Last updated: March 12, 2021

Overview

We first describe the experimental design used to estimate belowground vital rates and illustrate how we combine estimates from a seed bag burial experiment and viability trials to obtain parameters for a structured population model (Section 1). We use a conceptual diagram to connect the experimental design and data (Figure 6), and describe how estimates from the data are used to compose parameters for a structured population model (Tables 1 and 2).

We then describe the data and model construction for the seed bag experiment (Section 2). This section also includes the directed acyclic graphs (DAGs) for any likelihoods associated with the model construction. The section also describes how we estimated survival from reproduction to October when the seed bag experiments started.

We then describe the data and model construction for the viability trials and again include the DAGs (Section 3).

Finally, Section 4 describes how we use the parameter estimates from Section 2 and 3 to obtain the parameters for the structured population model from Section 1.

Note to self about how to turn this into Methods. The key pieces are Figure 6, columns 1, 3, and 6 of Table 1 and columns 1 & 3 in Table 2. Figure 6 illustrates the relationship between the seed bag burial experiment, viability trials, and survival/germination probabilities. Table 1 gives the mathematical relationship among the components of Figure 6. Table 2 translates the survival and germination components to parameters for the structured population model. I would to pare this down I would pare down Section 1 to describe the experiment and reduce the tables to the relevant parameters, likely combining them into 1 table (top section being the survival functions, bottom section being the model parameters). I would keep the description of the full model and the viability models. I would move any discussion of how the models were constructed and DAGs to a separate document. I would reduce the section on translating estimated parameters to model parameters to generally describe how this was done and move the rest to a separate document.

1 Experimental design and parameters

We used a field experiment in which we buried known numbers of seeds to estimate the probability of germination and survival for seeds of different ages (Figure 6A, gray panel). Buried bags were unearthed two times each during their first, second, or third year. Bags that were dug up in the first year were only used to count intact seeds during the first year; at the end fo the year bags were removed fro the field. The seed bag burial experiment was initiated in 3 consecutive years (3 rounds) at the 20 study populations. In round 1, bags were dug up in year 1, 2, and 3. In round 2, bags were dug up in year 1 and 2. In round 3, bags were dug up in year 1. There were thus 3 sets of counts for age 1 seeds, 2 sets of counts for age 2 seeds, and 1 set of counts for age 3 seeds.

In summary, during each experimental round we collected data on the number of seeds in the soil seed bank for up to 3 years. These estimates correspond to the number of seeds that remain in the soil seed bank. We thus define a persistent seed as one that remains intact in the seed bags (i.e. was countable in the seed bags); seeds were lost through physical destruction, decay, or germination. In (Figure 6A), points represent the average seed numbers associated with hypothetical counts in seed bags across an experimental round.

At the end of each experimental year, bags were brought to the lab and intact seeds were tested in a two-stage viability trial (Figure 6B, gray panel). In the lab, we conducted germination trials and viability assays on subsets of the seeds from each bag to estimate the viability of the 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. [Eckhart et al. 2011: 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 population 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 conducted viability trials in each year we conducted seed burial experiments.

We tested the viability of seeds in October, and were thus able to estimate the proportion of viable seeds (Figure 6B; filled points). We further 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 6B; open points).

We used the data from the seed bag burial experiment and viability trials to estimate age-specific seed survival and germination. We start with survival. To construct our estimates of seed survival, we first describe the discrete survival functions, probability mass functions, and hazards (e.g. Klein and Moeschberger 2003) associated with the persistence and viability of seeds in the soil seed bank (Table 1). We define a persistent seed as one that remains intact in the seed bags and a viable seed as one that is intact and viable (Thompson et al. 2003?).

Table 1: Seed persistence and viability in the soil seed bank

		Persistence			Viability		
Time	x_i	$S(x_i)$	$p(x_i)$	$h(x_j)$	$S(x_i)$	$p(x_i)$	$h(x_j)$
Oct_0	x_0	θ_0	1	—	$\phi_0 = \theta_0$	1	—
$\text{Jan}_{1,\text{total}}$	x_1	θ_1	$1-\theta_1$	$\frac{1-\theta_1}{1}$	$\phi_1 = \theta_1(\gamma_1 + (1-\gamma_1)\nu_1^{1/3})$	$1-\phi_1$	$\frac{1-\phi_1}{1}$
$\text{Jan}_{1,\text{intact}}$	x_2	θ_2	$\theta_1-\theta_2$	$\frac{\theta_1-\theta_2}{\theta_1}$	$\phi_2 = \theta_2\nu_1^{1/3}$	$\phi_1-\phi_2$	$\frac{\phi_1-\phi_2}{\phi_1}$
Oct_1	x_3	θ_3	$\theta_2-\theta_3$	$\frac{\theta_2-\theta_3}{\theta_2}$	$\phi_3 = \theta_3\nu_1$	$\phi_2-\phi_3$	$\frac{\phi_2-\phi_3}{\phi_2}$
$\text{Jan}_{2,\text{total}}$	x_4	θ_4	$\theta_3-\theta_4$	$\frac{\theta_3-\theta_4}{\theta_3}$	$\phi_4 = \theta_4(\gamma_2 + (1-\gamma_2)\nu_1(\nu_2/\nu_1)^{1/3})$	$\phi_3-\phi_4$	$\frac{\phi_3-\phi_4}{\phi_3}$
$\text{Jan}_{2,\text{intact}}$	x_5	θ_5	$\theta_4-\theta_5$	$\frac{\theta_4-\theta_5}{\theta_4}$	$\phi_5 = \theta_5\nu_1(\nu_2/\nu_1)^{1/3}$	$\phi_4-\phi_5$	$\frac{\phi_4-\phi_5}{\phi_4}$
Oct_2	x_6	θ_6	$\theta_5-\theta_6$	$\frac{\theta_5-\theta_6}{\theta_5}$	$\phi_6 = \theta_6\nu_2$	$\phi_5-\phi_6$	$\frac{\phi_5-\phi_6}{\phi_5}$
$\text{Jan}_{3,\text{total}}$	x_7	θ_7	$\theta_6-\theta_7$	$\frac{\theta_6-\theta_7}{\theta_6}$	$\phi_7 = \theta_7(\gamma_3 + (1-\gamma_3)\nu_2(\nu_3/\nu_2)^{1/3})$	$\phi_6-\phi_7$	$\frac{\phi_6-\phi_7}{\phi_6}$
$\text{Jan}_{3,\text{intact}}$	x_8	θ_8	$\theta_7-\theta_8$	$\frac{\theta_7-\theta_8}{\theta_7}$	$\phi_8 = \theta_8\nu_2(\nu_3/\nu_2)^{1/3}$	$\phi_7-\phi_8$	$\frac{\phi_7-\phi_8}{\phi_7}$
Oct_3	x_9	θ_9	$\theta_8-\theta_9$	$\frac{\theta_8-\theta_9}{\theta_8}$	$\phi_9 = \theta_9\nu_3$	$\phi_8-\phi_9$	$\frac{\phi_8-\phi_9}{\phi_8}$

The discrete survival function for persistent seeds (Table 1) represents seeds that do not leave the soil seed bank, either through physical destruction or germination, and summarizes the survival at marked intervals in the graph in (Figure 6A). The discrete survival function for viable seeds (Table 1) represents seeds that do not leave the soil seed bank through physical destruction, germination, or physiological death. To obtain the second survival function, we multiplied the probability of persistence at each time by the estimated or inferred viability. We also calculated the proportion of viable seeds before germination in January by assuming that all seedlings were viable before germination. The survival function for persistent seeds sets an upper bound for the survival function of viable seeds. Lower estimates of viability decrease the survival function of viable seeds, with effects that compound over time (Figure 6D).

The survival function for viable seeds (ϕ) is composed of estimates of persistence over time ($\theta.$), estimates of viability ($\nu.$), and estimates of germination conditional on persistence ($\gamma.$). We used the survival function and germination probabilities to define the parameters in the structured population model for *Clarkia xantiana* ssp. *xantiana*. Table 2 defines the age-specific germination probabilities and survival probabilities for the Leslie matrix in Eckhart et al. 2011 in terms of the survival function (Table 1) and germination probabilities.

Table 2: Structured model parameters conditional on persistence and viability

Parameter	Probability statement	Probability
g_1	$P(G_1 \| S_1)$	γ_1/ϕ_1
g_2	$P(G_2 \| S_3, S_2, G_1^c, S_1)$	γ_2/ϕ_4
g_3	$P(G_3 \| S_5, S_4, G_2^c S_3, S_2, G_1^c, S_1)$	γ_3/ϕ_7
s_1	$P(S_1)$	ϕ_1
s_2	$P(S_2 \| G_1^c, S_1)$	ϕ_3/ϕ_2
s_3	$P(S_3 \| S_2, G_1^c, S_1)$	ϕ_4/ϕ_3
s_4	$P(S_4 \| G_2^c S_3, S_2, G_1^c, S_1)$	ϕ_6/ϕ_5
s_5	$P(S_5 \| S_4, G_2^c S_3, S_2, G_1^c, S_1)$	ϕ_7/ϕ_6
s_6	$P(S_6 \| G_3^c, S_5, S_4, G_2^c S_3, S_2, G_1^c, S_1)$	ϕ_9/ϕ_8

Figure 6C illustrates the relationship among the pieces discussed here. Estimates of germination from the seed bag experiment correspond to the probability of germination conditional on persistence (e.g. γ_1). Multiplying these estimates by the probability of persistence in January before germination gives the unconditional probability of germination (e.g. $\theta_1 \times \gamma_1$). Finally, the probability of germination conditional on viability is estimated by incorporating loss of viability into the survival function (e.g. γ_1/ϕ_1).

2 Seed bag burial experiments

2.1 Data

We collected the following data from the seed bag burial experiments:

- $n_{\mathrm{g},ijkl}$ = observed count of intact seeds and seedlings in the seed bags in January in the i^{th} bag, from the j^{th} population, from the k^{th} experimental year, assumed to be measured perfectly

- $y_{\mathrm{g},ijkl}$ = observed count of germinated seedlings in the seed bags in January in the i^{th} bag, from the j^{th} population, from the k^{th} experimental year, for seeds of age l, assumed to be measured perfectly

- n_{ijk} = observed count of seeds in the seed bags at the start of the experiment in October in the i^{th} bag, from the j^{th} population, from the k^{th} experimental year, assumed to be measured perfectly

- y_{ijk}= observed count of ungerminated, total seeds in the seed bags in January (ungerminated seeds plus seedlings) and October in the i^{th} bag, from the j^{th} population, from the k^{th} experimental year, assumed to be measured perfectly

We started these seed burial experiments in three subsequent years (2005, 2006, 2007) to obtain multiple counts to estimate for seed survival and germination. These are two sets of data; the data with subscript g are used to estimate conditional germination and the data without a subscript are used to estimate survival.

2.2 Models

I build the models for each dataset before linking them. First, I describe a model for the age-specific, conditional germination probability. Second, I describe a model for the probability of seed survival.

2.2.1 Conditional germination

The joint likelihood for the germination model is given by

$$
\begin{aligned}
[\boldsymbol{\mu}_{\mathrm{g}}, \boldsymbol{\sigma}_{\mathrm{g}}, \boldsymbol{\mu}_{\mathrm{g}}^{\mathrm{pop}}, \boldsymbol{\sigma}_{\mathrm{g}}^{\mathrm{pop}} | \mathbf{y}_{\mathrm{g}}] \propto & \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\prod_{l=1}^{L} \mathrm{binomial}(y_{\mathrm{g},ijkl} | n_{\mathrm{g},ijkl}, \mathrm{logit}^{-1}(\alpha_{\mathrm{g},ijkl})) \\
& \times \mathrm{normal}(\alpha_{\mathrm{g},ijkl} | \mu_{\mathrm{g},jkl}, \sigma_{\mathrm{g},jkl}) \\
& \times \mathrm{normal}(\mu_{\mathrm{g},jkl} | \mu_{\mathrm{g},jl}^{\mathrm{pop}}, \sigma_{\mathrm{g},jl}^{\mathrm{pop}}) \\
& \times \text{half-normal}(\sigma_{\mathrm{g},jkl} | 0, 1) \\
& \times \mathrm{normal}(\mu_{\mathrm{g},jl}^{\mathrm{pop}} | 0, 1) \text{half-normal}(\sigma_{\mathrm{g},jl}^{\mathrm{pop}} | 0, 1).
\end{aligned}
\tag{1}
$$

in which J indexes the 20 populations in the study; K indexes the 3 experimental years; L indexes the 3 ages; I indexes the observations within each set of data. The model has a binomial likelihood and logit-link for the latent probability of germination. The latent probability of germination at each age is described by two hierarchical levels: the first level is the experimental years and the second (upper) level is the population. The variance parameters have weakly informative, half-normal priors and the mean parameters have weakly informative, normal priors. See Appendix: Priors for details on the choice of weakly informative priors.

2.2.2 Survival

I represented the seed mortality process using a deterministic function, the Weibull survival function. Smits 2015 uses the following parameterization of the Weibull survival function:

$$S(t) = \exp(-(\frac{t}{\beta})^{\alpha}) \tag{2}$$

for which β is corresponds to the inverse of the scale parameter in the exponential distribution, accounting for the influence of the shape parameter. Formally, if λ is the scale parameter, $\lambda = \beta^{-\frac{1}{\alpha}}$ so that $\beta = 1/\lambda^{\alpha}$ Smits 2015, ProbOnto. To model the influence of group effects (random intercepts $\eta_i j$), $\beta_i j = \exp(-\frac{\eta_i j}{\alpha})$. This approach would also extend to incorporating covariates into the numerator to model their effect on the scale parameter.

Parameter alpha controls the shape of the decomposition trajectory and beta the rate of decomposition. This model can reduce to the exponential model when alpha is 1. If alpha is less than 1, the decomposition rate decreases through time. If alpha is greater than 1, the decomposition rate increases through time.

Counts of seeds were indexed by sites j, rounds of observation starting in a series of years k, and observation m. Here, the observation index tracks the time at which a particular observation. The joint likelihood for the survival model is then given by

$$
\begin{aligned}
\beta_{ijk} &= \exp(-\frac{\eta_{ijk}}{\alpha_j}) \\
g(\eta_{ijk}, \boldsymbol{\theta}, t_{ijkm}) &= \boldsymbol{\theta}\exp(-(\frac{t_{ijkm}}{\beta_{ijk}})^{\alpha_j}) \\
&= \boldsymbol{\theta}\exp(-(\frac{t_{ijkm}}{\exp(-\frac{\eta_{ijk}}{\alpha_j})})^{\alpha_j}) \\
[\boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\mu}^{\text{pop}}, \boldsymbol{\sigma}^{\text{pop}}|\mathbf{y}] &\propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\prod_{m=1}^{M} \text{binomial}(y_{ijkm}|n_{ijkm}, g(\boldsymbol{\theta}, \eta_{ijk}, \alpha_j, t_{ijkm})) \\
&\times \text{normal}(\eta_{ijk}|\mu_{jk}, \sigma_{jk}) \\
&\times \text{normal}(\mu_{jk}|\mu_j^{\text{pop}}, \sigma_j^{\text{pop}}) \\
&\times \text{half-normal}(\sigma_{jk}|0, 1) \\
&\times \text{normal}(\mu_j^{\text{pop}}|0, 1)\text{half-normal}(\sigma_j^{\text{pop}}|0, 1) \\
&\times \text{gamma}(\alpha_j|2, 2).
\end{aligned} \tag{3}
$$

The Weibull has a shape and scale parameter; we estimate a shape parameter for each population (α_j) Smits 2015. This is equivalent to assuming the rate of change in survivorship is a population-level property but that the scale varies from year to year within each population (η_{ijk}).

The survival data has to account for the loss of seeds due to mortality as well as germination. Because we model germination as conditionally dependent on the number of intact seeds at the time of germination, we estimate the germination rate independently. We combine that information with the survival time model. We use the survival function associated with a Weibull distribution as the deterministic function for how seed survivorship changes over time. In the likelihood above, I write the influence of germination as $\boldsymbol{\theta}$ and describe how this connects to the germination model next.

2.2.3 Germination histories

I represent the germination history of seeds in the experiment with distributions γ_{ya}, describing the probability that a seed has germinated. The subscript corresponds to seeds of age $a = \ldots$ and experimental year $y = \ldots$.

The experiment was initiated in three consecutive years (2006, 2007, 2008). Seed bags from the first year were collected at ages 1, 2, and 3. Seed bags from the second year were collected at ages 1 and 2. Seed bags from the third year were collected at age 3. There are thus the following probability distributions:

$$\begin{aligned} \gamma_{1a} &= (\gamma_{11}, \gamma_{12}, \gamma_{13}) \\ \gamma_{2a} &= (\gamma_{21}, \gamma_{22}) \\ \gamma_{3a} &= (\gamma_{31}) \end{aligned} \tag{4}$$

I used these probability distributions to summarize the event histories leading up to each count. In total, I defined twelve probabilities, one for each seed count. These probabilities are defined by age-specific germination rates and the experimental design, specifically the timing of when seeds were unearthed.

Seeds from each experimental year and age were collected twice. First, the seed bags were unearthed in the January as seeds were germinating. At this point, we counted intact seeds and emerging seedlings and summed these to get a count of total seeds just before germination. The seed bags were then returned to the ground from January to October. At this point, the seed bags were unearthed again and removed from the field to get counts of intact seeds.

Event histories 1-6 correspond to seed bags buried in experimental year 1; event histories 7-10 correspond to seed bags buried in experimental year 2; event histories 11-12 correspond to seed bags buried in experimental year 3. The probabilities are composed of germination probabilities for seeds of age 1 ($\gamma_{\cdot\cdot1}$), age 2 ($\gamma_{\cdot\cdot2}$), and age 3 ($\gamma_{\cdot\cdot3}$). The germination probabilities are indexed by site j, experimental year k, and age l; for example, germination of 1-year old seeds at site j from the experimental year k is given by γ_{jkl}. The compound event histories are defined by:

$$\begin{aligned}
&\theta_{j1,1} = 1\\
&\theta_{j1,2} = (1-\gamma_{j1,1})\\
&\theta_{j1,3} = (1-\gamma_{j1,1})\\
&\theta_{j1,4} = (1-\gamma_{j1,1}) \times (1-\gamma_{j1,2})\\
&\theta_{j1,5} = (1-\gamma_{j1,1}) \times (1-\gamma_{j1,2})\\
&\theta_{j1,6} = (1-\gamma_{j1,1}) \times (1-\gamma_{j1,2}) \times (1-\gamma_{j1,3})\\
&\theta_{j2,1} = 1\\
&\theta_{j2,2} = (1-\gamma_{j2,1})\\
&\theta_{j2,3} = (1-\gamma_{j2,1})\\
&\theta_{j2,4} = (1-\gamma_{j2,1}) \times (1-\gamma_{j2,2})\\
&\theta_{j3,1} = 1\\
&\theta_{j3,2} = (1-\gamma_{j3,1})
\end{aligned} \tag{5}$$

These compound event histories ($\boldsymbol{\theta}$) link the germination model to the seed survival model.

They enter the likelihood statement in the definition of the deterministic process model. The probability of survival to time t_{ijkm} is now multiplied by the probability that a seed has not germinated by that time. That probability is defined by the compound event history leading up to time t_{ijkm}. The index m thus corresponds tracks the observation time and/or event history associated with that observation time.

2.2.4 Linking germination and seed survival

In the full model linking germination and seed survival data, counts of seeds are indexed by sites j, rounds of observation starting in a series of years k, seed ages l, and event histories m. The likelihood for the model is then

$$g(\eta_{ijk}, \theta_{jkm}, t_{ijkm}) = \theta_{jkm} \times \exp(-(\frac{t_{ijkm}}{\exp(-\frac{\eta_{ijk}}{\alpha_j})})^{\alpha_j})$$

$$\begin{aligned}
[\boldsymbol{\mu}_{\mathrm{g}}, \boldsymbol{\sigma}_{\mathrm{g}}, \boldsymbol{\mu}_{\mathrm{g}}^{\mathrm{pop}}, \boldsymbol{\sigma}_{\mathrm{g}}^{\mathrm{pop}}, \boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\mu}^{\mathrm{pop}}, \boldsymbol{\sigma}^{\mathrm{pop}} | \mathbf{y}_{\mathrm{g}}, \mathbf{y}] \propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\Bigg[\Big[\prod_{m=1}^{M} & \mathrm{binomial}(y_{ijkm}|n_{ijkm}, g(\theta_{jkm}, \eta_{ijk}, \alpha_j, t_{ijkm})) \\
& \times \mathrm{normal}(\eta_{ijk}|\mu_{jk}, \sigma_{jk}) \\
& \times \mathrm{normal}(\mu_{jk}|\mu_j^{\mathrm{pop}}, \sigma_j^{\mathrm{pop}})\text{half-normal}(\sigma_{jk}|0,1) \\
& \times \mathrm{normal}(\mu_j^{\mathrm{pop}}|0,1)\text{half-normal}(\sigma_j^{\mathrm{pop}}|0,1) \\
& \times \mathrm{gamma}(\alpha_j|2,2)\Big] \\
& \Big[\times \prod_{l=1}^{L} \mathrm{binomial}(y_{\mathrm{g},ijkl}|n_{\mathrm{g},ijkl}, \mathrm{logit}^{-1}(\alpha_{\mathrm{g},ijkl})) \\
& \times \mathrm{normal}(\alpha_{\mathrm{g},ijkl}|\mu_{\mathrm{g},jkl}, \sigma_{\mathrm{g},jkl}) \\
& \times \mathrm{normal}(\mu_{\mathrm{g},jkl}|\mu_{\mathrm{g},jl}^{\mathrm{pop}}, \sigma_{\mathrm{g},jl}^{\mathrm{pop}})\text{half-normal}(\sigma_{\mathrm{g},jkl}|0,1) \\
& \times \mathrm{normal}(\mu_{\mathrm{g},jl}^{\mathrm{pop}}|0,1)\text{half-normal}(\sigma_{\mathrm{g},jl}^{\mathrm{pop}}|0,1).\Big]\Bigg]
\end{aligned} \tag{6}$$

2.3 Estimating seed survival from reproduction to October

The model described so far uses data from the seed bag experiment to estimate conditional germination and seed survival through time. However, we can't use this data to estimate the probability of seeds surviving from reproduction to the time the seed bag experiment started in October (s_0). What we do to try to get this value is the following.

We use the counts of fruits per plant conducted in the permanent plots to get a total number of fruits per plot. We then multiply this count by the average number of seeds per fruit. We then have an estimate of the number of seeds produced per plot. We do this for aboveground data in 2007 and 2008. We link this data to the number of seedlings observed in the same plots in the following year (2008 and 2009, respectively). We thus take the total number of seeds produced in year t as the number of trials in a binomial experiment for which the outcome is the number of seedlings observed in year $t+1$. The probability is the product of survival from reproduction to October (s_0, estimated here), survival from October to January, and germination. We have estimates of the latter two probabilities thanks to the seed bag experiments. We link the three components and estimate the remaining term s_0.

For the time being, we assume that the majority of seedlings in a plot emerge from seeds produced the previous year (read: germination of 1 year old seeds). With this assumption, we do not model germination from older seeds. We are also unable to assess the contribution of seeds from outside the plot to the total number of seeds available. Because the number of seedlings sometimes exceed the total number of seeds, we summed across transects to get counts for the number of seeds and observed seedlings.

In the this model, counts of seeds are indexed by sites j and years k. The likelihood for the model is then

$$
\begin{aligned}
[\boldsymbol{\mu}_{\mathrm{s}}, \boldsymbol{\sigma}_{\mathrm{s}}, \boldsymbol{\mu}_{\mathrm{s}}^{\mathrm{pop}}, \boldsymbol{\sigma}_{\mathrm{s}}^{\mathrm{pop}} | \mathbf{y}_{\mathrm{s}}] \propto \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} & \mathrm{binomial}(y_{\mathrm{g},ijk} | n_{\mathrm{g},ijk}, \boldsymbol{\theta}\boldsymbol{\eta}\mathrm{logit}^{-1}(\alpha_{\mathrm{s},ijk})) \\
& \times \mathrm{normal}(\alpha_{\mathrm{s},ijk} | \mu_{\mathrm{g},jk}, \sigma_{\mathrm{s},jk}) \\
& \times \mathrm{normal}(\mu_{\mathrm{s},jk} | \mu_{\mathrm{s},j}^{\mathrm{pop}}, \sigma_{\mathrm{s},j}^{\mathrm{pop}}) \\
& \times \text{half-normal}(\sigma_{\mathrm{s},jk} | 0, 1) \\
& \times \mathrm{normal}(\mu_{\mathrm{s},j}^{\mathrm{pop}} | 0, 1) \text{half-normal}(\sigma_{\mathrm{s},j}^{\mathrm{pop}} | 0, 1).
\end{aligned}
\tag{7}
$$

This likelihood can be augmented to the one for the seed bag burial experiment.

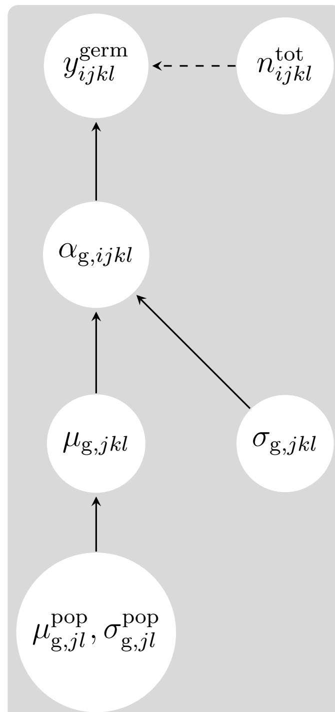
2.4 Simulations

I conducted simulations to check whether the model formulation for the seed bag experiment above was able to recover parameters of known values. I generated data similar to what we observed in the experiment but, for the time, ignoring the hierarchical structure of the data).

I simulated data with known combinations of the shape and scale parameters for the Weibull, age-dependent germination probabilities, and aboveground plot dynamics. I checked for convergence and recovered the simulated parameters. I also performed graphical posterior predictive checks and calculated Bayesian p-values to assess how the model might perform when fit to real data.

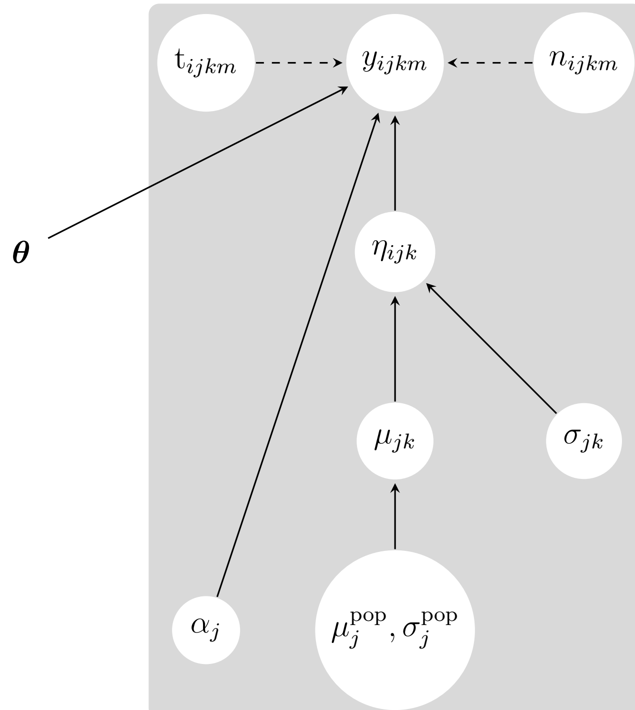
2.5 Directed acyclic graphs

2.5.1 Germination



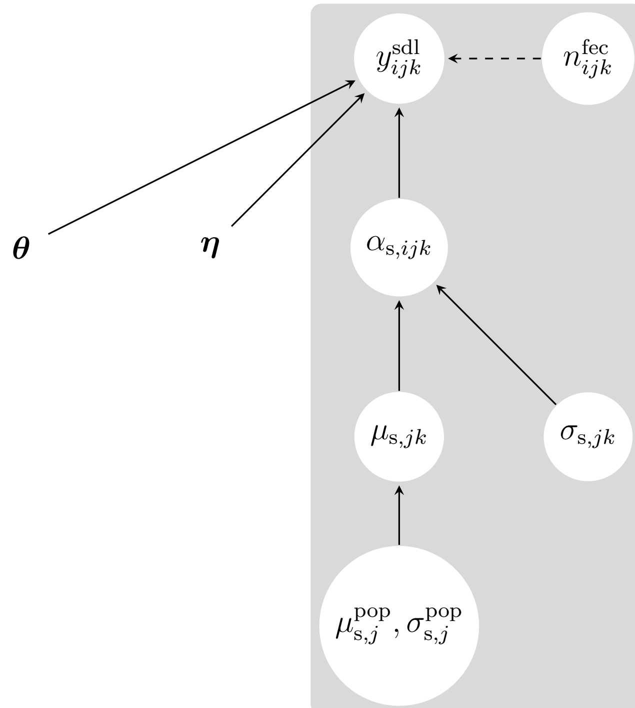
(a) Directed acyclic graph for the hierarchical model for conditional germination. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

2.5.2 Seed survival



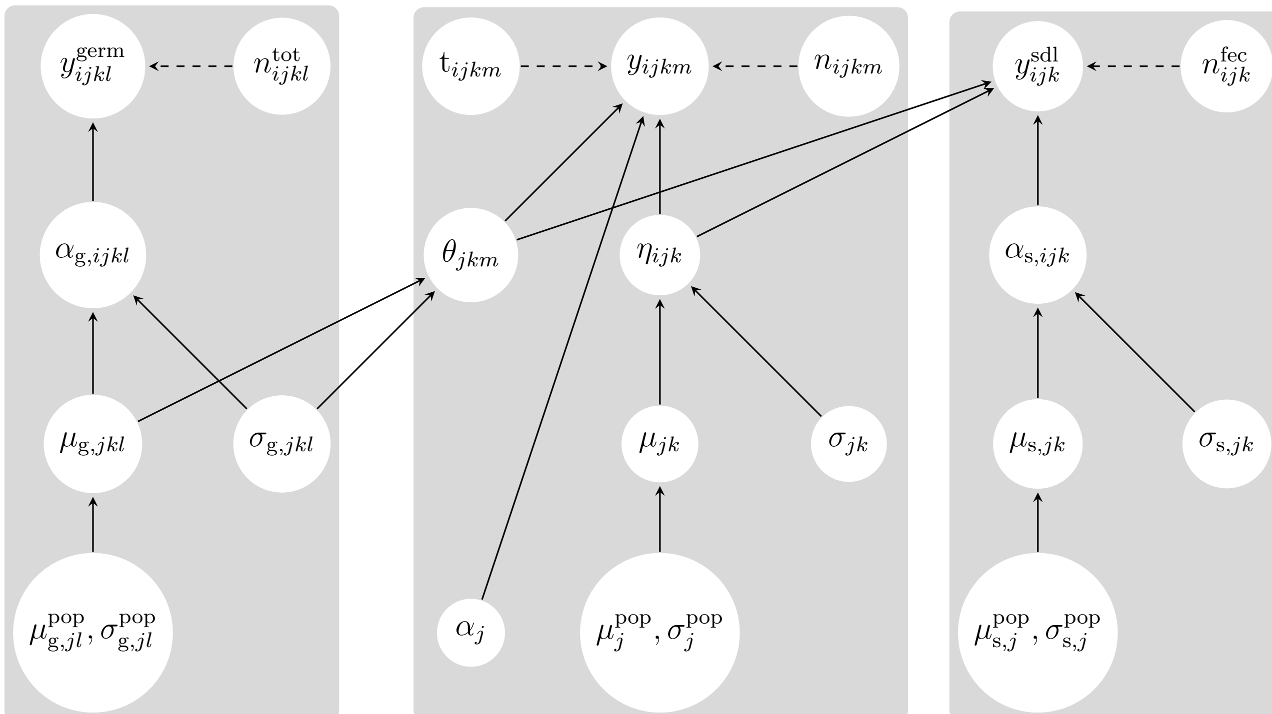
(a) Directed acyclic graphs for the models of seed survival. The parameter $\boldsymbol{\theta}$ links the conditional germination model to the seed survival model. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

2.5.3 Seed survival from June to October



(a) Directed acyclic graphs for the models of seed survival from reproduction to October. The parameter $\boldsymbol{\theta}$ links the model to the conditional germination model. The parameter $\boldsymbol{\eta}$ links the model to the seed survival model. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

2.5.4 Full model: germination, seed survivorship, initial seed survival



(a) Directed acyclic graphs for the full model. Models for each set of data are encapsulated in a gray box; links among the datasets are shown by arrows that cross over between boxes. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

3 Viability trials

3.1 Data

We collected the following data during the viability trials:

- n^{germ}_{ijkl} = observed count of seeds at the start of the germination trial for the i^{th} bag, from the j^{th} population, in the k^{th} experimental year, for seeds of age l, assumed to be measured perfectly

- y^{germ}_{ijk} = observed count of germinated seedlings in the germination trial for the i^{th} bag, from the j^{th} population, in the k^{th} year, for seeds of age l, assumed to be measured perfectly

- n^{viab}_{ijk} = observed count of seeds at the start of the viability trial for the i^{th} bag, from the j^{th} population, in the k^{th} year, for seeds of age l, assumed to be measured perfectly

- y^{viab}_{ijk} = observed count of viable seedlings in the viability trial for the i^{th} bag, from the j^{th} population, in the k^{th} year, for seeds of age l, assumed to be measured perfectly

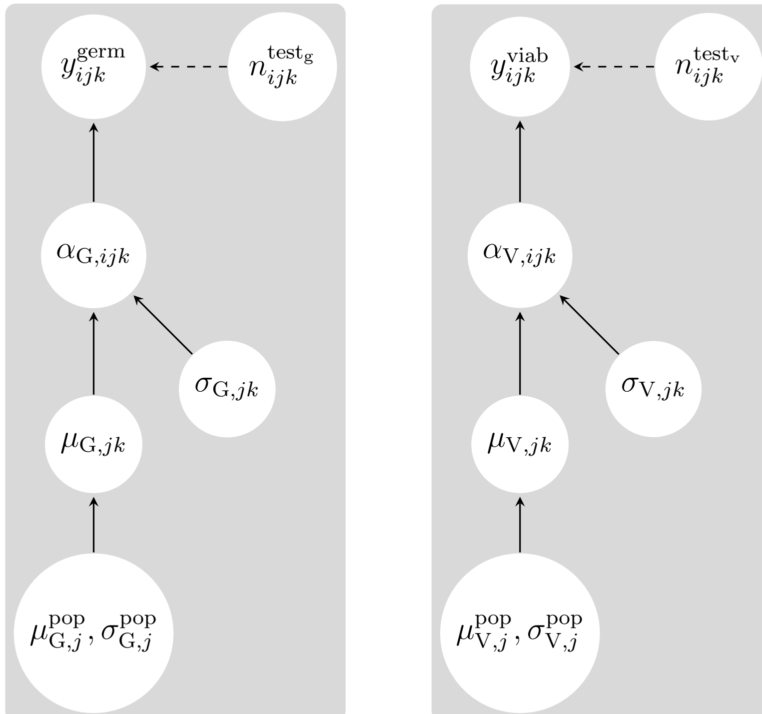
3.2 Model

All data from viability trials is in the form of binomial trials: we have counts of seeds at the start and end of an experimental window of time. All models have the same structure for seeds in bag i in population j in experimental year k. If the number of seeds starting the trial (trials) is n_{ijk} and the number of seeds at the end of the trial (successes) is y_{ijk}, we write a model that has a population-level mean and year-level means drawn from the population-level distribution. Broadly, this is two-level hierarchical model with a population-level mean, and year-level means drawn from the population-level distribution. The probability of success for each bag is drawn from this year- and population-level distribution. The model uses a binomial likelihood. The joint posterior for viability trials thus has the following general form:

$$
\begin{aligned}
[\boldsymbol{\alpha_G}, \boldsymbol{\mu_G}, \boldsymbol{\sigma_G}, \boldsymbol{\mu_G^{\mathrm{pop}}}, \boldsymbol{\sigma_G^{\mathrm{pop}}} | \boldsymbol{y^{\mathrm{tot}}}] \propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} & \operatorname{binomial}(y_{ijk}^{\mathrm{germ}} | n_{ijk}^{\mathrm{test_g}}, \operatorname{logit}^{-1}(\alpha_{G,ijk})) \\
& \times \operatorname{normal}(\alpha_{G,ijk} | \mu_{G,jk}, \sigma_{G,jk}) \\
& \times \operatorname{normal}(\mu_{G,jk} | \mu_{G,j}^{\mathrm{pop}}, \sigma_{G,j}^{\mathrm{pop}}) \\
& \times \text{half-normal}(\sigma_{G,jk} | 0, 1) \\
& \times \operatorname{normal}(\mu_{G,j}^{\mathrm{pop}} | 0, 1000) \text{half-normal}(\sigma_{G,j}^{\mathrm{pop}} | 0, 1).
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
[\boldsymbol{\alpha_V}, \boldsymbol{\mu_V}, \boldsymbol{\sigma_V}, \boldsymbol{\mu_V^{\mathrm{pop}}}, \boldsymbol{\sigma_V^{\mathrm{pop}}} | \boldsymbol{y^{\mathrm{tot}}}] \propto \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} & \operatorname{binomial}(y_{ijk}^{\mathrm{viab}} | n_{ijk}^{\mathrm{test_v}}, \operatorname{logit}^{-1}(\alpha_{V,ijk})) \\
& \times \operatorname{normal}(\alpha_{V,ijk} | \mu_{V,jk}, \sigma_{V,jk}) \\
& \times \operatorname{normal}(\mu_{V,jk} | \mu_{V,j}^{\mathrm{pop}}, \sigma_{V,j}^{\mathrm{pop}}) \\
& \times \text{half-normal}(\sigma_{V,jk} | 0, 1) \\
& \times \operatorname{normal}(\mu_{V,j}^{\mathrm{pop}} | 0, 1000) \text{half-normal}(\sigma_{V,j}^{\mathrm{pop}} | 0, 1).
\end{aligned} \tag{9}
$$

3.3 Directed acyclic graphs



(a) Directed acyclic graphs for the hierarchical models for viability trials. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

4 Estimated parameters to model parameters

We now use the parameter estimates from Section 2 and 3 to obtain the parameters for the structured population model from Section 1. Below, we describe how we obtained population-level estimates for the parameters in Table 2.

First, we use the full model for germination, seed survival, and initial seed survival to obtain population-level estimates for the survival functions for seed persistence and age-specific germination. We use the posterior distributions for the population-level mean of the inverse-scale parameter (μ_j^{pop}) and the shape parameter (α_j) to derive (semantics: marginalize to obtain?) the population-level parameters for the Weibull survival function. We transform the population-level posterior distributions for the survival probability from the latent scale to $[0, 1]$ by taking the inverse logit; this transforms the parameters into the probability of success.

We then use the posterior distributions for the population-level mean of the probability of germination on the latent scale ($\mu^{\text{pop}}_{\text{g},jl}$) to derive the age-specific germination probabilities. We transform the population-level posterior distributions for the germination probability from the latent scale to $[0, 1]$ by taking the inverse logit; this transforms the parameters into the probability of success.

We use the product of the Weibull survival function and the germination histories (as in Section 2) to obtain the survival functions associated with persistence of seeds in the soil seed bank ($\theta.$ in Table 1). At this point, we switch from continuous time to the discrete times in Table 1) by calculating the survival at those times.

We then estimate viability with the posterior distributions for the population-level mean of the probability of germination and viability from the viability trials. For estimates of viability in October, we transform the values from the latent scale ($\mu^{\text{pop}}_{\text{G},j}$, $\mu^{\text{pop}}_{\text{V},j}$) to the probability scale (θ_G, θ_V. We estimate the overall probability of viability, ν, as $\theta_g + \theta_v(1 - \theta_g)$. This weights the estimates relative to the probability of germination (eg. if no seeds germinate the estimate of viability will mostly come from the viability test). For estimates of viability in January, we interpolate between estimates of viability on the latent scale before transforming interpolated values to the probability scale. This allows us to retain the uncertainty associated with our estimates of viability. We transform the population-level posterior distributions for the overall probability of viability from the latent scale to $[0, 1]$ by taking the inverse logit; this transforms the parameters into the probability of success.

With all of these pieces in hand, we calculate values of ϕ, which are the probability of seed persistence and viability. We then estimate the probability of germination conditional on viability as in Table 2.

Figures

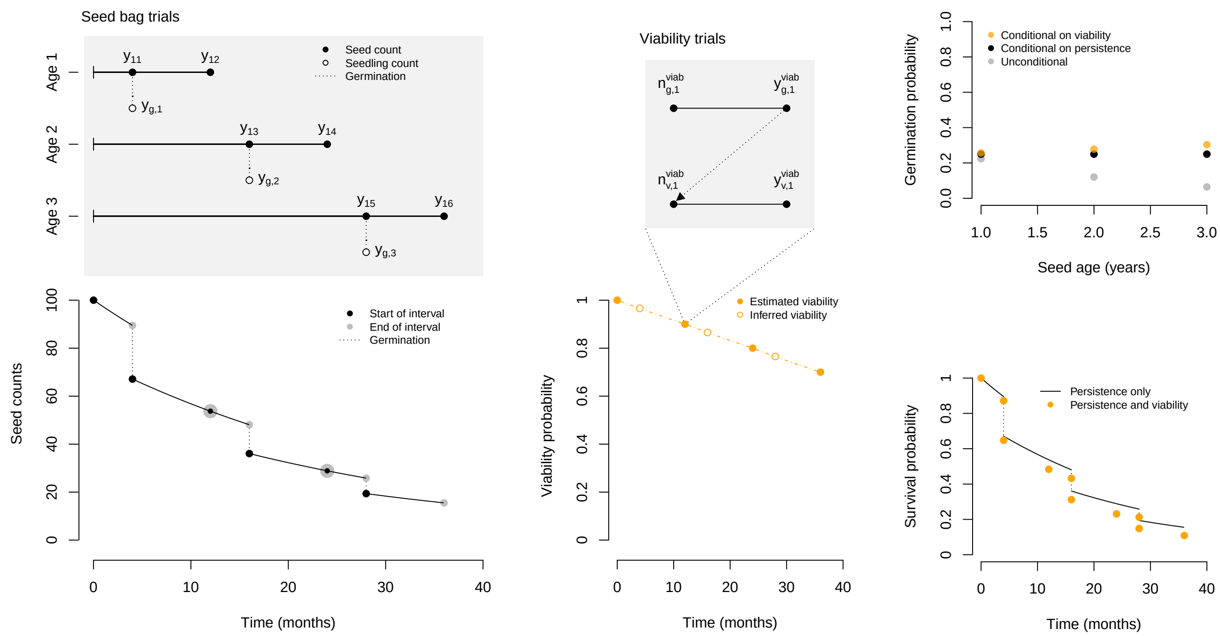


Figure 6: Summary of the seed bag burial experiments and viability trials. Figure will be labeled as (A: seed bag trials , B: viability trials, C: germination probability, D: survival probability. . (A) The gray panel contains a graphical representation of the seed bag trials. Seeds were buried at the start of each experiment (100 seeds in month 0). Seed bags were unearthed and intact seeds $(y_{\cdot\cdot})$ and germinants $(y_{g,\cdot}$ 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 probably 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).