

1 **[Working title]: Intraspecific variation in range-wide seed bank dynamics is not**
2 **consistent with density-independent model of bet hedging**

3 Gregor-Fausto Siegmund and Monica Geber

4 Last updated: March 20, 2021

Writing list

- 5 1. Read Cohen & Ellner to identify the role of complete reproductive failure in the original
6 models for the evolution of bet hedging
- 7 2. Write description of modeling for the seed bag experiment
- 8 3. Write about strong and weaker test of the hypothesis (partial pooling, no partial pool-
9 ing)
- 10 4. Write paragraph about history of studies of bet hedging via seed bank, with emphasis
11 on how this study addresses this question at an intraspecific level. Lower level of
12 variation in intraspecific germination fraction. ([link](#))
- 13 5. Write models for belowground vital rates
- 14 6. Revise description for density-independent model for germination probability
- 15 7. Write and implement model checking process

Introduction

General introduction to seed banks/dormancy – Clarkia as a case where seed banks have been posited to be important – bet hedging theory and predictions – review of tests of bet hedging in plants – what we did

Seed banks can buffer plant populations against environmental change and stochasticity (Eager et al. (2014); Paniw et al. (2017)), increase effective population size (Nunney (2002); Waples (2006)), and maintain genetic diversity (Mccue and Holtsford (1998)). Dormancy can 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)).

What drives the evolution of delayed germination? 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? These models 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 (see: Supplementary Material: Theoretical background for hypotheses). 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 germination fraction decreases. Bet hedging should evolve to maximize the long-term geometric population growth rate (as compared to the arithmetic population growth rate) (Cohen (1966, 1968); Ellner (1985a,b)).

Cohen (1966) emphasizes the role of particularly bad years. This is highlighted by the inequality in equation (12), which states that for the optimal germination strategy to be

bet hedging, it is sufficient that the harmonic mean is less than the survival probability of seeds that do not germinate. The minimum fitness has a strong impact on the harmonic mean of fitness (see definition of harmonic mean on Wikipedia), and a single year of very low fitness would tend to make the harmonic mean small. In particular, the harmonic mean is 0 when any of its values is 0. The probability of complete reproductive failure is thus particularly important. Years with no seedling survival in the plots or no germination are those in which fitness would be zero; there are only 4/20 populations in which this does not happen. We fit two types of models to the aboveground data; one with partial pooling and one with no pooling. The first was an attempt to correct for sampling variation; the latter an attempt to estimate per-capita reproductive success as-is, providing a more extreme estimate of interannual variation but one that reflects the combined effect of sampling variation and true variation.

Revisit theory to see particular role of complete reproductive failure versus low fitness years more generally.

In paragraph on theory, discuss density-independent models Cohen (1966, 1968), density-dependent models Ellner (1985a,b), predictive germination Cohen (1967), and unstructured versus structured models Easterling and Ellner (2000) (Tuljaparkur 1993, Valleriani and Tielborger 2006).

Paragraph about the history of empirical studies of bet hedging.

Population vital rates are known to vary across *Clarkia 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.]

A previous study with *Clarkia xantiana* 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.

See also the comments in Lewis 1953 on lifespan of seeds, absence of dormancy, and cued germination. Lewis 1963 Presidential address at SSE on catastrophic selection is also relevant. Discusses extinction and possible absence of a seed bank in populations of *Clarkia* species. Specifically mentions a population in the upper Kern Valley that had both whites and pinks; the pinks went extinct during a drought year while the whites maintained their population size. This work was followed up by others on different species in *Clarkia* genus. I think its important to consider the possibility that *C. xantiana* being studied in 1963 was a *xantiana/parviflora* ssp. which may have different dynamics than the outcrossing *xantiana/xantiana*. Furthermore: the upper parts of the range have the highest rates of germination which may also trade off with low survival. It would be interesting to put that observation in the range-wide context!

Knies et al. 2004 studied 2 *C. xantiana* populations; seeds of outcrosser are heavier than selfers. Identify heavier seeds (.1 mg heavier at SC than WH population). Seeds from both populations were bullet-shaped but differ in shape of exotesta cells; WH cells are gourd shaped, SC shapes are even width and spherical. Seed size/number trade-off: smaller size for dispersal; greater mass increases germination probability. Smaller seed size also facilitates seed integration into soil seed bank. Small seed size is also correlated with a persistent

seed bank (Thompson et al. 1993). Larger seeds may germinate under more conditions and perhaps have smaller seed bank.

Moeller et al. 2011 population genetic study.

Eckhart and Geber 1999 identify more populations than those of Lewis, populations that they have since monitored intensively.

Price et al. 1985 followed a population for 2 years, did an elimination experiment. determined a dormant seed bank. Snow 1960 reports eliminating plants and not observing plants the following year.

Geber and Eckhart 2005 observed a 30-fold difference in lifetime fitness [calculated as number of germinants times survival, flower number, fruit set and seed set] between two transplant years (one wet, one dry). Fitness was very low in the dry year of the study (0.02-0.03 seeds per seed planted). They observed a 2-4 fold difference between years in germination, seedling survival, and flower number. They did not observe a difference in fruit or seed set. Western populations showed evidence of regional adaptation: survival and flower number. Germination declined from the western to eastern transplant sites [in contrast to what we observe here?] Germination is consistently higher in subspecies *xantiana* than *parviflora*. From west to east, rainfall declines and becomes more variable. Low survivorship is the product of end-of-season drought, small mammal herbivory (Geber and Eckhart 2005, Benning et al. 2019), competition by higher vegetation cover of grasses (survivorship increases with plant size), fungal rust infection. Plants from western populations are larger. Note that the transplant experiments took place on igneous soils, and later work by Vince has shown that soil type is important.

Eckhart et al. 2010 IJPS: geographic patterns in precipitation, decline from west to east, mirrored geographic pattern in midday water potential (reduced performance). Metasedimentary soils were more resistant and finer. Soil characteristic curves indicated that metasedimentary soils can hold more water but that it becomes unavailable to plants at higher water

contents. Higher water potential on steeper, eroding slopes; this soil was softer and coarser. Decline in soil water potential (more negative) decreases plant size and fruit number. However, soil water potential is unrelated to seedling survival in the study years. Metasedimentary rocks may compound effect of decline in precipitation. Metasedimentary rocks may hold more water but it may be more available to plants in igneous soils.

Bet hedging should evolve to maximize the long-term geometric population growth rate (as compared to the arithmetic population growth rate) (Cohen (1966, 1968); Ellner (1985a,b)). Seed banks are more likely to be selected in populations which experience higher levels of interannual variation in per-capita reproductive success. To investigate this empirical relationship, I will estimate the correlation between interannual variation in per-capita reproductive success 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 per-capita reproductive success. [What is the relationship between interannual variation in fitness and dormancy? This is a question about whether fitness variation and dormancy are positively correlated as expected ? this is what selects for bet hedging.]

Seed dormancy and persistence in the soil seed bank may be a bet-hedging strategy that is favored by environmental uncertainty. If this is the case in *C. xantiana*, I think we might expect increased seed survival and decreased germination at the eastern range edge (precipitation is more variable at eastern populations in winter and spring). This seems to be opposite of what was observed in the previous study ? though I could also be misinterpreting this. This is could be one reason to revisit this question with a new analysis.

Why using the Clarkia system? We expect to see differences among populations in either seed survivorship or variance in fitness. 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. [need to be clear

about this] also know that the environment changes across the range; that will affect the mean fitness

I expect bet hedging stages to be sensitive to variation in an environmental cue. For something to be an adaptive strategy, it should respond to variation in the environment to capitalize on good years such as ones with high precipitation. To investigate the sensitivity of vital rates to the environment, I estimate the slope of the relationship between environmental cues and vital rates at individual sites using a random coefficients model that estimates covariation between intercepts and slopes. I predict that I will estimate variation both in the intercept and slope for populations, and that variation in the modeled cue [to be determined] will be positively related to the estimated slope. [What is the relationship between dormancy and environmental cues? This is a question about whether we can make any inferences about the mechanism responsible for bet hedging in this system.]

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 (1985 <i>a,b</i>)	Cohen (1967)	Tuljapurkar and Is- tock 1997 Easterling and Ellner (2000) Val- leriani and Tielborger 2006
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 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

1 Methods

1.1 Clarkia life history

Clarkia xantiana is a winter annual that germinates with late fall and winter rains. In our study region, the Kern Valley in the southern Sierra Nevada Mountains, historically happens between October and late February or early March. Seedlings grow throughout the winter and spring, and surviving plants flower in late spring and early summer, April into early July. Pollinated fruits set seed in the early summer, June to July. Seeds of *C. xantiana* are produced in early summer, with fruits that dry out and gradually split open. Most seeds appear to be shed from fruits within 3-4 months after production, but can remain on the plant for up to a year. Seeds are small (< 1 mm in width) and have no structures to aid in aerial dispersal.

We represent the *Clarkia xantiana* as a life cycle graph (Figure 1A) that describes transitions from October of year t to October of year $t + 1$ in terms of underlying vital rates. The census period occurs when the entire population is seeds, and corresponds to the time at which seed bags are placed into the field (see below). Seeds are grouped into three stages: age 0 seeds, which were produced in the current year; age 1 seeds, which were produced in the previous year; age 2+ seeds, which were produced two or more years ago. Persistence of seeds in the seed bank is represented by transitions from

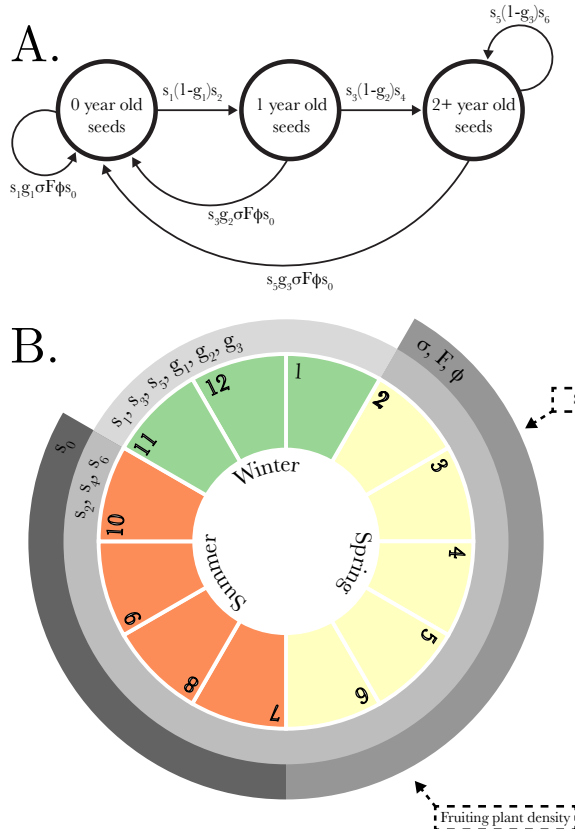


Figure 1: Diagram representations of *Clarkia* life cycle.

180 younger to older seeds. Production of new
 181 seeds is captured by transition to the age 0
 182 seed state.

183 Transitions in the life cycle graph are the product of age-specific seed survival and germi-
 184 nation, and aboveground seedling survival to fruiting, fruit production, and seeds per fruit.
 185 Seed-related rates are represented separately for age 0, 1, and 2+ seeds. Germination for
 186 each age class is given as g_1 , g_2 , and g_3 , respectively. Seed survival from seed production to
 187 the first October is given as s_0 , and survival from October to February is given as s_1 , s_3 ,
 188 and s_5 for age 0, 1, and 2+ seeds, respectively. Survival from February to October is given
 189 as s_2 , s_4 , and s_6 for age 0, 1, and 2+ seeds, respectively. We assume that vital rates remain
 190 unchanged after age 2. We also we assume that all plants experience the same vital rates
 191 upon germination seed age at germination does not affect seedling survival to fruiting (σ),
 192 fruits per plant (F), or seeds per fruit (ϕ).

The life cycle graph (Figure 1A) corresponds to the annual projection matrix

$$\mathbf{A} = \begin{bmatrix} s_1 g_1 \sigma F \phi s_0 & s_3 g_2 \sigma F \phi s_0 & s_5 g_3 \sigma F \phi s_0 \\ s_1 (1 - g_1) s_2 & 0 & 0 \\ 0 & s_3 (1 - g_2) s_4 & s_5 (1 - g_3) s_6 \end{bmatrix} \quad (1)$$

193 that summarizes transitions between stages.

194 [Eckhart et al. 2004](#) conducted a transplant experiment in two years in which seed set
 195 varied ten-fold. There are additional studies that did elimination experiments (with other
 196 species of *Clarkia*) and suggested limited role/role for seed bank because plants did or did
 197 not come up with following year.

198 [Lewis 1962](#) suggested that catastrophic selection played a major role in the dynamics
 199 of *Clarkia* populations. In [13 years](#) of data on 20 populations across the species range,

we observed high seedling mortality and/or low germination in at least one year in all but four populations. Although most populations had at least some flowering plants in most years, there were also four populations in which no flowering plants were observed in the entire population. None of these populations, however, went extinct during the period of the study. Seeds can remain viable in the field and laboratory for at least 5 years, contrary to previous suggestions that seeds are unlikely to persist past a few years (Lewis 1962).

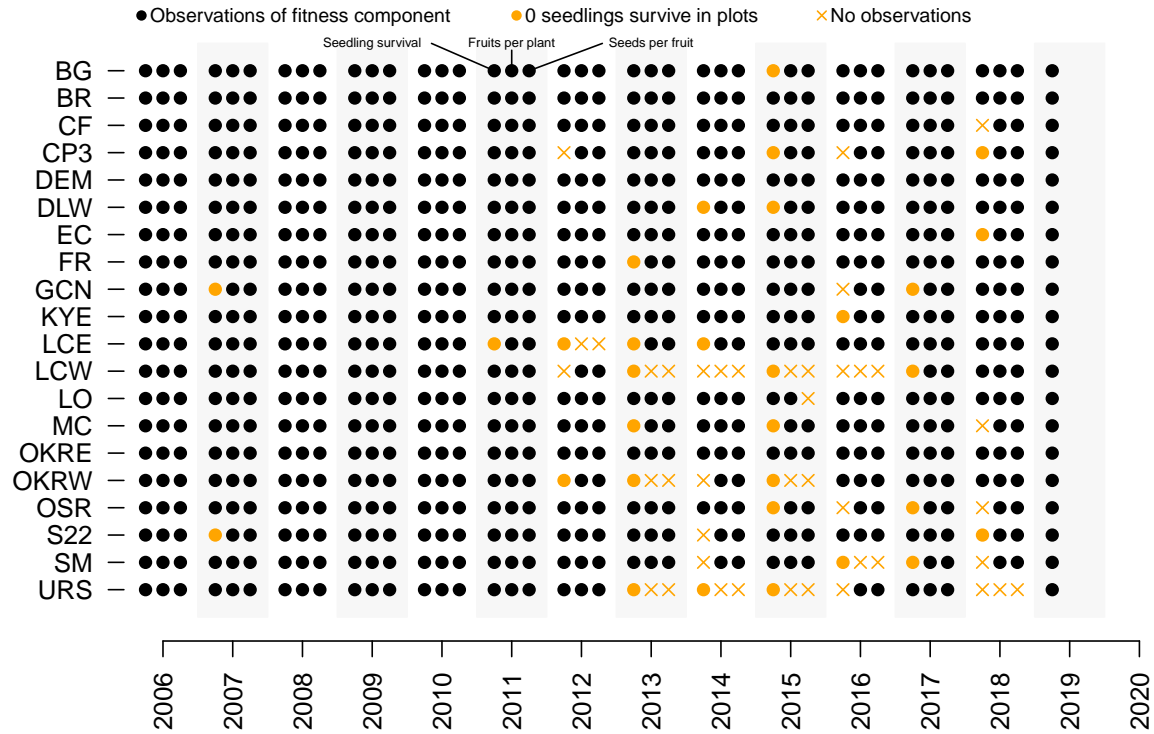


Figure 2: Summary of the aboveground observations, low fitness, and no observations.

1.2 Creating the dataset

We used field experiments and surveys to assemble observations of below- and above-ground demography for 20 populations of *Clarkia xantiana* (Table 2). Specifically, we used experiments to estimate transitions in the seed bank and surveys to estimate components of per-capita reproductive success. These demographic data have previously been used to test hypotheses about the geography of demography (Eckhart et al. 2011) and species distributions (Pironon et al. 2018). Here, we sought to obtain population-level estimates of germination and seed survival, and yearly estimates of per-capita reproductive success.

To estimate transitions in the seed bank, we used observations from a seed bag burial experiment conducted in all populations from 2006-2009 (Figure ??). The experiment has been previously described in Eckhart et al. 2011 and we reanalyze the data here. Geber and collaborators buried seeds in bags and unearthed 1, 2, or 3 years after being buried to count seedlings and intact, viable seeds. The experiment was repeated in 3 consecutive years and ended in 2009. We thus have 3 sets of observations associated with 1 year old seeds, 2 sets of observations associated with 2 year old seeds, and 1 set of observations associated with 3 year old seeds. We use data from the experiment to estimate age-specific germination and seed survival (see Joint model for seed vital rates) but note that we test predictions of bet-hedging theory that are based on an unstructured seed bank and use only the relevant subset of transitions in our analysis (see Computing vital rates).

To estimate per-capita reproductive success, we combine censuses of seedlings and fruiting plants, surveys of fruits per plant, and lab counts of seeds per fruit. To assess the survival of seedlings to fruiting plants, we counted seedlings and fruiting plants in 30 0.5 m² permanent plots from 2006–2018 (Eckhart et al. 2011). To assess seed production by plants that survive to reproduction, we counted fruits per plant on individual plants in permanent plots, and on additional haphazardly chosen plants throughout the population. We then attempted to

230 obtain 20-30 fruits per population, which we used to count seeds per fruit (Eckhart2011).

Table 2: Summary of data sets used to estimate parameters.

	Parameter data	Description	Data set	Time span
	SEED VITAL RATES	—	—	—
	Seed survival and germination	Seed bag burial	\mathbf{Y}_1	2006-2009
	Seed viability	Viability trials	\mathbf{Y}_2	2006-2009
	Seed survival and germination	Seed pots	\mathbf{Y}_3	2013-2019
	SEEDLING SURVIVAL	—	—	—
	Seedling survival to fruiting	Field surveys	\mathbf{Y}_4	2006-2019
232	FRUITS PER PLANT	—	—	—
	Total fruit equivalents per plant	Field surveys	\mathbf{Y}_5	2006-2012
	Undamaged and damaged fruits per plant	Field surveys	\mathbf{Y}_6	2013-2019
	Total fruit equivalents per plant	Extra plots	\mathbf{Y}_7	2006-2012
	Undamaged and damaged fruits per plant	Extra plots	\mathbf{Y}_8	2013-2019
	SEEDS PER FRUIT	—	—	—
	Seeds per undamaged fruit	Lab counts	\mathbf{Y}_9	2006-2019
	Seeds per damaged fruit	Lab counts	\mathbf{Y}_{10}	2013-2019

1.3 Model

233 We use observational and experimental data from 20 populations of *Clarkia xantiana* to
 234 estimate transition probabilities across the life cycle. We fit multilevel models to obtain
 235 population-specific estimates for belowground vital rates, and year- and population-specific
 236 estimates for aboveground vital rates. Because we were interested in describing the life
 237 histories of individual populations, we built separate models for each population. Our gen-
 238 eral approach applies a common model structure to partially pool observations from each
 239 population.

240 We first explicitly describe our formulation in terms linear mixed models before defining
 241 the joint posterior (Evans et al. 2010, Ogle and Barber 2020). We assume that the latent
 242 mean of observations in year j at a population k , θ_{jk} , is drawn from a normal distribution
 243 with mean $\theta_{0,k}$ and variance σ_j^2 .

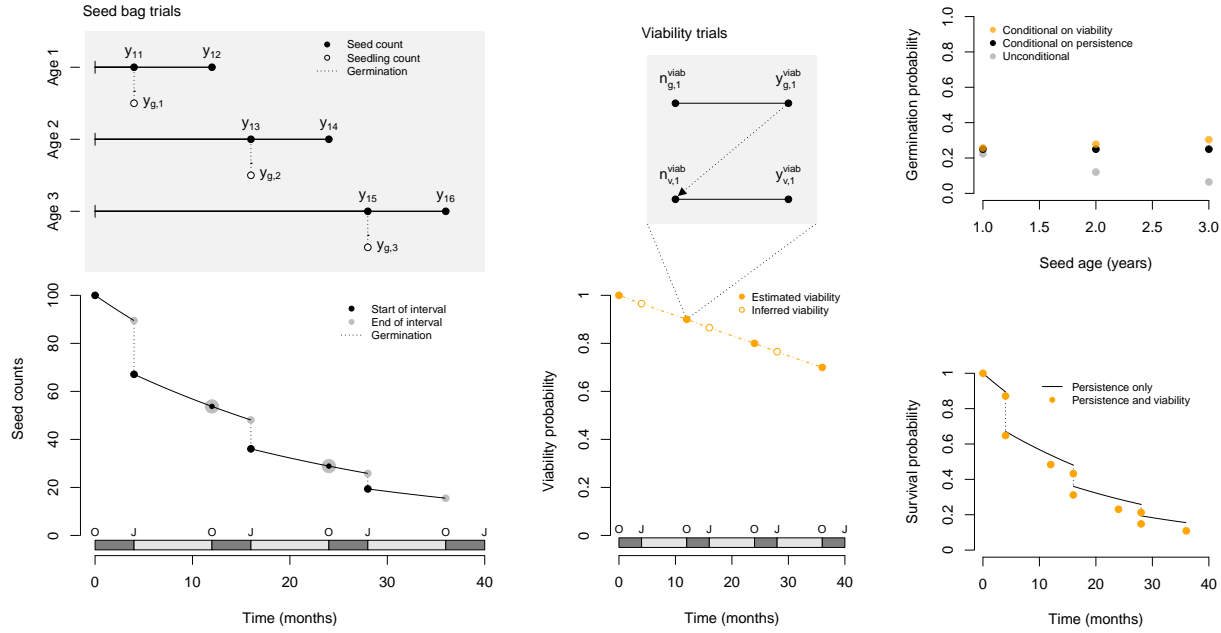


Figure 3: Summary of the seed bag burial experiments and viability trials. Figure will be labeled as (A-B: seed bag trials, C-D: viability trials, E-F: germination probability and survival probability. Add month markers to the y-axis in panels B, D, F. (A) The gray panel contains a graphical representation of the seed bag trials. Seeds were buried at the start of each experiment (100 seeds in month 0). Seed bags were unearthed and intact seeds ($y_{..}$) and germinants ($y_{g,}$ counted. The graph below the panel shows a hypothetical survival function associated with persistence of seeds in the soil seed bank. (B) The gray panel contains a graphical representation of the viability trials. Seeds were tested in two rounds; germination trials were performed and then some or all of the ungerminated seeds were tested for viability. The graph below the panel shows hypothetical data from a series of viability trials and the interpolated, inferred viabilities at times when viability was unobserved. (C) Age-specific germination probability is summarized in three ways. (D) The graph shows the survival function for persistence of seeds in the soil seed bank (black line) and the estimated discrete survival probabilities for persistence and viability of seeds (orange points).

$$\theta_{jk} = \theta_{0,k} + \epsilon_{(jk)}. \quad (2)$$

effects can be written as $\epsilon_{(jk)} \sim N(0, \varsigma^2)$. For the moment, we focus on describing the hierarchical structure of the model but note that we use link functions for transformation to parameters that are appropriate for the likelihoods we use to model different sets of observations (e.g. binomial for seed bag experiments; Poisson for counts of seed per fruit). We note that such a linear mixed effects model with random intercepts for years is one method commonly used to model interannual variation in demographic rates (e.g. Metcalf et al. 2015). Using hierarchical centering, the same model is rewritten as

$$\theta_{jk} = \alpha_{(jk)}. \quad (3)$$

The mean θ_{jk} , is now drawn from a normal distribution with mean $\alpha_{(jk)}$ and variance σ_j^2 . We place a prior on $\alpha_{(jk)}$ such that $\alpha_{(jk)} \sim N(\theta_{0,k}, \varsigma^2)$. The expressions are related by $\alpha_{(jk)} = \theta_{0,k} + \epsilon_{(jk)}$. We thus draw year-level means from the population-level means.

For a single population (ie. suppressing subscript k), we write the the posterior proportional to the joint distribution as

$$[\theta_j, \theta_0, \sigma_j^2, \varsigma^2 | y_{ij}] \propto [y_{ij} | \theta_j, \sigma_j^2][\theta_j | \theta_0, \varsigma^2][\theta_0][\sigma_j^2][\varsigma^2]. \quad (4)$$

The distribution of the observations y_{ij} is conditional on the year-specific parameters θ_j and σ_j^2 . In turn, the year-specific parameter θ_j is conditional on the population-specific parameters θ_0 and ς^2 . We placed priors on all parameters found only on the

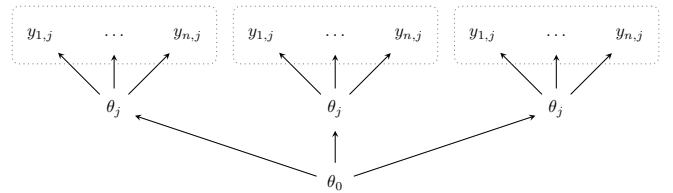


Figure 4: Graph depicting the general structure for the hierarchical models, for one population. Observations from each year, y_{ij} , are shown grouped and outlined by dotted lines. The observations are drawn from year-level parameters, θ_j , which in turn are drawn from a population-level parameter, θ_0 . The graph omits variance terms.

right hand side of conditional statements
 $(\theta_0, \sigma_j^2, \varsigma^2)$. In practice, we implemented
this model by specifying the population- and
year-levels of the model with normal distri-
butions; for example, $[\theta_j | \theta_0, \varsigma^2]$ is $\theta_j \sim N(\theta_0, \varsigma^2)$. The model thus describes a structure in
which years are nested within populations.

1.4 Model statements, implementation, and fitting

We include the expression for the posterior proportional to the joint distribution, and corre-
sponding directed acyclic graphs, in [Appendix: Joint Posterior](#). Priors for all parameters are
defined in [Table Priors](#). We applied the following principles for specifying priors: (1) we used
weakly informative priors that avoided placing probability mass on biologically implausible
values ([Gelman; Lemoine; Wesner and Pomeranz](#)), (2) we placed positive, unbounded priors
on variance components ([REF](#)), (3) we conducted prior predictive checks to assess the scale
of priors after parameter transformation ([Hobbs and Hooten; Gabry; Wesner and Pomer-
anz](#)), and (4) we simulated prior predictive distributions to confirm that the joint likelihood
generated data within the observed range ([Gabry; Conn; Hobbs and Hooten](#)). We provide
additional detail regarding our choice of priors in [Appendix: Priors](#).

We prepared data for analysis using the tidyverse and tidybayes packages ([CITE](#)) in R
[VERSION; CITE](#). We wrote, fit all models, and estimated posterior distributions using
JAGS [VERSION](#) with rjags ([Plummer 2016](#)). We randomly generated initial conditions for
all parameters with a prior by drawing from the corresponding probability distribution in R
before passing the initial values to rjags. We ran three chains for [XX,000](#) iterations. The first
[XX,000](#) samples were discarded as burn-in and we sampled the following [XX,000](#) iterations.
We did not thin the chains ([Elder and Miller 2016](#)).

We assessed convergence of the MCMC samples with visual inspection of trace plots, by calculating the Brooks-Gelman-Rubin diagnostic (R-hat), and by calculating the Heidelberg-Welch diagnostic (Elder and Miller 2016). The Gelman-Rubin diagnostic is used to assess convergence between chains and the Heidelberg-Welch for stationarity within chains. Trace plots for all chains, histograms of R-hat, and the percentage of chains that passed the HW test are shown in the appendix.

To evaluate our model's fit to the data, we performed model checks that are described in full in Appendix: Model Checking. We used our posterior distribution to simulate replicate datasets based on the parameters of our model. We compared samples from the simulated datasets to the real, observed datasets using both graphical, visual checks and by calculating Bayesian p -values for test statistics calculated for the observed and simulated data. In the following section, we describe how we used the models we fit to obtain the parameters that describe the *Clarkia* life history. While we do not perform model checks for these derived quantities (e.g. winter seed survival accounting for the combined effect of seed decay and loss of viability) because we combine the output of multiple models, the model checks are still essential to determine whether our inferences are reasonable.

1.5 Computing vital rates

1.5.1 Belowground vital rates - seed bank paper

We used the age-specific germination probabilities, survival function, and viability estimates to account for viability in estimates for the probability of germination and survival. We first discretized the survival function to times at which we observed germination and counted seeds (January and October). Estimates of survival over these intervals are the probability that a seed remains intact, but does not account for loss of viability. Next, we used viability estimates from October to calculate viability for January by interpolation (Figure 3D). We

tested the viability of seeds in October, and were thus able to estimate the proportion of viable seeds (Figure 3B; filled points). We inferred the viability of intact seeds in January by assuming that seeds lost viability at a constant rate (exponential decay). Further, we interpolated between estimates by assuming that viability changed at a constant rate between years, and that all seeds were viable at the start of the experiment (Figure 3B; open points).

We combined the discretized survival function and viability estimates to construct a survival function for the probability that a seed remains intact and viable (Table 3, column X). Specifically, we multiplied the posteriors of the discretized survival and viability estimates. Because we combined estimates, some portions of the posterior for seed survival probability was than 1, especially for later seed ages. We restricted the posterior to be less than 1 by truncating the distribution and resampling to redistribute the probability mass. We take this step to retain parameter uncertainty about survival probability in cases where combining the estimates implies a high probability of survival. The survival function for viable seeds (ϕ) is composed of estimates of persistence over time (θ), estimates of viability (ν), and estimates of germination conditional on persistence (γ).

We used the discretized survival function and age-specific germination probability to obtain the estimates of germination and seed survival required to test predictions from betting theory. Table ?? defines the age-specific germination probabilities and survival probabilities for the structured model in Eckhart et al. 2011 in terms of the survival function and age-specific germination probabilities. Figure 3E-F illustrate the relationship among the various probabilities of germination and seed survival. Estimates from the seed bag experiment correspond to the probability of germination or survival conditional on persistence (e.g. γ_1). Multiplying these estimates by the probability of persistence up to a certain time gives the unconditional probability (e.g. $\theta_1 \times \gamma_1$). Finally, the probability conditional on persistence and viability is estimated by incorporating loss of viability into the survival function (e.g. γ_1/ϕ_1), and defines the parameters in the structured population model.

Table 3: Seed persistence and viability in the soil seed bank

	Persistence	Persistence & viability
Time (x_i)	$S(x_i)$	$S(x_i)$
Oct ₀	θ_0	$\phi_0 = \theta_0$
Jan _{1,total}	θ_1	$\phi_1 = \theta_1(\gamma_1 + (1 - \gamma_1)\nu_1^{1/3})$
Jan _{1,intact}	θ_2	$\phi_2 = \theta_2\nu_1^{1/3}$
Oct ₁	θ_3	$\phi_3 = \theta_3\nu_1$
Jan _{2,total}	θ_4	$\phi_4 = \theta_4(\gamma_2 + (1 - \gamma_2)\nu_1(\nu_2/\nu_1)^{1/3})$
Description	Parameter	Probability
July-October	s_0	
October-January	s_1	ϕ_1
1-year old germination	g_1	γ_1/ϕ_1
January-October	s_2	ϕ_3/ϕ_2
October-January	s_3	ϕ_4/ϕ_3

1.5.2 Per-capita reproductive success

In order to make our analysis comparable to previous empirical studies of bet hedging, we calculated per-capita reproductive success as the product of the probability of seedling survival to fruiting, fruits per plant, and seeds per fruit. We thus calculate per-capita reproductive success as the number of seeds produced per seedling, on average (e.g. Venable 2007, Gremer et al. 2014).

We used a consistent method to estimate seedling survival to fruiting throughout the experiment, and use the population- and year-level means ($\mu_{S,jk}$) in our calculation. Because we estimated fruit production in 2 different ways during the study, we chose to use total fruit equivalents (TFE) per plant as our common estimate of fruit production. From 2006–2012, we used $\mu_{TFE,jk}$ as estimated in the statistical model. From 2013–2018, we used the ratio of seeds per damaged to undamaged fruit to calculate a proportion of damaged fruits to add to undamaged fruit counts, as in

$$\text{TFE} = \text{undamaged fruits} + \frac{\text{seeds per damaged fruit}}{\text{seeds per undamaged fruit}} \times \text{damaged fruits.} \quad (5)$$

We used posterior distributions for population- and year-level parameters (e.g. $\mu_{\text{US},jk}$) for these calculations and obtained estimates of $\mu_{\text{TFE},jk}$ for 2013–2018. Finally, we used estimates of seeds per undamaged fruit ($\mu_{\text{US},jk}$) as our estimate of seeds per fruit.

In terms of parameters from our statistical models, per-capita reproductive success F_{jk} at population j in year k is calculated as

$$F_{jk} = \phi_{jk} \times \lambda_{\text{TFE},jk} \times \lambda_{\text{US},jk}, \quad (6)$$

where

$$\begin{aligned} \phi_{jk} &= \text{logit}^{-1}(\mu_{\text{S},jk}) \\ \lambda_{\text{TFE},jk} &= \exp(\mu_{\text{TFE},jk}) \\ \lambda_{\text{US},jk} &= \exp(\mu_{\text{US},jk}). \end{aligned} \quad (7)$$

Our multilevel models for aboveground vital rates pooled data more strongly in years with relatively little data. A benefit of this approach is that it implicitly corrects for variation in sample size (e.g. an observation of 0/37 seeds surviving is given more weight than an observation of 0/1 seeds surviving). While this is beneficial for distinguishing between spurious estimates and true temporal variation in reproductive success, it may also underestimate variation in reproductive success. At the extreme, estimates in years without any data are pooled to the population-level means. Years with zero seedling survivorship would thus have

estimates for fruits per plant that are pooled towards the population-mean (because there were no fruiting plants on which to count fruits).

We calculated the posterior mode of annual estimates for each parameter in Equation 6 before multiplying to obtain the per-capita reproductive success in that year. Using the posterior mode is equivalent to taking the BLUP of a linear model, and allowed us to estimate vital rates in years with small sample sizes.

Analysis

Strict and less strict tests of the bet-hedging model. We consider two models for per-capita reproductive success. In the first, we use partial pooling to correct for sampling bias in estimates of seedling survival, fruits per plant, and seeds per fruit. However, our model with partial pooling pools years with few plants towards the overall population mean, which will reduce the variance in per-capita reproductive success. We thus also considered a second model in which we did not pool to the population-level. In this model, we instead estimated seedling survival, fruits per plant, and seeds per fruit each year separately and did not include a population-level effect (in other words, we did not nest year in population). This would have the effect of letting the prior have a stronger effect each year. We conducted model checks for both of these. Years without data would be missing or true NAs. Finally, we could also consider a model without pooling and in which the observed estimate is uncorrected.

Correlation between germination probability and seed survival

Increased seed survivorship is predicted to decrease the optimal germination probability (Cohen (1966); Ellner (1985*a*)). I assessed whether the observed germination probability was negatively correlated with seed survival (Gremer and Venable (2014)). I calculated the probability that seeds which do not germinate in January remain in the seed bank until the

following January (s_2s_3). I obtained the posterior distribution for the correlation between germination and seed survival by calculating the correlation of g_1 and s_2s_3 at each iteration of the MCMC output (Hobbs and Hooten 2015, p 194-5). Calculating the sample correlation in draws from the posterior. (quote Stan list-serv). Results of this analysis are shown in Figure 8. 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. The bottom panel shows the posterior distribution of correlation between the probability of germination and seed survival. The median correlation is negative (-0.07) and the 95% credible interval overlaps 0.

Ref: <https://discourse.mc-stan.org/t/computing-correlations-from-the-posterior/2633>

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

Increased variance in per-capita reproductive success is predicted to decrease the optimal germination probability (Cohen (1966); Ellner (1985a)). I assessed whether the observed germination probability was negatively correlated with variance in per-capita reproductive success (Venable (2007)). Per-capita reproductive success F_{jk} at population j in year k was calculated as:

$$F_{jk} = \phi_{jk} \times \lambda_{jk}^F \times \lambda_{jk}^P \quad (8)$$

where

$$\begin{aligned} \phi_{jk} &= \text{logit}^{-1}(\alpha_{0,j}^S + \beta_{jk}^S) \\ \lambda_{jk}^F &= \exp(\alpha_{0,j}^F + \beta_{jk}^F) \\ \lambda_{jk}^P &= \exp(\alpha_{0,j}^P + \beta_{jk}^P) \end{aligned} \quad (9)$$

To calculate the temporal variation in per-capita reproductive success for each population, I sampled the posterior distribution of reproductive success for each year and calculated the geometric SD of per capita reproductive success. I obtained the posterior distribution for the correlation between germination and geometric SD of per capita reproductive success at each iteration of the MCMC output [Hobbs and Hooten 2015, p 194-5](#). [Calculating the sample correlation in draws from the posterior](#). (quote Stan list-serv). Using this approach, I obtained a distribution of correlation estimates. 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. Results of this analysis are shown in Figure 9.

Density-independent model for germination probability

We use estimates of seed survival and reproductive success to investigate the adaptive value of delayed germination ([Gremer and Venable \(2014\)](#)). We parameterize a model of population growth rate and calculate the optimal germination strategy for different combinations of seed survival and reproductive success. We describe *C. xantiana*'s life cycle and calculate population growth rate with the equation:

$$\lambda = g_1 Y(t) s_0 s_1 + (1 - g_1) s_2 s_3 \quad (10)$$

The parameters in this equation were fit in models corresponding to equations (??), (??), and (??). 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 (equation (8)). Temporal variation is incorporated into the model by varying the per-capita reproductive success, $Y(t)$, between years.

For each population, I numerically calculate the optimal germination probability for

the observed variation in reproductive success and seed survival. In each case, I sampled parameters in the equation for density-independent growth (equation (10)) from the fitted vital rate distributions. I resampled the fitted values of per-capita reproductive success ($Y(t)$) to obtain a sequence of 1000 years. I used this same sequence of $Y(t)$ and the fitted 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 maximizes the geometric mean of the population growth rate. I repeat the simulations 50 times for each population, resampling the sequence of per-capita reproductive success, $Y(t)$, each time. I then calculated the mean of the optimal germination fractions.

Models in which per-capita reproductive success is density-independent predict that germination probability should respond to variance in fitness (Cohen (1966)). To evaluate the density-independent model, I compared modeled germination probabilities to predicted germination optima. I plot this comparison in Figure 10. 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.

Density-dependent model for germination fraction

consequences of a germination strategy for an individual's fitness depend on the strategies being used by other individuals in the population (Gremer and Venable 2014).

Here is what I think the general strategy would be if I wanted to test this model. Because we don't have data on all the species in the plot, we are solely focusing on the strength of intraspecific competition, which may vary with regards to how good of a proxy for over-all competition the seedling experiences. In years with high grass germination there may be strong competition from other species (for example). Our data on seeds per fruit comes

431 from haphazard collections of fruits; there is no information in density-dependence in this
432 estimate. Our data on fruits per plant could be informed by the number of seedlings or
433 adult plants in the plot. However, then we are getting one step removed from where the
434 competition happens (among seedlings). We thus start by estimating seedling survival to
435 fruiting as a function of density, assuming that this is the stage at which density-dependence
436 is strongest.

437 -We use counts of seedlings in the plot to incorporate the number of seedlings into our
438 model for seedling survival to fruiting. We assume 'low density' is a single plant, and so
439 obtain an estimate for the probability of survival at low densities as the marginal posterior
440 probability of survival in a plot with a single plant. In the end, we obtain an estimate of
441 seedling survival per fruit at low density in a given year (K in Gremer and Venable) and a
442 competition coefficient (from our logistic regression).

Age-structured model for germination fraction

443 Valleriani and Tielborger build on Tuljaparkur and Easterling/Ellner to show that an age-
444 structured seed bank can modify the expectations for how dormancy should evolve. Here,
445 we would want to show that there is (1) age-structure in the seed bank and (2) that the
446 importance of age structure varies across the species range.

Results

Correlation between germination probability and seed survival

447 I examined the correlation between germination probability and seed survival in the seed
448 bank. Results of this analysis are shown in Figure 8. The bottom panel shows the posterior
449 distribution of correlation between modeled germination probability and the probability of
450 seed survival; the median correlation is negative (-0.07) and the 95% credible interval overlaps

451 0. There is no correlation between germination and seed survival.

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

452 I examined the correlation between germination probability and variance in per-capita repro-
453 ductive success. Results of this analysis are shown in Figure 9. The bottom panel shows the
454 posterior distribution of correlation between modeled germination probability and geometric
455 SD in per-capita reproductive success; the median correlation is positive (0.28) and the 95%
456 credible interval overlaps 0.

Optimal germination probability predicted by a density-independent model

457 ...

Discussion

458 ...

Figures

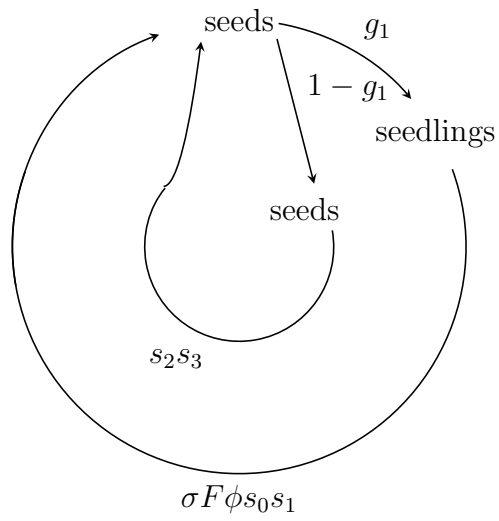


Figure 5: Life cycle diagram for *Clarkia xantiana*. I would like to draw a better life cycle diagram.

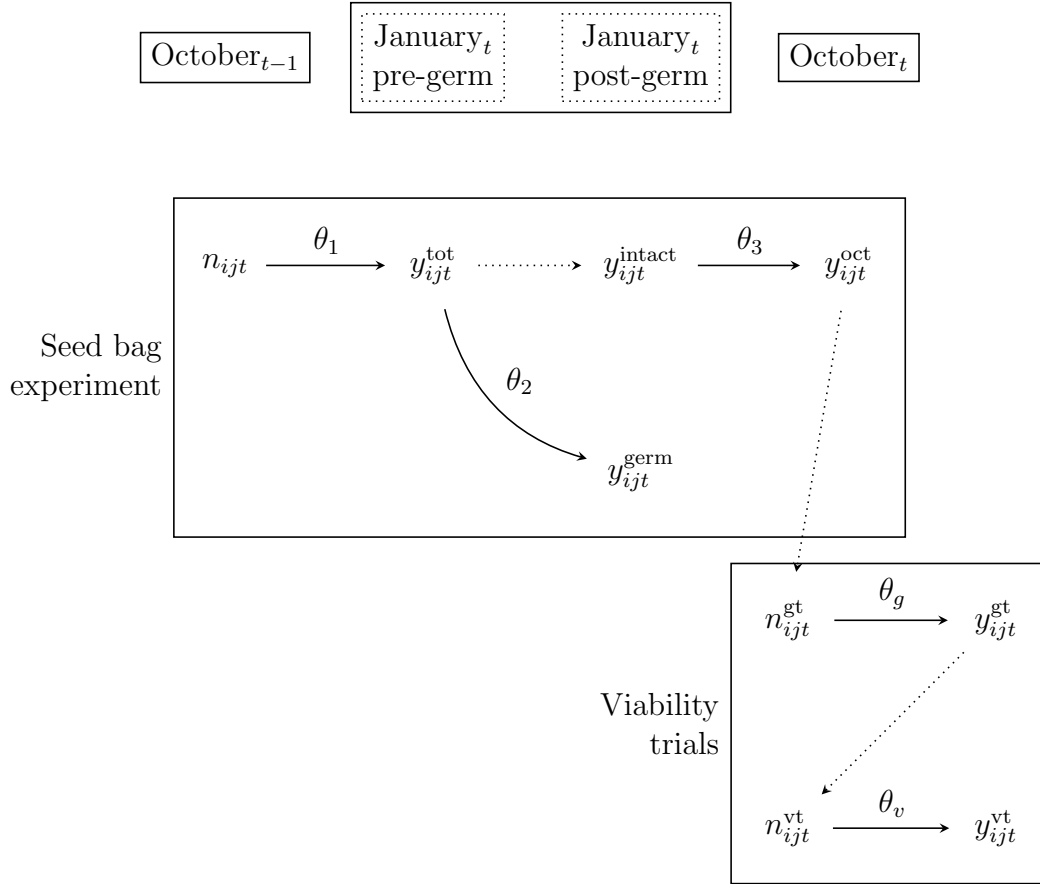


Figure 6: Diagram of data from the seed bag experiments and viability trials. There are two boxes: one for the seed bag experiment and one for the viability trials. In the seed bag experiment, I split January into two steps, one for just before germination and one for just after. Solid arrows represent probabilities estimated with a binomial experiment and are labeled with corresponding parameters. Dotted arrows represent cases where the seeds at the head of the arrow include some, possibly all, seeds at the tail of the arrow.

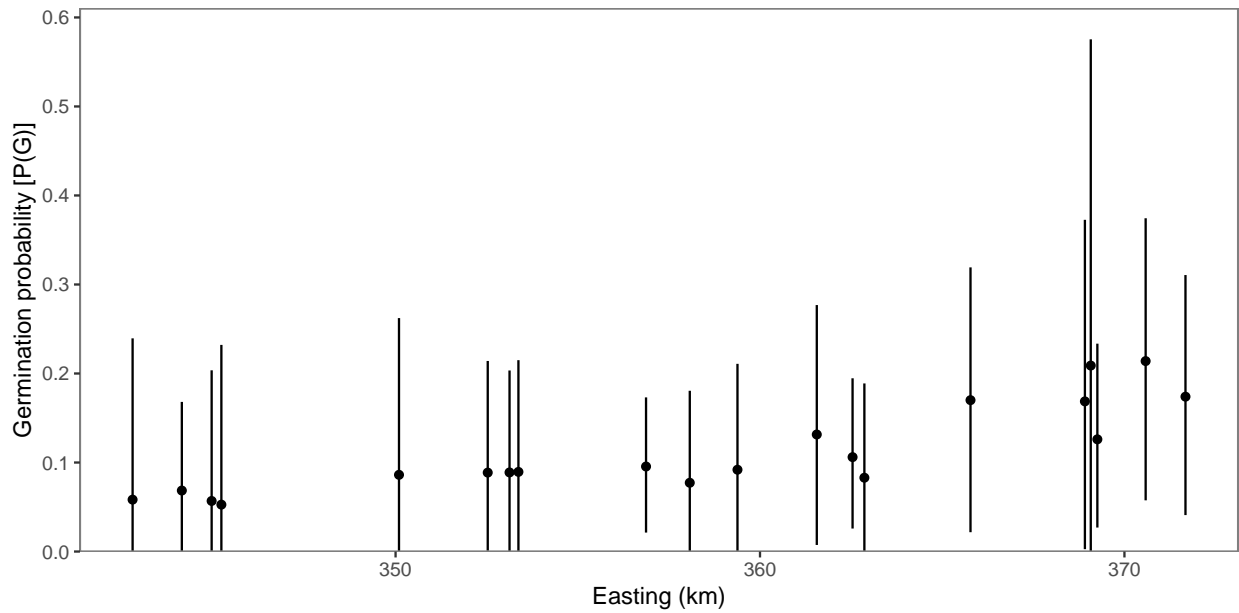


Figure 7: Germination probability plotted against easting (km). The plot shows the marginal posterior distribution for germination probability at each site. The points are the median of the posterior. The thinner line represents the 95% credible interval and the thicker line represents the 50% credible interval.

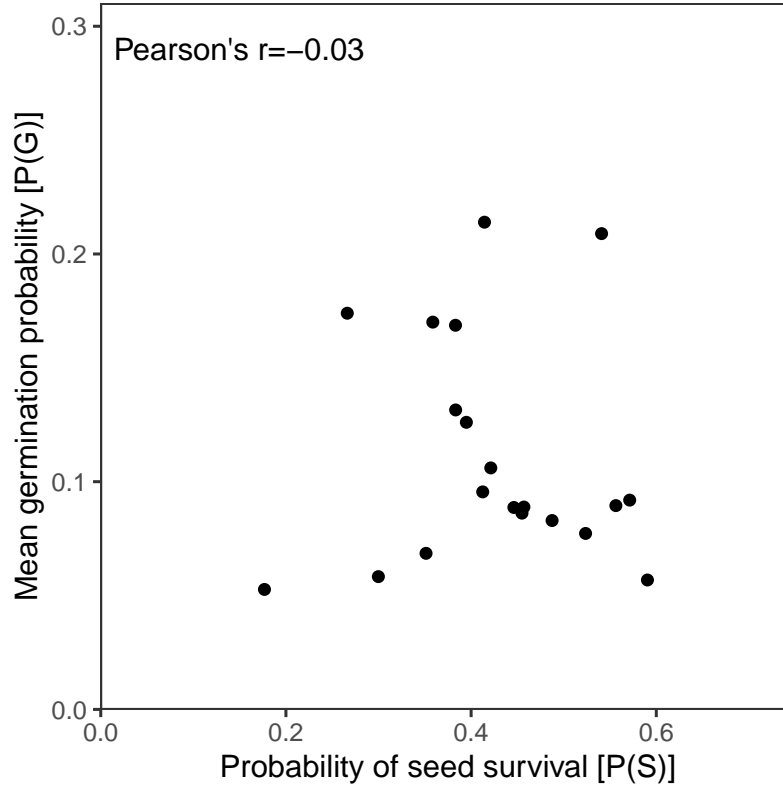


Figure 8: The top panel shows the observed germination probability plotted against probability of seed survival. The bottom panel shows the posterior distribution of correlation between observed germination probability and the probability of seed survival; the correlation is negative (-0.03).

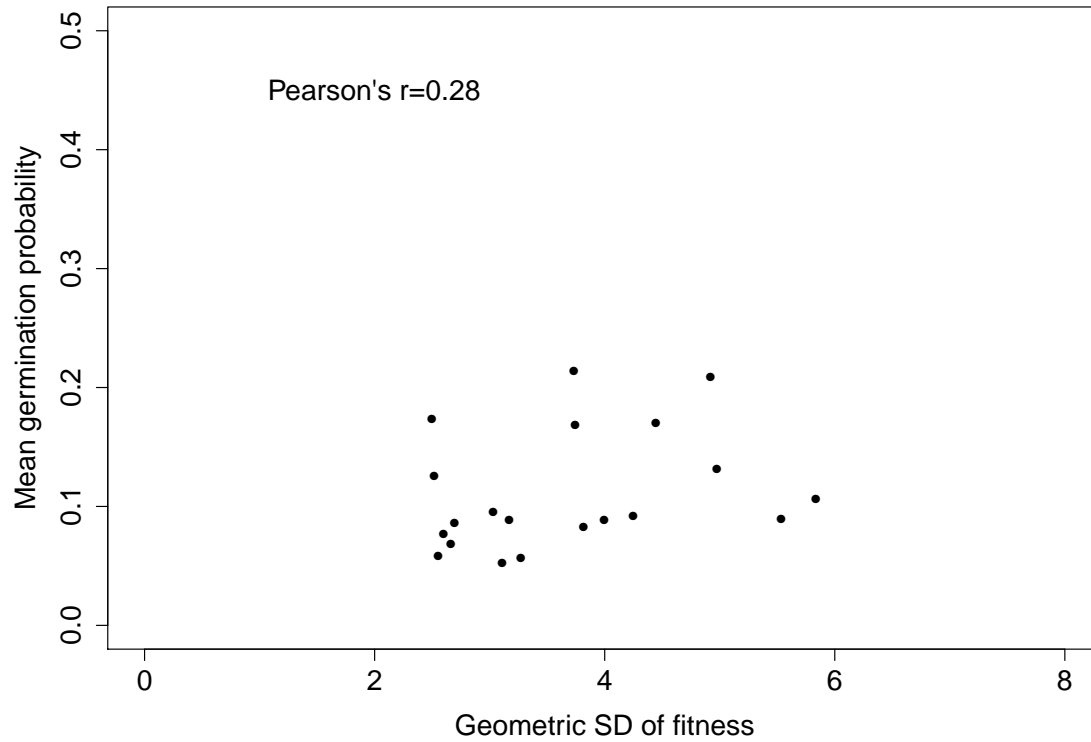


Figure 9: The top panel shows the observed germination probability plotted against the temporal variation in per capita reproductive success. The bottom panel shows the posterior distribution of correlation between observed germination probability and geometric SD of per capita reproductive success; the median correlation is negative (-0.16) but the 95% credible interval overlaps 0.

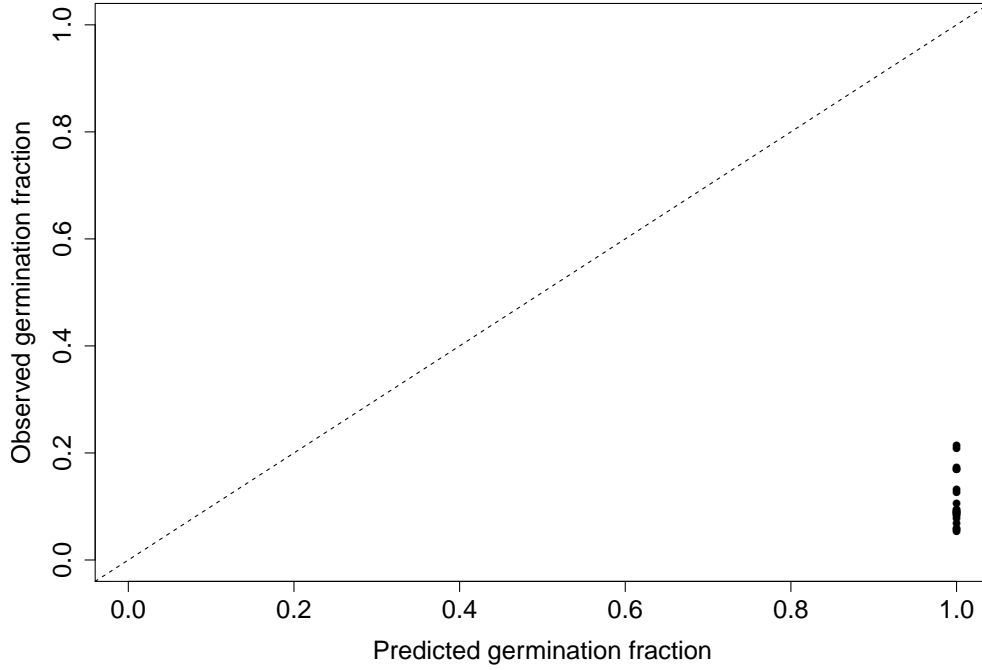


Figure 10: Observed germination probability plotted against the optimal germination probability predicted by a density-independent model. For each population, the observed germination probability is the obtained from the model for seed bank vital rates. Each point is the population-specific median of 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

- Cohen, D. 1966. Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology*, **12**:119–129.
- Cohen, D. 1967. Optimizing Reproduction in a Randomly 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.
- Cohen, D. 1968. A General Model of Optimal Reproduction in a Randomly Varying Environment. *The Journal of Ecology*, **56**:219.
- Eager, E. A., R. Rebarber, and B. Tenhumberg. 2014. 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**:1809–1834.
- Easterling, M. R. and S. P. Ellner. 2000. Dormancy strategies in a random environment: Comparing structured and unstructured models. *Evolutionary Ecology Research*, **2**:387–407.
- Eckhart, V. M., M. A. Geber, W. F. Morris, E. S. Fabio, P. Tiffin, and D. A. Moeller. 2011. The Geography of Demography: Long-Term Demographic Studies and Species Distribution Models Reveal a Species Border Limited by Adaptation. *The American Naturalist*, **178**:S26–S43.
- Ellner, S. 1985*a*. ESS germination strategies in randomly varying environments. I. Logistic-type models. *Theoretical Population Biology*, **28**:50–79.
- Ellner, S. 1985*b*. ESS germination strategies in randomly varying environments. II. Reciprocal Yield-Law models. *Theoretical Population Biology*, **28**:80–116.

480 Evans, M. E. K. and J. J. Dennehy. 2005. Germ Banking: Bet-Hedging and Variable Release
481 From Egg and Seed Dormancy. *The Quarterly Review of Biology*, **80**:431–451.

482 Gremer, J. R., S. Kimball, and D. L. Venable. 2016. Within-and among-year germination
483 in Sonoran Desert winter annuals: bet hedging and predictive germination in a variable
484 environment. *Ecology Letters*, **19**:1209–1218.

485 Gremer, J. R. and D. L. Venable. 2014. Bet hedging in desert winter annual plants: optimal
486 germination strategies in a variable environment. *Ecology Letters*, **17**:380–387.

487 Heinrich, L., J. Müller, A. Tellier, and D. Živković. 2018. Effects of population- and seed
488 bank size fluctuations on neutral evolution and efficacy of natural selection. *Theoretical*
489 *Population Biology*, **123**:45–69.

490 Mccue, K. A. and T. P. Holtsford. 1998. Seed bank influences on genetic diversity in the
491 rare annual *Clarkia springvillensis* (Onagraceae). *American Journal of Botany*, **85**:30–36.

492 Nunney, L. 2002. The Effective Size of Annual Plant Populations: The Interaction of a Seed
493 Bank with Fluctuating Population Size in Maintaining Genetic Variation. *The American*
494 *Naturalist*, **160**:195.

495 Paniw, M., P. F. Quintana-Ascencio, F. Ojeda, and R. Salguero-Gómez. 2017. Accounting for
496 uncertainty in dormant life stages in stochastic demographic models. *Oikos*, **126**:900–909.

497 Ritland, K. 1983. The joint evolution of seed dormancy and flowering time in annual plants
498 living in variable environments. *Theoretical Population Biology*, **24**:213–243.

499 Venable, D. L. 2007. Bet hedging in a guild of desert annuals. *Ecology*, **88**:1086–1090.

500 Waples, R. S. 2006. Seed Banks, Salmon, and Sleeping Genes: Effective Population Size
501 in Semelparous, Age-Structured Species with Fluctuating Abundance. *The American*
502 *Naturalist*, **167**:118.

Supplementary material

Theoretical background for hypotheses.

Explanation of key papers that develop theoretical results about seed banks. The document describes results from these papers that are relevant to understanding and interpreting the data in this manuscript. Link to document: <https://github.com/gregor-fausto/clarkiaSeedBanks/blob/master/products/appendices/appendix-cohen-results/appendix-x-cohen.pdf>

Data summary.

Summary tables for all datasets used in the manuscript. The document summarizes the types of data collected. The document provides a table summarizing each dataset (e.g. sample size per each site and year). Link to document: <https://github.com/gregor-fausto/clarkiaSeedBanks/blob/master/products/tables/data-summary.pdf>

Data processing workflow.

Description of workflow for processing the data used in the analysis. The document describes how comma-separated value (.csv) and Excel (.xls and .xlsx) files were read and processed in R. Link to document: <https://github.com/gregor-fausto/clarkiaSeedBanks/blob/master/library/dataProcessingWorkflow.md>

Method for estimating seed bank parameters using conditional probabilities.

The document explains how we compose conditional probabilities to calculate probabilities of survival and germination of seeds in the seed burial experiment. Link to document: <https://>

518 github.com/gregor-fausto/clarkiaSeedBanks/blob/master/products/appendices/appendix-conditional-probability.pdf

519 [appendix-x-conditional-probability.pdf](https://github.com/gregor-fausto/clarkiaSeedBanks/blob/master/products/appendices/appendix-conditional-probability.pdf)