Stochastic stable population growth in integral projection models: Theory and application

Article *in* Journal of Mathematical Biology · March 2007

DOI: 10.1007/s00285-006-0044-8 · Source: PubMed

2 authors, including:

Mark Rees
The University of Sheffield
150 PUBLICATIONS **7,358** CITATIONS

J. Math. Biol. (2007) 54:227–256
DOI 10.1007/s00285-006-0044-8

Mathematical Biology

Stochastic stable population growth in integral projection models: theory and application

Stephen P. Ellner · Mark Rees

Received: 8 March 2006 / Revised: 11 October 2006 /
Published online: 23 November 2006

Abstract Stochastic matrix projection models are widely used to model age- or stage-structured populations with vital rates that fluctuate randomly over time. Practical applications of these models rest on qualitative properties such as the existence of a long term population growth rate, asymptotic log-normality of total population size, and weak ergodicity of population structure. We show here that these properties are shared by a general stochastic integral projection model, by using results in (Eveson in D. Phil. Thesis, University of Sussex, 1991, Eveson in Proc. Lond. Math. Soc. **70**, 411–440, 1993) to extend the approach in (Lange and Holmes in J. Appl. Prob. **18**, 325–344, 1981). Integral projection models allow individuals to be cross-classified by multiple attributes, either discrete or continuous, and allow the classification to change during the life cycle. These features are present in plant populations with size and age as important predictors of individual fate, populations with a persistent bank of dormant seeds or eggs, and animal species with complex life cycles. We also present a case-study based on a 6-year field study of the Illyrian thistle, *Onopordum illyricum*, to demonstrate how easily a stochastic integral model can be parameterized from field data and then applied using familiar matrix software and methods. Thistle demography is affected by multiple traits

Research supported by NSF grant OCE 0326705 in the NSF/NIH Ecology of Infectious Diseases program and the Cornell College of Arts and Sciences (SPE), and NERC grant NER/A/S/2002/00940 (MR).

S. P. Ellner (✉)
Cornell University, Department of Ecology and Evolutionary Biology, Ithaca, NY, USA
e-mail: spe2@cornell.edu

M. Rees
University of Sheffield, Deparment of Animal and Plant Sciences, Sheffield, UK
e-mail: m.rees@sheffield.ac.uk

(size, age and a latent "quality" variable), which would be difficult to accomodate in a classical matrix model. We use the model to explore the evolution of size- and age-dependent flowering using an evolutionarily stable strategy (ESS) approach. We find close agreement between the observed flowering behavior and the predicted ESS from the stochastic model, whereas the ESS predicted from a deterministic version of the model is very different from observed flowering behavior. These results strongly suggest that the flowering strategy in *O. illyricum* is an adaptation to random between-year variation in vital rates.

Keywords Stochastic demography · Integral projection models · Structured populations · Hilbert's projective metrix · *Onopordum illyricum*

Mathematics Subject Classification (2000) 92D25 · 60H25 · 37H15 · 47B65

1 Introduction

Matrix projection models are probably the most widely used model for structured biological populations [3,38,14], especially in data-limited situations where it is important to have a simple parameter-sparse model. The deterministic model is

$$\mathbf{n}(t+1) = A\mathbf{n}(t), \tag{1}$$

which generalizes the classical Leslie matrix model for an age-structured population. The population vector $\mathbf{n}(t)$ lists the numbers of individuals in a finite set of categories (stages of the life cycle, size classes, etc.) and the $(i,j)^{th}$ entry in the projection matrix A gives the average per-capita contribution from individuals in category j at time t to category i at time $t+1$, either by survival or reproduction. Relatively simple matrix models have figured prominently in management planning for endangered species (e.g. sea turtles [8,27,28], harvested salmon stocks [30], plants [36]) and are coming into use for invasive species management (e.g., [35,46,47]).

Population forecasting and risk assessment require a recognition that survival and fecundity rates in most natural populations vary greatly over time. Admitting this variability into (1) leads to the stochastic matrix projection model

$$\mathbf{n}(t+1) = A(t)\mathbf{n}(t) \tag{2}$$

where $\{A(t), t = 0, 1, 2, \ldots\}$ is a stochastic sequence of non-negative matrices. These models are widely used to estimate extinction risk in threatened and endangered species (e.g. [3,18,38,36]) and to study the evolution of life histories in temporally varying environments (e.g. [1,49]).

Applications of model (2) depend on "stable population growth": properties associated with the existence of a long-term population growth rate and population structure. For the deterministic model (1) with a power-positive matrix A, the long-term population trend (by the Perron–Frobenius Theorem) is

$$\lim_{t\to\infty} t^{-1} \log \|\mathbf{n}(t)\|_1 = \lambda \tag{3}$$

and

$$\lim_{t\to\infty} \mathbf{n}(t)/\|\mathbf{n}(t)\|_1 = w/\|w\|_1 . \tag{4}$$

Here λ is the dominant eigenvalue of the projection matrix A, w is the corresponding right eigenvector [3,38], and $\|w\|_1$ denotes the L_1 norm

$$\|w\|_1 = \|(w_1, w_2, \ldots, w_n)\|_1 = \sum_{i=1}^{n} |w_i|.$$

Under suitable assumptions the stochastic model (2) still has (3) with λ almost surely constant [48]. In addition, a Central Limit Theorem for fluctuations about the long-term growth rate implies that the total population $\|\mathbf{n}(t)\|_1$ has an asymptotically lognormal distribution. Instead of (4) there is asymptotic convergence to a joint stationary distribution for $A(t)$ and the population structure $\mathbf{u}(t) = \mathbf{n}(t)/\|\mathbf{n}(t)\|_1$.

A matrix model is based on classifying individuals into a finite set of discrete categories. This is natural for age-based population modeling, but in many species the variables that most strongly affect individual demographic performance vary continuously. In fact most of the empirical applications of matrix population models in the standard monograph [3] are based on some continuous measure of individual size. An integral projection model (IPM) allows individuals to be classified by continuous variables, or by a mix of continuous and discrete (e.g., size and age). In the simplest case of a single continuous individual-level variable $x \in [L, U]$, the deterministic model is

$$n(y, t+1) = \int_L^U K(y, x) n(x, t) \mathrm{d}x \tag{5}$$

where the kernel $K(y, x)$ gives the contribution of type-x individuals at time t to type-y individuals at time $t+1$. Much of the basic theory for the deterministic matrix model (1) has recently been extended to a general class of deterministic IPMs [15]. A stochastic IPM has randomly time-varying kernels; in empirical applications these are specified typically as $K(y, x, \theta(t))$ where $\theta(t)$ is a vector-valued stochastic process representing variation over time in the environmental conditions affecting the population.

When individual growth, survival, and birth rates are smooth functions of continuously varying traits, an IPM is a direct translation of statistical models for demographic rates [15]. As a result, when multiple individual-level traits are important an IPM can be parameterized much more parsimoniously than a matrix model [15]. Many situations that previously required individual-based

simulations can be handled with an IPM using standard matrix software and established analytic and numerical methods for structured population models. Recent applications include [4,5,42,44,43]. An IPM can also be used to model the dynamics of physiologically structured populations in discrete time. This may be a useful alternative to the typical PDE models for physiologically structured populations, or more general continuous-time integral models [9,10], when a model is being parameterized from individual-level observations (A. de Roos, personal communication). Because data necessarily come at a discrete set of sampling times, a discrete-time model can be fitted without having to solve the inverse problem of inferring unmeasured continuous-time rates from discrete-time observations.

In this paper, our first purpose is to show that the "stable population growth" properties of stochastic matrix models all continue to hold in a general density-independent stochastic integral projection model. The basis for our analysis is Lange and Holmes's paper [34] on stochastic matrix models. In contrast to previous treatments [6,7,19] Lange and Holmes avoided matrix-specific calculations by using compactness arguments and the fact that multiplication by a positive matrix is a contraction map under Hilbert's projective metric (defined below). This resembles Birkhoff's approach to positive integral operators [2]. Using recent work of Eveson on the projective metric for positive operators [16,17], we show that the contractive properties used in [34] also hold for stochastic IPMs. After that the proofs in [34] apply almost verbatim.

Our second purpose is to demonstrate how the mathematical results can be used to help understand natural systems. We extend a size-, age- and quality-dependent IPM for the thistle *Onopordum illyricum* [15] to include stochastic variation in demography. This case-study shows how simple it is to completely parametrize a stochastic IPM from real field data, using standard statistical software for fitting regression models. We then use the model to explore the evolution of size- and age-dependent flowering in *Onopordum* using the evolutionarily stable strategy (ESS) approach. This application also demonstrates how the mechanics of working with a fitted IPM can be accomplished using standard matrix methods and software.

The justification for this paper is the utility of the results, rather than any beauty or depth to the mathematics. Subsequent papers (in preparation) will focus on numerical methods, applications to population forecasting and risk assessment (including sensitivity/elasticity analysis and small-fluctuations approximations), and additional aspects of model construction such as methods for parameterizing models that include between-year correlations in environmental variables affecting demographic rates.

2 General model and assumptions

The basic premise of an integral projection model is that the state of an individual during its lifetime follows a Markov chain on some space $\mathbf{X}$, with possibly time-varying transition probabilities. For example,

- In many plant populations age and size are both important predictors for survival, growth and fecundity. If $[L_a, U_a]$ is the range of possible sizes at age a and M the maximum possible age, then $\mathbf{X}$ would be the set of intervals $\Omega_a = [L_a, U_a], a = 0, 1, 2, \ldots M$, each sitting in its own copy of the real line.
- If a species has a persistent bank of dormant seeds, we might add to $\mathbf{X}$ a discrete point representing the state of being a dormant seed, a series of points representing dormant seeds of different ages, a continuous interval representing seeds buried at different depths, or a set of intervals for a depth $\times$ age classification of buried seeds.
- If individual performance is affected by a second continuous trait—perhaps a measure of individual condition or local habitat quality – the intervals $[L_a, U_a]$ are replaced by rectangles $[L_a, U_a] \times [\, m_a, M_a]$ representing possible joint values of size and the second trait.

Reproduction is also Markovian: at each time t the number of offspring produced by an individual and the initial states of offspring are chosen from probability distributions that depend on the parent's state at time t but not on the parent's prior history. However, an IPM is a population-level mean field model: it ignores demographic stochasticity and tracks the expected number of individuals and their state distribution.

A nontechnical statement of our general model is as follows. In the basic model (5) the individual-level state space is an interval (e.g., a range of possible values for body size or condition). In our general model the set $\mathbf{X}$ of possible individual states can include any finite set of points $\{x_j\}_{j=1}^n$, closed intervals $[L_j, U_j]$, rectangles $[L_j, U_j] \times [\, m_j, M_j]$, cubes, etc. The population dynamics are governed by a set of continuous kernels $K_{i,j}(y, x)$ representing all possible transitions (survival, growth, fecundity) within and among the various components of the state space. The model is made stochastic by introducing an *environment process* $\theta(t)$ that affects demographic rates. The kernels are then $K_{i,j}(y, x, \theta(t))$, which can be (loosely) interpreted as giving the number of state-y individuals in the i^{th} component of $\mathbf{X}$ at time $t+1$, per state-x individual in the j^{th} component at time t, when the environment state in year t is $\theta(t)$. We assume that the environment sequence $\theta(t)$ is stationary (the probability distribution of different environment states is the same each year) and ergodic (the distant future becomes completely independent of the present). The kernels must be *power-positive*, meaning that there is a time interval m such that the m-step kernels representing the population changes over that time interval are positive on the entire state space, for any possible values of the environment sequence. As in matrix models, power-positivity is only possible if post-reproductive stages or states are removed (those without any chance of reproducing now or in the future)—the population is modeled as if such individuals are already dead. However, as in [34] post-reproductives can be included under a suitable power-positivity assumption on the kernels for potentially reproductive states. Similarly, our results also hold for "power-smooth" models where the m-step population dynamics are described by smooth kernels, even if the one-step dynamics are not. This situation often occurs if some individual-level state variables

remain constant over an individual's lifetime. These extensions are explained in Sect. 3.1, which is unavoidably somewhat technical.

In the rest of this section we give a technical and slightly more general statement of our model and assumptions. Readers without the necessary background (measure-theoretic probability and analysis) can go directly from here to Sect. 4 where we present the conclusions.

The set $\mathbf{X}$ of possible individual states is assumed to be a compact metric space; this is general enough for virtually any empirically-based structured population model [15]. The IPM is defined relative to a measure space $(\mathbf{X}, \mathcal{B}, \mu)$ where $\mathcal{B}$ is the Borel σ-field on $\mathbf{X}$ and μ is a finite measure. The state of the population at time t is given by a nonnegative function $n(x,t)$ in $L_1(\mathbf{X})$, the space of measurable functions $f(x)$ such that $\int_{\mathbf{X}} |f(x)| \mathrm{d}\mu(x) < \infty$ (here and below we use $L_1(\mathbf{X})$ as a shorthand for $L_1(\mathbf{X}, \mathcal{B}, \mu)$). The interpretation of $n(x,t)$ is that $\int_A n(x,t) \mathrm{d}\mu(x)$ is the number of individuals whose state at time t is in the set A.

The population dynamics are specified by projection kernels $K_t(y,x)$ which give the expected contribution of a state-x individual at time t to state-y individuals at time $t+1$. The general time-varying IPM is then

$$n(y,t+1) = \int_{\mathbf{X}} K_t(y,x) n(x,t) \mathrm{d}\mu(x). \tag{6}$$

The kernel K_t typically can be decomposed as $K_t = P_t + F_t$ where P_t represents survival and state transitions by survivors, and F_t represents fecundity, but these might be decomposed further (e.g., sexual versus vegetative reproduction).

Model (6) is slightly more general than our formulation of deterministic IPMs in [15] in that the kernels can be defined relative to any finite Borel measure, but everything in [15] still applies (with minor notation changes) in this setting. The model could be generalized further by using finite measures to describe the population state, as in [9,10,20]. However this would cost us the ability to give conditions for long-term stable growth that are relatively simple to state and interpret (see in particular the discussion of eigenvector existence and uniqueness for the "forward process" in Sect. 4.3 of [20]); we return to this issue in the Discussion.

Compactness of $\mathbf{X}$ is not just for convenience. IPMs on an unbounded state space are more like models for spatial spread on an infinite domain, and can have travelling wave solutions rather than the convergence to stable structure typical of matrix population models [15]. A compact state space will typically result if the model is constrained so that it cannot produce individuals too dissimilar from those actually observed in the population [15]. For example, the modeler can truncate the state space and transition kernels at limits placed several standard deviations beyond the range of observations.

The model (6) is density-independent, because the population state has no effect on the projection kernel. Density-dependent models would be expected to behave very differently, typically converging to a stationary distribution for n rather than a stable pattern of population growth. Hardin et al. [24], extending

results for unstructured population models in [13], prove that this occurs in density-dependent integral models representing spatially distributed populations. However several of their assumptions typically would not hold in models for populations structured by age, size or individual-level physiological traits. The results of [24] can probably be extended to more general density-dependent IPMs, but we leave that for another paper.

Stable population growth requires some assumptions about the projection kernels.

1. *Continuity* The kernels are of the form $K_t(y,x) = K(y,x,\theta(t))$ for $\theta(t)$ in a compact metric space Θ, where $K(y,x,\theta)$ is a continuous real-valued nonnegative function on $\mathbf{X}\times\mathbf{X}\times\Theta$. We write $K(\theta)$ as shorthand for the kernel $K(\cdot,\cdot,\theta)$ and $\mathbf{K}(\theta)$ for the corresponding operator.
2. *Stationarity* The sequence of environment states $\{\theta(t), t = \ldots,-1,0,1,2,\ldots\}$ is generated by a stationary ergodic stochastic process.
3. *Power positivity* There exists an integer $m > 0$ such that

$$K(\theta_m)\circ K(\theta_{m-1})\circ\cdots\circ K(\theta_1) \gg 0 \tag{7}$$

whenever $\theta_1,\theta_2,\cdots,\theta_m$ are all in Θ. Here "$\gg 0$" means that the function is everywhere positive, and $\circ$ indicates the composition operator that corresponds to iteration of (6), i.e.

$$(K_2\circ K_1)(y,x) = \int_X K_2(y,z)K_1(z,x)\mathrm{d}\mu(z). \tag{8}$$

It follows from (7) that

$$\int_{\mathbf{X}} K(y,x,\theta)\mathrm{d}\mu(x) > 0 \tag{9}$$

for all $y\in\mathbf{X}$ and $\theta\in\Theta$ (if this were to fail for some y_0,θ_0, then (7) would fail if $\theta_m=\theta_0$). Invoking compactness and continuity, it follows that there exist positive constants α_i,β_i such that

$$\alpha_1 \le K(\theta_m)\circ K(\theta_{m-1})\circ\cdots\circ K(\theta_1) \le \alpha_2 \tag{10}$$

whenever $\theta_1,\theta_2,\ldots,\theta_m$ are all in Θ, and

$$\beta_1 \le \int_{\mathbf{X}} K(y,x,\theta)\mathrm{d}\mu(x) \le \beta_2 \tag{11}$$

for all $y\in\mathbf{X}$ and $\theta\in\Theta$. Equation (10) corresponds to the typical assumption for stochastic matrix models that the matrices lie almost surely in an *ergodic set*. The kernel depends on the choice of measure μ. Careful choice of μ may be

important in some cases, because it may affect whether or not the kernels are power-positive. For notational convenience, define $\Psi = \Theta^m$ and for $\psi \in \Psi$ we let $\mathbf{Q}(\psi)$ and $Q(\psi)$ denote the operator and kernel, respectively, corresponding to the kernel iterate in (7).

For asymptotic lognormality of the total population size, the environment process has to "forget its past" quickly enough. Specifically, let $\mathcal{F}_j^k, j \le k$ be the σ-algebra generated by $(\theta(j), \theta(j+1), \ldots, \theta(k))$ and

$$\phi(n) = \sup_k \sup \left\{ |P(B|A) - P(B)| : A \in \mathcal{F}_{-\infty}^k, B \in \mathcal{F}_{k+n}^{\infty}, P(A) > 0 \right\}. \quad (12)$$

If

$$\sum_{n=1}^{\infty} \phi_n^{1/2} < \infty \quad (13)$$

then the environment process is said to be **uniform mixing**.

3 Contractive properties

In this section, we show that iteration of the model eventually produces uniform geometric contraction in Hilbert's projective metric (defined below), which implies stable population growth. This section is the most technical part of the paper. It is safe to skip it and go directly to the conclusions and applications in Sect. 4 and thereafter, unless you want to check the proofs. The conclusions here are essentially the same as ([34], Sect. 3) for the stochastic matrix model. However the infinite-dimensional case requires some new arguments that do not use special properties of Euclidean space such as compactness of the unit sphere.

Because of compactness and continuity the iteration (6) takes $L_1(\mathbf{X})$ into $C(\mathbf{X})$, the space of bounded continuous functions on $\mathbf{X}$–note that $C(\mathbf{X}) \subset L_1(\mathbf{X})$ under our assumptions. We can therefore regard the model as a linear operator on either $L_1(\mathbf{X})$ or on $C(\mathbf{X})$. Unless otherwise noted, functions on $\mathbf{X}$ are assumed to lie in $C(\mathbf{X})$ and operators (on spaces of functions mapping $\mathbf{X} \to \mathbb{R}$) are assumed to map all of $L_1(\mathbf{X})$ into $C(\mathbf{X})$. We always use the sup-norm $\|f\|_{\sup} = \sup\{|f(x)| : x \in \mathbf{X}\}$ on $C(\mathbf{X})$ and the partial order induced by the cone $C^+(\mathbf{X})$ of nonnegative functions, i.e. $f \ge g$ means that $f - g \in C^+(\mathbf{X})$. For this section, without loss of generality we assume that $\mu(\mathbf{X})$=1.

For the case of interest here, $C(\mathbf{X})$ ordered by $C^+(\mathbf{X})$, the projective metric $\rho(f,g)$ can be defined as follows (see e.g. [17] for the general definition). For $f, g \in C(\mathbf{X})$ let

$$S(f,g) = \sup\{(f(x)/g(x) : x \in \mathbf{X}\}, \quad I(f,g) = \inf\{f(x)/g(x) : x \in \mathbf{X}\}.$$

If $I(f,g) > 0$ and $S(f,g)$ is finite then f and g are called *comparable* and their projective distance is defined to be

$$\rho(f,g) = \log \frac{S(f,g)}{I(f,g)}. \tag{14}$$

Actually ρ is just a semi-metric because $\rho(f,cf) = 0$ for any positive scalar c; more generally $\rho(cf,dg) = \rho(f,g)$ if $c,d > 0$ and f and g are comparable. However, ρ can induce a genuine metric on subsets that cannot contain two distinct multiples of the same function. Let $C^{++}(\mathbf{X})$ denote the set of everywhere-positive functions in $C(\mathbf{X})$ and let

$$\Sigma = \{f \in C^{++}(\mathbf{X}) : \|f\|_1 = 1\}. \tag{15}$$

Lemma 1 *$\rho(f,g)$ defines a metric on Σ which is equivalent to $\|f-g\|_{\sup}$, in the sense that convergence in either metric implies convergence in the other for a sequence of functions in Σ.*

Proof This is a conclusion of Theorem 1.2.1 in [17] so we just need to verify the necessary assumptions. See pp. 412–414 in [17] for the definitions of any unfamiliar terms. We are working in the space $\mathbf{B} = C(\mathbf{X})$ ordered by the cone $V = C^{+}(\mathbf{X})$. We need to verify that

1. $C^{+}(\mathbf{X})$ is an almost Archimedean cone in $C(\mathbf{X})$.
2. Every pair of elements in Σ is regularly comparable.
3. Σ is a subset of a component C_u of $C^{+}(\mathbf{X})$, such that the cone V_u induced by C_u is *normal* in the subspace $\mathbf{B}_u$ generated by C_u.
4. V_u has nonempty interior in $\mathbf{B}_u$.

The cone $C^{+}(\mathbf{X})$ is normal and therefore almost Archimedean, because $0 \le f \le g$ implies that $\|f\|_{\sup} \le \|g\|_{\sup}$. All functions in Σ have a finite, positive infimum and supremum, hence all pairs of functions in Σ are comparable. If $f,g \in \Sigma$ either they are equal–in which case $I(f,g) = S(f,g) = 1$–or else there is an x_0 at which one of them is strictly smaller than the other, say $f(x_0) < g(x_0)$, so there must also be an x_1 at which the reverse is true. In that case $I(f,g) < 1 < S(f,g)$. Consequently, all $f,g \in \Sigma$ have $I(f,g) \le 1 \le S(f,g)$, which is the definition of being regularly comparable.

For the last two items listed above we choose $u(x) \equiv 1$. Then

- The component C_u is by definition the set of functions in V comparable to u, which is $C^{++}(\mathbf{X})$.
- The subspace $\mathbf{B}_u$ is the set of functions $f \in \mathbf{B}$ such that $-ru \le f \le ru$ for some $r \ge 0$ ([17], p. 412) which is all of $C(\mathbf{X})$.
- The cone induced by C_u is by definition $V_u = V \cap \mathbf{B}_u$, which is $C^{+}(\mathbf{X})$.

So $\Sigma \subset C_u$ by definition. $V_u = C^{+}(\mathbf{X})$ is normal in $B_u = C(\mathbf{X})$ as noted above, and its interior in B_u under the sup-norm is clearly nonempty. □

There are also some explicit inequalities between ρ and other metrics. For example, the cone of non-negative functions is normal in both $L_1(\mathbf{X})$ and $C(\mathbf{X})$, so Theorem 3.4.4 of [16] implies that for all $f, g \in \Sigma$

$$\begin{aligned} \|f - g\|_1 &\leq 3(e^{\rho(f,g)} - 1), \\ \|f - g\|_{\sup} &\leq 3 \min\{\|f\|_{\sup}, \|g\|_{\sup}\}(e^{\rho(f,g)} - 1). \end{aligned} \tag{16}$$

In the opposite direction, applying Theorem 3.4.1 of [16] to $C(\mathbf{X})$ ordered by $C^+(\mathbf{X})$, we have that if $f, g \in C^{++}(\mathbf{X})$ and $\|f - g\|_{\sup} < \inf(g)$ then

$$\rho(f, g) \leq \log \frac{\inf(g) + \|f - g\|_{\sup}}{\inf(g) - \|f - g\|_{\sup}}. \tag{17}$$

Contraction properties are based on Lemmas below showing that the normalized population distribution eventually enters and then remains in a compact set. Define $T \boxdot n = Tn / \|Tn\|_1$ for $n \neq 0$. The dynamics of the population structure are then

$$u_{t+1} = \mathbf{K}(\theta_t) \boxdot u_t. \tag{18}$$

Let $W = \left\{n \in C^+(\mathbf{X}) : \|n\|_1 = 1\right\}$, and

$$\begin{aligned} U = \{\mathbf{Q}(\psi) \boxdot w : \psi \in \Psi, w \in W\} = \big\{&\mathbf{K}(\theta_m) \boxdot \mathbf{K}(\theta_{m-1}) \boxdot \cdots \mathbf{K}(\theta_1) \boxdot \\ &w : \theta_1, \theta_2, \ldots, \theta_m \in \Theta, w \in W\big\}. \end{aligned}$$

The population structure $n / \|n\|_1$ enters U within m time steps, and remains in U thereafter. The compact set is $\bar{U}$, the closure of U in $C(\mathbf{X})$.

Lemma 2 *If $u \in \bar{U}$ then $\frac{\alpha_1}{\alpha_2} \leq u(x) \leq \frac{\alpha_2}{\alpha_1}$ for all $x \in \mathbf{X}$.*

Proof If $u \in U$ then $u = \mathbf{Q}(\psi)w / \|\mathbf{Q}(\psi)w\|_1$ for some $\psi \in \Psi, w \in W$. Writing out

$$\mathbf{Q}(\psi)w(x) = \int_X Q(x, z, \psi)w(z) \mathrm{d}\mu(z)$$

and using (10) we have that

$$\alpha_1 \leq \mathbf{Q}(\psi)w(x) \leq \alpha_2 \tag{19}$$

for all $x \in \mathbf{X}$. Integrating over $\mathbf{X}$ gives

$$\alpha_1 \leq \|Q(\psi)w\|_1 \leq \alpha_2, \tag{20}$$

so the lemma holds for all $u \in U$. Convergence in $C(\mathbf{X})$ implies pointwise convergence, so the lemma also holds for all $\bar{u} \in \bar{U}$. □

Lemma 3 *$\bar{U}$ is compact in $C(\mathbf{X})$.*

Proof U is pointwise compact by the last lemma, so by the Arzela–Ascoli Theorem we only need to show that the elements of U are equicontinuous. Consider the function

$$M(x,y,z,\psi) = |Q(z,x,\psi) - Q(y,x,\psi)| \tag{21}$$

on $\mathbf{X}^3 \times \Psi$. M is continuous on a compact metric space, therefore uniformly continuous. Let $\delta(x,y)$ denote the metric on $\mathbf{X}$ and define

$$f(r) = \sup\{|Q(z,x,\psi) - Q(y,x,\psi)| : \psi \in \Psi, x,y,z \in \mathbf{X}, \delta(y,z) \le r\}. \tag{22}$$

Uniform continuity of M implies that f is finite and $f(r) \to 0$ as $r \to 0$.

If $u \in U$ we have $u = \mathbf{Q}(\psi) \boxdot w$ for some $\psi \in \Psi, w \in W$. Let $\tilde{u} = Q(\psi)w$ so $u = \tilde{u}/\left\|\tilde{u}\right\|_1$. Then

$$\left|\tilde{u}(z) - \tilde{u}(y)\right| \le \int_X f(\delta(z,y))w(x)\mathrm{d}\mu(x) = f(\delta(z,y)) \tag{23}$$

so with (20) we have $|u(z) - u(y)| \le f(\delta(z,y))/\alpha_1$. □

The next Theorem is the main contraction result. The statement is the same as as Theorem 2 of [34], which is very similar to a previous result of Joel Cohen.

Theorem 1 *There exist constants $0 \le r < 1$ and $k_1 \ge 0$ such that*

$$\rho(\mathbf{K}(\theta_n) \boxdot \cdots \boxdot \mathbf{K}(\theta_1) \boxdot u, \mathbf{K}(\theta_n) \boxdot \cdots \boxdot \mathbf{K}(\theta_1) \boxdot v) \le k_1 r^n \tag{24}$$

for all $\theta_1, \theta_2, \ldots, \theta_n \in \Theta, u, v \in \bar{U}$ (for $n = 0$ the lefthand side of (24) is interpreted as $\rho(u,v)$).

Proof By (10) any $Q(\psi), \psi \in \Psi$ satisfies the assumptions of Corollary 2.6.1 in [16] with κ (as defined there) being $\le \alpha_2/\alpha_1$, so there exists a constant $q, 0 \le q < 1$ such that if u, v are any two comparable elements of $C^+(\mathbf{X})$ and $\psi \in \Psi$, then

$$\rho(\mathbf{Q}(\psi) \boxdot u, \mathbf{Q}(\psi) \boxdot v) \le q\rho(u,v). \tag{25}$$

q is called the Birkhoff contraction ratio, and the bound on κ implies that $q \le \frac{\alpha_2 - \alpha_1}{\alpha_2 + \alpha_1}$ ([16], p. 437). Choose any $u, v \in \bar{U}$ and any sequence $\theta_1, \theta_2, \ldots$ from Θ. Define

$$u_j = \mathbf{K}(\theta_j) \boxdot \cdots \boxdot \mathbf{K}(\theta_1) \boxdot u, \quad \text{and} \quad v_j = \mathbf{K}(\theta_j) \boxdot \cdots \boxdot \mathbf{K}(\theta_1) \boxdot v.$$

For $n \leq m$ we have $S(u_n, v_n) \leq \left(\frac{\alpha_2}{\alpha_1}\right)^2$ and $I(u_n, v_n) \geq \left(\frac{\alpha_1}{\alpha_2}\right)^2$ by Lemma 2, so $\rho(u_n, v_n) \leq 4 \log \frac{\alpha_2}{\alpha_1}$. For $n > m$ we can write $n = jm + k$, so by (25)

$$\rho(u_n, v_n) \leq q^j \rho(u_k, v_k) \leq 4q^j \log \frac{\alpha_2}{\alpha_1}.$$

These clearly imply (24) with $r = q^{1/m}$. □

Lemma 4 *There exists a constant $k_2 \geq 0$ such that*

$$\frac{\|\mathbf{K}(\theta_n) \cdots \mathbf{K}(\theta_1) v_1\|_1}{\|\mathbf{K}(\theta_n) \cdots \mathbf{K}(\theta_1) v_2\|_1} \leq 1 + k_2 r^m \tag{26}$$

whenever

$$v_i = \mathbf{K}(\theta_m) \boxdot \cdots \boxdot \mathbf{K}(\theta_1) \boxdot u_i, u_i \in \bar{U}, \theta_i \in \Theta. \tag{27}$$

If $n = 0$ then set $\mathbf{K}(\theta_n) \cdots \mathbf{K}(\theta_1) v_i = v_i$, and if $m = 0$ then set $v_i = u_i$. From (27) *it follows that*

$$\log \frac{\|\mathbf{K}(\theta_n) \cdots \mathbf{K}(\theta_1) v_1\|_1}{\|\mathbf{K}(\theta_n) \cdots \mathbf{K}(\theta_1) v_2\|_1} \leq k_2 r^m. \tag{28}$$

Proof By (24) $\rho(v_1, v_2) \leq k_1 r^m$ for $m \geq 0$, so by (16) $\|v_1 - v_2\|_{\sup} \leq 3\frac{\alpha_2}{\alpha_1} k_1 r^m$. Because the v_i are in U, this last inequality and Lemma 2 imply that for some constant k_2 we have $v_1 \leq (1 + k_2 r^m) v_2$. This is Eq. 5 in [34], and the rest of the proof is identical. □

Lemma 5 *Let $w_1, w_2 \geq 0$ be nonzero bounded measurable functions on $\mathbf{X}$. Then there is a constant k_3, depending only on w_1 and w_2 such that*

$$\frac{\langle w_1, \mathbf{K}(\theta_m) \cdots \mathbf{K}(\theta_1) u_1 \rangle}{\langle w_2, \mathbf{K}(\theta_m) \cdots \mathbf{K}(\theta_1) u_2 \rangle} \leq k_3 \tag{29}$$

whenever $\theta_j \in \Theta$ and $u_1, u_2 \in U$.

Proof Identical to the proof of Lemma 4 in [34]. □

3.1 Extensions

We mention here two extensions that may be useful in empirical applications. The first is inclusion of senescent (post-reproductive) individuals in the model. Suppose that the state space is the union of disjoint compact sets $\mathbf{X}_r$ and $\mathbf{X}_s$ representing potentially reproductive and senescent individuals–these could be 2 copies of $\mathbf{X}$ with individuals moving into $\mathbf{X}_s$ when they senesce. The

model is then specified by four component operators $\mathbf{K}_{i,j}(\theta), i,j \in \{r,s\}$ with $\mathbf{K}_{i,j}$ describing $\mathbf{X}_j \to \mathbf{X}_i$ transitions and $\mathbf{K}_{r,s} \equiv 0$ by definition. Symbolically,

$$\begin{pmatrix} n_r(t+1) \\ n_s(t+1) \end{pmatrix} = \begin{pmatrix} \mathbf{K}_{r,r}(\theta) & 0 \\ \mathbf{K}_{s,r}(\theta) & \mathbf{K}_{s,s}(\theta) \end{pmatrix} \begin{pmatrix} n_r(t) \\ n_s(t) \end{pmatrix} \tag{30}$$

where $n_r(t)$ and $n_s(t)$ are the state-distribution functions on $\mathbf{X}_r$ and $\mathbf{X}_s$. The corresponding components $\mathbf{K}_{i,j}^{(m)}$ of any m-step operator can be computed by using "matrix multiplication" to iterate (30).

Model (30) cannot be power-positive because initial conditions with only senescent individuals lead to extinction. Suppose instead that:

- There is an upper limit to the post-reproductive lifespan, i.e. there exists some q such that

$$\mathbf{K}_{s,s}(\theta_q) \circ \mathbf{K}_{s,s}(\theta_{q-1}) \circ \cdots \circ \mathbf{K}_{s,s}(\theta_1) \equiv 0$$

  for all $\theta_1, \theta_2, \ldots, \theta_q \in \Theta$.
- There is an $m \geq q$ such that the m-step kernel components $K_{r,r}^{(m)}(y,x)$ and $K_{s,r}^{(m)}(y,x)$ are both positive on their entire domains for all possible $\theta_1, \theta_2, \ldots, \theta_m \in \Theta$. That is: so long as some reproductive individuals are present initially, $n(t)$ will be positive on all of $\mathbf{X}$ for $t \geq m$.

Then (again following [34]) the results of this section and the conclusions stated in the next section still apply to populations started with some individuals in $\mathbf{X}_r$. We define $W = \left\{n \in C^+(\mathbf{X}) : \|n\|_1 = 1, n \equiv 0 \text{ on } \mathbf{X}_s\right\}$, and the constants α_1, α_2 are defined for the m-step kernel projecting from $\mathbf{X}_r$ at time t to $\mathbf{X}$ at time $t+m$,

$$K^{(m)}(y,x) = \begin{cases} K_{r,r}^{(m)}(y,x), & y \in X_r,\ x \in X_r \\ K_{s,r}^{(m)}(y,x), & y \in X_s,\ x \in X_r \end{cases}. \tag{31}$$

The contractive properties are derived as above using the fact that the $s \to r$ and $s \to s$ components of any m-step operator are identically zero, so only the uniformly positive kernel (31) contributes to the m-step dynamics.

The second extension is directed at models where some traits have deterministic dynamics. For example, the genotype (for Mendelian traits) or breeding value (for quantitative traits) remains constant across an individual's lifetime; changes only occur when parents are replaced by offspring. This is not a problem for discrete traits (e.g. a finite set of genotypes), but a continuous trait that remains fixed or changes deterministically over an individual's lifetime is not represented by a smooth kernel. These situations may be still handled in our framework if there is a maximum lifespan M. If parent-to-offspring trait transmission is described by a smooth kernel, then typically the $(M+1)$-step iterate will have a smooth kernel satisfying our assumptions. The conclusions in the next section then apply to the process observed at times $0, M+1, 2(M+1), \ldots$, and therefore to the entire process.

4 Conclusions for the general integral model

With the results of the previous section, plus the straightforward verification that certain functions are continuous, the rest of [34] now carries over. The general long-term growth properties of stochastic matrix projection models therefore also hold for density-independent stochastic IPMs. In this section we summarize those properties.

The first set of properties concerns the long-run growth and population structure. Specifically, under the assumptions in Sect. 2 we have (paraphrasing Chap. 2 of [48])

A. The logarithm of the total population size $N(t) = \|n(t)\|_1$ has a long-term growth rate $\log \lambda_s$ which is constant with probability 1,

$$\log \lambda_s = \lim_{t\to\infty} t^{-1} \log N(t) = \lim_{t\to\infty} t^{-1} E \log N(t). \tag{32}$$

The same is true for any part of the population $N_w(t) = \langle w, n(t)\rangle$ where $w \geq 0$ is a nonzero bounded measurable function on $\mathbf{X}$,

B. Starting from any nonzero initial population structure $n(x,0) = n_0(x)$ the population structure $u(t) = n(t)/\|n(t)\|_1$ converges to a time-dependent stationary random sequence of structure vectors $\hat{u}(t)$ which is independent of n_0.

C. The joint sequence of environment states and stationary population structures $(\theta_0, \hat{u}(0), \theta_1, \hat{u}(1), \theta_2, \ldots)$ is a stationary ergodic process.

D. The long-term growth rate can be computed as the average one-step growth rate, $\log \lambda_s = E \log \left\|\mathbf{K}(\theta)\hat{u}\right\|_1$ where the expectation is with respect to the joint stationary measure in the last item.[1]

The intuitive picture behind these properties is that

1. A joint stationary distribution (property C above) must exist due to compactness (see Lemma 1 in [19] or Chap. 1 Lemma 2.2 in [33]).
2. Because of contraction in the projective metric, two copies of the population structure process running under the same sequence of environment states converge exponentially fast onto each other. So starting from anywhere (copy 1) the population structure converges onto the stationary process (copy 2). The stationary distribution is therefore unique, and it must be ergodic because the initial state has no effect on long-term behavior.

In general, it isn't or is not possible to find an explicit formula for the stationary distribution, even if the environment process is simple (e.g., independent and identically distributed over time). However, the results above ensure that the long-run properties of interest really do exist and that they can be estimated by

[1] In [48] properties C and D are only stated as holding under the assumption that the environment process is a countable state Markov chain, but the arguments in [34] show that they hold for any stationary ergodic environment.

simulation: start anywhere, iterate the model for a long time, and discard initial transients.

The other important property is the asymptotic lognormal distribution of population size. As with matrix models, this occurs unless the environment process has long-memory autocorrelation. The intuitive picture behind this property is that the log of total population size is approximately a random walk. In the unstructured model $n(t+1) = \lambda(t)n(t)$ with no between-year correlations in the growth rates $\lambda(t)$, $\log n(t)$ really is a random walk so its distribution is asymptotically Normal by the Central Limit Theorem. In a structured population this is not quite true because the population structure induces correlations between the population growth rates at successive times. But because of property B, the structure-induced autocorrelations decay rapidly. The random walk approximation is then valid over time scales longer than the correlation time of population structure, unless there are long-memory correlations in the environment process that by themselves induce long-memory correlations in the population growth rate.

The conclusion (equivalent to Theorem 9 in [34]) is:

E. If the environment process $\theta(t)$ is uniformly mixing (as defined above in Eq. (13)), then the asymptotic distribution of the total population size or any part of the population is lognormal, i.e.

$$(\log N(t) - t \log \lambda_s) / \sqrt{t} \Rightarrow \text{Normal}(0, \sigma^2) \tag{33}$$

for some $\sigma \geq 0$, where $\Rightarrow$ denotes convergence in distribution and $N(t)$ is defined in item A above.

Uniform mixing holds if the environment process is an ergodic finite-state Markov chain, but this is not necessarily true for ergodic Markov chains on general state spaces [37]. A range of results similar to (33) could be stated under weaker assumptions about the mixing rate of the environment process, leading to less detailed conclusions about the convergence rate in (33). Central Limit Theorems for general stationary processes involve scaling by the standard deviation of partial sums. Using these in place of the Central Limit Theorm for uniform-mixing processes would produce generalizations of Ishitani's Central Limit Theorem for subadditive processes [29], as observed in [21]. These would in turn imply results like (33), but without an explicit expression for the denominator on the left-hand side.

5 Stochastic IPM for *O. illyricum*

We turn now to the second purpose of this paper: demonstrating how the results of the previous section make the stochastic integral model a practical option—often with significant advantages—for applications that until now have been addressed using stochastic matrix models or individual-based simulations. One advantage—shared by deterministic IPMs and discussed at length in [15]—is that IPMs are more parsimonious and require less data for parameterization

than a matrix model when individual demographic performance is affected by multiple continuous traits. Another advantage—specific to stochastic models and illustrated here—is the ease with which stochastic variability can be modeled and estimated from empirical data, just by including "year" as a categorical predictor variable in regression models for demographic rates.

Our case-study to illustrate these points is a stochastic IPM for the Illyrian thistle, *Onopordum illyricum*, based on a six-year field study [41]. As in [41] and [15] we will use the model to understand the selective forces responsible for this plant's flowering "strategy"—the functional relationship between plant age, size, and flowering probability. Previously this question had to be addressed through an individual-based simulation model for the population dynamics and the evolution of parameters characterizing the flowering strategy [41]. Here, we show how statistical analysis of the field data translates directly into a stochastic IPM for the population, and how the general properties derived in the previous section make it straightforward to use the model for identifying evolutionarily stable (ESS) flowering strategies. The field study and data analysis have been described in detail [41] so we repeat here only the information needed to understand how the data are used to construct an IPM.

5.1 Field study

Onopordum illyricum is a monocarpic perennial (reproduction is fatal), which reproduces only by seed. These form a seed bank (up to 190 seeds m^{-2}) with a typical half-life of 2–3 years. There were two study sites in southern France, a horse- and cattle-grazed pasture near Viols-en-Laval, and a sheep-grazed semi-arid steppe habitat in the Plaine du Crau. Twenty 1×2 m quadrats were placed at random within a 40×40m area of high plant density at each site. Sampling ran from August 1987 to August 1992, which included the complete lifetime of the 1987 seedling cohort. Plants were censused four times yearly (August, November, March and May) to monitor their growth, survival, and seed production. At each census the location and diameters (the longest and its perpendicular) were recorded for each plant in the quadrats. The log-transformed maximum of the November, March and May rosette areas was used as the measure of plant size x. Additional visits were made each summer to collect all flowering heads within the study quadrats. All apparently viable seeds were counted and then scattered randomly in their quadrat of origin.

5.2 Data analysis

Statistical models were fitted using data from both study sites, with site effects included when significant. For simplicity we present results from only one site, Plaine du Crau. Most analyses only used the data from 1988 to 1991, because sizes were not recorded in 1987 and death could not be distinguished from seasonal disappearance in 1992. The final models and parameter estimates are summarized in Table 1.

Table 1 Statistical models and parameter estimates describing the individual-level demography of *Onopordum illyricum*

Demographic process	Model
Growth	1988: $\bar{y} = 0.38 + 1.19x + 0.48a - 0.026x^2 - 0.084ax$ 1989: $\bar{y} = 1.82 + 0.97x + 0.48a - 0.026x^2 - 0.084ax$ 1990: $\bar{y} = 2.37 + 0.94x + 0.48a - 0.026x^2 - 0.084ax$ 1991: $\bar{y} = 0.39 + 1.23x + 0.48a - 0.026x^2 - 0.084ax$
Variance about the growth curve	$\sigma^2 = 38.5 \exp(-0.69\bar{y})$ $n = 656, P < 0.0001$
Survival probability	1988: $\text{Logit}(s) = -2.52 + q + 1.30x - 1.27a$ 1989: $\text{Logit}(s) = -3.57 + q + 1.30x - 1.27a$ 1990: $\text{Logit}(s) = -0.52 + q + 1.30x - 1.27a$ 1991: $\text{Logit}(s) = -2.90 + q + 1.30x - 1.27a$ $n = 1397, P < 0.0001$
Flowering probability	$\text{Logit}(p_f) = -24.01 + 2.91x + 0.84a$ $n = 721, P < 0.0001$
Fecundity (seeds per flowering plant)	$f_n = \exp(-11.84 + 2.27x)$ $n = 49, P < 0.0001$
Distribution of seedling size	Gaussian with mean=1.06, variance=3.37 truncated at zero $n = 389$
Distribution of seedling quality	Gaussian with mean zero and standard deviation $\sigma_s = 1.30$

The models are functions of log rosette area x, age a and individual quality q. The predicted values are the conditional mean ($\bar{y}$) and variance (σ^2) of log size next year given current size, survival probability s, flowering probability p_f, and fecundity f_n

Because we have only 4 years of data, we use a "fixed effects" approach to modeling between-year environmental variability. That is, statistical models for demographic rates were fitted with "year" as a categorical independent variable treated as unordered. A demographic model fitted this way implies 4 year-specific demographic models that result from holding "year" constant at one level—these are the year-specific regression equations presented in Table 1. These equations are components of the survival-growth kernel P, so they imply 4 year-specific kernels $P(y, x, \theta), \theta \in \{1, 2, 3, 4\}$, as described below.

In contrast, we use a "random effects" approach to analyze between-individual variability because we have repeated observations on many individual plants. That is, rather than fitting a series of individual-specific models, we estimate the distribution of parameter values across individuals (specifically, the mean and variance of the intercept parameter in a logistic regression model for survival).

Seedling size was well described by a normal distribution truncated at zero (see Fig. 1a in [15]). The mean of the seedling size distribution varied from year to year ($P < 0.001$). However this had little effect on the model predictions and so for consistency with our earlier analysis this was ignored. Annual changes in plant size were described by a complicated regression model with site × size × year ($P < 0.0001$) and age × size ($P < 0.02$) interaction terms, and a quadratic size term ($P < 0.001$).

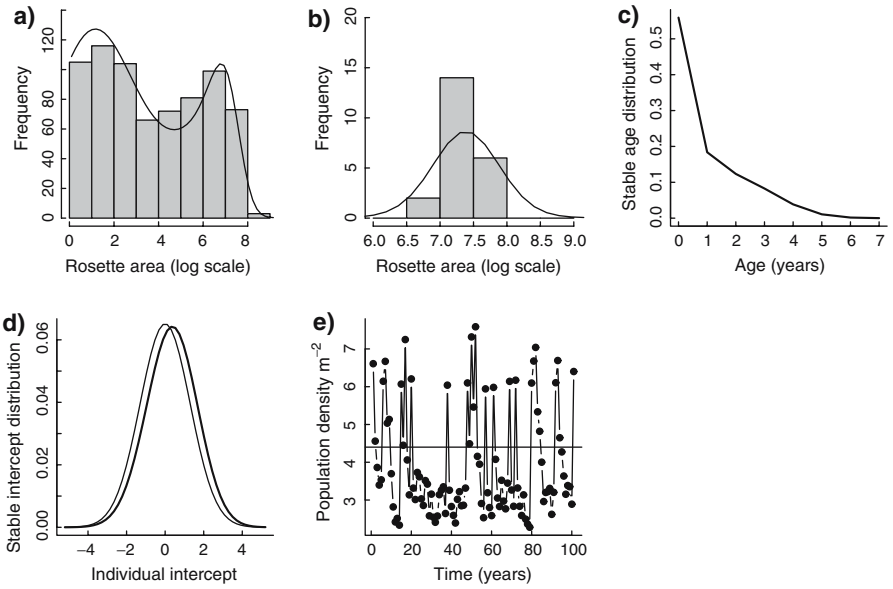

Fig. 1 Output from the *Onopordum* model, showing **a** the stationary distribution of plant sizes (*line*) and the observed distribution of plant sizes (*grey bars*), **b** the stationary distribution of size at flowering (*line*) and observed distribution of size at flower (*grey bars*), **c** the stationary age distribution, **d** the stationary distribution of individual survival intercepts (*bold curve*) and the distribution of survival intercepts in recruits (*light curve*), and **e** a time series of total population sizes and the observed average population size (*horizontal solid line*). Stationary distributions were estimated by running the model for 5,000 years and averaging with the first 1,000 years discarded

Survival probability was modeled as a mixed logistic regression with size, age, site, and year as independent variables and a Gaussian distribution of individual-specific intercepts. Plant size was the most important predictor of survival ($P < 0.0001$), and the next most important was the site by year interaction ($P < 0.0001$). Plant age was also highly significant ($P < 0.0001$), and there was significant between-individual heterogeneity in the intercept ($P < 0.002$). Survival probability increases with plant size and decreases with age. The standard deviation of the intercept distribution σ_s, quantifies the variability among individuals. This variability may reflect differences in the local competitive environment, abiotic conditions, genetic differences or other properties that remain constant over an individual's life.

Flowering probability was modeled by standard logistic regression because there was no evidence of between-individual variation ($P > 0.1$). Flowering probability increased with plant size ($P < 0.0001$) and age ($P < 0.008$), but site effects were not significant ($P > 0.1$). Seed production is strongly size-dependent ($P < 0.0002$) and highly variable, although year effects were not significant ($P > 0.05$).

Intraspecific competition with neighbors had very little influence on growth and survival. In contrast, despite seed production being highly variable (0–2,750

per quadrat), the number of recruits was remarkably constant and independent of seed production the previous year ($P > 0.2$). Our density-dependent model therefore assumes that population growth is limited by microsite availability.

In summary, growth, survival and the probability of flowering are all size- and age-dependent. In addition growth, survival and fecundity vary from year to year and there is substantial variation between individuals in survival that is not related to size, age or year. Per-capita seed production by flowering plants is strongly size-dependent, but does not fluctuate from year to year—variation in fecundity is driven by variation in the probability that a plant of a given size "now" will survive the winter and flower "next year".

5.3 General structure of the *Onopordum* model

The demographic models imply an IPM in which individuals are cross-classified by size, age and individual quality (expressed by the intercept parameter in the logistic regression for survival). Size and quality are continuous variables, and age is discrete. The individual state space $\mathbf{X}$ therefore is a union of rectangles Ω_a defined by the possible size (x) and quality (q) ranges for an individual of age a.

Survival-growth transitions send $\Omega_a \to \Omega_{a+1}$ and are represented by kernels $P_{a,q}(y,x)$ where x and y are log-transformed rosette area:

$$n_{a+1}(y,q,t+1) = \int_{L_a}^{U_a} P_{a,q}(y,x,t) n_a(x,q,t) \mathrm{d}x, \quad a = 1,2,\ldots,M-1. \tag{34}$$

Note that there is no integration over q on the right-hand side because q is constant over the lifetime (by assumption). The kernels $P_{a,q}$ are derived from the fitted demographic models for survival, flowering and growth. In the notation of Table 1,

$$P_{a,q}(y,x,t) = s(x,a,q,t)(1 - p_f(x,a))g(y,x,a,t). \tag{35}$$

That is, to reach size y from size x the individual must survive, not flower (because flowering is fatal), and change size from $x \to y$. Formulas for s and p_f are given in Table 1, and the growth kernel g is given by the conditional size distribution from Table 1: $y \sim$ Normal with mean $\bar{y}(x,a,t)$ and variance $\sigma^2 = 38.5 \exp(-0.69\bar{y})$.

We presume that offspring size and quality variation are primarily caused by fine-scale variation in properties of the microsite where the individual becomes established. Therefore, the model assumes complete "mixing at birth"[15]—the distribution of offspring size and quality are the same for all parent plants.

Fecundity goes from all components to those with $a = 0$,

$$n_0(y,q,t+1)=p_e(t)\varphi_0(y)\alpha(q)\left[\sum_a \iint_{\Omega_a} f_n(x)p_f(x,a)s(x,a,q',t)n_a(x,q',t)\mathrm{d}x\mathrm{d}q'\right], \tag{36}$$

where φ_0 is the probability density for offspring size and α is the probability density for offspring quality, as specified in Table 1. The bracketed sum in (36) is the total seed production by all parents, and this is multiplied by the offspring size $\times$ quality distribution $\varphi_0(y)\alpha(q)$ and the establishment probability $p_e(t)$ to give the distribution of new recruits. Equation (36) implicitly defines the set of fecundity kernels $p_e(t)F_a(y,x,q,q',t)$, each corresponding to one term in the sum, which specify the per-capita number of new offspring of quality q produced by age-a parents of quality q'.

The between-year variability in survival, growth and fecundity is modeled by sampling at random from the 4 year-specific survival-growth and fecundity kernels estimated from the field study. That is, for each year t we randomly draw $\theta(t) \in \{1,2,3,4\}$ with equal probabilities, and set

$$P_{a,q}(y,x,t) = P_{a,q}(y,x,\theta(t)), \quad F_a(y,x,q,q',t) = F_a(y,x,q,q',\theta(t)).$$

This corresponds to the *matrix selection* approach to constructing stochastic matrix models ([3,38] Sect. 14.5.3), and can be regarded as a nonparametric bootstrap from the set of estimated kernels.

To complete the model we need to specify the seedling establishment probability $p_e(t)$. This is the one stage of the life cycle where there is density-dependence, so for the purpose finding ESS flowering strategies we need to consider both a resident population with one flowering strategy, and a rare invading population with a different flowering strategy but otherwise identical to the resident.

In the field study, total seedling recruitment each year was found to be independent of population size the previous year. So in the model the total number of new recruits each year, $R(t)$, is a sequence of values drawn independently at random from the estimated numbers of recruits in each year of the field study. Because invaders and residents differ only in their flowering strategy, all seedlings compete on an equal footing for the available microsites. Thus $p_e(t)$ for both resident and invader is given by $R(t)$ divided by the total seed production. In an ESS analysis we regard the invader as being so rare that the "total seed production" is really the resident's total seed production $S_r(t)$. So for both resident and invader

$$p_e(t) = R(t+1)/S_r(t) \tag{37}$$

where the resident's seed production $S_r(t)$ is given by the sum in (36) computed for the resident population using the resident's flowering parameters.

Equation (37) is density dependent because $S_r(t)$ depends on the number of residents as well as on their age×size×quality distribution. Because the number of residents and their state-distribution depend on the resident's flowering strategy, (37) imposes a feedback between flowering strategy and fecundity that affects the selective pressures acting on the population. So although resident and invader fecundity are equally affected by between-seedling competition each year, the density dependence cannot be ignored when we use the model to predict ESS flowering strategies.

However, the model for the invader is a density-independent IPM in which the "environment" process is $\theta_i(t) = (R(t+1), n_r(t), \theta(t))$, where $n_r(t)$ is the state distribution function of the resident population in year t. The year-type $\theta(t) \in \{1,2,3,4\}$ affects the invader's survival-growth kernel, and $R(t+1)$ and $n_r(t)$ jointly determine the establishment probability $p_e(t)$. In Appendix A we show that $\theta_i(t)$ satisfies our general assumptions about the environment process. The invader population therefore has a long-term growth rate λ_s, whose value determines whether or not invasion is successful. The ESS flowering strategy is defined by the property that no invader with a different strategy can achieve $\lambda_s > 1$.

Our computer implementation of the model is described in Appendix B. On the computer the IPM is implemented using matrix operations, but this is just a way of computing integrals and does not discretize the life cycle as in a conventional matrix population model. The computer implementation assumes a finite maximum age of $M = 7$ years. The maximum age observed in the field was 5 years, but the study was terminated after the last individual from the initial cohort died and some plants might live longer. With a maximum age of 7 years, the kernels for 8-year transitions satisfy our general assumptions even though the year-to-year survival, Eq. (34), is not described by a smooth kernel in (x,q)-space because q is constant over time (see Sect. 3.1 and the Discussion regarding this kind of situation).

6 Results for *O. illyricum*

6.1 Resident population

The density-dependent model for the resident population provides an accurate description of the bimodal distribution of sizes observed in the population (Fig. 1a) and the distribution of flowering sizes (Fig. 1b). It predicts that the population will dominated by new recruits (Fig. 1c), and shows that there is a change in the distribution of survival intercepts relative to new recruits (Fig. 1d). This shift reflects the increased survival of individuals with larger survival intercepts. The model also provides a reasonable description of the average population density (Fig. 1e).

6.2 ESS flowering strategies

We are interested in how the probability of flowering varies with plant size and age. This relationship is described by a logistic regression (Table 1) with intercept β_0 and slope coefficients for the effect of plant size and age, β_a and β_s, respectively. As in Rees et al. [41], when all three parameters are allowed to vary the predicted strategy is a sharp threshold: all plants of a given age should flower with probability 0 or 1, depending on whether their size is below or above a threshold that decreases with age. The observed size-dependence of flowering is gradual, representing possibly a constraint or else a decision that depends on plant size at some time between censuses [41]. We therefore imposed gradual size dependence by holding the size slope fixed at its estimated value and allowing the intercept, β_0, and age slope, β_a to evolve (see [4]).

Previous analyses have shown that ignoring temporal variation in the environment results in predicted ESS flowering strategies that are very different from that observed in the field [41] and that this difference is statistically highly significant [15].

To estimate the ESS for the stochastic model we first simulated a sequence of values for the resident population and environment process $(\theta(t), R(t+1), n_r(t))$, and then used these to compute the seed establishment probabilities, equation (37). Having done this we computed the invader's kernels for each year, and then numerically maximized the invader's long term growth rate $\log(\lambda_s)$ using the Nelder–Mead simplex algorithm. The flowering strategy (β_0, β_a) that maximizes the invader's long term growth rate then becomes the new resident. This process is repeated until successive values of $\max(\lambda_s)$ converge on 1 to some specified tolerance (0.0001), showing that no other strategy can increase when rare. The final flowering strategy (β_0, β_a) is then taken to be the putative ESS.

The results (Fig. 2a) show that there is good agreement between the observed flowering strategy and the predicted ESS from the stochastic model, with the ESS having a slightly lower flowering probability for all sizes observed in the field. This comparison is overly rigorous, as it ignores uncertainty in the predicted ESS caused by uncertainty in the values of other model parameters. Even so, the predicted ESS is within the range of statistical uncertainty on the actual flowering strategy. In contrast, even when we allow for estimation errors in the other demographic processes, the ESS flowering strategy predicted by the deterministic IPM is significantly different from the actual strategy ($P < 0.01$, [15]). There is however a small discrepancy between the ESS flowering strategy found here and that found previously [41] using individual-based simulations. This difference is largely a result of the earlier study using a different fecundity function, resulting in the gains made from flowering larger accruing more slowly and so selecting for flowering at smaller sizes or earlier. The previous fecundity model was based on an indirect measure of fecundity, the area of flowering receptacle matured, while our current model uses direct measurements of individual plant size and seed production.

To confirm that the stochastic ESS identified using the numerical procedure described above was in fact an ESS we constructed the landscape of invasion

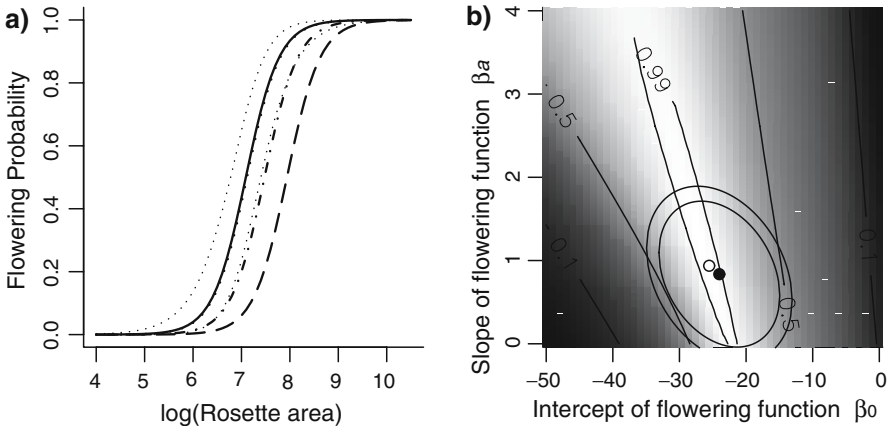

Fig. 2 **a** Estimated actual relationship and predicted ESS relationship between rosette area (on log scale) and flowering probability in *Onopordum*. The *solid* and *dotted curves* show the estimated relationship between size and flowering probability for an individual of age 3, surrounded by pointwise 95% confidence bands, based on flowering data from the field study. The confidence bands are the pointwise 2.5 and 97.5 percentiles from 5,000 draws of flowering model parameters from the multivariate Normal distribution implied by the variance–covariance matrix of parameter estimates for the fitted model. The dashed curve shows the ESS flowering probability predicted by the constant-environment IPM described in [15]. The *dash-dot* curve shows the predicted ESS from the stochastic-environment IPM described in this paper. The ESS predicted using a stochastic-environment IPM with the fecundity function used in [41] overplots almost exactly on top of the estimated flowering probability (*solid curve*). **b** Invasion fitness as a function of the parameters defining the invader's flowering strategy in *Onopordum*. Invasion fitness is the value of λ_s for an invader assuming the resident uses the predicted ESS flowering strategy (β_0, β_a). The curves marked 0.5 and 0.99 are contours of invasion fitness. The filled point shows the value of (β_0, β_a) estimated from flowering data in the field study, and the surrounding ellipses are the 95 and 99% confidence regions, calculated using the standard quadratic approximation to the likelihood (i.e., assuming that the likelihood is χ^2 distributed with 3 degrees of freedom). The open point is the predicted ESS flowering strategy. The estimated ESS necessarily has the highest invasion fitness, because it is defined by the property of having the highest invasion fitness against itself, but the fitness achieved by the estimated actual strategy is within 1% of the maximum

fitness λ_s for mutant flowering strategies invading the putative ESS (Fig. 2b). This landscape shows that the ESS was uninvadable, as all mutants have $\lambda_s < 1$. As the age-dependence of the flowering strategy becomes stronger, so the fitness landscape becomes shallower. This suggests that age and size-dependent flowering strategies are to some extend exchangeable, as a range of strategies with similar fitness can be constructed by quite different combinations of size- and age-dependence (Fig. 2b).

7 Discussion

Our general results in this paper are "the skeleton of a fairly general theory for random rates"([48], p. 33) in density-independent integral projection models. As is the case for matrix projection models these general results are essential for

substantive applications [3,48], such as our analysis here of flowering strategies in *O. illyricum*.

The deterministic *Onopordum* model developed in [15] illustrates how complex demographic transitions involving size, age and quality dependence can be directly translated into an IPM and the resulting model used to address evolutionary questions. In this paper, we have extended that model to include yearly variation in the model parameters. For an IPM this is relatively straightforward, because terms describing yearly variation in parameters of the survival and fecundity kernels can be fitted and tested for statistical significance using standard regression methods and statistical software. This greatly facilitates the construction of stochastic models, in particular relative to the lengthy procedures for fitting a conventional stochastic matrix model to the same type of data (see Chap. 8 in [38]).

The comparison of the ESS predictions of constant and stochastic models (Fig. 2a) clearly demonstrates that including environmental variability improves the predictive ability of the model. However, in order to understand how stochasticity shapes the ESS it is necessary to understand how different forms of stochasticity influence the ESS and to separate the effects of changes in mean demographic rates—caused by non-linear averaging—from variance effects caused by non-equilibrium dynamics [42]. Effects of non-equilibrium dynamics result from different flowering strategies being intrinsically favored in different years (fluctuating selection) and from the fitness of different flowering strategies fluctuating as a result of variation in population structure. These effects can operate in different directions and so in order to assess the role of stochastic variation each must be quantified; for an example of this type of analysis, see [42].

Some kind of stationarity assumption is required for the existence of long-term properties such as an asymptotic growth rate. Two types of assumptions have been used for stochastic matrix models:

1. The matrix sequence is an ergodic Markov chain or a function of an ergodic Markov chain.
2. The matrix sequence is an ergodic stationary stochastic process.

Both of these imply that the distant future becomes independent of the past, but neither of them implies the other. The first assumption requires that future values depend directly on only a finite number of past values. The second allows infinite memory, but requires that any effects of initial transients have already washed out completely, so that this year and all subsequent years have the same marginal distribution of environment states. Although finite memory seems intuitively reasonable, there is evidence that many environmental variables affecting natural populations have long-memory dependence (reviewed in [22,23]), which is revealed by slower than exponential decay in the between-year autocovariance $\mathrm{Cov}(\theta(t), \theta(t-\tau))$ as $\tau \to \infty$. This cannot occur in ergodic Markov chains, so stationarity assumptions such as ours seem to be more realistic for modeling natural populations.

Some recent studies have shown, for stochastic matrix models, that both the between-year [50] and within-year [40] correlations among matrix entries can

have a significant impact on model predictions—it is not enough to correctly specify the marginal distribution of each matrix entry. One advantage of the integral projection framework is that within-year correlations happen "automatically" as a result of variation in underlying parameters. For example, a higher slope or intercept parameter for growth in a given year will affect individuals of all sizes. Many correlations among entries in a matrix model are of this type—a good year is good for everyone, so the matrix entries for growth are all above average while those for reductions in size are below average.

In this paper, within-year correlations are naturally incorporated by constructing kernels from each year of the study—analagous to the "matrix selection" approach for stochastic models where each year's data is used to construct a year-specific projection matrix. This approach increases model realism by avoiding distributional assumptions on environmental variability, but simultaneously makes it difficult to assess the sensitivity of model predictions to between-year correlations [18] and may overestimate the between-year variance when annual parameters are drawn from a stationary distribution (see Sect. 6.5 in [45]).

An alternative approach is "element selection" where yearly matrices are constructed by selecting matrix elements from a fitted joint probability distribution [38]. For matrix models this can be a complicated procedure—the variance–covariance matrix for all pairwise correlations between matrix entries must be estimated, and then values of matrix entries must be simulated from the joint distribution in a way that satisfies sum constraints (e.g., the total survival probability for any class of individuals must be ≤ 1 [32]). However, for IPMs the corresponding approach can be implemented easily by using mixed-effects models to describe the temporal variation in demography, and constructing yearly kernels by sampling from the fitted parameter distributions (Rees and Ellner, in preparation). In this approach between-year parameter correlations have to be explicitly estimated, and are treated as a parameter of the model making it straightforward to assess their effects. The same is true in principle for matrix models, but the large number of possible correlations (between all $(a_{ij}(t), a_{kl}(t-1))$ pairs) makes it difficult to distinguish between genuinely significant correlations and type-I errors.

Here and in [15] we have formulated the integral projection model at a level of generality in between our original Euclidean space version [11,12] and the fully measure-theoretic versions of Grafen [20] and Diekmann et al. [9,10]. In all of these approaches, individual state is a Markov process on a state space general enough to allow cross-classification of individuals by an arbitrary list of continuous and discrete attributes. Our formulation sacrifices some generality by assuming that state transitions are specified by smooth kernels rather than a general transition probability [37]. Our goal was to be general enough for empirical applications but concrete enough that the assumptions have intuitive meaning and can be verified in practice for specific models, if necessary by brute-force computation (see pp. 416–417 in [15]). The rationale for smoothness is that everything in the real world is stochastic to some extent, and if the set of possible states is continuous then so is the stochasticity. It is then reasonable to

use smooth probability densities to model state transitions. This rationale fails for continuous attributes that remain fixed over the lifetime, such as individual quality in our *Onopordum* model, or that are best modeled as changing deterministically. This is not just a matter of technical difficulties in proofs; it reflects real phenomena, such as bistability in deterministic trait dynamics, that would rule out the existence of any long-term properties invariant over all initial conditions and subsequent sample paths. For *Onorpordum* the situation is rescued by assuming a finite maximum lifespan and a smooth density for offspring quality, so smoothness prevails eventually. Even when the data do not mandate using age as an individual-level state variable, making the model age-structured with a large maximum age M (and all vital rates actually independent of age) makes it simple to calculate age-based summary statistics such as the distribution of age at first reproduction, and M can be increased until numerical values stabilize. But it would still be useful to identify a slightly more general IPM formulation that readily accomodates a mix of deterministic and stochastic trait dynamics without becoming intractable or harder to understand.

Acknowledgments We thank Simon Eveson for providing unpublished portions of his doctoral thesis. For comments on the manuscript we are grateful to the Cornell EEB Eco-theory lunch bunch (including B. Chan, E.G. Cooch, M. Cortez, R. Dore, P. Hurtado, L.E. Jones, J. Mandel, V. Pasour, J. Reilly, E. Van Nostrand, and E. Zipkin) and an anonymous referee.

A Appendix: The resident population

Here, we analyze the resident process to show that the invader's environment process $\theta_i(t)$ as defined in the text is stationary and ergodic.

The resident dynamics are simplified by the decoupling of total recruitment from seed production. Combining this with mixing-at-birth, each year's cohort of new recruits is a random multiple of the offspring (size $\times$ quality) distribution, i.e. $R(t)\phi_0(y,q)$ where the number of new recruits $R(t)$ is an *iid* sequence. Let $\mathbf{P}(t)$ denote the operator on population distributions corresponding to the environment-dependent survival-growth kernels $P(\theta_t)$. The resident population dynamics are then

$$n_r(t+1) = R(t+1)\phi_0 + \mathbf{P}(t)n_r(t). \tag{38}$$

This is a time-varying linear autoregressive model, but note that it is density-dependent (in the usual meaning) because the per capita resident fecundity is inversely proportional to the resident population density. The formal solution of (38) is

$$\begin{aligned} n_r(t) = (&R(t) + R(t-1)\mathbf{P}(t-1) + R(t-2)\mathbf{P}(t-1)\mathbf{P}(t-2) \\ &+R(t-3)\mathbf{P}(t-1)\mathbf{P}(t-2)\mathbf{P}(t-3) + \cdots)\,\phi_0. \end{aligned} \tag{39}$$

The stationary resident process $\tilde{n}_r(t)$ is defined by (39) with the series continued into the infinite past. To justify this construction, note that there is a maximum

survival probability p^*, i.e.

$$\int_X P(y,x,\theta)\mathrm{d}\mu(y) \le p^* \tag{40}$$

with $p^* < 1$ by compactness. Consequently, the L_1 operator norm of any $\mathbf{P}(t)$ is at most p^*. For any realizations of the R and θ processes the series in (39) is therefore Cauchy in $L_1(\mathbf{X})$ and hence convergent. Any partial sum of the right-hand side in (39) is a well-defined random variable, so the limiting infinite sum is also measurable. The limit is clearly stationary. It is ergodic because $n_r(t)$ is a function of the ergodic process $Z(k) \equiv (R(-k), \theta(-k))$ at times $k = -t, -t+1, -t+2, \ldots$ (see Sect. 9.6 in [31]; Z is ergodic because years are independent and identically distributed by assumption). And because the R and θ processes are both *iid*, the invader's environment process with stationary resident, $\theta_i(t) = (R(t+1), \tilde{n}_r(t), \theta(t))$, is also ergodic.

B Appendix: Computer implementation of the *Onopordum* model

All calculations were done in R [39], version 2.1 or above. Integrals were evaluated numerically using midpoint rule as in [15], because it is fast enough and more robust than higher order methods. Applying midpoint rule to the integration over quality in (36) we get

$$n_0(y,q_j,t+1) = p_e(t)\varphi_0(y)\alpha_j \left[\sum_{a,k} \int_{L_a}^{U_a} f_n(x)p_f(x,a)s(x,a,q_k,t)n_a(x,q_k,t)\mathrm{d}x \right], \tag{41}$$

where q_k is the k^{th} mesh point for numerical integration in q and α_j is the fraction of offspring whose quality lies in the mesh interval containing q_j.

We can regard Eqs. (34) and (41) as defining a model in which age and quality are both discrete while size is continuous, and the one-step dynamics are represented by a set of smooth kernels; in [15] this was our initial formulation of the deterministic *Onopordum* model. The state space for the discrete-q model is the union of components $\Omega_{a,k}$ for individuals of age a and quality q_k, each of which is an interval $[L_a, U_a]$ of possible sizes for an age-a individual. Survival-growth transitions are block-structured: individuals in $\Omega_{a,k}$ either die or else move to $\Omega_{a+1,k}$. We therefore used a series of matrices to represent each survival-growth kernels, stacking them all into a 5-dimensional array $\mathbf{P}$ whose entry in location i,j,a,k,t is $hP(t)_{a,k}(x_i,x_j)$ where the x's are the quadrature mesh points and $h = x_{k+1} - x_k$. The current population state is stored in a 3-dimensional array $\mathbf{N}$ where $N[j,a,k] = n_{a,k}(x_j,t)$. This is performed for survival-growth transitions by computing for each age $\times$ quality combination the matrix–vector product $\mathbf{P}[\cdot,\cdot,a,k,t]\mathbf{N}[\cdot,a,k]$, which gives $\mathbf{N}[\cdot,a+1,k]$ at time t+1.

Births are less structured because offspring size and quality are independent of parent age, size and quality. We stored values of per-capita total seedling production $p_e(t)f_n(x_j)p_f(x_j,a)s(x_j,a,q_k,t)$ in an array $\mathbf{B}(t)$ with the same structure as $\mathbf{N}$. The total seedling production is computed as $S(t) = h \times \text{sum}(\mathbf{B}(t).\mathbf{N})$ where $\mathbf{B}(t).\mathbf{N}$ is the element-by-element product of $\mathbf{B}(t)$ and $\mathbf{N}$. The size $\times$ quality distribution for seedlings is then given by $S(t)J(t)$, where the $(j,k)^{th}$ entry in the matrix $\mathbf{J}$(t) is the fraction of seedlings of quality-class k and size x_j, calculated from the distributions in Table 1. This gives all the age-0 entries of $\mathbf{N}$ at time $t+1$, $\mathbf{N}[\cdot,0,\cdot] = S(t)J(t)$.

In all calculations we used 50 mesh points for size, 50 for quality, and imposed a maximum possible age of 7 years as discussed in the text. Increasing the maximum age had virtually no effect on model output, as very few individuals survive past age 5 in the model. For all numerical calculations we used 5,000 iterates, discarding the first 1,000.

References

1. Benton, T.G., Grant, A.: How to keep fit in the real world: elasticity analyses and selection pressures on life histories in a variable environment. Am. Nat. **147**, 115–139 (1996)
2. Birkhoff, G.: Extensions of Jentzch's Theorem. Trans. Am. Math. Soc. **85**, 219–227 (1957)
3. Caswell, H.: Matrix Population Models. Sinauer, Sunderland (2001)
4. Childs, D.Z., Rees, M., Rose, K.E., Grubb, P.J., Ellner, S.P.: Evolution of complex flowering strategies: an age and size-structured integral projection model. Proc. R. Soc. **B 270**, 1829–1839 (2003)
5. Childs, D.Z., Rees, M., Rose, K.E., Grubb, P.J., Ellner, S.P.: Evolution of size dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. Proc. R. Soc. **B 271**, 425–434 (2004)
6. Cohen, J.E.: Ergodicity of age structure in populations with Markovian vital rates. I. Countable states. J. Am. Stat. Assoc. **71**, 335–339 (1976)
7. Cohen, J.E.: Ergodicity of age structure in populations with Markovian vital rates. 2. General states. Adv. Appl. Prob. **9**, 18–37 (1977)
8. Crowder, L.B., Crouse, D.T., Heppell, S.S., Martin, T.H.: Predicting the impact of turtle excluder devices on loggerhead sea-turtle populations. Ecol. Appl. **4**, 437–445 (1994)
9. Diekmann, O., Gyllenberg, M, Metz, J.A.J., Thieme, H.R. On the formulation and analysis of general deterministic structured population models I. Linear Theory. J. Math. Biol. **36**, 349–388 (1998)
10. Diekmann, O., Gyllenberg, M., Huang, H., Kirkilionis, M., Metz, J.A.J., Thieme, H.R.: On the formulation and analysis of general deterministic structured population models II. Nonlinear Theory. J. Math. Biol. **43**, 157–189 (2001)
11. Easterling, M.R.: The integral projection model: theory, analysis and application. Doctoral thesis, North Carolina State University, Raleigh (1998)
12. Easterling, M. R., Ellner, S.P., Dixon, P.M.: Size-specific sensitivity: applying a new structured population model. Ecology **81**, 694–708 (2000)
13. Ellner, S.: Asymptotic behavior of some stochastic difference equation population models. J. Math. Biol. **19**, 169–200 (1984)
14. Ellner, S.P., Guckenheimer, J.: Dynamics Models in Biology. Princeton University Press, Princeton (2006)
15. Ellner, S.P., Rees, M.: Integral projection models for species with complex demography. Am. Nat. **167**, 410–428 (2006)
16. Eveson, S.P.: Theory and application of Hilbert's projective metric to linear and nonlinear problems in positive operator theory. D. Phil. Thesis, University of Sussex (1991)
17. Eveson, S.P.: Hilberts' projective metric and the spectral properties of positive linear operators. Proc. Lond. Math. Soc. **70**, 411–440 (1993)

18. Fieberg, J., Ellner, S.P.: Stochastic matrix models for conservation and management: a comparative review of methods. Ecol. Lett. **4**, 244–266 (2001)
19. Furstenburg, H., Kesten, H.: Products of random matrices. Ann. Math. Stat. **31**, 457–469 (1960)
20. Grafen, A.: A theory of Fisher's reproductive value. J. Math. Biol. **53**, 15–60 (2006)
21. Hall, P., Heyde, C.C.: Martingale limit theory and its applications. Academic, New York (1980)
22. Halley, J.M.: Ecology,evolution, and 1/*f*-noise. Trends Ecol. Evol. **11**, 33–37 (1996)
23. Halley, J.M., Inchausti, P.: The increasing importance of 1/*f*-noises as models of ecological variability. Fluct. Noise. Lett. **4**, R1–R26 (2004)
24. Hardin, D.P., Takáč, P., Webb, G.F.: Asymptotic properties of a continuous-space discrete time population model in a random environment. J. Math. Biol. **26**, 361–374 (1988)
25. Hardin, D.P., Takáč, P., Webb, G.F.: A comparison of dispersal strategies for survival of spatially heterogeneous populations. SIAM J. Appl. Math. **48**, 1396–1423 (1988)
26. Hardin, D.P., Takáč, P., Webb, G.F.: Dispersion population models discrete in time and continuous in space. J. Math. Biol. **28**, 406–409 (1990)
27. Heppell, S.S., Crowder L.B., Crouse D.T.: Models to evaluate headstarting as a management tool for long-lived turtles Ecol. Appl. **6**, 556–565 (1996)
28. Heppell, S.S., Crouse, D.R, Crowder, L.B.: Using matrix models to focus research and management efforts in conservation. In: Ferson, S., Burgman, M., (eds.) Quantitative Methods for Conservation Biology, pp. 148-168. Springer, Berlin Heidelberg New York (1998)
29. Ishitani, H.: A Central Limit Theorem for the subadditive process and its application to products of random matrices. Publ Res Inst Math Sci Kyoto University **12**, 565–575 (1977)
30. Kareiva, P., Marvier, M., McClure, M.: Recovery and management options for spring/summer Chinook salmon in the Columbia River basin. Science **290**, 977–979 (2000)
31. Karlin, S., Taylor, H.M.: A First Course in Stochastic Processes, 2nd ed. Academic, New York (1975)
32. Kaye, T.N., Pyke, D.A.: The effect of stochastic technique on estimates of population viability from transition matrix models. Ecology **84**, 1464–1476
33. Kifer, Y.: Ergodic Theory of Random Transformations. Birkhäuser, Boston (1986)
34. Lange, K, Holmes,W.: Stochastic stable population growth. J. Appl. Prob. **18**, 325–344 (1981)
35. McEvoy, P.B., Coombs, E.M.: Biological control of plant invaders: regional patterns, field experiments, and structured population models. Ecol. Appl. **9**, 387–401 (1999)
36. Menges, E.S.: Population viability analyses in plants: challenges and opportunities. Trends Ecol. Evol. **15**, 51–56 (2000)
37. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, Berlin Heidelberg New York (1993)
38. Morris, W., Doak, D.: Quantitative Conservation Biology. Sinauer, Sunderland (2002)
39. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0, URL `http://www.R-project.org` (2005)
40. Ramula, S., Kehtilä, K.: Importance of correlations among matrix entries in stochastic models in relation to number of transition matrices. Oikos **111**, 9–18 (2005)
41. Rees, M., Sheppard, A., Briese, D. Mangel, M.: Evolution of size-dependent flowering in *Onopordum illyricim*: a quantitative assessment of the role of stochastic selection pressures. Am. Nat. **154**, 628–651 (1999)
42. Rees, M., Childs, D.Z., Rose, K.E., Grubb, P.J.: Evolution of size dependent flowering in a variable environment: partitioning the effects of fluctuating selection. Proc. R. Soc. **B 271**, 471–475 (2004)
43. Rees, M., Childs, D.Z., Metcalf, J.C., Rose, K.E., Sheppard, A.W., Grubb, P.J.: Seed dormancy and delayed flowering in monocarpic plants: selective interactions in a stochastic environment. Am. Nat. **168**, E53–E71 (2006)
44. Rose, K.E., Louda, S., Rees, M.: Demographic and evolutionary impacts of native and invasive insect herbivores: a case study with Platte thistle, Cirsium canescens. Ecology **86**, 453–465 (2005)
45. McCulloch, C.E., Searle, S.R.: Generalized, Linear, and Mixed Models. Wiley, New York (2001)
46. Shea, K., Kelly, D.: Estimating biocontrol agent impact with matrix models: *Carduus nutans* in New Zealand. Ecol. Appl. **8**, 824–832 (1998)
47. Shea, K., Kelly, D., Sheppard, A.W., Woodburn, T.L.: Context-dependent biological control of an invasive thistle. Ecology **86**, 3174–3181 (2005)

48. Tuljapurkar, S.: Population Dynamics in Variable Environments. Springer, Berlin Heidelberg New york (1990)
49. Tuljapurkar, S., Wiener, P.: Escape in time: stay young or age gracefully? Ecol. Model. **133**, 143–159 (2000)
50. Tuljapurkar, S., Haridas, C.V.: Temporal autocorrelation and stochastic population growth. Ecol. Lett. **9**, 327–337 (2006)