Elsevier Editorial System(tm) for Ecological Modelling
Manuscript Draft

Manuscript Number: ECOMOD-18-694R2

Title: Consequences of neglecting cryptic life stages from demographic models

Article Type: VSI:Matrix Population Models

Keywords: Comparative demography; cryptic life stages; matrix population models; plant population ecology; seed bank; seed dormancy

Corresponding Author: Dr. Vuong Nguyen, Ph.D.

Corresponding Author's Institution: The Save Sight Institute

First Author: Vuong Nguyen, Ph.D.

Order of Authors: Vuong Nguyen, Ph.D.; Yvonne Buckley; Roberto Salguero-Gomez; Glenda Wardle

Vuong Nguyen

3 October 2018

Dear Prof Brian Fath and editorial board of *Ecological Modelling*,

Please find attached our manuscript entitled "**Consequences of continuing to exclude cryptic life stages from demographic models"** which we are submitting in consideration for publication in Ecological Modelling as part of the special issue on **"Theory and Practice in Matrix Population Modelling".**

This work identifies a methodological oversight that continues to plague the world of demographic matrix modelling and believe that it will be a perfect fit for the upcoming special issue. The findings are an '*inconvenient truth'* for modellers and empirical demographers alike and we are candidly admitting that the work has been subject to peer review already. Reviewers were complementary about the importance of the topic and the clarity of the writing but held divergent view on the modelling based on their philosophical point of view. We provide substantial supplementary material to ensure that our methodology and analyses can be thoroughly checked.

In this manuscript, we explore the impact of excluding cryptic life stages stage from demographic models, focusing specifically on dormant seed banks as a common example of such a stage. We conducted a review of published plant matrix population models from 2008-2011 and found 47% had excluded the seed bank stage without providing justification; despite a call made over a decade ago for a better justification of seedbank exclusion (Doak et al, 2002), the field has not changed in this regard. This finding provided the appropriate motivation to investigate how this exclusion, and subsequent uncertainty in demographic parameters for the seed bank, will affect the model outcomes. We used a combined simulation and empirical approach, targeting published population models in which we found the study species had the potential for seed dormancy but the author(s) excluded the seed bank without justification, and (ii) ran Monte Carlo simulations with both uninformed and informed priors for seed bank vital rates based on literature estimates. Subsequently, we found that the relative importance of the seed bank to the long-term stochastic performance of the populations is context-specific, depending on the species and types of environmental cues present. Demographic data on the seed bank can be difficult to obtain, and our combined use of simulations and external information from the literature to assess the potential for dormancy and identify priors for vital rates provides a useful starting point for assessing whether a seed bank can be excluded without consequence or is actually important.

Yours sincerely,
Vuong Nguyen

Desert Ecology Research Group
School of Biological Sciences
Faculty of Science
Rm 325, Heydon-Laurence Building A08
The University of Sydney
Sydney, NSW 2006, Australia

T +61 2 9351 7627
F +61 2 9351 4119
E pngu4751@sydney.edu.au
www.sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

THE UNIVERSITY OF
SYDNEY

On behalf of the authors

Point by point response form

E.1) I appreciate your positive responses to the numerous Reviewer's remarks in the form of a 12-page Table, yet I'm quite surprised that you have not written any word of gratitude to acknowledge Reviewer's 262 (!) 'PDF' Comments made in your original ms. By the way, the probability density function should better be abbreviated as 'p.d.f.'

Response: We thank the editor for reminding us of this oversight and we have acknowledged the anonymous reviewer for their substantial edits/comments. The abbreviation for probability density function has also been edited to p.d.f. as recommended.

E.2) The Highlight: 'Uncertainty in seed bank vital rates the major driver influencing model predictions' lacks the predicate and looks too hopeless. It should better be changed to something like: >Inclusion of the seed bank may change the model predictions made otherwise.<

Response: We have changed the highlight as recommended by the editor.

E.3) You give a misleading comment on ergodicity in lines 220–221: 'the condition that it is possible to transit from each stage to all other stages of the life cycle graph' In fact, this condition is called strong connectedness of the directed graph, while the ergodicity, as Caswell defined it following Cohen (1979), is its consequence. Also, the life cycle graph has no 'stages' – just nodes and arcs between them.

Response: We have corrected these term s to "strong connectedness" and "nodes" where appropriate.

E.4) The next sentence (222) should better be given in Present Indefinite. There still are other places where the Past Indefinite verb about the current study is neighbouring with refs to others, thus invoking ambiguity.

Response: We have read through the manuscript and fixed further instances of the past indefinite verb being used to describe the current study and the outcomes.

E.5) There other places in English, too, to be corrected, in particular, wrong commas (265) and missed commas/full stops after displayed formulae.

Response: Commas have been added/removed where appropriate

E.6) Reviewer's point 6.4) (about elasticity matrices) has been well responded in Revision 1, but Table S2 lacks an explanation of what Ø means.

Response: We have included the following the caption for Table S2 "The Ø symbol indictates that the elasticity is not defined for the corresponding element of the plant matrix model since the link was not present in the life cycle graph."

E.7) The correctness of notation λ1 is accompanied with its greater vulnerability to misprints than just λ, such as λ1 (402, 430, elsewhere) or λ1 (253), hence it needs greater care.

Response: Notation for all instances of λ has been checked for correct use of superscripts and italicisation.

Consequences of neglecting cryptic life stages from demographic models

Vuong Nguyen[1]*, Yvonne M. Buckley[2,3], Roberto Salguero-Gómez[2,4,5] and Glenda M. Wardle[1]

[1]*Desert Ecology Research Group, School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia*

[2]*School of Biological Sciences, The University of Queensland, Queensland, 4072, Australia*

[3]*School of Natural Sciences, Department of Zoology, Trinity College Dublin, Dublin, Ireland*

[4]*Evolutionary Biodemography Laboratory, Max Planck Institute for Demographic Research, Rostock 18057, Germany*

[5]*Department of Zoology, University of Oxford, Oxford, UK*

*Correspondence author. E-mail: phuc.nguyen@sydney.edu.au

Abstract

Information on individuals from *all* stages of life is crucial to explore their ecology, evolution and conservation biology. However, the life cycles of many species contain cryptic life stages that are difficult to detect and track over time and are therefore omitted from demographic models. One example is the dormant seed bank, an evolutionary bet-hedging mechanism that buffers plant populations in variable environments. To evaluate this methodological oversight, we conduct~~ed~~ simulations to explore the effect of seed bank parameter uncertainties on demographic outputs such as the deterministic (λ_1) and stochastic population growth rate (λ_S), and extinction probabilities of 12 plant species. We have used uninformed and informed priors for seed bank parameters based on literature estimates, and reconstructed published models in which the seed bank was excluded without justification. Trials on removing the seed stage from models (6 species) explored the worst-case scenario for ignoring the seed bank. Inclusion of a seed bank and demographic uncertainty in seed bank parameters have little impact on stable populations ($\lambda_1 \approx 1$) with high post-seedling survival. When populations deviate from stability or demonstrate temporal demographic variation, greater changes in λ_1 and the range of possible growth rates caused by demographic uncertainty are observed. As expected, decreasing populations ($\lambda_1 < 1$) benefit from the inclusion of a seed bank through increases in the growth rate and extinction times, whereas increasing populations ($\lambda_1 > 1$) are slowed down. While germination estimates from the literature cannot accurately reflect those obtained in the field, they provide a starting point to assess the relative importance of a seed bank. The exclusion of the seed bank must be justified by confirming that dormancy is either non-existent or not important. Accounting for cryptic stages in demographic models will produce better informed management decisions for threatened or invasive species.

Keywords: Comparative demography, cryptic life stages, matrix population models, plant population ecology, seed bank, seed dormancy.

Highlights

- Inclusion of seed bank reduced extinction times and improved long-term persistence
- Published literature and simulations can assess the importance of the seed bank
- ~~Uncertainty in seed bank vital rates the major driver influencing model predictions~~Inclusion of the seed bank may change model predictions made otherwise

Introduction

Demographers have long recognised that individuals in a population span across different stages of their life cycle, from propagules up to reproductive adults, and that these stages contribute to its persistence in different ways – failing to account for all these stages can therefore be misleading (Caswell, 2001). However, the life cycles of many species contain cryptic life stages that may be difficult to detect or sample. These cryptic life stages occur over a range of taxonomic groups, and may include hibernating or dormant individuals, clonal individuals, eggs, mobile propagules, or simply stages occupying inaccessible locations (Geiser, 2004, Katzner et al., 2011, Shefferson, 2009, Tanner, 2001). The dormant seed bank is perhaps one of the most common examples of the cryptic life stage whereby seeds enter prolonged dormancy (>1 year) within the soil or above ground in woody cones, and can make up a significant portion of the population (Baskin and Baskin, 2014, Harper, 1977). By spreading germination through time, seed dormancy acts as a bet-hedging mechanism mitigating the effects of unfavourable years in exchange for reduced proliferation during favourable years (Gremer et al., 2012, Venable, 2007). In extreme cases, seed banks can prevent local extinctions when above-ground populations are killed off entirely (Stöcklin and Fischer, 1999). Seed dormancy has also been shown to promote species co-existence in variable environments (Warner and Chesson, 1985) and act as a temporal reservoir for genetic diversity (Vitalis et al., 2004).

With such substantial consequences for population dynamics, it would be a major oversight to exclude the seed bank, or any cryptic life stage, from population estimates and demographic analyses. This argument is indeed not new: the need to incorporate the seed bank into demographic models has been recognised for decades (Harper, 1977, Kalisz and McPeek, 1992). Matrix population models, which divide individuals into their respective stage classes, are the most common tool for modelling plant populations – the online database of plant matrix models, COMPADRE v5.0.0 contains 8,906 published matrices for 759 species (COMPADRE Plant Matrix Database, v5.0.0., 2019, Salguero-Gómez et al., 2015). These models have a long history in the management of threatened and invasive species (Crone et al., 2011). Yet, a previous review of 70 published matrix population models found almost half (43%) had excluded the seed bank stage without justification (Doak et al., 2002). Surprisingly, we have found that this proportion has remained unchanged in the last decade (47%; Table 1). We acknowledge that including the seed bank is not straightforward as the estimation of vital rates such as survival and germination generally rely on additional laboratory experiments (Crone et al., 2011), and even these would inflate actual *in situ* estimates (Gross and Mackay, 2014). Furthermore, unless such experiments are rigorous enough to include age-specific vital rates (Kalisz and McPeek, 1992), the seed bank has been generally assumed to be unstructured. Excluding the seed bank may be justified if it is discovered the species, or population, does not possess dormant seeds (see Baskin and Baskin, 2014 for review on the presence of seed dormancy), or if it is mathematically and empirically justified in a modified model (Logofet et al. 2017), however, these cases are the minority compared to those that do not provide justification (Doak et al., 2002; Table 1). Unjustified exclusion can have real-world consequences if, for example, the seed bank of an invasive species recolonises after the adult individuals have been eliminated (Regan et al., 2011), or if the population size and persistence of threatened species is underestimated (Doak et al., 2002).

The practical issues of including the seed bank and its exclusion from demographic models raise two important questions. First, how can we bridge the gap in our understanding of whether a seed bank is necessary to adequately model the population dynamics in the absence of field data? Second, how does uncertainty in the presence of a seed stage and its transition rates translate into uncertainty in the model? Addressing these questions using a simulation-based approach can be useful for investigating a range of scenarios and assumptions but may not be biologically meaningful without application to real-world examples. An empirical approach, on the other hand, may be limited in its scope if it were to focus on a single case study, as there are clear difficulties in obtaining individual-level records for published models. Using published matrix models that have excluded the seed bank without justification and information on dormancy and germination rates from the literature, we implement a combination of these approaches and perform a comparative analysis across a range of species and life forms (12 species including eight trees, three shrubs and one herb). We conduct Bayesian Monte Carlo simulations on these published models using uninformed and informed priors (based on literature estimates) for seed bank vital rates to gauge the level of uncertainty when the seed bank presence and parameters are unknown. To complement these addition simulations, we also investigate the implications of assuming the absence of a persistent seed bank by intentionally removing the seed bank from models which did include it.

We make a number of predictions based on the evolutionary advantages of seed dormancy. First, we might expect the relative importance of seed banks to be greater for species in unpredictable environments with variable post-seedling survival compared to stable populations of long-lived species with consistently high adult survival (Venable, 2007). Furthermore, the deterministic growth rate (λ_1), which assumes vital rates remain constant, should increase during poor demographic years ($\lambda_1 < 1$) and decrease during productive years ($\lambda_1 > 1$) to reflect the function of the seed bank as a bet hedging mechanism against unfavourable years (Gremer et al., 2012, Kalisz and McPeek, 1992). Similarly, elasticities (the relative influence of each demographic process to the population growth rate; de Kroon et al., 2000) for seed bank transition rates should be greater during poor years. We should also observe increases in times to local extinction and stochastic growth rates (λ_S)~~,~~ which incorporate temporal variation in vital rates. Using both an empirical and simulation-based approach, we provide more robust projections of the demographic fates of populations and help to motivate the inclusion of the seed bank into demographic models.

Materials and Methods

Database of population matrix models

At the time this study was conducted, the COMPADRE Plant Matrix Database had not yet become open-access online (Salguero-Gómez et al., 2015), with the most recent update consisting of 204 terrestrial plant species up to 2007, which we hereafter refer to as the "MPM2007" database constructed by the Vegetation Function Working Group (see Burns et al., 2010 for further information). We have used the ISI Web of Science to identify published plant matrix population models from 2008-2011, which we hereafter refer~~red~~ to as the "MPM2008-2011" database.~~,~~ These studies ~~which~~ would later go on to be added to the COMPADRE database. Examples of keywords used in our search include "matrix model*", "transition matrix", "demographic model*", "population model*", "population viability analysis", "population growth rate", and "population dynamics". The addition of * at the end of a search term acts as a wildcard and allows for alternative forms of a given word, e.g. model(s), model(ling). We have found 55 studies with matrix models for 77 terrestrial plant species including 38 trees, 27 herbs, 8 shrubs and 4 succulents.

Confirmation of the seed bank

Using Baskin and Baskin (2014) and the species name as a search term in ISI Web of Science, we have verified whether a seed bank was reported for species in which seed bank exclusion was unjustified (by not mentioning the seed bank; Table 1). While each population may not necessarily possess a seed bank, we can still determine the impact of its presence on model outcomes. In total, we have constructed new matrix population models for the 12 plant species for which published germination data on either the species or the genus were available and the original matrices were provided (see Supplementary File S1 for all original and reconstructed matrices).

Seed germination submodel

The seed bank may be modelled as structured or unstructured. In the unstructured seed bank model, it is assumed that the seed bank does not age and is potentially immortal. In the structured seed bank model, the seed bank progressively ages~~,~~ with potentially different survival, viability and germination rates~~,~~ until it reaches time step *n*, at which point the vital rates remain constant or the seeds die if they do not germinate (see Figure 1B for a life cycle example of the structured seed bank and 1C for the unstructured seed bank). The maximum length of time the seed bank can remain viable is variable across species and is difficult to quantify - some seeds have been documented germinate after centuries of dormancy (Fenner, 2000).

Addition and removal of seed bank stages

Adding, removing, or excluding stages without justification alters the dimensionality of the model~~,~~ and this will affect the demographic parameter estimates. Merging or further dividing up stage classes directly alters vital rates for each class being affected and the speed of the life cycle, where decreasing the dimensionality reduces the number of stages to progress through the life cycle and vice versa (Salguero-Gómez and Plotkin, 2010). As the seed bank functions as a bet hedging mechanism, delayed movement through the life cycle is to be expected and this is reflected in our hypotheses. Since the seed bank is a separate, discrete stage class that involves no splitting or merging, transi~~a~~tion rates for the other classes should remain unaffected by its inclusion. However, seedling fecundity would need to be lowered accordingly to account for observed seedlings emerging from the seed bank to avoid overestimating fecundity. We use s~~S~~eedling counts ~~were used~~ to estimate fecundity in all models where the seed bank was excluded without justification. For simplicity, we ~~have~~ assume~~d~~ no clonal reproduction into the seedling stage such that the observed seedlings only consist of two components: the individuals that germinated immediately between year *t* and *t* + 1 and those that germinated from the dormant seed bank from prior years. The number of observed seedlings can thus be described by the following equation:

$$f_{i,j} = F_j(t)vgs + F_j(t\text{-}1)v(1\text{-}g)(1\text{-}d)g_{b1} + \mathrm{F}_j(t\text{-}2)\mathrm{v}(1\text{-}g)(1\text{-}g_{b1})(1\text{-}d)^2 g_{b2} + \ldots, \tag{1}$$

where the first term on the right hand side represents immediate germinants, the second term represents germination of seeds from the seed bank produced at the previous time step~~, and~~ the third term represents germination of seeds from the seed bank produced two time steps ago, and so on. Notations:

$f_{i,j}$ = per-capita fecundity into stage class *i* by stage class *j* in one time step;
$F_j(t)$ = total fecundity by stage class *j* at time *t;*
v = seed viability rate during the step *t* to *t* + 1;
g = non-seed bank germination rate at the period *t* to *t* + 1;
g_{b1} = germination rate from seeds that remained dormant in the seed bank for 1 time step;

g_{b2} = germination rate from seeds that remained dormant in the seed bank for 2 time steps;

s = seedling survival rate until the next census;

d = seed bank mortality rate

See Fig. 1 for more information.

To simplify the inclusion of the seed bank stage under limited information, we have made the following assumptions. First, the seedling survival rate (s) from the period t to $t + 1$, is equal to the seed bank survival rate ($1 - d$). Second, we start with simulations that focus on the contribution of the seed bank from $t - 1$. This unstructured submodel captures the predominant seed bank contribution since the contribution of seeds from previous years will diminish exponentially with each year~~,~~ if we assume the vital rates to be not seed age dependent. Alternative age-structured formulations are explored below. Additionally, we assume the fecundity from the previous year ~~was assumed~~ to equal the fecundity in the current year, i.e. $F_j(t) = F_j(t-1)$, hereafter denoted by F_j. Finally, the probability of germinating within the census year ~~was~~ is equal to the probability of germinating from the seed bank, i.e. $g = g_b$ (Kalisz and McPeek, 1992). Using the published information on germination, we simplify Eq. 1:

$$f_{i,j} = F_j v g s + F_j v(1-g)(1-d)g, \tag{2}$$

where the first term on the right-hand side represents the number of seedlings germinating immediately between t and $t + 1$, and the second term corresponds to germination from the seed bank. For estimates of the deterministic and stochastic growth rates, we set seed viability (v), seedling survival (s) and seed survival ($1 - d$) equal to 1 unless this information was provided in the literature alongside germination estimates. Monte Carlo simulations (described below) then investigate~~d~~ drawing these values from both uninformed and informed prior distributions. Taking the observed $f_{i,j}$ from the original matrix, we solve Eq. 2 for F_j and calculate the per-capita fertility contributions to seedlings and seed bank (See Supplementary Method S1 for an example). Following the same approach, we also explore alternative model formulations and the effect of different assumptions but only perform~~ed~~ Monte Carlo simulations for the unstructured scenario, which is the simplest and most common approach to including the seed bank. The alternatives include the addition of an age-structured seed bank distinguishing seeds from t - 2 and t - 3 and more conservative estimates on the effect of the seed bank by further assuming the seeds do not survive beyond the first, second or third year (See Supplementary Method S1 and Table S1 for the alternative formulations and population growth rates for all models).

Researchers excluding the seed bank from a model make the implicit assumption that it is not present in the population. Thus we intentionally remove the seed bank from selected models to explore the implications of this assumption. Given the large number of potential models to choose from, we have decided to focus on cases in which the seed bank was known to be important (elasticities of λ_1 to seed survival, germination or fecundity > 0.1) to demonstrate the worst case scenario that might be possible when the seed bank is ignored. We have identified six such species from the MPM2007 and MPM2008-2011 databases that fulfil this criteria whilst retaining ~~ergodicity~~strong connectedness, the condition that it is possible to transit from each ~~stage~~ node to all other ~~stages~~ nodes of the life cycle graph (Caswell, 2001), such as those in Figure 1. We use ~~I~~information on seed viability, seedling survival, and seed survival ~~were used~~ if provided, but ~~were~~ otherwise we set them to 1, and we no longer assume that $g = g_b$ since g_b can be obtained from the original matrix. We estimate t~~T~~he rate of seed germination within the studied period, g, ~~was then estimated~~ based on partitioning the total reproductive output (Eqs. 1

and 2) to recalculate only the fecundity into the seedling stage. To further ensure that no artefacts from changes to reproduction or mortality ~~were ~~are introduced, the removal protocol ~~was ~~is the reverse of the addition protocol, ~~, and a~~and we ~~-~~check ~~of ~~that the ~~recovery of the initial matrix after these manipulations was performed~~can be recovered by reapplying the addition protocol (See Supplementary Method S2 for an example).

Analyses

We use R (v 3.1.1; The R Foundation for Statistical Computing) to calculate ~~P~~population growth rates (λ_1), elasticities, stochastic simulations exploring temporal variation in vital rates and Monte Carlo simulations exploring uncertainty in seed bank vital rates (~~were calculated using R (v 3.1.1; The R Foundation for Statistical Computing; ~~see Supplementary File S2 for code to perform Monte Carlo simulations).

~~Monte Carlo simulations used an~~We use uninformed priors for germination, viability, seedling survival and seed survival for Monte Carlo simulations by drawing these rates~~and thus these rates were drawn~~ from a uniform distribution~~ for the simulations~~. Although we might expect covariance among vital rates for the below- and above-ground individuals (Morris and Doak, 2002), introducing this structure is complicated particularly when dealing with the seed bank. For example, models that do include the seed bank generally estimate seed bank vital rates through separate germination experiments such that the covariance between the seed bank and the above-ground population is lost in the process. While the impact of this covariance structure on the model outcomes would be of interest, we instead focus here on the exclusion of the seed bank from demographic models and~~,~~ leave~~ing~~ investigations regarding correlation structure for future research.

We perform 10,000 simulations, recalculating the fecundities and reconstructing the matrix for each set of parameters, and report the mean growth rate with confidence intervals obtained by the 2.5% and 97.5% quantiles. We then determine the effect of using informed priors via literature estimates for germination, and occasionally viability when available, on the resulting probability density function (~~PDF~~p.d.f.) of the distribution of growth rates. Publications on germination often contain multiple treatments and an estimated mean germination rate for each treatment. We have specified our prior distributions depending on the type of demographic year to reflect the function of the seed bank as a bet hedging mechanism (the highest germination rate available when $\lambda_1 > 1$ and the lowest when $\lambda_1 < 1$; Gremer et al., 2012, Kalisz and McPeek, 1992). However, if only one germination rate was provided, we use this value for all population matrix models regardless of the demographic year. If only the mean matrix was provided (arithmetic mean matrix across all years and sites), we use the average germination over all treatments. We model both germination and viability using the beta distribution as it varies between 0 and 1, with the distribution mode equal to the germination rate. We ~~have~~ also calculate~~d~~ λ_1 for fixed values of the germination rate (0.01-1 by increments of 0.01) and viability rate (0.1, 0.5, and 1) to capture the total range of possible values when only these parameters are allowed to vary. The matrices in which the germination rate was set to that obtained from the literature ~~were~~ are used to calculate single point estimates of λ_1, elasticity matrices, λ_S and extinction probability, with viability, seed survival and seedling survival set to 1 if no information was available. The initial population for all stochastic and extinction probability calculations hereafter consists of 1,000 individuals at the stable stage distribution, calculated as the normalized dominant right eigenvector (Caswell, 2001) of the mean matrix. We have calculated s~~S~~tochastic growth rates ~~were calculated~~ via random choice from a set of available

Formatted: Font: Italic

Formatted: Subscript

Formatted: Font: Italic

Formatted: Subscript

matrices (Kaye and Pyke, 2003), where each matrix has an equal probability of being selected at each time step. The initial population ~~was~~ is projected over 500,000 time steps with the first 5,000 omitted to remove transient dynamics (Kaye and Pyke, 2003, Morris and Doak, 2002). We use ~~T~~the mean and variance of the growth rates ~~were used~~ to obtain an estimate for λ_S and 95% confidence intervals.

We also compare extinction probabilities between models with and without seed banks. Times to extinction (defined as population size ≤ 1 individual) ~~were~~ are calculated for matrices predicting population decline. We project an initial population for individual matrices over 100 years and compared the difference in times to extinction with and without the seed bank to quantify the effect of the seed bank on population persistence. We then conduct stochastic simulations to calculate the cumulative probability of quasi-extinction (Morris and Doak, 2002) with a quasi-extinction threshold set to 100 individuals (i.e. one tenth of the initial population size). We perform 100 runs, where a single run consists of 5000 iterations each projected over 100 years, and calculate the mean fraction of projections that fell below the quasi-extinction threshold during or before the year *t* (Morris and Doak, 2002).

Results

Addition of the seed bank and demographic uncertainty

Inclusion of an unstructured seed bank show~~ed~~ fairly small differences in λ for most species regardless of the assumptions made (See Fig. 2 and Table 2 to compare the growth rates). However, the potential range of growth rates when demographic parameters are uncertain can be large depending on the species and the amount of temporal variation observed. The 8 tree species, for example, have~~d~~ stable populations ($\lambda_1 \approx 1$)~~,~~ and variation in seed bank parameters has~~d~~ little influence in all of the models with the exception of *Illicium anisatum*

(Schisandraceae). Subsequently, using informed priors for these species ha~~d~~s little impact on the ~~PDF~~ p.d.f. of the distribution of possible growth rates. In contrast, the shorter-lived shrub species, *Atriplex acanthocarpa* (Amaranthaceae) and *Sambucus racemosa* (Adoxaceae), ~~were~~ are more sensitive to the introduction of a seed bank and demographic uncertainty in their vital rates. Here, using an informed prior on germination narrow~~ed~~s the ~~PDF~~ p.d.f. of the distribution of possible growth rates for the *A. acanthocarpa* 1996-1997 matrix during an extremely productive year (Fig. 3). Conversely, the use of prior information during the poorer 1997-1999 years appear~~ed~~s to increase the spread of possible growth rates. Despite this, poor years tend to have a narrower distribution of possible growth rates and ~~were~~ are therefore less sensitive to changes in demographic parameters compared to good years (Figs. 3 and 4). Generally, the ~~PDF~~ p.d.f. of growth rates ~~were~~are skewed, with peaks being concentrated when germination rates ~~were~~are high and greater variation in λ_1 when germination is low (Figs. 3 and 4). However, we note these simulations only investigate~~d~~ the unstructured seed bank scenario. Further uncertainties exist with regards to the longevity and structure of the seed bank, both of which we~~re~~ only briefly explore~~d~~ in this study (Supplementary Table S1).

Elasticity analyses in most plant species, particularly the trees, reveal~~ed~~ survival of the adult or late juvenile stages to be the most important demographic process affecting λ_1. The addition of the seed bank did not change this pattern except in the case of *A. acanthocarpa* during 1997-98 and 1998-99 when the populations were declining (Supplementary Table S2). In the 1997-98 period, the elasticity of λ_1 to stasis in the third adult stage ~~was~~ is drastically reduced (0.5759 to 0.0647) following the inclusion of the seed bank, whereas the elasticity of λ_1 to stasis in the seed bank ~~was~~ is high (0.7919). Similarly, the elasticity of λ_1 to stasis in the first adult stage during the 1998-99 period, ~~was~~ is reduced from 0.3983 to 0.12 after seed bank addition, while the elasticity of λ_1 to survival in the seed bank ~~was~~ is high (0.5039).

Formatted: Font: Not Italic

Removal of the seed bank and uncertainty in seed bank presence

Removal of the seed bank ~~led to~~ further increases population ~~increase~~ growth when $\lambda_1 > 1$ since germination is no longer delayed, and ~~to~~ reduces~~d~~ population persistence when $\lambda_1 < 1$ since the buffering effect is no longer present (Table 3). The notable exception is the wetland herb *Helenium virginicum* (Asteraceae), ~~which experienced~~where removing the seed bank decreases ~~a comparatively high decrease in~~ λ_1 (1.1145 to 0.6985) ~~when the seed bank was removed~~ even though the population was increasing. This unusual response is likely a combination of extremely high fecundity into the seed bank and within-the-seed-bank stasis rate (1025 and 0.987 respectively), and the comparatively low fecundity and survival of the smallest stage (0.0075 and 0.199 respectively; see Supplementary File S1 for the original and updated matrices). Decreases in λ_1 ~~were~~ is also particularly substantial for *Carduus nutans* (Asteraceae) and *Digitalis purpurea* (Plantaginaceae), but ~~were~~ is much smaller or unchanged for the remaining species when the seed bank ~~was~~ is removed for poor years.

Extinction probabilities

Inclusion of the seed bank using point estimates for germination generally results~~ed~~ in longer times to local extinction (Supplementary Figure S1). Extinction times for *A. acanthocarpa* for the 1997-98 and 1998-99 matrices ~~received~~ have relatively high increases in extinction times (> 25 years) when the seed bank is included. In addition, certain populations of *A. canescens* and the perennial herb *D. purpurea* ~~were~~ are no longer predicted to go extinct in the next 100 years. The populations of the short-lived herb *C. nutans* ~~received~~ have comparatively smaller increases in extinction times. In one population, however, including the seed bank doubles~~d~~ the time to extinction from 8 to 16 years, which is relatively high when compared to the life expectancy of the species ($\eta_E \cong$2-3 years, based on the fundamental matrix estimation; Caswell, 2001).

The seed bank generally improves~~d~~ the outlook of stochastic projections for the estimation of extinction probability, but patterns vary~~ied~~ across species (Supplementary Figure S2). The cumulative probability of extinction at 100 years decrease~~d~~s drastically for *A. acanthocarpa* from 0.19 to 0.01 following inclusion of the seed bank. The extinction probability for *A. canescens* ~~showed~~ has an interesting pattern where the matrix model with the seed bank results initially in a higher extinction probability than the original in the first 38 years and eventually in a lower extinction probability, but in the remaining two cases, the seed bank has~~d~~ little effect on the extinction probability. Both populations of *D. purpurea* ~~were~~ are predicted to go extinct within 100 years~~. One population received a~~, however, there is a modest increase in extinction time (16 to 30 years) for one population and ~~while the increase was~~ a smaller increase in extinction time for the other (4 to 9 years).

Discussion

The presence of cryptic life stages and their exclusion from demographic models continues to be an impediment to understand and model the population dynamics. These cryptic stages are diverse, ranging from clonality, vegetative dormancy, hibernation, diapause, aestivation, etc., as is their potential impact on population dynamics and model outcomes. Using the seed bank as a well-known example of a cryptic life stage, we demonstrate the use of simulations as a means to reduce uncertainty in seed bank dynamics and to reveal the relative importance of the seed bank and its effect on model outcomes. Ignoring the seed bank in a population can produce variable degrees of uncertainty in the estimates of growth rates and extinction probabilities depending on the observed temporal variation in key vital rates. For example, models with high seedling survival and low temporal variation in demographic parameters ~~were~~ are more robust to seed bank uncertainty compared to models with high temporal variation in growth rates and demographic parameters. Furthermore, the potential impact of the seed bank can be influenced by the assumptions underpinning the models and the accuracy of the vital rates that are obtained.

Dealing with the seed bank when its presence is unknown or uncertain

Anticipating the species for which the seed bank will be an important factor is challenging due to the uncertainty in the demographic parameters of the seed bank stage across species, populations and time. Here, we provide a protocol outlining the different approaches one can take when conducting a plant population study (Fig. 5). First, the potential presence of a seed bank can be determined through a literature search of seed germination and dormancy values (Baskin and Baskin, 2014, is of particular note). Second, if a seed bank in the population is confirmed, the best option would be to obtain field data over the study period to provide estimates for the seed bank vital rates. Without this field data, characterising the potential role of the seed bank becomes exceptionally difficult and assumptions regarding fecundity, seed survival, viability and longevity need to be made. Prior knowledge from the literature can narrow down these possibilities to some extent, although seed characteristics and seed age composition can vary substantially between sites. Also, knowledge from the literature or from field and glasshouse trails may not necessarily reflect those of the study site and must therefore be interpreted with caution (Gross and Mackay, 2014). Finally, simulations can be used to complement these literature estimates to efficiently investigate the numerous possible assumptions regarding seed bank vital rates and determine plausible effects of a potential seed bank. Garcia et al. (2010), for example, performed simulations covering feasible information obtained in the field or from other publications for the long-lived orchid *Cypripedium calceolus* (Orchidaceae) to justify excluding the seed bank from their models. Likewise, for many of our species, without the benefit of greater observed temporal variation, ignoring the seed bank would appear to ~~have been~~be of minimal consequence for the original study, but it is still crucial that an investigation is performed to provide this justification. Otherwise, effective implementation and assessment of management strategies will be hindered as the demographic model is potentially missing a key component of the life cycle.

Uncertainty in the seed bank and its effect on model outcomes

Demographic uncertainty in the estimation of vital rates appears to be the major driver in influencing the model predictions but varies by species. ~~For example, l~~Long-lived tree species ~~were~~ are virtually unaffected by both the inclusion of the seed bank~~,~~ and variation in their demographic parameters. This is not entirely surprising as high adult survival and reduced mortality when environmental conditions are poor can be an alternative strategy to improve the long-term fitness (Clauss and Venable, 2000). In such species, it is likely that the seed bank will be of little consequence, although it would still be prudent to confirm this suspicion (e.g., using simulations) before ignoring the seed stage. Additionally, this does not necessarily mean the seed bank will be unimportant if the population were to experience a large disturbance. In contrast, in populations experiencing greater fluctuations in demographic parameters and population growth rates, the contribution of the seed bank is more visible as is evident for *A. acanthocarpa*, where previously high elasticities for juvenile and adult survival ~~were shifted~~shift towards survival in the seed bank during poor demographic years (Table S2).

Temporal variation and the type of demographic year are also factors that will affect of demographic uncertainty. Alterations in germination and viability rates during poor demographic years ($\lambda_1 < 1$) have~~d~~ relatively little impact on the growth rate compared to favourable years when λ_1 exceeded 1 (Fig. 4). Large reproductive failures during poor years would have resulted in low seed bank and seedling input, reducing the impact of varying seed bank vital rates on λ_1. In contrast, we would expect variations in germination and viability rates during poor years to have a much larger impact on long-term persistence and environmental stochasticity. Nonetheless, with the inclusion of prior information from the literature, the ~~PDF~~ p.d.f. of the distribution of growth rates during these poor demographic years shifts towards one (Fig. 3). We note however that this is assuming dormancy is higher (low germination) during poor demographic years and lower during productive years (Kalisz and McPeek, 1992), but seed dormancy needs not necessarily be correlated with fitness. Interestingly, the use of literature did not necessarily reduce uncertainty in the range of possible growth rates and in fact becomes more variable when dormancy is high (Fig. 3). This is an important result as an uninformed simulation may underestimate the level of uncertainty in demographic output if prior information suggests potentially high levels of dormancy. In contrast, when dormancy is suspected to be low, the impact of the seed bank will not be as great, thus variation in its demographic parameters will be of much less consequence.

Formatted: Font: Not Italic

In addition to variations in demographic parameters, changes to how the seed bank is included can impact the model structure and outcomes. Our uncertainty simulations investigate a 1-year seed bank as a counterfactual example to illustrate the need for including a seed bank. We acknowledge that while single point estimates of λ_1 under the unstructured scenario showed minor increases compared to the original models, this gradually reduces when additional age classes are introduced (See Table S1 to compare the growth rates). This is likely because a higher proportion of emergent seedlings are now assumed to have germinated from seeds produced in previous years, thus compounding the effect of delayed germination. Similarly, if we assume seeds do not survive beyond the first, second or third year, the effect of the seed bank is almost negligible (See Table S1). However, such estimates are overly conservative and most models that do include the seed bank do not make this assumption.

Another interesting scenario that can be tested is that of no seed bank.~~, and o~~Our removal simulations provides insights into the worst-case scenarios when the seed bank is ignored despite being crucial to persistence (as indicated by high elasticities to seed bank transitions). As expected, the implications of ignoring the seed bank in these cases were generally much more evident~~,~~ with decreases in λ_1 of up to 0.4, although these reflect the worst-case scenarios and changes in λ_1 would not generally be so dramatic. The assumption of fecundity being equal to that of the previous year ($F_j(t) = F_j(t)$) was necessary due to lack of information, however, interannual variation is likely to play an important role for many species. When variance in the fecundity is high, the contribution of the seed bank would be overestimated in good demographic years and underestimated in bad demographic years. This would likely have similar consequences on any changes in λ_1, thus we might expect the benefit of having a seed bank during poor years to also be underestimated.

Analysis of extinction probabilities over shorter time periods can provide additional insights not considered under asymptotic analyses. Comparing the difference in extinction times and cumulative extinction probabilities between models with and without the seed bank can provide an empirical measure of the uncertainty in population persistence when the seed bank is ignored, and this measure can vary substantially between years and species. Although increases in extinction times can appear trivial, they must also be considered in context of the life expectancy. For example, ~~T~~the seed bank had little effect on the times to extinction for the

Formatted: Font: Not Italic

Formatted: Font: Not Italic

populations of *C. nutans*~~ for example~~. However, given the short life expectancy of this species, even a small increase in extinction time may prove crucial.

Conclusions

There is still much progress to be made ~~with regards to~~concerning the seed bank~~,~~ and it is but one of many cryptic life stages with the potential to create biases in our model outputs. Adult plants, for example, can undergo vegetative dormancy, whereby individuals do not sprout above ground for one or more growing seasons, resulting in overestimated mortality rates and incorrect assignment of false birth events to re-emerging individuals (Logofet et al., 2019, Shefferson, 2009). Demographic studies on animals can be additionally problematic as certain stages of the life cycle may occupy inaccessible locations or be absent from sampling designs. Katzner et al. (2011) found raptor populations to be drastically underestimated as monitoring programs only captured the breeding portion of the population. In both plant and animal studies, incorrect estimates of population size, persistence, and occupancy in demographic models can result in suboptimal management decisions and make it difficult to assess the success of these management actions (Regan et al., 2011). While ~~this study~~we have focused exclusively on plant population models in this study, characterising the extent of exclusion of cryptic life stages in demographic models for animals is just as important and open for investigation with the availability of the COMADRE Animal Matrix Database (Salguero-Gómez *et al.*, 2016).

Acknowledgements

The authors would like to acknowledge the critical review by an anonymous reviewer that has greatly improved the manuscript, ~~thank~~ E. Crone and J. Ehrlén for helpful comments in previous versions of this manuscript, and S. Ramula and members of the ARC-NZ Research Network for Vegetation Function Plant Population Syndromes working group.

Author Contributions

The first and last author conceived and designed the study. The first author collected the data and performed the analysis. All authors contributed to the study methodology and writing of the manuscript.

References

Abe S, Motai H, Tanaka H, Shibata M, Kominami Y, Nakashizuka T (2008). Population maintenance of the short-lived shrub *Sambucus* in a deciduous forest. Ecology, 89, 1155-1167.

Adams VM, Marsh DM, Knox JS (2005). Importance of the seed bank for population viability and population monitoring in a threatened wetland herb. Biol. Conserv., 124, 425-436.

Angert AL (2006). Demography of central and marginal populations of monkey flowers (*Mimulus cardinalis* and *M. lewisii*). Ecology, 87, 2014-2025.

Baskin CC, Baskin GM (2014). Seeds: Ecology, biogeography, and evolution of dormancy and germination, Academic Press, San Diego, California.

Burns JH, Blomberg SP, Crone EE, Ehrlén J, Knight TM, Pichancourt JP, Ramula S, Wardle GM, Buckley YM (2010). Empirical tests of life-history evolution theory using phylogenetic analysis of plant demography. J. Ecol., 98, 334-344.

Caswell H (2001). Matrix population models: construction, analysis and interpretation, Sinauer Associates, Inc., Sunderland, Massachusetts.

Chien PD, Zuidema PA, Nghia NH (2008). Conservation prospects for threatened Vietnamese tree species: results from a demographic study. Popul. Ecol., 50, 227-237.

Clauss MJ, Venable DL (2000). Seed germination in desert annuals: an empirical test of adaptive bet hedging. Am. Nat., 155, 168-186.

COMPADRE Plant Matrix Database. 2019. *COMPADRE Plant Matrix Database v5.0.0.* [Online]. Max Planck Institute for Demographic Research (Germany). [Accessed 20th May 2019].

Crone EE, Menges ES, Ellis MM, Bell T, Bierzychudek P, Ehrlén J, Kaye TN, Knight TM, Lesica P, Morris WF, Oostermeijer G, Quintana-Ascencio PF, Stanley A, Ticktin T, Valverde T, Williams JL (2011). How do plant ecologists use matrix population models? Ecol. Lett., 14, 1-8.

de Kroon H, Plaisier A, J.V. G (1987). Density dependent simulation of the population dynamics of a perennial grassland species, *Hypochaeris radicata*. Oikos, 50, 3-12.

de Kroon H, van Groenendael J, Ehrlen J (2000). Elasticities: a review of methods and model limitations. Ecology (Wash. D. C.), 81, 607-618.

Doak DF, Thomson D, Jules ES (2002). Population viability analysis for plants: understanding the demographic consequences of seed banks for population health. In: Beissinger SR, McCullough DR (eds.) Population Viability Analysis. The University of Chicago Press, Chicago, Illinois, USA.

Fenner M (2000). Seeds: the ecology of regeneration in plant communities, CABI Publishing, Wallingford, UK.

Garcia MB, Goni D, Guzman D (2010). Living at the edge: local versus positional factors in the long-term population dynamics of an endangered orchid. Conserv. Biol., 24, 1219-1229.

Geiser F (2004). Metabolic rate and body temperature reduction during hibernation and daily torpor. Annu. Rev. Physiol., 66, 239-274.

Gremer JR, Crone EE, Lesica P (2012). Are dormant plants hedging their bets? Demographic consequences of prolonged dormancy in variable environments. Am. Nat., 179, 315-327.

Gross C, Mackay D (2014). Two decades of demography reveals that seed and seedling transitions limit population persistence in a translocated shrub. Annals of Botany.

Harper JL (1977). Population biology of plants, Academic Press, London.

Jongejans E, Shea K, Sheppard AW (2006). What controls the population dynamics of the invasive thistle *Carduus nutans* in its native range? J. Appl. Ecol., 43, 877-886.

Kalisz S, McPeek MA (1992). Demography of an age-structured annual: resampled projection matrices, elasticity analyses, and seed bank effects. Ecology, 73, 1082-1093.

Katzner TE, Ivy JAR, Bragin EA, Milner-Gulland E, DeWoody JA (2011). Conservation implications of inaccurate estimation of cryptic population size. Anim. Conserv., 14, 328-332.

Kaye TN, Pyke DA (2003). The effect of stochastic technique on estimates of population viability from transition matrix models. Ecology, 84, 1464-1476.

Kisanuki H, Nakasu M, Nakai A, Yurugi Y (2008). Predicting the population dynamics of three understory broad-leaved evergreen species under the influence of Sika deer in primary and secondary forests of mid-western Japan. J. For. Res., 13, 52-58.

Kouassi KI, Barot S, Gignoux J, Bi IAZ (2008). Demography and life history of two rattan species, *Eremospatha macrocarpa* and *Laccosperma secundiflorum*, in Cote d'Ivoire. J. Trop. Ecol., 24, 493-503.

Logofet DO, Kazantseva ES, Belova IN, Onipchenko VG (2019). Disappointing viability forecat for a local population of *Androsace albana* in a random environment. Zh. Obshch. Biol., 80, 1-13.

Mondragon D (2009). Population viability analysis for *Guarianthe aurantiaca*, an ornamental epiphytic orchid harvested in Southeast Mexico. Plant Species Biol., 24, 35-41.

Morris WF, Doak DF (2002). Quantitative conservation biology: theory and practice of population viability analysis, Sinauer Associates, Sunderland, Massachusetts.

Parker IM (2000). Invasion dynamics of *Cytisus scoparius*: a matrix model approach. Ecol. Appl., 10, 726-743.

Quitete Portela RDC, Bruna EM, Maes dos Santos FA (2010). Demography of palm species in Brazil's Atlantic forest: a comparison of harvested and unharvested species using matrix models. Biodivers. Conserv., 19, 2389-2403.

R Core Team. 2015. *R: A language and environment for statistical computing* [Online]. Vienna, Austria: R Foundation for Statistical Computing. [Accessed 3rd October 2018].

Regan TJ, Chades I, Possingham HP (2011). Optimally managing under imperfect detection: a method for plant invasions. J. Appl. Ecol., 48, 76-85.

Salguero-Gómez R, Jones OR, Archer CR, Buckley YM, Che-Castaldo J, Caswell H, Hodgson D, Scheuerlein A, Conde DA, Brinks E, de Buhr H, Farack C, Gottschalk F, Hartmann A, Henning A, Hoppe G, Römer G, Runge J, Ruoff T, Wille J, Zeh S, Davison R, Vieregg D, Baudisch A, Altwegg R, Colchero F, Dong M, de Kroon H, Lebreton J-D, Metcalf CJE, Neel MM, Parker IM, Takada T, Valverde T, Vélez-Espino LA, Wardle GM, Franco M, Vaupel JW (2015). The COMPADRE Plant Matrix Database: an open online repository for plant demography. J. Ecol., 103, 202-218.

Salguero-Gómez R, Plotkin JB (2010). Matrix dimensions bias demographic inferences: implications for comparative plant demography. Am. Nat., 176, 710-722.

Shefferson RP (2009). The evolutionary ecology of vegetative dormancy in mature herbaceous perennial plants. J. Ecol., 97, 1000-1009.

Sletvold N, Rydgren K (2007). Population dynamics in *Digitalis purpurea*: the interaction of disturbance and seed bank dynamics. J. Ecol., 95, 1346-1359.

Stöcklin J, Fischer M (1999). Plants with longer-lived seeds have lower local extinction rates in grassland remnants 1950-1985. Oecologia, 120, 539-543.

Tanner JE (2001). The influence of clonality on demography: patterns in expected longevity and survivorship. Ecology, 82, 1971-1981.

Venable DL (2007). Bet hedging in a guild of desert annuals. Ecology, 88, 1086-1090.

Verhulst J, Montana C, Mandujano MC, Franco M (2008). Demographic mechanisms in the coexistence of two closely related perennials in a fluctuating environment. Oecologia (Berl.), 156, 95-105.

Vitalis R, Glémin S, Olivieri I (2004). When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. The American Naturalist, 163, 295-311.

Warner RR, Chesson PL (1985). Coexistence mediated by recruitment fluctuations: a field guide to the storage effect. Am. Nat., 125, 769-787.

Tables

Table 1. Comparison of studies from the MPM2008-2011 database with a previous review by (Doak et al., 2002) examining the inclusion of the seed bank in plant population models. Studies were classified as *justified exclusion* if a valid reason was provided for not including the seed bank (e.g. the authors found no evidence of a seed bank) and *unjustified exclusion* if the possibility of a seed bank was not explored or mentioned. Studies from COMPADRE were not analysed to determine whether the exclusion of the seed bank was justified.

Population models and seed banks	**Doak et al. (2002)**	**MPM2008-2011**	**COMPADRE v.5.0.0**
Included	34 (48.6%)	16 (29.1%)	292 (34.0%)
Justified exclusion	6 (8.6%)	13 (23.6%)	568 (66.0%)
Unjustified exclusion	30 (42.9%)	26 (47.3%)	
Total studies	70	55	860

Table 2. Comparison of population growth rates (λ_1) following the addition of a seed bank. *Pop.* refers to populations of the same species in the same study. *Avg.* represents the growth rate from the mean matrix for each population. *Stoch.* represents the mean stochastic growth rate from populations projected over 10,000 times. SB denotes single point estimates from models with germination and viability equal to that obtained from the literature. Growth rates from Monte Carlo simulations were obtained from the mean of 10,000 trials with parameters for germination, viability, seed survival, and seedling survival drawn from uniform distributions when priors are uninformed, while germination and viability were drawn from beta distributions when the priors were informed. 95% confidence intervals for Monte Carlo simulations and stochastic growth rates are displayed in round brackets.

Study	**Life Form**	**Pop**	**Period**	**No SB**	**SB**	**Monte Carlo sim. uninformed prior**	**Monte Carlo sim. informed prior**
Single time period							
Chien et al. (2008)							
Calocedrus macrolepus	Tree	1	1	0.9683	0.9715	0.9700 (0.9683, 0.9785)	0.9699 (0.9683, 0.9752)
Parashorea chinensis	Tree	1	1	0.9948	0.9955	0.9963 (0.9948, 1.0040)	0.9954 (0.9948, 0.9978)
Pinus kwangtungensis	Tree	1	1	0.9776	0.9786	0.9786 (0.9776, 0.9834)	0.9782 (0.9776, 0.9799)
Kisanuki et al. (2008)							
Illicium anisatum	Tree	1	1	1.1271	1.1647	1.1344 (1.1258, 1.1765)	1.1408 (1.1263, 1.1778)
		2	1	1.1864	1.2276	1.1931 (1.1828, 1.2363)	1.2008 (1.1840, 1.2416)
Abe et al. (2008)							
Sambucus racemosa	Shrub	1	1	1.1696	1.1690	1.1789 (1.1661, 1.2355)	1.1951 (1.1674, 1.2639)
		2	1	1.0007	1.0012	1.0109 (1.0007, 1.0622)	1.0255 (1.0008, 1.0963)
		3	1	2.2227	2.2031	2.1588 (2.0468, 2.2210)	2.1459 (2.0551, 2.2125)
Kouassi et al. (2008)							
Eremospatha macrocarpa	Tree	1	1	0.9792	0.9811	0.9806 (0.9792, 0.9887)	0.9805 (0.9792, 0.9842)

Laccosperma secundiflorum	Tree	1	1	0.9606	0.9661	0.9638 (0.9607, 0.9815)	0.9621 (0.9607, 0.9658)
Multiple time periods							
Mondragon (2009)							
Guarianthe aurantiaca	Herb	1	1	0.9851	0.9881	0.9917 (0.9851, 1.0253)	0.9872 (0.9851, 0.9929)
		1	2	0.9895	0.9908	0.9929 (0.9895, 1.0107)	0.9905 (0.9895, 0.9931)
			Avg.	0.9869	0.9891	-	-
			Stoch.	0.9869 (0.9868, 0.9869)	0.9891 (0.9891, 0.9892)	-	-
Quitete Portela et al. (2010)							
Astrocaryum aculeatissimum	Tree	1	1	1.0093	1.0106	1.0111 (1.0093, 1.0209)	1.0100 (1.0093, 1.0112)
		1	2	0.9964	0.9968	0.9971 (0.9964, 1.0001)	0.9966 (0.9964, 0.9972)
			Avg.	0.9998	1.0085	-	-
			Stoch.	0.9998 (0.9998, 0.9998)	1.0005 (1.0005, 1.0005)	-	-
Geonoma schottiana	Tree	1	1	1.0043	1.0240	1.0114 (1.0050, 1.0442)	1.0213 (1.0050, 1.0712)
		1	2	0.9823	1.0002	1.0115 (1.0050, 1.0453)	1.0216 (1.0050, 1.0721)
			Avg.	0.9949	1.0150	-	-
			Stoch.	1.0050 (1.0050, 1.0050)	1.0146 (1.0145, 1.0148)	-	-
Verhulst et al. (2008)							
Atriplex acanthocarpa	Shrub	1	1	2.4526	2.3455	2.3016 (2.0678, 2.4476)	2.3541 (2.0759, 2.4489)
		1	2	0.7301	0.8305	0.7418 (0.7302, 0.8288)	0.7508 (0.7303, 0.8362)
		1	3	0.7436	0.8862	0.7712 (0.7438, 0.8951)	0.7924 (0.7444, 0.9051)
			Avg.	1.3461	1.3351	-	-
			Stoch.	1.1486 (1.1460, 1.1512)	1.1792 (1.1773, 1.1811)	-	-
Atriplex canescens	Shrub	1	1	1.7211	1.7012	1.6883 (1.6184, 1.7315)	1.6791 (1.6160, 1.7175)
		1	2	0.9451	0.9451	0.9451 (0.9451, 0.9453)	0.9451 (0.9451, 0.9451)
		1	3	0.8805	0.8835	0.8810 (0.8805, 0.8823)	0.8805 (0.8805, 0.8807)
			Avg.	1.1346	1.1648	-	-
			Stoch.	1.0775 (1.0730, 1.0779)	1.1140 (1.1120, 1.1159)	-	-

Table 3. Comparison of population growth rates following removal of the seed bank. Studies are classified on the basis of variability in transition elements that was reported or number of time periods that were provided. Pop. refers to different populations of the same species in the same study. *Avg.* represents the growth rate from the mean matrix. *Stoch.* represents the mean stochastic growth rate from populations projected over 10,000 times. Confidence intervals for stochastic simulations when the seed bank was removed are shown in brackets.

Study	Life Form	Pop.	Period	SB	No SB
No variability reported					
Adams et al. (2005)					
Helenium virginicum	Herb	1	1	1.1145	0.6985
de Kroon et al. (1987)					
Hypochaeris radicata	Herb	1	1	0.2973	0.2154
Variability, single time period					
Jongejans et al. (2006)					
Carduus nutans	Herb	1	1	0.4680	0.3894
		2	1	0.6270	0.3771
		3	1	0.6020	0.5524
Multiple time periods					
Angert (2006)					
Mimulus lewisii	Herb	1	1	0.7778	0.7778
		1	2	0.6863	0.6815
		1	3	0.9959	0.9943
		1	Avg.	0.8541	0.8332
		1	Stoch.	0.8423	0.8181
				(0.8410, 0.8435)	(0.8222, 0.8233)
		2	1	1.0154	1.0053
		2	2	0.6978	0.6917

		2	3	1.3203	1.3746
		2	Avg.	1.0158	1.0410
		2	Stoch.	0.9642	0.9949
				(0.9769, 0.9787	(0.9935, 0.9963)
		3	1	1.1052	1.1094
		3	2	1.1197	1.1450
		3	3	0.9523	0.9507
		3	Avg.	1.0510	1.0583
		3	Stoch.	1.0472	1.0537
				(1.0470, 1.0500)	(1.0544, 1.0557)
		4	1	1.5841	1.6704
		4	2	1.0459	1.0470
		4	3	1.1227	1.3083
		4	Avg.	1.3112	1.3448
		4	Stoch.	1.2919	1.3198
				(1.2908, 1.2930)	(1.3192, 1.3205)
Parker (2000)					
Cytisus scoparius	Herb	1	1	2.1963	2.3331
		1	2	2.2629	2.3858
		1	Avg.	2.5019	2.7128
		1	Stoch.	2.4919	2.6927
				(2.4879, 2.4960)	(2.6868, 2.6986)
		2	1	1.7097	1.7407
		2	2	1.2542	1.2406
		2	Avg.	1.5681	1.5552
		2	Stoch.	1.5507	1.5351
				(1.5495, 1.5518)	(1.5343, 1.5359)
		3	1	1.0040	1.0040
		3	2	0.9295	0.9287
		3	Avg.	0.9748	0.9746
		3	Stoch.	0.9740	0.9738
				(0.9734, 0.9746)	(0.9700, 0.9776)

Sletvold and Rydgren (2007)					
Digitalis purpurea	Herb	1	1	0.7533	0.3174
		1	2	0.7831	0.5536
		1	Avg.	0.7715	0.5071
		1	Stoch.	0.7713	0.5008
				(0.7712, 0.7713)	(0.5002, 0.5013)
		2	1	0.6831	0.2983
		2	2	1.0198	0.9379
		2	Avg.	0.8565	0.6701
		2	Stoch.	0.8388	0.5794
				(0.8384, 0.8392)	(0.5787, 0.5802)

Figures

Fig. 1. Life cycle graphs corresponding to: **A)** a plant life cycle consisting of two stages: the seedlings, Se, and the adults, Ad; **B)** a plant life cycle consisting of seeds that have remained in the seed bank for 1 year, S1, up to *n* years, Sn (Eq. 1), and **C)** a plant life cycle with an unstructured seed bank, SB, (Eq. 2). The transition rate $f_{\mathrm{Se,Ad}}$ gives the fecundity into the seedling stage while $f_{\mathrm{S1,Ad}}$ and $f_{\mathrm{SB,Ad}}$ gives the fecundity into the seed bank. **D)** Details of the derivation of Eq. 1 in which the parameters are contained in boxes (see formal notation in the text). Tracking the vital rates required to reach the seedling stage from the adult stage either directly or via the seed bank demonstrates how each term in Eq. 1 was derived.

Fig. 2. Scatterplot comparing population growth rates between models when the seed bank is included (y-axis) and when the seed bank is not included (x-axis). “Addition” refers to models in which we added the seed bank the model (Table 2), while “Removal” refers to models where we removed the seed bank (Table 3). Growth rates for models where the seed bank was added were taken from Monte Carlo simulations with germination prior guided by the literature. Points above the one-to-one line indicate the seed bank has a positive effect on λ while points below indicate a negative effect. Points within the red section are declining populations ($\lambda_1 < 1$) while those outside are increasing populations ($\lambda_1 > 1$).

Fig. 3. Probability density function of the distribution of population growth rates for *Atriplex acanthocarpa* from **A)** 1996-1997, **B)** 1997-1998, and **C)** 1998-1999 based on 10,000 Monte Carlo simulations calculated using an uninformed prior on germination (uniform distribution) and an informed prior on germination (beta distribution with the mode equal to the germination

rate obtained from the literature). Red dashed lines indicate the mean growth rate using uninformed and informed priors as reported in Table 2.

Fig. 4. Population growth rate (λ_1) as a function of germination rate from the seed bank for **A)** *Atriplex acanthocarpa*, **B)** *Atriplex canescens*, **C)** *Illicium anisatum*, and **D)** *Geonoma schottiana*. Population growth rates were calculated at seed viability rates of 1 (green line), 0.5 (blue line), and 0.1 (red line) for the germination rates between 0 and 1 at intervals of 0.01. The germination rate for *G. schottiana* is restricted between 0 and 0.69 to not exceed the total seed survival estimate obtained from the literature.

Formatted: Font: Not Italic

Fig. 5. Flow chart describing different approaches to dealing with a potential seed bank when modelling plant population dynamics. When the presence of the seed bank is uncertain, it is advised that researchers explore this possibility either through literature confirmation, or investigating the soil reserves for viable seeds. By not investigating, the exclusion is unjustified, resulting in the highest possible uncertainty. Justified exclusion, by either proving it does not exist or showing it is unimportant via simulations, gives the highest level of certainty in the model outcomes. Including the seed bank, either by using the literature or data obtained in the field would unavoidably result in some amount of uncertainty from estimating the vital rates, but is preferable to unjustified exclusion.

Fig. 1

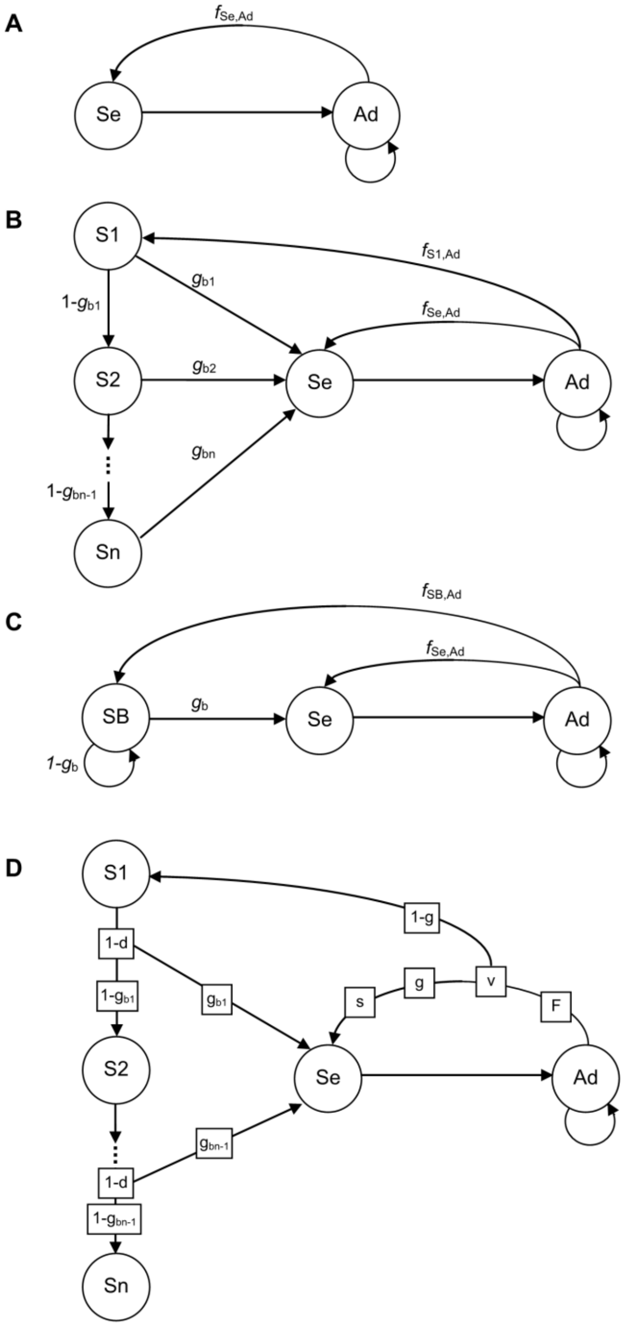

A
$f_{Se,Ad}$
Se
Ad
B
S1
$f_{S1,Ad}$
g_{b1}
1-g_{b1}
$f_{Se,Ad}$
S2
g_{b2}
Se
Ad
g_{bn}
1-g_{bn-1}
Sn
C
$f_{SB,Ad}$
$f_{Se,Ad}$
SB
g_b
Se
Ad
1-g_b
D
S1
1-g
1-d
g
v
s
F
1-g_{b1}
g_{b1}
S2
Se
Ad
g_{bn-1}
1-d
1-g_{bn-1}
Sn

Fig. 2

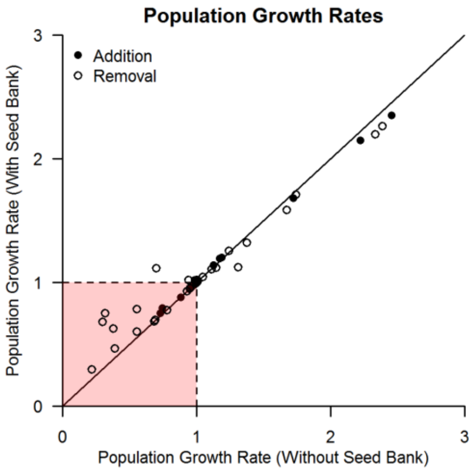

Population Growth Rates
Addition
Removal
Population Growth Rate (With Seed Bank)
Population Growth Rate (Without Seed Bank)
0
1
2
3
0
1
2
3

Fig. 3

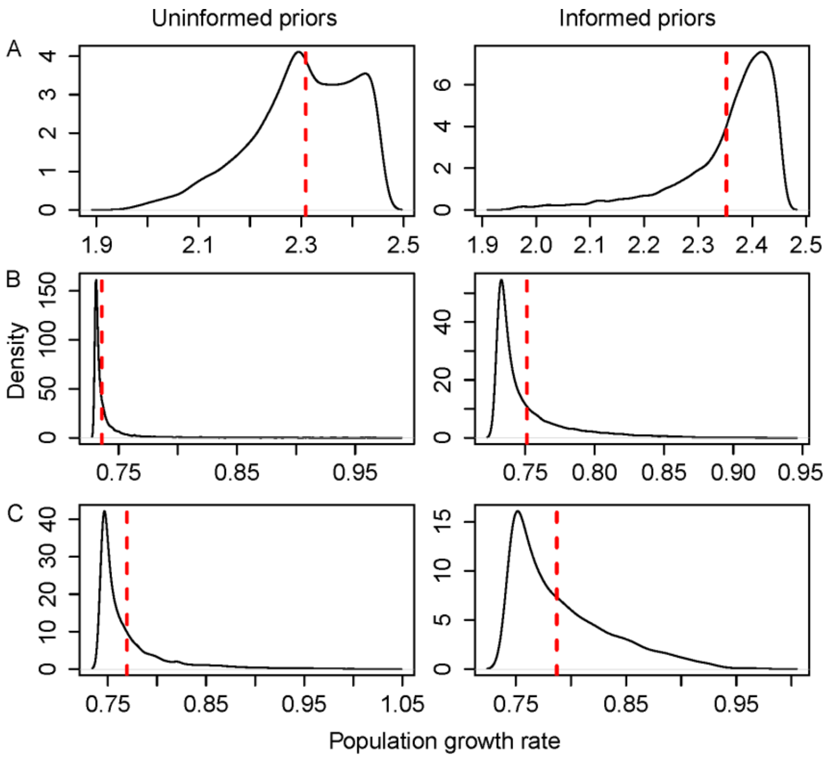
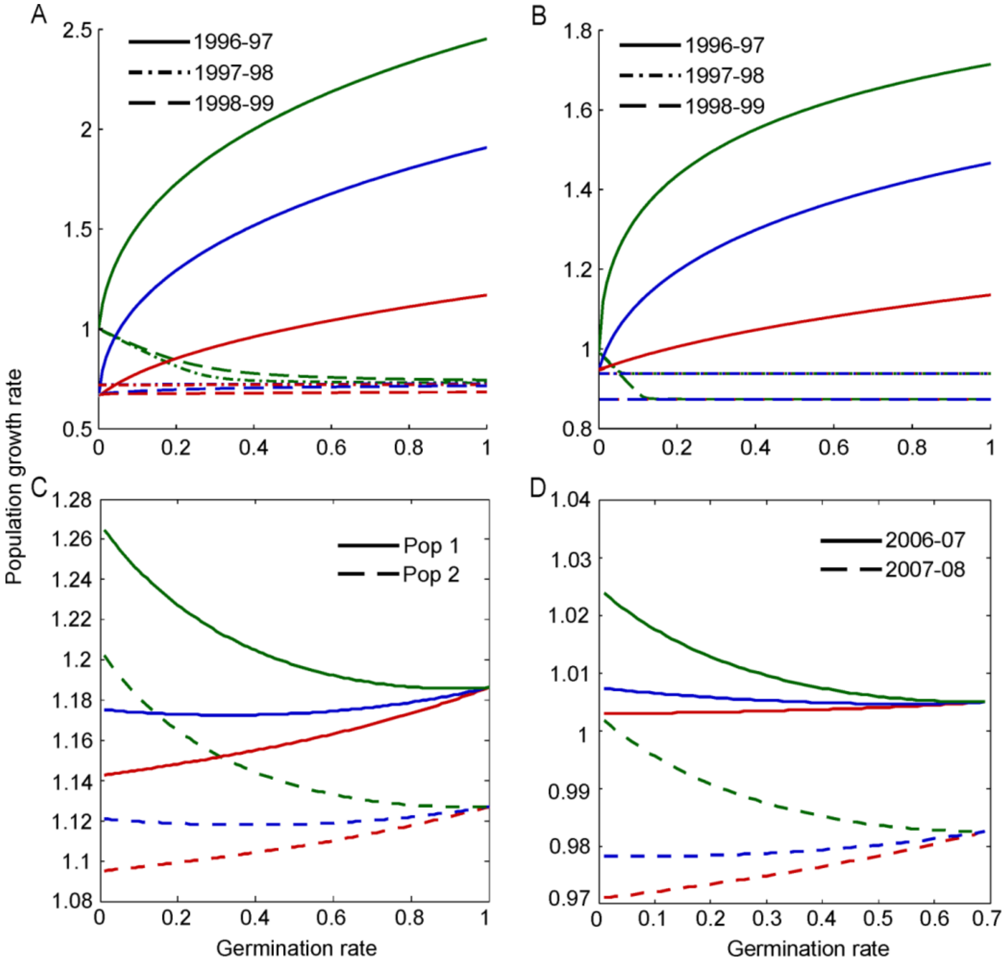
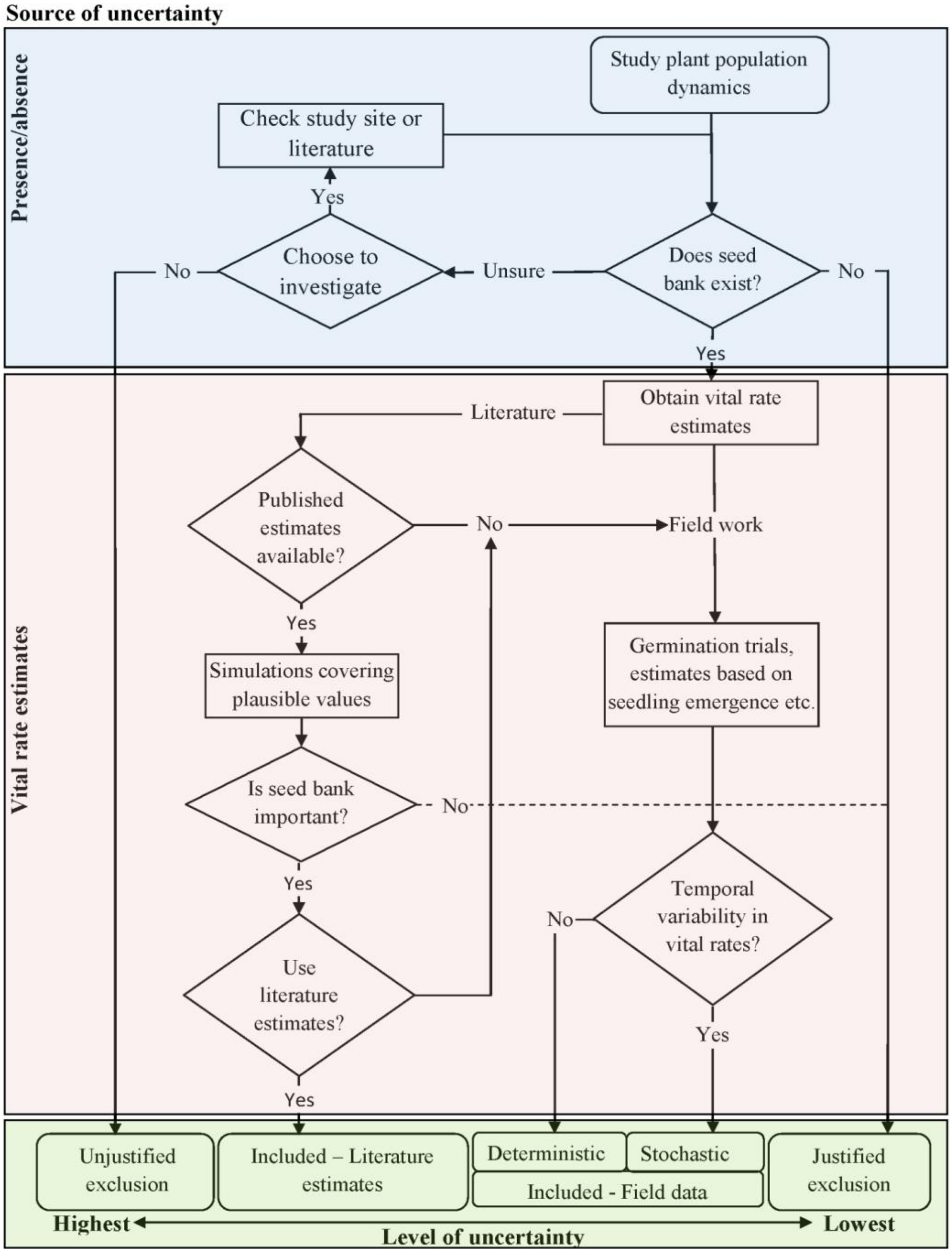

Uninformed priors
Informed priors
A
B
C
Density
0
1
2
3
4
0
2
4
6
1.9
2.1
2.3
2.5
1.9
2.0
2.1
2.2
2.3
2.4
2.5
0
50
100
150
0
20
40
0.75
0.85
0.95
0.75
0.80
0.85
0.90
0.95
0
10
20
30
40
0
5
10
15
0.75
0.85
0.95
1.05
0.75
0.85
0.95
Population growth rate

Fig. 4


A
2.5
2
1.5
1
0.5
0
0.2
0.4
0.6
0.8
1
1996-97
1997-98
1998-99
B
1.8
1.6
1.4
1.2
1
0.8
0
0.2
0.4
0.6
0.8
1
1996-97
1997-98
1998-99
Population growth rate
C
1.28
1.26
1.24
1.22
1.2
1.18
1.16
1.14
1.12
1.1
1.08
0
0.2
0.4
0.6
0.8
1
Pop 1
Pop 2
Germination rate
D
1.04
1.03
1.02
1.01
1
0.99
0.98
0.97
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2006-07
2007-08
Germination rate

660 Fig. 5


Source of uncertainty
Presence/absence
Study plant population dynamics
Check study site or literature
Yes
No
Choose to investigate
Unsure
Does seed bank exist?
No
Yes
Vital rate estimates
Obtain vital rate estimates
Literature
Published estimates available?
No
Field work
Yes
Simulations covering plausible values
Germination trials, estimates based on seedling emergence etc.
Is seed bank important?
No
Yes
Temporal variability in vital rates?
No
Use literature estimates?
Yes
Yes
Unjustified exclusion
Included – Literature estimates
Deterministic
Stochastic
Included - Field data
Justified exclusion
Highest
Lowest
Level of uncertainty

661

Supporting Information

Table S1

Population growth rates from seed bank models, Monte Carlo simulations and alternative formulations with 2 and 3 year longevity

Table S2

Elasticity matrices for 1997 and 1998 for *Atriplex acanthocarpa* with and without the seed bank

Figure S1

Log population size projections for single point matrices with and without the seed bank

Figure S2

Cumulative extinction probabilities of quasi-extinction

Method S1

Example demonstrating methodology for including the seed bank stage and alternative formulations

Method S2

Example of seed bank stage removal from a matrix model for *Digitalis purpurea*

File S1

Excel file containing original and updated matrices used for addition and removal simulations

- Inclusion of seed bank reduced extinction times and improved long-term persistence
- Published literature and simulations can assess the importance of the seed bank
- Inclusion of the seed bank may change model predictions made otherwise

Consequences of neglecting cryptic life stages from demographic models

Vuong Nguyen[1]*, Yvonne M. Buckley[2,3], Roberto Salguero-Gómez[2,4,5] and Glenda M. Wardle[1]

[1]*Desert Ecology Research Group, School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia*

[2]*School of Biological Sciences, The University of Queensland, Queensland, 4072, Australia*

[3]*School of Natural Sciences, Department of Zoology, Trinity College Dublin, Dublin, Ireland*

[4]*Evolutionary Biodemography Laboratory, Max Planck Institute for Demographic Research, Rostock 18057, Germany*

[5]*Department of Zoology, University of Oxford, Oxford, UK*

*Correspondence author. E-mail: phuc.nguyen@sydney.edu.au

Abstract

Information on individuals from *all* stages of life is crucial to explore their ecology, evolution and conservation biology. However, the life cycles of many species contain cryptic life stages that are difficult to detect and track over time and are therefore omitted from demographic models. One example is the dormant seed bank, an evolutionary bet-hedging mechanism that buffers plant populations in variable environments. To evaluate this methodological oversight, we conduct simulations to explore the effect of seed bank parameter uncertainties on demographic outputs such as the deterministic (λ_1) and stochastic population growth rate (λ_S), and extinction probabilities of 12 plant species. We have used uninformed and informed priors for seed bank parameters based on literature estimates, and reconstructed published models in which the seed bank was excluded without justification. Trials on removing the seed stage from models (6 species) explored the worst-case scenario for ignoring the seed bank. Inclusion of a seed bank and demographic uncertainty in seed bank parameters have little impact on stable populations ($\lambda_1 \approx 1$) with high post-seedling survival. When populations deviate from stability or demonstrate temporal demographic variation, greater changes in λ_1 and the range of possible growth rates caused by demographic uncertainty are observed. As expected, decreasing populations ($\lambda_1 < 1$) benefit from the inclusion of a seed bank through increases in the growth rate and extinction times, whereas increasing populations ($\lambda_1 > 1$) are slowed down. While germination estimates from the literature cannot accurately reflect those obtained in the field, they provide a starting point to assess the relative importance of a seed bank. The exclusion of the seed bank must be justified by confirming that dormancy is either non-existent or not important. Accounting for cryptic stages in demographic models will produce better informed management decisions for threatened or invasive species.

Keywords: Comparative demography, cryptic life stages, matrix population models, plant population ecology, seed bank, seed dormancy.

Highlights

- Inclusion of seed bank reduced extinction times and improved long-term persistence
- Published literature and simulations can assess the importance of the seed bank
- Inclusion of the seed bank may change model predictions made otherwise

Introduction

Demographers have long recognised that individuals in a population span across different stages of their life cycle, from propagules up to reproductive adults, and that these stages contribute to its persistence in different ways – failing to account for all these stages can therefore be misleading (Caswell, 2001). However, the life cycles of many species contain cryptic life stages that may be difficult to detect or sample. These cryptic life stages occur over a range of taxonomic groups, and may include hibernating or dormant individuals, clonal individuals, eggs, mobile propagules, or simply stages occupying inaccessible locations (Geiser, 2004, Katzner et al., 2011, Shefferson, 2009, Tanner, 2001). The dormant seed bank is perhaps one of the most common examples of the cryptic life stage whereby seeds enter prolonged dormancy (>1 year) within the soil or above ground in woody cones, and can make up a significant portion of the population (Baskin and Baskin, 2014, Harper, 1977). By spreading germination through time, seed dormancy acts as a bet-hedging mechanism mitigating the effects of unfavourable years in exchange for reduced proliferation during favourable years (Gremer et al., 2012, Venable, 2007). In extreme cases, seed banks can prevent local extinctions when above-ground populations are killed off entirely (Stöcklin and Fischer, 1999). Seed dormancy has also been shown to promote species co-existence in variable environments (Warner and Chesson, 1985) and act as a temporal reservoir for genetic diversity (Vitalis et al., 2004).

With such substantial consequences for population dynamics, it would be a major oversight to exclude the seed bank, or any cryptic life stage, from population estimates and demographic analyses. This argument is indeed not new: the need to incorporate the seed bank into demographic models has been recognised for decades (Harper, 1977, Kalisz and McPeek, 1992). Matrix population models, which divide individuals into their respective stage classes, are the most common tool for modelling plant populations – the online database of plant matrix models, COMPADRE v5.0.0 contains 8,906 published matrices for 759 species (COMPADRE Plant Matrix Database, v5.0.0., 2019, Salguero-Gómez et al., 2015). These models have a long history in the management of threatened and invasive species (Crone et al., 2011). Yet, a previous review of 70 published matrix population models found almost half (43%) had excluded the seed bank stage without justification (Doak et al., 2002). Surprisingly, we have found that this proportion has remained unchanged in the last decade (47%; Table 1). We acknowledge that including the seed bank is not straightforward as the estimation of vital rates such as survival and germination generally rely on additional laboratory experiments (Crone et al., 2011), and even these would inflate actual *in situ* estimates (Gross and Mackay, 2014). Furthermore, unless such experiments are rigorous enough to include age-specific vital rates (Kalisz and McPeek, 1992), the seed bank has been generally assumed to be unstructured. Excluding the seed bank may be justified if it is discovered the species, or population, does not possess dormant seeds (see Baskin and Baskin, 2014 for review on the presence of seed dormancy), or if it is mathematically and empirically justified in a modified model (Logofet et al. 2017), however, these cases are the minority compared to those that do not provide justification (Doak et al., 2002; Table 1). Unjustified exclusion can have real-world consequences if, for example, the seed bank of an invasive species recolonises after the adult individuals have been eliminated (Regan et al., 2011), or if the population size and persistence of threatened species is underestimated (Doak et al., 2002).

The practical issues of including the seed bank and its exclusion from demographic models raise two important questions. First, how can we bridge the gap in our understanding of whether a seed bank is necessary to adequately model the population dynamics in the absence of field data? Second, how does uncertainty in the presence of a seed stage and its transition rates translate into uncertainty in the model? Addressing these questions using a simulation-based approach can be useful for investigating a range of scenarios and assumptions but may not be biologically meaningful without application to real-world examples. An empirical approach, on the other hand, may be limited in its scope if it were to focus on a single case study, as there are clear difficulties in obtaining individual-level records for published models. Using published matrix models that have excluded the seed bank without justification and information on dormancy and germination rates from the literature, we implement a combination of these approaches and perform a comparative analysis across a range of species and life forms (12 species including eight trees, three shrubs and one herb). We conduct Bayesian Monte Carlo simulations on these published models using uninformed and informed priors (based on literature estimates) for seed bank vital rates to gauge the level of uncertainty when the seed bank presence and parameters are unknown. To complement these addition simulations, we also investigate the implications of assuming the absence of a persistent seed bank by intentionally removing the seed bank from models which did include it.

We make a number of predictions based on the evolutionary advantages of seed dormancy. First, we might expect the relative importance of seed banks to be greater for species in unpredictable environments with variable post-seedling survival compared to stable populations of long-lived species with consistently high adult survival (Venable, 2007). Furthermore, the deterministic growth rate (λ_1), which assumes vital rates remain constant, should increase during poor demographic years ($\lambda_1 < 1$) and decrease during productive years ($\lambda_1 > 1$) to reflect the function of the seed bank as a bet hedging mechanism against unfavourable years (Gremer et al., 2012, Kalisz and McPeek, 1992). Similarly, elasticities (the relative influence of each demographic process to the population growth rate; de Kroon et al., 2000) for seed bank transition rates should be greater during poor years. We should also observe increases in times to local extinction and stochastic growth rates (λ_S) which incorporate temporal variation in vital rates. Using both an empirical and simulation-based approach, we provide more robust projections of the demographic fates of populations and help to motivate the inclusion of the seed bank into demographic models.

Materials and Methods

Database of population matrix models

At the time this study was conducted, the COMPADRE Plant Matrix Database had not yet become open-access online (Salguero-Gómez et al., 2015), with the most recent update consisting of 204 terrestrial plant species up to 2007, which we hereafter refer to as the "MPM2007" database constructed by the Vegetation Function Working Group (see Burns et al., 2010 for further information). We have used the ISI Web of Science to identify published plant matrix population models from 2008-2011, which we hereafter refer to as the "MPM2008-2011" database. These studies would later go on to be added to the COMPADRE database. Examples of keywords used in our search include "matrix model*", "transition matrix", "demographic model*", "population model*", "population viability analysis", "population growth rate", and "population dynamics". The addition of * at the end of a search term acts as a wildcard and allows for alternative forms of a given word, e.g. model(s), model(ling). We have found 55 studies with matrix models for 77 terrestrial plant species including 38 trees, 27 herbs, 8 shrubs and 4 succulents.

Confirmation of the seed bank

Using Baskin and Baskin (2014) and the species name as a search term in ISI Web of Science, we have verified whether a seed bank was reported for species in which seed bank exclusion was unjustified (by not mentioning the seed bank; Table 1). While each population may not necessarily possess a seed bank, we can still determine the impact of its presence on model outcomes. In total, we have constructed new matrix population models for the 12 plant species for which published germination data on either the species or the genus were available and the original matrices were provided (see Supplementary File S1 for all original and reconstructed matrices).

Seed germination submodel

The seed bank may be modelled as structured or unstructured. In the unstructured seed bank model, it is assumed that the seed bank does not age and is potentially immortal. In the structured seed bank model, the seed bank progressively ages with potentially different survival, viability and germination rates until it reaches time step *n*, at which point the vital rates remain constant or the seeds die if they do not germinate (see Figure 1B for a life cycle example of the structured seed bank and 1C for the unstructured seed bank). The maximum length of time the seed bank can remain viable is variable across species and is difficult to quantify - some seeds have been documented germinate after centuries of dormancy (Fenner, 2000).

Addition and removal of seed bank stages

Adding, removing, or excluding stages without justification alters the dimensionality of the model and this will affect the demographic parameter estimates. Merging or further dividing up stage classes directly alters vital rates for each class being affected and the speed of the life cycle, where decreasing the dimensionality reduces the number of stages to progress through the life cycle and vice versa (Salguero-Gómez and Plotkin, 2010). As the seed bank functions as a bet hedging mechanism, delayed movement through the life cycle is to be expected and this is reflected in our hypotheses. Since the seed bank is a separate, discrete stage class that involves no splitting or merging, transition rates for the other classes should remain unaffected by its inclusion. However, seedling fecundity would need to be lowered accordingly to account for observed seedlings emerging from the seed bank to avoid overestimating fecundity. We use seedling counts to estimate fecundity in all models where the seed bank was excluded without justification. For simplicity, we assume no clonal reproduction into the seedling stage such that the observed seedlings only consist of two components: the individuals that germinated immediately between year t and $t + 1$ and those that germinated from the dormant seed bank from prior years. The number of observed seedlings can thus be described by the following equation:

$$f_{i,j} = F_j(t)vgs + F_j(t-1)v(1-g)(1-d)g_{b1} + \mathrm{F}_j(t-2)\mathrm{v}(1-g)(1-g_{b1})(1-d)^2 g_{b2} + \ldots, \qquad (1)$$

where the first term on the right hand side represents immediate germinants, the second term represents germination of seeds from the seed bank produced at the previous time step, the third term represents germination of seeds from the seed bank produced two time steps ago, and so on.

Notations:

$f_{i,j}$ = per-capita fecundity into stage class i by stage class j in one time step;

$F_j(t)$ = total fecundity by stage class j at time t;

v = seed viability rate during the step t to $t + 1$;

g = non-seed bank germination rate at the period t to $t + 1$;

g_{b1} = germination rate from seeds that remained dormant in the seed bank for 1 time step;

g_{b2} = germination rate from seeds that remained dormant in the seed bank for 2 time steps;

s = seedling survival rate until the next census;

d = seed bank mortality rate

See Fig. 1 for more information.

To simplify the inclusion of the seed bank stage under limited information, we have made the following assumptions. First, the seedling survival rate (s) from the period t to $t + 1$, is equal to the seed bank survival rate ($1 - d$). Second, we start with simulations that focus on the contribution of the seed bank from $t - 1$. This unstructured submodel captures the predominant seed bank contribution since the contribution of seeds from previous years will diminish exponentially with each year if we assume the vital rates to be not seed age dependent. Alternative age-structured formulations are explored below. Additionally, we assume the fecundity from the previous year to equal the fecundity in the current year, i.e. $F_j(t) = F_j(t\text{-}1)$, hereafter denoted by F_j. Finally, the probability of germinating within the census year is equal to the probability of germinating from the seed bank, i.e. $g = g_b$ (Kalisz and McPeek, 1992). Using the published information on germination, we simplify Eq. 1:

$$f_{i,j} = F_j v g s + F_j v(1\text{-}g)(1\text{-}d)g, \qquad (2)$$

where the first term on the right-hand side represents the number of seedlings germinating immediately between t and $t + 1$, and the second term corresponds to germination from the seed bank. For estimates of the deterministic and stochastic growth rates, we set seed viability (v), seedling survival (s) and seed survival ($1 - d$) equal to 1 unless this information was provided in the literature alongside germination estimates. Monte Carlo simulations (described below) then investigate drawing these values from both uninformed and informed prior distributions. Taking the observed $f_{i,j}$ from the original matrix, we solve Eq. 2 for F_j and calculate the per-capita fertility contributions to seedlings and seed bank (See Supplementary Method S1 for an example). Following the same approach, we also explore alternative model formulations and the effect of different assumptions but only perform Monte Carlo simulations for the unstructured scenario, which is the simplest and most common approach to including the seed bank. The alternatives include the addition of an age-structured seed bank distinguishing seeds from t - 2 and t - 3 and more conservative estimates on the effect of the seed bank by further assuming the seeds do not survive beyond the first, second or third year (See Supplementary Method S1 and Table S1 for the alternative formulations and population growth rates for all models).

Researchers excluding the seed bank from a model make the implicit assumption that it is not present in the population. Thus we intentionally remove the seed bank from selected models to explore the implications of this assumption. Given the large number of potential models to choose from, we have decided to focus on cases in which the seed bank was known to be important (elasticities of λ_1 to seed survival, germination or fecundity > 0.1) to demonstrate the worst case scenario that might be possible when the seed bank is ignored. We have identified six such species from the MPM2007 and MPM2008-2011 databases that fulfil this criteria whilst retaining strong connectedness, the condition that it is possible to transit from each node to all other nodes of the life cycle graph (Caswell, 2001), such as those in Figure 1. We use information on seed viability, seedling survival, and seed survival if provided, but otherwise we set them to 1, and we no longer assume that $g = g_b$ since g_b can be obtained from the original matrix. We estimate the rate of seed germination within the studied period, g, based on partitioning the total reproductive output (Eqs. 1 and 2) to recalculate only the fecundity into the seedling stage. To further ensure that no artefacts from changes to reproduction or mortality are introduced, the removal protocol is the reverse of the addition protocol, and we check that the can be recovered by reapplying the addition protocol (See Supplementary Method S2 for an example).

Analyses

We use R (v 3.1.1; The R Foundation for Statistical Computing) to calculate population growth rates (λ_1), elasticities, stochastic simulations exploring temporal variation in vital rates and Monte Carlo simulations exploring uncertainty in seed bank vital rates (see Supplementary File S2 for code to perform Monte Carlo simulations).

We use uninformed priors for germination, viability, seedling survival and seed survival for Monte Carlo simulations by drawing these rates from a uniform distribution. Although we might expect covariance among vital rates for the below- and above-ground individuals (Morris and Doak, 2002), introducing this structure is complicated particularly when dealing with the seed bank. For example, models that do include the seed bank generally estimate seed bank vital rates through separate germination experiments such that the covariance between the seed bank and the above-ground population is lost in the process. While the impact of this covariance structure on the model outcomes would be of interest, we instead focus here on the exclusion of the seed bank from demographic models and leave investigations regarding correlation structure for future research.

We perform 10,000 simulations, recalculating the fecundities and reconstructing the matrix for each set of parameters, and report the mean growth rate with confidence intervals obtained by the 2.5% and 97.5% quantiles. We then determine the effect of using informed priors via literature estimates for germination, and occasionally viability when available, on the resulting probability density function (p.d.f.) of the distribution of growth rates. Publications on germination often contain multiple treatments and an estimated mean germination rate for each treatment. We have specified our prior distributions depending on the type of demographic year to reflect the function of the seed bank as a bet hedging mechanism (the highest germination rate available when $\lambda_1 > 1$ and the lowest when $\lambda_1 < 1$; Gremer et al., 2012, Kalisz and McPeek, 1992). However, if only one germination rate was provided, we use this value for all population matrix models regardless of the demographic year. If only the mean matrix was provided (arithmetic mean matrix across all years and sites), we use the average germination over all treatments. We model both germination and viability using the beta distribution as it varies between 0 and 1, with the distribution mode equal to the germination rate. We also calculate λ_1 for fixed values of the germination rate (0.01-1 by increments of 0.01) and viability rate (0.1, 0.5, and 1) to capture the total range of possible values when only these parameters are allowed to vary. The matrices in which the germination rate was set to that obtained from the literature are used to calculate single point estimates of λ_1, elasticity matrices, λ_S and extinction probability, with viability, seed survival and seedling survival set to 1 if no information was available. The initial population for all stochastic and extinction probability calculations hereafter consists of 1,000 individuals at the stable stage distribution, calculated as the normalized dominant right eigenvector (Caswell, 2001) of the mean matrix. We have calculated stochastic growth rates via random choice from a set of available matrices (Kaye and Pyke, 2003), where each matrix has an equal probability of being selected at each time step. The initial population is projected over 500,000 time steps with the first 5,000 omitted to remove transient dynamics (Kaye and Pyke, 2003, Morris and Doak, 2002). We use the mean and variance of the growth rates to obtain an estimate for λ_S and 95% confidence intervals.

We also compare extinction probabilities between models with and without seed banks. Times to extinction (defined as population size ≤ 1 individual) are calculated for matrices predicting population decline. We project an initial population for individual matrices over 100 years and compared the difference in times to extinction with and without the seed bank to quantify the effect of the seed bank on population persistence. We then conduct stochastic simulations to calculate the cumulative probability of quasi-extinction (Morris and Doak, 2002) with a quasi-extinction threshold set to 100 individuals (i.e. one tenth of the initial population size). We perform 100 runs, where a single run consists of 5000 iterations each projected over 100 years, and calculate the mean fraction of projections that fell below the quasi-extinction threshold during or before the year t (Morris and Doak, 2002).

Results

Addition of the seed bank and demographic uncertainty

Inclusion of an unstructured seed bank show fairly small differences in λ for most species regardless of the assumptions made (See Fig. 2 and Table 2 to compare the growth rates). However, the potential range of growth rates when demographic parameters are uncertain can be large depending on the species and the amount of temporal variation observed. The 8 tree species, for example, have stable populations ($\lambda_1 \approx 1$) and variation in seed bank parameters has little influence in all of the models with the exception of *Illicium anisatum* (Schisandraceae). Subsequently, using informed priors for these species has little impact on the p.d.f. of the distribution of possible growth rates. In contrast, the shorter-lived shrub species, *Atriplex acanthocarpa* (Amaranthaceae) and *Sambucus racemosa* (Adoxaceae), are more sensitive to the introduction of a seed bank and demographic uncertainty in their vital rates. Here, using an informed prior on germination narrows the p.d.f. of the distribution of possible growth rates for the *A. acanthocarpa* 1996-1997 matrix during an extremely productive year (Fig. 3). Conversely, the use of prior information during the poorer 1997-1999 years appears to increase the spread of possible growth rates. Despite this, poor years tend to have a narrower distribution of possible growth rates and are therefore less sensitive to changes in demographic parameters compared to good years (Figs. 3 and 4). Generally, the p.d.f. of growth rates are skewed, with peaks being concentrated when germination rates are high and greater variation in λ_1 when germination is low (Figs. 3 and 4). However, we note these simulations only investigate the unstructured seed bank scenario. Further uncertainties exist with regards to the longevity and structure of the seed bank, both of which we only briefly explore in this study (Supplementary Table S1).

Elasticity analyses in most plant species, particularly the trees, reveal survival of the adult or late juvenile stages to be the most important demographic process affecting λ_1. The addition of the seed bank did not change this pattern except in the case of *A. acanthocarpa* during 1997-98 and 1998-99 when the populations were declining (Supplementary Table S2). In the 1997-98 period, the elasticity of λ_1 to stasis in the third adult stage is drastically reduced (0.5759 to 0.0647) following the inclusion of the seed bank, whereas the elasticity of λ_1 to stasis in the seed bank is high (0.7919). Similarly, the elasticity of λ_1 to stasis in the first adult stage during the 1998-99 period, is reduced from 0.3983 to 0.12 after seed bank addition, while the elasticity of λ_1 to survival in the seed bank is high (0.5039).

Removal of the seed bank and uncertainty in seed bank presence

Removal of the seed bank further increases population growth when $\lambda_1 > 1$ since germination is no longer delayed, and reduces population persistence when $\lambda_1 < 1$ since the buffering effect is no longer present (Table 3). The notable exception is the wetland herb *Helenium virginicum* (Asteraceae), where removing the seed bank decreases λ_1 (1.1145 to 0.6985) even though the population was increasing. This unusual response is likely a combination of extremely high fecundity into the seed bank and within-the-seed-bank stasis rate (1025 and 0.987 respectively), and the comparatively low fecundity and survival of the smallest stage (0.0075 and 0.199 respectively; see Supplementary File S1 for the original and updated matrices). Decreases in λ_1 is also particularly substantial for *Carduus nutans* (Asteraceae) and *Digitalis purpurea* (Plantaginaceae), but is much smaller or unchanged for the remaining species when the seed bank is removed for poor years.

Extinction probabilities

Inclusion of the seed bank using point estimates for germination generally results in longer times to local extinction (Supplementary Figure S1). Extinction times for *A. acanthocarpa* for the 1997-98 and 1998-99 matrices have relatively high increases in extinction times (> 25 years) when the seed bank is included. In addition, certain populations of *A. canescens* and the perennial herb *D. purpurea* are no longer predicted to go extinct in the next 100 years. The populations of the short-lived herb *C. nutans* have comparatively smaller increases in extinction times. In one population, however, including the seed bank doubles the time to extinction from 8 to 16 years, which is relatively high when compared to the life expectancy of the species ($\eta_E \cong 2$-3 years, based on the fundamental matrix estimation; Caswell, 2001).

The seed bank generally improves the outlook of stochastic projections for the estimation of extinction probability, but patterns vary across species (Supplementary Figure S2). The cumulative probability of extinction at 100 years decreases drastically for *A. acanthocarpa* from 0.19 to 0.01 following inclusion of the seed bank. The extinction probability for *A. canescens* has an interesting pattern where the matrix model with the seed bank results initially in a higher extinction probability than the original in the first 38 years and eventually in a lower extinction probability, but in the remaining two cases, the seed bank has little effect on the extinction probability. Both populations of *D. purpurea* are predicted to go extinct within 100 years, however, there is a modest increase in extinction time (16 to 30 years) for one population and a smaller increase in extinction time for the other (4 to 9 years).

Discussion

The presence of cryptic life stages and their exclusion from demographic models continues to be an impediment to understand and model the population dynamics. These cryptic stages are diverse, ranging from clonality, vegetative dormancy, hibernation, diapause, aestivation, etc., as is their potential impact on population dynamics and model outcomes. Using the seed bank as a well-known example of a cryptic life stage, we demonstrate the use of simulations as a means to reduce uncertainty in seed bank dynamics and to reveal the relative importance of the seed bank and its effect on model outcomes. Ignoring the seed bank in a population can produce variable degrees of uncertainty in the estimates of growth rates and extinction probabilities depending on the observed temporal variation in key vital rates. For example, models with high seedling survival and low temporal variation in demographic parameters are more robust to seed bank uncertainty compared to models with high temporal variation in growth rates and demographic parameters. Furthermore, the potential impact of the seed bank can be influenced by the assumptions underpinning the models and the accuracy of the vital rates that are obtained.

Dealing with the seed bank when its presence is unknown or uncertain

Anticipating the species for which the seed bank will be an important factor is challenging due to the uncertainty in the demographic parameters of the seed bank stage across species, populations and time. Here, we provide a protocol outlining the different approaches one can take when conducting a plant population study (Fig. 5). First, the potential presence of a seed bank can be determined through a literature search of seed germination and dormancy values (Baskin and Baskin, 2014, is of particular note). Second, if a seed bank in the population is confirmed, the best option would be to obtain field data over the study period to provide estimates for the seed bank vital rates. Without this field data, characterising the potential role of the seed bank becomes exceptionally difficult and assumptions regarding fecundity, seed survival, viability and longevity need to be made. Prior knowledge from the literature can narrow down these possibilities to some extent, although seed characteristics and seed age composition can vary substantially between sites. Also, knowledge from the literature or from field and glasshouse trails may not necessarily reflect those of the study site and must therefore be interpreted with caution (Gross and Mackay, 2014). Finally, simulations can be used to complement these literature estimates to efficiently investigate the numerous possible assumptions regarding seed bank vital rates and determine plausible effects of a potential seed bank. Garcia et al. (2010), for example, performed simulations covering feasible information obtained in the field or from other publications for the long-lived orchid *Cypripedium calceolus* (Orchidaceae) to justify excluding the seed bank from their models. Likewise, for many of our species, without the benefit of greater observed temporal variation, ignoring the seed bank would appear to be of minimal consequence for the original study, but it is still crucial that an investigation is performed to provide this justification. Otherwise, effective implementation and assessment of management strategies will be hindered as the demographic model is potentially missing a key component of the life cycle.

Uncertainty in the seed bank and its effect on model outcomes

Demographic uncertainty in the estimation of vital rates appears to be the major driver in influencing the model predictions but varies by species. Long-lived tree species are virtually unaffected by both the inclusion of the seed bank and variation in their demographic parameters. This is not entirely surprising as high adult survival and reduced mortality when environmental conditions are poor can be an alternative strategy to improve the long-term fitness (Clauss and Venable, 2000). In such species, it is likely that the seed bank will be of little consequence, although it would still be prudent to confirm this suspicion (e.g., using simulations) before ignoring the seed stage. Additionally, this does not necessarily mean the seed bank will be unimportant if the population were to experience a large disturbance. In contrast, in populations experiencing greater fluctuations in demographic parameters and population growth rates, the contribution of the seed bank is more visible as is evident for *A. acanthocarpa*, where previously high elasticities for juvenile and adult survival shift towards survival in the seed bank during poor demographic years (Table S2).

Temporal variation and the type of demographic year are also factors that will affect of demographic uncertainty. Alterations in germination and viability rates during poor demographic years ($\lambda_1 < 1$) have relatively little impact on the growth rate compared to favourable years when λ_1 exceeded 1 (Fig. 4). Large reproductive failures during poor years would have resulted in low seed bank and seedling input, reducing the impact of varying seed bank vital rates on λ_1. In contrast, we would expect variations in germination and viability rates during poor years to have a much larger impact on long-term persistence and environmental stochasticity. Nonetheless, with the inclusion of prior information from the literature, the p.d.f. of the distribution of growth rates during these poor demographic years shifts towards one (Fig. 3). We note however that this is assuming dormancy is higher (low germination) during poor demographic years and lower during productive years (Kalisz and McPeek, 1992), but seed dormancy needs not necessarily be correlated with fitness. Interestingly, the use of literature did not necessarily reduce uncertainty in the range of possible growth rates and in fact becomes more variable when dormancy is high (Fig. 3). This is an important result as an uninformed simulation may underestimate the level of uncertainty in demographic output if prior information suggests potentially high levels of dormancy. In contrast, when dormancy is suspected to be low, the impact of the seed bank will not be as great, thus variation in its demographic parameters will be of much less consequence.

In addition to variations in demographic parameters, changes to how the seed bank is included can impact the model structure and outcomes. Our uncertainty simulations investigate a 1-year seed bank as a counterfactual example to illustrate the need for including a seed bank. We acknowledge that while single point estimates of λ_1 under the unstructured scenario showed minor increases compared to the original models, this gradually reduces when additional age classes are introduced (See Table S1 to compare the growth rates). This is likely because a higher proportion of emergent seedlings are now assumed to have germinated from seeds produced in previous years, thus compounding the effect of delayed germination. Similarly, if we assume seeds do not survive beyond the first, second or third year, the effect of the seed bank is almost negligible (See Table S1). However, such estimates are overly conservative and most models that do include the seed bank do not make this assumption.

Another interesting scenario that can be tested is that of no seed bank. Our removal simulations provides insights into the worst-case scenarios when the seed bank is ignored despite being crucial to persistence (as indicated by high elasticities to seed bank transitions). As expected, the implications of ignoring the seed bank in these cases were generally much more evident with decreases in λ_1 of up to 0.4, although these reflect the worst-case scenarios and changes in λ_1 would not generally be so dramatic. The assumption of fecundity being equal to that of the previous year ($F_j(t) = F_j(t)$) was necessary due to lack of information, however, interannual variation is likely to play an important role for many species. When variance in the fecundity is high, the contribution of the seed bank would be overestimated in good demographic years and underestimated in bad demographic years. This would likely have similar consequences on any changes in λ_1, thus we might expect the benefit of having a seed bank during poor years to also be underestimated.

Analysis of extinction probabilities over shorter time periods can provide additional insights not considered under asymptotic analyses. Comparing the difference in extinction times and cumulative extinction probabilities between models with and without the seed bank can provide an empirical measure of the uncertainty in population persistence when the seed bank is ignored, and this measure can vary substantially between years and species. Although increases in extinction times can appear trivial, they must also be considered in context of the life expectancy. For example, the seed bank had little effect on the times to extinction for the populations of *C. nutans*. However, given the short life expectancy of this species, even a small increase in extinction time may prove crucial.

Conclusions

There is still much progress to be made concerning the seed bank and it is but one of many cryptic life stages with the potential to create biases in our model outputs. Adult plants, for example, can undergo vegetative dormancy, whereby individuals do not sprout above ground for one or more growing seasons, resulting in overestimated mortality rates and incorrect assignment of false birth events to re-emerging individuals (Logofet et al., 2019, Shefferson, 2009). Demographic studies on animals can be additionally problematic as certain stages of the life cycle may occupy inaccessible locations or be absent from sampling designs. Katzner et al. (2011) found raptor populations to be drastically underestimated as monitoring programs only captured the breeding portion of the population. In both plant and animal studies, incorrect estimates of population size, persistence, and occupancy in demographic models can result in suboptimal management decisions and make it difficult to assess the success of these management actions (Regan et al., 2011). While we have focused exclusively on plant population models in this study, characterising the extent of exclusion of cryptic life stages in demographic models for animals is just as important and open for investigation with the availability of the COMADRE Animal Matrix Database (Salguero-Gómez *et al.*, 2016).

Acknowledgements

The authors would like to acknowledge the critical review by an anonymous reviewer that has greatly improved the manuscript, E. Crone and J. Ehrlén for helpful comments in previous versions of this manuscript, and S. Ramula and members of the ARC-NZ Research Network for Vegetation Function Plant Population Syndromes working group.

Author Contributions

The first and last author conceived and designed the study. The first author collected the data and performed the analysis. All authors contributed to the study methodology and writing of the manuscript.

References

Abe S, Motai H, Tanaka H, Shibata M, Kominami Y, Nakashizuka T (2008). Population maintenance of the short-lived shrub *Sambucus* in a deciduous forest. Ecology, 89, 1155-1167.

Adams VM, Marsh DM, Knox JS (2005). Importance of the seed bank for population viability and population monitoring in a threatened wetland herb. Biol. Conserv., 124, 425-436.

Angert AL (2006). Demography of central and marginal populations of monkey flowers (*Mimulus cardinalis* and *M. lewisii*). Ecology*,* 87, 2014-2025.

Baskin CC, Baskin GM (2014). Seeds: Ecology, biogeography, and evolution of dormancy and germination, Academic Press, San Diego, California.

Burns JH, Blomberg SP, Crone EE, Ehrlén J, Knight TM, Pichancourt JP, Ramula S, Wardle GM, Buckley YM (2010). Empirical tests of life-history evolution theory using phylogenetic analysis of plant demography. J. Ecol., 98, 334-344.

Caswell H (2001). Matrix population models: construction, analysis and interpretation, Sinauer Associates, Inc., Sunderland, Massachusetts.

Chien PD, Zuidema PA, Nghia NH (2008). Conservation prospects for threatened Vietnamese tree species: results from a demographic study. Popul. Ecol., 50, 227-237.

Clauss MJ, Venable DL (2000). Seed germination in desert annuals: an empirical test of adaptive bet hedging. Am. Nat., 155, 168-186.

COMPADRE Plant Matrix Database. 2019. *COMPADRE Plant Matrix Database v5.0.0.* [Online]. Max Planck Institute for Demographic Research (Germany). [Accessed 20th May 2019].

Crone EE, Menges ES, Ellis MM, Bell T, Bierzychudek P, Ehrlén J, Kaye TN, Knight TM, Lesica P, Morris WF, Oostermeijer G, Quintana-Ascencio PF, Stanley A, Ticktin T, Valverde T, Williams JL (2011). How do plant ecologists use matrix population models? Ecol. Lett., 14, 1-8.

de Kroon H, Plaisier A, J.V. G (1987). Density dependent simulation of the population dynamics of a perennial grassland species, *Hypochaeris radicata*. Oikos, 50, 3-12.

de Kroon H, van Groenendael J, Ehrlen J (2000). Elasticities: a review of methods and model limitations. Ecology (Wash. D. C.), 81, 607-618.

Doak DF, Thomson D, Jules ES (2002). Population viability analysis for plants: understanding the demographic consequences of seed banks for population health. In: Beissinger SR, McCullough DR (eds.) Population Viability Analysis. The University of Chicago Press, Chicago, Illinois, USA.

Fenner M (2000). Seeds: the ecology of regeneration in plant communities, CABI Publishing, Wallingford, UK.

Garcia MB, Goni D, Guzman D (2010). Living at the edge: local versus positional factors in the long-term population dynamics of an endangered orchid. Conserv. Biol., 24, 1219-1229.

Geiser F (2004). Metabolic rate and body temperature reduction during hibernation and daily torpor. Annu. Rev. Physiol., 66, 239-274.

Gremer JR, Crone EE, Lesica P (2012). Are dormant plants hedging their bets? Demographic consequences of prolonged dormancy in variable environments. Am. Nat., 179, 315-327.

Gross C, Mackay D (2014). Two decades of demography reveals that seed and seedling transitions limit population persistence in a translocated shrub. Annals of Botany.

Harper JL (1977). Population biology of plants, Academic Press, London.

Jongejans E, Shea K, Sheppard AW (2006). What controls the population dynamics of the invasive thistle *Carduus nutans* in its native range? J. Appl. Ecol., 43, 877-886.

Kalisz S, McPeek MA (1992). Demography of an age-structured annual: resampled projection matrices, elasticity analyses, and seed bank effects. Ecology, 73, 1082-1093.

Katzner TE, Ivy JAR, Bragin EA, Milner-Gulland E, DeWoody JA (2011). Conservation implications of inaccurate estimation of cryptic population size. Anim. Conserv., 14, 328-332.

Kaye TN, Pyke DA (2003). The effect of stochastic technique on estimates of population viability from transition matrix models. Ecology, 84, 1464-1476.

Kisanuki H, Nakasu M, Nakai A, Yurugi Y (2008). Predicting the population dynamics of three understory broad-leaved evergreen species under the influence of Sika deer in primary and secondary forests of mid-western Japan. J. For. Res., 13, 52-58.

Kouassi KI, Barot S, Gignoux J, Bi IAZ (2008). Demography and life history of two rattan species, *Eremospatha macrocarpa* and *Laccosperma secundiflorum*, in Cote d'Ivoire. J. Trop. Ecol., 24, 493-503.

Logofet DO, Kazantseva ES, Belova IN, Onipchenko VG (2019). Disappointing viability forecat for a local population of *Androsace albana* in a random environment. Zh. Obshch. Biol., 80, 1-13.

Mondragon D (2009). Population viability analysis for *Guarianthe aurantiaca*, an ornamental epiphytic orchid harvested in Southeast Mexico. Plant Species Biol., 24, 35-41.

Morris WF, Doak DF (2002). Quantitative conservation biology: theory and practice of population viability analysis, Sinauer Associates, Sunderland, Massachusetts.

Parker IM (2000). Invasion dynamics of *Cytisus scoparius*: a matrix model approach. Ecol. Appl., 10, 726-743.

Quitete Portela RDC, Bruna EM, Maes dos Santos FA (2010). Demography of palm species in Brazil's Atlantic forest: a comparison of harvested and unharvested species using matrix models. Biodivers. Conserv., 19, 2389-2403.

R Core Team. 2015. *R: A language and environment for statistical computing* [Online]. Vienna, Austria: R Foundation for Statistical Computing. [Accessed 3rd October 2018].

Regan TJ, Chades I, Possingham HP (2011). Optimally managing under imperfect detection: a method for plant invasions. J. Appl. Ecol., 48, 76-85.

Salguero-Gómez R, Jones OR, Archer CR, Buckley YM, Che-Castaldo J, Caswell H, Hodgson D, Scheuerlein A, Conde DA, Brinks E, de Buhr H, Farack C, Gottschalk F, Hartmann A, Henning A, Hoppe G, Römer G, Runge J, Ruoff T, Wille J, Zeh S, Davison R, Vieregg D, Baudisch A, Altwegg R, Colchero F, Dong M, de Kroon H, Lebreton J-D, Metcalf CJE, Neel MM, Parker IM, Takada T, Valverde T, Vélez-Espino LA, Wardle GM, Franco M, Vaupel JW (2015). The COMPADRE Plant Matrix Database: an open online repository for plant demography. J. Ecol., 103, 202-218.

Salguero-Gómez R, Plotkin JB (2010). Matrix dimensions bias demographic inferences: implications for comparative plant demography. Am. Nat., 176, 710-722.

Shefferson RP (2009). The evolutionary ecology of vegetative dormancy in mature herbaceous perennial plants. J. Ecol., 97, 1000-1009.

Sletvold N, Rydgren K (2007). Population dynamics in *Digitalis purpurea*: the interaction of disturbance and seed bank dynamics. J. Ecol., 95, 1346-1359.

Stöcklin J, Fischer M (1999). Plants with longer-lived seeds have lower local extinction rates in grassland remnants 1950-1985. Oecologia, 120, 539-543.

Tanner JE (2001). The influence of clonality on demography: patterns in expected longevity and survivorship. Ecology, 82, 1971-1981.

Venable DL (2007). Bet hedging in a guild of desert annuals. Ecology, 88, 1086-1090.

Verhulst J, Montana C, Mandujano MC, Franco M (2008). Demographic mechanisms in the coexistence of two closely related perennials in a fluctuating environment. Oecologia (Berl.), 156, 95-105.

Vitalis R, Glémin S, Olivieri I (2004). When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. The American Naturalist, 163, 295-311.

Warner RR, Chesson PL (1985). Coexistence mediated by recruitment fluctuations: a field guide to the storage effect. Am. Nat., 125, 769-787.

Tables

Table 1. Comparison of studies from the MPM2008-2011 database with a previous review by (Doak et al., 2002) examining the inclusion of the seed bank in plant population models. Studies were classified as *justified exclusion* if a valid reason was provided for not including the seed bank (e.g. the authors found no evidence of a seed bank) and *unjustified exclusion* if the possibility of a seed bank was not explored or mentioned. Studies from COMPADRE were not analysed to determine whether the exclusion of the seed bank was justified.

Population models and seed banks	**Doak et al. (2002)**	**MPM2008-2011**	**COMPADRE v.5.0.0**
Included	34 (48.6%)	16 (29.1%)	292 (34.0%)
Justified exclusion	6 (8.6%)	13 (23.6%)	568 (66.0%)
Unjustified exclusion	30 (42.9%)	26 (47.3%)	
Total studies	70	55	860

Table 2. Comparison of population growth rates (λ_1) following the addition of a seed bank. *Pop*. refers to populations of the same species in the same study. *Avg*. represents the growth rate from the mean matrix for each population. *Stoch*. represents the mean stochastic growth rate from populations projected over 10,000 times. SB denotes single point estimates from models with germination and viability equal to that obtained from the literature. Growth rates from Monte Carlo simulations were obtained from the mean of 10,000 trials with parameters for germination, viability, seed survival, and seedling survival drawn from uniform distributions when priors are uninformed, while germination and viability were drawn from beta distributions when the priors were informed. 95% confidence intervals for Monte Carlo simulations and stochastic growth rates are displayed in round brackets.

Study	Life Form	Pop	Period	No SB	SB	Monte Carlo sim. uninformed prior	Monte Carlo sim. informed prior
Single time period							
Chien et al. (2008)							
Calocedrus macrolepus	Tree	1	1	0.9683	0.9715	0.9700 (0.9683, 0.9785)	0.9699 (0.9683, 0.9752)
Parashorea chinensis	Tree	1	1	0.9948	0.9955	0.9963 (0.9948, 1.0040)	0.9954 (0.9948, 0.9978)
Pinus kwangtungensis	Tree	1	1	0.9776	0.9786	0.9786 (0.9776, 0.9834)	0.9782 (0.9776, 0.9799)
Kisanuki et al. (2008)							
Illicium anisatum	Tree	1	1	1.1271	1.1647	1.1344 (1.1258, 1.1765)	1.1408 (1.1263, 1.1778)
		2	1	1.1864	1.2276	1.1931 (1.1828, 1.2363)	1.2008 (1.1840, 1.2416)
Abe et al. (2008)							
Sambucus racemosa	Shrub	1	1	1.1696	1.1690	1.1789 (1.1661, 1.2355)	1.1951 (1.1674, 1.2639)
		2	1	1.0007	1.0012	1.0109 (1.0007, 1.0622)	1.0255 (1.0008, 1.0963)
		3	1	2.2227	2.2031	2.1588 (2.0468, 2.2210)	2.1459 (2.0551, 2.2125)
Kouassi et al. (2008)							
Eremospatha macrocarpa	Tree	1	1	0.9792	0.9811	0.9806 (0.9792, 0.9887)	0.9805 (0.9792, 0.9842)

Laccosperma secundiflorum	Tree	1	1	0.9606	0.9661	0.9638 (0.9607, 0.9815)	0.9621 (0.9607, 0.9658)
Multiple time periods							
Mondragon (2009)							
Guarianthe aurantiaca	Herb	1	1	0.9851	0.9881	0.9917 (0.9851, 1.0253)	0.9872 (0.9851, 0.9929)
		1	2	0.9895	0.9908	0.9929 (0.9895, 1.0107)	0.9905 (0.9895, 0.9931)
			Avg.	0.9869	0.9891	-	-
			Stoch.	0.9869 (0.9868, 0.9869)	0.9891 (0.9891, 0.9892)	-	-
Quitete Portela et al. (2010)							
Astrocaryum aculeatissimum	Tree	1	1	1.0093	1.0106	1.0111 (1.0093, 1.0209)	1.0100 (1.0093, 1.0112)
		1	2	0.9964	0.9968	0.9971 (0.9964, 1.0001)	0.9966 (0.9964, 0.9972)
			Avg.	0.9998	1.0085	-	-
			Stoch.	0.9998 (0.9998, 0.9998)	1.0005 (1.0005, 1.0005)	-	-
Geonoma schottiana	Tree	1	1	1.0043	1.0240	1.0114 (1.0050, 1.0442)	1.0213 (1.0050, 1.0712)
		1	2	0.9823	1.0002	1.0115 (1.0050, 1.0453)	1.0216 (1.0050, 1.0721)
			Avg.	0.9949	1.0150	-	-
			Stoch.	1.0050 (1.0050, 1.0050)	1.0146 (1.0145, 1.0148)	-	-
Verhulst et al. (2008)							
Atriplex acanthocarpa	Shrub	1	1	2.4526	2.3455	2.3016 (2.0678, 2.4476)	2.3541 (2.0759, 2.4489)
		1	2	0.7301	0.8305	0.7418 (0.7302, 0.8288)	0.7508 (0.7303, 0.8362)
		1	3	0.7436	0.8862	0.7712 (0.7438, 0.8951)	0.7924 (0.7444, 0.9051)
			Avg.	1.3461	1.3351	-	-
			Stoch.	1.1486 (1.1460, 1.1512)	1.1792 (1.1773, 1.1811)	-	-
Atriplex canescens	Shrub	1	1	1.7211	1.7012	1.6883 (1.6184, 1.7315)	1.6791 (1.6160, 1.7175)
		1	2	0.9451	0.9451	0.9451 (0.9451, 0.9453)	0.9451 (0.9451, 0.9451)
		1	3	0.8805	0.8835	0.8810 (0.8805, 0.8823)	0.8805 (0.8805, 0.8807)
			Avg.	1.1346	1.1648	-	-
			Stoch.	1.0775 (1.0730, 1.0779)	1.1140 (1.1120, 1.1159)	-	-

Table 3. Comparison of population growth rates following removal of the seed bank. Studies are classified on the basis of variability in transition elements that was reported or number of time periods that were provided. Pop. refers to different populations of the same species in the same study. *Avg.* represents the growth rate from the mean matrix. *Stoch.* represents the mean stochastic growth rate from populations projected over 10,000 times. Confidence intervals for stochastic simulations when the seed bank was removed are shown in brackets.

Study	**Life Form**	**Pop.**	**Period**	**SB**	**No SB**
No variability reported					
Adams et al. (2005)					
Helenium virginicum	Herb	1	1	1.1145	0.6985
de Kroon et al. (1987)					
Hypochaeris radicata	Herb	1	1	0.2973	0.2154
Variability, single time period					
Jongejans et al. (2006)					
Carduus nutans	Herb	1	1	0.4680	0.3894
		2	1	0.6270	0.3771
		3	1	0.6020	0.5524
Multiple time periods					
Angert (2006)					
Mimulus lewisii	Herb	1	1	0.7778	0.7778
		1	2	0.6863	0.6815
		1	3	0.9959	0.9943
		1	Avg.	0.8541	0.8332
		1	Stoch.	0.8423	0.8181
				(0.8410, 0.8435)	(0.8222, 0.8233)
		2	1	1.0154	1.0053
		2	2	0.6978	0.6917

		2	3	1.3203	1.3746
		2	Avg.	1.0158	1.0410
		2	Stoch.	0.9642	0.9949
				(0.9769, 0.9787	(0.9935, 0.9963)
		3	1	1.1052	1.1094
		3	2	1.1197	1.1450
		3	3	0.9523	0.9507
		3	Avg.	1.0510	1.0583
		3	Stoch.	1.0472	1.0537
				(1.0470, 1.0500)	(1.0544, 1.0557)
		4	1	1.5841	1.6704
		4	2	1.0459	1.0470
		4	3	1.1227	1.3083
		4	Avg.	1.3112	1.3448
		4	Stoch.	1.2919	1.3198
				(1.2908, 1.2930)	(1.3192, 1.3205)
Parker (2000)					
Cytisus scoparius	Herb	1	1	2.1963	2.3331
		1	2	2.2629	2.3858
		1	Avg.	2.5019	2.7128
		1	Stoch.	2.4919	2.6927
				(2.4879, 2.4960)	(2.6868, 2.6986)
		2	1	1.7097	1.7407
		2	2	1.2542	1.2406
		2	Avg.	1.5681	1.5552
		2	Stoch.	1.5507	1.5351
				(1.5495, 1.5518)	(1.5343, 1.5359)
		3	1	1.0040	1.0040
		3	2	0.9295	0.9287
		3	Avg.	0.9748	0.9746
		3	Stoch.	0.9740	0.9738
				(0.9734, 0.9746)	(0.9700, 0.9776)

Sletvold and Rydgren (2007)					
Digitalis purpurea	Herb	1	1	0.7533	0.3174
		1	2	0.7831	0.5536
		1	Avg.	0.7715	0.5071
		1	Stoch.	0.7713	0.5008
				(0.7712, 0.7713)	(0.5002, 0.5013)
		2	1	0.6831	0.2983
		2	2	1.0198	0.9379
		2	Avg.	0.8565	0.6701
		2	Stoch.	0.8388	0.5794
				(0.8384, 0.8392)	(0.5787, 0.5802)

Figures

Fig. 1. Life cycle graphs corresponding to: **A)** a plant life cycle consisting of two stages: the seedlings, Se, and the adults, Ad; **B)** a plant life cycle consisting of seeds that have remained in the seed bank for 1 year, S1, up to *n* years, Sn (Eq. 1), and **C)** a plant life cycle with an unstructured seed bank, SB, (Eq. 2). The transition rate $f_{\mathrm{Se,Ad}}$ gives the fecundity into the seedling stage while $f_{\mathrm{S1,Ad}}$ and $f_{\mathrm{SB,Ad}}$ gives the fecundity into the seed bank. **D)** Details of the derivation of Eq. 1 in which the parameters are contained in boxes (see formal notation in the text). Tracking the vital rates required to reach the seedling stage from the adult stage either directly or via the seed bank demonstrates how each term in Eq. 1 was derived.

Fig. 2. Scatterplot comparing population growth rates between models when the seed bank is included (y-axis) and when the seed bank is not included (x-axis). “Addition” refers to models in which we added the seed bank the model (Table 2), while “Removal” refers to models where we removed the seed bank (Table 3). Growth rates for models where the seed bank was added were taken from Monte Carlo simulations with germination prior guided by the literature. Points above the one-to-one line indicate the seed bank has a positive effect on λ while points below indicate a negative effect. Points within the red section are declining populations ($\lambda_1 < 1$) while those outside are increasing populations ($\lambda_1 > 1$).

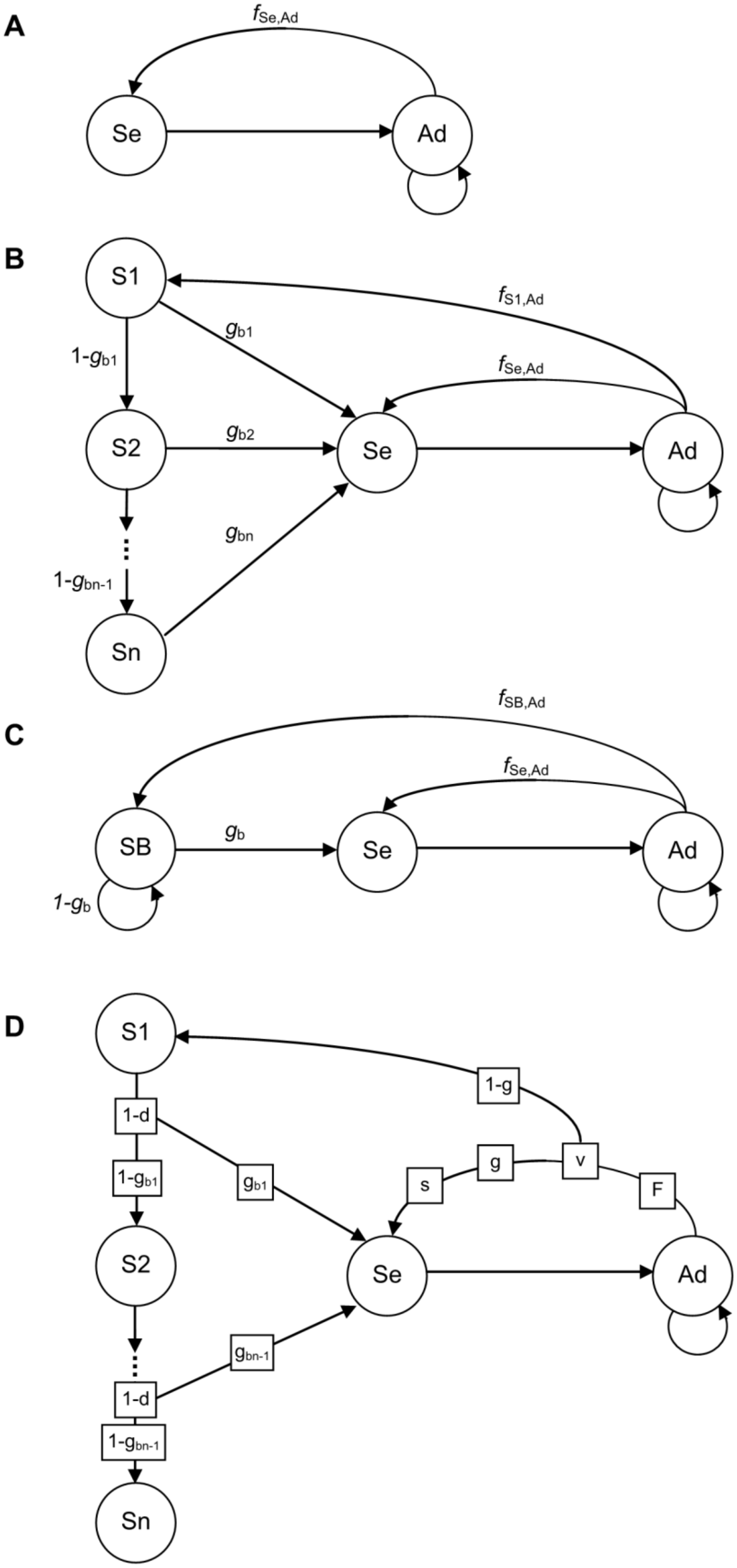
Fig. 3. Probability density function of the distribution of population growth rates for *Atriplex acanthocarpa* from **A)** 1996-1997, **B)** 1997-1998, and **C)** 1998-1999 based on 10,000 Monte Carlo simulations calculated using an uninformed prior on germination (uniform distribution) and an informed prior on germination (beta distribution with the mode equal to the germination

rate obtained from the literature). Red dashed lines indicate the mean growth rate using uninformed and informed priors as reported in Table 2.

Fig. 4. Population growth rate (λ_1) as a function of germination rate from the seed bank for **A)** *Atriplex acanthocarpa*, **B)** *Atriplex canescens*, **C)** *Illicium anisatum*, and **D)** *Geonoma schottiana*. Population growth rates were calculated at seed viability rates of 1 (green line), 0.5 (blue line), and 0.1 (red line) for the germination rates between 0 and 1 at intervals of 0.01. The germination rate for *G. schottiana* is restricted between 0 and 0.69 to not exceed the total seed survival estimate obtained from the literature.

Fig. 5. Flow chart describing different approaches to dealing with a potential seed bank when modelling plant population dynamics. When the presence of the seed bank is uncertain, it is advised that researchers explore this possibility either through literature confirmation, or investigating the soil reserves for viable seeds. By not investigating, the exclusion is unjustified, resulting in the highest possible uncertainty. Justified exclusion, by either proving it does not exist or showing it is unimportant via simulations, gives the highest level of certainty in the model outcomes. Including the seed bank, either by using the literature or data obtained in the field would unavoidably result in some amount of uncertainty from estimating the vital rates, but is preferable to unjustified exclusion.

Fig. 1



A
$f_{Se,Ad}$
Se
Ad
B
S1
$f_{S1,Ad}$
g_{b1}
1-g_{b1}
$f_{Se,Ad}$
S2
g_{b2}
Se
Ad
g_{bn}
1-g_{bn-1}
Sn
C
$f_{SB,Ad}$
$f_{Se,Ad}$
SB
g_b
Se
Ad
1-g_b
D
S1
1-g
1-d
g
v
s
1-g_{b1}
g_{b1}
F
S2
Se
Ad
g_{bn-1}
1-d
1-g_{bn-1}
Sn

Fig. 2

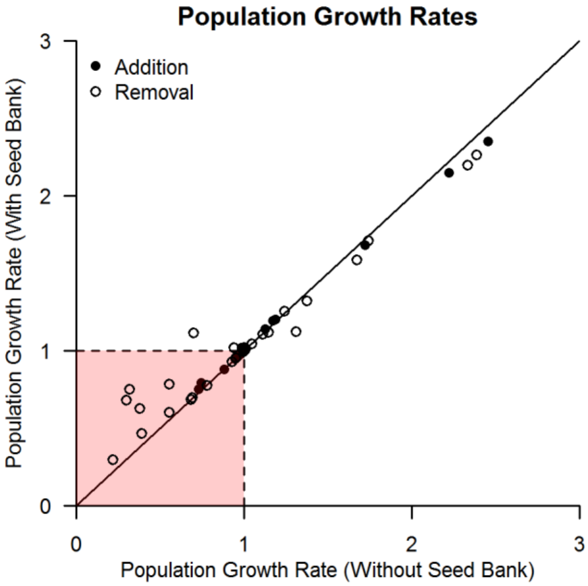

Population Growth Rates
Addition
Removal
Population Growth Rate (With Seed Bank)
Population Growth Rate (Without Seed Bank)
0
1
2
3
0
1
2
3

Fig. 3

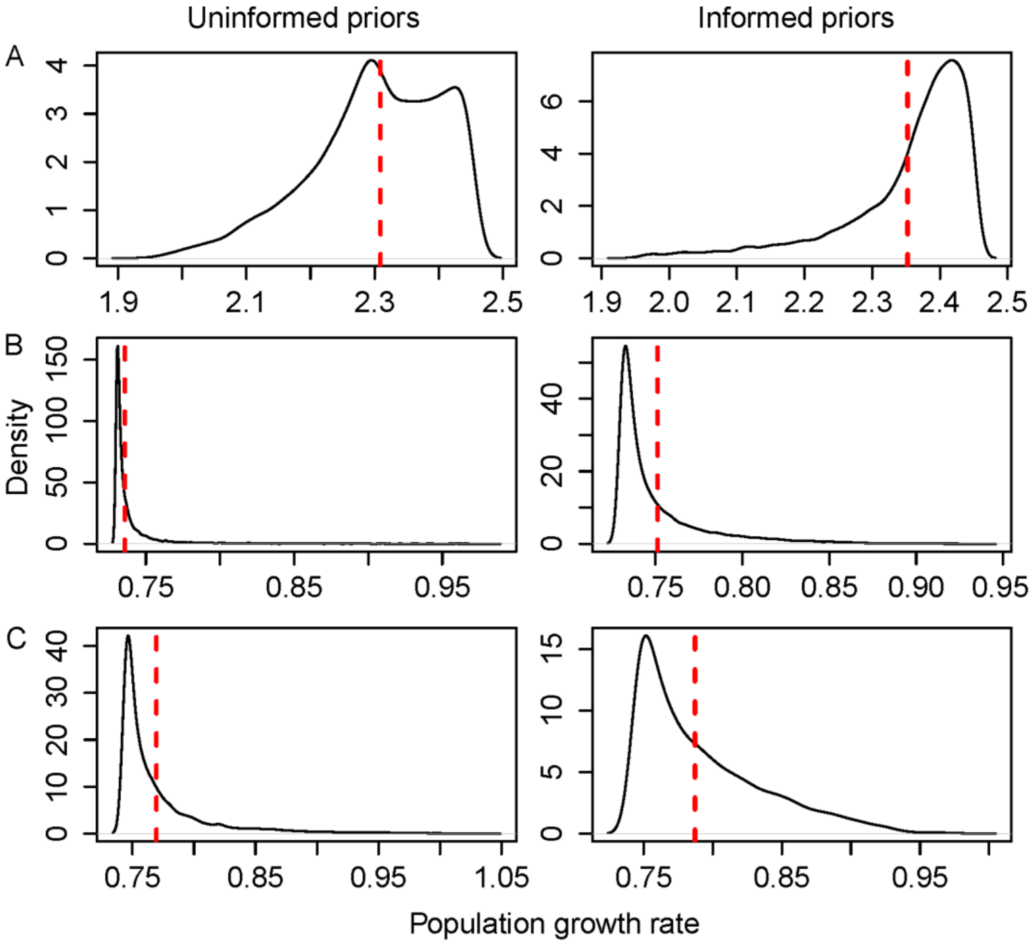

Uninformed priors
Informed priors
A
B
C
Density
Population growth rate
0
1
2
3
4
0
2
4
6
1.9
2.1
2.3
2.5
1.9
2.0
2.1
2.2
2.3
2.4
2.5
0
50
100
150
0
20
40
0.75
0.85
0.95
0.75
0.80
0.85
0.90
0.95
0
10
20
30
40
0
5
10
15
0.75
0.85
0.95
1.05
0.75
0.85
0.95

Fig. 4

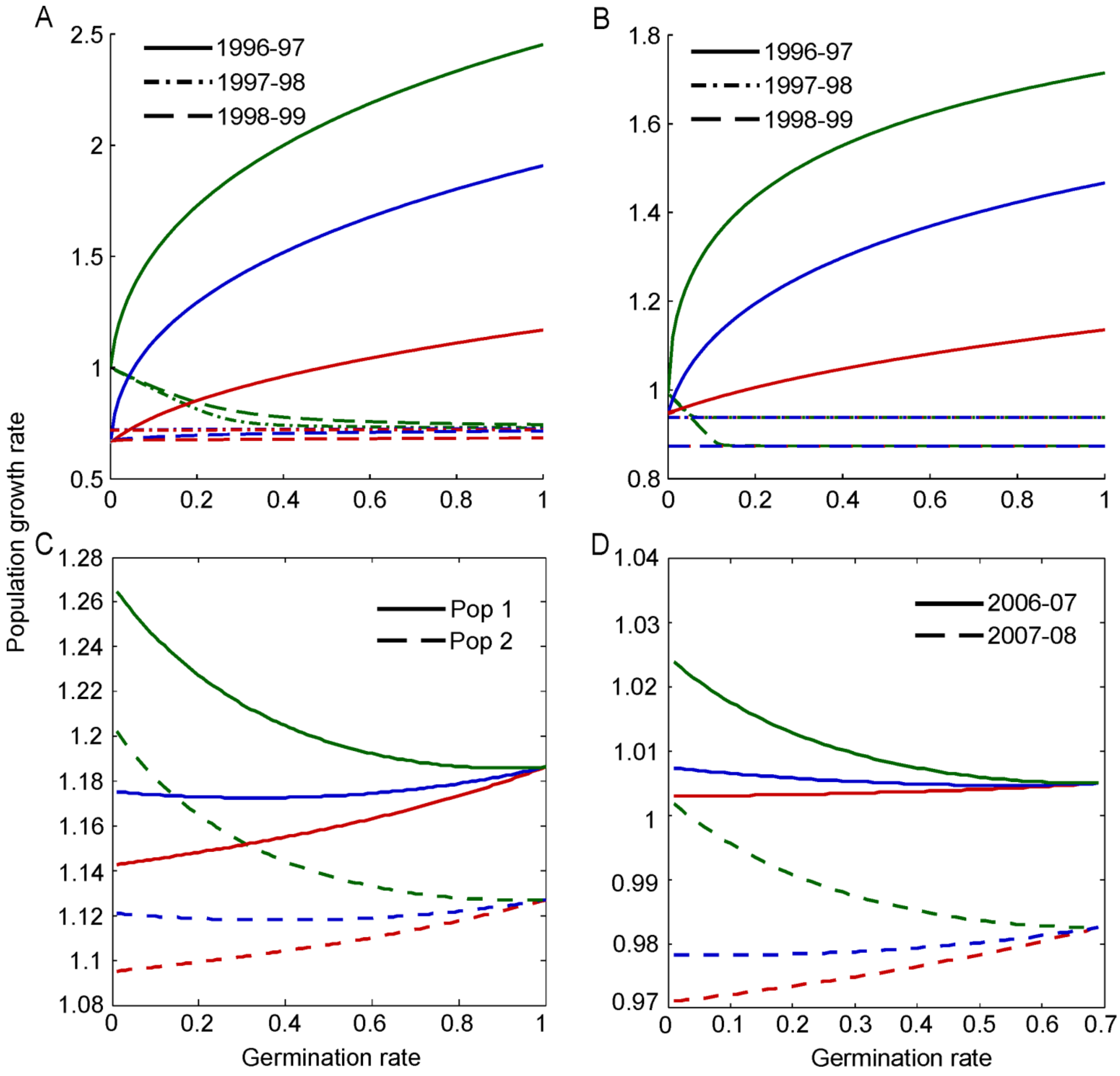
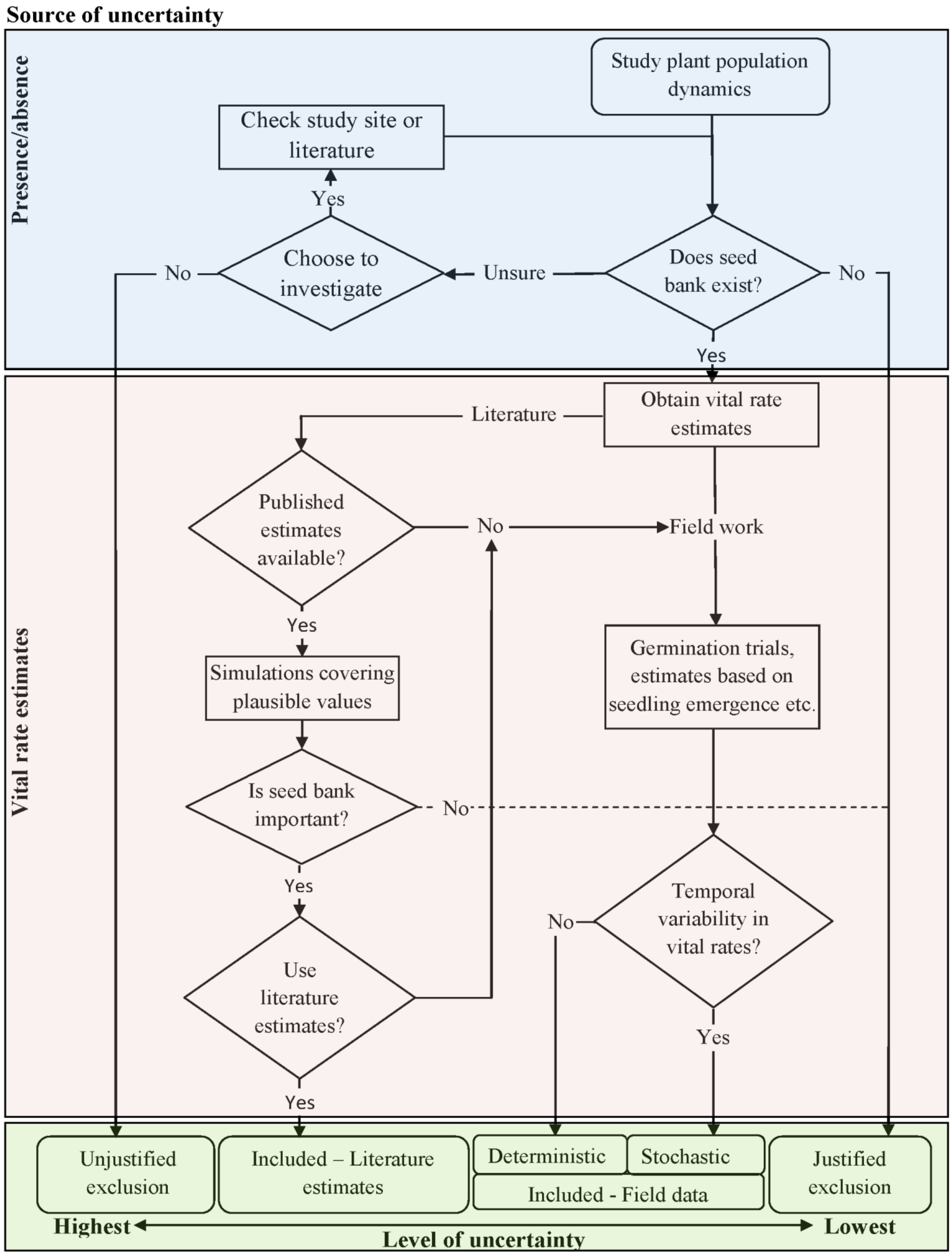

A
1996-97
1997-98
1998-99
B
1996-97
1997-98
1998-99
C
Pop 1
Pop 2
D
2006-07
2007-08
Population growth rate
Germination rate
Germination rate
2.5
2
1.5
1
0.5
1.8
1.6
1.4
1.2
1
0.8
1.28
1.26
1.24
1.22
1.2
1.18
1.16
1.14
1.12
1.1
1.08
1.04
1.03
1.02
1.01
1
0.99
0.98
0.97
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7

Fig. 5


Source of uncertainty
Presence/absence
Study plant population dynamics
Check study site or literature
Yes
Choose to investigate
No
Unsure
Does seed bank exist?
No
Yes
Vital rate estimates
Obtain vital rate estimates
Literature
Published estimates available?
No
Field work
Yes
Simulations covering plausible values
Germination trials, estimates based on seedling emergence etc.
Is seed bank important?
No
Yes
Temporal variability in vital rates?
No
Use literature estimates?
Yes
Yes
Unjustified exclusion
Included – Literature estimates
Deterministic
Stochastic
Included - Field data
Justified exclusion
Highest
Lowest
Level of uncertainty

Supporting Information

Table S1

Population growth rates from seed bank models, Monte Carlo simulations and alternative formulations with 2 and 3 year longevity

Table S2

Elasticity matrices for 1997 and 1998 for *Atriplex acanthocarpa* with and without the seed bank

Figure S1

Log population size projections for single point matrices with and without the seed bank

Figure S2

Cumulative extinction probabilities of quasi-extinction

Method S1

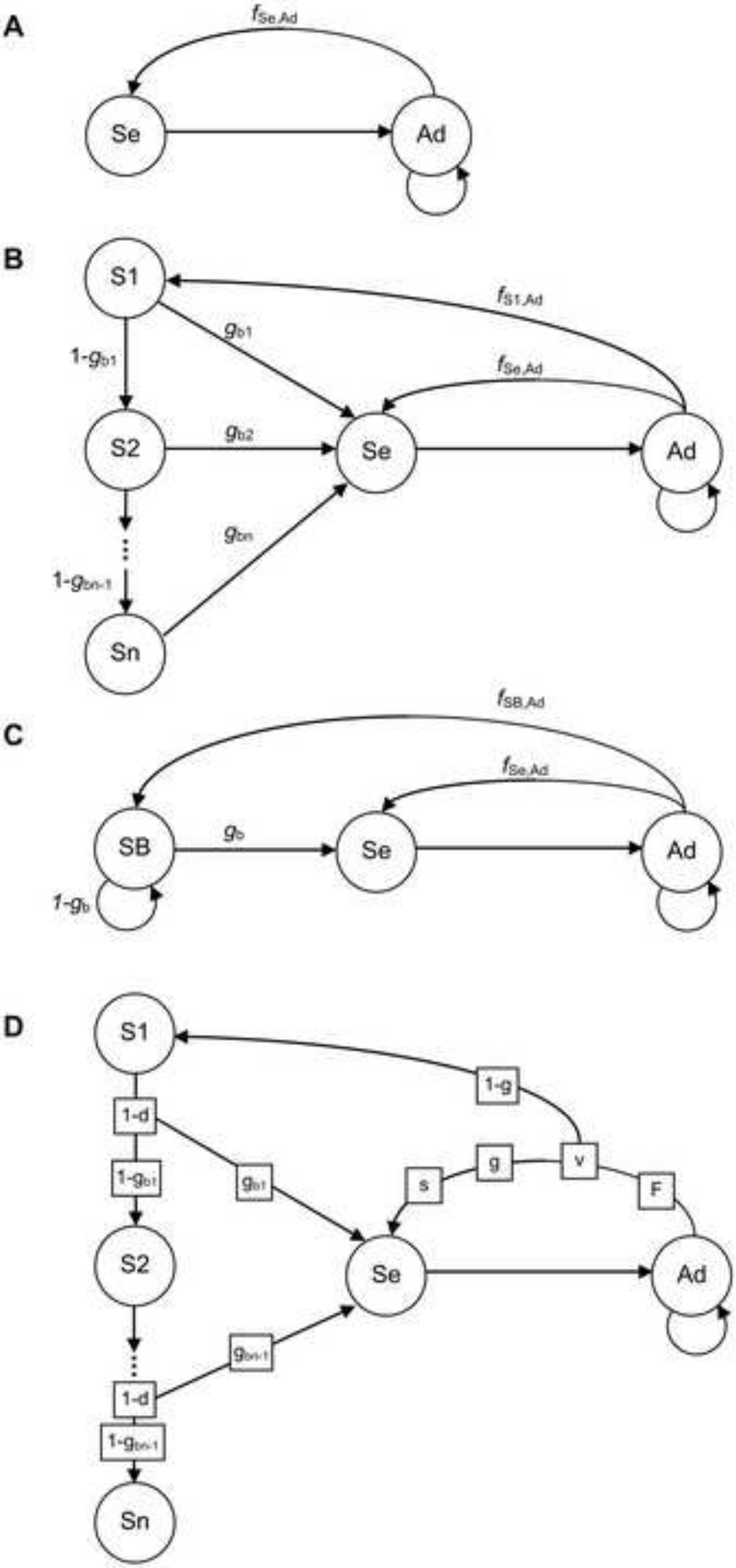
Example demonstrating methodology for including the seed bank stage and alternative formulations

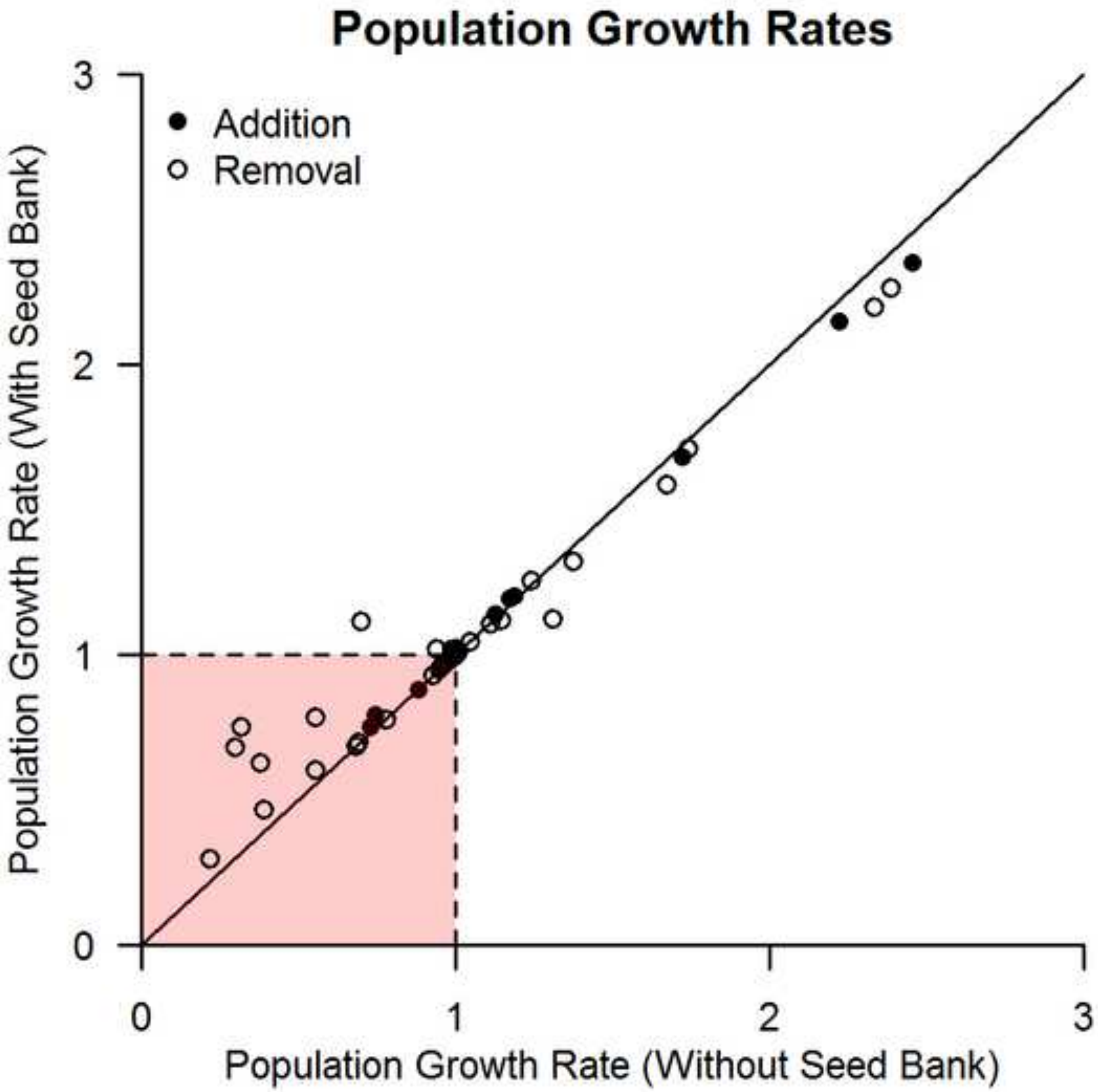
Method S2

Example of seed bank stage removal from a matrix model for *Digitalis purpurea*

File S1

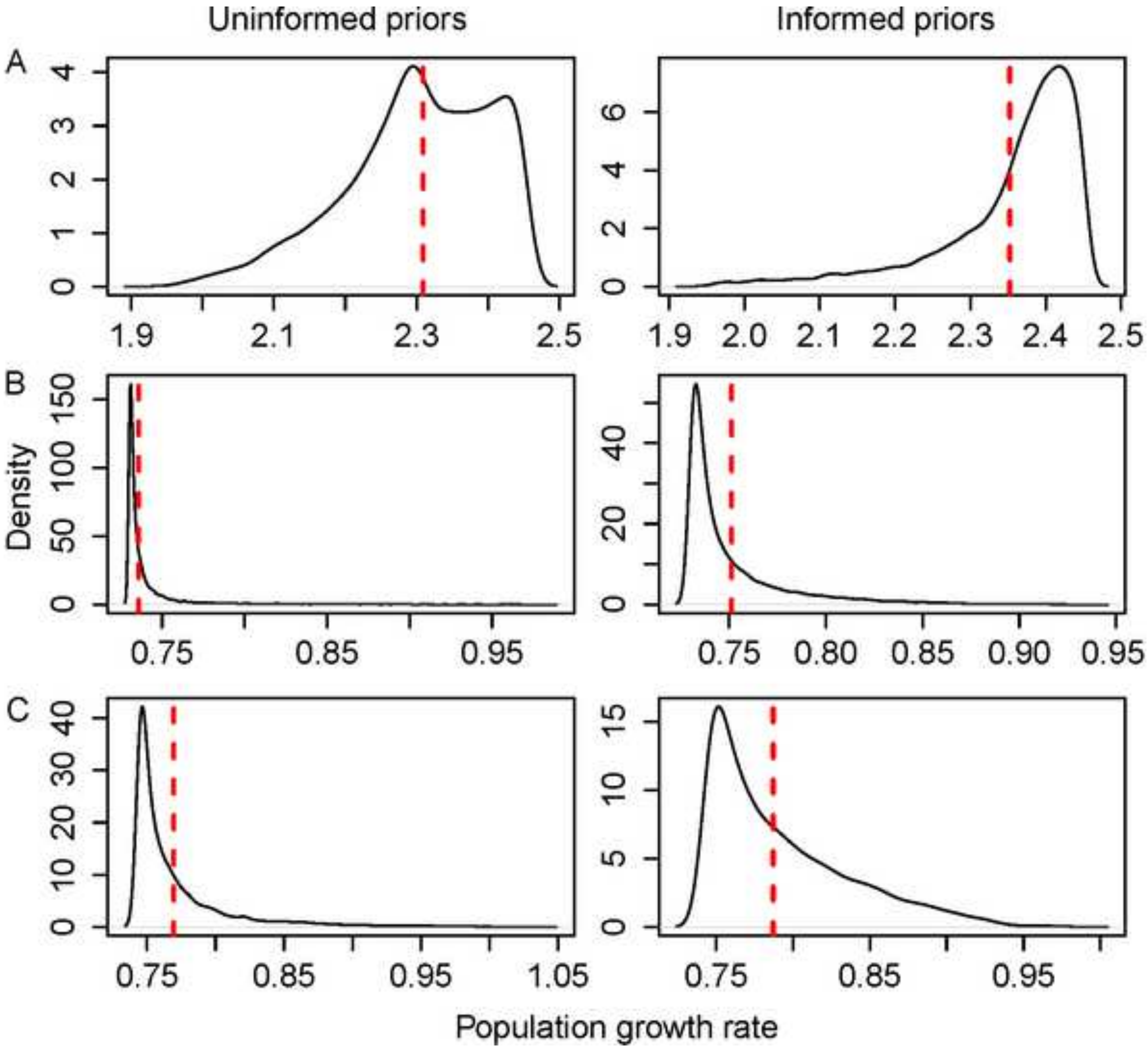
Excel file containing original and updated matrices used for addition and removal simulations


A
$f_{Se,Ad}$
Se
Ad
B
S1
$f_{S1,Ad}$
g_{b1}
$1-g_{b1}$
$f_{Se,Ad}$
S2
g_{b2}
Se
Ad
g_{bn}
$1-g_{bn-1}$
Sn
C
$f_{SB,Ad}$
$f_{Se,Ad}$
SB
g_b
Se
Ad
$1-g_b$
D
S1
1-g
1-d
$1-g_{b1}$
g_{b1}
v
g
s
F
S2
Se
Ad
g_{bn-1}
1-d
$1-g_{bn-1}$
Sn


Population Growth Rates
● Addition
○ Removal
Population Growth Rate (With Seed Bank)
Population Growth Rate (Without Seed Bank)
0
1
2
3
0
1
2
3

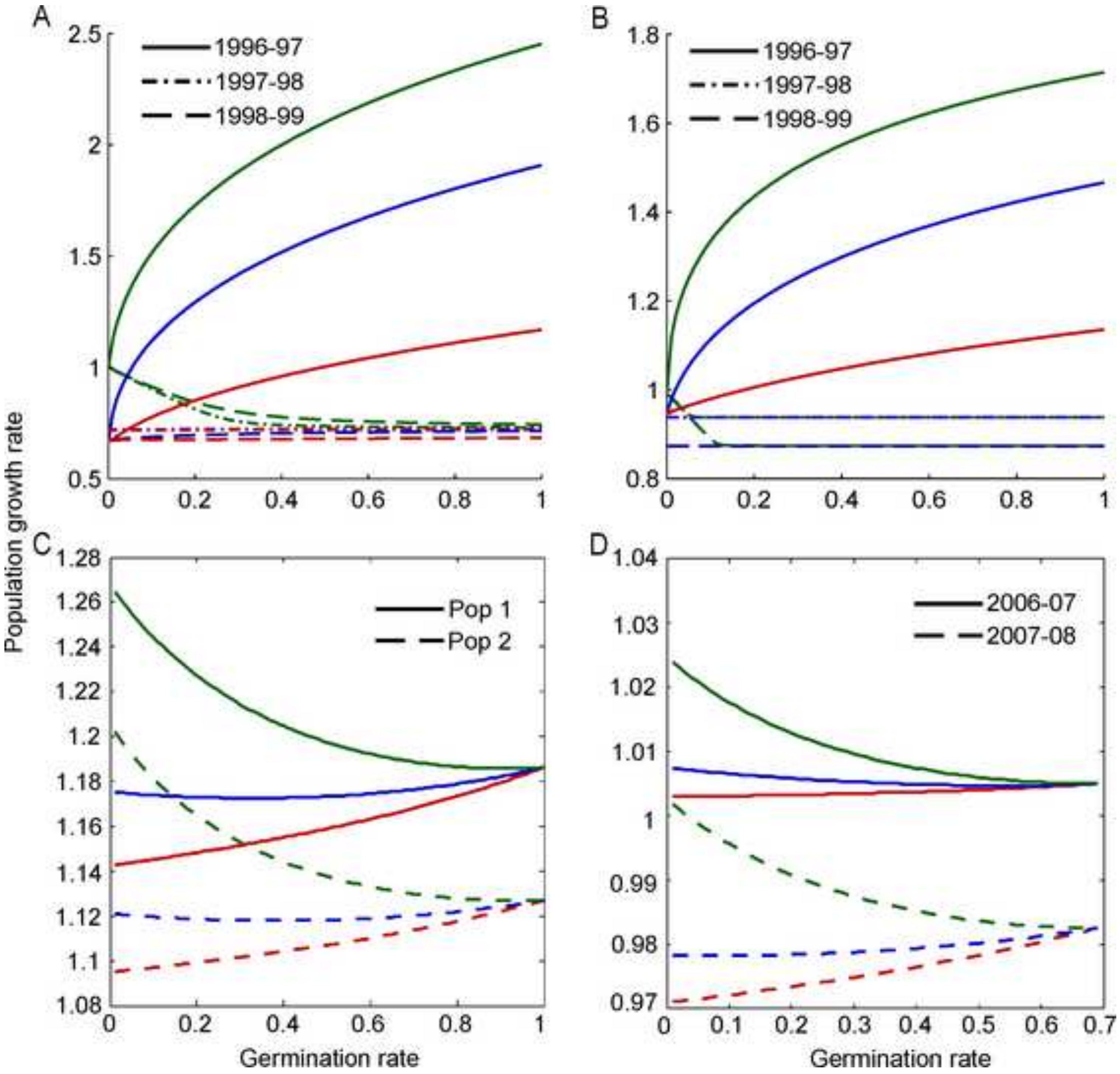
Figure

Click here to download high resolution image



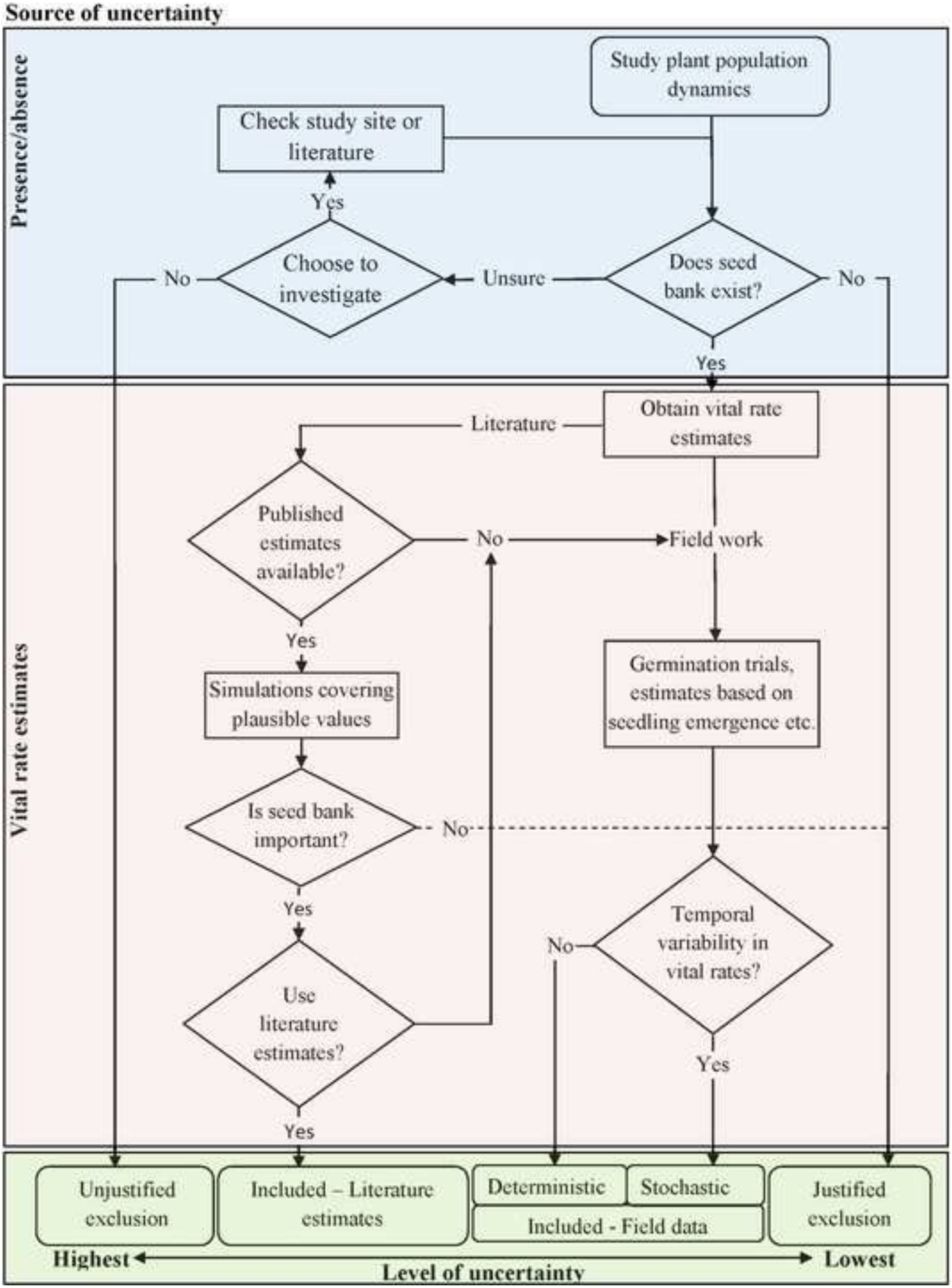
Figure

Click here to download high resolution image



Figure

Click here to download high resolution image



Supplementary material for online publication only

Credit Author Statement

VN contributed to the conceptualization, methodology, formal analysis, investigation and writing of the original draft, review and editing.

YB contributed to the conceptualisation, methodology and writing of the original draft, review and editing.

RSG contributed to the methodology and writing of the original draft, review and editing.

GW contributed to the conceptualisation, methodology, resources and writing of the original draft, review and editing.