

$$= \pi^\top(x)(E[\eta(x)] \circ E[\eta(x)] - (\pi^\top(x)E[\eta(x)])^2) \quad (45)$$

This variance decomposition is well-known in probability theory (Rényi 1970: Chapter 5.6, Theorem 1), forms the basis of the analysis of variance in statistics (e.g.

Kempthorne 1957), and is a familiar tool in the analysis of mixture models (Frühwirth-Schnatter 2006). It is a valuable tool in quantifying the relative contributions of heterogeneity and individual stochasticity to the variance in demographic outcomes (e.g. Caswell 2009, Edwards 2011, Hartemink et al. 2017, Jenouvrier et al. 2018, Hartemink and Caswell 2018). It is a natural tool for the analysis of age × stage-classified models, because the incorporation of stage-dependence is a natural way to include the effects of heterogeneity in age-specific parameters.

*1. Beyond longevity.*—The Expressions 43 and 45 require only the expressions for the mean and the variance of  $\eta$ , for each stage at each age of interest. Although we have described it in terms of variance in longevity, it can be applied to any quantity for which those quantities are available. For example, a stage-classified calculation has applied the analysis to age at first reproduction and lifetime reproductive output (Jenouvrier et al. 2018). The extension to age × stage-classified models will, in some cases, require an extension of occupancy time theory (Roth and Caswell 2018), in other cases the variance terms are directly available.

*2. Survivorship revisited.*—The fundamental matrix  $\tilde{\mathbf{N}}$  provides an alternative way to compute survivorship, dependent on the stage of the individual at birth (Keyfitz and Caswell 2005). Starting with  $\tilde{\mathbf{N}}$ , we first aggregate the rows by summing stages within age classes; the result gives the mean number of visits to each age class. But an age class can be visited at most once, so the mean number of visits equals the probability of ever visiting that age class, which is the survivorship. The column block corresponding to starting age class 1 contains the survivorship of each stage at that initial age; this is

$$\mathbf{L} = (\mathbf{I}_\omega \otimes \mathbf{1}_s^\top)\tilde{\mathbf{N}}(\mathbf{e}_1 \otimes \mathbf{I}_s), \quad \omega \times s \quad (46)$$

The matrix  $\mathbf{L}$ , of dimension  $\omega \times s$ , contains as columns the survivorship functions for each stage. The survivorship for a mixture of initial stages, specified by a mixing distribution vector  $\pi_s$ , is given by

$$\begin{aligned} \ell(x|\pi) &= \mathbf{L}\pi_s \\ &= (\mathbf{I}_\omega \otimes \mathbf{1}_s^\top)\tilde{\mathbf{N}}(\mathbf{e}_1 \otimes \pi) \end{aligned} \quad (47)$$

#### Age and stage at death

In classical age-structured life tables, the probability distribution of age at death is  $b(x) = \ell(x) - \ell(x+1)$ . In age × stage-classified analysis, there is a joint distribution of age and stage at death, for an individual of any initial age-stage combination. To obtain this joint distribution, we construct a mortality matrix  $\tilde{\mathbf{M}}$  in the absorbing Markov chain (Eq. 33) such that the absorbing states are defined by the combination of age and stage at death. To do so, define  $\mathbf{q}_i$  as the vector of stage-specific mortality probabilities for age class  $i$ , obtained from  $\mathbf{U}_i$  as

$$\mathbf{q}_i^\top = \mathbf{1}_s^\top - \mathbf{1}_s^\top \mathbf{U}_i. \quad (48)$$

Then the mortality matrix  $\tilde{\mathbf{M}}$  is

$$\tilde{\mathbf{M}} = \sum_{i=1}^{\omega} \mathbf{E}_{ii} \otimes \mathcal{D}(\mathbf{q}_i), \quad s\omega \times s\omega \quad (49)$$

inheriting the block structure of  $\tilde{\mathbf{U}}$ .

As a simple extension of the result for age- or stage-classified models (Caswell 2001, 2009, 2012) all the joint distributions of age and stage at death are contained in the matrix

$$\tilde{\mathbf{B}} = \tilde{\mathbf{M}}\tilde{\mathbf{N}} \quad s\omega \times s\omega. \quad (50)$$

Each column of  $\tilde{\mathbf{B}}$  contains the joint distribution of age and stage at death for an individual in one of the possible age-stage categories.

Marginalizing by summing stages within age classes, or age classes within stages, gives the marginal age and stage distributions as the columns of the matrices

$$\mathbf{B}_{\text{age}} = (\mathbf{I}_{\omega} \otimes \mathbf{1}_s^{\top})\tilde{\mathbf{B}} \quad \omega \times s\omega \quad (51)$$

$$\mathbf{B}_{\text{stage}} = (\mathbf{1}_{\omega}^{\top} \otimes \mathbf{I}_s)\tilde{\mathbf{B}} \quad s \times s\omega. \quad (52)$$

The distributions resulting from a mixed cohort, with mixture distribution  $\pi_{s\omega}$  are

$$\tilde{\mathbf{b}}(\pi) = \tilde{\mathbf{B}}\pi_{s\omega} \quad s\omega \times 1 \quad (53)$$

$$\mathbf{b}_{\text{age}}(\pi) = \mathbf{B}_{\text{age}}\pi_{s\omega} \quad \omega \times 1 \quad (54)$$

$$\mathbf{b}_{\text{stage}}(\pi) = \mathbf{B}_{\text{stage}}\pi_{s\omega} \quad s \times 1. \quad (55)$$

As this collection of results makes clear, the combination of age and stage dependence of all the vital rates creates rich opportunities to examine the distributions of age and stage at death.

#### *Life disparity: years of life lost*

An important application of  $\tilde{\mathbf{B}}$  is to the calculation of the life disparity, or the mean number of life years lost to mortality, an index introduced by Vaupel and Canudas Romo (2003). In an age-classified model, an individual that dies at age  $x$ , which happens with probability  $b(x)$ , loses some remaining years of life that would have been available had the individual not, in fact, died. The lost years of life are unknown, but their expectation is the remaining life expectancy at age  $x$ . Thus the expectation of the life lost due to mortality is obtained by integrating over all ages at death

$$e^{\dagger} = \int_0^{\infty} e(x)b(x)dx \quad (56)$$

where  $e(x)$  is remaining life expectancy at age  $x$ .

The mean life lost plays another role: it is a measure of the variation in age at death called *life disparity*. If all individuals were to die at the same age, say  $x^*$ , then  $b(x)$  would be a delta function at  $x^*$  and  $e(x^*)$  would be 0, and (Eq. 56) would give  $e^{\dagger} = 0$ . As a measure of variability in longevity,  $e^{\dagger}$  is highly correlated with the standard deviation of

longevity and other measures of variation. Vaupel et al. (2011) analyzed patterns of  $e^{\dagger}$  across countries in relation to life expectancy as a way to analyze premature death.

In an age  $\times$  stage-classified model, we define a vector

$$\tilde{\boldsymbol{\eta}}^{\dagger} = (\tilde{\boldsymbol{\eta}}_1^{\top} \tilde{\mathbf{B}})^{\top} \quad s\omega \times 1. \quad (57)$$

The vector  $\tilde{\boldsymbol{\eta}}_1$  contains the expected remaining longevity of every age-stage combination; the columns of  $\tilde{\mathbf{B}}$  contain the probabilities of death at each age-stage combination. Thus the vector  $\tilde{\boldsymbol{\eta}}^{\dagger}$  contains the expected life lost due to mortality, for individuals of every age and stage combination. These age  $\times$  stage-specific expectations can be combined according to a mixing distribution  $\boldsymbol{\pi}$  (dimension  $s\omega \times 1$ ) as

$$\boldsymbol{\eta}^{\dagger}(\boldsymbol{\pi}) = \boldsymbol{\pi}^{\top} \tilde{\boldsymbol{\eta}}^{\dagger}. \quad (58)$$

The result  $\boldsymbol{\eta}^{\dagger}(\boldsymbol{\pi})$  is a scalar, the mean life lost due to mortality in a cohort composed of age-stage combinations in proportions given by  $\boldsymbol{\pi}$ .

#### POPULATION DYNAMICS: GROWTH AND STRUCTURE

We turn now from probabilistic outcomes for individuals within cohorts to the dynamics of populations. Population dynamics are the outcome of survival, development, and fertility throughout the life cycle. The first two are accounted for by  $\tilde{\mathbf{U}}$ ; the third is accounted for by  $\tilde{\mathbf{F}}$ . The basic projection of population dynamics is

$$\tilde{\mathbf{n}}(t+1) = (\tilde{\mathbf{U}} + \tilde{\mathbf{F}})\tilde{\mathbf{n}}(t) \quad (59)$$

$$= \tilde{\mathbf{A}}\tilde{\mathbf{n}}(t) \quad (60)$$

with specified initial population  $\tilde{\mathbf{n}}(0) = \tilde{\mathbf{n}}_0$ .

#### *Stable population theory and demographic ergodicity*

If  $\tilde{\mathbf{A}}$  is time-invariant, we expect the population to converge, from any non-negative and non-zero initial population, to exponential growth at a rate  $\lambda$  given by the dominant eigenvalue of  $\tilde{\mathbf{A}}$ , and a structure proportional to the corresponding right eigenvector  $\tilde{\mathbf{w}}$ . The reproductive value vector will be given by the corresponding left eigenvector  $\tilde{\mathbf{v}}$  of  $\tilde{\mathbf{A}}$ .

When normalized to sum to 1, the vector  $\tilde{\mathbf{w}}$  gives the joint age  $\times$  stage distribution in the stable population. The marginal stable age and stage distributions are given by

$$\mathbf{w}_{\text{age}} = (\mathbf{I}_{\omega} \otimes \mathbf{1}_s^{\top})\tilde{\mathbf{w}} \quad (61)$$

$$\mathbf{w}_{\text{stage}} = (\mathbf{1}_{\omega}^{\top} \otimes \mathbf{I}_s)\tilde{\mathbf{w}}. \quad (62)$$

The reproductive value vector  $\tilde{\mathbf{v}}$  is not a probability distribution; instead, the entries of  $\tilde{\mathbf{v}}$  give the asymptotic relative sizes of populations initialized with a single individual of the corresponding age-stage combination. The reproductive value of a mixture of ages and stages is

$$v(\boldsymbol{\pi}) = \boldsymbol{\pi}^{\top} \tilde{\mathbf{v}} \quad (63)$$

where  $\boldsymbol{\pi}$  is the mixing distribution.

*Ergodicity.*—Our expectation of ergodic behavior (the convergence to exponential growth and a stable structure from any initial condition) may be disappointed, because  $\tilde{\mathbf{A}}$  may not be irreducible (a matrix is irreducible if there exists a pathway from every state to every other state). The Perron-Frobenius theorem guarantees that the state space of a reducible matrix can be decomposed into subspaces, each of which leads to different asymptotic behavior (see Csetenyi and Logofet 1989, Caswell 2001: Section 4.5, Stott et al. 2010, Caswell 2012: Appendix B). For example, life cycles with post-reproductive age classes are reducible and the behavior of such a population depends on whether the initial population contains only post-reproductive individuals (eventual extinction) or contains some reproductive individuals (potential population growth).

The  $i$ -state space of an age × stage-classified model will certainly be reducible if it contains some impossible combinations, which appear as rows and columns of zeros in  $\tilde{\mathbf{A}}$  (see Eqs. 8 and 9). These states cannot be reached from, nor do they lead to, any other states; hence  $\tilde{\mathbf{A}}$  is reducible. However, if we make the eminently reasonable decision not to start with a population composed solely of impossible individuals, these states will have no effect on eventual dynamics.

It is more challenging to deal with reducibility generated by the complicated pathways through the life cycle produced by the interaction of age-specific and stage-specific processes. Stochastic ergodic theorems may provide some guidance (Cohen 1982). Ergodicity can, in any specific case, be determined using the patterns of zero and non-zero entries of the reproductive value vector  $\tilde{\mathbf{v}}$  (Caswell 2012). If  $\tilde{\mathbf{v}}$  has positive entries for all non-impossible states, then population dynamics converge to  $\lambda$  and  $\tilde{\mathbf{w}}$  from any initial condition not restricted to impossible states. If  $\tilde{\mathbf{v}}$  has zero entries for some non-impossible states, then an initial condition restricted to those states would not converge to  $\tilde{\mathbf{w}}$ , but rather to the distribution given by a different eigenvector. This possibility can easily be checked for any specified matrix.

*Net reproductive rate.*—The familiar net reproductive rate in age-classified models,

$$R_0 = \int_0^\infty \ell(x)m(x)dx \tag{64}$$

serves three functions (Caswell 2009): it is the mean number of (usually female) offspring produced over a lifetime, it is the per-generation growth rate of the population, and it is an indicator function that distinguishes population growth ( $R_0 > 1$ ) from population decline ( $R_0 < 1$ ). Cushing and Zhou (1994) generalized this concept to stage-classified models by showing that

$$R_0 = \max \text{eig}(\mathbf{R}) \tag{65}$$

where  $\mathbf{R}$  is the next-generation matrix,

$$\mathbf{R} = \mathbf{F}\mathbf{N}. \tag{66}$$

If  $\mathbf{F}$  contains only a single type of offspring, then Eq. 65 satisfies all three functions of the net reproductive rate. But

if multiple types of offspring exist, then  $R_0$  defined by Eq. 65 satisfies the last two (per-generation growth rate and growth indicator function), but it is not the mean lifetime number of offspring. Indeed, if multiple types of offspring exist, there is no single “number of offspring” to calculate.

The Cushing-Zhou result (Eq. 65) applies equally to age × stage-classified models, so

$$R_0 = \max \text{eig}(\tilde{\mathbf{R}}) \tag{67}$$

where  $\tilde{\mathbf{R}} = \tilde{\mathbf{F}}\tilde{\mathbf{N}}$ . However, taking advantage of the structure of the population vector, we can carry the analysis further. Let  $\mathbf{F}_{ij}$  and  $\mathbf{N}_{ij}$  denote the age blocks in  $\tilde{\mathbf{F}}$  and  $\tilde{\mathbf{N}}$ , respectively. For example, for  $\omega = 3$ ,

$$\tilde{\mathbf{F}} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} & \mathbf{F}_{13} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \tilde{\mathbf{N}} = \begin{pmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} & \mathbf{N}_{13} \\ \mathbf{N}_{21} & \mathbf{N}_{22} & \mathbf{N}_{23} \\ \mathbf{N}_{31} & \mathbf{N}_{32} & \mathbf{N}_{33} \end{pmatrix}. \tag{68}$$

Then

$$\tilde{\mathbf{R}} = \begin{pmatrix} \sum_{i=1}^\omega \mathbf{F}_{1i}\mathbf{N}_{i1} & \sum_{i=1}^\omega \mathbf{F}_{1i}\mathbf{N}_{i2} & \sum_{i=1}^\omega \mathbf{F}_{1i}\mathbf{N}_{i3} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{69}$$

$$= \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{70}$$

The matrix  $\tilde{\mathbf{R}}$  is block upper-triangular, and hence its dominant eigenvalue is

$$R_0 = \max \text{eig}(\mathbf{R}_{11}). \tag{71}$$

The matrix  $\mathbf{R}_{11}$  is of dimension  $s \times s$ ; its  $(i,j)$  entry is

$$\mathbf{R}_{11}(i,j) = E(\text{lifetime stage } i \text{ offspring} | \text{starting in stage } j). \tag{72}$$

It can be extracted easily from  $\tilde{\mathbf{R}}$  as

$$\mathbf{R}_{11} = \mathbf{Z} \tilde{\mathbf{R}} \mathbf{Z}^T \tag{73}$$

where

$$\mathbf{Z} = (\mathbf{I}_s | \mathbf{0}_{s \times (\omega-1)s}) \tag{74}$$

The remaining lifetime reproductive output at age class  $j$  is obtained by applying the following analyses to the block  $\mathbf{R}_{1j}$ .

*Mean lifetime reproduction.*—As was the case for stage-classified models, if there exists only a single stage of offspring, then  $R_0$  calculated from the Cushing-Zhou theorem also gives the mean lifetime reproduction. However, when there are multiple offspring stages, the matrix  $\mathbf{R}_{11}$  contains information on the production of all these types of offspring. Just as we did for the age-specific fertility function in Eqs. 29–32,

we can aggregate lifetime reproductive output in several ways.

1. Weighted mean lifetime reproduction. Let  $\mathbf{c}_s$  be a vector of weights, of dimension  $s \times 1$ . Then

$$\mathbf{r}_{\text{weighted}} = (\mathbf{c}^\top \mathbf{R}_{11})^\top \quad (75)$$

The vector  $\mathbf{r}_{\text{weighted}}$  is a weighted combination of the rows of  $\mathbf{R}_{11}$ , and  $\mathbf{c}$ , a vector whose entries assign weights to the different types of offspring. The  $i$ th entry of  $\mathbf{r}_{\text{weighted}}$  gives the mean weighted lifetime reproduction by an individual starting life in stage  $i$ .

2. Lifetime reproduction of a mixed cohort. A cohort starting as a mixture specified by a mixing distribution  $\boldsymbol{\pi}$  will have a lifetime reproductive output of

$$\mathbf{r}_{\text{mixed}} = \mathbf{R}_{11} \boldsymbol{\pi} \quad (76)$$

This vector is a weighted combination of the columns of  $\mathbf{R}_{11}$ ; its  $i$ th entry is the mean lifetime production of type  $i$  offspring by the mixed cohort.

3. Mixed and weighted lifetime reproduction. A scalar measure of mean lifetime reproduction is given by

$$\rho = \mathbf{c}^\top \mathbf{R}_{11} \boldsymbol{\pi}. \quad (77)$$

In the special case where only a single type of offspring is produced,  $\rho = R_0$ . But in general, lifetime offspring production in age-stage models is a more diverse and nuanced concept than  $R_0$ . Note that much more information about lifetime reproductive output, including variances, higher moments, and sensitivity analysis can be obtained using Markov chains with rewards (Caswell 2011, van Daalen and Caswell 2015, 2017). The application of these methods to age  $\times$  stage-classified models will be explored elsewhere.

#### POPULATION DYNAMICS: GENERATION TIME

Generation time, which is defined in several ways, is an important demographic measure of the time scale on which a population operates (Gaillard et al. 2005, Lebreton 2005, Steiner et al. 2014). In conservation standards established by the International Union for the Conservation of Nature, population decline measured in units of generation time helps establish the threat level for a species.

Three measures of generation time are in common use in age-classified demography (Coale 1972, Caswell 2001): the growth rate generation time, which is the time required to grow by a factor of  $R_0$ ,

$$T = \frac{\log R_0}{\log \lambda} \quad (78)$$

the cohort generation time, which is the mean age of the production of offspring by an individual,

$$\Gamma = \frac{\int x m(x) \ell(x) dx}{\int m(x) \ell(x) dx} \quad (79)$$

and the stable age generation time, which is the mean age of the parents of the offspring produced in a population at the stable age distribution  $c(x)$ ,

$$A = \frac{\int x m(x) c(x) dx}{\int m(x) c(x) dx}. \quad (80)$$

Of these three, the cohort generation time is the most appropriate as a measure of the time scale on which a life history operates, because it is explicitly calculated as a property of an individual over its lifetime. The cohort generation time is denoted as  $\mu$  by Coale (1972), but that symbol is also used for the mortality rate. It is not unusual for such notational conflicts to arise when discussing demographic calculations that span a variety of indices or processes.

Three complications arise in extending the familiar calculations for age-classified demography to age  $\times$  stage models: (1) Multiple kinds of offspring may be produced, following different age schedules, and thus each appearing to have a different generation time; (2) Individuals that start life in different offspring stages may survive, mature, and reproduce differently, and hence have different generation times; (3) Unlike age-classified models, in which a cohort is always of a fixed age, a cohort may start life as a mixture of stages.

The cohort generation time for stage-classified models was derived in (Caswell 2009: Appendix A.5). What follows is the extension to age  $\times$  stage-classified models, in which both age and stage trajectories must be taken into account.

Begin with a cohort  $\tilde{\mathbf{n}}_0$  (dimension  $s\omega \times 1$ ); being a cohort, it contains only age class 1 individuals, but may have any initial stage distribution. Normalize the cohort so that  $\|\tilde{\mathbf{n}}_0\| = 1$ . The survivors of this cohort at age  $x$  are

$$\tilde{\mathbf{n}}(x) = \tilde{\mathbf{U}}^x \tilde{\mathbf{n}}_0 \quad (81)$$

and the offspring, in the first age class, produced by these survivors are

$$\boldsymbol{\phi}(x) = (\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{F}} \tilde{\mathbf{U}}^x \tilde{\mathbf{n}}_0 \quad s \times 1 \quad (82)$$

where  $\mathbf{e}_1$  is a unit vector of length  $\omega$ . Summing this quantity over all ages gives the vector of total lifetime reproduction,

$$\boldsymbol{\phi}_{\text{life}} = \sum_{x=0}^{\infty} \boldsymbol{\phi}(x) \quad (83)$$

$$= (\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{F}} \sum_{x=0}^{\infty} \tilde{\mathbf{U}}^x \tilde{\mathbf{n}}_0 \quad (84)$$

$$= (\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{F}} \tilde{\mathbf{N}} \tilde{\mathbf{n}}_0. \quad (85)$$

The mean age of the production of the offspring over the lifetime is proportional to

$$\sum_{x=0}^{\infty} x \boldsymbol{\phi}_{\text{life}}(x) = (\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{F}} \sum_{x=0}^{\infty} x \tilde{\mathbf{U}}^x \tilde{\mathbf{n}}_0. \quad (86)$$

But

$$\sum_{x=0}^{\infty} x \tilde{\mathbf{U}}^x = \tilde{\mathbf{N}} \tilde{\mathbf{U}} \tilde{\mathbf{N}}. \quad (87)$$

The proportionality constant is required to make the entries of  $\Phi_{\text{lifc}}$  sum to 1 as probability distributions; with this normalization the vector of cohort generation times, for each starting stage of offspring, is

$$\Gamma = \mathcal{D}(\Phi_{\text{lifc}})^{-1}(\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{F}} \tilde{\mathbf{N}} \tilde{\mathbf{U}} \tilde{\mathbf{n}}_0 \quad s \times 1. \quad (88)$$

The  $i$ th entry of  $\Gamma$  is the mean age of production of offspring of type  $i$  by a cohort of individuals with initial stage distribution  $\tilde{\mathbf{n}}_0$ , and accounting for all stage transitions that occur as the individuals age. Entries of  $\Gamma$  corresponding to stages that never appear as offspring will be undefined (0/0) in Eq. 88; they should be set equal to 0.

SENSITIVITY ANALYSIS OF AGE × STAGE-CLASSIFIED MODELS

In this paper, each step in the construction and analysis of the age × stage-classified model has been written in terms of matrix operations. This is not only for notational, analytical, and computational efficiency. It also makes possible the systematic calculation of the sensitivity of any output to changes in any set of parameters, using matrix calculus. Some previous studies have applied matrix calculus to vec-permutation models that are similar or equivalent to age × stage-classified models, including stage-classified epidemics (Klepac and Caswell 2011), spatial models (Strasser et al. 2012), the sensitivity of population growth rate in age-stage models (Caswell 2012, Caswell and Shyu 2012, Caswell and Salguero-Gómez 2013), and the effects of age and frailty (Roth and Caswell 2016). These results can be extended to a general sensitivity analysis of age × stage-classified models; details will be presented elsewhere.

Let  $\xi$  be any output variable (see Table 1), scalar- or vector-valued, calculated from  $\tilde{\mathbf{A}}$ , and let  $\theta$  be a vector of parameters that affect any of the  $\mathbf{U}_i$  and/or  $\mathbf{F}_i$ . In matrix calculus notation, the sensitivity of  $\xi$  to  $\theta$  is

$$\begin{aligned} \frac{d\xi}{d\theta^\top} &= \underbrace{\left( \frac{d\xi}{d\text{vec}^\top \tilde{\mathbf{A}}} \right)}_1 \left[ \underbrace{\left( \frac{d\text{vec} \tilde{\mathbf{U}}}{d\text{vec}^\top \tilde{\mathbf{U}}} \right)}_2 \right. \\ &\times \sum_{i=1}^{\omega} \underbrace{\left( \frac{d\text{vec} \mathbf{U}_i}{d\text{vec}^\top \mathbf{U}_i} \right)}_3 \underbrace{\left( \frac{d\text{vec} \mathbf{U}_i}{d\theta^\top} \right)}_4 + \underbrace{\left( \frac{d\text{vec} \tilde{\mathbf{F}}}{d\text{vec}^\top \tilde{\mathbf{F}}} \right)}_2 \\ &\times \sum_{i=1}^{\omega} \underbrace{\left( \frac{d\text{vec} \mathbf{F}_i}{d\text{vec}^\top \mathbf{F}_i} \right)}_3 \underbrace{\left( \frac{d\text{vec} \mathbf{F}_i}{d\theta^\top} \right)}_5 \left. \right] \quad (89) \end{aligned}$$

Each of the numbered terms in (Eq. 89) has its own function.

Term 1

Expresses the dependence of the outcome variable  $\xi$  on the matrix  $\tilde{\mathbf{A}}$  or its components  $\tilde{\mathbf{F}}$  and  $\tilde{\mathbf{U}}$ . this matrix is obtained by differentiating  $\xi$  with respect to  $\tilde{\mathbf{U}}$  and  $\tilde{\mathbf{F}}$ .

For example, if  $\xi = \lambda$ , then Term 1 is the familiar derivative of  $\lambda$  to the entries of the matrix  $\tilde{\mathbf{A}}$ ,

$$\frac{d\lambda}{d\text{vec}^\top \tilde{\mathbf{A}}} = \mathbf{w}^\top \otimes \mathbf{v}^\top. \quad (90)$$

But suppose instead that interest focuses on the net reproductive rate, as defined by Eq. 67, and let  $\mathbf{y}$  and  $\mathbf{x}$  be the right and left eigenvectors of  $\tilde{\mathbf{R}}$  corresponding to  $R_0$ , scaled so that  $\mathbf{y}^\top \mathbf{x} = 1$ . Then  $\xi = R_0$ , and Term 1 is the derivative of the net reproductive rate with respect to  $\tilde{\mathbf{A}}$ , which is

$$\frac{dR_0}{d\text{vec}^\top \tilde{\mathbf{A}}} = \frac{dR_0}{d\text{vec}^\top \tilde{\mathbf{R}}} \frac{d\text{vec} \tilde{\mathbf{R}}}{d\text{vec}^\top \tilde{\mathbf{A}}} \quad (91)$$

$$= \frac{dR_0}{d\text{vec}^\top \tilde{\mathbf{R}}} \left( \frac{d\text{vec} \tilde{\mathbf{R}}}{d\text{vec}^\top \tilde{\mathbf{F}}} + \frac{d\text{vec} \tilde{\mathbf{R}}}{d\text{vec} \tilde{\mathbf{N}}} \frac{d\text{vec} \tilde{\mathbf{N}}}{d\text{vec}^\top \tilde{\mathbf{U}}} \right) \quad (92)$$

$$= (\mathbf{y}^\top \otimes \mathbf{x}^\top) \{ (\tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega}) + (\mathbf{I}_{s\omega} \otimes \tilde{\mathbf{F}}) (\tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{N}}) \} \quad (93)$$

$$= (\mathbf{y}^\top \otimes \mathbf{x}^\top) \{ (\tilde{\mathbf{N}}^\top \otimes \mathbf{I}_{s\omega}) + (\tilde{\mathbf{N}}^\top \otimes \tilde{\mathbf{R}}) \} \quad (94)$$

(extending calculations from Caswell [2009] to the age × stage case).

The point of these two examples is that Term 1 is the only part of the sensitivity calculation that depends on what the dependent variable may be ( $\lambda$ ,  $R_0$ , or anything else).

Term 2

These are constant matrices, depending only on the vec-permutation matrix  $\mathbf{K}$  the age transition matrix  $\mathbb{D}$  and the age assignment matrix  $\mathbb{H}$ . Since  $\tilde{\mathbf{U}}$  and  $\tilde{\mathbf{F}}$  are given by Eqs. 11 and 12, the two derivatives in Term 2 are

$$\frac{d\text{vec} \tilde{\mathbf{U}}}{d\text{vec}^\top \tilde{\mathbf{U}}} = (\mathbf{I}_{s\omega} \otimes \mathbf{K}^\top \mathbb{D} \mathbf{K}) \quad (95)$$

$$\frac{d\text{vec} \tilde{\mathbf{F}}}{d\text{vec}^\top \tilde{\mathbf{F}}} = (\mathbf{I}_{s\omega} \otimes \mathbf{K}^\top \mathbb{H} \mathbf{K}). \quad (96)$$

Notice that these matrices are independent of the dependent variable or on the identity of the parameters  $\theta$ .

Term 3

These are the derivatives of the block diagonal matrices  $\mathbb{U}$  and  $\mathbb{F}$  to their diagonal entries  $\mathbf{U}_i$  and  $\mathbf{F}_i$ . As such they depend only on the arrangement of the matrices on the diagonal. A new way to obtain these derivatives is based on Eqs. 14 and 15 of Caswell and van Daalen (2016). Define the matrices  $\mathbf{P}_i$  and  $\mathbf{Q}_i$ , of dimension  $\omega s \times s$  and  $s \times \omega s$ , respectively,

$$\mathbf{P}_i = \begin{pmatrix} \mathbf{0}_{s(i-1) \times s} \\ \mathbf{I}_s \\ \mathbf{0}_{s(\omega-i) \times s} \end{pmatrix} \quad \mathbf{Q}_i = (\mathbf{0}_{s \times (i-1)s} \quad \mathbf{I}_s \quad \mathbf{0}_{s \times (\omega-i)s}). \quad (97)$$

Then

$$\frac{d\text{vec}\mathbf{U}}{d\text{vec}\mathbf{U}_i} = \frac{d\text{vec}\mathbf{F}}{d\text{vec}\mathbf{F}_i} = \mathbf{Q}_i^\top \otimes \mathbf{P}_i. \quad (98)$$

Again, note that this term is independent of the dependent variable and the parameters being perturbed.

#### Terms 4 and 5

These terms contain the biologically interesting part of the calculation, because it is at this point that we find the dependence of the transition and survival matrices  $\mathbf{U}_i$  and the fertility matrices  $\mathbf{F}_i$  on the parameter vector  $\boldsymbol{\theta}$ . This is the point where the age-stage specific processes of survival, growth, development, reproduction, etc. are determined.

The possibilities are limited only by imagination. As an example, suppose that interest focuses on a parameter that imposes a multiplicative perturbation on the stage-specific survival probabilities at a specified age or ages. We can write

$$\mathbf{U}_i = \mathbf{G}_i \boldsymbol{\Sigma}_i \quad (99)$$

where  $\boldsymbol{\Sigma}_i$  is a diagonal matrix with a vector of survival probabilities  $\boldsymbol{\sigma}$  on the diagonal and  $\mathbf{G}_i$  is a matrix of transition probabilities conditional on survival. From this expression, it can be shown that

$$\frac{d\text{vec}\mathbf{U}_i}{d\boldsymbol{\sigma}_i^\top} = (\mathbf{I}_s \otimes \mathbf{G}_i) \mathcal{D}(\text{vec } \mathbf{I}_s) (\mathbf{I}_s \otimes \mathbf{I}_s). \quad (100)$$

Now incorporate the parameter vector by writing  $\boldsymbol{\sigma}_i = \hat{\boldsymbol{\sigma}}_i \circ \boldsymbol{\theta}$ , where  $\hat{\boldsymbol{\sigma}}_i$  is the baseline vector of stage-specific survival probabilities at age class  $i$ . Then

$$\frac{d\text{vec}\mathbf{U}_i}{d\boldsymbol{\theta}^\top} = \frac{d\text{vec}\mathbf{U}_i}{d\boldsymbol{\sigma}_i^\top} \frac{d\boldsymbol{\sigma}_i}{d\boldsymbol{\theta}^\top} \quad (101)$$

$$= \frac{d\text{vec}\mathbf{U}_i}{d\boldsymbol{\sigma}_i^\top} \mathcal{D}(\hat{\boldsymbol{\sigma}}_i). \quad (102)$$

Because this perturbation is hypothesized to affect only survival, the derivatives of  $\mathbf{F}_i$  to  $\boldsymbol{\theta}$  in Eq. 89 are all zero.

This calculation is an arbitrary example. The reader can no doubt think up possible applications (e.g.,  $\boldsymbol{\theta}$  might describe harvest or bycatch impacts that reduce the survival probability of certain stages of a threatened species, or it might reflect allocation of resources to survival, with an associated cost in fertility, ...).

The causal pathways from a parameter vector  $\boldsymbol{\theta}$  to some demographic outcome  $\xi$  can be enormously complex in an age  $\times$  stage-classified model. The great value of the vec-permutation formulation is that all these pathways are taken into account, and that attention can focus on the effects of parameters on the age  $\times$  stage-specific vital rates in terms 4 and 5, and on the way that the outcome is computed from the age  $\times$  stage matrix  $\tilde{\mathbf{A}}$  in term 1. The connection between these two, for any outcome for any

species described by any set of stages, is given by the series of constant, and easily computable, matrices in terms 2 and 3.

#### A MODEL SPECIES EXAMPLE

Here, we construct and analyze an age  $\times$  stage-classified matrix model for a hypothetical model species, inspired by (although not identical to) poecilogonous marine invertebrates (Levin et al. 1987, Krug 2007, Knott and McHugh 2012). These are species that produce two dramatically different types of larval offspring. One type, so-called planktrophic larvae, feed during the larval stage and receive only a small parental investment. The other type, lecithotrophic larvae, do not feed, and rely on a large parental investment in the form of a yolk reserve. Lecithotrophic larvae are more costly to produce and are produced in lower numbers than are planktrophic larvae. However, lecithotrophic larvae have a higher survival probability than their planktrophic siblings, an advantage that persists into post-larval life (Levin et al. 1987).

Our model species has four stages: large and small juveniles, and adults that began life as either large or small juveniles. We will refer to these as large and small adults, although large-born and small-born might be more accurate. MATLAB code for the calculations is available as Supporting information.

#### The age $\times$ stage-classified model

The life cycle of our model species contains small and large juveniles and small and large adults. Stage-specific demography is defined in terms of survival probability  $\sigma$ , maturation probability (conditional on survival)  $\gamma$ , and fertility  $f$ . Each of these parameters is a function of both stage and age. The resulting life cycle graph is shown in Fig. 1; the model parameters are described in Table 3. The life cycle graph describes the stages; age dependence enters by making the survival, maturation, and fertility parameters in the graph into functions of age as well as stage.

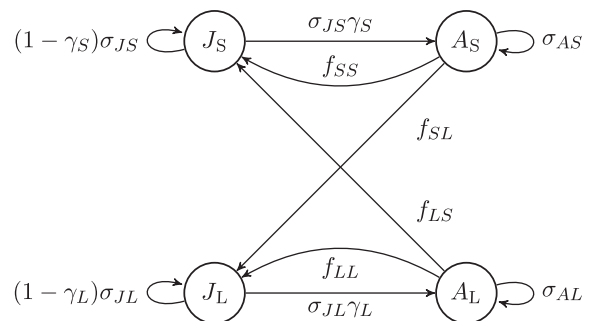


FIG. 1. Life cycle graph for the example species. The nodes represent the four stages: small juvenile ( $J_S$ ), large juvenile ( $J_L$ ), small adult ( $A_S$ ) and large adult ( $A_L$ ). The arrows represent survival, maturation, and reproduction. The per-time-step probabilities of survival ( $\sigma_{JS}$ ,  $\sigma_{JL}$ ,  $\sigma_{AS}$ ,  $\sigma_{AL}$ ) and maturation ( $\gamma_S$ ,  $\gamma_L$ ) and the per-time-step production of new individuals ( $f_{SS}$ ,  $f_{SL}$ ,  $f_{LS}$ ,  $f_{LL}$ ) all depend on age,  $x$ .

TABLE 3. Model parameters used in the life cycle graph in Fig. 1.

Notation	Description
Per-time step survival probability	
$\sigma_{JS}(j)$	<i>small juvenile</i>
$\sigma_{JL}(j)$	<i>large juvenile</i>
$\sigma_{AS}(j)$	<i>small adult</i>
$\sigma_{AL}(j)$	<i>large adult</i>
Per-time step maturation probability	
$\gamma_S(j)$	<i>small juvenile</i>
$\gamma_L(j)$	<i>large juvenile</i>
Per-time step production	
$f_{SS}(j)$	<i>of small juveniles by small adults</i>
$f_{SL}(j)$	<i>of large juveniles by small adults</i>
$f_{LS}(j)$	<i>of small juveniles by large adults</i>
$f_{LL}(j)$	<i>of large juveniles by large adults</i>

Note: The index  $j$  refers to the age class of the individual; for reproduction it refers to the age class of the parents.

*Life history description.*—Because the two types of offspring differ in the resources allocated to them by the parent, offspring type has consequences throughout the life cycle. Both juvenile and adult survival probability are lower for small individuals (Fig. 2a and b). Small juveniles also begin to mature later than large ones and their maximum maturation probability is lower (Fig. 2c).

Small individuals are less costly to produce, so they are produced at a higher rate, by both small and large adults. We have incorporated a certain heritability of type, which results in large adults producing relatively more large juveniles and small adults producing relatively more small juveniles (Fig. 2d and e).

*Matrix construction*

We construct the age × stage model from the stage transition matrices  $\mathbf{U}_x$ , the reproduction matrices  $\mathbf{F}_x$ , the age transition matrix  $\mathbf{D}$  and the age assignment matrix  $\mathbf{H}$ . The life cycle contains  $s = 4$  stages (stage 1, small juveniles; 2, large juveniles; 3, small adults; and 4, large adults), and we consider  $\omega = 50$  age classes.

The matrix  $\mathbf{U}_x$  describes the transition and survival probabilities for individuals in age class  $x$ ,

$$\mathbf{U}_x = \begin{pmatrix} \sigma_{SJ}(1-\gamma_S) & 0 & 0 & 0 \\ 0 & \sigma_{LJ}(1-\gamma_L) & 0 & 0 \\ \sigma_{SJ}\gamma_S & 0 & \sigma_{SA} & 0 \\ 0 & \sigma_{LJ}\gamma_L & 0 & \sigma_{LA} \end{pmatrix} (x), \quad x = 1, \dots, 50. \tag{103}$$

The  $s \times s$  matrices  $\mathbf{F}_x$  describe the stage-specific per capita production of new individuals by reproduction, for individuals in age class  $x$ ,

$$\mathbf{F}_x = \begin{pmatrix} 0 & 0 & f_{SS} & f_{LS} \\ 0 & 0 & f_{SL} & f_{LL} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} (x), \quad x = 1, \dots, 50. \tag{104}$$

The first and second rows of  $\mathbf{F}_x$  describe per capita production of small and large juveniles, respectively.

The  $\omega \times \omega$  matrix  $\mathbf{D}_i$  advances the age class of surviving individuals in stage  $i$ . All of the  $\mathbf{D}_i$  are identical, given by Eq. 6. The  $\omega \times \omega$  matrices  $\mathbf{H}_i$  that assign newborn individuals to the first age class are all identical, given by Eq. 7. The matrices  $\mathbf{U}$ ,  $\mathbf{F}$ ,  $\mathbf{D}$ , and  $\mathbf{H}$  are created by putting  $\mathbf{U}_i$ ,  $\mathbf{F}_i$ ,  $\mathbf{D}_i$ , and  $\mathbf{H}_i$  on the diagonals, as in Eq. 10. Each of these matrices is of dimension  $s\omega \times s\omega$ , which for our model is  $200 \times 200$ . Finally, the age × stage-classified projection matrices  $\tilde{\mathbf{U}}$ ,  $\tilde{\mathbf{F}}$ , and  $\tilde{\mathbf{A}}$  are constructed following Eqs. 11–13.

*Population growth and structure*

It is convenient to begin with analyses of population growth, because the stable population structure will provide a family of mixing distributions that will be used several times in the cohort analyses to follow.

*Stable population structure.*—Using the parameters defined by Fig. 2, the population growth rate given by the dominant eigenvalue of  $\tilde{\mathbf{A}}$  is  $\lambda = 1.0083$ . The joint distribution of age and stage in the stable population is contained in the corresponding eigenvector  $\tilde{\mathbf{w}}$ , normalized to sum to 1. Fig. 3 shows the joint distribution of age and stage and the marginal age distribution  $\mathbf{w}_{\text{age}}$  and stage distribution  $\mathbf{w}_{\text{stage}}$ , calculated from Eqs. 61 and 62.

The joint distribution of age and stage in the stable population is dominated by young and small individuals (Fig. 3a). Adults (small and large) appear only at later ages. The marginal age distribution decays rapidly with age; most individuals are younger than 10. The marginal stage distribution is dominated by small juveniles (~60%). Large juveniles are much scarcer (~15%). Large adults, however, are more common (~20%) than small adults (~5%).

Even with only four stages, the interaction between age and stage in the stable population is complicated. The conditional stage distribution within each age class is shown in Fig. 3d. Up until an age of about 10, the population is dominated by small juveniles, but at later ages, it is dominated by large adults. The pattern results from the differences in survival probability and development time between large and small juveniles and adults.

*A mixing distribution for cohort initiation.*—Cohort dynamics begin with a group of newborn individuals, all in the first age class, but in possibly different stages. Calculations of survivorship, longevity, reproduction, and the age and stage at death need the stage distribution of the cohort as a mixing distribution. The mixing distribution  $\boldsymbol{\pi}$  (dimension  $s \times 1$ ) is obtained by extracting the entries of  $\tilde{\mathbf{w}}$  corresponding to age class 1 and normalizing so that the resulting vector sums to 1

$$\boldsymbol{\pi} = \frac{(\mathbf{e}_1^\top \otimes \mathbf{I}_s) \tilde{\mathbf{w}}}{(\mathbf{e}_1 \otimes \mathbf{I}_s^\top) \tilde{\mathbf{w}}} \tag{105}$$

$$= (0.85 \quad 0.15 \quad 0 \quad 0)^\top \tag{106}$$



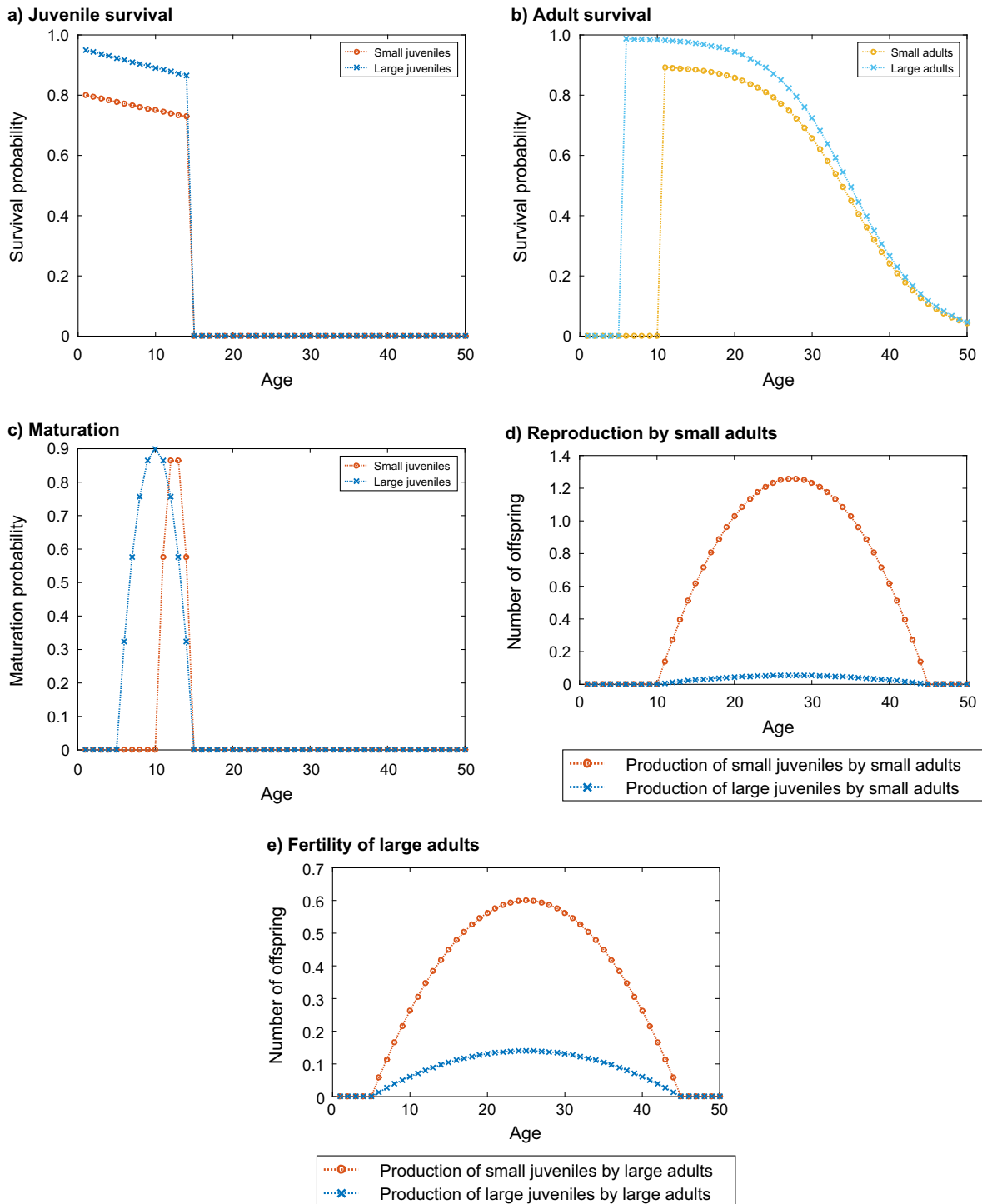


FIG. 2. The age  $\times$  stage-specific life history parameters for the example species.

where  $\mathbf{e}_1$  is a unit vector of length  $\omega$ . Thus about 85% of the new offspring in the stable population are small.

*Reproductive value.*—Reproductive value is given by the left eigenvector  $\bar{\mathbf{v}}$  of  $\tilde{\mathbf{A}}$  (dimension  $s\omega \times 1$ ). It is a relative measure of contribution to future population size, customarily scaled relative to that of a newborn individual. In our model

species, there are two such types, so Fig. 4 shows reproductive value scaled relative to that of a small juvenile.

The survival and maturation advantages of large juveniles relative to small juveniles translate into a dramatic increase in reproductive value. Juvenile reproductive value increases with age, as individuals get closer to maturation. After maturation, large adults have a reproductive advantage.

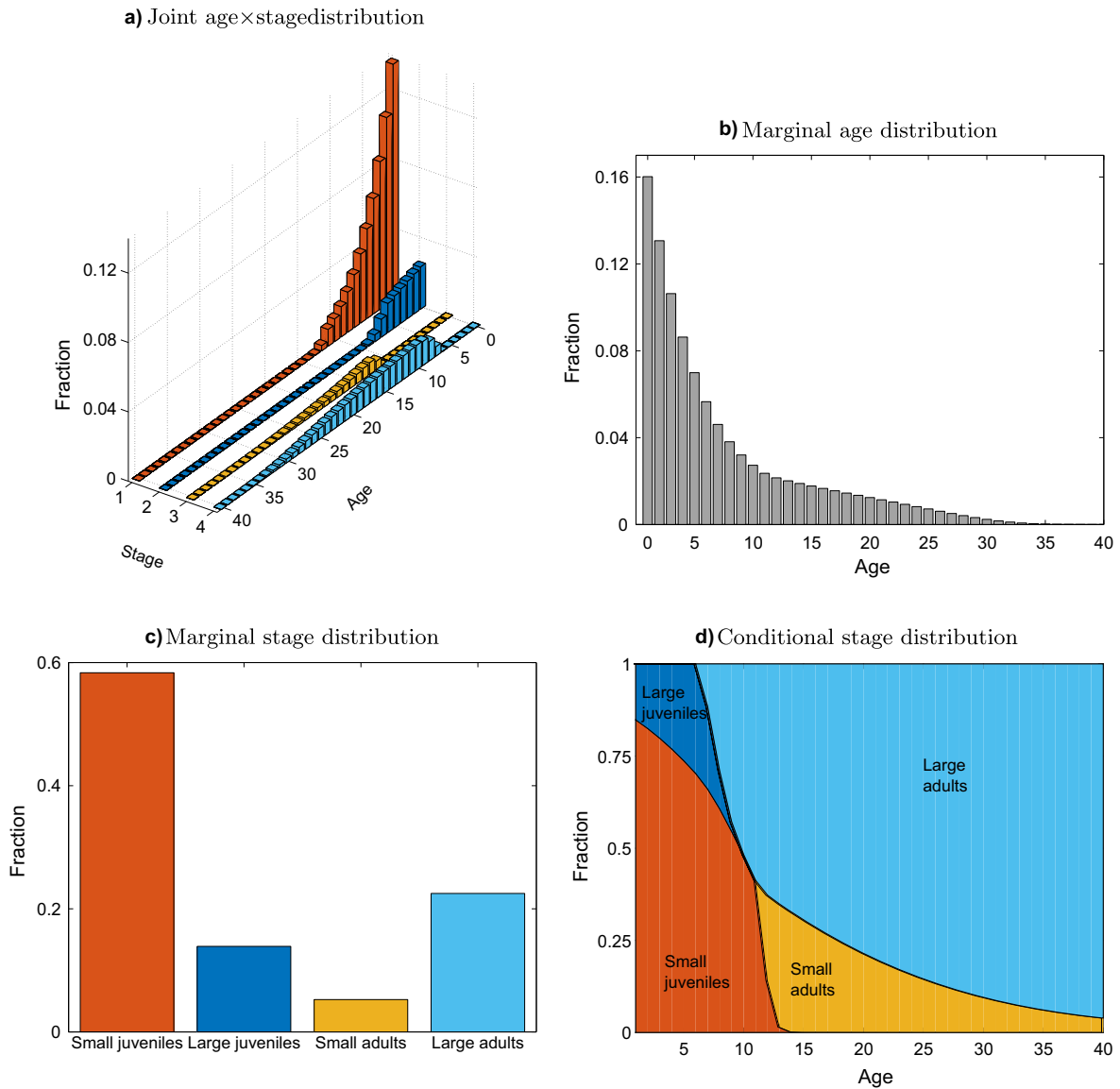


FIG. 3. The stable population structure for the model species, showing (a) the joint distribution by age and stage, (b) the marginal stable age distribution, (c) the marginal stable stage distribution, and (d) the conditional stable stage distribution.

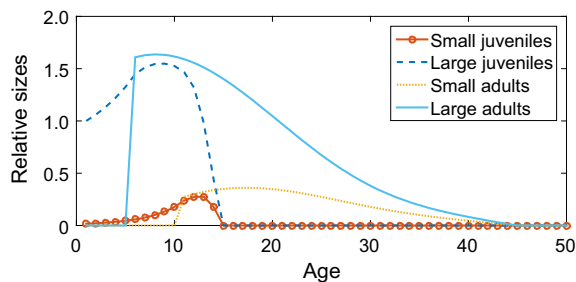


FIG. 4. The age-specific relative reproductive value of each stage for the example species, scaled relative to that of a large juvenile at birth.

Cohort analyses

*Survivorship.*—The stage-specific survivorship vector  $\ell(x)$  is calculated by projecting a cohort with Eq. 17 and

counting the proportion of survivors. Because our model species has two possible initial stages, small and large juveniles, the vector  $\ell(x)$  has only two non-zero entries. Fig. 5 shows the survivorship of small and large individuals. Survivorship declines rapidly at young ages, when individuals are subject to the high mortality rates of the juvenile stages, and then declines more slowly after maturation. Large individuals have a dramatic survivorship advantage. The survivorship of a mixed cohort, starting with a mixing distribution  $\pi$  calculated from the stable structure in Eq. 106, lies between the survivorship curves for small and large individuals, but much closer to that of small individuals. This reflects the dominance of the mixing distribution by small juveniles.

*Longevity.*—The remaining life expectancy (mean remaining longevity at a given age) of every stage is given by the vector  $\bar{\eta}_1$ , of dimension  $s\omega \times 1$ , calculated from the fundamental

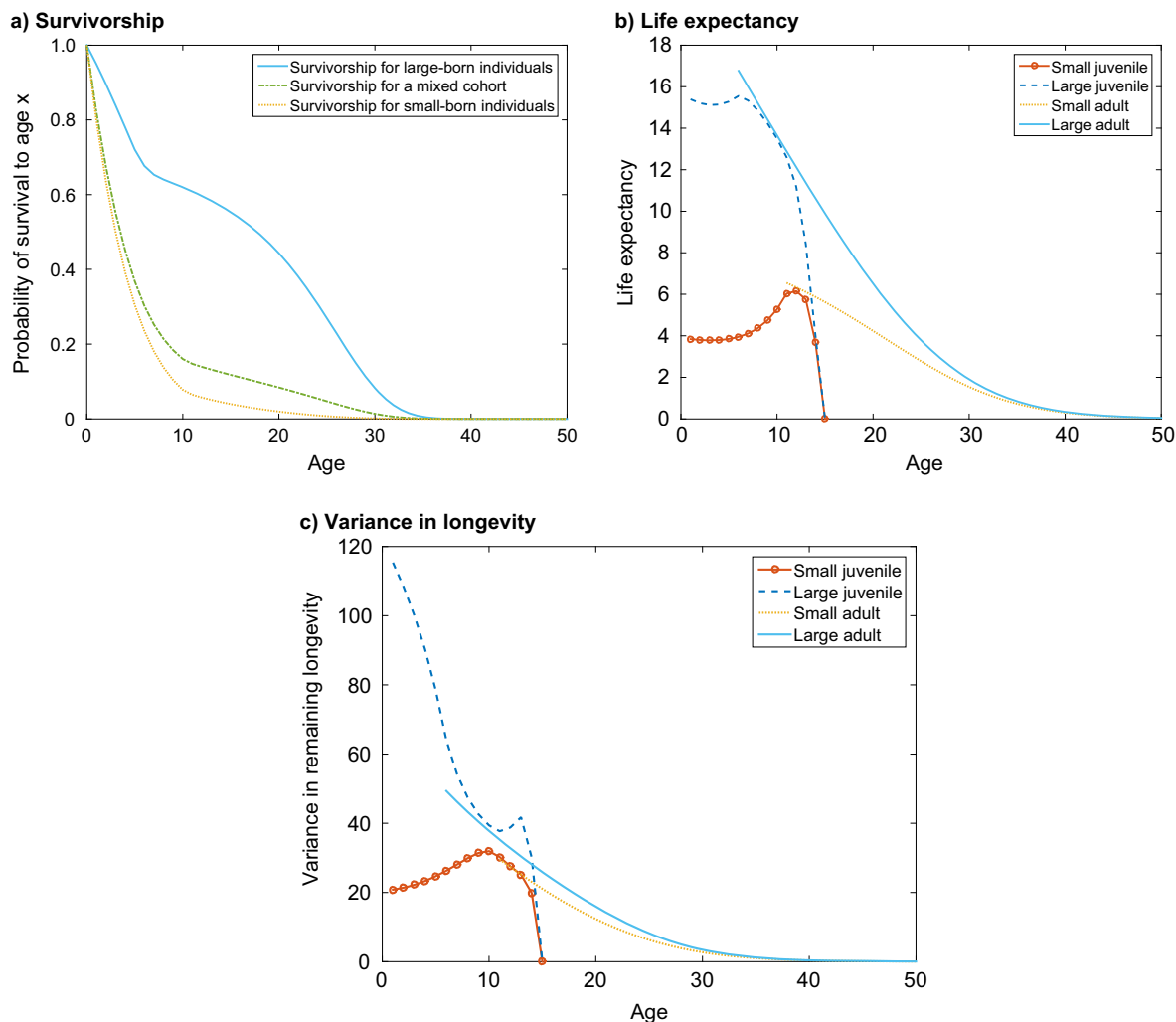


FIG. 5. Properties of the example species. (a) Survivorship  $\ell(x)$  for small and large individuals, and for a mixed cohort with proportions given by the mixing distribution  $\pi$ . (b) Remaining life expectancy as a function of age for each of the four stages. (c) Variance in remaining longevity as a function of age for each of the four stages.

matrix using Eq. 35 (Fig. 5c). Life expectancy of juveniles increases with age as they approach maturity, after which they will experience higher survival probability. The life expectancy of the adult stages decreases smoothly with age.

The variance in longevity, calculated using Eq. 37, decreases with age for large juveniles, but shows a peak just before the end of maturation, at which point individuals either die very soon, or mature and probably survive for quite some time, hence the large variance (Fig. 5c). Small juveniles show a similar peak in variance. For adults, variance decreases steadily with age.

The mean and variance of longevity for a mixed cohort are calculated using Eqs. 40 and 41 and the mixing distribution vector  $\pi_1$  given by Eq. 106. The resulting mean and variance in longevity, at birth, are shown in Table 4.

When the variance in longevity of the mixed cohort is decomposed into that due to heterogeneity among stages and that due to stochasticity within stages, using Eqs. 43 and 45, the variance due to stochasticity was 35.2 and the variance due to heterogeneity was 17.2. Thus the percentage

of the variance attributable to the heterogeneity among stages is 32.8%.

*Fertility.*—The entries of  $\mathbf{F}_x$  in Eq. 104 give the mean production of each type of offspring by parents of each type, at each age, as shown in Fig. 2d and e. We summarize this information with the fertility indices defined in Eqs. 29–32.

TABLE 4. Mean and variance in longevity, and decomposition into contributions within and between stages at birth.

Outcome	Small-born	Large-born	Mixed	%
Mean longevity	4.8	16.4	6.6	
Variance in longevity	20.8	115.4	52.3	
Variance between			17.2	32.8
Variance within			35.2	67.2

Note: Mixing distribution defined by the reproductive output of the stable population.

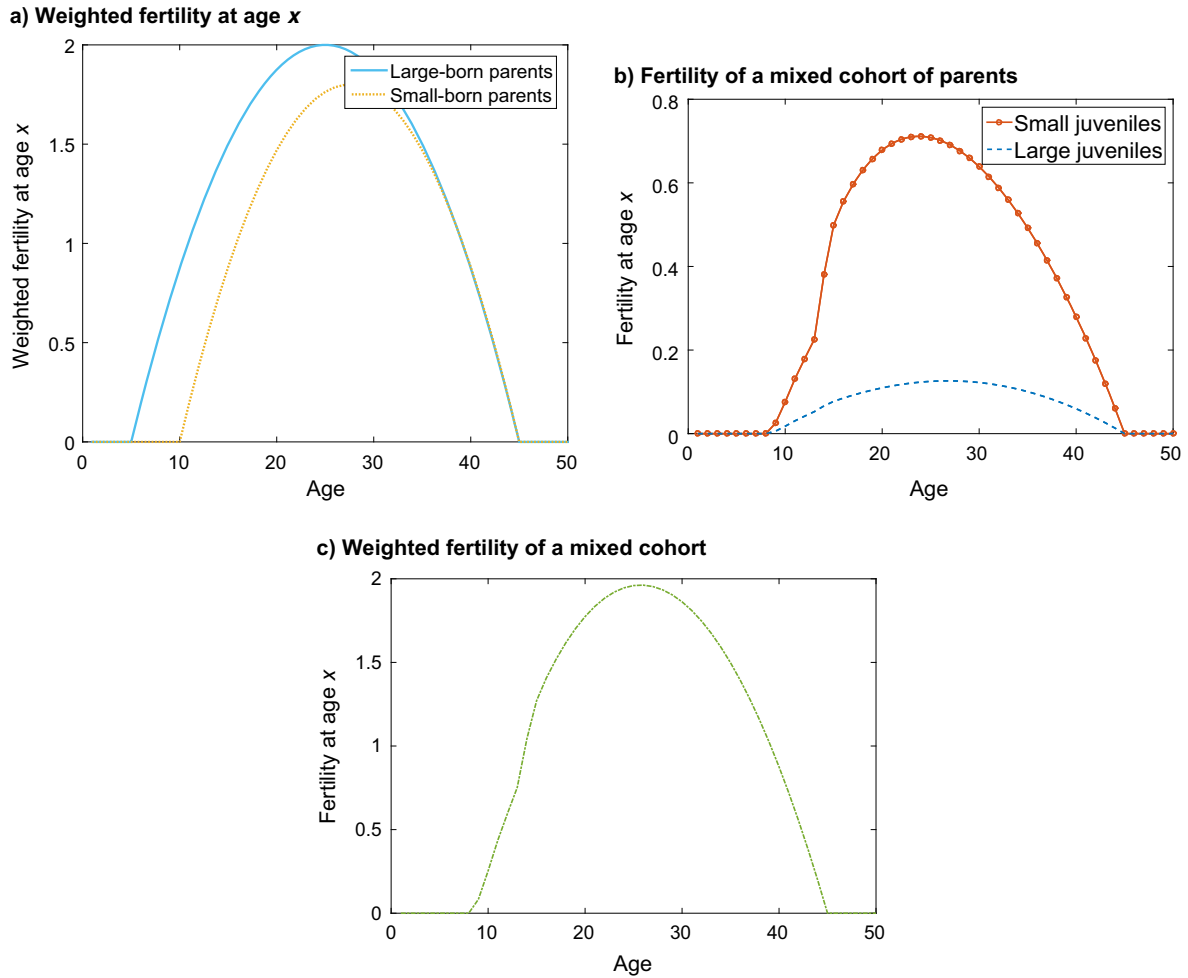


FIG. 6. Properties of the example species. (a) Weighted age-specific fertility  $\mathbf{f}_{\text{weighted}}(x)$  of small and large adults, with weights defined by the relative costs of small and large offspring (large 10 times more expensive than small). (b) The age-specific fertility  $\mathbf{f}_{\text{mixed}}(x)$  of a mixed cohort of small and large parents, with the mixing distribution given by the reproduction of the stable population structure. (c) The age-specific weighted fertility  $f(x)$  of the mixed cohort.

As an example of a weighted fertility schedule, we suppose that a large offspring is 10 times as costly as a small one, so we set

$$\mathbf{c} = \begin{pmatrix} 1 & 10 & 0 & 0^\top \end{pmatrix}. \quad (107)$$

The resulting fertility schedule  $\mathbf{f}_{\text{weighted}}$  is shown in Fig. 6a. The mixed fertility schedule, given by Eq. 30 is shown in Fig. 6b. And, finally, the weighted and mixed schedule from Eq. 32 is shown in Fig. 6c.

*R<sub>0</sub> and lifetime reproductive output.*—The matrix  $\mathbf{R}_{11}$  containing mean lifetime reproduction for our model species is

$$\begin{pmatrix} 0.334 & 4.656 & 0 & 0 \\ 0.014 & 1.087 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (108)$$

A large juvenile will, as a result of the combined age- and stage-specific life history characteristics of the model

species, produce many more offspring (by a 14-fold difference for small offspring and a 78-fold difference for large offspring). The net reproductive rate, calculated as the dominant eigenvalue of  $\mathbf{R}_{11}$ , is

$$R_0 = 1.166. \quad (109)$$

Suppose that we simply want to count total numbers of offspring produced. Then we use a weighting vector  $\mathbf{c} = \mathbf{1}_s$ , and obtain

$$\mathbf{r}_{\text{weighted}} = (0.35 \quad 5.74 \quad 0 \quad 0)^\top. \quad (110)$$

We see that, in terms of lifetime number of offspring, large individuals have a 16-fold advantage over small individuals.

To compute the lifetime reproduction of a mixed cohort, we use the mixing distribution  $\boldsymbol{\pi}$  calculated from the stable population in Eq. (106), to obtain

$$\mathbf{r}_{\text{mixed}} = (0.99 \quad 0.18 \quad 0 \quad 0)^\top. \quad (111)$$

Combining this particular choice of weighting and mixing, we obtain the mean lifetime reproduction, by a

mixed cohort, and counting total offspring of both types, as

$$\rho = 1.167. \quad (112)$$

The analysis of an age  $\times$  stage-classified model untangles the interacting effects of age- and stage-specific parameters. Even in a simple model species, large-born individuals have a higher survivorship than small-born individuals, and their earlier maturation increases their survival advantage because they can escape juvenile mortality sooner. The two stages have very different lifetime reproductive output, not simply as a result of differences in production (Fig. 2d and e), but as an interacting effect of juvenile survival, adult survival, and maturation. Our model example also demonstrates how the composition of a population will influence the demographic calculations. Mixed cohorts and populations and different choices of weighting vectors determine how a mixed population will perform in terms of survival, reproduction and population growth. These mixture calculations make possible the analysis of heterogeneity in initial population structure, and in the structure that develops as cohorts age within a population.

#### DISCUSSION

Our goal in this paper has been the development of a systematic methodology, of wide applicability, for the analysis of age  $\times$  stage-classified matrix population models.

#### Themes

Several themes have appeared repeatedly. Because the state space of an age  $\times$  stage-classified model is two-dimensional, results that are familiar as scalars in age-classified or stage-classified models now become vectors or matrices. The operations of marginalization and mixing thus play a critical role.

Consider any demographic quantity calculated from an age  $\times$  stage-classified model. If this quantity is a joint function of age and stage (e.g., the age and stage at death), then the marginal distributions (age at death and stage at death) are obtained as in Eqs. 51 and 52. If, on the other hand, a quantity is *conditional on* the age and stage of an individual (e.g., mean longevity), then the properties of this quantity for a heterogeneous cohort of individuals in different stages is calculated as a mixture, as in Eqs. 40 and 41. Using the total variance theorem, the variance in this quantity is partitioned into components due to stochasticity and to the heterogeneity created by the mixture, as in Eqs. 43 and 45. Any age  $\times$  stage-classified analysis should keep these concepts in mind.

Another theme that appears throughout our methodology is the interaction of age  $\times$  stage-dependence across three levels: the level of the individual (as in the life table functions), the level of the cohort (as in calculations of life expectancy and lifetime reproduction), and the level of the population. Although these levels are implicit in both population ecology and human demography, they appear explicitly here as a part of the methodology.

Variance is a recurring theme here. There is an increasing appreciation of the importance of variance created by stochastic differences between individuals (individual stochasticity). The inclusion of variance calculations in the approach here makes possible the decomposition of variances into components due to demographic differences (i.e., heterogeneity) among stages and ages, and to stochastic factors within stages.

#### Estimation and data

It is no surprise that age  $\times$  stage-classified models require an extra dimension of data: age-specific rates at every stage, stage-specific rates at every age. Such detailed data are rare (but see van Groenendael and Slim 1988) but we believe they will become more common as the importance of long-term individual data is recognized (Clutton-Brock and Sheldon 2010). The appropriate estimation procedures will, as always, depend on the properties of the species, the possibilities for monitoring, logistical difficulties, and available statistical methods. For that reason, we make no assumptions about how the stages are chosen or how the matrices  $U_i$  and  $F_i$  are estimated.

That being said, the generality of the method suggests some promising directions for future investigation. For example, multistate capture–recapture methods (Lebreton et al. 2009) are designed to estimate survival and transition probabilities from imperfect capture histories. In human demography and biomedical survival analysis, multistate event history methods are used to estimate age- and stage-specific rates of mortality (Andersen and Keiding 2015), but the data are often less fragmentary than in capture–recapture analysis. Possible connections between these approaches might help to estimate parameters for age  $\times$  stage-classified models.

In this context, for the special case of stages defined by continuous traits, integral projection models (IPMs) are essentially statistically sophisticated tools for estimating high-dimensional matrices. An IPM requires kernel functions for growth, survival, and fertility. If these kernels were estimated separately for age classes 1,  $\dots$ ,  $\omega$ , the resulting model would be analogous to those considered here (Ellner et al. 2016: Chapter 6). The fertility kernel are special cases of the set of  $F_i$ , and the growth and survival kernels special cases of the  $U_i$ , when partitioned into  $G_i$  and  $\Sigma_i$  as in Eq. 99. When discretized for analysis, the set of IPMs would provide a set of matrices to which our methods can be applied. Note, however, that stages need not be defined by discretization of continuous variables.

#### Data requirements

Estimating stage-specific survival, transitions, and fertility is challenging. Obtaining these estimates as a function of age is even more so. However, we have made no assumptions about how the  $U_i$  and  $F_i$  vary with age, so even partial age dependence can be incorporated. Perhaps fertility varies with age but survival does not (or cannot be estimated). Or perhaps age dependence can be detected only over sets of age classes (survival and transitions described for age classes 1, 2–5, and  $>5$ , while fertility is

described for age classes 1–5, 6–10, and >10). The end result in any of these cases is a set of matrices  $U_i$  and  $F_i$ , some of which are identical, but the analyses apply without any additional calculations.

In the limit, our method can be applied when  $U$  and  $F$  do not vary with age at all, in order to explore what would happen if age-dependent effects were introduced (e.g., Caswell and Salguero-Gómez 2013).

### Applications

Age × stage-classified matrix models are, in the end, matrix models. As such, they can be applied in the same way as age-classified or stage-classified models, to, inter alia, conservation of threatened species, control of pests, ecotoxicology, and epidemiology. These applications will require a model that specifies a management-relevant outcome from a management-relevant demographic structure. That might be a simple juvenile–adult model (as in our artificial species) or it might focus on the details of a particular species in a particular place at a particular time. In the latter case, one builds a model based on data from that place and time and applies the relevant methods to analyze it.

If both age and stage are significant components of the  $i$ -state (which is the rationale for constructing an age × stage-classified model in the first place), then applications will sometimes be well served by including the additional information. (But not necessarily. There is no reason to expect that the most complicated model is the most useful for any specific application.) Having a methodology for age × stage models that parallels that for age- and stage-classified models will facilitate model selection for applications.

### Heterogeneity

The role of heterogeneity in population dynamics is a problem of increasing interest (e.g. Vindenes et al. 2008, Steiner and Tuljapurkar 2012, Caswell 2014, Vindenes and Langangen 2015, Cam et al. 2016). From the perspective of an age-classified model, stage structure is a form of unobserved heterogeneity. It distorts age-specific outcomes because of intra-cohort selection (Vaupel et al. 1979, Vaupel and Yashin 1983). It is an additional source of variance among individuals in demographic fate above and beyond that generated by the age schedules of fertility and mortality. From the perspective of a stage-classified model, age is a form of unobserved heterogeneity. It distorts the dynamics of stage structures and stage durations, particularly in short-term transient responses. It can also increase variance above that created just by the outcome of stage-specific processes.

The age × stage-classified methodology presented here, especially the variance decomposition analysis, can contribute to the analysis of heterogeneity, because it makes it possible to directly measure the consequences of any form of heterogeneity that can be incorporated in the model, for any outcome for which a variance is available. This includes, but is not limited to, the incorporation of latent, unobserved factors such as frailty (Caswell 2014, Hartemink et al. 2017, Jenouvrier et al. 2018, Hartemink and Caswell 2018).

### Age, stage, and phenotype: incorporating heritability

The models analyzed here include stages of both parents and offspring, and so it is natural to wonder if somehow the relation of parent to offspring, and the associated concept of heritability, could be incorporated. There are three ways in which the methods could be extended to include phenotypes and heritability. Two of these correspond to the models of Coulson et al. (2010); the third is new.

1. Let stages represent phenotype categories. Then  $F_i$  would become a map from parent phenotype to offspring phenotype, for parents of age class  $i$ , encapsulating the notion of heritability. The matrix  $U_i$  would become a phenotype transition and survival matrix for age class  $i$ ; its structure would depend on whether phenotype traits do or do not change over an individual lifetime. The result would be an age × phenotype model.
2. Replace age classes with phenotype classes. This would lead to a stage × phenotype model. The matrices  $U_i$  would describe the survival and stage transitions of individuals in phenotype class  $i$ . The  $F_i$  would describe the production of offspring of all stages by individuals of all stages in phenotype class  $i$ . In the age × stage model, the matrix  $H_i$  is a map from parent age to offspring age. It has the trivial form of Eq. 7 because all newborn individuals start life in age class 1, regardless of the age of their parent. In the stage × phenotype model,  $H_i$  would become a parent-offspring phenotype map, encapsulating the heritability of the phenotypic trait. Perfect heritability would imply that  $H_i = I$ . In the complete absence of heritability, all columns of  $H_i$  would contain an identical probability vector. In this model, the matrices  $D_i$  would no longer describe age transitions, but would be replaced by phenotype transition matrices. Careful thought would be required to distinguish stages, or more generally  $i$ -states, from phenotypic traits (e.g., body size might be a stage, while body size at birth, or the coefficients in a growth equation, would be phenotypic traits). See Chevin (2015) for related discussion.
3. Age × stage × phenotype. The third possibility would be to incorporate phenotype as a third  $i$ state dimension, creating a fully age × stage × phenotype-classified model. The projection matrix for such a model, of dimension  $wsp$  (where  $p$  is the number of phenotype classes) is obtained from a three-dimensional array of matrices by a multigenerational generalization of the vec operator. Such models are called hyperstate matrix models; their construction and analysis is presented in Roth and Caswell (2016).

The development of age × phenotype, phenotype × stage, and especially age × stage × phenotype models is an open research problem.

### Extensions of age × stage-classified demography

The approach outlined here suggests extensions of the models; here we mention a few of these.

The construction of  $\bar{n}$  from the array  $\mathcal{N}$  yields a population vector in which stages are grouped within age classes.

This gives a sort of priority to stage-specific processes, treating age as a secondary characteristic. The alternative, treating age as primary and stage as a secondary kind of heterogeneity, would be obtained by transposing  $\mathcal{N}$  and creating a population vector, call it  $\tilde{\mathbf{n}}'$ , given by

$$\tilde{\mathbf{n}}' = \text{vec}\mathcal{N}^\top = \mathbf{K}_{s,o}\tilde{\mathbf{n}}. \quad (113)$$

The dynamics of  $\tilde{\mathbf{n}}'$  are given by a projection matrix

$$\tilde{\mathbf{A}}' = \mathbf{K}_{s,o}(\tilde{\mathbf{U}} + \tilde{\mathbf{F}})\mathbf{K}_{s,o}^\top. \quad (114)$$

Then all the results in this paper can be extended to this alternative arrangement by transforming the matrices as in Eq. 114. In doing so, the block-Leslie structure of  $\tilde{\mathbf{A}}$  in Eq. 14, in which stage matrices are arranged in the pattern of an age-classified matrix is replaced by a structure with age-classified matrices arranged in the pattern of the stage transitions.

Yet another interesting extension is to models structured by stage and age *within* a stage. Stage-classified models lead to geometrically distributed occupancy times. But if some time must elapse before an individual can advance to the next stage, some internal structure must be imposed to keep track of how long the individual has been in the stage; see Birt et al. (2009) for a recent presentation. These models are related to the age  $\times$  stage-classified models considered here, and it will be valuable to connect the two approaches.

We have presented methods here for linear, time-invariant models, but one value of our approach is that it permits direct extension to more sophisticated models. The population dynamics in Eq. 60 are an ordinary matrix population model, with the projection matrix  $\tilde{\mathbf{A}}$  written as the sum of  $\tilde{\mathbf{U}}$  and  $\tilde{\mathbf{F}}$ . Incorporating density dependence would result in a nonlinear model in which  $\tilde{\mathbf{A}}$  would be a function of  $\tilde{\mathbf{n}}(t)$ . The possible combinations of age- and stage-specific density effects make it a daunting prospect to construct such models.

However, the methodology here builds the projection matrices from the block-diagonal matrices like Eq. 10, each of which incorporates one of the sets of processes determining the dynamics (stage development, age transitions, fertility, and age assignment for newborn individuals). Even more, the block-diagonal matrices in turn are composed of well-defined matrices  $\mathbf{U}_i$  and  $\mathbf{F}_i$  describing the processes operating on each age-stage combination. It is at the level of these component matrices that the density effects or time variation actually operate, and the formulation presented here makes those matrices directly accessible for analysis. Something like this has been approached for a stage-classified epidemic model by Klepac and Caswell (2011).

Stability analysis, reactivity analysis, and sensitivity analysis of nonlinear age  $\times$  stage-classified models will then be approachable via extensions, similar to those in Eq. 89, of the matrix calculus approaches developed for nonlinear age- and stage-classified models (Caswell 2008, Verdy and Caswell 2008).

Our methodology invites similar extensions to time-varying and stochastic environments, by making it easily possible to explore the effects of stochasticity in age- and stage-specific survival, transitions, and fertility. The same applies to

spatial structure. However, adding location to the model will require increasing the dimensionality of the  $i$ -state space (just as did the addition of phenotype discussed above). The population vector for an age  $\times$  stage-classified model is obtained by applying the vec operator to the two-dimensional matrix  $\mathcal{N}$  in Eq. 1. The population vector for the age  $\times$  stage  $\times$  location model is obtained by applying a multidimensional version of the vec operator to the three-dimensional analogue of Eq. 1. Such arrays are called hypermatrices, and the higher-dimensional models referred to as hyperstate matrix models; their construction and analysis are detailed in Roth and Caswell (2016). It should come as absolutely no surprise that adding additional dimensions to the  $i$ -state space requires increasing amounts of additional data, but the framework presented here makes it possible, in principle, to do so. This list can be extended. We invite the reader to do so.

#### ACKNOWLEDGMENTS

This research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013), ERC Advanced Grant 322989, and also by NSF Grant DEB-1257545, and NWO-ALW Project ALWOP.2015.100. We thank Roberto Salguero-Gomez and three anonymous reviewers for comments on an earlier version of the paper. Discussions in the Theoretical and Computational Ecology Department at the University of Amsterdam provided the environment for the development of this paper.

#### LITERATURE CITED

- Andersen, P. K., and N. Keiding. 2015. Event history analysis in continuous time. Pages 310–319 in J. Wright, editor. International encyclopedia of the social and behavioral sciences. Volume 8. Second edition. Elsevier, Amsterdam, The Netherlands.
- Birch, L. C. 1948. The intrinsic rate of natural increase of an insect population. *Journal of Animal Ecology* 17:15–26.
- Birt, A., R. M. Feldman, D. M. Cairns, R. N. Coulson, M. Tchakerian, W. Xi, and J. M. Guldin. 2009. Stage-structured matrix models for organisms with non-geometric development times. *Ecology* 90:57–68.
- Cam, E., L. M. Aubry, and M. Authier. 2016. The conundrum of heterogeneities in life history studies. *Trends in Ecology & Evolution* 31:872–886.
- Caswell, H. 2001. *Matrix population models: construction, analysis, and interpretation*. Second edition. Sinauer Associates, Sunderland, Massachusetts, USA.
- Caswell, H. 2006. Applications of Markov chains in demography. Pages 319–334 in A. N. Langville, and W. J. Stewart, editors. MAM2006: Markov anniversary meeting. Bosen Books, Raleigh, North Carolina, USA.
- Caswell, H. 2008. Perturbation analysis of nonlinear matrix population models. *Demographic Research* 18:59–116.
- Caswell, H. 2009. Stage, age and individual stochasticity in demography. *Oikos* 118:1763–1782.
- Caswell, H. 2011. Beyond R0: demographic models for variability of lifetime reproductive output. *PLoS ONE* 6:e20809.
- Caswell, H. 2012. Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology* 5:403–417.
- Caswell, H. 2013. Sensitivity analysis of discrete Markov chains via matrix calculus. *Linear Algebra and its Applications* 438:1727–1745.
- Caswell, H. 2014. A matrix approach to the statistics of longevity in heterogeneous frailty models. *Demographic Research* 31:553–592.

- Caswell, H., and R. Salguero-Gómez. 2013. Age, stage and senescence in plants. *Journal of Ecology* 101:585–595.
- Caswell, H., and E. Shyu. 2012. Sensitivity analysis of periodic matrix population models. *Theoretical Population Biology* 82:329–339.
- Caswell, H., and S. F. van Daalen. 2016. A note on the vec operator applied to unbalanced block-structured matrices. *Journal of Applied Mathematics* 2016:1–3.
- Charlesworth, B. 1994. *Evolution in age-structured populations*. Second edition. Cambridge University Press, Cambridge, UK.
- Chevin, L.-M. 2015. Evolution of adult size depends on genetic variance in growth trajectories: a comment on analyses of evolutionary dynamics using integral projection models. *Methods in Ecology and Evolution* 6:981–986.
- Clutton-Brock, T., and B. C. Sheldon. 2010. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology and Evolution* 25:562–573.
- Coale, A. J. 1972. *The growth and structure of human populations: a mathematical approach*. Princeton University Press, Princeton, New Jersey, USA.
- Cochran, M. E., and S. Ellner. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecological Monographs* 62:345–364.
- Cohen, J. E. 1982. Multiregional age-structured populations with changing vital rates: weak and stochastic ergodic theorems. Pages 477–503 in K. C. Land, and A. Rogers, editors. *Multidimensional mathematical demography*. Academic Press, New York, New York, USA.
- Cole, L. C. 1954. The population consequences of life history phenomena. *Quarterly Review of Biology* 29:103–137.
- Coulson, T., S. Tuljapurkar, and D. Z. Childs. 2010. Using evolutionary demography to link life history theory, quantitative genetics and population ecology. *Journal of Animal Ecology* 79:1226–1240.
- Csetenyi, A. I., and D. O. Logofet. 1989. Leslie model revisited: some generalizations to block structures. *Ecological Modelling* 48:277–290.
- Cushing, J. M., and Y. Zhou. 1994. The net reproductive value and stability in matrix population models. *Natural Resources Modelling* 8:297–333.
- De Roos, A. M. 2008. Demographic analysis of continuous-time life-history models. *Ecology Letters* 11:1–15.
- Deevey, E. S. 1947. Life tables for natural populations of animals. *Quarterly Review of Biology* 22:283–314.
- Dublin, L. I., and A. J. Lotka. 1937. Uses of the life table in vital statistics. *American Journal of Public Health* 27:481–491.
- Easterling, M. R., S. P. Ellner, and P. M. Dixon. 2000. Size-specific sensitivity: applying a new structured population model. *Ecology* 81:694–708.
- Edwards, R. D. 2011. Changes in world inequality in length of life: 1970–2000. *Population and Development Review* 37:499–528.
- Edwards, R. D., and S. Tuljapurkar. 2005. Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review* 31:645–674.
- Ellner, S. P., D. Z. Childs, and M. Rees. 2016. *Data-driven modelling of structured populations*. Springer, New York, New York, USA.
- Engelman, M., H. Caswell, and E. M. Agree. 2014. Why do lifespan variability trends for the young and old diverge? *Demographic Research* 30:1367–1396.
- Feeney, G. M. 1970. Stable age by region distributions. *Demography* 7:341–348.
- Feichtinger, G. 1971. *Stochastische modelle demographischer prozesse*. Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, Berlin, Germany.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Oxford University Press, Oxford, UK.
- Frühwirth-Schnatter, S. 2006. *Finite mixture and Markov switching models*. Springer-Verlag, New York, New York, USA.
- Gaillard, J.-M., N. G. Yoccoz, J.-D. Lebreton, C. Bonenfant, S. Devillard, A. Loison, D. Pontier, and D. Allaine. 2005. Generation time: a reliable metric to measure life-history variation among mammalian populations. *American Naturalist* 166:119–123.
- Goodman, L. A. 1969. The analysis of population growth when the birth and death rates depend on several factors. *Biometrics* 25:659–681.
- Hamilton, W. D. 1966. The moulding of senescence by natural selection. *Journal of Theoretical Biology* 12:12–45.
- Hartemink, N., and H. Caswell. 2018. Variance in animal longevity: contributions of heterogeneity and stochasticity. *Population Ecology*. <https://doi.org/10.1007/s10144-018-0616-7>.
- Hartemink, N., T. I. Missov, and H. Caswell. 2017. Stochasticity, heterogeneity, and variance in longevity in human populations. *Theoretical Population Biology* 114:107–117.
- Henderson, H. V., and S. R. Searle. 1981. The vec-permutation matrix, the vec operator and Kronecker products: a review. *Linear and Multilinear Algebra* 9:271–288.
- Honeycutt, A. A., J. P. Boyle, K. R. Broglio, T. J. Thompson, T. J. Hoerger, L. S. Geiss, and K. M. V. Narayan. 2003. A dynamic Markov model for forecasting diabetes prevalence in the United States through 2050. *Health Care Management Science* 6:155–164.
- Horvitz, C. C., and S. Tuljapurkar. 2008. Stage dynamics, period survival, and mortality plateaus. *American Naturalist* 172:203–215.
- Hutchinson, G. E. 1978. *An introduction to population ecology*. Yale University Press, New Haven, Connecticut, USA.
- Iosifescu, M. 1980. *Finite Markov processes and their applications*. Wiley, New York, New York, USA.
- Jenouvrier, S., L. M. Aubry, C. Barbraud, H. Weimerskirch, and H. Caswell. 2018. Interacting effects of unobserved heterogeneity and individual stochasticity in the life cycle of the Southern fulmar. *Journal of Animal Ecology* 87:212–222.
- Kempthorne, O. 1957. *An Introduction to Genetic Statistics*. John Wiley & Sons, New York, New York, USA.
- Keyfitz, N. 1968. *Introduction to the Mathematics of Population*. Addison-Wesley, Reading, Massachusetts, USA.
- Keyfitz, N., and H. Caswell. 2005. *Applied mathematical demography*. Third edition. Springer, New York, New York, USA.
- Klepac, P., and H. Caswell. 2011. The stage-structured epidemic: linking disease and demography with a multi-state matrix approach model. *Theoretical Ecology* 4:301–319.
- Knott, K. E., and D. McHugh. 2012. Introduction to symposium: poecilogony window on larval evolutionary transitions in marine invertebrates. *Integrative and Comparative Biology* 52:120–127.
- Krug, P. J. 2007. Poecilogony and larval ecology in the gastropod genus *Alderia*\*. *American Malacological Bulletin* 23:99–111.
- 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. *Theoretical Population Biology* 49:291–313.
- Lebreton, J.-D. 2005. Age, stages, and the role of generation time in matrix models. *Ecological Modelling* 188:22–29.
- Lebreton, J.-D., M. Khaladi, and V. Grosbois. 2000. An explicit approach to evolutionarily stable dispersal strategies: no cost of dispersal. *Mathematical Biosciences* 165:163–176.
- Lebreton, J.-D., J. D. Nichols, R. J. Barker, R. Pradel, and J. A. Spendlow. 2009. Modeling individual animal histories with multistate capture-recapture models. *Advances in Ecological Research* 41:87–173.
- Lefkovich, L. 1965. The study of population growth in organisms grouped by stages. *Biometrics* 21:1–18.
- Leslie, P. H., and T. Park. 1949. The intrinsic rate of natural increase of *Tribolium castaneum* Herbst. *Ecology* 30:469–477.



- Levin, L. A., H. Caswell, K. D. DePatra, and E. L. Creed. 1987. Demographic consequences of larval development mode: planktotrophy vs. lecithotrophy in *Streblospio benedicti*. *Ecology* 68:1877–1886.
- Lewontin, R. C. 1965. Selection for colonizing ability. Pages 77–91 in H. G. Baker and G. L. Stebbins, editors. *The genetics of colonizing species*. Academic Press, New York, New York, USA.
- Logofet, D. O. 2002. Three sources and three constituents of the formalism for a population with discrete age and stage structures (in Russian). *Mathematical Modelling* 14:11–22.
- Logofet, D. O. 2013. Complexity in matrix population models: polyvariant ontogeny and reproductive uncertainty. *Ecological Complexity* 15:43–51.
- Loichinger, E. 2015. Labor force projections up to 2053 for 26 EU countries, by age, sex, and highest level of educational attainment. *Demographic Research* 32:443–486.
- Lotka, A. J. 1939. *Théorie analytique des associations biologiques: analyse démographique avec application particulière à l'espèce humaine*. Actualités Scientifiques et Industrielles, Hermann et Cie, Paris, France (published in translation as: *Analytical theory of biological populations*. Translated by D. P. Smith and H. Rosert. Plenum Press, New York, NY. 1998).
- Magnus, J. R., and H. Neudecker. 1979. The commutation matrix: some properties and applications. *Annals of Statistics* 7:381–394.
- Metcalf, C. J. E., J. Lessler, P. Klepac, A. Morice, B. T. Grenfell, and O. N. Bjørnstad. 2012. Structured models of infectious disease: inference with discrete data. *Theoretical Population Biology* 82:275–282.
- Metz, J. A. J., and O. Diekmann. 1986. *The dynamics of physiologically structured populations*. Lecture Notes in Biomathematics. Springer-Verlag, Berlin, Germany.
- Nisbet, R. M., and W. Gurney. 1982. *Modelling fluctuating populations*. Wiley, Chichester, UK.
- Norton, H. T. J. 1928. Natural selection and Mendelian variation. *Proceedings of the London Mathematical Society* 28:1–45.
- Pearl, R., and J. R. Miner. 1935. Experimental studies on the duration of life. XIV. The comparative mortality of certain lower organisms. *Quarterly Review of Biology* 10:60–79.
- Rényi, A. 1970. *Probability theory*. North-Holland, Amsterdam, The Netherlands.
- Rogers, A. 1968. *Matrix analysis of interregional population growth and distribution*. University of California Press, Berkeley, California, USA.
- Rogers, A. 1975. *Introduction to multiregional mathematical demography*. Wiley, New York, New York, USA.
- Roth, W. E. 1934. On direct product matrices. *Bulletin of the American Mathematical Society* 40:461–468.
- Roth, G., and H. Caswell. 2016. Hyperstate matrix models: extending demographic state spaces to higher dimensions. *Methods in Ecology and Evolution* 7:1438–1450.
- Roth, G., and H. Caswell. 2018. Occupancy time in sets of states for demographic models. *Theoretical Population Biology* 120:62–77.
- Salguero-Gómez, R., et al. 2015. The COMPADRE plant matrix database: an open online repository for plant demography. *Journal of Ecology* 103:202–218.
- Salguero-Gómez, R., et al. 2016. COMADRE: a global data base of animal demography. *Journal of Animal Ecology* 85:371–384.
- Sánchez Gassen, N. E. 2014. Germany's future electors: developments of the German electorate in times of demographic change. Springer, Wiesbaden, Germany.
- Sinko, J. W., and W. Streifer. 1967. A new model for age-size structure of a population. *Ecology* 48:910–918.
- Slobodkin, L. B. 1961. *Growth and regulation of animal populations*. Holt, Reinhart, and Winston, New York, New York, USA.
- Stearns, S. C. 1992. *The evolution of life histories*. Oxford University Press, Oxford, UK.
- Steiner, U. K., and S. Tuljapurkar. 2012. Neutral theory for life histories and individual variability in fitness components. *Proceedings of the National Academy of Sciences of the United States of America* 109:4684–4689.
- Steiner, U. K., S. Tuljapurkar, T. Coulson, and C. Horvitz. 2012. Trading stages: life expectancies in structured populations. *Experimental Gerontology* 47:773–781.
- Steiner, U. K., S. Tuljapurkar, and T. Coulson. 2014. Generation time, net reproductive rate, and growth in stage-age-structured populations. *American Naturalist* 183:771–783.
- Steinsaltz, D., and S. N. Evans. 2004. Markov mortality models: implications of quasistationarity and varying initial distributions. *Theoretical Population Biology* 65:319–337.
- Stott, I., S. Townley, D. Carslake, and D. J. Hodgson. 2010. On reducibility and ergodicity of population projection matrix models. *Methods in Ecology and Evolution* 1:242–252.
- Strasser, C. A., M. G. Neubert, H. Caswell, and C. M. Hunter. 2012. Contributions of high- and low-quality patches to a metapopulation with stochastic disturbance. *Theoretical Ecology* 5:167–179.
- Tuljapurkar, S. 1990. *Population dynamics in variable environments*. Springer-Verlag, New York, New York, USA.
- Tuljapurkar, S., and H. Caswell. 1997. *Structured-population models in marine, terrestrial, and freshwater ecosystems*. Chapman and Hall, New York, New York, USA.
- Tuljapurkar, S., and R. D. Edwards. 2011. Variance in death and its implications for modeling and forecasting mortality. *Demographic Research* 24:497–526.
- Tuljapurkar, S., and C. C. Horvitz. 2006. From stage to age in variable environments: life expectancy and survivorship. *Ecology* 87:1497–1509.
- van Daalen, S., and H. Caswell. 2015. Lifetime reproduction and the second demographic transition: stochasticity and individual variation. *Demographic Research* 33:561–588.
- van Daalen, S. F., and H. Caswell. 2017. Lifetime reproductive output: individual stochasticity, variance, and sensitivity analysis. *Theoretical Ecology* 10:355–374.
- Van der Gaag, N., G. Bijwaard, J. de Beer, and L. Bonneux. 2015. A multistate model to project elderly disability in case of limited data. *Demographic Research* 32:75.
- van Groenendael, J. M., and P. Slim. 1988. The contrasting dynamics of two populations of *Plantago lanceolata* classified by age and size. *Journal of Ecology* 76:585–599.
- Van Raalte, A. A., and H. Caswell. 2013. Perturbation analysis of indices of lifespan variability. *Demography* 50:1615–1640.
- Vaupel, J. W., and V. Canudas Romo. 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., and A. I. Yashin. 1983. *The deviant dynamics of death in heterogeneous populations*. International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16:439–454.
- Vaupel, J. W., Z. Zhang, and A. A. van Raalte. 2011. Life expectancy and disparity: an international comparison of life table data. *British Medical Journal Open* 1:e000128.
- Verdy, A., and H. Caswell. 2008. Sensitivity analysis of reactive ecological dynamics. *Bulletin of Mathematical Biology* 70:1634–1659.
- Vindenes, Y., and Ø. Langangen. 2015. Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology Letters* 18:417–432.
- Vindenes, Y., S. Engen, and B.-E. Saether. 2008. Individual heterogeneity in vital parameters and demographic stochasticity. *American Naturalist* 171:455–467.
- Werner, P. A., and H. Caswell. 1977. Population growth rates and age versus stage-distribution models for teasel (*Dipsacus sylvestris* Huds.). *Ecology* 58:1103–1111.
- Willekens, F. 2014. *Multistate analysis of life histories with R*. Springer, New York, New York, USA.
- Wu, G., Y.-M. Wang, A. Yen, J.-M. Wong, H.-C. Lai, J. Warwick, and T. Chen. 2006. Cost effectiveness analysis of colorectal cancer screening with stool DNA testing in intermediate-incidence countries. *BMC Cancer* 6:136.

- Zhou, Y., H. Putter, and G. Doblhammer. 2016. Years of life lost due to lower extremity injury in association with dementia, and care need: a 6-year follow-up population-based study using a multi-state approach among German elderly. *BMC Geriatrics* 16:9.
- Zuidema, P. A., R. J. W. Brienen, H. J. During, and B. Güneralp. 2009. Do persistently fast-growing juveniles contribute disproportionately to population growth? A new analysis tool for matrix models and its application to rainforest trees. *American Naturalist* 174:709–719.

## SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecm.1306/full>