INDIVIDUAL-BASED INTEGRAL PROJECTION MODELS: THE ROLE OF SIZE-STRUCTURE ON EXTINCTION RISK AND ESTABLISHMENT SUCCESS

SEBASTIAN J. SCHREIBER AND NOAM ROSS

ABSTRACT.

(1) Matrix models or integral projection models (IPMs) are commonly used to study the dynamics of structured populations, where discrete or continuous traits influence survival, growth, or reproduction. When a population's size is small, as is often the case for threatened species or potentially invasive species arriving in novel habitats, extinction risk may be substantial due to demographic stochasticity.

(2) Branching processes, which are individual-based counterparts to matrix models and IPMs, allow one to quantify these risks of extinction. For discretely structured populations, the theory of multi-type branching processes provides analytic methods to compute how extinction risk changes over time and how it depends on the size and composition of the population. Building on prior work on continuous-state branching processes, we extend these analytic methods to individual-based models accounting for any mixture of discrete and continuous population structure.

(3) The individual-based IPMs are defined by probabilistic update rules at the level of the individual which determine how each individual with a given trait value dies, changes trait value (e.g. grows in size), or produces individuals with the same or other trait values. Probabilities of extinction are shown to be analytically determined by probability-generating functionals associated with the

individual-based IPMs. In particular, we present analytical expressions for how extinction probabilities change over time and depend on the initial abundance and trait distribution of the population. We illustrate how to numerically implement these methods using data from the short-lived desert shrub species *Cryptantha flava*, and provide a more general discussion of how to implement these methods to other data sets including those involving fluctuating environmental conditions.

(4) As most IPM studies have the necessary data to parameterize individual-based IPMs, these methods provide a computationally efficient means to explore how continuously structured populations differing in their evolutionary history and environmental context may differ in their vulnerability to extinction or ability to colonize new habitats.

INTRODUCTION

Computations of extinction probabilities or likelihoods of establishment success lie on opposing sides of a theoretician's coin and have been used to address theoretical and practical issues in conservation biology, restoration ecology, biological invasions and population genetics. Risks of extinction or establishment failure stem from populations consisting of a finite number of individuals, each of which faces a non-zero risk of mortality on any given day. These extinction risks are shaped, in part, by the size and composition of a population whose individuals may differ in age, size, geographical location, or other important characteristics influencing demography. When population structure is finite-dimensional (e.g. a finite number of age classes, stages, geographical locations), multi-type branching processes can model these extinction risks and, thereby, serve as the stochastic, finite-population counterpart of matrix models [Harris, 1963, Athreya and Ney, 2004, Caswell, 2001, Haccou et al., 2005]. These stochastic models have been used successfully to address a

diversity of questions concerning fixation probabilities of beneficial alleles [Patwa and Wahl, 2008], evolutionary emergence of pathogens [Antia et al., 2003, Park et al., 2013], extinction risk of small populations [Boyce, 1992, Gosselin and Lebreton, 2000, Fujiwara and Caswell, 2001, Erickson et al., in press], and establishment success in heterogeneous environments [Haccou and Iwasa, 1996, Haccou and Vatunin, 2003, Schreiber and Lloyd-Smith, 2009].

To parameterize matrix models or multi-type branching processes, individuals must be discretely categorized into a finite number of types. However, when collecting demographic data, researchers commonly measure continuous traits (e.g. mass, length, geographical location) about individuals and use continuum-based statistics to approximate "fine-grained" discrete-traits. Integral projection models (IPMs) allow one to account for this continuous population structure [Easterling et al., 2000]. These IPMs can be viewed as infinite-dimensional matrix models and can be numerically approximated by finite-dimensional matrix models. Consequently, many of the standard demographic concepts and methods for matrix models (e.g. stable state distributions, reproductive values, life table response experiments, sensitivity analysis) exist for IPMs [Easterling et al., 2000, Ellner and Rees, 2006, 2007, Rees and Ellner, 2009, Coulson, 2012, Ellner and Schreiber, 2012, Metcalf et al., 2013, Rees et al., 2014, Merow et al., 2014].

Here, we describe individual-based counterparts of IPMs using continuous-state branching processes [Harris, 1963]. For these finite population, stochastic models, we present an analytical method for computing extinction probabilities. As these methods are easily implemented numerically, they circumvent the need to use individual-based simulations and allow one to efficiently study how extinction probabilities or establishment failure depend on continuous as well as discrete population structure. We illustrate the application of these methods with an individual-based IPM of the short-lived desert shrub *Cryptantha flava* from Utah, USA [Salguero-Gómez et al., 2012].

The General Models and Methods

The Individual Based IPM. We consider an individual-based model where the set of all possible individual states (e.g. age, size, geographical location, etc.) lies in a compact metric space X. For a standard size-structured IPM, $X = [a, b]$ corresponds to the range of sizes measured in the field where a is the minimal size and b is the maximal size. For models with a mixture of age and size structure, X could be given by $\{1, \ldots, T\} \times [a, b]$ where T corresponds to the maximal age of an individual.

Following Harris [1963], we consider finite populations in which the state of the population at any point in time is characterized by the different states $(x_1, x_2, \ldots, x_k)$ of individuals within the population and the number of individuals in each state $(n_1, n_2, \ldots, n_k)$. Specifically, if there n_1 individuals in state x_1, n_2 individuals in state x_2, ..., n_k individuals in state x_k, then the state of the population is given by $\mathbf{s} = (n_1, n_2, \ldots, n_k; x_1, x_2, \ldots, x_k)$. The set of all possible population states is

$$\mathcal{S} = \{(\mathbf{n}; \mathbf{x}) = (n_1, \ldots, n_k; x_1 \ldots, x_k) : k, n_i \in \mathbb{N}, x_i \in X\} \cup \{0\}$$

where $\mathbb{N} = \{1, 2, 3, \ldots\}$ denotes the natural numbers and 0 is the extinction state corresponding to no individuals in the population.

Let $\mathbf{s}_t = (\mathbf{n}; \mathbf{x}) \in \mathcal{S}$ be the population state at time t. The dynamics of $\mathbf{s}_t$ are determined by a set of probabilistic rules that determine the contribution of each individual in the population to the population in next time step $t + 1$. These "contributions" may correspond to an individual surviving and changing state (e.g. growing in size, getting older, dispersing to another geographical location), or an individual having offspring. Consistent with standard branching process theory, each individual updates independently of all other individuals in the population.

The update rule for an individual in state x is given by a probability measure $m(x, d\mathbf{s})$ on the state space $\mathcal{S}$. Specially, the probability an individual in state x contributes $\mathbf{s}$ individuals to the

95 population in the next time step where $\mathbf{s}$ lies in a subset $A \subset \mathcal{S}$ equals

$$\mathbb{P}[\mathbf{s}_1 \in A | \mathbf{s}_0 = (1; x)] = \int_A m(x, d\mathbf{s})$$

96 where the left hand side reads "the probability the population state lies in A at time 1 after

97 initially having only one individual in state x at time 0." If the population state is currently

98 $\mathbf{s}_t = (n_1, \ldots, n_k; x_1, \ldots, x_k)$, then the state $\mathbf{s}_{t+1}$ is determined as follows:

    99 (1) for each of the n_1 individuals in state x_1, randomly and independently choose the number

    100   of replacement individuals from distribution $m(x_1, d\,\mathbf{s})$,

    101 (2) repeat step (1) for the states $x_2, \ldots, x_k$, and

    102 (3) determine the new population state $\mathbf{s}_{t+1}$ by identifying the states of all individuals and

    103   counting the total number of individuals in each of these states.

104 This iterative algorithm can be used to create individual based simulations of the individual based

105 IPM. As with any branching process, stochastic realizations of this process, with probability one,

106 either go to extinction in finite time, or the population abundance grows without bound. This latter

107 event is typically interpreted as a population becoming established or persisting.

108 **Probability-generating functionals and extinction probabilities.** We can characterize the

109 probabilistic state of the system using probability-generating functionals Ψ (pgfs). Unlike moment-

110 generating functionals as used by Harris [1963], the pgfs allow us to directly compute how extinction

111 probabilities change in time as well as compute asymptotic extinction probabilities.

112   To define the pgf Ψ, we introduce the following notation: given a continuous function $h : X \to \mathbb{R}$

113 and $\mathbf{s} \in \mathcal{S}$ (individuals' states x_i are real numbers, and the state s is within all possible states S),

114 let

$$h^{\mathbf{s}} = \begin{cases} \prod_{i=1}^k h(x_i)^{n_i} & \text{if } \mathbf{s} = (n_1, \ldots, n_k; x_1, \ldots, x_k) \\ 1 & \text{if } \mathbf{s} = 0. \end{cases}$$

If the function $h(x)$ is defined such that it corresponds to the probability that an individual of size x dies and has no offspring over the next year, then $h^{\mathbf{s}}$ corresponds to the probability that a population in state $\mathbf{s}$ goes extinct in the next year. The corresponding probability-generating functional is defined by

$$\Psi(h)(x) = \int h^{\mathbf{s}}\, m(x, d\mathbf{s}). \tag{1}$$

Given an individual in state x, $\Psi(h)(x)$ corresponds to the expected value of $h^{\mathbf{s}}$. This requires integrating over all possible states s that the population may take on in the next time step due to the different possible contributions of the individual in state x.

Two facts allow us to use Ψ to calculate extinction probabilities. First, when h_0 is the zero function (i.e. $h_0(x) = 0$ for all x), $\Psi(h_0)(x)$ is the probability the population goes extinct in one time step, given that initially it consisted of one individual in state x. That is,

$$\Psi(h_0)(x) = \int 1 \times m(x, d\mathbf{s}) = \mathbb{P}[\mathbf{s}_1 = 0 | \mathbf{s}_0 = (1; x)] \tag{2}$$

Secondly, we clam that the t-fold composition of Φ to h_0,

$$\Psi^t(h_0) = \underbrace{\Psi(\Psi(\ldots \Psi(h_0)\ldots))}_{t \text{ times}}$$

is the probability of extinction by time t given that the population initially consisted of one individual in state x:

$$\Psi^t(h_0)(x) = \mathbb{P}[\mathbf{s}_t = 0 | \mathbf{s}_0 = (1; x)] =: E_t(x) \tag{3}$$

128   To verify this claim, we argue by induction. Equation 2 implies that equation (3) holds for $t = 1$.

129   Now suppose that equation (3) holds at time t; we will show it holds at time $t + 1$. On the event

130   that $\mathbf{s}_1 = (n_1, \ldots, n_k; x_1, \ldots, x_k)$ is the population state at time 1, extinction occurs by time $t + 1$

131   only if each of the lineages of the $n_1 + \cdots + n_k$ individuals go extinct in the next t time steps. As

132   the fates of these lineages are independent of one another, it follows that

$$\mathbb{P}[s_{t+1} = 0 | \mathbf{s}_0 = (1, x), \mathbf{s}_1 = (n_1, \ldots, n_k; x_1, \ldots, x_k)] = \prod_{i=1}^{k} \mathbb{P}[\mathbf{s}_{t+1} = 0 | \mathbf{s}_1 = (1, x_i)]^{n_i}$$

$$= \prod_{i=1}^{k} \left(\Psi^t(h_0)(x_i) \right)^{n_i} = \Psi^t(h_0)^{\mathbf{s}_1}$$

133   where the second equality follows from our inductive hypothesis. By the law of total probability

$$E_{t+1}(x) = \mathbb{P}[s_{t+1} = 0 | \mathbf{s}_0 = (1; x)] = \int [\Psi^t(h_0)]^{\mathbf{s}} \, m(d\mathbf{s}, x)$$

134   which, by definition, equals $\Psi^{t+1}(h_0)(x)$ as claimed.

135   Equation (3) can be used to compute extinction probabilities iteratively. Furthermore, as indi-

136   viduals update independent of one another, the probability of the population going extinct by time

137   t for any initial condition $\mathbf{s} = (n_1, \ldots, n_k; x_1, \ldots, x_k)$ equals

$$(4) \qquad\qquad\qquad \mathbb{P}[\mathbf{s}_t = 0 | \mathbf{s}_0 = \mathbf{s}] = E_t^{\mathbf{s}}.$$

138   These analytic expressions allow us to efficiently compute extinction probabilities by constructing

139   a numerical approximation of the pgf Ψ and iterating it with an initial condition of a zero vector

140   which corresponds to the numerical approximation of the zero function.

141   As $E_0(x) \le E_1(x), \le E_2(x) \ldots$ for any $x \in X$ and $E_t(x) \le 1$ for all t, there is a well defined limit

142   corresponding to the probability of eventual extinction:

$$E_\infty(x) := \lim_{t \to \infty} E_t(x).$$

143   Using moment-generating functionals with suitable technical hypotheses, Harris [1963] showed that

144   $E_\infty(x) < 1$ for all x if the dominant eigenvalue of the mean-field IPM is greater than one, and

145   $E_\infty(x) = 1$ for all x otherwise.

146   AN ILLUSTRATION WITH A SHORT-LIVED PERENNIAL

147   We illustrate these general methods using an individual-based IPM for the yellow-flowered peren-

148   nial plant, Plateau Yellow Miner's Candle (*Cryptantha flava*), of the borage family (*Boraginaceae*).

149   For populations growing along the Colorado Plateau, USA, Salguero-Gómez et al. [2012] developed

150   an IPM using data collected from 2004 to 2010. Here, we use a subset of this data available in an R

151   package, IPMpack [Metcalf et al., 2014]. All code to for this example is archived at Zenodo [Schreiber

152   and Ross, 2015].

```
## Error in library(IPMpack):   there is no package called 'IPMpack'
```

```
## Error in eval(expr, envir, enclos):  object 'cryptantha' not found

## Error in eval(expr, envir, enclos):  object 'cryptantha' not found

## Error in seq(alpha, beta, length = 100):  object 'alpha' not found

## Error in plot(cryptantha$size, jitter(cryptantha$surv, factor = 0.1), :  object
'cryptantha' not found

## Error in lines(xs, Survival(xs), lwd = 4, col = "blue"):  object 'xs' not found

## Error in image(xs, xs, t(outer(xs, xs, Growth)), xlab = expression(size[t]), :  error
in evaluating the argument 'x' in selecting a method for function 'image':  Error:
object 'xs' not found

## Error in points(cryptantha$size, cryptantha$sizeNext, pch = 21, bg = rgb(0.2, :
object 'cryptantha' not found

## Error in plot(cryptantha$size, jitter(cryptantha$fec0, factor = 0.1), :  object
'cryptantha' not found

## Error in lines(xs, Flowering(xs), lwd = 4, col = "blue"):  object 'xs' not found

## Error in plot(cryptantha$size[cryptantha$fec0 == 1], cryptantha$fec1[cryptantha$fec0
== :  object 'cryptantha' not found

## Error in lines(xs, Fecundity_flowered(xs), lwd = 2):  object 'xs' not found
```

153   In the model, the state x of an individual is the size, which equals the total number of vegetative
154   and flowering rosettes. If $N_t(x)$ denotes the density of individuals of size x in year t, then Salguero-
155   Gómez et al. [2012] used an IPM of the form

$$N_{t+1}(x) = \int_a^b [s(y)G(y,x) + e(x)p(y)f(y)N_t(y)] \, dy$$

156   where $s(y)$ is the probability of surviving to the next year for individuals of size y, $G(y,x)\,dy$ is the
157   infinitesimal probability that a surviving individual of size y is size x in the next year, $p(y)$ is the

158   probability that an individual of size y flowers, $f(y)$ is the mean number of offspring produced by
159   an individual of size y, and $e(x)$ is the infinitesimal probability that an offspring is of size x at the
160   time of the annual census.

161   Following Salguero-Gómez et al. [2012], we use generalized linear models (GLMs) for most of the
162   functional forms of the demographic kernels. Computations were performed using the base GLM
163   function in R [R Core Team, 2015]. We used logistic regression (i.e. a GLM with the binomial
164   family) for determining $s(y)$ and $p(y)$, and a GLM with the Poisson family for modelling $f(y)$.
165   For the growth kernel, linear regression determined the expected size of an individual in the next
166   year and the actual size was assumed to be normally distributed about this mean. The variance
167   of this normal distribution, for simplicity, was assumed to be independent of the current size of
168   an individual. We modeled $e(y)$ with a gamma distribution fit to the the empirical distribution of
169   germinants. Figure **??** shows the data and fits for $s(y)$, $G(y, x)$, $p(x)$ and $f(x)$.

170   The kernels $s(x)$, $G(y, x)$, $p(x)$ and $e(x)$ provide us with all the information that the individual-
171   based IPM requires for probabilistic updating individuals for survival, growth, flowering, reproduc-
172   tion, and size of germinating individuals at first census. The fecundity kernel $f(y)$, however, only
173   specifies the mean number of offspring produced, but for an individual-based IPM, we need the
174   distribution of the number of offspring produced by an individual. Fortunately, this information is
175   built into the structure of the GLMs due to the assumptions in our model choice. As the fecundity
176   data was modeled using a Poisson family for the GLM, the mean number $f(x)$ of offspring also
177   specifies the distribution. More generally, one might use multi-parameter distributions such as a
178   zero-inflated Poisson or a negative binomial, in which case parameters in addition to the mean are
179   needed to specify the distribution of offspring number.

180   **Deriving the pgf Ψ.** To define Ψ, we observe that the contributions of an individual of size x to
181   the population in the next time step involves the sum of two independent random variables: the
182   contribution due to survival and growth and the contribution due to reproduction. We will identify

183 two pgfs, Ψ_g and Ψ_f, for each of these processes separately. Then, we make use of a fundamental

184 property of pgfs

> **Fundamental Property 1**: The pgf for a sum of independent random variables is the product of
> 185
> the pgfs of these random variables.

186 to get that

$$\Psi = \Psi_g \times \Psi_f.$$

187 To write down each of these pgfs, we make use of another fundamental property of pgfs:

> **Fundamental Property 2**: The pgf for a sum of N independent, identically distributed random
> 188
> variables X_i is the composition of the pgf for N with the pgf for the X_i.

189 For survival and growth, $\Psi_g(h)(x)$ corresponds to integrating $h^{\mathbf{s}}$ over all possible contributions

190 $\mathbf{s}$ from an individual of size x surviving and growing. These contributions are of two types: $\mathbf{s} = 0$

191 when the individual dies, and $\mathbf{s} = (1; y)$ when the individual survives and grows to size y. The

192 first event occurs with probability $1 - s(x)$ and the infinitesimal probability of the second event is

193 $s(x)G(y, x)dy$. As $h^{\mathbf{s}} = 1$ when $\mathbf{s} = 0$ and $h^{\mathbf{s}} = h(y)$ when $\mathbf{s} = (1; y)$, integrating over all possible

194 contributions due to survival and growth yields

$$\Psi_g(h)(x) = (1 - s(x)) \times 1 + s(x) \int h(y)G(y, x)dy.$$

195 For fecundity, $\Psi_f(h)(x)$ is given by integrating $h^{\mathbf{s}}$ over all possible states $\mathbf{s}$ corresponding to the

196 offspring produced by an individual of size x. To write this down, we begin by conditioning on the

197 event that an individual of size x flowers. On this event, the individual produces a Poisson number

198 N of offspring with mean $f(x)$. The pgf for N is given by $\phi(x, \xi) = \exp(-f(x)(\xi - 1))$ where ξ is

199 a dummy variable. The size of each of these offspring is drawn interdependently from the common

200 offspring distribution $e(y)dy$. Hence, the contribution of a flowering individual of size x is the sum

201 of N independent random variables with distribution $e(y)dy$. By **Fundamental Property 2** of

202   pgfs, we need to take the composition of the pgf ϕ for N with the pgf for a single offspring, namely

$$\Psi_e(h)(x) = \int h(y)e(y)dy.$$

203   Thus, we get the pgf associated with a flowering individual is

$$\Psi_{flowering}(h)(x) = \phi(x, \int h(y)e(y)dy)$$

204   To get the pgf for flowering and non-flowering contributions to fecundity, we observe that fecundity
205   contributions of an individual of size x is given by the sum of a Bernoulli number of flowering
206   individuals where the probability of success is $p(x)$. By the **Fundamental Property 2**, we need
207   to compose the pgf of a Bernoulli, which is $\theta(\xi) = 1 - p(x) + p(x)\xi$ where ξ is the dummy variable,
208   with the pgf $\Psi_{flowering}$ of a flowering individual. This composition yields

$$\Psi_f(h)(x) = \theta(\Psi_{flowering}(h)(x)) = 1 - p(x) + p(x)\phi(x, \int h(y)e(y)dy).$$

209   Muliplying Ψ_g and Ψ_f, we get the desired pgf Ψ:

$$\Psi(h)(x) = \left((1 - s(x)) \times 1 + s(x) \int h(y)G(y, x)dy \right) \left(1 - p(x) + p(x)\phi(x, \int h(y)e(y)dy) \right).$$

210   **Numerically Implementing and Using the pgf Ψ.** To approximate Ψ numerically, we discretize
211   a larger size interval $[a, 2b]$ using $n = 500$ equal sized intervals of width $\Delta x = (2b - a)/n$. We used
212   the larger size interval of $[a, 2b]$ to minimize the effects of eviction (see below). We created a vector
213   $\vec{x}$ corresponding to the midpoints of these intervals. Using this vector we discretized the survival
214   function as a vector $\vec{s} = s(\vec{x})$, the growth kernel as a matrix using the outer product G of the growth
215   kernel g with $\vec{x}$, the probability of flowering function as a vector $\vec{p} = p(\vec{x})$, the fecundity function
216   as a vector $\vec{f} = f(\vec{x})$, and the offspring size distribution as a vector $\vec{e} = e(\vec{x})$.
217   For the previously described methods to work, it is critical that column sums for G and the sum
218   of $\vec{e}$ equal one. For most IPMs, this will not occur automatically due to individuals being evicted

219 from the size interval $[a, 2b]$. There are a variety of ways to handle this issue [Williams et al., 2012].

220 As the offspring size vector nearly summed to one, we simply re-normalized it so that it summed

221 to one. For the growth matrix G, we treated eviction as mortality. To do this, we took one minus

222 the column sums of G, subtracted these sums from the survival vector, and then normalized the

223 column sums of G so that they added to one. When taking the product of survival and growth, the

224 resulting mean-field IPM is unaffected by these changes.

225   Using these discretized demographic components and the pgf ϕ for fecundity, we get the discretized

226 pgf Ψ, given by

$$\Psi_{discrete}(\vec{h}) = (1 - \vec{s} + \vec{s} \circ (G^T \vec{h})) \circ (1 - \vec{p} + \vec{p} \circ \phi(\vec{x}, \vec{e}^T \vec{h}))$$

227 where $\vec{h}$ corresponds to a discretized function i.e. a vector of length n, ^{T} denotes the transpose of

228 a matrix or vector, and $\circ$ denotes element by element multiplication.

```
## Error in seq(alpha, 2 * beta, length = n):  object 'alpha' not found
## Error in eval(expr, envir, enclos):  object 'xs' not found
## Error in eval(expr, envir, enclos):  could not find function "Survival"
## Error in dgamma(xs, shape = coef(recruit.size)[1], rate = coef(recruit.size)[2]):
object 'xs' not found
## Error in eval(expr, envir, enclos):  object 'e.vec' not found
## Error in eval(expr, envir, enclos):  could not find function "Flowering"
## Error in outer(xs, xs, Growth):  object 'xs' not found
## Error in colSums(g.mat):  error in evaluating the argument 'x' in selecting a method
for function 'colSums':  Error:  object 'g.mat' not found
## Error in eval(expr, envir, enclos):  object 'g.mat' not found
## Error in eval(expr, envir, enclos):  object 'q.evict' not found
```

229   Iterating Ψ yields how extinction probabilities $E_t(x)$ vary with time for a population initiated

230   with a single individual (Figure **??**). Intuitively, this figure illustrates that the probability of

231   extinction decreases with the size of founding individual, and that extinction probabilities increase

232   over time. Furthermore, $E_t(x)$ as t increases are approaching limiting extinction probabilities $E_\infty(x)$

233   which always equals one (Figure **??**A). This stems from the fact that the dominant eigenvalue of

234   the mean-field IPM is less than one, consistent with the results of Salguero-Gómez et al. [2012].

235   By increasing seed survivorship by a factor of three, the dominant eigenvalue of the mean-field

236   IPM becomes greater than one and the asymptotic extinction probabilities become less than one

237   (Figure **??**B). We approximated these asymptotic extinction probabilities by iterating Ψ until the

238   difference between E_t and E_{t+1} were below a specified error tolerance.

```
## Error in Psi(hs[t, ]):  object 's.vec' not found
## Error in as.matrix(x):  object 'xs' not found
## Error in Psi(h):  object 's.vec' not found
## Error in lines(xs, h, type = "l", lwd = 5, xlim = c(alpha, beta), col = viridis(7)[1]):
object 'xs' not found
## Error in Psi.plus(hs[t, ]):  object 's.vec' not found
## Error in as.matrix(x):  object 'xs' not found
## Error in Psi.plus(h):  object 's.vec' not found
## Error in lines(xs, h, type = "l", lwd = 5, xlim = c(alpha, beta), col = viridis(7)[1]):
object 'xs' not found
## Error in strwidth(legend, units = "user", cex = cex, font = text.font):  plot.new
has not been called yet
```

239   To scale things up to an entire population initially in state $\mathbf{s}_0$, recall that the extinction probability

240   at time t is $E_t^{\mathbf{s}_0}$. Figure **??** illustrates how the extinction probabilities over time vary for a population

241   with initially 100 individuals of the smallest size $x = 1$ (i.e. $\mathbf{s}_0 = (1; 100)$) versus a population with

242 5 or 8 individuals of the largest size $x = 60$ typically observed in the field (i.e., $\mathbf{s}_0 = (60; 5)$ or $(60; 8)$.
243 Figure **??** suggests that, from the extinction risk perspective, about 6 or 7 larger individuals are
244 equivalent to 100 of the smallest individuals. These types of comparisons may be particularly
245 useful when trying to assess whether planting small or larger individuals are more effective for
246 establishment success.

```
## Error in Psi(hs[t, ]):  object 's.vec' not found

## Error in Psi(hs[t, ]):  object 's.vec' not found

## Error in Psi(hs[t, ]):  object 's.vec' not found

## Error in cbind(out1, out2, out3):  object 'out1' not found

## Error in as.matrix(y):  object 'out' not found

## Error in strwidth(legend, units = "user", cex = cex, font = text.font):  plot.new
has not been called yet
```

247 We also used the data to see how extinction risk over different time frames depends on the size
248 and composition of the population. Specifically, for different founding population abundances N,
249 we randomly sampled N individuals from the data and computed extinction risk of this sampled
250 population over 5 and 10 year periods (Figure 1). For each founding population abundance N, we
251 considered 500 random samples of size N. Figure 1A illustrates that, on average, log extinction risk
252 decreases with the founding population size N and is greater for the 10 year period than the 5 year
253 period. For smaller founding popualation sizes N, there is substantial overlap in the distributions of
254 extinction times for the 5 and 10 year time frames. This overlap stems from founding populations
255 of mostly large individuals being more likely to persist at least 10 years than founding populations
256 of mostly small individuals persisting at least 5 years. Consistent with this explanation, Figure 1B
257 illustrates that mean size of an individual within a founding population has a strong negative
258 correlation with (log) extinction probability, and the slope of this correlation is steeper over shorter
259 time frames than longer time frames.

260   [Figure 1 about here.]

261   RECOMMENDATIONS, EXTENSIONS, AND FUTURE CHALLENGES

262   To implement the methods presented here, there are two main steps. First, one needs to identify

263   the main demographic processes of the population, the order in which these processes occur relative

264   to the censuses used for data collection, and develop the statistical models for the each of the

265   demographic processes. Rees et al. [2014] and Merow et al. [2014] provide excellent reviews on the

266   philosophical and methodological issues associated with this step. Typically, whenever a study has

267   sufficient data for constructing the mean-field IPM, there is no need to collect any additional data

268   to build the individual-based IPM. Unlike the mean-field IPMs, though, the individual-based IPMs

269   make use of the complete distributional information associated with fecundity. As with other areas

270   of stochastic demography, the shapes, not just the means, of these distributions may have significant

271   effects on the likelihood of extinction or establishment success [Lloyd-Smith et al., 2005]. Hence,

272   it is best to examine several options (e.g. Poisson versus negative Binomial versus zero-inflated

273   distributions) to identify which distribution does a better job of describing the fecundity data. As

274   in all areas of modeling, if there is significant uncertainty about the "best" choice of the model,

275   one should perform the analyses with each of the alternative fecundity distributions to identify the

276   sensitivity of predictions of extinction risk to these alternatives.

277   The second step, the focus of this paper, involves constructing the probability-generating func-

278   tional Ψ. For the uninitiated, this step may be intimidating. However, there are a three basic

279   principles that simplify this construction. First, while this pgf Ψ takes functions to functions, one

280   should focus on writing down $\Psi(h)(x)$, which involves understanding the contributions of a single

281   individual of size x to the next census. Second, one can often can break up these contributions

282   into a sum of independent contributions, find the pgfs associated with these simpler contributions,

283 and then use the two fundamental properties of pgfs to "stitch" together Ψ. Third, the distribu-

284 tions used to describe the number of offspring produced by an individual typically involve random

285 variables (e.g. Poisson, negative binomial) for which the associated pgfs are well-known. Finally,

286 whenever in doubt, find a collaborator that you trust to help put the pieces together correctly.

287 For the individual-based models considered here, we assumed the environment remains constant

288 over time. However, IPMs have and continue to be used to study the effects of fluctuating en-

289 vironmental conditions on population demography and life history evolution [Childs et al., 2004,

290 Dahlgren and Ehrlén, 2011, Rees and Ellner, 2009]. The methods described here easily extended to

291 fluctuating environments. Specifically, if Ψ_t is the pgf of the individual-based IPM associated with

292 year t, then the probability of extinction for a population initially in state $\mathbf{s}$ is

$$\Psi_0(\Psi_1(\ldots\Psi_t(h_0)))^{\mathbf{s}}$$

293 where h_0, as before, is the zero function. Note that the composition here is in the reverse order

294 of what one does when iterating the mean-field IPM model forward in time. While we know of

295 no formal proof, in the case of a stationary environment with suitable technical assumptions, we

296 conjecture the following limit theorem holds: the asymptotic extinction probabilities are strictly

297 less than one if and only if the stochastic growth rate (aka dominant Lyapunov exponent) of the

298 mean-field IPM is positive. For multi-type branching processes, this result was proven by Tanny

299 [1981].

300 Despite this relatively straightforward extension to temporally variable environments, many chal-

301 lenges remain. From the computational perspective, finding the efficient methods to deal with

302 multi-dimensional states variables (e.g. size and location, or multi-dimensional traits) continues to

303 be a challenge, as it is for mean-field IPMs. While the analytical methods presented here cover

304 multi-dimensional state variables, their numerical implementation involves approximating mutli-

305 dimensional integrals which can be computationally expensive. From the analytical perspective,

accounting for temporal correlations in individual growth or reproductive rates (e.g. individuals that grew larger than expected in one year being more likely to grow larger than expected in the next year) or correlations among individuals are particularly important challenges as strong correlations likely have large effects on extinction risk. Finally, and perhaps most importantly, the manners in which size-structured demography may shape extinction risk for real-world population remains to be understood. One might hope that by applying these methods to the many data sets for which IPMs have been developed, as well as future data sets, might provide a computationally efficient means to explore how size-structure for populations differing in their evolutionary history and environmental context influences their vulnerability to extinction.

REFERENCES

R. Antia, R.R. Regoes, J.C. Koella, and C.T. Bergstrom. The role of evolution in the emergence of infectious diseases. *Nature*, 426:658–661, 2003.

K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004.

M.S. Boyce. Population viability analysis. *Annual review of Ecology and Systematics*, pages 481–506, 1992.

H. Caswell. *Matrix Population Models*. Sinauer, Sunderland, Massachuesetts, 2001.

D.Z. Childs, M. Rees, K.E. Rose, P.J. Grubb, and S.P. Ellner. Evolution of size–dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271:425–434, 2004.

T. Coulson. Integral projections models, their construction and use in posing hypotheses in ecology. *Oikos*, 121:1337–1350, 2012.

329 J.P. Dahlgren and J. Ehrlén. Incorporating environmental change over succession in an integral
330   projection model of population dynamics of a forest herb. *Oikos*, 120:1183–1190, 2011.

331 M.R. Easterling, S.P. Ellner, and P.M. Dixon. Size-specific sensitivity: applying a new structured
332   population model. *Ecology*, 81:694–708, 2000.

333 S. P. Ellner and S. J. Schreiber. Temporally variable dispersal and demography can accelerate the
334   spread of invading species. *Theoretical Population Biology*, 82:283–298, 2012.

335 S.P. Ellner and M. Rees. Integral projection models for species with complex demography. *American
336   Naturalist*, 167:410–428, 2006.

337 S.P. Ellner and M. Rees. Stochastic stable population growth in integral projection models: theory
338   and application. *Journal of Mathematical Biology*, 54:227–256, 2007.

339 R.A. Erickson, J.C. Eager, E.A. amd Stanton, J.A. Beston, and Thogmartin W.E. Diffendorfer,
340   J.E. Assessing local population vulnerability with branching process models. *Ecosphere*, in press.

341 M. Fujiwara and H. Caswell. Demography of the endangered north atlantic right whale. *Nature*,
342   414:537–541, 2001.

343 F. Gosselin and J. Lebreton. Potential of branching processes as a modeling tool for conservation
344   biology. In *Quantitative methods for conservation biology*, pages 199–225. Springer, 2000.

345 P. Haccou and Y. Iwasa. Establishment probability in fluctuating environments: a branching process
346   model. *Theoretical Population Biology*, 50:254–280, 1996.

347 P. Haccou and V. Vatunin. Establishment success and extinction risk in autocorrelated environ-
348   ments. *Theoretical Population Biology*, 64:303–314, 2003.

349 P. Haccou, P. Jagers, and V.A. Vatutin. *Branching processes: variation, growth, and extinction of
350   populations*. Cambridge University Press, 2005.

351 T.E. Harris. *The theory of branching processes*. Springer-Verlag, 1963.

352 J. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the impact of
353   individual variation on disease emergence. *Nature*, pages 355–359, 2005.

C. Merow, J.P. Dahlgren, C.J.E. Metcalf, D.Z. Childs, M.E.K. Evans, E. Jongejans, S. Record, M. Rees, R. Salguero-Gómez, and S.M. McMahon. Advancing population ecology with integral projection models: a practical guide. *Methods in Ecology and Evolution*, 5:99–110, 2014.

C.J.E. Metcalf, S.M. McMahon, R. Salguero-Gómez, and E. Jongejans. Ipmpack: an r package for integral projection models. *Methods in Ecology and Evolution*, 4:195–200, 2013.

C.J.E. Metcalf, S.M. McMahon, R. Salguero-Gomez, E. Jongejans, and C. Merow. *IPMpack: Builds and analyses Integral Projection Models (IPMs).*, 2014. URL `http://CRAN.R-project.org/package=IPMpack`. R package version 2.1.

M. Park, C. Loverdo, S.J. Schreiber, and J.O. Lloyd-Smith. Multiple scales of selection influence the evolutionary emergence of novel pathogens. *Philiosophical Transactions of the Royal Soceity. B*, Vol.:368, 2013.

Z Patwa and Lindi M Wahl. The fixation probability of beneficial mutations. *Journal of The Royal Society Interface*, 5:1279–1289, 2008.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `http://www.R-project.org/`.

M. Rees and S.P. Ellner. Integral projection models for populations in temporally varying environments. *Ecological Monographs*, 79:575–594, 2009.

M. Rees, D.Z. Childs, and S.P. Ellner. Building integral projection models: a user's guide. *Journal of Animal Ecology*, 83:528–545, 2014.

R. Salguero-Gómez, W. Siewert, B.B. Casper, and K. Tielbörger. A demographic approach to study effects of climate change in desert plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367:3100–3114, 2012.

S.J. Schreiber and J.O. Lloyd-Smith. Invasion dynamics in spatially heterogenous environments. *American Naturalist*, 174:490–505, 2009.

S.J. Schreiber and N. Ross. *ibipm: Code for Individual-Based Integro-Projection Models.* Zenodo. http://dx.doi.org/10.5281/zenodo.32086, 2015. URL `http://dx.doi.org/10.5281/zenodo.32086`.

D. Tanny. On multitype branching processes in a random environment. *Advances in Applied Probability*, 13:464–497, 1981.

J.L. Williams, T.E.X. Miller, and S.P. Ellner. Avoiding unintentional eviction from integral projection models. *Ecology*, 93:2008–2014, 2012.

DEPARTMENT OF EVOLUTION AND ECOLOGY, ONE SHIELDS AVENUE, UNIVERSITY OF CALIFORNIA, DAVIS, CALIFORNIA 95616

E-mail address: `sschreiber@ucdavis.edu`

ECOHEALTH ALLIANCE, 460 WEST 34TH STREET, 17TH FLOOR, NEW YORK, NY 10001
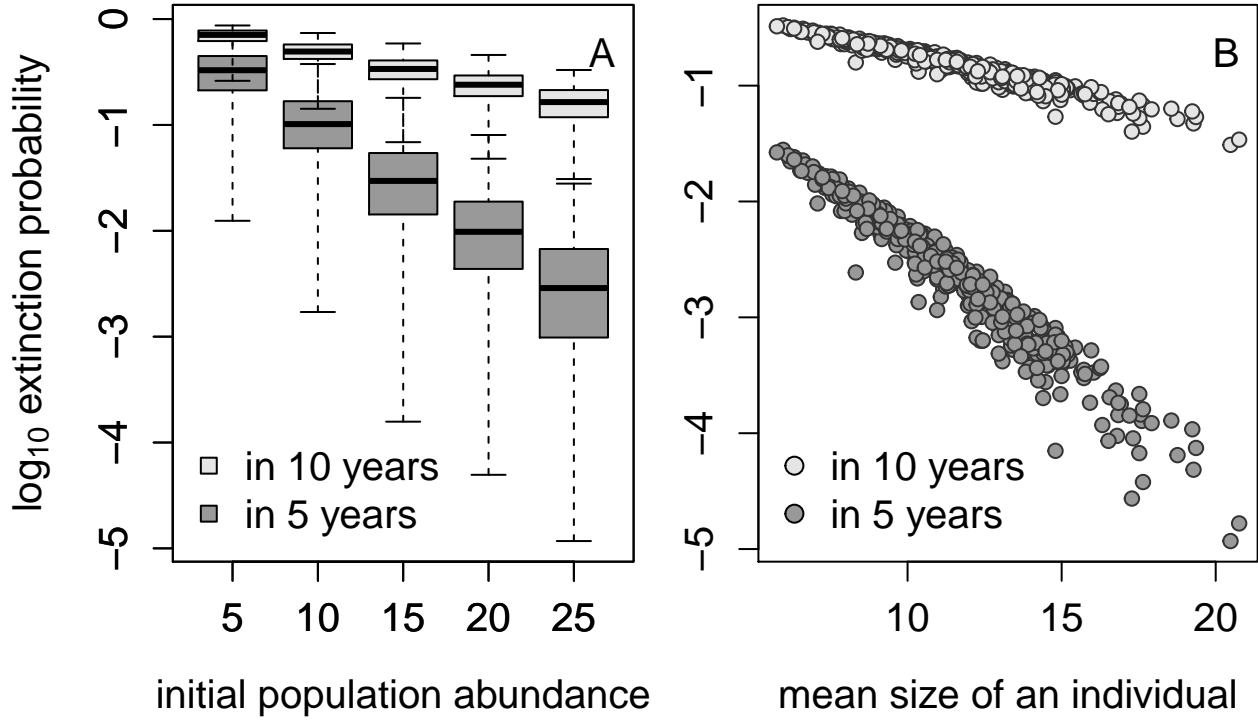
E-mail address: `ross@ecohealthalliance.org`

FIGURE 1. Extinction probabilities for founding populations of different abundance. For each founding population abundance N, 500 samples consisting of N randomly choosen individuals from the data set were used to create a founding population of N individuals. Extinction probablities by year 5 and year 10 were calculated for each of these sample populations. In A, extinction probability is plotted as a box plot for different N values. In B, extinction probability is plotted against the mean size of an individual for a population abundance $N = 25$.