UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Demographic variance in heterogeneous populations: matrix models and sensitivity analysis

Caswell, H.; Vindenes, Y.

DOI
10.1111/oik.04708

Publication date
2018

Document Version
Final published version

Published in
Oikos

License
CC BY

Link to publication

Citation for published version (APA):
Caswell, H., & Vindenes, Y. (2018). Demographic variance in heterogeneous populations: matrix models and sensitivity analysis. *Oikos*, *127*(5), 648-663. https://doi.org/10.1111/oik.04708

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, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:11 Oct 2026

OIKOS

Research

Demographic variance in heterogeneous populations: matrix models and sensitivity analysis

Hal Caswell and Yngvild Vindenes

H. Caswell (http://orcid.org/0000-0003-4394-6894) (h.caswell@uva.nl), Inst. for Biodiversity and Ecosystem Dynamics, Univ. of Amsterdam, PO Box 94248, NL-1090n GE Amsterdam, the Netherlands. – Y. Vindenes, Centre for Ecological and Evolutionary Synthesis, Univ. of Oslo, Oslo, Norway.

Oikos
127: 648–663, 2018
doi: 10.1111/oik.04708

Subject Editor: Jean-Michel Gaillard
Editor-in-Chief: Dries Bonte
Accepted 9 January 2018

The demographic consequences of stochasticity in processes such as survival and reproduction are modulated by the heterogeneity within the population. Therefore, to study effects of stochasticity on population growth and extinction risk, it is critical to use structured population models in which the most important sources of heterogeneity (e.g. age, size, developmental stage) are incorporated as i-state variables.

Demographic stochasticity in heterogeneous populations has often been studied using one of two approaches: multitype branching processes and diffusion approximations. Here, we link these approaches, through the demographic stochasticity in age- or stage-structured matrix population models. We derive the demographic variance, σ_d^2, which measures the per capita contribution to the variance in population growth increment, and we show how it can be decomposed into contributions from transition probabilities and fertility across ages or stages. Furthermore, using matrix calculus we derive the sensitivity of σ_d^2 to age- or stage-specific mortality and fertility. We apply the methods to an extensive set of data from age-classified human populations (long-term time-series for Sweden, Japan and the Netherlands; two hunter–gatherer populations, and the high-fertility Hutterites), and to a size-classified population of the herbaceous plant *Calathea ovandensis*. For the human populations our analysis reveals substantial temporal changes in the demographic variance as well as its main components across age.

These new methods provide a powerful approach for calculating the demographic variance for any structured model, and for analyzing its main components and sensitivities. This will make possible new analyses of demographic variance across different kinds of heterogeneity in different life cycles, which will in turn improve our understanding of mechanisms underpinning extinction risk and other important biological outcomes.

Introduction

The sources of inter-individual variance in population outcomes can be divided into two categories. One is heterogeneity: genuine differences in the properties of individuals. Heterogeneity is defined as any difference among individuals that results in differences among them in any of the vital rates (mortality, fertility, development) that

© 2018 The Authors. This an Online Open article
This work is licensed under the terms of a Creative Commons Attribution 4.0 International License (CC-BY) <http://creativecommons.org/licenses/by/4.0/>. The license permits use, distribution and reproduction in any medium, provided the original work is properly cited.

they experience. The differences could be genetic, physiological, behavioral, the result of environmental conditions (e.g. resource availability in early life), etc.

But heterogeneity is not the only source; inter-individual variance may also be due to stochasticity: variance due to random processes in the life cycle or the environment. (We are concerned with stochasticity in ecological dynamics; we do not address sampling errors in the estimation of parameters, which is a different issue.) Our understanding of stochasticity has progressed dramatically in recent years. Stochasticity in ecological models can be divided into three types.

- Individual stochasticity. This is variation among individuals created by random events in the processes of survival, reproduction and development (Caswell 2009, 2011, Caswell and Kluge 2015, van Daalen and Caswell 2015, 2017). Individual stochasticity has the same meaning as 'dynamic heterogeneity' in the independent usage of Tuljapurkar et al. (2009). Even in the complete absence of heterogeneity; i.e. when every individual is subject to exactly the same vital rates, individuals will differ in their demographic outcomes because of individual stochasticity. It is thus a candidate explanation for variance among individuals in, e.g. longevity and lifetime reproductive output. It has now been documented that individual stochasticity can explain a surprisingly large fraction of inter-individual variance (Steiner et al. 2010, Steiner and Tuljapurkar 2012, Hartemink et al. 2017, Jenouvrier et al. 2017, Hartemink and Caswell 2018).
- Demographic stochasticity. This is the result of embedding individual stochasticity in a population growth process, by including the (stochastic) creation of new individuals by reproduction. In models for demographic stochasticity, including the branching processes and diffusion models considered here, the random events at the individual level produce variance in population size over time. Demographic stochasticity is thus a candidate explanation for variance in population growth trajectories, growth rates, and population size, as well as a potential cause of population extinction. Because demographic stochasticity creates variation in population growth due to random individual events, its effects scale inversely with population size. It is particularly important for the viability of small populations in conservation, for the fate of new mutations, for the success or failure of introduced populations, and for the fate of epidemics following exposure to an infected individual.
- Environmental stochasticity. This is variation in population growth caused by stochastic changes over time in the environment, which affect all individuals in the population (Tuljapurkar 1990). We will not consider environmental stochasticity here.

The incorporation of individual heterogeneity has always been the main goal and the greatest accomplishment of demography. Age-classified models are so familiar that we forget that, in their absence, age differences were a source of heterogeneity that was unaccounted for in unstructured models. Size-classified models seem so ordinary now that we may forget that they were developed because size was a source of heterogeneity unaccounted for in age-classified models. Age-classified, stage-classified and multistate demographic models incorporate differences among individuals in the vital rates and yield the consequences for population behavior. (Regardless of the mathematical form of the models: renewal equations, matrix projections, integral equations, partial differential equations, etc.). Incorporating age, or size, or more elaborate differences among individuals into an analysis does not change their nature as heterogeneity in any signficant way.

The analysis of unobserved or latent heterogeneity, before it is incorporated into demographic models, is a long-standing challenge in demography (Yule 1910). Such heterogeneity is given a variety of names: frailty (Vaupel et al. 1979), individual quality (Wilson and Nussey 2010), fecundability (Sheps and Menken 1973), and the distinction between movers and stayers (Morrison et al. 1971). Recent advances in multistate matrix models and estimation of latent variables have made it possible to incorporate unobserved heterogeneity along with other forms of heterogeneity.

Once the extra source of heterogeneity is incorporated into a multistate model, then demographic analysis can quantify the effects of all types of heterogeneity operating together. The incorporation can be accomplished either by identifying and measuring the source of heterogeneity (e.g. health status or education in human demography) or by estimating it as a latent unobserved variable (Erişoğlu et al. 2012, Hartemink et al. 2017, Jenouvrier et al. 2017).

The rest of the paper is organized as follows. The section 'Two models for demographic stochasticity' recalls the definitions of the branching process and diffusion models for demographic stochasticity. The section 'Demographic variance from matrix models' presents a general result for calculating the demographic variance (to be defined below) from any matrix population model. The matrix formulation leads directly, in the section 'Sensitivity analysis of the demographic variance' to a sensitivity analysis that quantifies the effects of mortality, fertility, and transitions on the demographic variance and extinction probability. The next section (Examples) presents examples from several age-classified human populations and a stage-classified plant species.

Notation

Matrices are denoted by upper-case boldface letters (e.g. $\mathbf{U}$), and vectors by lower-case boldface letters (e.g. $\mathbf{n}$). Vectors are column vectors by default. Where necessary, the dimensions of matrices and vectors are denoted by subscripts; thus $\mathbf{I}_s$ is an identity matrix of order s and $\mathbf{1}_s$ is a $s \times 1$ vector of ones. The unit vector $\mathbf{e}_i$ is a vector with a 1 in the ith 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. When convenient,

Matlab notation will be used to refer to rows and columns of matrices; thus $\mathbf{X}(i, :)$ is the ith row and $\mathbf{X}(:, j)$ the jth column of $\mathbf{X}$. The diagonal matrix with $\mathbf{x}$ on the diagonal and zeros elsewhere is denoted $D(\mathbf{x})$. The symbol $\circ$ denotes the Hadamard, or element-by-element product; the symbol $\otimes$ denotes the Kronecker product. The symbol $\|\mathbf{x}\|$ denotes the 1-norm of the vector $\mathbf{x}$. The transpose of the matrix $\mathbf{X}$ is $\mathbf{X}^{\mathsf{T}}$. The number of age classes is ω and the number of stages is s. The expected value of x is $\mathrm{E}(x)$; the variance is $\mathrm{Var}(x)$. The covariance matrix of the entries of the vector $\mathbf{x}$ is $\mathrm{Cov}(\mathbf{x})$. The vec operator (vec $\mathbf{X}$) stacks the columns of $\mathbf{X}$ into a column vector.

Two models for demographic stochasticity

Two modelling approaches – branching processes and diffusion models – have been widely used to study effects of demographic stochasticity; our goal in this paper is to present a new synthesis of these approaches and develop new analyses from them. There is a close link between matrix population models and multitype branching process (MBP) models (Pollard 1966, Caswell 2001). Information on the higher moments of the population can be obtained from the matrix model with only a small amount of extra information (or, failing that, a few extra approximations; Caswell 2001). Our goal here is to provide an explicit link between the MBP formulation of matrix population models and diffusion models for demographic stochasticity, and to use that link to develop a sensitivity analysis for demographic variance. Our results are generally applicable to age- or stage-classified models.

Moments of multitype branching processes

Demographic stochasticity was the subject of what may be the first stochastic population model: the branching process introduced by Galton and Watson (1874) to analyze the extinction of British surnames. The Galton–Watson branching process describes a population in which all individuals are identical, and independently produce a random number of 'offspring' according to a specified probability distribution. The 'offspring' may, depending on the context, include the survival of the 'parent' individual as well as the production of genuinely new offspring. The analysis is carried out by a now classical application of probability generating functions.

Structured populations, in which individuals differ on the basis of age or stage, are analyzed using multitype branching processes (MPB). Demographic applications of these models were developed for age-classified models by Pollard (1966) and extended to stage-classified models by Caswell (2001, chapter 15). In these models, an individual of a given type is characterized by a multivariate probability distribution of the numbers of 'offspring' of all types produced over a time interval. Again, offspring include the possible survival and stage transition of the starting individual. Analysis based on the probability generating functions of these multivariate distributions gives the properties of future population size and the probabilities of extinction. The analysis also provides expressions for the moments of the population vector as a function of age or stage. The familiar population projection matrix is the expected value operator of a MBP; this makes it possible to develop a MBP from any stage- or age-classified matrix model.

Let $\mathbf{n}(t)$ denote the population vector, with expectation $\mathrm{E}(\mathbf{n}(t))$ and covariance matrix $\mathbf{G} = \mathrm{Cov}(\mathbf{n}(t))$. The population projection matrix is $\mathbf{A} = \mathbf{U} + \mathbf{F}$, where $\mathbf{U}$ contains transition probabilities and $\mathbf{F}$ contains fertilities.

Define the vectors giving the per-capita production of offspring of all stages (recall that there are s stages) by an individual of stage j as

$$\boldsymbol{\xi}_{\mathrm{U}}^{(j)} = \text{offspring produced by transitions} \qquad s \times 1 \qquad (1)$$

$$\boldsymbol{\xi}_{\mathrm{F}}^{(j)} = \text{offspring produced by reproduction} \qquad s \times 1 \qquad (2)$$

Both $\boldsymbol{\xi}_{\mathrm{U}}$ and $\boldsymbol{\xi}_{\mathrm{F}}$ are random variables. The total offspring prodution is $\boldsymbol{\xi}^{(j)} = \boldsymbol{\xi}_{\mathrm{U}}^{(j)} + \boldsymbol{\xi}_{\mathrm{F}}^{(j)}$.

The variability introduced by individual stochasticity depends on the covariance matrices of the $\boldsymbol{\xi}^{(j)}$. At this point, approximations of various kinds must be made. These are well known, and are described at length in (Caswell 2001, chapter 15), so we discuss them only briefly here.

We will rely on a set of approximations tailored to the demographic information that is reported in the overwhelming majority of ecological studies. In idealized situations, some of the approximations can be avoided, and we return to this issue in the Discussion.

A matrix population model reports the expected value of the $\boldsymbol{\xi}_{\mathrm{F}}^{(j)}$ in the fertility matrix $\mathbf{F}$. A complete branching process model for the population would require the complete joint probability distribution of the numbers of all types of offspring produced by an individual of each stage. Here, we are focused on the mean and variance of the populations, so the calculations require only the means and covariances of offspring production. Confronted with only the expected values, the appropriate approximation is to use a biologically reasonable distribution that can be parameterized by its mean. For species producing at most one offspring, the appropriate choice is the Bernoulli distribution. For species producing many offspring, the Poisson distribution is an appropriate choice. More interesting distributions (e.g. mixture distributions such as the negative binomial) can be parameterized only if data on the variance of reproduction are available. For example, Kendall and Wittmann (2010) compared the fit of a variety of probability distributions to data on vertebrate annual reproductive success, and found that a generalized Poisson distribution gave the best fit. This distribution requires an extra parameter that adjusts the relationship between the variance and the mean, so it cannot be fit from data on the mean alone. We encourage researchers to present full distributions, not only means, of stage-specific reproductive output, to enable use of such distributions.

In this paper we explore both Bernoulli and Poisson models for reproduction. Study of other distributions, and of distributions appropriate to taxonomic groups other than vertebrates, are interesting and significant unsolved research problems.

The second approximation concerns the covariance between $\boldsymbol{\xi}_{\mathrm{U}}^{(j)}$ and $\boldsymbol{\xi}_{\mathrm{F}}^{(j)}$; that is, between reproduction and transitions. The usual demographic models provide no information on this, simply adding together transitions and fertility to obtain **A**. In the absence of this information, the appropriate approximation is to treat transitions and fertility as independent, in which case the covariance matrix of offspring production is the sum of the covariances due to fertility and those due to transition. This assumption about independence of the outcomes of random events does not imply that the rates of survival and of reproduction, across i-states, are independent.

We write the covariance matrix among the stages as

$$\mathbf{C}^{(j)} = \mathrm{Cov}\left(\boldsymbol{\xi}_{\mathrm{U}}^{(j)}\right) + \mathrm{Cov}\left(\boldsymbol{\xi}_{\mathrm{F}}^{(j)}\right) \tag{3}$$

$$= \mathbf{C}_{\mathrm{U}}^{(j)} + \mathbf{C}_{\mathrm{F}}^{(j)} \tag{4}$$

We collect all this covariance information in a matrix **D** given by

$$\mathbf{D} = \begin{pmatrix} \mathrm{vec}\,\mathbf{C}^{(1)} & \cdots & \mathrm{vec}\,\mathbf{C}^{(s)} \end{pmatrix} \tag{5}$$

This information suffices to project the mean population vector E(**n**) and its covariance matrix **G**:

$$\begin{pmatrix} \mathrm{E}(\mathbf{n}) \\ \mathrm{vec}\,\mathbf{G} \end{pmatrix}(t+1) = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{D} & \mathbf{A}\otimes\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathrm{E}(\mathbf{n}) \\ \mathrm{vec}\,\mathbf{G} \end{pmatrix}(t) \tag{6}$$

(Pollard 1966, Caswell 2001). This is the extension of the matrix population model to incorporate demographic stochasticity. The key to the analysis is covariance matrices that appear in **D**, which we now derive.

Covariance due to transitions

The transition part of Eq. 4 is easy; if **U** is the transition probability matrix among the stages, then the distribution of offspring is multinomial, with a covariance matrix

$$\mathbf{C}_{\mathrm{U}}^{(j)} = \mathcal{D}\left(\mathbf{U}(:,j)\right) - \mathbf{U}(:,j)\mathbf{U}(:,j)^{\mathsf{T}} \tag{7}$$

Covariance due to fertility

The covariance matrix $\mathbf{C}_{\mathrm{F}}$ in Eq. 4 depends on the types of offspring produced and the probability distribution of the numbers produced. Here we consider two common cases: one in which a single offspring is produced, and one in which multiple offspring, of multiple types, are produced. Other reproductive modes can be constructed similarly.

Bernoulli reproduction

Many long-lived organisms, including whales, albatrosses, elephants and humans, typically produce a single offspring at a time. Suppose stages are numbered so that new offspring are stage 1. Then **F** contains expected fertilities f_j in the first row and zeros elsewhere, and the number of offspring is a Bernoulli random variable. The covariance matrix is

$$\mathbf{C}_{\mathrm{F}}^{(j)} = \begin{pmatrix} f_j(1-f_j) & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \tag{8}$$

Poisson reproduction

A simple model for reproduction in organisms that produce potentially many offspring, of many different types, is obtained by assuming that the number of offspring of type i produced by a parent of type j is a Poisson random variable with mean f_{ij}. The variance of the Poisson distribution is equal to the mean; thus, if the types of offspring are independent, then the covariance matrix satisfies

$$\mathbf{C}_{\mathrm{F}}^{(j)} = \begin{pmatrix} f_{1j} & & & \\ & f_{2j} & & \\ & & \ddots & \\ & & & f_{sj} \end{pmatrix} \tag{9}$$

Diffusion models

At the opposite extreme from the discrete-state, discrete-time and vector-valued multitype branching process models are diffusion processes, which are continuous-state, continuous-time, scalar-valued stochastic processes. There is a rich but mathematically difficult literature on diffusion models (Cox and Miller 1965, Karlin and Taylor 1981), which has been applied to a variety of models in population ecology (Lande et al. 2003) and population genetics (Kimura 1964, Nagylaki 1989).

Diffusion models can be derived in several ways. Start, quite generally, with a continuous-time Markov process for a continuous variable $x(t)$. If it can be solved, this original model produces a probability density function $\phi(x, t)$ of population size at time t. Such a solution may not be possible or known, and Bailey (1964) shows it can be approximated by a diffusion model. Considering the probability of a small change in x over a short time interval, and expanding the resulting change in $\phi(x, t)$ in a Taylor series, yields a diffusion equation for ϕ:

$$\frac{\partial\phi(x,t)}{\partial t} = -\frac{\partial\left[\mu(x)\phi(x,t)\right]}{\partial x} + \frac{1}{2}\frac{\partial^2\left[\nu(x)\phi(x,t)\right]}{\partial x^2} \tag{10}$$

Here the function $\mu(x)$ defines the growth of the mean, and $\nu(x)$ the growth of the variance, of x. In applications to

population growth, x is some measure of population size, and $\mu(x)$ and $\nu(x)$ are taken to be proportional to x (Bailey 1964). That is, population size influences the growth rate of both mean population and the variance. We will denote the proportionality constant relating x and $\nu(x)$ as σ_d^2.

The diffusion coefficients are obtained from the original model (which is being approximated as a diffusion) as

$$\mu(x)\Delta t \approx \mathrm{E}[\Delta x \mid x] \tag{11}$$

$$\begin{aligned} \nu(x)\Delta t &\approx \mathrm{Var}[\Delta x \mid x] \\ &= \sigma_d^2\, \mathrm{E}[x] \end{aligned} \tag{12}$$

where $\mathrm{E}[\Delta x|x]$ and $\mathrm{Var}[\Delta x|x]$ are obtained from the model being approximated.

Because we are focusing here on variance in the growth increment that is due to demographic stochasticity, we have followed standard practice (Lande et al. 2003) and written the coefficients in terms of σ_d^2, referred to as the demographic variance.

There are two main strengths of the diffusion approximation. First, the diffusion approximation is usually described by fewer parameters than is the full process. This simplifies the model, and the calculation of μ and ν from Eq. 11 and Eq. 12 provides valuable information about which aspects of the process are important to consider for describing the dynamics. Second, it is also possible to simulate realizations of the diffusion process, using $\mu(x)$ and $\nu(x)$ to add a random perturbation to the growth of the population (Vindenes et al. 2010, 2011).

Extinction probability

Analytical solutions, which are known for many diffusion processes (Karlin and Taylor 1981), provide information on the time required to reach a given state (first passage time). Extinction, in a population described by a diffusion model, is analyzed as just such a passage time problem. If N is the population size, then a natural extinction boundary exists at $N_e = 1$. The probability of extinction depends on the initial population size N_0, the growth rate $r = \log \lambda$ of the mean population size, and the demographic variance σ_d^2, according to

$$P_e = \begin{cases} \exp\left[\dfrac{-2r(N_0 - N_e)}{\sigma_d^2}\right] & r > 0 \\ 1 & r \le 0 \end{cases} \tag{13}$$

(Engen et al. 2005). For a specified value of $r > 0$ and a given initial population, extinction probability increases with increasing demographic variance.

Demographic variance from matrix models

We now connect the structure of multitype branching processes with the framework of diffusion models, by calculating the demographic variance σ_d^2 from the matrix model. To do so, the dimensionality of the MBP must be reduced to a scalar population size, in spite of the fact that population structure may be fluctuating. The key is to measure population size not by total numbers, but by the total reproductive value (Engen et al. (2005) for age-classified models, Vindenes et al. (2008) for stage-classified models). It is known that the total reproductive value grows exponentially regardless of the age or stage distribution of the population (Fisher 1930).

Begin with the matrix $\mathbf{A}$, with dominant eigenvalue λ and corresponding right and left eigenvalues $\mathbf{w}$ and $\mathbf{v}$, where $\mathbf{w}$ gives the stable stage distribution and $\mathbf{v}$ the reproductive value distribution. We scale the eigenvectors so that $\mathbf{w}$ sums to 1 and $\mathbf{w}^\mathsf{T}\mathbf{v} = 1$. The total reproductive value is $V(t) = \mathbf{v}^\mathsf{T}\mathbf{n}(t)$.

Define a vector $\mathbf{q}$ whose ith entry is

$$q_i = \mathbf{v}^\mathsf{T}\mathbf{C}^{(i)}\mathbf{v} \tag{14}$$

Then it can be shown that

$$\mathrm{Var}\left[V(t+1) \mid V(t)\right] = \mathbf{q}^\mathsf{T}\,\mathrm{E}\left[\mathbf{n}(t)\right] \tag{15}$$

(see section 'Derivations' for the derivation). The entries of $\mathbf{q}$ are the contributions to the variance from each stage, so the conditional variance in $V(t+1)$ depends on the stage distribution of the population. Following Engen et al. (2007) and Vindenes et al. (2008), we will use the stable stage distribution, so that $\mathbf{n}(t) \approx \mathbf{w}N$. Under this assumption, $\mathrm{E}[n_i] = w_i \mathrm{E}[V]$, so that Eq. 15 can be written

$$\mathrm{Var}\,[V(t+1) \mid V(t)] = \mathbf{q}^\mathsf{T}\mathbf{w}\ \mathrm{E}\,[V(t)] \tag{16}$$

Comparing this to Eq. 12, which defines a conditional variance in terms of σ_d^2 and an expected population, we obtain the demographic variance in terms of $\mathbf{w}$ and $\mathbf{q}$,

$$\sigma_d^2 = \mathbf{q}^\mathsf{T}\mathbf{w} \tag{17}$$

This result provides the demographic variance for any age- or stage-classified model from the eigenvectors of $\mathbf{A}$ and the variances and covariances of the stage-specific transition probabilities and reproductive outputs.

Partitioning the components of demographic variance

Because Eq. 4 expresses the covariances in terms of the matrices $\mathbf{U}$ and $\mathbf{F}$, the demographic variance can be partitioned into contributions from these life cycle components. The demographic variance (Eq. 17) is a weighted mean, over the stable stage distribution, of the variances of the offspring production of each stage, $\mathbf{v}^\mathsf{T}\mathbf{C}^{(i)}\mathbf{v}$. Each of those variances contains a term due to transitions and a term due to reproduction. Thus

$$\sigma_d^2 = \sum_{i=1}^{s} \underbrace{w_i \mathbf{v}^\mathsf{T}\mathbf{C}_\mathrm{U}^{(i)}\mathbf{v}}_{\text{transitions}} + \sum_{i=1}^{s} \underbrace{w_i \mathbf{v}^\mathsf{T}\mathbf{C}_\mathrm{F}^{(i)}\mathbf{v}}_{\text{reproduction}} \tag{18}$$

Sensitivity analysis of the demographic variance

We turn now to the problem of the sensitivity of σ_d^2 to changes in parameters affecting the means and covariances of the vital rates. This calculation is a systematic application of the matrix calculus approach to sensitivity analysis; see (Caswell 2007, 2008, 2009, 2012). Derivations of all results are given in the section 'Derivations'.

Differentiating Eq. 17 and applying the vec operator gives

$$d\sigma_d^2 = \mathbf{q}^\mathsf{T} d\mathbf{w} + \mathbf{w}^\mathsf{T} d\mathbf{q} \tag{19}$$

The analysis proceeds by determining the differentials $d\mathbf{w}$ and $d\mathbf{q}$; the latter will require differentials of both $\mathbf{v}$ and of covariances.

Derivatives of w

From Caswell (2008) we have $d\mathbf{w}$ subject to the constraint $\mathbf{1}_s^\mathsf{T}\mathbf{w} = 1$,

$$d\mathbf{w} = \left(\lambda\mathbf{I} - \mathbf{A} + \mathbf{w}\mathbf{1}_s^\mathsf{T}\mathbf{A}\right)^{-1} \times\left[\left(\mathbf{w}\otimes\mathbf{I}\right) - \left(\mathbf{w}^\mathsf{T}\otimes\mathbf{w}\mathbf{1}_s^\mathsf{T}\right)\right] d\,\mathrm{vec}\,\mathbf{A} \tag{20}$$

Derivatives of q

To differentiate $\mathbf{q}$, start by rewriting Eq. 45 as

$$\mathbf{q} = \sum_i \mathbf{e}_i \mathbf{v}^\mathsf{T}\mathbf{C}^{(i)}\mathbf{v} \tag{21}$$

Differentiating and applying the vec operator gives

$$d\mathbf{q} = \sum_i \left(\mathbf{v}^\mathsf{T}\mathbf{C}^{(i)\mathsf{T}}\otimes\mathbf{e}_i\right) d\mathbf{v} + \sum_i \left(\mathbf{v}^\mathsf{T}\otimes\mathbf{e}_i\mathbf{v}^\mathsf{T}\right) d\,\mathrm{vec}\,\mathbf{C}^{(i)} + \sum_i \mathbf{e}_i\mathbf{v}^\mathsf{T}\mathbf{C}^{(i)} d\mathbf{v} \tag{22}$$

For this, we need two pieces: the differential of the left eigenvector $\mathbf{v}$ and the differential of the covariance matrices $\mathbf{C}_i$. We address these in turn.

Derivatives of v

An extension of the methods in Caswell (2008) yields the differential of the $\mathbf{v}$ scaled so that $\mathbf{v}^\mathsf{T}\mathbf{w} = 1$:

$$d\mathbf{v} = \left(\lambda\mathbf{I} - \mathbf{A}^\mathsf{T} + \lambda\mathbf{v}\mathbf{w}^\mathsf{T}\right)^{-1}\left[\left(\mathbf{I}\otimes\mathbf{v}^\mathsf{T}\right) - \left(\mathbf{v}\mathbf{w}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T}\right)\right] d\,\mathrm{vec}\,\mathbf{A} - \lambda\left(\lambda\mathbf{I} - \mathbf{A}^\mathsf{T} + \lambda\mathbf{v}\mathbf{w}^\mathsf{T}\right)^{-1}\left(\mathbf{v}\otimes\mathbf{v}^\mathsf{T}\right) d\mathbf{w} \tag{23}$$

where $d\mathbf{w}$ is given by Eq. 20.

Derivatives of the covariance matrices C(i)

The derivatives of the covariance matrices $\mathbf{C}^{(i)} = \mathbf{C}_\mathrm{U}^{(i)} + \mathbf{C}_\mathrm{F}^{(i)}$ include a number of possibilities, because the covariance matrix will change due to changes in $\mathbf{U}$ and $\mathbf{F}$, as well as due to changes that may be incorporated in parametric form.

Covariances due to transitions

We start with the covariances due to transitions, which are given by Eq. 7. For each parental stage i, we have

$$d\,\mathrm{vec}\,\mathbf{C}_\mathrm{U}^{(i)} = \left\{\mathbf{Z}\left(1_s\otimes\mathbf{I}\right) - \left(\mathbf{U}(:,i)\otimes\mathbf{I}\right) - \left(\mathbf{I}\otimes\mathbf{U}(:,i)\right)\right\} \times\left(\mathbf{e}_i^\mathsf{T}\otimes\mathbf{I}\right) d\,\mathrm{vec}\,\mathbf{U} \tag{24}$$

where $\mathbf{Z} = \mathcal{D}(\mathrm{vec}\,\mathbf{I})$.

Covariances due to reproduction

In both Bernoulli and Poisson reproduction, the covariance matrix $\mathbf{C}^{(j)}$ is obtained from the fertility matrix $\mathbf{F}$. For the moment, supress the indication of the parent stage i, and let $\mathbf{f}$ denote column j of $\mathbf{F}$.

Writing $\mathbf{C}_\mathrm{F}^{(i)}$ as a function of $\mathbf{F}$, and thus of $\mathbf{f}$, it can be shown that

$$d\,\mathrm{vec}\,\mathbf{C}_\mathrm{F}^{(j)} = \begin{cases} \mathbf{Z}(\mathbf{I}\otimes 1)\mathcal{D}(1-2\mathbf{f})\left(\mathbf{e}_j\otimes\mathbf{I}\right) d\,\mathrm{vec}\,\mathbf{F} & \text{Bernoulli} \\ \mathbf{Z}(\mathbf{I}\otimes 1)\left(\mathbf{e}_j\otimes\mathbf{I}\right) d\,\mathrm{vec}\,\mathbf{F} & \text{Poisson} \end{cases} \tag{25}$$

where $\mathbf{Z} = \mathcal{D}(\mathrm{vec}\,\mathbf{I})$.

The Bernoulli and Poisson models for reproduction are convenient because they permit the calculation of variances from the mean fertility matrix $\mathbf{F}$ routinely reported in demographic studies. But if data were available on the mean and the variance of reproduction by each stage, one could use these empirical measures, and calculate the sensitivities using the derivatives of the matrices $\mathbf{C}_\mathrm{F}^{(i)}$ with respect to the variance of fertility.

Sensitivity of extinction probability

The extinction probability in Eq. 13 depends on the initial population size and the extinction threshold. Since population size is a continuous variable in the diffusion approximation, extinction must be interpreted as quasi-extinction relative to this threshold.

Consider a population of initial size (total reproductive value) V_0, and quasi-extinction threshold of V_e, and write $\tilde{V} = V_0 - V_e$. Assume that $r = \log\lambda$ is greater than 0 (otherwise extinction is certain); the probability of extinction is

$$P_e = \exp\left(\frac{-2r\tilde{V}}{\sigma_d^2}\right) \tag{26}$$

(cf. Eq. 13). A parameter vector $\boldsymbol{\theta}$ will, in general, affect both r and σ_d^2. Differentiating P_e and simplifying yields

$$\frac{1}{P_e}\frac{dP_e}{d\boldsymbol{\theta}^\mathsf{T}} = \frac{-2\tilde{V}}{\sigma_d^2}\frac{dr}{d\boldsymbol{\theta}^\mathsf{T}} \quad \frac{2r\tilde{V}}{(\sigma_d^2)^2}\frac{d\sigma_d^2}{d\boldsymbol{\theta}^\mathsf{T}} \tag{27}$$

The first of these terms shows how increases in r reduce extinction probability; the sensitivity of r depends only on the mean matrix $\mathbf{A}$ and can be calculated by well known formulae (Caswell 2001). The second term shows how increases in σ_d^2 increase the extinction probability; the results in previous sections for the sensitivity of σ_d^2 make it straightforward to compute this term.

Implementation: a protocol for sensitivity analysis

We now have all the information necessary to calculate the sensitivity of the demographic variance to any set of parameters that affect **U** and/or **F**. Let $\boldsymbol{\theta}$ be such a vector of parameters. The following protocol integrates the results from the previous sections. Starting with the matrices **U** and **F**, and the dominant eigenvalue of **A** and its eigenvectors **w** and **v**,

1. calculate $d\mathbf{w}$ from Eq. 20,
2. calculate $d\mathbf{v}$ from Eq. 23,
3. calculate $d\mathrm{vec}\,\mathbf{C}_{\mathrm{U}}^{(i)}$, for $i = 1, \ldots, s$, from Eq. 24,
4. calculate $d\mathrm{vec}\,\mathbf{C}_{\mathrm{F}}^{(i)}$, for $i = 1, \ldots, s$, from Eq. 25,
5. calculate $d\mathbf{q}$ from Eq. 22, using $d\mathbf{v}$ and the $d\mathrm{vec}\,\mathbf{C}^{(i)}$, and finally
6. calculate $d\sigma_d^2$ from Eq. 19 using $d\mathbf{q}$ and $d\mathbf{w}$.

The steps apply to any parameter vector θ. To obtain the derivative of σ_d^2 to $\boldsymbol{\theta}$, replace the differentials $d\mathrm{vec}\,\mathbf{U}$ and $d\mathrm{vec}\,\mathbf{F}$ with the derivatives $\frac{d\,\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}}$ and $\frac{d\,\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}}$, respectively. These derivatives will, of course, be unique to the choice of parameters, the structure of **U** and **F**, and their dependence on $\boldsymbol{\theta}$.

The expressions for each of these steps may appear complicated, but the matrices involved in each are either simple constants, composed of identity matrices and vectors of ones, or are obtained directly from the population projection matrix **A** or its components **U** and **F**.

Examples

The results presented here apply to a wide range of life cycles (age-classified, stage-classified, multistate) and reproductive tactics (single offspring, multiple offspring, single or multiple types of offspring). In this section we explore a few examples to reveal some intruiging patterns. We will calculate the values of the demographic variance and the contributions to demographic variance from the (co)variances of transitions and of fertility. We will also present the sensitivity of σ_d^2 to fertility and to mortality, and do so for both age-classified and stage-classified examples, including both Bernoulli- and Poisson-distributed numbers of offspring.

Our intention is not a rigorous comparative study, but some of the patterns that begin to appear suggest that such a study would be worthwhile.

Age-classified human populations

The large collections of age-specific mortality and fertility data for human populations, sometimes over long time periods, make them an attractive option for studying demographic stochasticity in age-classified populations. While the details would vary in other species, the human life cycle is typical of a long-lived, slowly maturing, monovular (ignoring uncommon multiple births) species.

We will present results for nine populations studied, in a different but related context, by van Daalen and Caswell (2017). These include time series for the Netherlands (1950–2009), Sweden (1891–2010), and Japan (1947–2009), as well as two hunter–gatherer populations (the Ache of subtropical Paraguay, Gurven and Kaplan 2007, Hill and Hurtado 1996, and the Hadza of the Tanzanian savanna, Blurton Jones 2011), and the Hutterites of North America. The Netherlands, Sweden and Japan show patterns over time that are typical of developed countries (the demographic transition), in which reductions in mortality lead to increases in life expectancy, followed some time later by reductions in fertility. The hunter–gatherer populations have higher mortality, lower life expectancy and higher fertility than the developed countries. The Hutterites, an Anabaptist religious community in the United States and Canada, make an interesting comparison. They had the highest total fertility recorded for any human population (Eaton and Mayer 1953), but had (or were assumed to have) experienced mortality rates similar to those of the US around 1950.

Female life table data for the Netherlands, Sweden and Japan were obtained from the Human Mortality Database (Human Mortality Database 2013) and the Human Fertility Database (Human Fertility Database 2013). Data for the Ache were obtained from Gurven and Kaplan (2007) and Hill and Hurtado (1996), and for the Hadza from Blurton Jones (2011). The fertility and mortality schedules for the Hutterites were taken from Eaton and Mayer (1953). Fertility was divided by 2 to count female offspring per female.

Analysis

Matrices **U** and **F** were created for each population in each year. Fertility was treated as a Bernoulli process. Demographic variance was calculated from Eq. 17, including the contributions from survival and fertility using Eq. 7 and Eq. 8 for the covariance matrices and Eq. 18 for the decomposition. The sensitivity of σ_d^2 to age-specific mortality and fertility was calculated from Eq. 19.

In addition to σ_d^2, its components, and its sensitivity to the life history, other demographic properties were computed. Population growth rate λ was calculated as the dominant eigenvalue of **A**. The net reproductive rate R_0 was calculated as the dominant eigenvalue of the next generation matrix $\mathbf{F}(\mathbf{I} - \mathbf{U})^{-1}$. The mean longevity (life expectancy) and the variance in longevity were computed from the fundamental matrix (e.g. Eq. 38 and 41 of Caswell (2009)). The variance in lifetime reproductive output was taken from van Daalen and Caswell (2017).

Life expectancy, R_0, and λ are statistics describing the mean consequences of the life cycle, in terms of mortality, fertility, and their interaction. The variance in longevity and the variance in lifetime reproductive output describe the results of individual stochasticity. One of our goals is to see how these quantities are related to σ_d^2, which describes the result of demographic stochasticity.

Results: σ_d^2*, contributions and sensitivity*

Figure 1 shows time series results for Sweden from 1891 to 2010. Over this time period, life expectancy increased by 58%, from 53 to 84 years, while lifetime reproductive output decreased by 34%, from 1.5 to 0.99 female offspring per female. The demographic variance σ_d^2 increased, with fluctuations (Fig. 1a) almost doubled, from 0.05 to 0.09. The contribution of fertility to σ_d^2 increased from 70% to 99% (Fig. 1b). In 1891, the contribution of infant and child mortality to σ_d^2 was visible, but it had disappeared by 2010 (Fig. 1c–d).

Corresponding time series for the Netherlands and Japan are shown in the Supplementary material Appendix 1, and the values at the beginning and end of the time series are shown in Table 1. In both countries, an increase in life expectancy and a reduction in fertility was accompanied by an increase in σ_d^2 and a shift to almost complete domination of σ_d^2 by contributions from fertility.

The sensitivity of σ_d^2 to mortality (for Sweden in 1891 and 2010; Fig. 2) is positive for ages up to about 40, then becomes negative, and goes to zero at old ages. The sensitivity to fertility is negative over most of the life span. Sensitivities of σ_d^2 to both mortality and fertility increased in absolute value from 1891 to 2010. Very similar patterns appear for the Netherlands (Supplementary material Appendix 1 Fig. A2) and Japan (Supplementary material Appendix 1 Fig. A4).

Figure 3 and 4 show the contributions to σ_d^2 and the sensitivity of σ_d^2 to mortality and fertility for the Ache and Hadza hunter–gatherer populations. The values of σ_d^2 (0.04 and 0.06) are in the same range of the values for the early developed countries. However, the contributions from fertility and from survival are almost exactly evenly matched (Table 1).

The Hutterite population, with its extremely high fertility, but more contemporary mortality, has a low value of σ_d^2 (0.023), almost all of which is contributed by fertility.

A stage-classified example

Calathea ovandensis is an herbaceous understory plant native to Central America. Horvitz and Schemske (1995) developed a size-classified model with eight stages (seeds, seedlings, juveniles and pre-reproductives, and small, medium, large

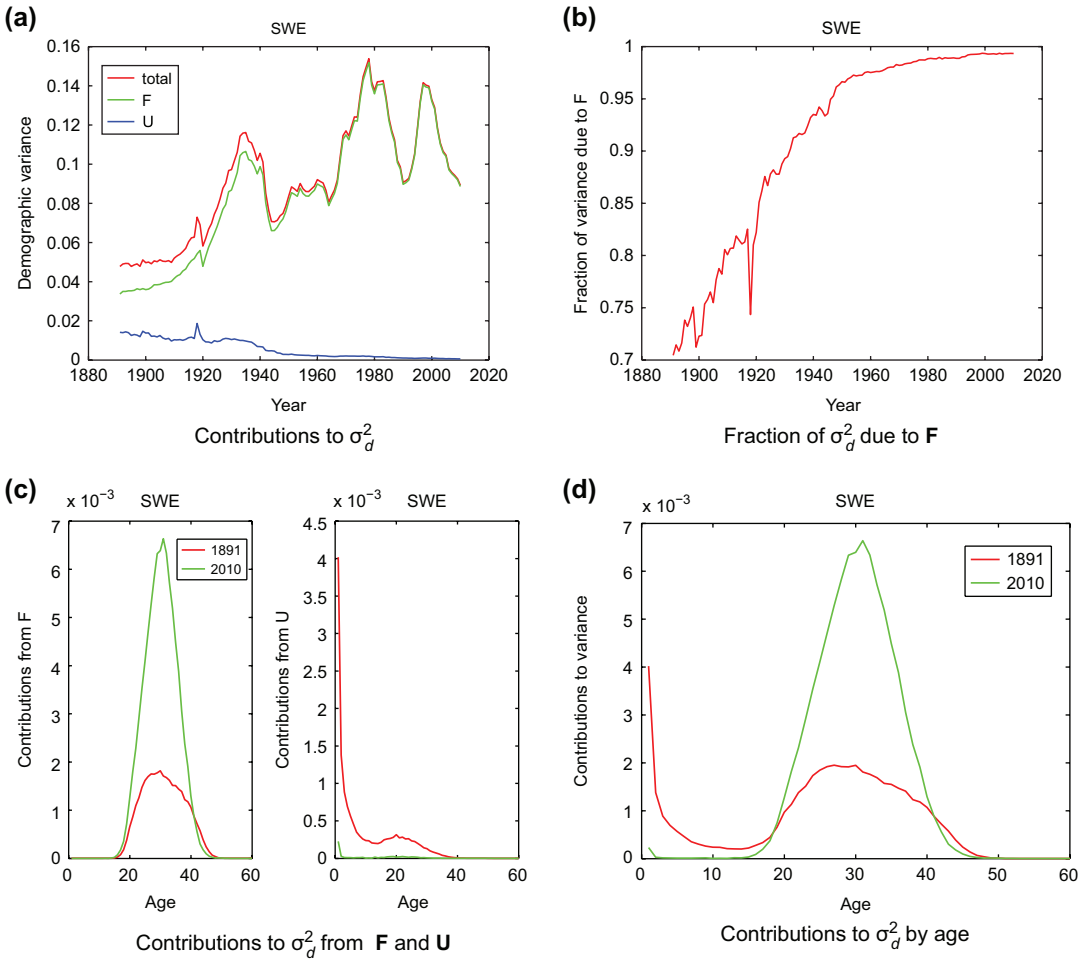

Figure 1. Demographic variance for Sweden 1891–2010. (a) Contributions to σ_d^2 from survival (**U**) and from fertility (**F**). (b) The fraction of the overall demographic variance due to fertility. (c) The age pattern of contributions from fertility and survival, for 1891 and 2010. (d) The age pattern of contributions to demographic variance, combining both fertility and survival.

Table 1. Summary results for human populations: the demographic variance σ_d^2 and its components $\sigma_d^2 F$ and $\sigma_d^2 U$ due to fertility and survival, the population growth rate λ, the net reproductive rate R_0, the mean $E(\eta_1)$ and variance $V(\eta_1)$ of longevity at birth, and the variance $V(\rho)$ in lifetime reproductive output due to individual stochasticity (van Daalen and Caswell 2017).

Population	σ_d^2	$\sigma_d^2 F$ (%)	$\sigma_d^2 U$ (%)	λ	R_0	$E(\eta_1)$	$V(\eta_1)$	$V(\rho)$
SWE 1891	0.048	0.034 (70.4)	0.014 (29.6)	1.013	1.50	53.0	944	5.60
SWE 2010	0.089	0.089 (99.3)	0.001 (0.7)	0.999	0.986	84.0	160	1.79
JPN 1947	0.045	0.033 (73.7)	0.012 (26.3)	1.019	1.75	54.2	830	6.10
JPN 2009	0.171	0.170 (99.4)	0.001 (0.6)	0.987	0.67	86.9	178	1.26
NLD 1950	0.051	0.049 (95.4)	0.002 (4.6)	1.013	1.482	73.1	331	2.91
NLD 2009	0.102	0.101 (99.3)	0.001 (0.7)	0.996	0.888	83.1	179	1.61
Ache	0.042	0.021 (49.5)	0.021 (50.5)	1.027	2.238	38.0	812	17.23
Hadza	0.057	0.027 (47.4)	0.031 (54.4)	1.016	1.566	34.6	989	11.3
Hutterites	0.023	0.021 (91.3)	0.002 (8.7)	1.044	3.763	70.0	433	8.58

and extra-large reproductives). Based on combined data over a range of sites and several years, they estimated a population projection matrix

$$\mathbf{A} = \begin{pmatrix} 0.498 & 0 & 0.593 & 7.139 & 14.271 & 24.695 & 34.903 & 40.544 \\ 0.097 & 0.011 & 0.019 & 0 & 0 & 0 & 0 & 0 \\ 0.004 & 0.044 & 0.338 & 0.070 & 0.025 & 0.007 & 0.009 & 0 \\ 0 & 0.001 & 0.136 & 0.429 & 0.174 & 0.097 & 0.043 & 0.043 \\ 0 & 0 & 0.036 & 0.384 & 0.603 & 0.426 & 0.299 & 0.217 \\ 0 & 0 & 0.002 & 0.025 & 0.113 & 0.239 & 0.171 & 0.283 \\ 0 & 0 & 0 & 0.009 & 0.027 & 0.155 & 0.325 & 0.196 \\ 0 & 0 & 0 & 0.003 & 0.006 & 0.045 & 0.128 & 0.239 \end{pmatrix} \tag{28}$$

(Horvitz and Schemske 1995).

Using this matrix, and assuming a Poisson distribution for fertility, we computed σ_d^2 using Eq. 17, and the contributions from the fertility matrix $\mathbf{F}$ and from the matrix $\mathbf{U}$ that includes transitions and survival. The resulting demographic variance for *Calathea* is

$$\sigma_d^2 = 3.07 \tag{29}$$

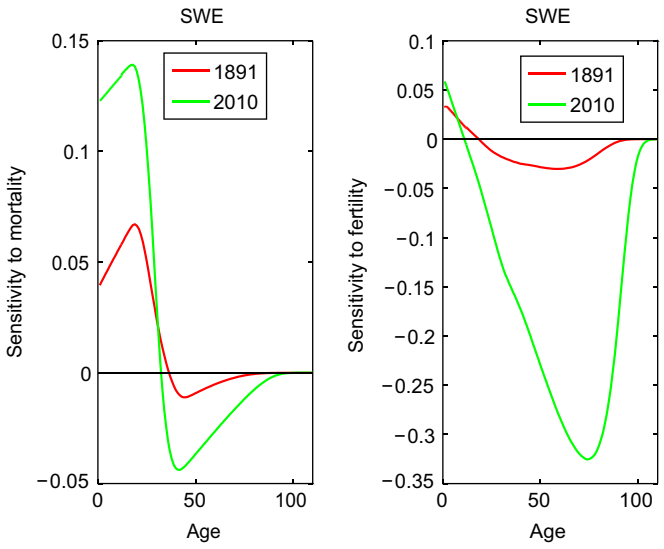

Figure 2. Sensitivity of the demographic variance σ_d^2 to age-specific mortality and fertility, for Sweden in 1891 and 2010.

99.8% of which was contributed by transitions and survival; only 0.2% was due to stochasticity in fertility.

We computed the sensitivity of σ_d^2 to changes in stage-specific fertility and mortality using (19). For the latter calculation, the derivatives of $\mathbf{U}$ to mortality were obtained by factoring $\mathbf{U}$,

$$\mathbf{U} = \mathbf{G}\boldsymbol{\Sigma} \tag{30}$$

where $\boldsymbol{\Sigma}$ has survival probabilities on the diagonal and $\mathbf{G}$ contains transition probabilities conditional on survival.

$$\frac{d\,\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\mu}^{\mathsf{T}}} = -\left(\mathbf{I}_s \otimes \mathbf{G}\right)\mathcal{D}(\mathrm{vec}\,\mathbf{I}_s)\left(\mathbf{I}_s \otimes \mathbf{1}_s\right)\boldsymbol{\Sigma} \tag{31}$$

The stage-specific contribution and sensitivity results are shown in Fig. 6. The largest contributions come from transitions and survival of stages 3–5. The fertility contributions are tiny, but of them the largest is from stage 5. The sensitivity patterns are quite different from those of the human populations. The sensitivity of σ_d^2 to mortality is negative in the smallest stage, and becomes positive later. The sensitivity to fertility is largest in stage 1 (which cannot reproduce), and declines across the larger stages.

Discussion

This methodological study has had several specific goals. The first and most general is to link matrix population models, multitype branching processes, and diffusion models as an approach to the study of demographic stochasticity. The second is to use the linkage to develop a sensitivity analysis of the demographic variance to parameters affecting any part[s] of the life cycle. The third is to develop methods to quantify the contributions to the demographic variance of fertility and transitions, throughout the life cycle. And finally, we set out to apply the methods to a variety of animal and plant population models as a first exploration of the patterns that emerge.

The diffusion approximation omits all the details; it is constructed from nothing more than a mean and a variance. That simplicity makes it a powerful technique because it is able

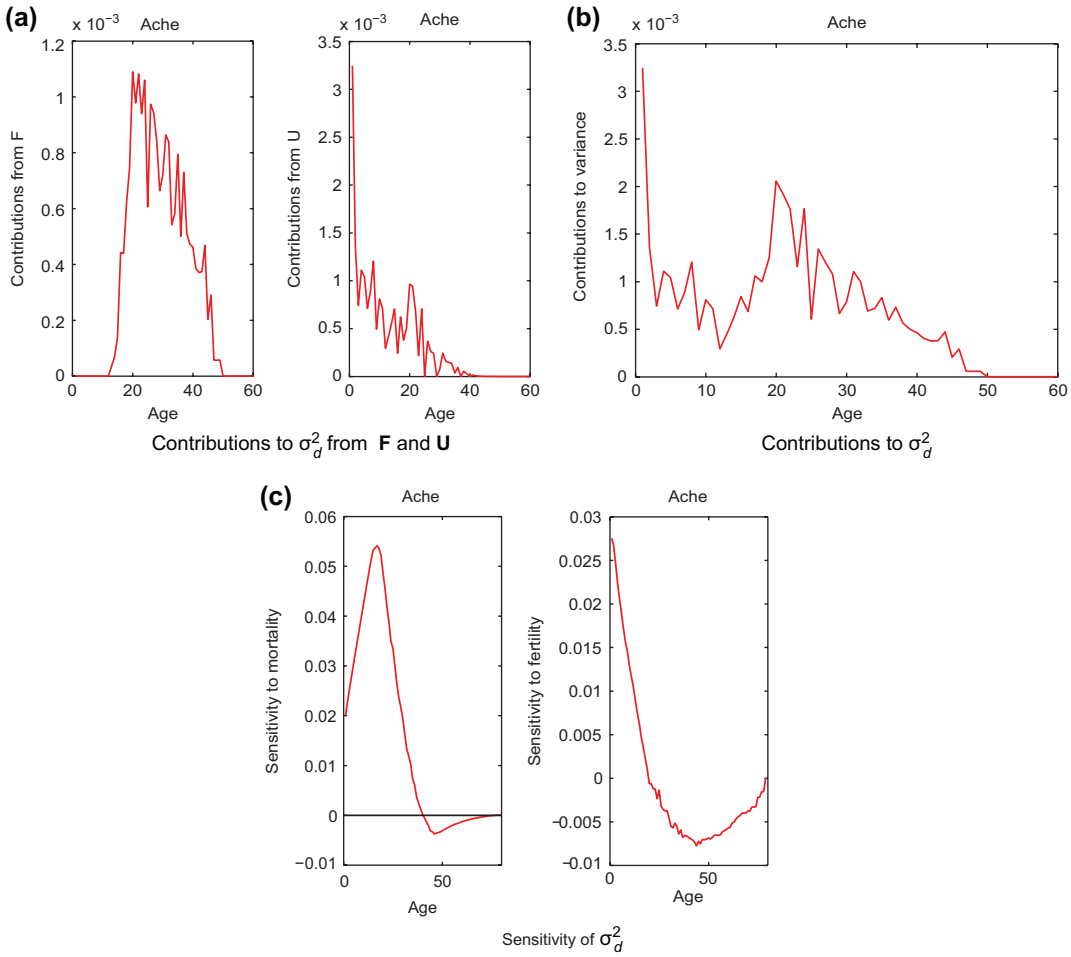

Figure 3. Demographic variance for Ache hunter–gatherers. (a) Contributions to σ_d^2 from age-specific survival (**U**) and from age-specific fertility (**F**). (b) Age-specific contributions to σ_d^2 combining fertility and survival. (c) The sensitivity of demographic variance to age-specific mortality and fertility.

to derive a great deal of stochastic information (e.g. extinction probability) from such a minimal specification. But that power is also its weakness, because the diffusion approximation obscures all the demographic and life history information that underlies population dynamics. It contains no stages, no transitions, no growth, no mortality schedules, no maturation, no metamorphosis, no size-dependent fertility, and thus no heterogeneity of any kind. As a result, interpretation of the results in terms of the life history traits that we know and love is impossible. For that reason, results that connect the diffusion approximation to age- and stage-structured models that do contain explicit life history information are so valuable.

Our results make it possible to compute the demographic variance from any age- or stage-classified population projection matrix and demonstrate the link between matrix population models, multitype branching processes, and diffusion models. Applications based on the diffusion approximation, including conservation biology (Lande et al. 2003), population dynamics (Lande et al. 2003, Engen et al. 2007, 2009, Vindenes et al. 2008), extinction (Lande 1993, Engen et al. 2005), and fixation probability (Vindenes et al. 2010, Cvijović et al. 2015) are now directly applicable to the huge accumulation of published matrix population models (over 7200 projection matrices, for over 950 species of plants and animals; see Salguero-Gómez et al. (2015, 2016)). It also makes it possible to use powerful matrix calculus methods to carry out sensitivity analyses.

Future research using comparative analyses will reveal patterns across different kinds of heterogeneity and life histories. For comparative analyses to be revealing, sufficient variation must exist in the variables to be analyzed. For these analyses, we note that among the small set of age-classified human populations, σ_d^2 varies by almost four-fold. The size-classified plant population extends that range to 68-fold. Among the human populations, fertility contributions range from about 50% to almost 100% of σ_d^2. Again, the plant population is extremely different: fertility contributes only 0.02% of σ_d^2.

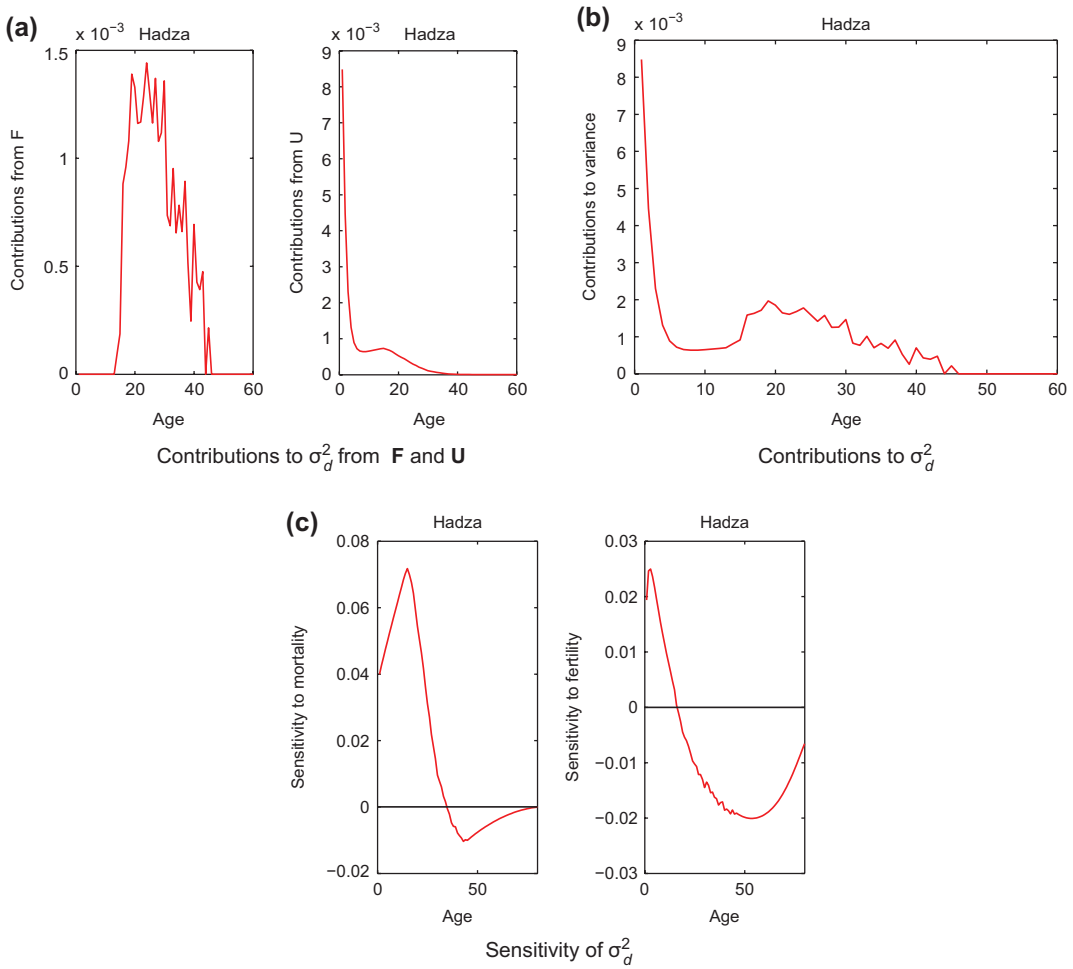

Figure 4. Demographic variance for Hadza hunter–gatherers. (a) Contributions to σ_d^2 from age-specific survival (**U**) and from age-specific fertility (**F**). (b) Age-specific contributions to σ_d^2 combining fertility and survival. (c) The sensitivity of demographic variance to age-specific mortality and fertility.

The empirical examples we included here are only a beginning, but we already see some interesting patterns.

- In human populations where extensive demographic data are available over a long time period, demographic variance has changed markedly over time, with a general increase in Sweden and Japan, while in the Netherlands a large increase in the 70s has been followed by a decline since the peak in the 80s.
- Among these human populations there has been a dramatic shift in contributions to σ_d^2 from survival and fertility. In populations with lower survival, the contributions from survival and fertility are roughly equal (especially in hunter–gatherer populations with life expectancy on the order of 35 years). As survival, and hence life expectancy, increases, the contributions shift to almost complete dominance by fertility in contemporary developed countries.
- The sensitivity of demographic variance to changes in mortality is consistently positive for early ages, and negative at older ages. The sensitivity to changes in fertility may be slightly positive at youngest ages (where fertility is zero anyway), but becomes negative during the reproductive years. These patterns are sufficiently consistent that one could conjecture that they are a common feature of age-classified populations.
- These comparisons of demographic variance across human populations cover a range of conditions that would be regarded as significant environmental variation in any mammal species (life expectancy varying by a factor of 2.5, lifetime reproductive output by a factor of 5). The resulting patterns differ from the interspecific comparisons among mammals by Vindenes and Engen (2017), who found a negative relationship between demographic variance and life expectancy. The explanation, as usual for comparative studies, will undoubtedly reflect the details of life history variation among, on the one hand, populations of one species, and populations of various related species.

The stage-classified model for *Calathea ovandensis* reveals quite different patterns. The demographic variance itself is 1–2 orders of magnitude larger than the values for the human

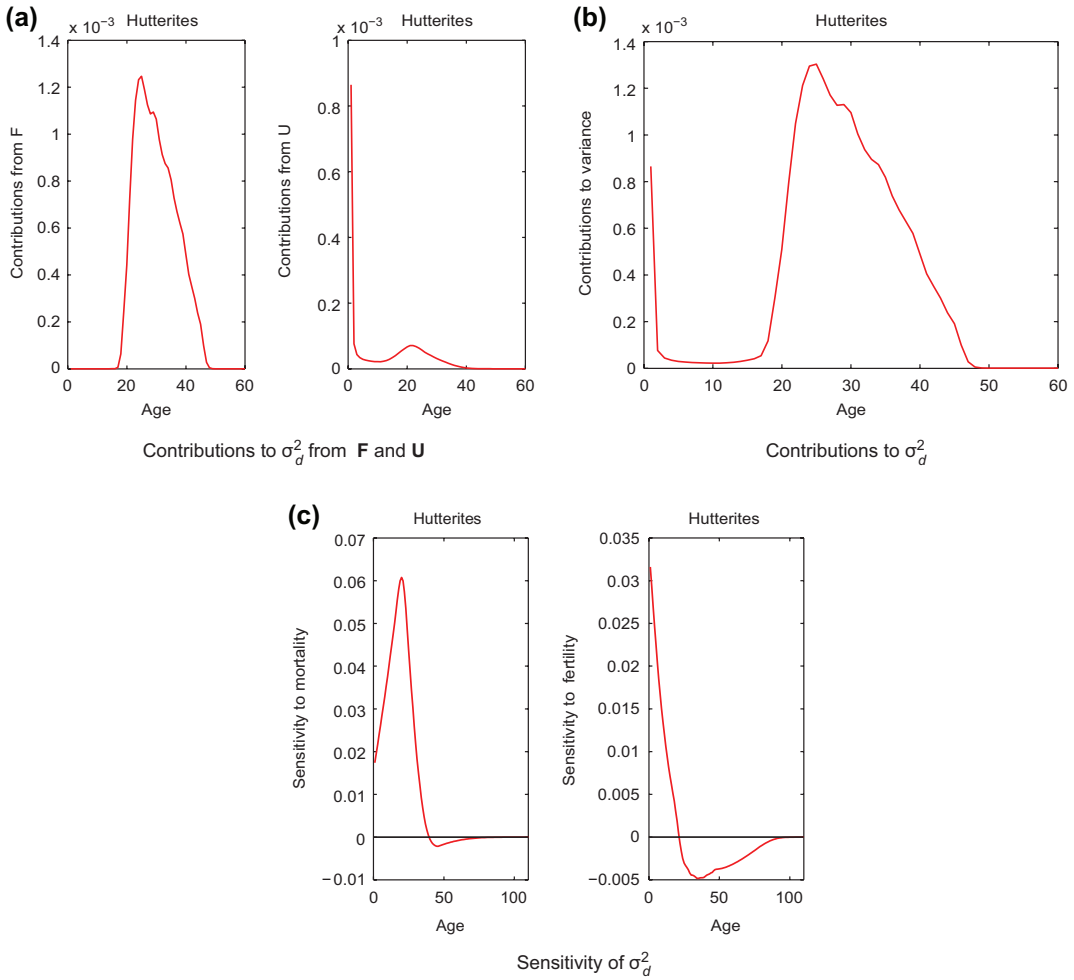

Figure 5. Demographic variance for Hutterites. (a) Contributions to σ_d^2 from age-specific survival (**U**) and from age-specific fertility (**F**). (b) Age-specific contributions to σ_d^2 combining fertility and survival. (c) The sensitivity of demographic variance to age-specific mortality and fertility.

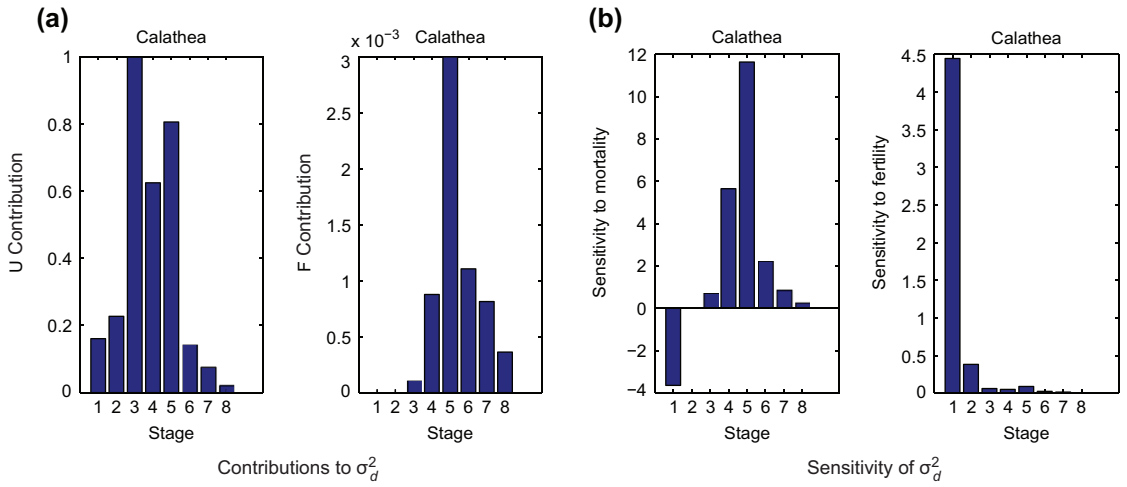

Figure 6. (a) The contributions to demographic variance for *Calathea ovandensis* from stage-specific fertility and stage-specific transitions and survival. (b) The sensitivity of demographic variance to stage-specific fertility and stage-specific mortality.

populations and is almost completely due to variance in survival and transitions, perhaps because the size-classified model offers so many more pathways through the life cycle and because the life history combines high fertility and low seedling survival.

The contributions from each stage or age class to the demographic variance depend on the expected contribution from each individual as well as the proportion of that stage in the stable distribution. Thus, stages where the variance contribution from each individual is small, for instance because the survival probability is high, may still have a large impact on the demographic variance if they constitute a high proportion of the stable distribution. If the survival probability approaches zero or one, the contribution to demographic stochasticity will approach zero even for stages constituting a high proportion of the population.

Future research directions

The methods we have introduced suggest a number of future research directions. We describe a few of those here.

Our examples have focused on age and size, two of the most important and most frequently studied sources of demographic heterogeneity. Other sources of heterogeneity will also affect the demographic variance (e.g. spatial structure, morphological or behavioral trait variation, genetic variation). These dimensions of heterogeneity can be incorporated into multistate or hyper-state matrix models (Caswell 2014, Roth and Caswell 2016, Caswell et al. 2017) and our analyses applied to them, to explore the effects of these sources. The analyses we present here can be applied when integral projection models are treated as a source of matrices by discretizing the state variable (Ellner et al. 2016). This will open new analyses of more fine-scaled variation in the contributions to demographic variance across continuous traits.

It is inviting to use model life tables (Coale and Demeny 1983, Barlow and Boveng 1991) or parametric survival and fertility schedules (Brass 1960, Siler 1979) to explore age-classified life cycles. Relating the demographic variance to the pace and shape of life histories (Baudisch 2011) might reveal new patterns in age-classified populations operating on different time scales. In the same vein, a systematic study of the consequences of mixture models for fertility (e.g. the negative binomial as a gamma-distributed mixture of Poisson distributions) would be a valuable application.

The diffusion approximation and the demographic variance σ_d^2 reveal the population consequences of stochastic individual events. The relation between σ_d^2 and measures of individual stochasticity warrants further investigation. Our results in Table 1 show that it is possible for σ_d^2 to increase while individual variance in longevity and lifetime reproductive output are declining. Identifying the conditions under which this happens will require extensive new analyses.

The data required to construct these models consists of the means and covariances of the production of all types of individuals, by all types of individuals. The means are precisely the entries of the standard population projection matrix. The only extra information required is the set of covariances. Because the transitions of existing individuals follow multinomial distributions, their covariances can be calculated directly from the means. The covariances for fertility are not part of the usual presentation of a matrix population model, and for that reason we have shown how to proceed from assuming a distribution (Bernoulli or Poisson) for fertility. But a sufficiently detailed individual data set would include information from which these could be estimated empirically. We encourage researchers to present data on variances in reproductive output as well as means, when available.

Derivations

In this section, we collect the derivations of the major results of analysis of the matrix model in terms of the demographic variance σ_d^2.

Derivation: demographic variance

We define λ as the dominant eigenvalue of $\mathbf{A}$, and $\mathbf{w}$ and $\mathbf{v}$ as the right and left eigenvectors corresponding to λ, scaled so that eigenvectors so that $\mathbf{w}$ sums to 1 and $\mathbf{w}^\mathsf{T}\mathbf{v}=1$. We note that, from the properties of the Kronecker product,

$$(\mathbf{v}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T})(\mathbf{A}\otimes\mathbf{A})=\lambda^2(\mathbf{A}\otimes\mathbf{A}) \tag{32}$$

The total reproductive value is

$$V(t)=\mathbf{v}^\mathsf{T}\mathbf{n}(t) \tag{33}$$

with expectation and variance

$$\mathrm{E}\left[V(t)\right]=\mathbf{v}^\mathsf{T}\mathrm{E}\left[\mathbf{n}(t)\right] \tag{34}$$

$$\mathrm{Var}\left[V(t)\right]=\mathbf{v}^\mathsf{T}\mathbf{G}(t)\mathbf{v} \tag{35}$$

$$=(\mathbf{v}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T})\,\mathrm{vec}\,\mathbf{G}(t) \tag{36}$$

The expectation of $V(t)$ grows exponentially at the rate λ:

$$\mathrm{E}\left[V(t+1)\right]=\mathbf{v}^\mathsf{T}\mathbf{A}\mathrm{E}\left[\mathbf{n}(t)\right] \tag{37}$$

$$=\lambda\mathrm{E}\left[V(t)\right] \tag{38}$$

The variance of $V(t)$ grows asymptotically at the rate λ^2,

$$\mathrm{Var}\left[V(t+1)\right]=(\mathbf{v}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T})\,\mathrm{vec}\,\mathbf{G}(t+1) \tag{39}$$

$$=(\mathbf{v}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T})\Big\{\left[\mathbf{D}E\left[\mathbf{n}(t)\right]\right]+(\mathbf{A}\otimes\mathbf{A})\,\mathrm{vec}\,\mathbf{G}(t)\Big\} \tag{40}$$

$$=(\mathbf{v}^\mathsf{T}\otimes\mathbf{v}^\mathsf{T})\mathbf{D}E\left[\mathbf{n}(t)\right]+\lambda^2\mathrm{Var}\left[V(t)\right] \tag{41}$$

Rewrite $\mathbf{D}$ as

$$\mathbf{D} = \sum_i \left(\operatorname{vec}\mathbf{C}^{(i)}\right)\mathbf{e}_i^{\mathsf{T}} \tag{42}$$

where $\mathbf{e}_i$ is the ith unit vector of length s. Substituting Eq. 42 into Eq. 41 gives

$$\begin{aligned}\operatorname{Var}[V(t+1)] &= \sum_i \left(\mathbf{v}^{\mathsf{T}} \otimes \mathbf{v}^{\mathsf{T}}\right)\left(\operatorname{vec}\mathbf{C}^{(i)}\right)\mathbf{e}_i^{\mathsf{T}}\operatorname{E}[\mathbf{n}(t)] \\ &\quad + \lambda^2 \operatorname{Var}[V(t)] \qquad (43)\\ &= \sum_i \operatorname{vec}\left(\mathbf{v}^{\mathsf{T}}\mathbf{C}^{(i)}\mathbf{v}\right)\mathbf{e}_i^{\mathsf{T}}\operatorname{E}[\mathbf{n}(t)] \\ &\quad + \lambda^2 \operatorname{Var}[V(t)] \qquad (44)\end{aligned}$$

Define a vector $\mathbf{q}$ whose ith entry is

$$q_i = \mathbf{v}^{\mathsf{T}}\mathbf{C}^{(i)}\mathbf{v} \tag{45}$$

Then

$$\operatorname{Var}[V(t+1)] = \mathbf{q}^{\mathsf{T}}\operatorname{E}[\mathbf{n}(t)] + \lambda^2 \operatorname{Var}[V(t)] \tag{46}$$

Conditioning on $V(t)$ sets the last term to 0, leading to Eq. 15, from which Eq. 16 and 17 follow.

Derivation: sensitivity of σ_d^2

The sensitivity analysis of σ_d^2 requires a systematic application of matrix calculus; the approach has been described in some detail in Caswell (2007, 2008, 2009, 2012). A good mathematical introduction is in Abadir and Magnus (2005) and a detailed description in Magnus and Neudecker (1988).

Differentiating Eq. 17 and applying the vec operator gives

$$d\sigma_d^2 = \mathbf{q}^{\mathsf{T}}d\mathbf{w} + \mathbf{w}^{\mathsf{T}}d\mathbf{q} \tag{47}$$

The analysis proceeds by determining the differentials $d\mathbf{w}$ (which is given in Eq. 20) and $d\mathbf{q}$ (given in Eq. 22); the latter requires differentials of both $\mathbf{v}$ and of covariances. We consider each of these in turn.

Derivatives of v

The reproductive value vector $\mathbf{v}$ is the left eigenvector of $\mathbf{A}$, scaled so that $\mathbf{v}^{\mathsf{T}}\mathbf{w} = 1$. The scaling requires a slight modification of the result in Caswell (2008). We note that the scaled vector $\mathbf{v}$ satisfies

$$\mathbf{v}^{\mathsf{T}} = \frac{\mathbf{v}^{\mathsf{T}}\mathbf{A}}{\mathbf{v}^{\mathsf{T}}\mathbf{A}\mathbf{w}} \tag{48}$$

Differentiating Eq. 48, using the rule for ratios, gives

$$\begin{aligned}d\mathbf{v}^{\mathsf{T}} = \frac{1}{\left(\mathbf{v}^{\mathsf{T}}\mathbf{A}\mathbf{w}\right)^2}\Big\{&\left(\mathbf{v}^{\mathsf{T}}\mathbf{A}\mathbf{w}\right)\left[\left(d\mathbf{v}^{\mathsf{T}}\right)\mathbf{A} + \mathbf{v}^{\mathsf{T}}(d\mathbf{A})\right] \\ &-\left[\left(d\mathbf{v}^{\mathsf{T}}\right)\mathbf{A}\mathbf{w} + \mathbf{v}^{\mathsf{T}}(d\mathbf{A})\mathbf{w} + \mathbf{v}^{\mathsf{T}}\mathbf{A}(d\mathbf{w})\right]\left(\mathbf{v}^{\mathsf{T}}\mathbf{A}\right)\Big\}\end{aligned} \tag{49}$$

Substituting $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$ and $\mathbf{v}^{\mathsf{T}}\mathbf{A} = \lambda\mathbf{v}^{\mathsf{T}}$ gives

$$\begin{aligned}\lambda^2 d\mathbf{v}^{\mathsf{T}} = &\lambda\left(d\mathbf{v}^{\mathsf{T}}\right)\mathbf{A} + \lambda\mathbf{v}^{\mathsf{T}}(d\mathbf{A}) - \lambda^2\left(d\mathbf{v}^{\mathsf{T}}\right)\mathbf{w}\mathbf{v}^{\mathsf{T}} \\ &- \lambda\mathbf{v}^{\mathsf{T}}(d\mathbf{A})\mathbf{w}\mathbf{v}^{\mathsf{T}} - \lambda^2\mathbf{v}^{\mathsf{T}}(d\mathbf{w})\mathbf{v}^{\mathsf{T}}\end{aligned} \tag{50}$$

Applying the vec operator to both sides and solving for $d\mathbf{v}$ gives Eq. 23 where $d\mathbf{w}$ is given by Eq. 20.

Derivatives of the covariance matrices $\mathbf{C}^{(i)}$

To differentiate $\mathbf{C}^{(i)}$ requires the differentials of $\mathbf{C}_{\mathrm{U}}^{(i)}$ and $\mathbf{C}_{\mathrm{F}}^{(i)}$. We consider each in turn.

Covariances due to transitions

The covariance due to transitions is given by Eq. 7. For notational convenience, we supress the dependence on the parental stage i, and define $\mathbf{u} = \mathbf{U}(:, i)$, in terms of which

$$\mathbf{C}_{\mathrm{U}} = \mathbf{I} \circ \left(\mathbf{u}\mathbf{1}_s^{\mathsf{T}}\right) - \mathbf{u}\mathbf{u}^{\mathsf{T}} \tag{51}$$

Differentiating Eq. 51 yields

$$d\mathbf{C}_{\mathrm{U}} = \mathbf{I} \circ \left[(d\mathbf{u})\mathbf{1}_s^{\mathsf{T}}\right] - (d\mathbf{u})\mathbf{u}^{\mathsf{T}} - \mathbf{u}\left(d\mathbf{u}^{\mathsf{T}}\right) \tag{52}$$

Applying the vec operator gives the result in terms of $d\mathbf{u}$,

$$d\operatorname{vec}\mathbf{C}_{\mathrm{U}} = \left\{\mathbf{Z}\left(\mathbf{1}_s \otimes \mathbf{I}\right) - \left(\mathbf{u} \otimes \mathbf{I}\right) - \left(\mathbf{I} \otimes \mathbf{u}\right)\right\}d\mathbf{u} \tag{53}$$

where $\mathbf{Z} = \mathcal{D}(\operatorname{vec}\mathbf{I})$. Return now to noting that $\mathbf{u} = \mathbf{U}(:, i)$; i.e.

$$\mathbf{u} = \mathbf{U}\mathbf{e}_i \tag{54}$$

Then

$$d\mathbf{u} = (\mathbf{e}_i^{\mathsf{T}} \otimes \mathbf{I})d\operatorname{vec}\mathbf{U} \tag{55}$$

Substituting this into Eq. 53 we obtain Eq. 24.

Covariances due to reproduction

In both Bernoulli and Poisson reproduction, the co-variance matrix $\mathbf{C}_{\mathrm{F}}^{(j)}$ is obtained from the fertility matrix $\mathbf{F}$. For the moment, supress the indication of the parent stage i, and let $\mathbf{f}$ denote column j of $\mathbf{F}$.

The covariance matrix is

$$\mathbf{C}_{\mathrm{F}} = \mathcal{D}(\operatorname{Var}(\boldsymbol{\xi})) \tag{56}$$

$$= \mathbf{I} \circ \left(\mathbf{1}\operatorname{Var}(\boldsymbol{\xi})^{\mathsf{T}}\right) \tag{57}$$

Differentiating and applying the vec operator yields

$$d\operatorname{vec}\mathbf{C}_{\mathrm{F}} = \mathbf{Z}(\mathbf{I} \otimes \mathbf{1})d\operatorname{Var}(\boldsymbol{\xi}) \tag{58}$$

In the case of Bernoulli reproduction, $\mathrm{Var}(\boldsymbol{\xi}) = \mathbf{f} \circ (\mathbf{1} - \mathbf{f})$, and

$$d\mathrm{Var}\,(\boldsymbol{\xi}) = \mathcal{D}(\mathbf{1} - 2\mathbf{f})d\mathbf{f} \tag{59}$$

In the case of independent Poisson distributed numbers of offspring, $\mathrm{Var}(\boldsymbol{\xi}) = \mathbf{f}$ and

$$d\mathrm{Var}\left(\boldsymbol{\xi}\right) = d\mathbf{f} \tag{60}$$

Recalling that $\mathbf{f} = \mathbf{F}(:, j)$ gives

$$d\mathbf{f} = (\mathbf{e}_j \otimes \mathbf{I})d\mathrm{vec}\mathbf{F} \tag{61}$$

Substituting Eq. 59 and Eq. 61 into Eq. 58 gives $d\mathrm{vec}\,\mathbf{C}_\mathrm{F}$ for Bernoulli reproduction. Substituting Eq. 60 and Eq. 61 into Eq. 58 gives $d\mathrm{vec}\,\mathbf{C}_\mathrm{F}$ for Poisson reproduction. Together, these results give Eq. 25.

Acknowledgements – HC acknowledges support from European Research Council Advanced Grant 322989 and US National Science Foundation Grant DEB-1257545. YV was supported by a grant from the Research Council of Norway (244404). We thank Jean-Michel Gaillard for helpful comments.

References

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

Bailey, N. T. J. 1964. The elements of stochastic processes with applications to the natural sciences. – Wiley.

Barlow, J. and Boveng, P. 1991. Modeling age-specific mortality for marine mammal populations. – Mar. Mammal Sci. 7: 50–65.

Baudisch, A. 2011. The pace and shape of ageing. – Methods Ecol. Evol. 2: 375–382.

Blurton Jones, N. 2011. Hadza demography and sociobiology. – <www.sscnet.ucla.edu/anthro/faculty/blurton-jones/hadza-part-1.pdf>.

Brass, W. 1960. The graduation of fertility distributions by polynomial functions. – Popul. Studies 14: 148–162.

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

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

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

Caswell, H. 2009. Stage, age and individual stochasticity in demography. – Oikos 118: 1763–1782.

Caswell, H. 2011. Beyond R_0: 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. – Theor. Ecol. 5: 403–417.

Caswell, H. 2014. A matrix approach to the statistics of longevity in heterogeneous frailty models. – Demographic Res. 31: 553–592.

Caswell, H. and Kluge, F. A. 2015. Demography and the statistics of lifetime economic transfers under individual stochasticity. – Demographic Res. 32: 563–588.

Caswell, H. et al. 2017. Age × stage-classified demographic analysis: a comprehensive approach. – Ecol. Monogr. in press.

Coale, A. J. and Demeny, P. 1983. Regional model life tables and stable populations, 2nd edn. – Academic Press.

Cox, D. R. and Miller, H. D. 1965. The theory of stochastic processes. – Methuen.

Cvijović, I. et al. 2015. Fate of a mutation in a fluctuating environment. – Proc. Natl Acad. Sci. USA 112: E5021–E5028.

Eaton, J. W. and Mayer, A. J. 1953. The social biology of very high fertility among the Hutterites: the demography of a unique population. – Human Biol. 25: 206–264.

Ellner, S. P. et al. 2016. Data-driven modelling of structured populations. – Springer.

Engen, S. et al. 2005. Extinction in relation to demographic and environmental stochasticity in age-structured models. – Math. Biosci. 195: 210–227.

Engen, S. et al. 2007. Using reproductive value to estimate key parameters in density-independent age-structured populations. – J. Theor. Biol. 244: 308–317.

Engen, S. et al. 2009. Reproductive value and the stochastic demography of age-structured populations. – Am. Nat. 174: 795–804.

Erişoğlu, U. et al. 2012. Mixture model approach to the analysis of heterogeneous survival data. – Pakistan J. Stat. 28: 115–130.

Fisher, R. A. 1930. The genetical theory of natural selection. Revised 1956, Dover Publications edn. – Oxford Univ. Press.

Galton, F. and Watson, H. W. 1874. On the probability of extinction of families. – J. Anthropol. Inst. 4: 138–144.

Gurven, M. and Kaplan, H. 2007. Longevity among hunter–gatherers: a cross-cultural examination. – Popul. Dev. Rev. 33: 321–365.

Hartemink, N. and Caswell, H. 2018. Variance in animal longevity: contributions of heterogeneity and stochasticity. – Popul. Ecol. in press.

Hartemink, N. et al. 2017. Stochasticity, heterogeneity and variance in longevity in human populations. – Theor. Popul. Biol. 114: 107–117.

Hill, K. and Hurtado , A. M. 1996. Aché life history: the ecology and demography of a foraging people. – Aldine de Gruyter, NY.

Horvitz, C. C. and Schemske, D. W. 1995. Spatiotemporal variation in demographic transitions of a tropical understory herb: projection matrix analysis. – Ecol. Monogr. 65: 155–192.

Human Fertility Database 2013. Max Planck Inst. for Demographic Research (Germany) and the Vienna Inst. of Demography (Austria). – <www.humanfertility.org>.

Human Mortality Database 2013. Univ. of California, Berkeley (USA), and Max Planck Inst. for Demographic Research (Germany). – <www.mortality.org>.

Jenouvrier, S. et al. 2017. Interacting effects of unobserved heterogeneity and individual stochasticity in the life cycle of the southern fulmar. – J. Anim. Ecol. 87: 212–222.

Karlin, S. and Taylor, H. E. 1981. A decond course in stochastic processes. – Academic Press.

Kendall, B. E. and Wittmann, M. E. 2010. A stochastic model for annual reproductive success. – Am. Nat. 175: 461–468.

Kimura, M. 1964. Diffusion models in population genetics. – J. Appl. Probability 1: 177–232.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. – Am. Nat. 142: 911–927.

Lande, R. et al. 2003. Stochastic population dynamics in ecology and conservation. – Oxford Univ. Press.
Magnus, J. R. and Neudecker, H. 1988. Matrix differential calculus with applications in statistics and econometrics. – Wiley.
Morrison, D. G. et al. 1971. The effect of nonhomogeneous populations on Markov steady-state probabilities. – J. Am. Stat. Assoc. 66: 268–274.
Nagylaki, T. 1989. The diffusion model for migration and selection. – In: Hastings, A. (ed.), Some mathematical questions in biology: models in population biology. Lectures on mathematics in life sciences. Am. Mat. Soc., 20: 55–75.
Pollard, J. H. 1966. On the use of the direct matrix product in analysing certain stochastic population models. – Biometrika 53: 397–415.
Roth, G. and Caswell, H. 2016. Hyperstate matrix models: extending demographic state spaces to higher dimensions. – Methods Ecol. Evol. 7: 1438–1450.
Salguero-Gómez, R. et al. 2015. The COMPADRE plant matrix database: an open online repository for plant demography. – J. Ecol. 103: 202–218.
Salguero-Gómez, R. et al. 2016. COMADRE: a global data base of animal demography. – J. Anim. Ecol. 85: 371–384.
Sheps, M. C. and Menken, J. A. 1973. Mathematical models of conception and birth. – Univ. of Chicago Press.
Siler, W. 1979. A competing-risk model for animal mortality. – Ecology 60: 750–757.
Steiner, U. K. and Tuljapurkar, S. 2012. Neutral theory for life histories and individual variability in fitness components. – Proc. Natl Acad. Sci. USA 109: 4684–4689.
Steiner, U. K. et al. 2010. Dynamic heterogeneity and life history variability in the kittiwake. – J. Anim. Ecol. 79: 436–444.
Tuljapurkar, S. 1990. Population dynamics in variable enviroinments. Number 85 in lecture notes in biomathematics. – Springer.
Tuljapurkar, S. et al. 2009. Dynamic heterogeneity in life histories. – Ecol. Lett. 12: 93–106.
van Daalen, S. and Caswell, H. 2015. Lifetime reproduction and the second demographic transition: stochasticity and individual variation. – Demographic Res. 33: 561–588.
van Daalen, S. F. and Caswell, H. 2017. Lifetime reproductive output: individual stochasticity, variance and sensitivity analysis. – Theor. Ecol. 3: 355–374.
Vaupel, J. W. et al. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. – Demography 16: 439–454.
Vindenes, Y. and Engen, S. 2017. Demographic stochasticity and temporal autocorrelation in the dynamics of structured populations. – Oikos 126: 462–471.
Vindenes, Y. et al. 2008. Individual heterogeneity in vital parameters and demographic stochasticity. – Am. Nat. 171: 455–467.
Vindenes, Y. et al. 2010. Fixation of slightly beneficial mutations: effects of life history. – Evolution 64: 1063–1075.
Vindenes, Y. et al. 2011. Integral projection models for finite populations in a stochastic environment. – Ecology 92: 1146–1156.
Wilson, A. J. and Nussey, D. H. 2010. What is individual quality? An evolutionary perspective. – Trends Ecol. Evol. 25: 207–214.
Yule, G. U. 1910. On the distribution of deaths with age when the causes of death act cumulatively, and similar frequency distributions. – J. R. Stat. Soc. 73: 26–38.

Supplementary material (available online as Appendix oik-04708 at <www.oikosoffice.lu.se/appendix/oik-04708>). Appendix 1.