Occupancy time in sets of states for demographic models

Roth, G.; Caswell, H.

Occupancy time in sets of states for demographic models

Gregory Roth *, Hal Caswell

Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Netherlands

ARTICLE INFO

ABSTRACT

As an individual moves through its life cycle, it passes through a series of states (age classes, size classes, reproductive states, spatial locations, health statuses, etc.) before its eventual death. The occupancy time in a state is the time spent in that state over the individual's life. Depending on the life cycle description, the occupancy times describe different demographic variables, for example, lifetime breeding success, lifetime habitat utilisation, or healthy longevity.

Models based on absorbing Markov chains provide a powerful framework for the analysis of occupancy times. Current theory, however, can completely analyse only the occupancy of single states, although the occupancy time in a set of states is often desired. For example, a range of sizes in a size-classified model, an age class in an age×stage model, and a group of locations in a spatial stage model are all sets of states.

We present a new mathematical approach to absorbing Markov chains that extends the analysis of life histories by providing a comprehensive theory for the occupancy of arbitrary sets of states, and for other demographic variables related to these sets (e.g., reaching time, return time). We apply this approach to a matrix population model of the Southern Fulmar (*Fulmarus glacialoides*). The analysis of this model provides interesting insight into the lifetime number of breeding attempts of this species.

Our new approach to absorbing Markov chains, and its implementation in matrix oriented software, makes the analysis of occupancy times more accessible to population ecologists, and directly applicable to any matrix population models.

1. Introduction

The life of an individual is a sequence of events. Birth and death are events common to every individual, but the sequence between birth and death – unique to each individual – consists of a potentially endless list of random events (e.g., surviving, developing, mating, reproducing, growing, dispersing, moving among social or occupational classes, or changing health status). Each event corresponds to a change in the state of the individual, resulting in a stochastic pathway that ends eventually in death. A central role in the analysis of these pathways is played by the concepts of *occupancy time* (the time spent in, or the number of visits to, a state over the individual's lifetime). Occupancy is a property of the stochastic pathway of an individual, and occupancy times define the time spent in each of the possible states during the lifetime. In particular, the longevity of an individual is measured by the sum of all these occupancy times. The interpretation of occupancy times depends on the identity of the transient states and the nature of the absorption. Thus, when the states are health status, occupancy time represents years of life while healthy, not

healthy, disabled, etc. When the states are spatial locations, occupancy time represents time spent in different places. When the states are marital status, occupancy times measure the part of the lifetime spent single, married, divorced, remarried, etc. When the states are employment status, or breeding activities, or any other interesting categorisation of individuals, the interpretation follows the same lines. As for absorption, it may be death, in which case occupancy time is a "lifetime" measure in the literal sense. But absorption can be defined as the first entrance to some state or set of states (e.g., occurrence of first breeding, or graduation, or metamorphosis, or hospitalisation, etc.).

Because the pathways are stochastic, occupancy time is a random variable. It is often described by its mean (e.g., life expectancy, expected lifetime reproduction). However, some individuals will live longer and some shorter, than the mean; some will mature later and some earlier than the mean. To characterise this variation, the probability distribution of occupancy time, or at least its moments, must be considered.

Models based on absorbing Markov chains provide a powerful framework for the analysis of occupancy times. An absorbing Markov chain describes the fate of an individual – under the assumption that the future of the individual, given its present, is independent of its past – evolving in a set of states and being eventually absorbed by the death state. The states may refer to

developmental states, physiological measures, behaviour types, locations, and so on. The set of transition rates between these states – described by a *transition matrix* – defined an absorbing Markov chain. As a population projection matrix describes the fate of a population, the transition matrix describes the fate of individuals in a population, and often is one component of a population projection matrix. The mathematical theory of absorbing Markov chain provides formulae for basic descriptive quantities of the absorbing Markov chain, based on its fundamental matrix (see e.g. Iosifescu (1980) and Kemeny and Snell (1961) for a mathematical perspective, and Caswell (2001) for a demographical perspective). Applied to demographic models, this theory provides simple and direct formulae for the probability distribution, the mean, variance, and all moments of longevity, the distribution of age or stage at death, the survivorship and mortality functions, causes of death, and a variety of measures of life disparity (e.g., Feichtinger, 1971; Cochran and Ellner, 1992; Caswell, 2001, 2006, 2009; Tuljapurkar and Horvitz, 2006; Horvitz and Tuljapurkar, 2008; Van Raalte and Caswell, 2013). Powerful sensitivity analyses are available for all these quantities (Caswell, 2006, 2009, 2011b, 2013).

Current theory, however, can completely analyse only the occupancy time of single states and the occupancy time of the whole state space. Our goal is to extend the analysis of life histories by providing a comprehensive theory for the occupancy of arbitrary sets of states. One type of set is a collection of states deemed biologically relevant for some purpose; we call these *super-states*. For example, a model based on reproductive behaviour might include states describing many details of the success, failure, timing, and number of offspring produced by breeding, but one might want to investigate the super-states created by aggregating these into "successful breeding" and "non-successful breeding" sets. A spatial model might describe habitats along an altitudinal gradient, and one might want to aggregate in order to compare the occupancy of low altitude and high altitude sites. A medical demography study might distinguish a variety of health conditions and treatments, but one might want to compare the occupancy of all states requiring hospitalisation and those not requiring hospitalisation. The utility of super-states will increase as more matrix models are created from the growth and survival kernels of integral projection models (e.g., Ellner et al., 2016). These matrices typically contain hundreds of size classes, no one of which is of particular interest, but sets of which (e.g., all trees large enough to reach the forest canopy) are of great interest.

A second type of sets of states arises in the context of multistate (or megamatrix) or hyperstate models (e.g., Rogers, 1975; Lebreton, 1996; Pascarella and Horvitz, 1998; Tuljapurkar et al., 2003; Roth and Caswell, 2016) in which individuals are classified by two or more criteria (age and stage, stage and location, etc.). One may want to analyse the occupancy of sets of states defined by integrating over one of these criteria; we call these *marginal sets*. For example, in a stage×size-classified model, the marginal set associated with the *juvenile* stage is the set containing the juvenile stage, integrated over all possible sizes.

The extension of occupancy time calculations to sets of states may seem trivial because the occupancy time in a set is the sum of the occupancy times in each state belonging to this set. Therefore, the mean occupancy time in a set is the sum of the means of the occupancy times in each state. However, this observation does not hold for the variance, for any higher moments, or for the probability distribution, because occupancy times in single states are not *independent* from each other. There are few analyses of the occupancy time in set of states, but they only focus on specific aspects of it. For example, Steiner and Tuljapurkar (2012) provide formulae for the mean and variance of the reproductive output using the joint generating function of the single state occupancy times. The reproductive output of an individual is closely related to

the occupancy time in the set of reproductive states (both are equal when fertility rates are ones in each reproductive state). Caswell (2011a) provides similar formulae using the theory of Markov chain with reward. The same theory is used by Caswell and Kluge (2015) to calculate the moments of lifetime accumulation of economic variables, which are also closely related to occupancy times. However, these studies do not provide the probability distribution of occupancy time in a set of states. In the mathematical literature, Sericola (2000) provides an iterative formula for the probability distribution of the partial (i.e. up to a fixed time) occupancy time in a set of states, but does not provide a closed formula for the total occupancy time. Here, we present a comprehensive approach to calculate the any moment and the probability distribution of the occupancy time in arbitrary sets of states. Our approach relies on the construction of a *sub Markov chain*, which describes the original Markov chain viewed through a filter that allows one to see only the states in the set of interest. As a consequence, all the statistics of the occupancy time in the set of interest may be calculated with the existing theory of absorbing Markov chain (Iosifescu, 1980), applied to the sub Markov chain.

The construction and the analysis of the sub chain extends the classical theory of absorbing Markov chain by providing not only several measures related to the occupancy of sets of states but also forms a basis for further calculations of measures related to sets of states, including

- The set occupancy time. Depending on the life cycle description, set occupancy times describe different demographic variables (e.g., lifetime breeding attempts in a model of reproductive behaviour, or lifetime habitat utilisation in a spatial model). We provide for the probability distribution, mean, and variance of the occupancy times.
- The correlation between the occupancy times in two different sets. This is an indicator of how the two sets are connected in the life cycle. As a particular case, we provide, for the first time, a formula for the correlation between the occupancy time in a state and the longevity of an individual. Depending on the life cycle description, this formula gives the correlation between different demographic variables and longevity (e.g., lifetime breeding attempts and longevity, lifetime reproduction and longevity, time to maturation and longevity).
- Properties of winners and losers. Relative to a particular target set, a *winner* is an individual that enters the set at least once in its life, and a *loser* is an individual that never enters the set. In a model classifying individuals by their developmental state, the winners might represent those individuals that successfully mature, and the losers those that do not. We provide the probability of becoming a winner, the distribution, mean, and variance of the time required for a winner to reach the set, and the longevity of a loser. After its first success, a winner may leave the set and never return, or it may return at some future time. We obtain the probability that a winner returns, and for those that do return, the probability distribution, mean, and variance of the return time.

Table 1 lists the demographic results to be presented and the equations in which they are derived. All the results are obtained directly from a single matrix, describing the transition probabilities among transient states. This matrix is obtainable from most population projection matrices (Caswell, 2001). Despite the large number of matrices and sometimes complicated expressions that appear in our derivations, our results are easily computed in matrix-oriented software. In the Supplementary Material, we provide the MATLAB code for calculating all of the demographic results listed in Table 1.

Table 1
Results of the set analysis in a Markov chain demographic model.

Output	Expression	Equation
Occupancy time in $\mathcal{B}$		
Mean and variance	$\mathbb{E}[\boldsymbol{\tau}_{\mathcal{B}}]$, $\mathrm{Var}(\boldsymbol{\tau}_{\mathcal{B}})$	(20), (24), (25), (26)
Moments	$\boldsymbol{\tau}_{\mathcal{B}}^{k}$	(23)–(25)
Distribution		(28)–(29)
Reaching the set $\mathcal{B}$		
Probability to reach	$\mathbf{p}_{a}$	(39)
Time to reach:	$\mathbf{t}_{\mathcal{B}}$	(41)
Mean and variance	$\mathbb{E}[\mathbf{t}_{\mathcal{B}}]$, $\mathrm{Var}(\mathbf{t}_{\mathcal{B}})$	(43)–(44)
Moments	$\mathbf{t}_{\mathcal{B}}^{k}$	(42)
Distribution		(45)
Returning to the set $\mathcal{B}$		
Probability to return	$\mathbf{p}_{r}$	(47)
Time to return:	$\boldsymbol{\mu}$	(48)
Mean and variance	$\mathbb{E}[\boldsymbol{\mu}]$, $\mathrm{Var}(\boldsymbol{\mu})$	(53)–(54)
Moments	$\boldsymbol{\mu}^{\mathbb{K}}$	(97), in the appendix
Distribution		(51)–(52)
Correlation between the occupancy times in two sets		
Correlation	$\mathrm{Corr}\left(\boldsymbol{\tau}_{\mathcal{B}_{1}}, \boldsymbol{\tau}_{\mathcal{B}_{2}}\right)$	(86)–(36)
Correlation between the occupancy time in $\mathcal{B}$ and the longevity		
Correlation	$\mathrm{Corr}\left(\boldsymbol{\tau}_{\mathcal{B}}, \boldsymbol{\eta}\right)$	(38)

Notation. Matrices are denoted by upper-case bold symbols, vectors by lower-case bold symbols. Vectors are column vectors by default. When a matrix has a subscript (e.g., $\mathbf{U}_{\mathbb{K}}$), we note its entries with a superscript, $u_{ij}^{\mathbb{K}}$. The vector $\mathbf{1}_{\xi}$ is the $\xi \times 1$ vector of ones, and the vector $\mathbf{0}_{\xi}$ is the $\xi \times 1$ vector of zeros. The matrix $\mathbf{I}_{\xi}$ is the identity matrix of size $\xi \times \xi$. The transpose of $\mathbf{A}$ is $\mathbf{A}^{\mathsf{T}}$. The diagonal matrix with the vector $\mathbf{z}$ on its diagonal is $\mathrm{diag}(\mathbf{z})$. The Hadamard product (component by component) of $\mathbf{A}$ and $\mathbf{B}$ is $\mathbf{A} \circ \mathbf{B}$. Any random variable x is defined on a probability space with probability $\mathbb{P}$, its expectation is denoted by $\mathbb{E}[x]$ and its variance is denoted by $\mathrm{Var}(x)$. The kth moment, $\mathbb{E}[\mathbf{y}^{k}]$, of a random vector $\mathbf{y}$ is denoted by $\mathbf{y}^{k}$.

2. Absorbing Markov chains as demographic models

An absorbing Markov chain describes the fate of a particle – under the assumption that the future of the particle, given its present, is independent of its past – evolving in a set of states and being eventually absorbed by an absorbing state. In a demographic model, the absorbing Markov chain describes the fate of an individual evolving through a set of states and being absorbed by the state representing its death. The absorbing Markov chain is determined by two key elements: the living states of the individuals, and the transition probabilities between those states, which depend on the entire set of vital rates.

Formally, we consider the finite set of living states $\mathcal{T} = \{1, \ldots, \omega\}$, and the inevitable state d, representing death. The living states and the death state are called respectively the *transient states* and *absorbing state* of the Markov chain. The state space of the Markov chain is their union,

$$\mathcal{S} = \mathcal{T} \cup \{d\}. \tag{1}$$

An absorbing Markov chain is uniquely defined by the one-step transition probabilities and an initial probability distribution. The transition probabilities are described by the matrix $\mathbf{P}$ of size $(\omega + 1) \times (\omega + 1)$. For each $i, j \in \mathcal{S}$, the $i - j$ entry of $\mathbf{P}$ is the probability that an individual in state j moves to state i. Therefore, the columns of $\mathbf{P}$ sum to one (i.e. $\mathbf{P}$ is a stochastic matrix). Note that in the mathematical literature, the transition matrix is given by the transpose of $\mathbf{P}$; however, our notation is widely used in the

demographical literature. The matrix $\mathbf{P}$ can be decomposed in four blocks:

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & \mathbf{0}_{\omega} \\ \hline \mathbf{m}^{\mathsf{T}} & 1 \end{array} \right) \quad \text{of size} \quad \left(\begin{array}{c|c} \omega \times \omega & \omega \times 1 \\ \hline 1 \times \omega & 1 \times 1 \end{array} \right), \tag{2}$$

where the matrix $\mathbf{U}$ of size $\omega \times \omega$ is the *transient transition matrix*, which describes the transition probabilities between the transient states, and the column vector $\mathbf{m}$ of size ω contains the *probabilities of death* from each transient state.

Since the column sums of $\mathbf{P}$ are equal to one, the column sums of matrix $\mathbf{U}$ are less than one (strictly less than one if the death probabilities are assumed to be non zero). Moreover, since there is a single absorbing state, the vector $\mathbf{m}$ is in fact a function of the matrix $\mathbf{U}$,

$$\mathbf{m}^{\mathsf{T}} = \mathbf{1}_{\omega}^{\mathsf{T}} - \mathbf{1}_{\omega}^{\mathsf{T}} \mathbf{U}. \tag{3}$$

As a consequence, the matrix $\mathbf{P}$ is function of the transient transition matrix $\mathbf{U}$, which is the only variable of the model.

An absorbing Markov chain from a population projection matrix. A population projection matrix can often be decomposed in two parts: one is the matrix $\mathbf{U}$ that describes the transitions of living individuals and the other is a matrix $\mathbf{F}$ that describes the rates of production of new individuals (see e.g. Caswell (2001)). Therefore, the population projection matrices contain an absorbing Markov chain, which is defined by the transient transition matrix $\mathbf{U}$.

Target states. The focus of this paper is the occupancy time in a set of living states, which we call the *target set*. To fix the idea, we define a target set $\mathcal{B}$ composed of β transient states, and we renumber the states so that the target states are the last β elements of $\mathcal{T}$, i.e.

$$\mathcal{B} = \{\alpha + 1, \alpha + 2, \ldots, \omega\}$$
$$\mathcal{B}^{c} = \{1, 2, \ldots, \alpha\},$$

where $\alpha = \omega - \beta$. According to this renumbering of the states, we re-arrange the entries of the matrix $\mathbf{P}$. This results in a new matrix $\mathbf{U}$ and a new vector $\mathbf{m}$. To avoid overloaded notation, we shall retain the symbols $\mathbf{U}$ and $\mathbf{m}$. It will be clear from the context which arrangement is being used.

Iosifescu (1980) derives formulae for the moments of the occupancy time of the Markov chain in a *single* target state (see also Caswell (2009)). The key element for those calculations is the *fundamental matrix*

$$\mathbf{N} = (\mathbf{I}_{\omega} - \mathbf{U})^{-1}. \tag{4}$$

The i-jth entry of the matrix $\mathbf{N}$ is the mean occupancy time in the state i for an individual starting in state j. Any higher moment of the occupancy time is derived from the fundamental matrix. In particular, the variances are given by the matrix,

$$\mathbf{V} = (2\mathbf{I}_{\omega} \circ \mathbf{N} - \mathbf{I}_{\omega})\mathbf{N} - \mathbf{N} \circ \mathbf{N}. \tag{5}$$

The mean occupancy time in the target set is the sum of the mean occupancy times in each target state. Hence, the mean is directly deduced from the fundamental matrix by summing the corresponding entries. However, this technique breaks down when calculating the variance of the occupancy time. Indeed, the variance of a sum of random variables is not the sum of the variance of each variable (unless the variables are independent). This complication motivates the need of the new technique presented in the next Section. This technique provides not only a straightforward formula for the variance, but also an entire set of new formulae for diverse measures of life history traits, as listed in Table 1.
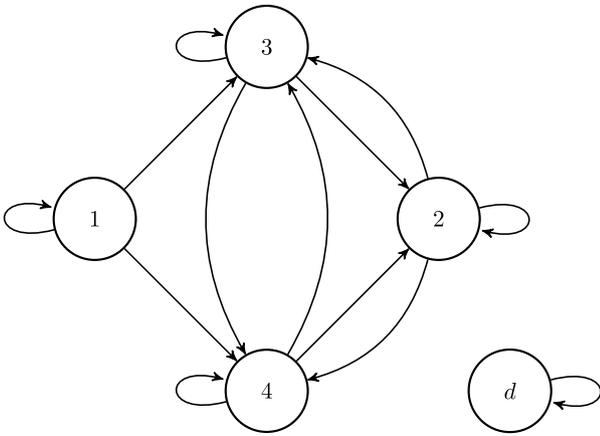
Fig. 1. Life cycle graph of Southern Fulmar. State 1 is the pre-breeder state. State 2 is the non-breeder state. State 3 is the successful breeder state. State 4 is the failed breeder state. State d is the death state. To lighten the graph we omit the transitions to the death state.

Example. The model and the results are illustrated throughout the paper with a model for the Southern Fulmar (*Fulmarus glacialoides*) derives by Jenouvrier et al. (2015). The life cycle graph shown in Fig. 1 is broken into four stages, which are defined at the end of the breeding season: (1) pre-breeders, who have yet to breed for the first time; this include fledged chicks from the previous season; (2) non-breeders, who have bred at least once before, but not in the current season; (3) successful breeders, who have successfully raised a chick during the current season, and (4) failed breeders, who have not successfully raised a chick during the current season because they failed to either hatch an egg or raise a chick. Hence, the set of transient states is $\mathcal{T} = \{1, 2, 3, 4\}$ and the individual state space is $\mathcal{S} = \{1, 2, 3, 4\} \cup \{d\}$. We define two target sets: $\mathcal{B}_a = \{2, 3, 4\}$ ($\beta = 3$ and $\alpha = 1$) and $\mathcal{B}_b = \{3, 4\}$ ($\beta = 2$ and $\alpha = 2$). The former represents the super-state *adult*, that is an individual in $\mathcal{B}_a$ is characterised by having bred at least once. The latter represents the super-state *breeding*; an individual in $\mathcal{B}_b$ is currently attempting to breed regardless of its success.

Outcome of a Markov chain model. An outcome of a demographic Markov chain model is function of the initial state of an individual, and its pathway through the life cycle. For example, the number of breeding attempt of a newborn Southern Fulmar is the result of its entire pathway − from birth to death. Here, the initial state is pre-breeder because any newborn is in this state. The number of breeding attempt of an adult is a result of its pathway, but only from its first passage in the super-state adult to death. Here, the initial state is successful breeder (or failed breeder) because an individual matures when it reaches one of these two states. For this reason, any result listed in Table 1 is a vector and each entry corresponds to a specific initial condition. For example, the occupancy time in $\mathcal{B}$ is given by the vector

$$\boldsymbol{\tau}_\mathcal{B} = (\tau_1, \ldots, \tau_\omega) \tag{6}$$

where τ_i is the occupancy time in the set $\mathcal{B}$ for an individual that is initially in state i.

3. Constructing the induced Markov chains

To derive our results, we construct three new Markov chains associated with the set of target states $\mathcal{B}$. These chains are created from the *original* Markov chain $\mathbf{P}$ and the target set $\mathcal{B}$.

- The *killed Markov chain*, with transition matrix $\mathbf{P}_\mathbb{K}$. This Markov chain is a copy of the original Markov chain that is stopped as soon as it enters in the target set. If the individual never enters the target set, then the pathway of the killed Markov chain is equivalent to the pathway of the initial Markov chain. See Section 3.1 for a formal definition.
- The *conditional Markov chain*, with transition matrix $\mathbf{P}_\mathbb{C}$. This Markov chain concerns only individuals that successfully reach the target set. It describes their pathway from their initial state to their first entrance to the target set. In other words, it describes these pathways of the killed Markov chain that do reach a target state. See Section 3.2 for a formal definition.
- The *sub-Markov chain*, with transition matrix $\mathbf{P}_\mathbb{S}$. This Markov chain is a copy of the original Markov chain viewed through a filter that allows one to see only the target set. The pathways of the sub-Markov chain corresponds to the pathways of the original Markov chain observed through this filter. See Section 3.3 for a formal definition.

To each pathway of the original Markov chain there corresponds a pathway in each induced Markov chain. Fig. 2 shows three pathways through the life cycle of the Southern Fulmar, and, below each of them, the corresponding pathways of the induced Markov chains, associated with the target set $\mathcal{B}_b$, consisting of the breeding states.

We now explain in detail how to construct the matrices $\mathbf{P}_\mathbb{K}$, $\mathbf{P}_\mathbb{C}$, and $\mathbf{P}_\mathbb{S}$ from the matrix $\mathbf{P}$; we illustrate this construction with the Southern Fulmar example. After the rearrangement of the matrix $\mathbf{P}$ according to the numbering of the states (see Section 2), the transition matrix $\mathbf{U}$ can be split in four block matrices, which contain the transition probabilities from $\mathcal{B}^c$ to $\mathcal{B}^c$, from $\mathcal{B}^c$ to $\mathcal{B}$, from $\mathcal{B}$ to $\mathcal{B}^c$, and from $\mathcal{B}$ to $\mathcal{B}$, respectively. These block matrices are denoted by $\mathbf{U}_\mathbb{K}$, $\mathbf{L}$, $\mathbf{K}$, and $\mathbf{Q}$, and appear within $\mathbf{U}$ as Eq. (7) given in (Box I).

These four matrices are essential to the construction of the killed, the conditional, and the sub Markov chains. All the other matrices used for this construction and for the derivation of the results are tabulated for easy reference in Table 2.

3.1. The killed Markov chain $\mathbf{P}_\mathbb{K}$

The transition matrix $\mathbf{P}_\mathbb{K}$ describes a copy of the original Markov chain that is stopped as soon as it enters $\mathcal{B}$. In this chain, the states in $\mathcal{B}$ and the death state d are now absorbing, and the new transient set is $\mathcal{B}^c$. The state space of the killed Markov chain is $\mathcal{S}$, and its transition probability matrix is

$$\mathbf{P}_\mathbb{K} = \left(\begin{array}{c|c} \mathbf{U}_\mathbb{K} & \mathbf{0}_{\alpha \times (\beta+1)} \\ \hline \mathbf{M}_\mathbb{K} & \mathbf{I}_{\beta+1} \end{array} \right) \quad \text{of size}$$

$$\left(\begin{array}{c|c} \alpha \times \alpha & \alpha \times (\beta+1) \\ \hline (\beta+1) \times \alpha & (\beta+1) \times \beta \end{array} \right). \tag{8}$$

The entry $u_{ij}^\mathbb{K}$ is the transition probability from non-target state j to non-target state i. The entry $m_{ij}^\mathbb{K}$ is the transition probability from state j to target state $\alpha + i$, for $i = 1, \ldots, \beta$, and to the death state, for $i = \beta + 1$. The matrix $\mathbf{U}_\mathbb{K}$ is extracted from the original transition matrix $\mathbf{U}$, as in Eq. (7). The matrix $\mathbf{M}_\mathbb{K}$ is composed of two blocks:

$$\mathbf{M}_\mathbb{K} = \left(\begin{array}{c} \mathbf{K} \\ \hline \mathbf{m}_\alpha^\mathsf{T} \end{array} \right). \tag{9}$$

The matrix $\mathbf{K}$ describes the transitions from the non target states to the target states; it is directly extracted from the original transition matrix $\mathbf{U}$, as in Eq. (7). The vector $\mathbf{m}_\alpha$ describes the transitions from

$$
\mathbf{U} = \begin{array}{c} \\ \mathcal{B}^c \left\{ \vphantom{\begin{array}{c}a\\a\\a\end{array}} \right. \\ \mathcal{B} \left\{ \vphantom{\begin{array}{c}a\\a\\a\end{array}} \right. \end{array} \overbrace{\left(\begin{array}{ccc|ccc} u_{11} & \cdots & u_{1\alpha} & u_{1,\alpha+1} & \cdots & u_{1\omega} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u_{\alpha 1} & \cdots & u_{\alpha\alpha} & u_{\alpha,\alpha+1} & \cdots & u_{\alpha\omega} \\ \hline u_{\alpha+1,1} & \cdots & u_{\alpha+1,\alpha} & u_{\alpha+1,\alpha+1} & \cdots & u_{\alpha+1,\omega} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u_{\omega,1} & \cdots & u_{\omega,\alpha} & u_{\omega,\alpha+1} & \cdots & u_{\omega,\omega} \end{array} \right)}^{\mathcal{B}^c \quad\quad\quad \mathcal{B}} =: \left(\begin{array}{c|c} \mathbf{U}_{\mathbb{K}} & \mathbf{L} \\ \hline \mathbf{K} & \mathbf{Q} \end{array} \right) \tag{7}
$$

Box I.

Table 2

Matrices used the construction and in the analysis of the Markov chain demographic model. ω and β denote the number of living states in the life cycle and the number of states in the set $\mathcal{B}$, respectively, and $\alpha = \omega - \beta$.

Notation	Expression	Size	Description
Matrices describing demographic model			
$\mathbf{P}$		$(\omega + 1) \times (\omega + 1)$	Transition probabilities
$\mathbf{U}$		$\omega \times \omega$	Living states transition probabilities
$\mathbf{m}$		$\omega \times 1$	Mortality probabilities
$\mathbf{N}$	$(\mathbf{I}_\omega - \mathbf{U})^{-1}$	$\omega \times \omega$	Fundamental matrix
Decomposition of U			
$\mathbf{U}_{\mathbb{K}}$	Eq. (7)	$\alpha \times \alpha$	$\mathcal{B}^c$-to-$\mathcal{B}^c$ transitions
$\mathbf{K}$	Eq. (7)	$\alpha \times \beta$	$\mathcal{B}^c$-to-$\mathcal{B}$ transitions
$\mathbf{L}$	Eq. (7)	$\alpha \times \beta$	$\mathcal{B}$-to-$\mathcal{B}^c$ transitions
$\mathbf{Q}$	Eq. (7)	$\beta \times \beta$	$\mathcal{B}$-to-$\mathcal{B}$ transitions
Matrices describing the killed MC			
$\mathbf{P}_{\mathbb{K}}$		$(\omega + 1) \times (\omega + 1)$	Transition probabilities
$\mathbf{U}_{\mathbb{K}}$		$\alpha \times \alpha$	Transient state transition probabilities
$\mathbf{M}_{\mathbb{K}}$	Eq. (9)	$(\beta + 1) \times \alpha$	Absorbing transition probabilities
$\mathbf{N}_{\mathbb{K}}$	$(\mathbf{I}_\alpha - \mathbf{U}_{\mathbb{K}})^{-1}$	$\alpha \times \alpha$	Fundamental matrix
Matrices describing the conditional MC			
$\mathbf{P}_{\mathbb{C}}$		$(\alpha + 1) \times (\alpha + 1)$	Transition probabilities
$\mathbf{U}_{\mathbb{C}}$	$\mathbf{D}_a \mathbf{U}_{\mathbb{K}} \mathbf{D}_a^{-1}$	$\alpha \times \alpha$	Transient state transition probabilities
$\mathbf{m}_{\mathbb{C}}$	$(\mathbf{1}_\alpha^{\mathsf{T}} \mathbf{K} \mathbf{D}_a^{-1})^{\mathsf{T}}$	$\alpha \times 1$	Absorbing transition probabilities
$\mathbf{N}_{\mathbb{C}}$	$(\mathbf{I}_\alpha - \mathbf{N}_{\mathbb{C}})^{-1}$	$\alpha \times \alpha$	Fundamental matrix
Matrices describing the sub MC			
$\mathbf{P}_{\mathbb{S}}$		$(\alpha + 1) \times (\alpha + 1)$	Transition probabilities
$\mathbf{U}_{\mathbb{S}}$	$\mathbf{AL} + \mathbf{Q}$	$\alpha \times \alpha$	Transient state transition probabilities
$\mathbf{m}_{\mathbb{S}}$	$\mathbf{1}_\beta^{\mathsf{T}} - \mathbf{1}_\beta^{\mathsf{T}} \mathbf{U}_{\mathbb{S}}$	$1 \times \alpha$	Mortality probabilities
$\mathbf{N}_{\mathbb{S}}$	$(\mathbf{I}_\beta - \mathbf{U}_{\mathbb{S}})^{-1}$	$\beta \times \beta$	Fundamental matrix
Other			
$\mathbf{A}$	$\mathbf{K}\mathbf{N}_{\mathbb{K}}$	$\beta \times \alpha$	Reaching a state in $\mathcal{B}$ probabilities
$\tilde{\mathbf{A}}$	$(\mathbf{A}|\mathbf{I}_\beta)$	$\beta \times \omega$	
$\mathbf{p}_a$	$\mathbf{A}^{\mathsf{T}} \mathbf{1}_\beta$	$\alpha \times 1$	Reaching $\mathcal{B}$ probabilities
$\mathbf{D}_a$	$\mathrm{diag}(\mathbf{p}_a)$	$\alpha \times \alpha$	
$\mathbf{p}_r$	$\mathbf{U}_{\mathbb{S}}^{\mathsf{T}} \mathbf{1}_\beta$	$\beta \times 1$	
$\mathbf{D}_r$	$\mathrm{diag}(\mathbf{p}_r)$	$\beta \times \beta$	
Transition probabilities given individual returns in $\mathcal{B}$			
$\mathbf{W}_{\mathrm{in}}$	$\mathbf{Q}\mathbf{D}_r^{-1}$	$\beta \times \beta$	$\mathcal{B}$ to $\mathcal{B}$ transitions
$\mathbf{W}_{\mathrm{out}}$	$\mathbf{D}_a \mathbf{L} \mathbf{D}_r^{-1}$	$\alpha \times \beta$	$\mathcal{B}$ to $\mathcal{B}^c$ transitions

the non target states to the death states; it is extracted from the original vector $\mathbf{m}$ describing the probabilities of death (Eq. (2)),

$$
\mathbf{m}_\alpha = (m_1, \ldots, m_\alpha). \tag{10}
$$

Example. Fig. 3 shows the graphs of the killed Markov chains associated with the target states $\mathcal{B}_b = \{3, 4\}$ and $\mathcal{B}_a = \{2, 3, 4\}$, respectively.

3.1.1. Absorption probabilities

In order to define the conditional Markov chain, we need to calculate, for each target state i and non-target state j, the probability that an individual starting in j passes through i before eventual

death. For $j = 1, \ldots, \alpha$, and $i = 1, \ldots, \beta$, let a_{ij} denote the probability that the killed chain initially in state j (non target state) is absorbed into the target state $\alpha + i$. Following Theorem 3.3 in Iosifescu (1980), those probabilities are described by the matrix

$$
\mathbf{A} = \mathbf{K}\mathbf{N}_{\mathbb{K}} \quad \text{of size } \beta \times \alpha, \tag{11}
$$

where $\mathbf{N}_{\mathbb{K}} = (\mathbf{I}_\alpha - \mathbf{U}_{\mathbb{K}})^{-1}$ is the fundamental matrix associated with the killed Markov chain. Note that the matrix $\mathbf{A}$ is a sub-matrix of the matrix $\mathbf{A}^* = \mathbf{M}_{\mathbb{K}} \mathbf{N}_{\mathbb{K}}$, which describes the entire distribution of fates for any starting non-target state, i.e. the probabilities that the killed Markov chain is stopped in any of the target states or in the death state. Therefore the columns of $\mathbf{A}^*$ sum to one.
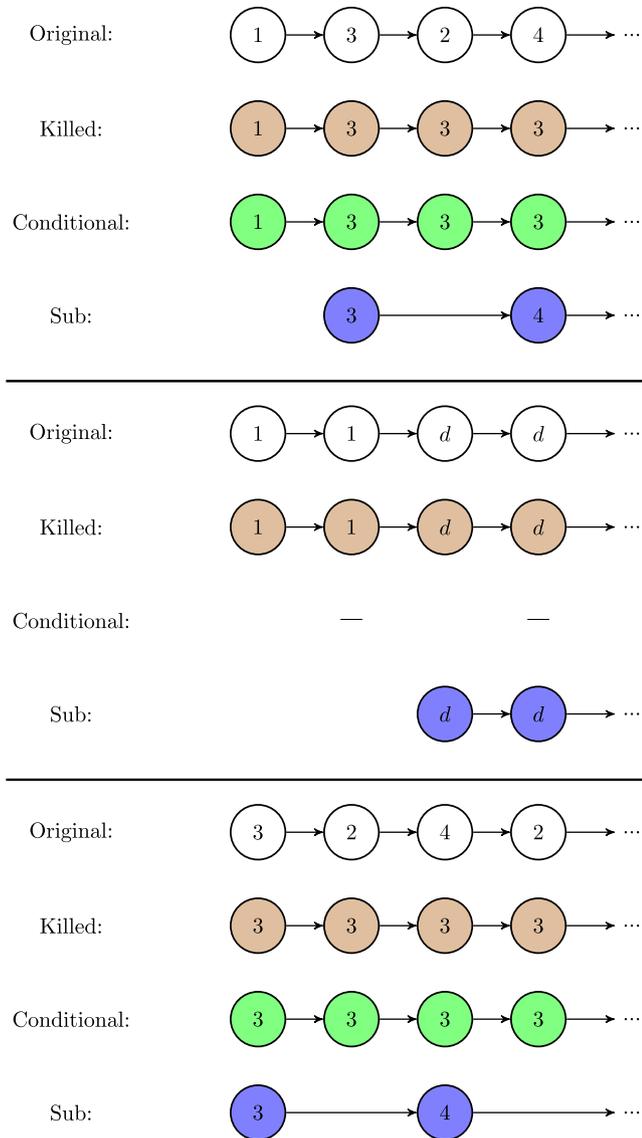
Fig. 2. The first row of each block is a realisation of the original Markov chain describing the Southern Fulmar life cycle. Under them are the corresponding realisations of the induced Markov chains associated with the target states {3, 4}: the killed Markov chain in brown, the conditional Markov chain in green, and the sub-Markov chain in blue. The realisations of the sub chain may have longer jumps because it describes the pathways of the original chain only in the target set. In the second block, the conditional chain has no realisation because the original chain does not enter any target state. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The probabilities of absorption in the set $\mathcal{B}$ – regardless of the specific state in which the chain is absorbed – are given by the vector $\mathbf{p}_a$ which satisfies the equation

$$\mathbf{p}_a^\mathsf{T} = \mathbf{1}_\beta^\mathsf{T} \mathbf{A}. \tag{12}$$

The ith entry of $\mathbf{p}_a$ is the probability that the killed Markov chain, initially in state i, is absorbed in a target state.

By definition of the killed Markov chain, a_{ij} is the probability that the first target state, visited by an individual initially in state j, is $\alpha + i$. Similarly, p_j^a is the probability that an individual initially in state j reaches a target state. In the Southern Fulmar example with target set $\mathcal{B}_b = \{3, 4\}$, the entry a_{11} is the probability that the first breeding attempt of a newborn is successful, and p_1^a is the probability that a newborn attempts breeding.

3.2. The conditional Markov chain $\mathbf{P}_\mathbb{C}$

The transition matrix $\mathbf{P}_\mathbb{C}$ describes the transition probabilities of the killed Markov chain, conditional on absorption in the target set. Therefore, the absorbing states are $\mathcal{B}$ and the new transient set is $\mathcal{B}^c$. The state space of the conditional Markov chain is $\mathcal{T}$, and its transition probability matrix is

$$\mathbf{P}_\mathbb{C} = \left(\begin{array}{c|c} \mathbf{U}_\mathbb{C} & \mathbf{0} \\ \hline \mathbf{M}_\mathbb{C} & \mathbf{I}_\beta \end{array} \right) \quad \text{of size} \quad \left(\begin{array}{c|c} \alpha \times \alpha & \alpha \times \beta \\ \hline \beta \times \alpha & \beta \times \beta \end{array} \right). \tag{13}$$

Note that the death state does not belong to the state space of the conditional Markov chain because all its trajectories are absorbed by the target set before death is reached. The entry $u_{ij}^\mathbb{C}$ is the conditional probability of the transition form state j to state i given that the individual will eventually enter in the set $\mathcal{B}$. The entry $m_{ij}^\mathbb{C}$ is the conditional transition probability from the state j to the target state $\alpha + i$, given that the individual will eventually enter in the set $\mathcal{B}$. The columns of $\mathbf{P}_\mathbb{C}$ sum to 1. The block matrices $\mathbf{U}_\mathbb{C}$ and $\mathbf{M}_\mathbb{C}$ are given by

$$\mathbf{U}_\mathbb{C} = \mathbf{D}_a \mathbf{U}_\mathbb{K} \mathbf{D}_a^{-1} \quad \text{and} \quad \mathbf{M}_\mathbb{C} = \mathbf{K} \mathbf{D}_a^{-1}, \tag{14}$$

where

$$\mathbf{D}_a = \mathrm{diag}\,(\mathbf{p}_a) \tag{15}$$

is a diagonal matrix with, on the diagonal, the probabilities of absorption in the target states, $\mathbf{p}_a$ (defined in Eq. (11)).

In Appendix A.2, we provide a formal proof that the matrix $\mathbf{P}_\mathbb{C}$, defined by Eqs. (13) and (14), describes the Markov chain whose trajectories are precisely those of the killed Markov chain that do not encounter death before entering the target set. This proof is deeply inspired by the proof, written by Iosifescu (1980) (Section 3.2.9), in the special case where the target set is a single state.

Example. Fig. 4 shows the graphs of the conditional Markov chains associated with the target states $\mathcal{B}_a = \{3, 4\}$ and $\mathcal{B}_a = \{2, 3, 4\}$, respectively.

3.3. The sub-Markov chain $\mathbf{P}_\mathbb{S}$

The transition matrix $\mathbf{P}_\mathbb{S}$ contains, for each pair of target states i and j, the probabilities that an individual currently in state i returns to the target set in state j or dies before returning to the target set. Note that the return may take more than one time step. The transient state space is $\mathcal{B}$, and d is the only absorbing state. The state space of the sub-Markov chain is $\mathcal{B} \cup \{d\}$. The transition probability matrix is

$$\mathbf{P}_\mathbb{S} = \left(\begin{array}{c|c} \mathbf{U}_\mathbb{S} & \mathbf{0} \\ \hline \mathbf{m}_\mathbb{S}^\mathsf{T} & 1 \end{array} \right) \quad \text{of size} \quad \left(\begin{array}{c|c} \beta \times \beta & \beta \times 1 \\ \hline 1 \times \beta & 1 \times 1 \end{array} \right). \tag{16}$$

The entry $u_{ij}^\mathbb{S}$ is the probability that an individual starting in state $\alpha + j$ reaches the state $\alpha + i$ without passing through any other state in $\mathcal{B}$. The entry $m_i^\mathbb{S}$ is the probability that an individual starting in state $\alpha + i$ dies before it reaches a state in $\mathcal{B}$. Since the columns of the matrix $\mathbf{P}_\mathbb{S}$ sum to one,

$$\mathbf{m}_\mathbb{S}^\mathsf{T} = \mathbf{1}_\beta^\mathsf{T} - \mathbf{1}_\beta^\mathsf{T} \mathbf{U}_\mathbb{S}, \tag{17}$$

and the only variable is the matrix $\mathbf{U}_\mathbb{S}$. In Appendix A.1, we derive the matrix $\mathbf{U}_\mathbb{S}$ from the matrix $\mathbf{P}$,

$$\mathbf{U}_\mathbb{S} = \mathbf{AL} + \mathbf{Q} \quad \text{of size } \beta \times \beta. \tag{18}$$

The matrix $\mathbf{Q}$ (defined in Eq. (7)) contains the probabilities that an individual starting in a target state makes a one-step transition to a target state. The product $\mathbf{AL}$ describes the probabilities that the individual first leaves, and then re-enters, the target set. The matrix $\mathbf{L}$ is defined in Eq. (7), and the matrix $\mathbf{A}$ is defined in Eq. (11).
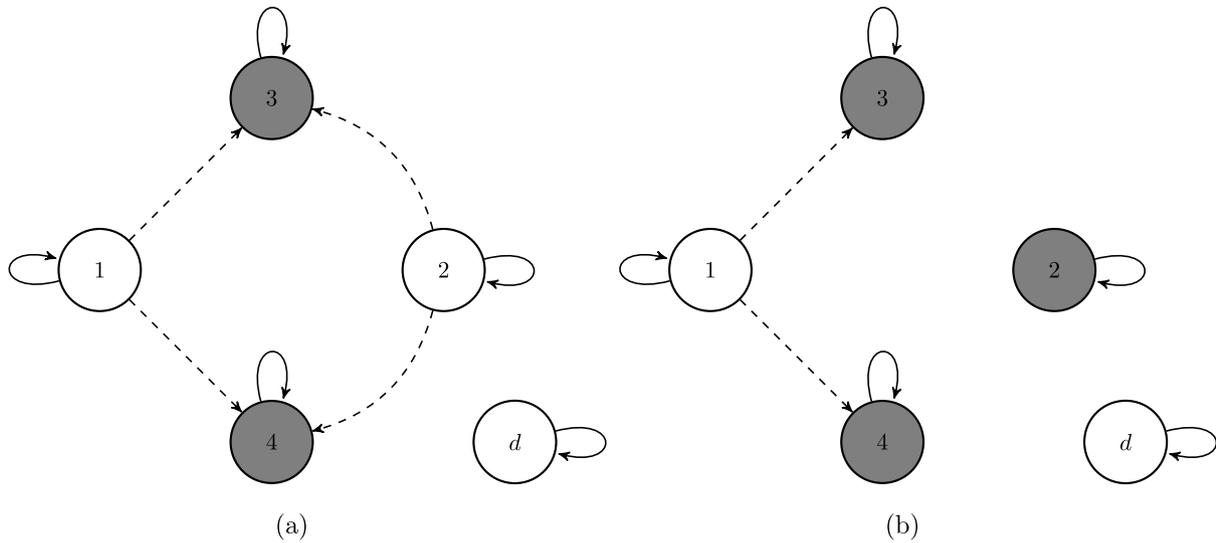
Fig. 3. Graphs of the killed Markov chain associated with the sets {3, 4} (a), and {2, 3, 4} (b), respectively. The grey circles are the target states, which are absorbing states for the killed Markov chain. The dashed arrows are the transitions into the absorbing states. To lighten the graph we omit the transitions to the death state.
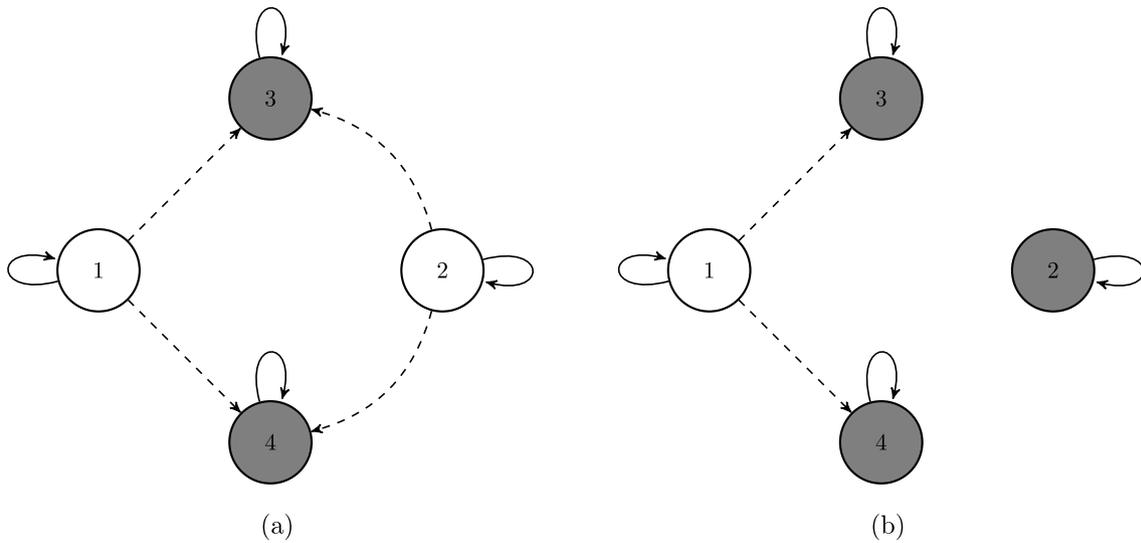

Fig. 4. Graphs of the conditional chain associated with the sets {3, 4} (a), and {2, 3, 4} (b), respectively. The grey circles are the target states. The dashed arrows are the transitions into the absorbing states.

Note that one time-step for the sub-Markov chain corresponds to a random number of time-steps for the initial Markov chain. However, the number of passages in a specific state in $\mathcal{B}$ starting in $\mathcal{B}$ is equivalent for both chains (see Section 4 for more details).

Example. Fig. 5 shows the graphs of the conditional Markov chains associated with the target states $\mathcal{B}_a = \{3, 4\}$ and $\mathcal{B}_b = \{2, 3, 4\}$, respectively.

4. Occupancy time in the target states

The occupancy time in the target states is the time spent in the target states over the individual's life. Because the pathways taken by individuals through their life is stochastic, the occupancy time is a random variable; each of its realisation is associated with the realised pathway of an individual. In this section, we calculate the distribution of the occupancy time in $\mathcal{B}$ and we provide formulae for its mean and variance, and for any of its higher moments.

Let τ_i denote the occupancy time in the set $\mathcal{B}$ for an individual initially in state i. The occupancy times are grouped into two

vectors, $\boldsymbol{\tau}_{\mathrm{in}}$ and $\boldsymbol{\tau}_{\mathrm{out}}$, depending on the initial state being in or out of the set $\mathcal{B}$, and both of the vectors are concatenated in the vector $\boldsymbol{\tau}_{\mathcal{B}}$,

$$\boldsymbol{\tau}_{\mathcal{B}} = (\underbrace{\tau_1, \ldots, \tau_\alpha}_{=:\boldsymbol{\tau}_{\mathrm{out}}}, \underbrace{\tau_{\alpha+1}, \ldots, \tau_\omega}_{=:\boldsymbol{\tau}_{\mathrm{in}}})^{\mathsf{T}}. \tag{19}$$

Let $\boldsymbol{\tau}_{\mathcal{B}}^k$, $\boldsymbol{\tau}_{\mathrm{out}}^k$, and $\boldsymbol{\tau}_{\mathrm{in}}^k$ denote the kth moment of the random vectors $\boldsymbol{\tau}_{\mathcal{B}}$, $\boldsymbol{\tau}_{\mathrm{out}}$, and $\boldsymbol{\tau}_{\mathrm{in}}$, respectively.

Moments of the occupancy time. First, we consider an individual initially in a state within the target set $\mathcal{B}$. To calculate its occupancy time in $\mathcal{B}$, we follow its pathway, and count the number of visits to target states. The parts of the pathway between two target-state visits are irrelevant for this count. Hence, it is sufficient to follow the pathway of the sub-Markov chain, and count its number of visits to $\mathcal{B}$. That is the number of steps in which the sub-Markov chain is in one of its transient states. Iosifescu (1980) (Theorem 3.2) provides a recursion formula for the moments of this number, which translates here into a formula for the moments of
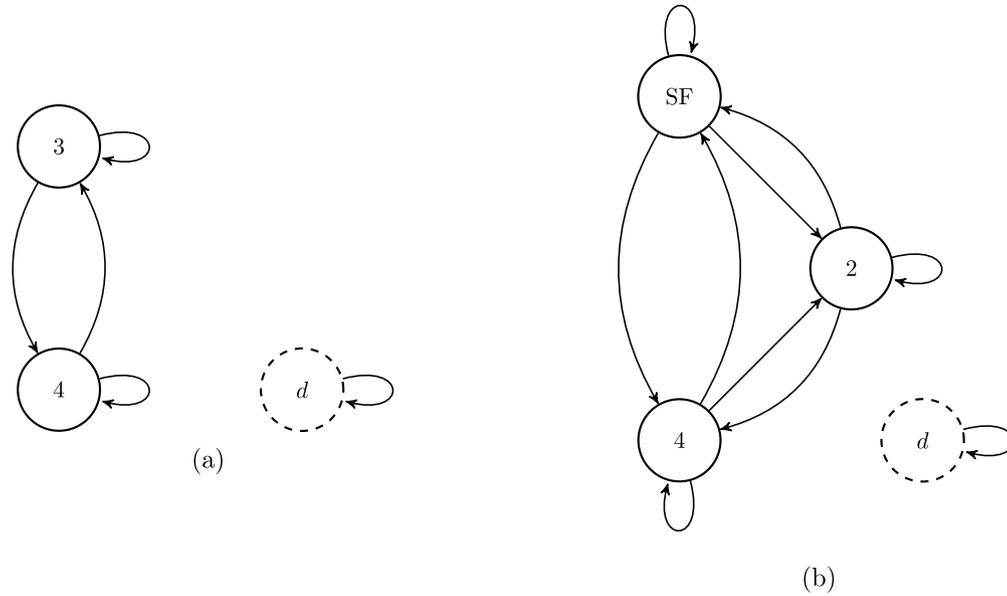
Fig. 5. Graphs of the sub-Markov chain associated with the sets {3, 4} (a), and {2, 3, 4} (b), respectively. The dashed arrows are the transitions into the absorbing states. The dashed circles are the target states. To lighten the graph we omit the transitions to the death state.

the occupancy time in the target set $\mathcal{B}$ for an individual initially in a target state:

$$\boldsymbol{\tau}_{\text{in}}^1 = \mathbf{N}_{\mathbb{S}}^{\mathsf{T}} \mathbf{1}_\beta \tag{20}$$

$$\boldsymbol{\tau}_{\text{in}}^2 = (\mathbf{N}_{\mathbb{S}}^{\mathsf{T}} - \mathbf{I}_\beta) 2 \boldsymbol{\tau}_{\text{in}}^1 + \boldsymbol{\tau}_{\text{in}}^1 \tag{21}$$

$$= (2\mathbf{N}_{\mathbb{S}}^{\mathsf{T}} - \mathbf{I}_\beta) \boldsymbol{\tau}_{\text{in}}^1 \tag{22}$$

$$\boldsymbol{\tau}_{\text{in}}^k = (\mathbf{N}_{\mathbb{S}}^{\mathsf{T}} - \mathbf{I}_\beta) \sum_{r=1}^{k-1} \binom{k}{r} \boldsymbol{\tau}_{\text{in}}^r + \boldsymbol{\tau}_{\text{in}}^1, \quad \text{for } k \geq 2, \tag{23}$$

where $\mathbf{N}_{\mathbb{S}} = (\mathbf{I}_\beta - \mathbf{U}_{\mathbb{S}})^{-1}$ is the fundamental matrix associated with the *sub* Markov chain.

From Eqs. (20) and (21), we obtain a formula for the variance of the occupancy time in $\mathcal{B}$ for an individual initially in a state within the target set $\mathcal{B}$,

$$\text{Var}(\boldsymbol{\tau}_{\text{in}}^{\mathsf{T}}) = \mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}} (2\mathbf{N}_{\mathbb{S}} - \mathbf{I}_\beta) - (\mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}}) \circ (\mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}}). \tag{24}$$

Second, we consider an individual initially in a non-target state. Since the sub-Markov chain is not defined outside the target states, we cannot directly apply the technique used above. However, we know that, either the individual never enters $\mathcal{B}$, in which case its occupancy time is zero, or it reaches some target state j. In the latter case, the probabilistic fate of the individual after it reaches j is equivalent to the fate of an individual that starts in the target state j; this is the *strong Markov property* (see e.g., Meyn and Tweedie (2009)). Thus, the conditional mean occupancy time, given the first-reached target state j, is the mean occupancy time of an individual initially in j. The unconditional mean occupancy time is the average, over the target states, of the means of the occupancy time of an individual initially in a target state, weighted by the probabilities to reach first these target states. This holds for all moments, and is translated in matrix notation into

$$\boldsymbol{\tau}_{\text{out}}^k = \mathbf{A}^{\mathsf{T}} \boldsymbol{\tau}_{\text{in}}^k \tag{25}$$

where the matrix $\mathbf{A}$ contains the probabilities of reaching the target states (defined in Eq. (11)), and the vector $\boldsymbol{\tau}_{\text{in}}^k$ is given by Eq. (23). In Appendix A.3, we provide a formal proof of Eq. (25).

From Eq. (25), we obtain a formula for the variance of the occupancy time in $\mathcal{B}$ for an individual initially in a non-target

state,

$$\text{Var}(\boldsymbol{\tau}_{\text{out}}^{\mathsf{T}}) = \mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}} (2\mathbf{N}_{\mathbb{S}} - \mathbf{I}_\beta) \mathbf{A} - (\mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}} \mathbf{A}) \circ (\mathbf{1}_\beta^{\mathsf{T}} \mathbf{N}_{\mathbb{S}} \mathbf{A}), \tag{26}$$

and

$$\text{Var}(\boldsymbol{\tau}_{\mathcal{B}}^{\mathsf{T}}) = \left[\text{Var}(\boldsymbol{\tau}_{\text{out}}^{\mathsf{T}}) : \text{Var}(\boldsymbol{\tau}_{\text{in}}^{\mathsf{T}}) \right]. \tag{27}$$

Distribution of the occupancy time. First, we calculate the distribution of the occupancy time in $\mathcal{B}$ for an individual initially in a state within the target set. Similarly to the moment calculations, this distribution is equivalent to the distribution of the number of steps in which the sub-Markov chain is in one of its transient states. Following Iosifescu (1980) (p. 104), we obtain

$$\left(\mathbb{P}\left(\tau_1^{\text{in}} = n\right), \dots, \mathbb{P}\left(\tau_\beta^{\text{in}} = n\right) \right)$$

$$= \begin{cases} \mathbf{0}_\beta^{\mathsf{T}} & \text{for } n = 0 \\ \mathbf{1}_\beta^{\mathsf{T}} (\mathbf{I}_\beta - \mathbf{U}_{\mathbb{S}}) \mathbf{U}_{\mathbb{S}}^{n-1} & \text{for } n \geq 1. \end{cases} \tag{28}$$

Second, we calculate the distribution of the occupancy time in $\mathcal{B}$ for an individual initially in a non-target state. Similarly to the moment calculations, we use the strong Markov property to deduce this distribution from the distribution of the occupancy time in $\mathcal{B}$ for an individual initially in a state within the target set, given in Eq. (28). The probability that the occupancy time is zero, however, requires a special attention. This probability is equal to the probability that the individual dies before it enters any target state; that is one minus the absorbing probability (see Eq. (12)). Hence, the distribution of the occupancy time in $\mathcal{B}$ for an individual initially in a non-target state is

$$\left(\mathbb{P}\left(\tau_1^{\text{out}} = n\right), \dots, \mathbb{P}\left(\tau_\alpha^{\text{out}} = n\right) \right)$$

$$= \begin{cases} \mathbf{1}_\alpha^{\mathsf{T}} - \mathbf{p}_a^{\mathsf{T}} & \text{for } n = 0 \\ \mathbf{1}_\beta^{\mathsf{T}} (\mathbf{I}_\beta - \mathbf{U}_{\mathbb{S}}) \mathbf{U}_{\mathbb{S}}^{n-1} \mathbf{A} & \text{for } n \geq 1. \end{cases} \tag{29}$$

5. Correlation between the occupancy times in two sets

Consider two subsets $\mathcal{B}_1$ and $\mathcal{B}_2$, of the transient set $\mathcal{T}$. In this section we calculate the correlation between the occupancy time in $\mathcal{B}_1$ and the occupancy time in $\mathcal{B}_2$. A positive (resp. negative)

correlation means that both of the occupancy times tend to have the same (resp. opposite) "behaviour", i.e. when one is greater than its mean, the other tends to be greater (resp. smaller) than its mean and vice versa when it is smaller. The correlation between $\tau_{\mathcal{B}_1}$ and $\tau_{\mathcal{B}_2}$ is defined as follows

$$\text{Corr}\left(\tau_{\mathcal{B}_1}, \tau_{\mathcal{B}_2}\right) = \frac{\text{Cov}\left(\tau_{\mathcal{B}_1}, \tau_{\mathcal{B}_2}\right)}{\sqrt{\text{Var}(\tau_{\mathcal{B}_1})\text{Var}(\tau_{\mathcal{B}_2})}}, \tag{30}$$

where the covariance between $\tau_{\mathcal{B}_1}$ and $\tau_{\mathcal{B}_2}$ is

$$\text{Cov}\left(\tau_{\mathcal{B}_1}, \tau_{\mathcal{B}_2}\right) = \mathbb{E}\left[\left(\tau_{\mathcal{B}_1} - \tau_{\mathcal{B}_1}^1\right)\left(\tau_{\mathcal{B}_2} - \tau_{\mathcal{B}_2}^1\right)\right]. \tag{31}$$

The denominator in the right hand side of Eq. (30) is directly calculated with the formulae (24) and (26) applied successively to $\mathcal{B}_1$ and $\mathcal{B}_2$. To calculate the numerator let us split the set $\mathcal{B}_1 \cup \mathcal{B}_2$ into three pairwise disjoint sets (possibly empty): the states in $\mathcal{B}_1$ that are not in $\mathcal{B}_2$:

$$\mathcal{C}_1 = \mathcal{B}_1 \cap \mathcal{B}_2^c, \tag{32}$$

the states in $\mathcal{B}_2$ that are not in $\mathcal{B}_1$:

$$\mathcal{C}_2 = \mathcal{B}_2 \cap \mathcal{B}_1^c, \tag{33}$$

and the states that are in $\mathcal{B}_1$ and in $\mathcal{B}_2$

$$\mathcal{C}_3 = \mathcal{B}_1 \cap \mathcal{B}_2. \tag{34}$$

Then we have

$$\text{Cov}\left(\tau_{\mathcal{B}_1}, \tau_{\mathcal{B}_2}\right) = \text{Cov}\left(\tau_{\mathcal{C}_1} + \tau_{\mathcal{C}_3}, \tau_{\mathcal{C}_2} + \tau_{\mathcal{C}_3}\right) \tag{35}$$

$$= \text{Cov}\left(\tau_{\mathcal{C}_1}, \tau_{\mathcal{C}_2}\right) + \text{Cov}\left(\tau_{\mathcal{C}_1}, \tau_{\mathcal{C}_3}\right)$$
$$+ \text{Cov}\left(\tau_{\mathcal{C}_2}, \tau_{\mathcal{C}_3}\right) + \text{Var}\left(\tau_{\mathcal{C}_3}\right). \tag{36}$$

The covariances in the right hand side of Eq. (36) are calculated with formula (86) applied to the appropriate sets. The variance $\text{Var}\left(\tau_{\mathcal{C}_3}\right)$ is calculated with formulae (24) and (26) applied to the set $\mathcal{C}_3$.

If the sets $\mathcal{B}_1$ and $\mathcal{B}_2$ are disjoint, Eq. (36) boils down to

$$\text{Cov}\left(\tau_{\mathcal{B}_1}, \tau_{\mathcal{B}_2}\right) = \frac{1}{2}\left[\text{Var}\left(\tau_{\mathcal{B}_1 \cup \mathcal{B}_2}\right)\right.$$
$$\left. - \text{Var}\left(\tau_{\mathcal{B}_1}\right) - \text{Var}\left(\tau_{\mathcal{B}_2}\right)\right] \tag{37}$$

(see Appendix A.4 for details). The variances $\text{Var}\left(\tau_{\mathcal{B}_1 \cup \mathcal{B}_2}\right)$, $\text{Var}\left(\tau_{\mathcal{B}_1}\right)$, and $\text{Var}\left(\tau_{\mathcal{B}_2}\right)$ are calculated with the formulae (24) and (26) applied to the sets $\mathcal{B}_1 \cup \mathcal{B}_2$, $\mathcal{B}_1$, and $\mathcal{B}_2$, respectively.

5.1. Correlation between occupancy time and longevity

In the particular case when the set $\mathcal{B}_2$ is the *entire* transient set, and $\mathcal{B}_1 = \mathcal{B}$ is another target set, the calculations above provide a formula for the correlation between occupancy time in $\mathcal{B}$ and the longevity.

Let η_i denote the longevity for an individual initially in state i. The longevity is the sum of the time spent in all the transient states before death, that is the occupancy time in the entire transient set $\mathcal{T}$ (see Caswell (2001)). Therefore the correlation between the occupancy time in the target states and the longevity of an individual can be calculated with the formula (30) applied to $\mathcal{B}_1 = \mathcal{B}$ and $\mathcal{B}_2 = \mathcal{T}$. In this particular case, the calculations above boil down to

$$\text{Corr}\left(\tau_{\mathcal{B}}, \eta\right) = \frac{\text{Cov}\left(\tau_{\mathcal{B}}, \tau_{\mathcal{B}^c}\right) + \text{Var}\left(\tau_{\mathcal{B}}\right)}{\sqrt{\text{Var}(\tau_{\mathcal{B}})\text{Var}(\eta)}}. \tag{38}$$

6. Behaviour of winners and losers

For want a better term, we call the individuals that eventually reach the target states *winners*. Individuals that never reach the target states are called *losers* (the terminology implies nothing about the desirability of the states). In the Southern Fulmar example with target set $\mathcal{B}_b$, the winners are individuals that eventually breed, successfully or not, and the losers are individuals that never breed. Two demographic variables are related to the behaviour of the winners: the time required to reach the target set, and the time to return to the target set. In the Southern Fulmar example with target set $\mathcal{B}_b$, these variables translate into the time to maturation, and the time interval between two breeding attempts. In this Section, we calculate the probability that an individual eventually becomes a winner, and the conditional probability distribution of the time to reach the target set, given that the individual eventually does so. We also provide formulae for the variance and any moments of this distribution. Then, we calculate the probability that a winner returns to the target set, after its first visit. Finally, we calculate the conditional probability distribution of the time to return to the target set, given that the individual eventually returns. We also provide formulae for the variance and any moments of this distribution.

6.1. Probability of reaching the target states

The probability that an individual eventually reaches the target set is the probability that the killed Markov chain is absorbed into a target state (see Section 3.1.1). Therefore the probabilities of being a winner, given the current state, are given by the entries of the vector $\mathbf{p}_a = \mathbf{1}^\mathsf{T}\mathbf{A}$,

$$\Pr\left(\text{individual in } i \text{ becoming a winner}\right) = p_i^a, \tag{39}$$

$$\Pr\left(\text{individual in } i \text{ becoming a loser}\right) = 1 - p_i^a. \tag{40}$$

6.2. Time to reach the target states

Let $t_i^{\mathcal{B}}$ denote the random time required to reach the target set $\mathcal{B}$ by a winner initially in the non target state i, and

$$\mathbf{t}_{\mathcal{B}} = (t_1^{\mathcal{B}}, \ldots, t_\alpha^{\mathcal{B}})^\mathsf{T}. \tag{41}$$

The time required by a winner to reach the target set is the occupancy time of the conditional Markov chain in its *entire* transient set. Indeed, by definition, the conditional Markov chain describes the transitions of winners before they enter the target set. Theorem 3.2 in Iosifescu (1980), applied to the conditional Markov chain, provides a recursion formula for the moments of the occupancy time, which translates here into a formula for the moments of the time required to reach the target states,

$$\mathbf{t}_{\mathcal{B}}^k = \left(\mathbf{N}_{\mathbb{C}}^\mathsf{T} - \mathbf{I}_\alpha\right)\sum_{r=1}^{k-1}\binom{k}{r}\mathbf{t}_{\mathcal{B}}^r + \mathbf{t}_{\mathcal{B}}^1, \tag{42}$$

where $\mathbf{N}_{\mathbb{C}} = (\mathbf{I}_\alpha - \mathbf{U}_{\mathbb{C}})^{-1}$ is the fundamental matrix associated with the conditional Markov chain. In particular, Eq. (42) yields the mean and variance of $\mathbf{t}_{\mathcal{B}}$,

$$\mathbb{E}[\mathbf{t}_{\mathcal{B}}^\mathsf{T}] = \mathbf{1}_\alpha^\mathsf{T}\mathbf{N}_{\mathbb{C}} \tag{43}$$

$$\text{Var}\left(\mathbf{t}_{\mathcal{B}}^\mathsf{T}\right) = \mathbf{1}_\alpha^\mathsf{T}\mathbf{N}_{\mathbb{C}}(2\mathbf{N}_{\mathbb{C}} - \mathbf{I}_\alpha) - \left(\mathbf{1}_\alpha^\mathsf{T}\mathbf{N}_{\mathbb{C}}\right) \circ \left(\mathbf{1}_\alpha^\mathsf{T}\mathbf{N}_{\mathbb{C}}\right). \tag{44}$$

Following Iosifescu (1980) (see p. 104), we also obtain a formula for the distribution of $\mathbf{t}_{\mathcal{B}}$

$$\left(\mathbb{P}\left(t_1^{\mathcal{B}} = n\right), \ldots, \mathbb{P}\left(t_\alpha^{\mathcal{B}} = n\right)\right) = \mathbf{1}_\alpha^\mathsf{T}(\mathbf{I}_\alpha - \mathbf{U}_{\mathbb{C}})\mathbf{U}_{\mathbb{C}}^{n-1}. \tag{45}$$

6.3. Return to the target states

A winner will eventually enter the target set but its life may not end there. It may leave $\mathcal{B}$ one time-step after its first entrance and never return, it may leave but return eventually, it may stay and leave later, or it may stay forever. An individual returns to the target set if it is in one of the target states at some time, and it visits a target state at a later time, but without passing through any target state in-between.

6.3.1. Return probabilities

The entries of the sub-Markov chain describe the probabilities for an individual to go – in possibly more than one time-step – from one target state to another, without passing through other target states in-between. Hence, the entry $u_{ij}^{\mathbb{S}}$ (Eq. (18)) is the probability that an individual in target state $\alpha + i$ returns to $\mathcal{B}$ in target state $\alpha + j$,

$$\Pr \text{ (individual currently in } i \text{ returns to } \mathcal{B} \text{ through } j) = u_{ji}^{\mathbb{S}}.$$
(46)

The probabilities of returning to $\mathcal{B}$ regardless in which state are given by the vector

$$\mathbf{p}_r = \mathbf{U}_{\mathbb{S}}^{\mathsf{T}} \mathbf{1}_{\beta}.$$
(47)

6.3.2. Return time

The return time to the target set is the time between two visits to this set. For a demographic Markov chain, it is the time interval between two instances of the demographic events described by the target set, e.g., the number of years between two breeding attempts. This time interval can be as small as 1 if the individual breeds the next year, but it can be infinite if the individual dies before breeding again. To avoid the later scenario, we calculate the conditional probability distribution of the return time, given that the individual does eventually return to $\mathcal{B}$. Caswell (2006) provides formulae for the distribution and the moments of the interval between demographic events that are described by a *single* target state. By using a different method, we generalise this result to events described by several target states.

Let μ_i denote the return time for an individual initially in target state $\alpha + i$, and

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{\beta})^{\mathsf{T}}.$$
(48)

To calculate the probability distribution of $\boldsymbol{\mu}$, we need to calculate the one-time-step transition probabilities from target states to target states, and from target states to non-target states, conditional on eventual return to the target set.

Let $\mathbf{W}_{\text{in}}$, of size $\beta \times \beta$, denote the matrix which contains the conditional probabilities that an individual initially in a target state moves to a target state in one step, given that it eventually returns to the target set. By definition of conditional probabilities, we have

$$\mathbf{W}_{\text{in}} = \mathbf{Q} \mathbf{D}_r^{-1} \quad \text{of size } \beta \times \beta,$$
(49)

where diagonal of $\mathbf{D}_r = \text{diag}(\mathbf{p}_r)$ contains the return probabilities, as defined in Eq. (47), and $\mathbf{Q}$ contains the unconditional probabilities that an individual initially in a target state moves to a target state, as defined in Eq. (7). Eq. (49) is derived in detail in Appendix A.5.

Let $\mathbf{W}_{\text{out}}$ of size $\alpha \times \beta$ denote the matrix which contains the conditional probabilities that an individual initially in a target state moves to a non-target state, given that it eventually returns to the target set. Similar to the calculation of the transition probabilities of the conditional Markov chain (Section 3.2), the calculation of $\mathbf{W}_{\text{out}}$ follows from the definition of conditional probabilities and the Markov properties of the original chain,

$$\mathbf{W}_{\text{out}} = \mathbf{D}_a \mathbf{L} \mathbf{D}_r^{-1} \quad \text{of size } \alpha \times \beta,$$
(50)

where $\mathbf{D}_r = \text{diag}(\mathbf{p}_r)$ and $\mathbf{D}_a = \text{diag}(\mathbf{p}_a)$, respectively, contain on their diagonal the probabilities of return to, and the probabilities of absorption in, the target states, as defined in Eqs. (11) and (47). The matrix $\mathbf{L}$ contains the unconditional probabilities that an individual initially in a target state moves to a non-target state, as defined in Eq. (7). Eq. (50) is derived in detail in Appendix A.5.

The probability distribution of $\boldsymbol{\mu}$ is derived from the matrices $\mathbf{W}_{\text{in}}$, $\mathbf{W}_{\text{out}}$, and the transition matrix $\mathbf{U}_{\mathbb{C}}$ of the conditional Markov chain. The return time of an individual initially in the target set is equal to one if its first transition is into the target set. The probabilities of those transitions are obtained by summing the columns of matrix $\mathbf{W}_{\text{in}}$, i.e.

$$\left(\mathbb{P}(\mu_1 = 1), \ldots, \mathbb{P}(\mu_{\beta} = 1) \right) = \mathbf{1}_{\beta}^{\mathsf{T}} \mathbf{W}_{\text{in}}.$$
(51)

To calculate the probabilities that the return time is n, with $n \geq 2$, we use the Markov property. An individual with a return time equal to n has first moved to a non-target state, and from there it has reached the target set in $n - 1$ steps. Hence, the probability that the return time is equal to n is the sum, over all non-target states, of the product of the probability that an individual moves to a non-target state and the probability that an individual reaches the target set from this non-target state, in $n - 1$ steps. Hence,

$$\left(\mathbb{P}(\mu_1 = n), \ldots, \mathbb{P}(\mu_{\beta} = n) \right)$$
$$= \left[\mathbf{1}_{\alpha}^{\mathsf{T}} (\mathbf{I}_{\alpha} - \mathbf{U}_{\mathbb{C}}) \mathbf{U}_{\mathbb{C}}^{n-2} \right] \mathbf{W}_{\text{out}}, \quad \text{for } n \geq 2.$$
(52)

The probabilities that an individual initially in a target state moves to a non-target set are given by the column-sums of the matrix $\mathbf{W}_{\text{out}}$. The bracketed term is the probability of reaching the target set in $n - 1$ steps, as given in Eq. (45).

In Appendix A.5, we also derive a formula for the moments of $\boldsymbol{\mu}$. In particular, this formula yields the mean and variance of $\boldsymbol{\mu}$,

$$\mathbb{E}[\boldsymbol{\mu}] = \mathbf{1}_{\beta} + \mathbf{W}_{\text{out}}^{\mathsf{T}} \mathbf{t}_{\mathcal{B}}^1$$
(53)

$$\text{Var}(\boldsymbol{\mu}) = \mathbf{W}_{\text{out}}^{\mathsf{T}} \mathbf{t}_{\mathcal{B}}^2 - \mathbf{W}_{\text{out}}^{\mathsf{T}} \mathbf{t}_{\mathcal{B}}^1 \circ \mathbf{W}_{\text{out}}^{\mathsf{T}} \mathbf{t}_{\mathcal{B}}^1.$$
(54)

7. Example: the Southern Fulmar

The Southern Fulmar is an ice-dependent seabird. This species breeds along the coast of Antarctica and outlying islands, and its individuals forage near the ice edge. Because of its breeding and foraging habitat, the vital rates of the Southern Fulmar are, to some extent, dependent on the sea ice condition. Using the life cycle illustrated in Fig. 1, Jenouvrier et al. (2015) estimated matrix population models for three ice conditions: favourable, ordinary, and unfavourable. From these models, we derive three transient state transition matrices, $\mathbf{U}_{\text{fav}}$, $\mathbf{U}_{\text{ord}}$, and $\mathbf{U}_{\text{unfav}}$, describing transitions and survival in a favourable, ordinary, and unfavourable ice conditions, respectively. These matrices are provided in the Supplementary Material.

We choose to focus on the super-state "breeding attempt"; that is, the target states $\mathcal{B}_{\text{b}} = \{3, 4\}$. This set plays an important role in the developmental description of an individual. The pre-breeder state describes juvenile individuals, the non-breeder state describes adult individuals that are skipping reproduction, and $\mathcal{B}_{\text{b}}$ captures the actively reproducing individuals. In this section, we first show the construction of the killed Markov chain, the conditional Markov chain, and the sub Markov chain for $\mathcal{B}_{\text{b}}$. Then, we illustrate the calculations listed in Table 1, describe their biological meaning for the target set $\mathcal{B}_{\text{b}}$, and compare them between the
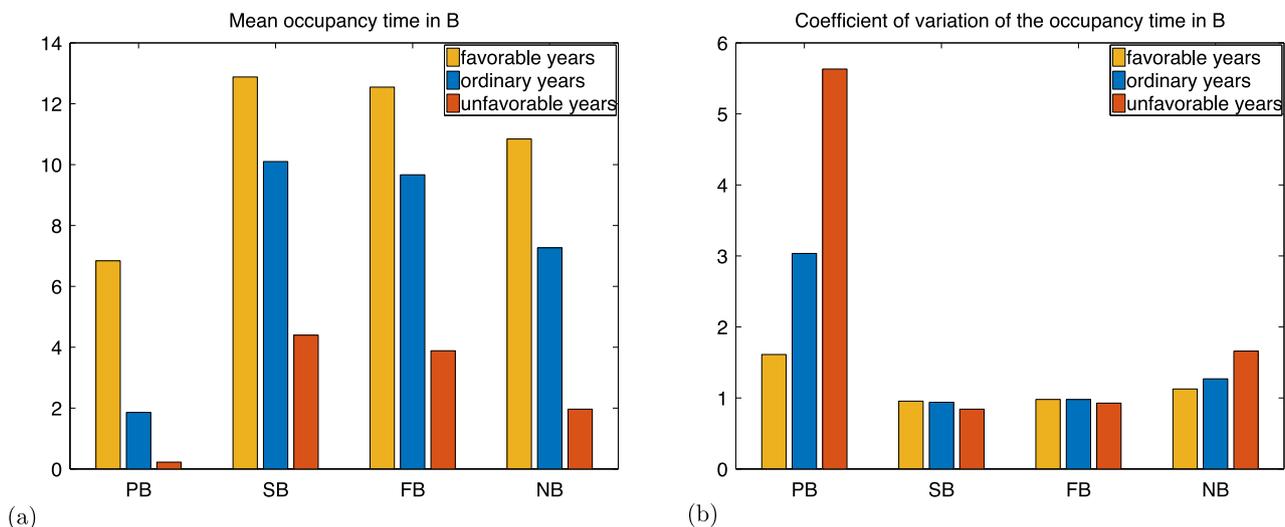
Fig. 6. Mean (a) and coefficient of variation (b) of the occupancy time in the set $\mathcal{B}_b$ for an individual initially in the states pre-breeder (PB), successful breeder (SB), failed breeder (FB), and non breeder (NB).

three ice conditions. In the Supplementary Material, we provide the MATLAB code for calculating all the measures presented here. Due to space constraints, we only show matrices associated with the ordinary ice conditions.

Original Markov chain. The transition probabilities between the transient states of the original Markov chain are extracted from Jenouvrier et al. (2015),

$$\mathbf{U}_{\text{ord}} = \begin{pmatrix} 0.9 & 0 & 0 & 0 \\ 0 & 0.63 & 0.07 & 0.18 \\ 0.01 & 0.18 & 0.67 & 0.49 \\ 0.01 & 0.09 & 0.19 & 0.24 \end{pmatrix}. \tag{55}$$

The death probability vector $\mathbf{m}$ is deduced from $\mathbf{U}_{\text{ord}}$,

$$\mathbf{m} = \mathbf{1}_4 - \mathbf{U}^\mathsf{T}\mathbf{1}_4 = \begin{pmatrix} 0.08 \\ 0.1 \\ 0.07 \\ 0.09 \end{pmatrix}. \tag{56}$$

Constructing the killed Markov chain. The killed chain is the original chain that is stopped as soon as it enters $\mathcal{B}_b$. The transient states of the killed chain are $\{1, 2\}$, and the absorbing states are $\{3, 4, d\}$. Fig. 3(a) shows the graph of the killed Markov chain. The variables required to describe the transition matrix $\mathbf{P}_\mathbb{K}$ (Eq. (8)) are the matrices $\mathbf{U}_\mathbb{K}$ and $\mathbf{K}$ (both extracted from $\mathbf{U}_{\text{ord}}$ as shown in Eq. (7)), and the vector $\mathbf{m}_2$ containing the death probabilities from the non-target states (as defined in Eq. (9)):

$$\mathbf{U}_\mathbb{K} = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.63 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 0.01 & 0.18 \\ 0.01 & 0.09 \end{pmatrix}, \text{ and}$$

$$\mathbf{m}_2 = \begin{pmatrix} 0.08 \\ 0.1 \end{pmatrix}. \tag{57}$$

Constructing the conditional Markov chain. The conditional Markov chain is the killed chain, conditional on eventual absorption in the target states. The transient states of the conditional chain are $\{1, 2\}$, and the absorbing state are $\{3, 4\}$. Fig. 4(a) shows the graph of the conditional Markov chain. The variables required to construct the transition matrix $\mathbf{P}_\mathbb{C}$ Eq. (8) are the matrices $\mathbf{U}_\mathbb{C}$ and $\mathbf{m}_\mathbb{C}$, both defined in Eq. (14),

$$\mathbf{U}_\mathbb{C} = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.63 \end{pmatrix}, \quad \mathbf{M}_\mathbb{C} = \begin{pmatrix} 0.061 & 0.25 \\ 0.037 & 0.12 \end{pmatrix}. \tag{58}$$

Constructing the sub Markov chain. The sub Markov chain is the original chain observed through a filter that makes visible only the target states and the death state. The transient states of the sub chain are the target states $\{3, 4\}$, and the absorbing state is $\{d\}$. Fig. 5(a) shows the graph of the sub-Markov chain. The variables required to construct the transition matrix $\mathbf{P}_\mathbb{S}$ (Eq. (16)) are the matrix $\mathbf{U}_\mathbb{S}$ (defined in Eq. (18)), and the vector $\mathbf{m}_\mathbb{S}$ (defined in Eq. (17),

$$\mathbf{U}_\mathbb{S} = \begin{pmatrix} 0.7 & 0.58 \\ 0.21 & 0.29 \end{pmatrix}, \quad \mathbf{m}_\mathbb{S} = \begin{pmatrix} 0.09 \\ 0.13 \end{pmatrix}. \tag{59}$$

Now, we illustrate the calculations listed in Table 1 for each ice condition.

Occupancy time in $\mathcal{B}_b$.

The occupancy time in the set $\mathcal{B}_b$ is the *lifetime number of breeding attempts*. Fig. 6 illustrates the mean and coefficient of variation of the lifetime number of breeding attempts (i.e. occupancy time in $\mathcal{B}_b$), calculated with formulae (20), (24), (25), and (26), applied to $\mathcal{B} = \mathcal{B}_b$.

The mean number of breeding attempt is extremely low under unfavourable ice conditions (Fig. 6). This is a consequence of the long time to maturation under unfavourable ice conditions. In fact, the probability of maturation in unfavourable conditions is only 0.0544. Thus, less than 1/20th of the newborns mature; 19/20 never attempt breeding. However, those that do mature have a relatively large number of breeding attempts (see Fig. 6(a), initial states successful breeder and failed breeder). This large difference between the mean number of breeding attempt of newborns and adults explains partially the large standard deviation of the number of breeding attempts of newborns, which is almost six times the mean, as shown in Fig. 6(b). In contrast, under favourable conditions, half of the newborns will reach maturity, which brings the mean number of breeding attempt for newborns closer to the mean number of breeding attempt for adults. As a consequence, the standard deviation of the number of breeding attempts is less than twice its mean.

Fig. 7 illustrates the probability distribution of the lifetime number of breeding attempt for adults under each ice condition, calculated by applying Eq. (28) to the set $\mathcal{B}_b$. These distributions are approximately geometric, with parameter $1 - p$, where $p = \frac{1}{2}(p_1^r + p_2^r)$ is the average probability of returning to $\mathcal{B}_b$ (see
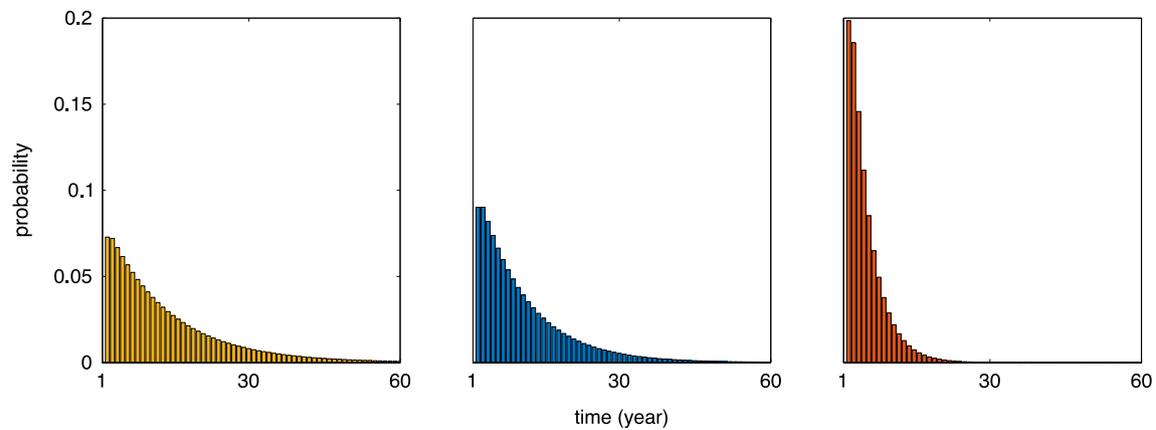
Fig. 7. Distributions of the occupancy time in the set $\mathcal{B}_b$ for an individual starting in the state successful breeder under favourable condition (left), ordinary condition (centre), and unfavourable condition (right).

Eq. (47)), which ranges from 0.9272 (favourable conditions) to 0.6951 (unfavourable conditions). The mean and coefficient of variation of this distribution are $\frac{1}{1-p}$ and $\sqrt{p}$, respectively. This explains why the relative decrease of the mean number of breeding events after maturation between favourable ice conditions and unfavourable conditions is approximately 0.7, but the relative decrease of the coefficient of variation is only 0.1, as shown in Fig. 7.

Correlation between occupancy time and longevity.

The correlation between the longevity of an individual and its number of breeding attempts provides information on how longevity affects the total number of breeding attempts. Using formula (38), we find a positive correlation (Fig. 8) for each ice condition. Indeed, we expect that on average a particularly large longevity results in a particularly large number of breeding attempt and vice versa. Because newborn individual cannot breed before maturity, the correlation between longevity and breeding attempts for newborn is smaller than the correlation for adults. Under favourable ice conditions, individuals spend most of their lives in $\mathcal{B}_b$, and the correlation approaches 1.

Time required to reach $\mathcal{B}_b$.

The target states in $\mathcal{B}_b$ describe active reproduction, and the time to reach $\mathcal{B}_b$ is the *time to maturation*. Fig. 9 illustrates the mean and coefficient of variation of this time, calculated with formulae (43) and (44) applied to $\mathcal{B} = \mathcal{B}_b$. The average maturation time under unfavourable ice conditions among individuals that do successfully mature is twice as large as that under favourable ice conditions.

Time required to return to $\mathcal{B}_b$.

This time corresponds to the time between two breeding attempts, which is a measure of inter-birth interval and an indicator of the individual's reproduction consistency. This time is 1 for individuals that never skip reproduction, and is infinite for individuals that reproduce only once. In terms of the Markov chain, this time is the return time to the set $\mathcal{B}_b$. The mean return time increases as the ice condition becomes less favourable (Fig. 10). Unlike for the maturation time, the coefficient of variation of the time between breeding increases significantly as well. One reason is that non-breeding individuals have a smaller probability of attempting breeding the following year under unfavourable ice conditions ($p_2^a = 0.19$) than under favourable ice conditions ($p_2^a = 0.66$). This results in more variation in the distribution of the time required to
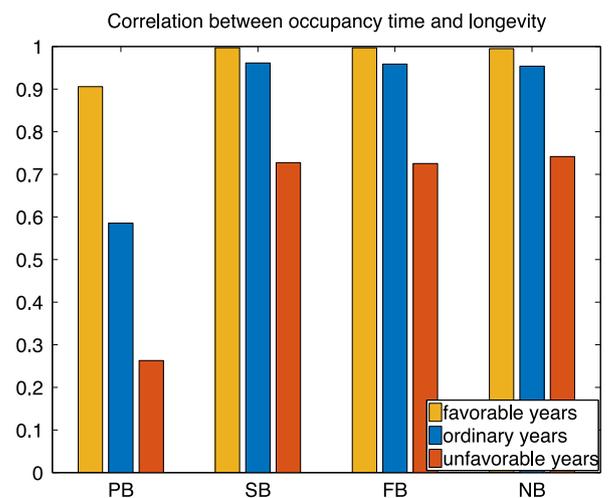

Fig. 8. Correlation between occupancy time in $\mathcal{B}_b$ and the expected longevity for an individual initially in the states pre-breeder (PB), successful breeder (SB), failed breeder (FB), and non breeder (NB).

attempt breeding after a non breeding season under unfavourable ice conditions than under favourable ice conditions, as shown in Figures S1 and S2 provided in the Supplementary Material.

8. Discussion

From birth to death, an individual goes through a sequence of random events (e.g., surviving, developing, mating, reproducing, growing, dispersing, moving among social or occupational classes, or changing health status). This results in a stochastic pathway through the individual state space. Many important life history traits are direct functions of these pathways (e.g., longevity, reproductive output, age at maturity) and are often related to the concept of occupancy time in a state. Absorbing Markov chains are a powerful model framework to analyse these stochastic pathways, but current theory can only calculate the occupancy times in a *single* state. However, often the occupancy time in a *set* of states is desired. For example, a size range in a size-classified model, an age class in a age×stage model, or a set of locations in a spatial model are all sets of states. We have presented a new mathematical approach to absorbing Markov chains that generalises the occupancy calculations to sets of states and enlarges the list of demographic measures that can be calculated from a demographic model.
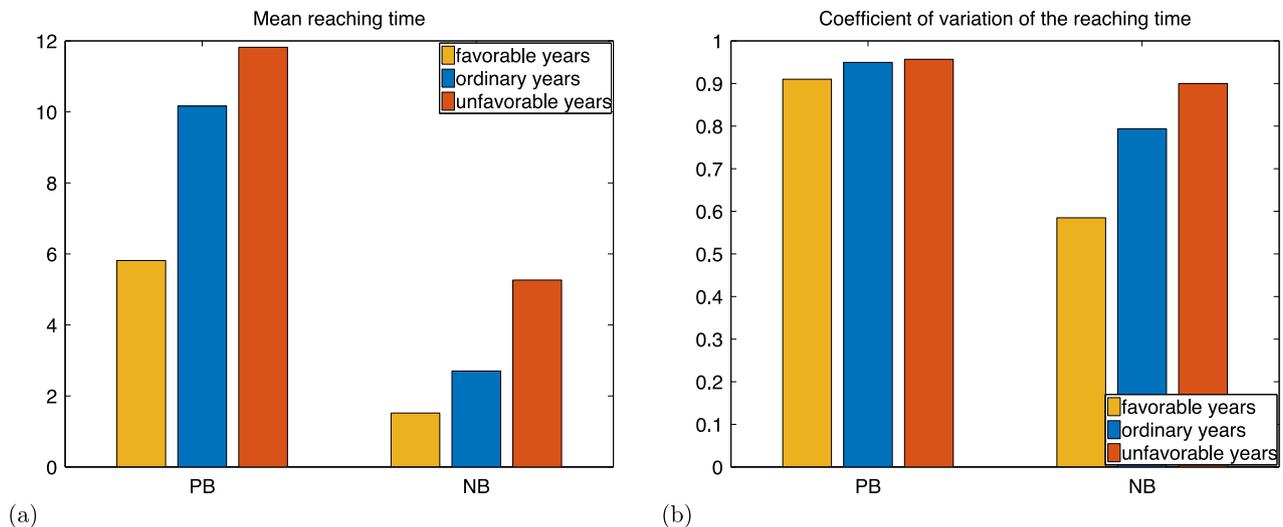
Fig. 9. Mean (a) and coefficient of variation (b) of the time to reach the set $\mathcal{B}_b$ for individual initially in the states pre-breeder (PB), and non breeder (NB).
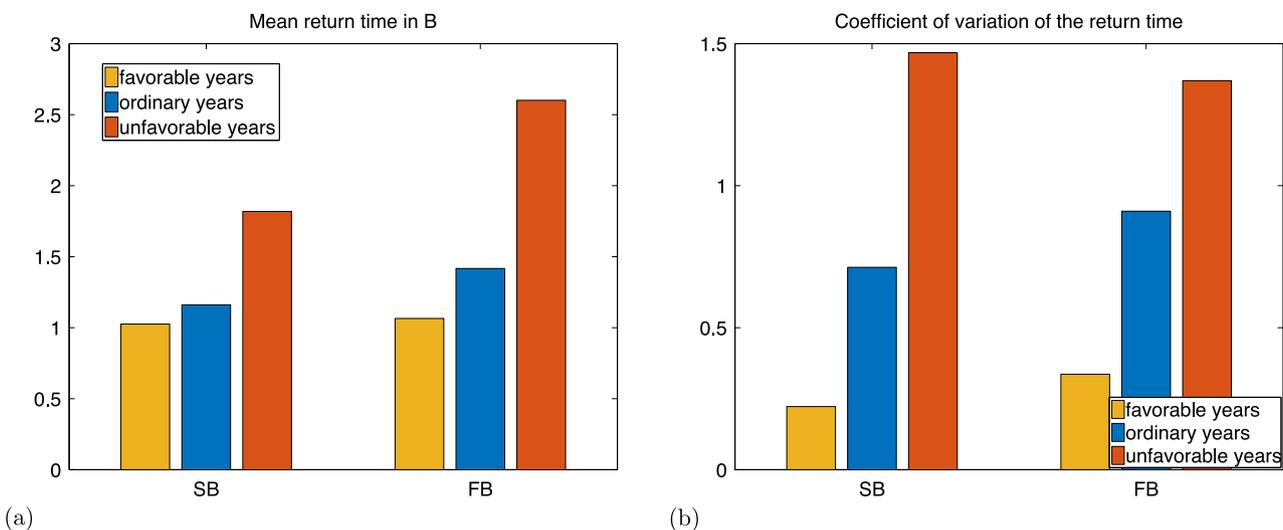

Fig. 10. Mean (a) and coefficient of variation (b) of the time to return to the set $\mathcal{B}_b$ for an individual initially in the states successful breeder (SB), and failed breeder (FB).

For any number of target states, we provide formulae to calculate any moments (e.g., mean and variance) and the probability distribution of

- the occupancy time in a target set
- the time required to reach a target set
- the time required to return to a target set.

Last but not least, we provide a formula to calculate the correlation between the occupancy times in two different sets. These formulae are straightforward to calculate in matrix oriented software.

In demographic models based on absorbing Markov chains, any life history trait that is a function of the individual's pathway is stochastic. The calculation of the resulting distribution of this trait among individuals is complicated, and is often approximated by its mean and, sometime, its variance. While the current theory can calculate any moment of the occupancy time in a single state, our generalisation to the set occupancy time provides also a formula for its probability distribution.

As the questions in population ecology become more sophisticated and data become more detailed, the use of multistate and hyperstate matrix models will become increasingly important (Roth and Caswell, 2016). For these models, it is often useful to consider marginal sets. For example, in a age×stage-structured model, the marginal set age x describes individuals in age class x independently of their stage. Similarly, an increasingly number of studies use integral projection models (Ellner et al., 2016) which are approximated – and then analysed – by matrix projection models with highly detailed i-state space (e.g., large number of size classes). For these models, the occupancy in a single state is often of no interest, while the occupancy in a set of state can be biological important (e.g., range of ages or sizes corresponding to a specific developmental stages). Our work generalises the occupancy calculations for these two highly powerful family of models.

In human demography, multistate models often combine age classes with health or social categories. Studies may focus on the *expected occupancy times* (e.g., Siegel (2012) and Clark et al. (2016)), but very little use is made of the Markov chain framework presented here. Applying our results to such human demographic models would significantly improve the calculations of the occupancy by providing the variance and its probability distribution.

The only data required for our calculations are the transition probabilities between the states that describe the individuals. These probabilities can be obtained from most population projection matrices. The availability of projection matrix data has been

improved by the databases COMPADRE for plant and COMADRE for animals (Salguero-Gómez et al., 2016, 2015). Hence, our approach to absorbing Markov chain makes the analysis of occupancy times accessible for a wide range of existing projection matrix models.

The power of our results is to provide easy-to-use formulae to calculate demographic measures that seem, a priori, to be complicated. All the indices listed in Table 1 can be calculated with matrix expressions, which are very easy to implement in matrix oriented softwares. In the Supplementary Material, we provide the MATLAB code that calculates the results listed in Table 1 from a given transition matrix $\mathbf{U}$ and a set of transient states $\mathcal{B}$. We also provide the MATLAB code for calculating the results of the Southern Fulmar example presented in Section 7.

In ecology and in human demography, discrete time models are widely used for their simplicity in their formulation and for their fit with data observed at specific times. However, the underlying processes of these dynamics are certainly continuous in time. Hence, extending the occupancy calculations to continuous time Markov chains would significantly enlarge the range of applications.

Acknowledgements

Appendix A

To derive our formulae, we need to define formally the stochastic process described by the transition matrices $\mathbf{P}$, $\mathbf{P}_\mathbb{K}$, $\mathbf{P}_\mathbb{C}$, and $\mathbf{P}_\mathbb{S}$. The Markov chain associated with the transition matrix $\mathbf{P}$ describes a stochastic process $\{\mathbf{X}_t\}_{t\geq 0}$ taking values in the i-state space $\mathcal{S}$ and satisfying the Markov property

$$\mathbb{P}\left(\mathbf{X}_{t+1} = i | \mathbf{X}_t = j, X_{t-1} = j_{t-1}, \ldots, \mathbf{X}_0 = j_0\right) = p_{ji} \quad (60)$$

for any $i, j, j_0, \ldots, j_{t-1} \in \mathcal{S}$. Likewise, the killed Markov chain, the conditional Markov chain, and the sub-Markov chain describe the stochastic processes $\{\mathbf{X}_t^\mathbb{K}\}_{t\geq 0}$ $\{\mathbf{X}_t^\mathbb{C}\}_{t\geq 0}$ $\{\mathbf{X}_t^\mathcal{B}\}_{t\geq 0}$, which satisfy the Markov property, respectively.

A.1. Derivation of the matrix $\mathbf{U}_\mathbb{S}$ and the vector $\mathbf{m}_\mathbb{S}$

Let i, j be two target states. The entry $u_{j-\alpha, i-\alpha}^\mathbb{S}$ is the probability that an individual initially in target state j to reach the state i without passing by any other states in $\mathcal{B}$. Define the stopping time $T = \min\{t \geq 1 | \mathbf{X}_t \in \mathcal{B}\}$. The time T is the random time – possibly infinite – at which the individual will enter the set $\mathcal{B}$. In particular, $\mathbf{X}_T$ is the state through which the individual enters for its first time in $\mathcal{B}$. Then we can rewrite $u_{j-\alpha, i-\alpha}^\mathbb{S}$ as

$$u_{j-\alpha, i-\alpha}^\mathbb{S} = \mathbb{P}_i\left(\mathbf{X}_T = j\right). \quad (61)$$

By definition of the absorbing probabilities (Eq. (11)), for a non target state $k \in \mathcal{B}^c$, we have

$$a_{j-\alpha, k} = \mathbb{P}_k\left(\mathbf{X}_T = j\right). \quad (62)$$

Using the Chapman–Kolmogorov equation (see e.g., Meyn and Tweedie (2009)), we obtain

$$\mathbb{P}_i\left(\mathbf{X}_T = j\right) = \mathbb{P}_i\left(\mathbf{X}_1 = j\right) \\ + \sum_{k \in \mathcal{B}^c} \mathbb{P}_i\left(\mathbf{X}_T = j | \mathbf{X}_1 = k\right)\mathbb{P}_i\left(\mathbf{X}_1 = k\right) \\ = \mathbb{P}_i\left(\mathbf{X}_1 = j\right) + \sum_{k \in \mathcal{B}^c} \mathbb{P}_k\left(\mathbf{X}_T = j\right)\mathbb{P}_i\left(\mathbf{X}_1 = k\right)$$

$$= \mathbb{P}_i\left(\mathbf{X}_1 = j\right) + \sum_{k \in \mathcal{B}^c} a_{\alpha-j, k}\mathbb{P}_i\left(\mathbf{X}_1 = k\right)$$

$$= u_{ji} + \sum_{k \in \mathcal{B}^c} a_{\alpha-j, k}u_{ki}. \quad (63)$$

In matrix notation, equation (63) is equivalent to Eq. (18), i.e.

$$\mathbf{U}_\mathbb{S} = \mathbf{AL} + \mathbf{Q} \quad (64)$$

where the matrices $\mathbf{L}$ and $\mathbf{Q}$ are extracted from the matrix $\mathbf{U}$, as in Eq. (7).

A.2. Proof that $\mathbf{X}_\mathbb{C}$ is a Markov chain

Iosifescu (1980) (section 3.2.9) proves that an absorbing Markov chain, with respect to the conditional probability that it is absorbed by a specific state, is still an absorbing Markov chain. Here, we generalise this statement to the condition that the chain is absorbed in a specific set of states.

Let us define the event $A = \{\mathbf{X}^\mathbb{K}$ is absorbed in the target set $\mathcal{B}\}$, i.e. the killed chain is absorbed in the target set. We consider the stochastic process $\mathbf{X}^\mathbb{C}$, living on the space $\mathcal{T}$, defined by

$$\mathbb{P}\left(\mathbf{X}_t^\mathbb{C} \in B\right) = \mathbb{P}\left(\mathbf{X}_t^\mathbb{K} \in B | A\right) \quad (65)$$

for any measurable set $B \subset \mathcal{S}$. To ease the notation, we write $\mathbb{P}\left(\mathbf{X}_t^\mathbb{K} \in B | A\right) = \mathbb{P}_A\left(\mathbf{X}_t^\mathbb{K} \in B\right)$.

By definition, the process $\mathbf{X}^\mathbb{C}$ corresponds to the killed Markov chain, where trajectories encountering death before target states are set aside. We first prove that $\mathbf{X}^\mathbb{C}$ is a Markov chain and then we show that its transition probabilities are describe by the matrix $\mathbf{P}_\mathbb{C}$ defined in Section 3.2. As a consequence, this proves that the conditional Markov chain is indeed a Markov chain and that it corresponds to the killed Markov chain, where trajectories encountering death before target states are set aside.

To prove that $\mathbf{X}^\mathbb{C}$ is a Markov chain, we only need to show that it satisfies the Markov property, i.e.

$$\mathbb{P}\left(\mathbf{X}_{t+1}^\mathbb{C} = i_{t+1} | \mathbf{X}_t^\mathbb{C} = i_t, \ldots, \mathbf{X}_0^\mathbb{C} = i_0\right) \\ = \mathbb{P}\left(\mathbf{X}_{t+1}^\mathbb{C} = i_{t+1} | \mathbf{X}_t^\mathbb{C} = i_t\right), \quad (66)$$

for $(i_0, \ldots, i_{t+1}) \in \mathcal{T}^{t+2}$.

Fix $(i_0, \ldots, i_{t+1}) \in \mathcal{T}^{t+2}$, and define the event $B_s = \{\mathbf{X}_s^\mathbb{K} = i_s, \ldots, \mathbf{X}_t^\mathbb{K} = i_0\}$, for $0 \leq s \leq t$.

From the definitions of conditional probabilities and the process $\mathbf{X}^\mathbb{C}$, we have

$$\mathbb{P}\left(\mathbf{X}_{t+1}^\mathbb{C} = i_{t+1} | \mathbf{X}_t^\mathbb{C} = i_t, \ldots, \mathbf{X}_0^\mathbb{C} = i_0\right) \\ = \frac{\mathbb{P}\left(\{\mathbf{X}_{t+1}^\mathbb{K} = i_{t+1}\} \cap B_0 \cap A\right)}{\mathbb{P}\left(A \cap B_0\right)} \quad (67)$$

If $(i_0, \ldots, i_t) \notin \mathcal{B}^{t+1}$, then

$$\mathbb{P}\left(A \cap B_s\right) = \mathbb{P}\left(A | B_s\right)\mathbb{P}\left(B_s\right) \\ = \mathbb{P}\left(A | \mathbf{X}_t^\mathbb{K} = i_t\right)\mathbb{P}\left(B_s\right) \\ p_{i_t}^a \mathbb{P}\left(B_s\right),$$

for any $0 \leq s \leq t$. The second equality is a consequence of the Markov property of the killed Markov chain, and the third equality follows from the definition of the absorbing probability vector $\mathbf{p}_a$ (see Eq. (12)).

Similarly,

$$\mathbb{P}\left(\{\mathbf{X}_{t+1}^\mathbb{K} = i_{t+1}\} \cap B_s \cap A\right) \\ = \mathbb{P}\left(\{\mathbf{X}_{t+1}^\mathbb{K} = i_{t+1}\} \cap A | B_s\right)\mathbb{P}\left(B_s\right) \quad (68)$$

$$= \mathbb{P}\left(\{\mathbf{X}_{t+1}^\mathbb{K} = i_{t+1}\} \cap A | \mathbf{X}_t^\mathbb{K} = i_t\right)\mathbb{P}\left(B_s\right) \quad (69)$$

$$= \mathbb{P}\left(B_s\right)\mathbb{I}_\mathcal{B}(i_{t+1})\mathbb{P}_{i_t}\left(\mathbf{X}_{t+1}^\mathbb{K} = i_{t+1}\right) \quad (70)$$

$$+ \; \mathbb{P}\,(B_s)\,(1 - \mathbb{I}_{\mathcal{B}}(i_{t+1}))\mathbb{P}\left(A|\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\right)$$
$$\times \, \mathbb{P}_{i_t}\left(\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\right) \tag{71}$$

$$= \mathbb{P}\,(B_s)\left[\mathbb{I}_{\mathcal{B}}(i_{t+1})\mathbb{P}_{i_t}\left(\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\right)\right.$$
$$\left. + \; (1 - \mathbb{I}_{\mathcal{B}}(i_{t+1}))p^a_{i_{t+1}}\mathbb{P}_{i_t}\left(\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\right)\right] \tag{72}$$

where $\mathbb{I}_{\mathcal{B}}(k)$ equals 1 if $k \in \mathcal{B}$ and 0 otherwise.

If $i_t \in \mathcal{B}$, then $B_s \subset A$, for $0 \le s \le t$, and

$$\frac{\mathbb{P}\left(\{\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\} \cap B_s \cap A\right)}{\mathbb{P}\,(A \cap B_s)}$$
$$= \frac{\mathbb{P}\left(\{\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\} \cap B_s\right)}{\mathbb{P}\,(B_s)} \tag{73}$$

$$= \mathbb{P}\left(\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}|B_s\right) \tag{74}$$

$$= \mathbb{P}\left(\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}|\mathbf{X}^{\mathbb{K}}_t = i_t\right). \tag{75}$$

Eqs. (72) and (75) imply that the ratio on the right hand side of Eq. (67) does not depend on s, for $0 \le s \le t$. In particular,

$$\mathbb{P}\left(\mathbf{X}^{\mathbb{C}}_{t+1} = i_{t+1}|\mathbf{X}^{\mathbb{C}}_t = i_t, \dots, \mathbf{X}^{\mathbb{C}}_0 = i_0\right)$$
$$= \frac{\mathbb{P}\left(\{\mathbf{X}^{\mathbb{K}}_{t+1} = i_{t+1}\} \cap B_t \cap A\right)}{\mathbb{P}\,(A \cap B_t)} \tag{76}$$

$$= \mathbb{P}\left(\mathbf{X}^{\mathbb{C}}_{t+1} = i_{t+1}|\mathbf{X}^{\mathbb{C}}_t = i_t\right). \tag{77}$$

This prove that $\mathbf{X}^{\mathbb{C}}$ satisfies the Markov property.

The transition probabilities of the Markov chain $\mathbf{X}^{\mathbb{C}}$ follow from the equations above. For $j \in \mathcal{T}$ and $i \notin \mathcal{B}$,

$$\mathbb{P}\left(\mathbf{X}^{\mathbb{C}}_{t+1} = j|\mathbf{X}^{\mathbb{C}}_t = i\right) = \frac{p^a_j p^{\mathbb{K}}_{ji}}{p^a_i}, \tag{78}$$

with the convention that $p^a_k = 1$ for $k \in \mathcal{B}$. And we have for $i, j \in \mathcal{B}$,

$$\mathbb{P}\left(\mathbf{X}^{\mathbb{C}}_{t+1} = j|\mathbf{X}^{\mathbb{C}}_t = i\right) = 1. \tag{79}$$

It follows form Eqs. (78) and (79) that the transition probabilities of the Markov chain $\mathbf{X}^{\mathbb{C}}$ are given by the matrix $\mathbf{P}_{\mathbb{C}}$ defined in Section 3.2, i.e.

$$\mathbf{P}_{\mathbb{C}} = \left(\begin{array}{c|c} \mathbf{U}_{\mathbb{C}} & \mathbf{0} \\ \hline \mathbf{M}_{\mathbb{C}} & \mathbf{I}_{\beta} \end{array}\right) \tag{80}$$

where

$$\mathbf{U}_{\mathbb{C}} = \mathbf{D}_a \mathbf{U}_{\mathbb{K}} \mathbf{D}_a^{-1} \quad \text{and} \quad \mathbf{M}_{\mathbb{C}} = \mathbf{K}\mathbf{D}_a^{-1}, \tag{81}$$

where $\mathbf{D}_a = \mathrm{diag}\,(\mathbf{p}_a)$ is a diagonal matrix with, on the diagonal, the probabilities of absorption in the target states.

A.3. Moments of occupancy times

To calculate the moments of the occupancy time for individuals initially outside the target set $\mathcal{B}$, we use the strong Markov property. If the individual never enters in $\mathcal{B}$, then its occupancy time is zero. If it does enter in $\mathcal{B}$, say through the state j, then the law of its occupancy time is equal to the law of the occupancy time for an individual starting in the state j. To fix the idea, consider the state $i \in \mathcal{B}^c$. Then

$$\mathbb{E}(\tau_i^m) = \sum_{\ell=1}^{\infty} \ell^m \mathbb{P}(\tau_i = \ell)$$
$$= \sum_{\ell=1}^{\infty} \ell^m \left(\sum_{j \in \mathcal{B}} \mathbb{P}(\tau_i = \ell | \text{enter in } \mathcal{B} \text{ in } j)\right.$$

$$\left. \times \; \mathbb{P}(\text{enter in } \mathcal{B} \text{ in } j)\right)$$
$$= \sum_{\ell=1}^{\infty} \ell^m \left(\sum_{j \in \mathcal{B}} \mathbb{P}(\tau_j = \ell)a_{ji}\right)$$
$$= \sum_{j \in \mathcal{B}} a_{ji} \sum_{\ell=1}^{\infty} \ell^m \mathbb{P}(\tau_j = \ell)$$
$$= \sum_{j \in \mathcal{B}} a_{ji}\mathbb{E}(\tau_j^m), \tag{82}$$

where a_{ji} is the probability that the killed Markov chain, starting in state i, is absorbed by the state j, as in Section 3.1.1. In matrix notation, Eq. (82) is equivalent to

$$\boldsymbol{\tau}^k_{\text{out}} = \mathbf{A}^{\mathsf{T}} \boldsymbol{\tau}^k_{\text{in}}. \tag{83}$$

A.4. Covariance between the occupancy times in two disjoint sets

Here, we calculate the covariance between the occupancy time in two disjoint subsets $\mathcal{B}_1$ and $\mathcal{B}_2$, of the transient set $\mathcal{T}$. As stated in the main text, the covariance between $\boldsymbol{\tau}_{\mathcal{B}_1}$ and $\boldsymbol{\tau}_{\mathcal{B}_2}$ is

$$\mathrm{Cov}\left(\boldsymbol{\tau}_{\mathcal{B}_1}, \boldsymbol{\tau}_{\mathcal{B}_2}\right) = \mathbb{E}\left[\left(\boldsymbol{\tau}_{\mathcal{B}_1} - \boldsymbol{\tau}^1_{\mathcal{B}_1}\right)\left(\boldsymbol{\tau}_{\mathcal{B}_2} - \boldsymbol{\tau}^1_{\mathcal{B}_2}\right)\right]. \tag{84}$$

We rewrite the covariance between $\boldsymbol{\tau}_{\mathcal{B}_1}$ and $\boldsymbol{\tau}_{\mathcal{B}_2}$ in terms of their variances and the variance of their sum,

$$\mathrm{Cov}\left(\boldsymbol{\tau}_{\mathcal{B}_1}, \boldsymbol{\tau}_{\mathcal{B}_2}\right) = \frac{1}{2}\left[\mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1} + \boldsymbol{\tau}_{\mathcal{B}_2}\right)\right.$$
$$\left. - \; \mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1}\right) - \mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_2}\right)\right]. \tag{85}$$

Since the sets $\mathcal{B}_1$ and $\mathcal{B}_2$ are disjoint, the occupancy time in the union $\mathcal{B}_1 \cup \mathcal{B}_2$ is the sum of the occupancy times in each of the subsets. Thus,

$$\mathrm{Cov}\left(\boldsymbol{\tau}_{\mathcal{B}_1}, \boldsymbol{\tau}_{\mathcal{B}_2}\right) = \frac{1}{2}\left[\mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1 \cup \mathcal{B}_2}\right) - \mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1}\right)\right.$$
$$\left. - \; \mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_2}\right)\right]. \tag{86}$$

The variances $\mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1 \cup \mathcal{B}_2}\right)$, $\mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_1}\right)$, and $\mathrm{Var}\left(\boldsymbol{\tau}_{\mathcal{B}_2}\right)$ are calculated with the formulae (24) and (26) applied to the sets $\mathcal{B}_1 \cup \mathcal{B}_2$, $\mathcal{B}_1$, and $\mathcal{B}_2$, respectively.

A.5. One-step transition probabilities from $\mathcal{B}$ given return

Let w^{in}_{ij} be the conditional probability that an individual in target state $\alpha + j$ moves to the target state $\alpha + i$, in one time-step, given that it eventually returns to the target set. Then,

$$w^{\text{in}}_{ji} := \mathbb{P}_{\alpha+i}\,(X_1 = \alpha + j|T < \infty)$$
$$= \frac{\mathbb{P}_{\alpha+i}\,(X_1 = \alpha + j, T < \infty)}{\mathbb{P}_{\alpha+i}\,(T < \infty)}$$
$$= \frac{\mathbb{P}_{\alpha+i}\,(X_1 = \alpha + j)}{\mathbb{P}_{\alpha+i}\,(T < \infty)}$$
$$= \frac{u_{\alpha+j,\alpha+i}}{\sum_{\ell \in \mathcal{B}} u^{\mathcal{B}}_{\ell i}}$$
$$= \frac{u_{\alpha+j,\alpha+i}}{p^r_i}, \tag{87}$$

where $\mathbf{p}_r$ describes the return probabilities, as defined in Eq. (47). Thus,

$$\mathbf{W}_{\text{in}} = \mathbf{Q}\mathbf{D}_r^{-1} \quad \text{of size } \beta \times \beta, \tag{88}$$

where $\mathbf{D}_r = \mathrm{diag}\,(\mathbf{p}_r)$ and the matrix $\mathbf{Q}$ is extracted from the matrix $\mathbf{U}$, as in Eq. (7). Let w_{ij}^{out} be the conditional probability that an individual in target state $\alpha + j$ moves to the non-target state i, in one time-step, given that it eventually returns to the target set. Then

$$
\begin{aligned}
w_{ji}^{\mathrm{out}} &:= \mathbb{P}_{\alpha+i}\,(X_1 = j | T < \infty) \\
&= \frac{\mathbb{P}_{\alpha+i}\,(T < \infty | X_1 = j)\,\mathbb{P}_{\alpha+i}(X_1 = j)}{\mathbb{P}_{\alpha+i}\,(T < \infty)} \\
&= \frac{\mathbb{P}_{\alpha+i}(X_1 = j)\mathbb{P}_{\alpha+j}\,(T < \infty)}{\mathbb{P}_{\alpha+i}\,(T < \infty)} \\
&= \frac{u_{j,\alpha+i}\sum_{\ell \in \mathcal{B}} a_{\ell j}}{\sum_{\ell \in \mathcal{B}} u_{\ell i}^{\mathcal{B}}} \\
&= \frac{u_{j,\alpha+i}p_j^a}{p_i^r},
\end{aligned}
\tag{89}
$$

where the vector p_a describes the probabilities of absorption in the target states, as defined in Eq. (11). Thus,

$$
\mathbf{W}_{\mathrm{out}} = \mathbf{D}_a \mathbf{L} \mathbf{D}_r^{-1} \quad \text{of size } \alpha \times \beta,
\tag{90}
$$

where $\mathbf{D}_r = \mathrm{diag}\,(\mathbf{p}_r)$, $\mathbf{D}_a = \mathrm{diag}\,(\mathbf{p}_a)$, and the matrix $\mathbf{L}$ is extracted form the matrix $\mathbf{U}$, as in Eq. (7).

Now, we derive the moments of μ, conditional on the individual returning to the target set. Let $\alpha + i$ be a target state. Then

$$
\mathbb{E}[\mu_i^k | T < \infty] = \sum_{n=1}^{\infty} n^k \mathbb{P}(\mu_i = n | T < \infty)
\tag{91}
$$

$$
\begin{aligned}
&= \mathbb{P}(\mu_i = 1 | T < \infty) \\
&\quad + \sum_{n=2}^{\infty} n^k \sum_{j \in \mathcal{B}^c} \mathbb{P}(t_j^{\mathcal{B}} = n - 1)\mathbb{P}_i(X_1 = j | T < \infty)
\end{aligned}
\tag{92}
$$

$$
\begin{aligned}
&= \mathbb{P}(\mu_i = 1 | T < \infty) \\
&\quad + \sum_{j \in \mathcal{B}^c} \mathbb{P}_i(X_1 = j | T < \infty) \\
&\quad \times \sum_{n=1}^{\infty} (n+1)^k \mathbb{P}(t_j^{\mathcal{B}} = n)
\end{aligned}
\tag{93}
$$

$$
\begin{aligned}
&= \mathbb{P}(\mu_i = 1 | T < \infty) \\
&\quad + \sum_{j \in \mathcal{B}^c} \mathbb{P}_i(X_1 = j | T < \infty) \\
&\quad \times \sum_{n=1}^{\infty} \sum_{r=0}^{k} \binom{k}{r} n^r \mathbb{P}(t_j^{\mathcal{B}} = n)
\end{aligned}
\tag{94}
$$

$$
\begin{aligned}
&= \mathbb{P}(\mu_i = 1 | T < \infty) \\
&\quad + \sum_{j \in \mathcal{B}^c} \mathbb{P}_i(X_1 = j | T < \infty) \sum_{r=0}^{k} \binom{k}{r} t_{\mathcal{B}_j}^r
\end{aligned}
\tag{95}
$$

$$
= 1 + \sum_{j \in \mathcal{B}^c} w_{ji}^{\mathrm{out}} \sum_{r=1}^{k} \binom{k}{r} t_{\mathcal{B}_j}^r.
\tag{96}
$$

Hence, in matrix notation,

$$
\mathbb{E}[\mu^k | T < \infty] = \mathbf{1}_\beta + \sum_{r=1}^{k} \binom{k}{r} \mathbf{W}_{\mathrm{out}}^\mathsf{T} \mathbf{t}_{\mathcal{B}}^r.
\tag{97}
$$

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.tpb.2017.12.007.

References

Caswell, H., 2001. Matrix Population Models: Construction, Analysis, and Interpretation, second ed. Sinauer Associates, Sunderland, MA, ISBN: 9780878930968.

Caswell, H., 2006. Applications of markov chains in demography. In: MAM2006: Markov Anniversary Meeting. Boson Books, Raleigh, North Carolina, pp. 319–334.

Caswell, H., 2009. Stage, age and individual stochasticity in demography. Oikos 118 (12), 1763–1782.

Caswell, H., 2011a. Beyond R_0: Demographic models for variability of lifetime reproductive output. PloS ONE 6 (6), e20809.

Caswell, H., 2011b. Perturbation analysis of continuous-time absorbing Markov chains. Numer. Linear Algebra Appl. 18 (6), 901–917.

Caswell, H., 2013. Sensitivity analysis of discrete Markov chains via matrix calculus. Linear Algebra Appl. 438 (4), 1727–1745. http://dx.doi.org/10.1016/j.laa.2011.07.046.

Caswell, H., Kluge, F.A., 2015. Demography and the statistics of lifetime economic transfers under individual stochasticity. Demographic Res. 32, 563–588. http://dx.doi.org/10.4054/DemRes.2015.32.19.

Clark, D.E., Ostrander, K.R., Cushing, B.M., 2016. A multistate model predicting mortality, length of stay, and readmission for surgical patients. Health Serv. Res. 51 (3), 1074–1094.

Cochran, M.E., Ellner, S.P., 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. Ecol. Monograph (ISSN: 00129615) 62 (3), 345–364.

Ellner, S.P., Childs, D.Z., Rees, M., 2016. Data-Driven Modelling of Structured Populations: A Practical Guide to the Integral Projection Model. Springer, ISBN: 9783319288918.

Feichtinger, G., 1971. Stochastische modelle demographischer prozesse. Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, Berlin, ISBN: 9783540054238.

Horvitz, C.C., Tuljapurkar, S., 2008. Stage dynamics, period survival, and mortality plateaus. Amer. Naturalist 172 (2), 203–215.

Iosifescu, M., 1980. Finite Markov Processes and Their Applications. Wiley, New York, New York.

Jenouvrier, S., Pron, C., Weimerskirch, H., 2015. Extreme climate events and individual heterogeneity shape life-history traits and population dynamics. Ecol. Monograph 85 (4), 605–624. http://dx.doi.org/10.1890/14-1834.1.

Kemeny, J.G., Snell, J.L., 1961. Finite continuous time Markov chains. Theory Probab. Appl. 6 (1), 101–105.

Lebreton, J.-D., 1996. Demographic models for subdivided populations: the renewal equation approach. Theoret. Popul. Biol. 49 (3), 291–313.

Meyn, S.P., Tweedie, R.L., 2009. Markov Chains and Stochastic Stability. Cambridge Univ. Press.

Pascarella, J.B., Horvitz, C.C., 1998. Hurricane disturbance and the population dynamics of a tropical understory shrub: Megamatrix elasticity analysis. Ecology 79 (2), 547–563.

Rogers, A., 1975. Introduction to Multiregional Mathematical Demography. Wiley, New York, New York.

Roth, G., Caswell, H., 2016. Hyperstate matrix models: extending demographic state spaces to higher dimensions. Methods Ecol. Evol. (ISSN: 2041-210X) 7 (12), 1438–1450.

Salguero-Gómez, R., Jones, O.R., Archer, C.R., Bein, C., Buhr, H., Farack, C., Gottschalk, F.and Hartmann, A., Henning, A., Hoppe, G.and Romer, G., Ruoff, T., Sommer, V., Wille, J., Voigt, J., Zeh, S., Vieregg, D., Buckley, Yvonne M, Che-Castaldo, J., Hodgson, D., Scheuerlein, A., Caswell, H.and Vaupel, J.W., 2016. COMADRE: a global data base of animal demography. J. Anim. Ecol. 85 (2), 371–384.

Salguero-Gómez, R., Jones, O.R., Archer, C.R., Buckley, Y.M., Che-Castaldo, J., Caswell, H., Hodgson, D., Scheuerlein, A., Conde, D.A., Brinks, E., et al., 2015. The compadre plant matrix database: an open online repository for plant demography. J. Ecol. 103 (1), 202–218.

Sericola, B., 2000. Occupation times in Markov processes. Comm. Statist. Stochastic Models (ISSN: 0882-0287) 16 (5), 479–510. http://dx.doi.org/10.1080/15326340008807601.

Siegel, J.S., 2012. The life table. In: The Demography and Epidemiology of Human Health and Aging. Springer, New York, pp. 135–216.

Steiner, U.K., Tuljapurkar, S., 2012. Neutral theory for life histories and individual variability in fitness components. Proc. Natl. Acad. Sci. 109 (12), 4684–4689.

Tuljapurkar, S., Horvitz, C.C., 2006. From stage to age in variable environments: life expectancy and survivorship. Ecology 87 (6), 1497–1509.

Tuljapurkar, S., Horvitz, C.C., Pascarella, J.B., 2003. The many growth rates and elasticities of populations in random environments. Amer. Naturalist 162 (4), 489–502.

Van Raalte, A.A., Caswell, H., 2013. Perturbation analysis of indices of lifespan variability. Demography 50 (5), 1615–1640.