Seed banks in *Clarkia xantiana*

Gregor-Fausto Siegmund and Monica Geber

Last updated: February 24, 2020

Introduction

Seed banks have important consequences for population persistence by acting as a buffer against environmental change and population stochasticity (Eager et al. (2014); Paniw et al. (2017)), increasing effective population size (Nunney (2002); Waples (2006)), and genetic diversity (Mccue and Holtsford (1998)). The presence of a seed bank can also affect the outcome of evolution (Heinrich et al. (2018); Ritland (1983)). Theory thus suggests that seed banks have ecological and evolutionary consequences (Evans and Dennehy (2005)).

A previous study suggests that the soil seed bank is important for population dynamics in *Clarkia xantiana* (Eckhart et al. (2011)). A separate set of seed burial experiments suggests that seeds of *C. xantiana* can remain viable in the soil for at least 10 years (Moeller personal communication). In the study of *C. xantiana* population dynamics that showed a decline of long-term stochastic population growth rate from west to east across the range, Eckhart et al. 2011 inferred a decrease in survival through winter (s1) and an increase in germination rate (g1) of first-year seeds from west to east.

Population vital rates are known to vary across *C. xantiana*'s geographic range. Population growth rates determine species abundance and distribution, and are ultimately what limit persistence beyond range edges. Geographic patterns to vital rates have so far been studied to help understand the demography of geography. Seed banks are a strategy that annual plants may use to buffer against environmental variation and may be part of population persistence. I will begin by characterizing geographic variation in belowground vital rates. [What is the geographic pattern to variation in germination or seed survival?] [I think this question could be expanded to make clear predictions and/or address another aspect such as variation in time.]

Bet hedging should evolve to maximize the long-term geometric population growth rate (as compared to the arithmetic population growth rate). Seed banks are more likely to

be selected for in populations which experience higher levels of interannual variation in fitness. To investigate this empirical relationship, I will estimate the correlation between interannual variation in fitness and the proportion of seeds that germinate in the winter immediately following seed production. I predict that germination is negatively correlated with interannual variation in fitness.

Table 1 outlines key references that develop theory and expectations for what drives the evolution of delayed germination. The table also includes some papers that have tested the theory empirically. The main set of papers I've included are ones that look at variation in germination among Sonoran Desert annuals (papers by Venable, Gremer). Finally, the table briefly lists predictions made by different models. I think examining the following relationships would be good starting points. The correlation between variance in fitness (seeds/seedling) and germination fraction should be negative – this is true under the density-independent and -dependent model. The correlation between seed survivorship and germination fraction should be negative under both a density-independent and -dependent model but the limit as survivorship approaches 1 differs. Finally, the correlation between mean seed yield and germination fraction will be positive if fitness is density-independent but is not necessarily positive if fitness is density-dependent.

The theory developed by Cohen (1966) frames the problem in the following terms. What is the optimal germination fraction for a given level of interannual variation in fitness and seed survivorship? The models developed by Cohen make it clear that the germination fraction that maximizes long-term population growth rate is a function of the distribution of fitness (characterized by the variation in fitness), the fitness values, and the rate of seed survivorship. For a given mean fitness, increasing the variance in fitness decreases the optimal germination fraction (Appendix Figure X). Increasing seed survivorship decreases the optimal germination fraction, and the degree to which it does so depends on the probability of a 'good year'. Specifically, as the probability of a high-fitness year decreases, the optimal

2

⁵⁴ germination fraction decreases.

⁵⁵   To empirically test this theory, one approach is to determine whether the observed ger-
⁵⁶ mination fraction is correlated with the

Table 1: Table 1: Models for germination delays: references and predictions

	Density-independent fitness	Density-dependent fitness	Predictive germination	Structured model
Key theory references	Cohen (1966, 1968)	Ellner (1985a,b)	Cohen (1967)	Easterling and Ellner (2000)
Key empirical tests	Venable (2007)	Gremer and Venable (2014)	Gremer et al. (2016)	...
Mean of seed yield	increase in $\bar{Y}$ increases G^*	increase in $\bar{K}$ can increase or decrease $\hat{G}$	...	...
CV of seed yield	increasing ρ_Y decreases G^*	increasing ρ_K or ρ_C decreases $\hat{G}$	...	...
Seed survivorship	increasing s decreases G^*; limit near $s=1$ is p	increasing s decreases $\hat{G}$; limit near $s=1$ is 0	...	...

Methods

Background on study system

[59] Monica Geber and collaborators have collected 12+ years of annual estimates for demographic data on a species of annual plant: survival of seedlings to fruiting adults, fruits per adult plant, and seeds per fruit. As part of the long-term work on *Clarkia xantiana*, there 3 sources of data on the transition between seeds in fruits and seedlings: 1 observational data set and 2 experimental data sets. Here, I present analyses of 1 observational data set and 1 experimental data set.

[65] Starting in 2007, there are (1) estimates of fruits/plant and seeds/fruit that provide an estimate of seed input into a plot and (2) estimates of germinants the following year. For most plots, the number of seeds entering a plot in year t-1 is much greater than the number of seedlings emerging in a plot in year t. However, this is not uniformly true, and there is also experimental data suggesting these seeds may survive in the seed bank for at least 10 years at some locations.

[71] There are two experiments conducted at non-overlapping points in time we use to estimate transitions in the seed bank. From 2006-2010, Geber and collaborators buried seeds in bags and periodically dug them up to count seedlings and intact, viable seeds. This data estimates transitions leading to germination or survival of seeds that are 1, 2, and 3 years old. Starting in 2013, Geber and collaborators placed seeds in pots and counted seedlings. This data estimates transitions of seeds in the soil seed bank but cannot separate germination and survival in the same way as the first experiment.

[78] Here, we analyze data from an experiment that involved burying seeds in seed bags (2005-2009).

Seed bag burial experiments

To determine how seed survival and germination varied among populations of *C. xantiana*, we conducted a series of seed burial experiments. We started these seed burial experiments in three subsequent years (2005, 2006, 2007) to obtain multiple estimates for seed survival and germination.

In June-July 2005, we collected seeds at each of the 20 populations included in this study. In October 2005, we buried 30 5×5-cm nylon mesh bags at each site. Each nylon mesh bag contained 100 seeds collected at that population. In January 2006, we removed 10 of these bags and counted the number of germinated seedlings and the number of ungerminated, intact seeds in each bag. We then returned the ungerminated, intact seeds to the resealed bag and returned the bag to the field. In October 2006, we removed these bags and counted the number of ungerminated, intact seeds. We collected the following data:

- n_{ijt} = observed count of seeds in the seed bags at the start of the experiment in October in the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\text{intact}ijt}$ = observed count of ungerminated, intact seeds in the seed bags in January in the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\text{germ}ijt}$ = observed count of germinated seedlings in the seed bags in January in the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\text{total}ijt}$ = observed count of ungerminated, intact seeds plus germinated seedlings in the seed bags in January in the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\text{surv}ijt}$ = observed count of ungerminated, intact seeds in the seed bags in October in the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

Viability trials

In the lab, we conducted germination trials and viability assays on subsets of the seeds from each bag to estimate the viability of the ungerminated, intact seeds. First, we placed up to 15 seeds from each bag on to moist filter paper in a disposable cup and observed germination over 10 days; we counted and removed germinants every 2 days.

After 10 days, all remaining ungerminated seeds (up to a total of 10 seeds) were sliced in half and individually placed into the wells of 96-well plates filled with a solution of tetrazolium chloride, which stains viable tissue red (ref). [this was from in Eckhart et al. 2011 - but not all ungerminated seeds were tested; most were] We covered the plates with foil. Each 96-well plate contained seed from at least one bag per site of a given seed-age class. Two or three tests of up to 15 seeds each were conducted for each bag. We checked and counted for viable seeds every 2 days for 10 days.

We collected the following data:

- $n_{\mathrm{germ}ijt}$ = observed count of seeds at the start of the $X^t h$ germination trial for the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\mathrm{germ}ijt}$ = observed count of germinated seedlings in the $X^t h$ germination trial for the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $n_{\mathrm{viab}ijt}$ = observed count of seeds at the start of the $X^t h$ viability trial for the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

- $y_{\mathrm{viab}ijt}$ = observed count of viable seedlings in the $X^t h$ viability trial for the ith bag, from the jth site, in the tth year, assumed to be measured perfectly

Parameter estimates for belowground transitions

Overall, the aim of our model was to combine the estimates of seed survival and germination that we obtained from the seed burial experiment in the field with the viability data from the lab trials. We seek to estimate (1) seed survival for different periods of the year or as a monthly rate, (2) germination of 0-, 1- and 2-year old seeds, and (3) viability of intact seeds unearthed in October.

Figure 1 illustrates the transitions in the first year the seed bags are buried. There are two boxes: one for the seed bag experiment and one for the viability trials. In the seed bag experiment, I split January into two steps, one for just before germination and one for just after. Solid arrows represent transitions and are labeled with corresponding vital rates. In the models, I have adopted $s_1 = \phi$, $g_1 = \gamma$, $s_2 = \rho$, and $v_1 = \upsilon$.

The probability that seeds from the start of the experiment remain intact in January is represented as s_1. In January, all seeds are intact (this includes viable and non-viable seeds). I assume that there is no decay during germination (i.e., seed loss does not instantaneously in January). The number of intact seeds before germination is equal to the number of seeds and seedlings after germination. At this point, the seeds transition into one of four possible states. Intact and viable seeds may have (1) germinated or (2) not germinated and remain dormant. All (3) other intact seeds are non-viable because (4) seeds that were not viable could not have germinated.

I represent two transitions between pre-germination seeds in January and post-germination seeds and seedlings in January. The first is for seeds that are viable and germinate; these become seedlings. The second is for seeds that do not germinate; these remain seeds and include both viable and non-viable seeds (the sum of $(1 - g_1)v_1^{\frac{1}{3}}$ and $(1 - g_1)(1 - v_1^{\frac{1}{3}})$. For the purposes of parameter estimation, I only represent the number of seedlings—viability is estimated separately.

146 I need to make some assumptions in order to incorporate the loss of viability into the

147 model. I assume that viability is lost at a constant rate, and that germination removes some

148 number of seeds from the pool of viable seeds but does not change the rate of decay. Some

149 fraction of the total seeds in January pre-germination is viable $(v_1^{\frac{1}{3}})$ and some of those viable

150 seeds germinate. I include viability in estimates of germination rate so as to not overestimate

151 the true germination rate. The number of seeds that remain intact are those that do not

152 germinate $(1 - g_1)$, which includes both viable $(v_1^{\frac{1}{3}})$ and non-viable $(1 - v_1^{\frac{1}{3}})$ seeds. Seeds

153 that germinate must be viable.

154 Here, I use viability in our germination estimates. For the full life cycle, this models the

155 transition probabilities for intact seeds and only incorporates viability in the germination

156 transition.

Seed survival and germination model

157 Overall, the aim of the model is to combine the estimates of seed survival and germination

158 that we obtained from the seed burial experiment in the field with the viability data from

159 the lab trials.

Model for viability trial data

160 Each bag i from site j in year k had n trials. The problem is most bags only had 2 trials so

161 it's difficult to estimate a variance. I want to estimate a bag-specific viability because that

162 is what I would use in the seed bag survival and germination model to put a bag-specific

163 viability rather than a site-specific viability. I fit the following model for trial h for bag i at

164 site j in year k:

For the germination trials, I estimated a probability of germination for each bag, θ^{g}.

Because the goal of the lab experiments was to provide an estimate of viability in each bag,
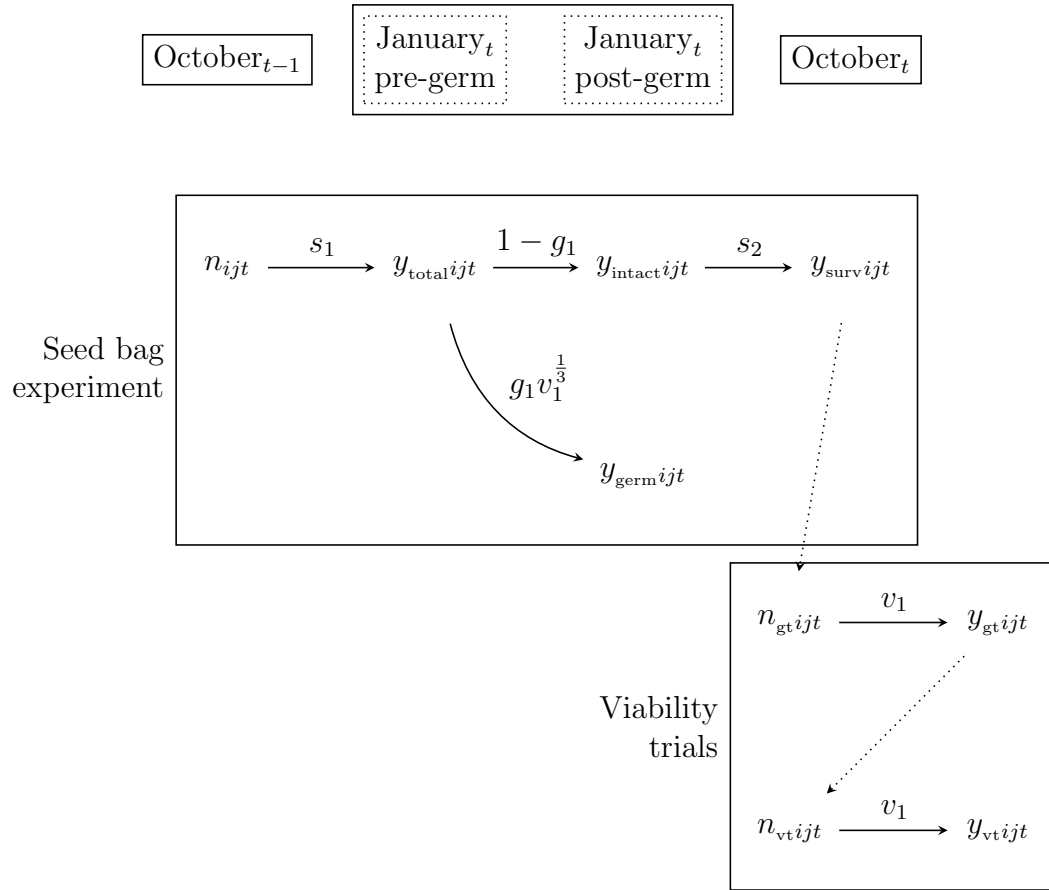
Figure 1: Diagram of data from the seed bag experiments and viability trials.

I was not interested in modeling the viability in each bag as coming from a population of viabilities at the population level. I provided each parameter with an uninformative beta (1,1) prior. I did not apply hyperpriors because (see above).

$$[\boldsymbol{\theta^g}|\boldsymbol{n^g}, \boldsymbol{y^g}] \propto$$
$$\prod_{j=1}^{J}\prod_{i=1}^{I}\prod_{h=1}^{H}\prod_{k=1}^{K}\text{binomial}(y_{hijk}^g|n_{hijk}^g, \theta_{ijk}^g)\text{beta}(\theta_{ijk}^g|1, 1) \tag{1}$$

165   I started considering applying the probabilities from the conditional probability tree.

166 This might make sense but then I realize that I've gone down this rabbit hole before. Not

167 all of the seeds that didn't germinate were tested. But I'm not sure how else to control for

168 that besides....

169   Where P(V) = P(V—Gc)*P(Gc) + P(V—G)*P(G)

170   If there are missing data, it might make sense to estimate a per-site/year viability from

171 hyperpriors. In other words, we would a distribution for theta defined by hyperpriors set

172 at the population level. Each population would get its own set of hyperpriors. Then when

173 we are looking to get the viability of a given bag in a given year, if the data are missing we

174 would draw from the hyperpriors for a site mean rather than propagate NAs.

175   The dataset has NAs in the viability stain column. These are there because the viability

176 trial started with 0 seeds. Viability trials started with 0 seeds when all/most of the seeds in

177 the germination trials germinated. The estimated probability for the (0,NA) set is 0.5; this

178 is because the result provides no new information. One approach might be to throw these

179 rows out entirely or to just take the germination probablity as the true probability. However,

180 because the probabliity of being viable if not germinating is multiplied by the probablility

181 of not germinating, the effect is quite small when most of the seeds germinate.

182   A figure that would help is one that plots the predicted vs. observed proportions for

183 the germination and viability experiments for one site. Maybe start with partial pool with

11

184 direct parameterization of theta via a beta prior on each bag. The goal of the experiment is

185 bag-specific viability. Reasons to include site would be to

Model for seed burial experiment data

To estimate the survival and germination of seeds in each population, I fit the following model for bag i at site j in year k.

$$[\boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\rho}, \boldsymbol{\alpha} | \boldsymbol{n}, \boldsymbol{n}_{\text{viab}}, \boldsymbol{y}_{\text{total}}, \boldsymbol{y}_{\text{germ}}, \boldsymbol{y}_{\text{surv}}, \boldsymbol{y}_{\text{viab}}] \propto$$

$$\prod_{j=1}^{J} \prod_{i=1}^{I} \text{binomial}(y_{ij}^{\text{tot}} | n_{ij}, f_1(\alpha_{0,j}^{s1}, \beta_j^{s1}))$$

$$\times \text{binomial}(y_{ij}^{\text{germ}} | y_{ij}^{\text{tot}}, f_2(\gamma_j, \alpha_{ijk}))$$

$$\times \text{binomial}(y_{ij}^{\text{surv}} | y_{ij}^{\text{tot}} - y_{ij}^{\text{germ}}, f_3(\alpha_{0,j}^{s2}, \beta_j^{s2}))$$

$$\times \text{binomial}(y_{ij}^{\text{viab}} | n_{ij}^{\text{viab}}, \text{logit}^{-1}(\alpha_{ijk})) \tag{2}$$

$$\times \text{normal}(\alpha_{0,j}^{s1} | 0, 10)\text{normal}(\alpha_{0,j}^{g1} | 0, 10)\text{normal}(\alpha_{0,j}^{s2} | 0, 10)$$

$$\times \text{normal}(\beta_j^{s1} | 0, \sigma_j^{s1})\text{normal}(\beta_j^{g1} | 0, \sigma_j^{g1})\text{normal}(\beta_j^{s2} | 0, \sigma_j^{s2})$$

$$\times \text{uniform}(\sigma_j^{s1} | 0, 100)\text{uniform}(\sigma_j^{g1} | 0, 100)\text{uniform}(\sigma_j^{s2} | 0, 100)$$

$$\times \text{normal}(\alpha_{ijk} | \mu_{jk}, \sigma_{jk})$$

$$\times \text{normal}(\mu_{jk} | 0, 100)\text{uniform}(\sigma_{jk} | 0, 100).$$

where

$$f_1(\alpha_{0,j}^{s1}, \beta_{jk}^{s1})) = \phi_{jk} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1})$$

$$f_2(\alpha_{0,j}^{g1}, \beta_{jk}^{g1}, \alpha_{ijk})) = \gamma_{jk} \times v_{ijk}^{\frac{1}{3}} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1}) \times (\text{logit}^{-1}(\alpha_{ijk}))^{\frac{1}{3}} \tag{3}$$

$$f_3(\alpha_{0,j}^{s2}, \beta_{jk}^{s2})) = \rho_{jk} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1})$$

186   I fit this model with JAGS in R.

12

Full model

To estimate the survival and germination of seeds in each population, I fit the following model for bag i at site j in year k:

$$[\phi, \gamma, \rho, \alpha | n, n_{\text{viab}}, y_{\text{total}}, y_{\text{germ}}, y_{\text{surv}}, y_{\text{viab}}] \propto$$

$$\prod_{j=1}^{J} \prod_{i=1}^{I} \text{binomial}(y_{ij}^{\text{tot}} | n_{ij}, \text{f}_1(\alpha_{0,j}^{s1}, \beta_j^{s1}))$$

$$\times \text{binomial}(y_{ij}^{\text{germ}} | y_{ij}^{\text{tot}}, \text{f}_2(\gamma_j, \alpha_{ijk}))$$

$$\times \text{binomial}(y_{ij}^{\text{surv}} | y_{ij}^{\text{tot}} - y_{ij}^{\text{germ}}, \text{f}_3(\alpha_{0,j}^{s2}, \beta_j^{s2}))$$

$$\times \text{binomial}(y_{ij}^{\text{viab}} | n_{ij}^{\text{viab}}, \text{logit}^{-1}(\alpha_{ijk})) \tag{4}$$

$$\times \text{normal}(\alpha_{0,j}^{s1} | 0, 10) \text{normal}(\alpha_{0,j}^{g1} | 0, 10) \text{normal}(\alpha_{0,j}^{s2} | 0, 10)$$

$$\times \text{normal}(\beta_j^{s1} | 0, \sigma_j^{s1}) \text{normal}(\beta_j^{g1} | 0, \sigma_j^{g1}) \text{normal}(\beta_j^{s2} | 0, \sigma_j^{s2})$$

$$\times \text{uniform}(\sigma_j^{s1} | 0, 100) \text{uniform}(\sigma_j^{g1} | 0, 100) \text{uniform}(\sigma_j^{s2} | 0, 100)$$

$$\times \text{normal}(\alpha_{ijk} | \mu_{jk}, \sigma_{jk})$$

$$\times \text{normal}(\mu_{jk} | 0, 100) \text{uniform}(\sigma_{jk} | 0, 100).$$

where

$$\text{f}_1(\alpha_{0,j}^{s1}, \beta_{jk}^{s1})) = \phi_{jk} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1})$$

$$\text{f}_2(\alpha_{0,j}^{g1}, \beta_{jk}^{g1}, \alpha_{ijk})) = \gamma_{jk} \times v_{ijk}^{\frac{1}{3}} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1}) \times (\text{logit}^{-1}(\alpha_{ijk}))^{\frac{1}{3}} \tag{5}$$

$$\text{f}_3(\alpha_{0,j}^{s2}, \beta_{jk}^{s2})) = \rho_{jk} = \text{logit}^{-1}(\alpha_{0,j}^{s1} + \beta_{jk}^{s1})$$

¹⁸⁷ These models relate the average seed survival in each population to the seed survival

¹⁸⁸ observed in a given year through a linear model. The function has a parameter for the

¹⁸⁹ average seed survival in each population, and a parameter for the seed survival of each

¹⁹⁰ population in each year. I treated seed survival as a binomially-distributed random variable

because the data come from an experiment in which we buried a known number of seeds and counted seeds. Thus for the ith observation, I parameterized a binomial distribution in terms of a probability and a known number of trials. I sampled a 95% credible interval from the posterior predictive distribution for seed survival for each site and compared it to the seed burial experiments to assess whether the model from the seed burial experiments could predict the seed survival in seed burial experiments.

I fit this model with JAGS in R.

Seedling survival to fruiting

The data consist of counts of seedlings and fruiting plants in 0.5 m^2 plots at 20 sites from 2006–present. Each site was visited in February and June to count the number of seedlings and fruiting plants, respectively. Seedlings and plants in each plot are counted by a single person at each visit.

For now, we assume that the data on seedlings is measured perfectly (i.e. no under- or over-counts of seedlings). However, there are at least two possible sources of error: (1) measurement error that arises because we failed to count seedlings that were present and (2) error that arises because seedlings germinated after we visited the site. Germination phenology varies may vary from year to year but also by geography; higher elevation sites may have delayed phenology. We may want to develop a model that relates our estimate of seedlings to the true number of seedlings in a plot because we sometimes observe more fruiting plants than seedlings. For now, I ignored data that involved undercounting by filtering out those rows in the dataset.

We assume that the data on fruiting plants is measured perfectly (i.e. we did not under- or over- count) because plants stand out from the background vegetation in June. Our model estimates the proportion of seedlings that survive to become fruiting plants. Define:

- n_{ijt} = observed counts of seedlings in the ith plot, from the jth site, from the kth year

- y_{ijt} = observed counts of fruiting plants in the ith plot, from the jth site, from the kth year, assumed to be measured perfectly

We estimated survival across all sites taking into account both temporal and between-site variability with the following model. In this model, $\alpha_{0,j}^S$ is the logit mean survival probability at site j, β_{jk}^S are independent identically distributed random variables drawn from normal distributions with mean 0 and site-specific temporal variance parameters σ_j^S. Writing the site-specific logit survival as a fixed effect means that each site parameter estimate is estimated separately with no shared variance term. The site-specific temporal variance parameter is written as a random effect, which means that each site has has year components that are drawn from a distribution with a shared variance term. I estimated the probability of surviving to fruiting using data from plots at sites in different years:

$$
\begin{aligned}
[\boldsymbol{\alpha_0^S}, \boldsymbol{\beta^S}, \sigma^S | \boldsymbol{n}, \boldsymbol{y}] \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} & \text{binomial}(y_{ijk}|n_{ijt}, \text{f}(\alpha_{0,j}^S, \beta_{jk}^S)) \\
& \times \text{normal}(\beta_{jk}^S|0, \sigma_j^S) \\
& \times \text{normal}(\alpha_{0,j}^S|0, 10)\text{uniform}(\sigma_j^S|0, 100)
\end{aligned}
\tag{6}
$$

where

$$
\text{f}(\alpha_{0,j}^S, \beta_k^S)) = \text{logit}^{-1}(\alpha_{0,j}^S + \beta_{jk}^S)
\tag{7}
$$

15

Fruits per plant

From 2006–2012, "we recorded...the number of fruits per plant for up to 15-20 plants per 0.5 m² " (Eckhart et al. (2011)). For each plant, we counted the number of undamaged fruits. We then took the damaged fruits and visually stacked them end to end to estimate how many additional undamaged fruits that was equivalent to (e.g. two half fruits corresponded to one undamaged fruit). We used these counts to estimate the number fruits produced per plant.

We seek to estimate (1) the number of fruits produced per plant and (2) the proportion of fruits that are damaged per plant. Define:

- y_{ijk}^{TFE} = observed counts of total fruit equivalents per plant on the ith plant, from the jth site, from the kth year, assumed to be measured perfectly

- n_{ijk} = observed counts of total fruits per plant (sum of $\boldsymbol{y_{ijk}}$) on the ith plant, from the jth site, from the kth year, assumed to be measured perfectly

To assess what probability distribution to use when fitting this model, I fit a power model with an intercept to the mean and variance using the 'nls' function in R, which returned an exponent of 1.99. The fit is close to quadratic which means a negative binomial is likely to be an appropriate distribution (Lindn and Mntyniemi (2011)). We estimated fruits per plant across all sites taking into account both temporal and between-site variability with the following model. I first worked only with data on total fruit equivalents on a plant (2006-2012). I estimated total fruit equivalents per plant as:

$$[\boldsymbol{\alpha_0^F}, \boldsymbol{\beta^F}, \sigma^F | \boldsymbol{n}, \boldsymbol{y}] \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} \text{negative binomial}(y_{ijk}^{\text{TFE}} | f_1(\alpha_{0,j}^F, \beta_{jk}^F), \kappa^F))$$

$$\times \text{normal}(\alpha_{0,j}^F | 0, 10) \qquad (8)$$

$$\times \text{normal}(\beta_{jk}^F | 0, \sigma_j^F)$$

$$\times \text{normal}(\sigma_j^F | 0, 100)$$

where

$$f_1(\alpha_{0,j}^F, \beta_{jk}^F)) = \lambda_{jk}^F = \exp(\alpha_{0,j}^F + \beta_{jk}^F) \qquad (9)$$

$$\text{negative binomial}(y_{ijk}^{\text{TFE}} | \frac{\kappa^F}{\kappa^F + \lambda_{jk}^F}, \kappa^F) \qquad (10)$$

where the negative binomial is parameterized with probability parameter p and dispersion parameter r [negative binomial(p, r)]. In this case $p = \frac{\kappa}{\kappa + \mu}$.

Seeds per fruit

From 2006–2012, "we collected one fruit from each of 20-30 haphazardly selected plants distributed across each population (but outside plots, to avoid influencing seed input within them) to estimate the mean number of seeds produced per fruit" (Eckhart et al. 2011). We collected fruits that were undamaged in the field, and fruits were broken open to count seeds. For each population in each year, we attempted to obtain 20-30 counts of seeds produced per undamaged fruit.

255   We seek to estimate (1) the number of seeds per undamaged fruit. Define:

256   • y_{ijk}^{und} = observed counts of seeds per undamaged fruit in the ith fruit, from the jth site,

257   from the kth year, assumed to be measured perfectly

258   • λ_{jk} = true, unobserved mean number of seeds per undamaged fruit from the jth site,

259   from the kth year

260   To assess what probability distribution to use when fitting this model, I fit a power model

261   with an intercept to the mean and variance using the 'nls' function in R, which returned an

262   exponent of 1.38. The fit is greater than linear but less than quadratic which means that

263   neither a Poisson nor negative binomial are likely to be entirely appropriate distributions for

264   the data (Lindn and Mntyniemi (2011)). I might try the parameterization in that reference

265   but for now I am using the negative binomial because the data are overdispersed. We

266   estimated seeds per fruit across all sites taking into account both temporal and between-site

267   variability with the following model. I first worked only with data from undamaged fruits

268   from the years (2006-2012). I estimated seeds per fruit as:

$$
\begin{aligned}
[\boldsymbol{\alpha_0^P}, \boldsymbol{\beta^P}, \sigma^P | \boldsymbol{n}, \boldsymbol{y}] \propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} & \text{negative binomial}(y_{ijk}^{und} | \text{f}_1(\alpha_{0,j}^P, \beta_{jk}^P), \kappa^P)) \\
& \times \text{normal}(\alpha_{0,j}^P | 0, 10) \times \\
& \times \text{normal}(\beta_{jk}^P | 0, \sigma_j^P) \\
& \text{normal}(\sigma_j^P | 0, 100)
\end{aligned}
\tag{11}
$$

269   where

$$
\text{f}_1(\alpha_{0,j}^P, \beta_{jk}^P)) = \lambda_{jk}^P = \exp(\alpha_{0,j}^P + \beta_{jk}^P)
\tag{12}
$$

18

$$\text{negative binomial}(y_{ijk}^{\text{und}} | \frac{\kappa^P}{\kappa^P + \lambda_{jk}^P}, \kappa^F P) \tag{13}$$

270 where the negative binomial is parameterized with probability parameter p and dispersion

271 parameter r [negative binomial(p, r)]. In this case $p = \frac{\kappa}{\kappa + \mu}$.

Correlation between germination probability and variance in per-capita reproductive success

272 I assessed whether the observed germination probability was negatively correlated with vari-

273 ance in per-capita reproductive success (Venable (2007)). Reproductive success F_{jk} at site

274 j in year k was calculated at the per year and per site level as follows:

$$F_{jk} = \phi_{jk} \times \lambda_{jk}^F \times \lambda_{jk}^P \tag{14}$$

275 where

$$\phi_{jk} = \text{logit}^{-1}(\alpha_{0,j}^S + \beta_{jk}^S)$$
$$\lambda_{jk}^F = \exp(\alpha_{0,j}^F + \beta_{jk}^F) \tag{15}$$
$$\lambda_{jk}^P = \exp(\alpha_{0,j}^P + \beta_{jk}^P)$$

276 To calculate the temporal variation in fitness for each population, I sampled the posterior

277 distribution of reproductive success for each year and calculated the geometric SD of per

278 capita reproductive success. For each population, I obtained a posterior distribution for the

19

geometric SD of per capita reproductive success; I used this and the posterior distribution of germination probability from model XX to calculate the correlation between germination and fitness variance. Using this approach, I obtained a distribution of correlation estimates. Results of this analysis are shown in Figure 2. Bet hedging models predict that germination probability should be negatively correlated with temporal variance in fitness; 95% credible intervals that do not overlap zero provide support for this prediction.

Correlation between germination probability and seed survival

I assessed whether the observed germination probability was negatively correlated with seed survival (Gremer and Venable (2014)). I calculated seed survival as $s_2 s_3$ as the product of these vital rates is the probability that seeds which do not germinate in January remain in the seed bank until the following January. I used the posteriors of g_1 and $s_2 s_3$ to calculate the correlation between germination and seed survival. Using this approach, I obtained a distribution of correlation estimates. Results of this analysis are shown in Figure 3. Bet hedging models predict that germination probability should be negatively correlated with seed survival; 95% credible intervals that do not overlap zero provide support for this prediction.

Density-independent model for germination probability

We used estimates of seed survival and reproductive success to investigate the adaptive value of delayed germination (Gremer and Venable (2014)). Briefly, we parameterize a model of population growth rate and calculate the optimal germination strategy for different combinations of seed survival and reproductive success. We use the following equation to summarize *Clarkia xantiana*'s life cycle and calculate population growth rate:

$$\lambda_j = g_1 Y(t) s_0 s_1 + (1 - g_1) s_2 s_3 \tag{16}$$

The parameters in this equation were fit in models corresponding to equations (1, 3, 5, 8). Seed survival rates (s_0, s_1, s_2, s_3) are population-level estimates. Per capita reproductive success $(Y(t))$ is calculated as the product of seedling survival to fruiting, fruits per plant, and seeds per fruit. Variation is incorporated into the model by having $Y(t)$ vary between years.

I numerically calculated the optimal germination probability for the observed level of variation in reproductive success and seed survival in each population. For each population, I randomly selected values 1000 from the posterior distribution for reproductive success. I used this same sequence of $Y(t)$ and the observed seed survival probabilities to calculate long-term stochastic population growth rates (λ_s) at each germination probability along an evenly spaced grid of possible germination probabilities (G) between 0 and 1. The optimal germination probability is estimated as the value of G that maximized geometric mean of the population growth rate. I repeated the simulations 50 times for each population, resampling from the posterior distribution for reproductive success each time. I calculated the mean of the optimal germination fractions.

A model in which per-capita reproductive success is density-independent predicts that germination probability should respond to variance in fitness (Cohen (1966)). To evaluate a density-independent model for germination probability, I compared observed germination probability to predicted germination optima. I plot this comparison in Figure 4. The dotted line indicates a 1:1 relationship between observations and predictions. Values below the line indicate that the model predicts higher germination probabilities than observed; values above the line would indicate that the model predicts lower germination probabilities than
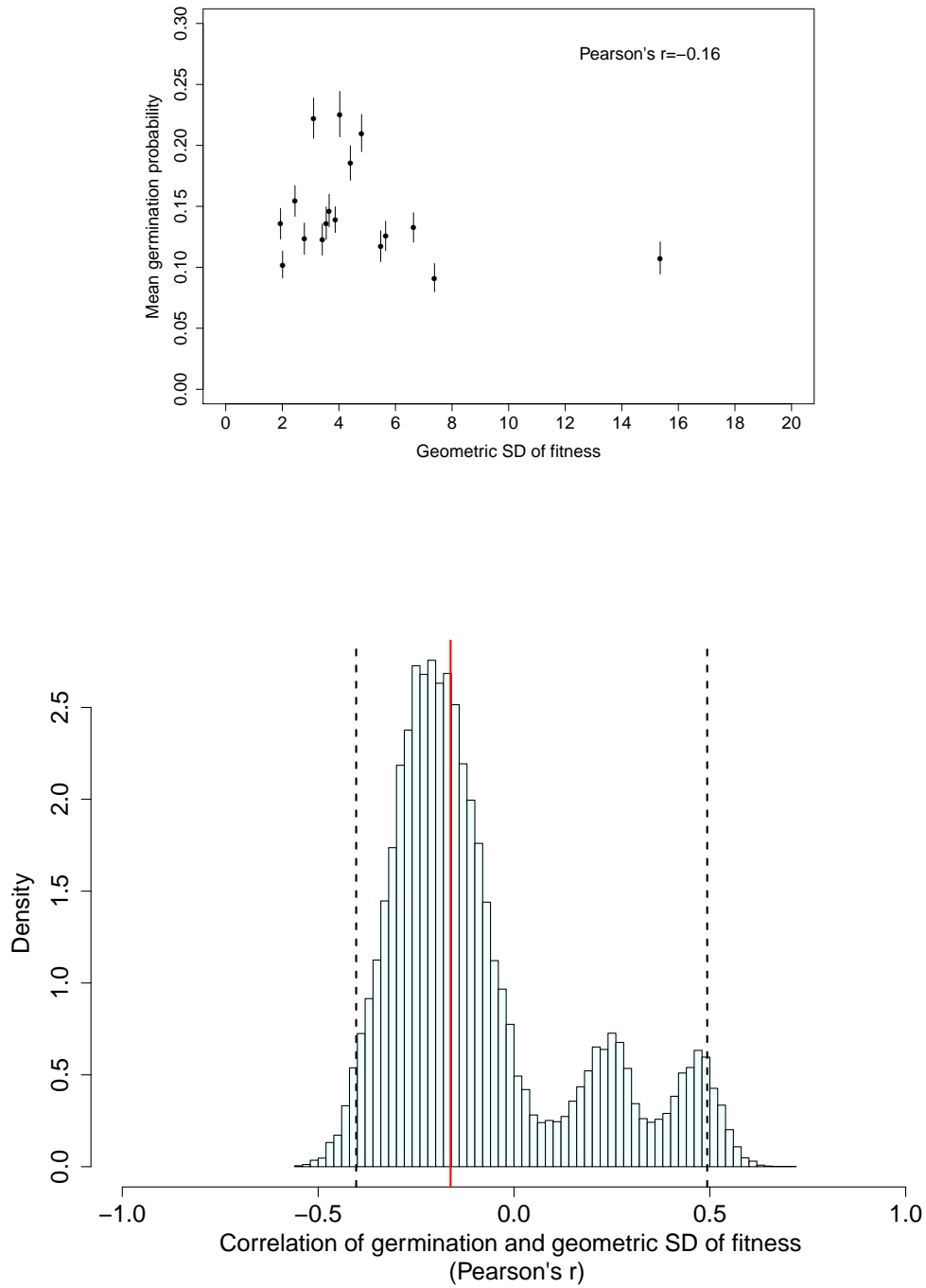
21

observed.

Results

Figure 2: The top panel shows the observed germination probability plotted against the temporal variation in per capita reproductive success. The bottom panel shows the posterior distribution of correlation between observed germination probability and geometric SD of per capita reproductive success; the median correlation is negative (-0.16) but the 95% credible interval overlaps 0.
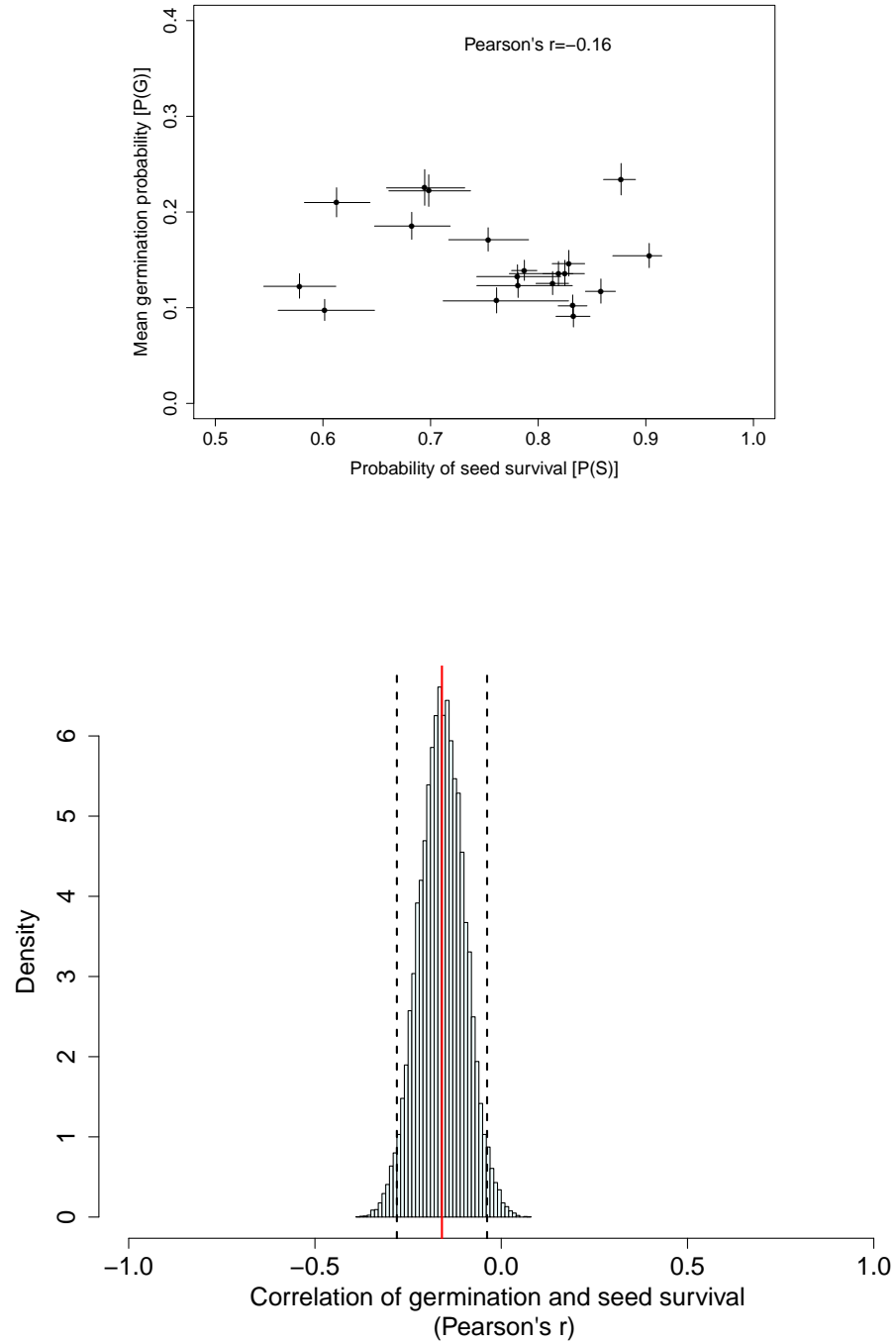
Figure 3: The top panel shows the observed germination probability plotted against probability of seed survival. The bottom panel shows the posterior distribution of correlation between observed germination probability and the probability of seed survival; the median correlation is negative (-0.16) and the 95% credible interval does not overlap 0.
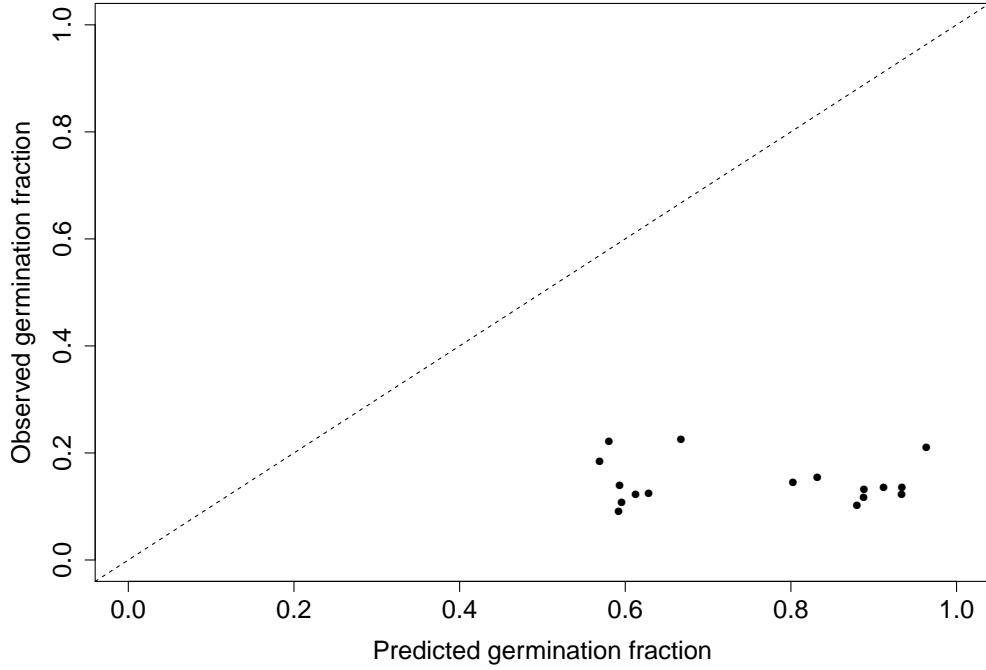
Figure 4: Observed germination probability plotted against the optimal germination probability predicted by a density-independent model. For each population, the observed germination probability is the obtained from the model for seed bank vital rates. Each point is the population-specific median of the posterior of g_1 for a model fit to data from seed bag experiments from 2006–2009. Data was pooled across years. The dotted line indicates a 1:1 relationship between observations and predictions. Values below the line indicate that the model predicts higher germination probabilities than observed; values above the line would indicate that the model predicts lower germination probabilities than observed.

References

D. Cohen. Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology*, 12(1):119–129, Sept. 1966. ISSN 00225193. doi: 10.1016/0022-5193(66)90188-3. URL https://linkinghub.elsevier.com/retrieve/pii/0022519366901883.

D. Cohen. Optimizing Reproduction in a Randoruly Varying Environment when a Correlation May Exist between the Conditions at the Time a Choice has to be Made and the Subsequent Outcome. *Journal of Theoretical Biology*, 16:1–14, 1967.

D. Cohen. A General Model of Optimal Reproduction in a Randomly Varying Environment. *The Journal of Ecology*, 56(1):219, Mar. 1968. ISSN 00220477. doi: 10.2307/2258075. URL https://www.jstor.org/stable/2258075?origin=crossref.

E. A. Eager, R. Rebarber, and B. Tenhumberg. Modeling and Analysis of a Density-Dependent Stochastic Integral Projection Model for a Disturbance Specialist Plant and Its Seed Bank. *Bulletin of Mathematical Biology*, 76(7):1809–1834, 2014. ISSN 15229602. doi: 10.1007/s11538-014-9978-y.

M. R. Easterling and S. P. Ellner. Dormancy strategies in a random environment: Comparing structured and unstructured models. *Evolutionary Ecology Research*, 2:387–407, 2000.

V. M. Eckhart, M. A. Geber, W. F. Morris, E. S. Fabio, P. Tiffin, and D. A. Moeller. The Geography of Demography: Long-Term Demographic Studies and Species Distribution Models Reveal a Species Border Limited by Adaptation. *The American Naturalist*, 178(S1):S26–S43, Oct. 2011. ISSN 0003-0147, 1537-5323. doi: 10.1086/661782. URL https://www.journals.uchicago.edu/doi/10.1086/661782.

S. Ellner. ESS germination strategies in randomly varying environments. II. Reciprocal Yield-Law models. *Theoretical Population Biology*, 28(1):80–116,

Aug. 1985a. ISSN 00405809. doi: 10.1016/0040-5809(85)90023-1. URL https://linkinghub.elsevier.com/retrieve/pii/0040580985900231.

S. Ellner. ESS germination strategies in randomly varying environments. I. Logistic-type models. *Theoretical Population Biology*, 28(1):50–79, Aug. 1985b. ISSN 00405809. doi: 10.1016/0040-5809(85)90022-X. URL https://linkinghub.elsevier.com/retrieve/pii/004058098590022X.

M. E. K. Evans and J. J. Dennehy. Germ Banking: BetHedging and Variable Release From Egg and Seed Dormancy. *The Quarterly Review of Biology*, 80(4):431–451, 2005. ISSN 0033-5770. doi: 10.1086/498282.

J. R. Gremer and D. L. Venable. Bet hedging in desert winter annual plants: optimal germination strategies in a variable environment. *Ecology Letters*, 17 (3):380–387, Mar. 2014. ISSN 1461023X. doi: 10.1111/ele.12241. URL http://doi.wiley.com/10.1111/ele.12241.

J. R. Gremer, S. Kimball, and D. L. Venable. Within-and among-year germination in Sonoran Desert winter annuals: bet hedging and predictive germination in a variable environment. *Ecology Letters*, 19(10):1209–1218, Oct. 2016. ISSN 1461023X. doi: 10.1111/ele.12655. URL http://doi.wiley.com/10.1111/ele.12655.

L. Heinrich, J. Mller, A. Tellier, and D. ivkovi. Effects of population- and seed bank size fluctuations on neutral evolution and efficacy of natural selection. *Theoretical Population Biology*, 123:45–69, 2018. ISSN 10960325. doi: 10.1016/j.tpb.2018.05.003. URL https://doi.org/10.1016/j.tpb.2018.05.003.

A. Lindn and S. Mntyniemi. Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology*, 92(7):1414–1421, July 2011. ISSN 0012-9658. doi: 10.1890/10-1831.1. URL http://doi.wiley.com/10.1890/10-1831.1.

K. A. Mccue and T. P. Holtsford. Seed bank influences on genetic diversity in the rare annual Clarkia springvillensis (Onagraceae). *American Journal of Botany*, 85(1):30–36, 1998. ISSN 00029122. doi: 10.2307/2446551.

L. Nunney. The Effective Size of Annual Plant Populations: The Interaction of a Seed Bank with Fluctuating Population Size in Maintaining Genetic Variation. *The American Naturalist*, 160(2):195, 2002. ISSN 00030147. doi: 10.2307/3079137. URL https://www.jstor.org/stable/10.2307/3079137.

M. Paniw, P. F. Quintana-Ascencio, F. Ojeda, and R. Salguero-Gmez. Accounting for uncertainty in dormant life stages in stochastic demographic models. *Oikos*, 126(6):900–909, 2017. ISSN 16000706. doi: 10.1111/oik.03696.

K. Ritland. The joint evolution of seed dormancy and flowering time in annual plants living in variable environments. *Theoretical Population Biology*, 24(3):213–243, 1983. ISSN 10960325. doi: 10.1016/0040-5809(83)90026-6.

D. L. Venable. Bet hedging in a guild of desert annuals. *Ecology*, 88(5):1086–1090, May 2007. ISSN 0012-9658. doi: 10.1890/06-1495. URL http://doi.wiley.com/10.1890/06-1495.

R. S. Waples. Seed Banks, Salmon, and Sleeping Genes: Effective Population Size in Semelparous, Age-Structured Species with Fluctuating Abundance. *The American Naturalist*, 167(1):118, 2006. ISSN 00030147. doi: 10.2307/3491252. URL https://www.jstor.org/stable/10.2307/3491252.