Last updated: January 18, 2021

Seed banks present a challenge for models of plant population demography. Individual seeds can not be tracked and it is likely that there is greater uncertainty associated with seed bank vital rates. Ecologists have turned to various methods to assess survival and germination from the seed bank, including experimental burials and seed addition experiments.

The models below represent the joint likelihood for data from the seed bag experiments. All data from seed bags and 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 for the parameters $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$ have the same structure for seeds in bag i in year j in population 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. The probability of success for each bag is drawn from this year- and population-level distribution:

I compared convergence diagnostics (R-hat, effective sample size) for centered and non-centered parameterizations of the model. Here, I use the centered parameterization because this led to improved convergence. In each model, we obtain the population-level posterior distribution probability of success (the θs) by marginalizing across years and taking the inverse logit.

The seed bank experiments might used to obtain age-specific survival and recruitment terms. So the terms could be survival to month 3 (equal to age-specific survival at midpoint), germination at month 3, survival from month 3 to month 12 (equal to age-specific survival at midpoint), germination at month 15, survival from month 15 to 24, germination at month 27, survival from month 27 to 36. There are some errors in the data still, need to look at those. For germination there is no clear pattern to germination, either for individual sites or across sites. There are monotonic increases, monotonic decreases, and nonmonotonic patterns. Might also help to look at viability across age.

The seed pot experiments seem like they would be amenable to the kind of functions fit in the paper by Rees and Long. In those cases there are only 3 time points which would limit the application of the empirical seedling recruitment curves. But perhaps overlaying them per population would allow us to figure out whether there is consistency across years or whether particular cohorts follow different curves, and how this varies in space?

I will estimate survival/mortality rates at 4 time points (0-3 months, 3-12 months, 15-24 months, 27-36 months), and germination at 3 time points (3 months, 15 months, 27 months). This will allow me to figure out what kind of model (cf Rees and Long) might be most appropriate for the data, especially for the emergence data from the seed pot trials.

Table 1: Summary of models in Rees and Long (1993).

Model	Description
Exponential	...
Compound exponential	...
Weibull	...
Log logistic	...

I've now re-read Rees and Long (1993) and I think that's not the correct approach for the seed bag data. They present an analysis of recruited seedlings, fitting distributions to 5 years of data on emerged seedlings.

Instead, I think it would be useful to look at models for litter decomposition. I read Olson (1963) and Cornwell et al. (2014), which both present logic for analyzing litter decomposition experiments. Cornwell et al. (2014) have a series of functions in Table 1 that could serve as "process models". The approach I'm going to try is to use a sampling model that represents the binomial experiment (seeds in seed bags) and a process model that represents the decay of survival probability.

We start with

$$\begin{aligned} y_{it} &\sim \text{binomial}(n_i, \theta) \\ \theta &\sim \text{beta}(1, 1) \end{aligned} \tag{1}$$

to say that the observations of the number of seeds counted in bag i at time t are represented as y_{it} drawn from a binomial distribution. In this case, n_i is the number of trials, the number of seeds starting the experiment; y_{it} is the number of successes, the number of seeds remaining. The θ is the probability of success on a single trial.

$$[\theta|\boldsymbol{y}, \boldsymbol{n}] \propto \text{binomial}(y_{it}|n_i, \theta)\text{beta}(\theta|1, 1) \tag{2}$$

We then have sampling variability that is implicit in the binomial. There are thus two sources of uncertainty. Uncertainty arising from sampling and uncertainty arising because of ....

The case where each bag has its own mean survival at each time, acknowledging variation among bags:

$$\begin{aligned} y_{it} &\sim \text{binomial}(n_i, \theta_{it}) \\ \theta_{it} &\sim \text{beta}(\alpha, \beta) \end{aligned} \tag{3}$$

Giving the following :

$$\begin{aligned} [\boldsymbol{\theta}, \alpha, \beta|\boldsymbol{y}, \boldsymbol{n}] \propto \text{binomial}(y_{it}|n_i, \theta_{it})\text{beta}(\theta_{it}|\alpha, \beta) \\ \text{gamma}(\alpha|0.001, 0.001)\text{gamma}(\beta|0.001, 0.001) \end{aligned} \tag{4}$$

We now want to model the process "change in being intact with time" as

$$g(k, t_i) = \exp(-kt_i) \tag{5}$$

This deterministic model represents the average proportion of seeds that are expected to be intact as a function of time under a negative exponential process. The function $g(k, t_i)$ is the overall mean probability of being intact at time t_i. This is the mean of the beta distribution; the variance is the variation in probability of being intact at time t_i that arises from differences among time. The uncertainty that arises from sampling - which we can estimate because of replication - is distinct from this process variance. I think the process variance is all effects that create variation beyond the seed age, as represented with a negative exponential function.

The next problem is that the process is not simply one of decay, decomposition, or mortality. Instead, there are annual events interspersed into this, namely germination. In this way I think the situation resembles the complement of case 2 in Olson (1963). Modeling

germination and survival jointly would account for the full data. Here's one idea:

$$
\begin{aligned}
h_1(k, t_1) &= \exp(-kt_1) \\
h_2(k, t_1) &= g_1 \exp(-kt_1) \\
h_3(k, t_2) &= (1 - g_1) \exp(-kt_2) \\
h_4(k, t_3) &= (1 - g_1) \exp(-kt_3) \\
h_5(k, t_3) &= (1 - g_1)(g_2) \exp(-kt_3) \\
h_6(k, t_4) &= (1 - g_1)(1 - g_2) \exp(-kt_4) \\
h_7(k, t_5) &= (1 - g_1)(1 - g_2) \exp(-kt_5) \\
h_8(k, t_5) &= (1 - g_1)(1 - g_2) g_3 \exp(-kt_5) \\
h_9(k, t_6) &= (1 - g_1)(1 - g_2)(1 - g_3) \exp(-kt_6)
\end{aligned}
\tag{6}
$$

Perhaps it would be productive to break the probability into two processes, one accounting for mortality and another accounting for decay. The one accounting for mortality would follow one of the following processes. First, we consider a negative exponential mortality trajectory.

$$g(k, t_i) = \exp(-kt_i) \tag{7}$$

Then, we consider a continuous exponential mortality trajectory:

$$g(a, b, t_i) = \frac{1}{(1 + bt_i)^a} \tag{8}$$

Then, we consider a mortality trajectory for Weibull residence times:

$$g(\alpha, \beta, t_i) = \exp -(\frac{t_i}{\beta})^\alpha \tag{9}$$

The mortality process would be multiplied with the process "change in removal from population with age due to germination", represented as

$$
\begin{aligned}
h_1(\boldsymbol{g}) &= 1 \\
h_2(\boldsymbol{g}) &= g_1 \\
h_3(\boldsymbol{g}) &= (1 - g_1) \\
h_4(\boldsymbol{g}) &= (1 - g_1) \\
h_5(\boldsymbol{g}) &= (1 - g_1)g_2 \\
h_6(\boldsymbol{g}) &= (1 - g_1)(1 - g_2) \\
h_7(\boldsymbol{g}) &= (1 - g_1)(1 - g_2) \\
h_8(\boldsymbol{g}) &= (1 - g_1)(1 - g_2)g_3 \\
h_9(\boldsymbol{g}) &= (1 - g_1)(1 - g_2)(1 - g_3)
\end{aligned}
\tag{10}
$$

In turn we would consider models that combine age-dependent and -independent germination functions and mortality functions. The table below shows the models we considered, the parameters, and the number of parameters for each model. The product of the mortality and germination process is a deterministic function that equals the average proportion of seeds that are expected to be intact or germinated. This average proportion is the mean of a beta distribution with process variance that captures variation beyond the effect of seed age on mortality as represented with a mortality process and germination. We might also consider a germination process where germination is randomly distributed and independent for each bag?

Table 2: Seed bag dataset models

Model	Mortality process	Germination process	Parameters	Parameter number
A1	Negative exponential	Constant	k, g	2
A2	Negative exponential	Age-dependent	k, g_{1-3}	4
B1	Compound exponential	Constant	a, b, g	3
B2	Compound exponential	Age-dependent	a, b, g_{1-3}	5
C1	Weibull residence time	Constant	α, β, g	3
C2	Weibull residence time	Age-dependent	α, β, g_{1-3}	5

The following equations correspond to the full conditional distributions for models at the first time point. Subsequent models incorporate germination in the function $m(\dots, \sigma^2)$. Specifically, the mean is the product of the mortality up to time t_i, represented by $g(\dots)$, germination up to time t_i, represented by $h(\dots)$. We will use the information from the seed bag experiments to determine whether there is an appropriate mortality process with which to model the recruitment data from the seed pots. What I would like to do with this is get a mortality and germination process that's appropriate for these sites and use the posterior for the parameters in that process as informed priors to model the seed recruitment from the seed pots.

The base assumption is often that seed mortality proceeds at a constant, absolute rate. Although this assumption has repeatedly been challenged (Lonsdale 1988, Rees and Long 1993) there remain limited assessments for how appropriate it is. A negative exponential model corresponds to the decay of radioactive isotopes. As described in Cornwell and Weedon (2014), this model has a constant decay parameter k that implies the decaying material be treated as a homogeneous mass with each seed having an equal probability of mortality throughout time. But we know that seeds exhibit phenotypic variation in dormancy-related traits. The negative exponential decay model may be appropriate if seed survival is random in time and largely unrelated to seed characteristics.

The continuous exponential model is one in which seed mortality is described by a continuous quality distribution.

The Weibull residence time model represents the seed pool as a distribution of survival

times. 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.

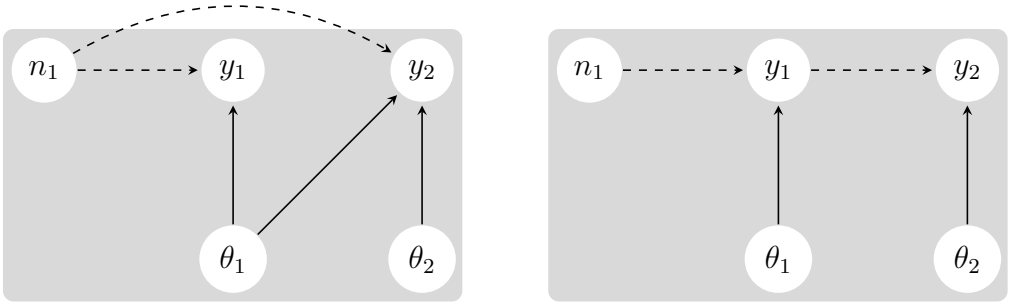
Finally, we also consider the model proposed in Gremer and Venable (2014), namely one in which seed mortality is seasonal and age-dependent.

$$\begin{aligned} g(k, t_i) &= \exp(-kt_i) \\ [k, \alpha, \beta | \boldsymbol{y}, \boldsymbol{n}] &\propto \text{binomial}(y_{it} | n_i, \theta_{it}) \text{beta}(\theta_{it} | m(g(k, t_i), \sigma^2)) \\ &\quad \times \text{gamma}(k | 0.001, 0.001) \\ &\quad \times \text{inverse gamma}(\sigma^2 | 0.001, 0.001) \end{aligned} \tag{11}$$

$$\begin{aligned} g(a, b, t_i) &= \frac{1}{(1 + bt_i)^a} \\ [k, \alpha, \beta | \boldsymbol{y}, \boldsymbol{n}] &\propto \text{binomial}(y_{it} | n_i, \theta_{it}) \text{beta}(\theta_{it} | m(g(a, b, t_i), \sigma^2)) \\ &\quad \times \text{gamma}(a | 0.001, 0.001) \text{gamma}(b | 0.001, 0.001) \\ &\quad \times \text{inverse gamma}(\sigma^2 | 0.001, 0.001) \end{aligned} \tag{12}$$

$$\begin{aligned} g(\alpha, \beta, t_i) &= \exp -(\frac{t_i}{\beta})^\alpha \\ [k, \alpha, \beta | \boldsymbol{y}, \boldsymbol{n}] &\propto \text{binomial}(y_{it} | n_i, \theta_{it}) \text{beta}(\theta_{it} | m(g(\alpha, \beta, t_i), \sigma^2)) \\ &\quad \times \text{gamma}(\alpha | 0.001, 0.001) \times \text{gamma}(\beta | 0.001, 0.001) \\ &\quad \times \text{inverse gamma}(\sigma^2 | 0.001, 0.001) \end{aligned} \tag{13}$$

Identifiability within years

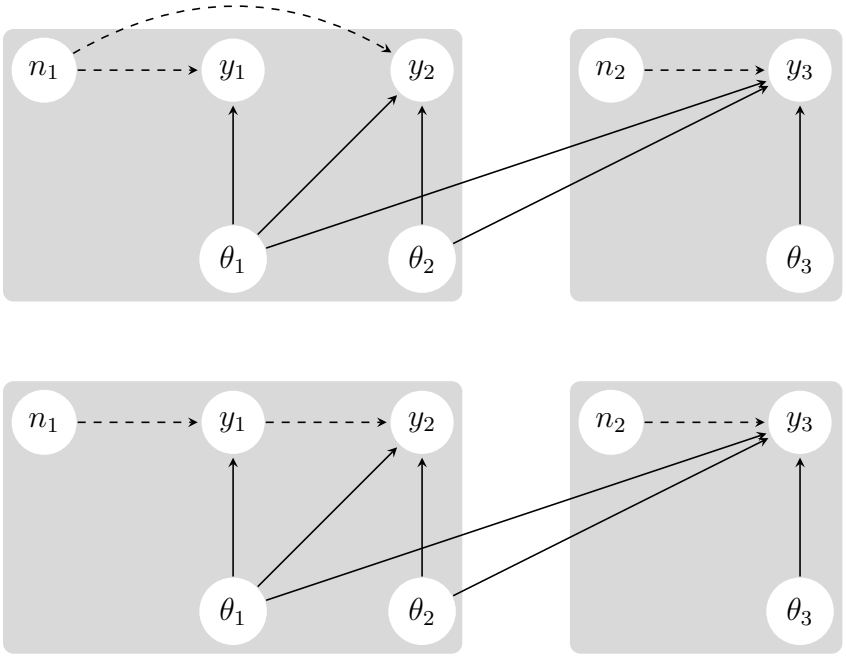


(a) Directed acyclic graphs for the hierarchical models for seed bag rates. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

$$
\begin{aligned}
[\theta_1, \theta_2 | \boldsymbol{y_1}, \boldsymbol{y_2}] \propto & \text{ binomial}(y_1 | n_1, \text{logit}^{-1}(\alpha_1)) \\
& \times \text{binomial}(y_2 | n_1, \text{logit}^{-1}(\alpha_1) \times \text{logit}^{-1}(\alpha_2)) \\
& \times \text{normal}(\alpha_1 | \mu_1, \sigma_1) \text{normal}(\alpha_2 | \mu_2, \sigma_2) \\
& \times \text{normal}(\mu_1 | 0, 1000) \text{half-normal}(\sigma_1 | 0, 1) \\
& \times \text{normal}(\mu_2 | 0, 1000) \text{half-normal}(\sigma_2 | 0, 1).
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
[\theta_1, \theta_2 | \boldsymbol{y_1}, \boldsymbol{y_2}] \propto & \text{ binomial}(y_1 | n_1, \text{logit}^{-1}(\alpha_1)) \\
& \times \text{binomial}(y_2 | y_1, \text{logit}^{-1}(\alpha_2)) \\
& \times \text{normal}(\alpha_1 | \mu_1, \sigma_1) \text{normal}(\alpha_2 | \mu_2, \sigma_2) \\
& \times \text{normal}(\mu_1 | 0, 1000) \text{half-normal}(\sigma_1 | 0, 1) \\
& \times \text{normal}(\mu_2 | 0, 1000) \text{half-normal}(\sigma_2 | 0, 1).
\end{aligned} \tag{15}
$$

Identifiability across years

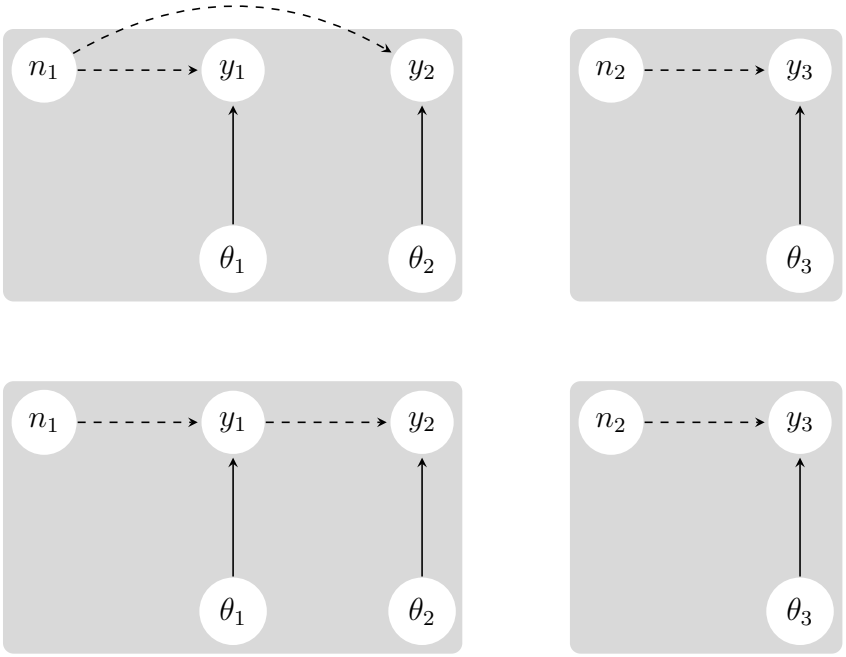


(a) Directed acyclic graphs for the hierarchical models for seed bag rates. Solid arrows depict the relationships among random variables, and dashed arrows depict the deterministic relationships.

$$
\begin{aligned}
[\theta_1, \theta_2 | \boldsymbol{y_1}, \boldsymbol{y_2}] \propto & \text{ binomial}(y_1 | n_1, \text{logit}^{-1}(\alpha_1)) \\
& \times \text{binomial}(y_2 | n_1, \text{logit}^{-1}(\alpha_1) \times \text{logit}^{-1}(\alpha_2)) \\
& \times \text{normal}(\alpha_1 | \mu_1, \sigma_1) \text{normal}(\alpha_2 | \mu_2, \sigma_2) \\
& \times \text{normal}(\mu_1 | 0, 1000) \text{half-normal}(\sigma_1 | 0, 1) \\
& \times \text{normal}(\mu_2 | 0, 1000) \text{half-normal}(\sigma_2 | 0, 1).
\end{aligned} \tag{16}
$$

$$
\begin{aligned}
[\theta_1, \theta_2 | \boldsymbol{y_1}, \boldsymbol{y_2}] \propto & \text{ binomial}(y_1 | n_1, \text{logit}^{-1}(\alpha_1)) \\
& \times \text{binomial}(y_2 | y_1, \text{logit}^{-1}(\alpha_2)) \\
& \times \text{normal}(\alpha_1 | \mu_1, \sigma_1) \text{normal}(\alpha_2 | \mu_2, \sigma_2) \\
& \times \text{normal}(\mu_1 | 0, 1000) \text{half-normal}(\sigma_1 | 0, 1) \\
& \times \text{normal}(\mu_2 | 0, 1000) \text{half-normal}(\sigma_2 | 0, 1).
\end{aligned} \tag{17}
$$

Exponential decay



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