Biological Conservation 306 (2025) 111141

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Sampling design and estimates of observation error greatly reduce quasi-extinction probability in plant populations

Héctor Miranda-Cebrián [a,*], Daniel F. Doak [b], María Begoña García [a]

[a] Instituto Pirenaico de Ecología (IPE - CSIC), Avenida de Montañana 1005, 50059 Zaragoza, Spain
[b] University of Colorado, Environmental Studies, 397 UCB, Boulder, CO 80309, USA

ARTICLE INFO

Keywords:
PVA
Population trends
Sampling methods
Demography
Citizen science
Monitoring

ABSTRACT

Estimates of population dynamics and risk of extinction are sensitive to both mean rates of annual change and also the variation in these rates caused by environmental stochasticity. The analytical machinery to incorporate the latter into estimates of long-term stochastic growth and quasi-extinction risk are well developed for count-based population data. However, analytical methods rarely account for the effects of observation error during the sampling process, which can inflate apparent stochasticity and thus alter estimates of population behavior. Here, we applied a Bayesian stochastic population model to estimate the growth rates and quasi-extinction risk of over 157 plant populations monitored through a collaborative science program in NE Spain, and calculated the effect of incorporating direct measures of the observation error into our estimates. We found that including the observation error into models reduced the estimated temporal variation of all populations, which in turn resulted in modest increases in estimated long-term growth rates but considerable reductions in quasi-extinction risk. In this study we show how adjusting sampling designs to the size, detectability and density of plant populations, and repeating surveys in one or more years substantially improves estimates of population growth and viability, thus contributing to guide a better conservation practice.

1. Introduction

All organisms show fluctuations in their population sizes and growth rates over time, caused by changes in abiotic conditions, biological interactions, and density effects. These drivers alter vital rates such as fecundity and survival, which in turn alter population growth. This stochastic variation in realized population growth rates is generally predicted to have a negative effect on long-term stochastic growth rate (Dennis et al., 1991; Gillespie, 1977; Lande, 1993; Tuljapurkar, 1990). Furthermore, this variability has a direct negative influence on the extinction risk of a population, as higher variance in growth rates increases the chance of a population having several "bad" years in a row and thus its probability of hitting quasi-extinction thresholds (Lande, 1993; Lande and Orzack, 1988; Dennis et al., 1991). This is particularly concerning for small populations because their size may not be sufficient to buffer the effects of prolonged negative growth (Shaffer, 1981). Hence, accurately estimating the variability of growth rates through time, in addition to mean growth, is a fundamental requirement for an effective management of biodiversity in general, and for population viability analysis of threatened species in particular (Boyce et al., 2006).

The most common data sets available to estimate extinction risk or stochastic dynamics are based on time series of population counts (Global Population Dynamics Database, Prendergast et al., 2010). For either density independent or density dependent dynamics, the use of these types of population size estimates to calculate mean and variance in growth rates and to predict population behavior is well-established (Dennis et al., 1991; Morris and Doak, 2002). However, part of the observed variation in numbers over time results from the observation process itself, rather than underlying changes in real numbers or densities themselves (Wilson et al., 2011). Neither the sampling process nor the observers can be perfect, and they are susceptible to biases that give us an imperfect picture of the actual population size by, for example, only sampling part of the population, missing individuals during the counting process, or counting the same individual twice. Several frameworks for taking observation errors into account when estimating demographic parameters from count based data have been proposed (Buonaccorsi and Staudenmayer, 2009; Carpenter et al., 1994; Dail and Madsen, 2011; Holmes, 2001), However, although measuring and considering observation error is common for viability analysis of animal populations, it is seldom measured during monitoring studies or taken

---

* Corresponding author.
E-mail address: hmiranda@ipe.csic.es (H. Miranda-Cebrián).

https://doi.org/10.1016/j.biocon.2025.111141
Received 7 May 2024; Received in revised form 24 March 2025; Accepted 1 April 2025

into account in plant PVA studies (Kellner and Swihart, 2014; McLoughlin and Messier, 2004; but see Kéry, 2004). When not accounted for, the observation error is folded into estimated environmental stochasticity in population growth (Holmes, 2001), leading to biased and inaccurate estimates of population growth and extinction risk.

Estimating the variation associated with observation error should allow us to decompose the observed variance in growth rates into true (process) variance, and variance due to observation errors, reducing the estimated stochastic variance of a population's size and improving our estimates of long-term growth rates and their variance (Dennis et al., 2010; Staples et al., 2004). This has direct implications for conservation practice. First, by reducing the estimated variance in population growth rates over time we will generally increase the estimated long-term growth rate of that population, and reduce the estimated probability of extinction or decline (See and Holmes, 2015). In turn, these improved assessments of population viability will allow us to develop better conservation plans and policies to help us manage the current biodiversity crisis (Cowie et al., 2022), especially in a context of limited resources for conservation actions that may require focusing our effort on some organisms over others (Ono et al., 2019).

Here, we use the abundances of 157 plant populations obtained through the long-term "Adopt a Plant" collaborative citizen science program in Aragon Province, NE Spain (García et al., 2021) to explore how accounting for the observation error (OE) measured directly during the sampling process can affect our estimates of population growth and long-term quasi-extinction risk. Similar to other plant monitoring programs like the Threatened Plants Project (Walker et al., 2017), the Adopt a Plant program monitors dozens of species and populations across a wide range of environments. The most distinctive aspects of this project are its long duration (volunteers commit to a 10-year period and some populations have been monitored uninterrupted since 2010) and the use of repeated surveys in some years to estimate the observation error in the monitoring for most species. As we show, the combination of simple resampling done as part of field surveys and a straightforward Bayesian modeling approach can substantially improve the parameter estimates resulting from monitoring, with significant effects on estimated population dynamics.

To illustrate this approach, we first develop a stochastic exponential population model which estimates the long-term growth rate of a population and its stochastic variance, while taking into account the variance associated with observation error (OE) to calculate the probabilities of quasi-extinction. Then, we use those estimates for each population and compare them with those from a model that does not consider OE. Finally, we compare the growth rates and quasi-extinction risk probabilities calculated across different groups of species, including those surveyed using different fieldwork abundance metrics.

Thus, here we strive to test if accounting for observation errors estimated from repeating surveys substantially alters stochastic growth or quasi-extinction risk across different abundance metrics. We also use simulated data sets to compare our approach to other methods of estimating population abundance change in the face of sampling errors.

2. Methods

2.1. Plant population data

The "Adopt a plant" program is a collaborative citizen science initiative that aims to monitor the population trends and threats of hundreds of plant populations in NE Spain (García et al., 2021). Many of the species are listed as threatened and others are considered major structural elements for habitats of interest by the European Union. The sampling personnel involves almost 400 people, with the majority being volunteers and forest and park rangers, although a few are naturalists with higher botanical skills. All participants are assigned one or more monitoring units (MU), which they will visit once a year for a period of at least 10 years. MUs are always surveyed by the same team to avoid differences between observers and each of them includes populations of one or more plant species that are monitored each year at a similar phenological stage to ensure consistency across years and keep plant detectability as constant as possible. In addition, all plant species monitored by this program have multi-annual life cycles, which helps in relocating individuals from one year to another (Elzinga et al., 2015). Participants are trained on site by the managing scientists during the first sampling year to ensure that they conduct the surveying process in a replicable and consistent manner following a specific sampling design established by the managing scientists, and that they are able to consistently identify their focal plant (see García et al., 2021 for more details). Plant abundance in each population is estimated using plots or along transects, and in some cases using small quadrats and gridded plots along transects. These plots and transects are fixed during the first year, and are revisited each subsequent sampling season (Appendix C). Plots are placed in order to sample across the different environmental conditions and plant densities within the population, avoiding extensive sampling of areas with extremely high or low density. All the data gathered by the participants is quality checked by the managing team before being included in the database. The median number of sampling years is 9, with a minimum of 4 and a maximum of 14, while populations per species ranges from one to six.

The abundance metric used varies between plant species and sites, as it is adjusted in the field to be representative of the variability in species occupancy and to reduce sampling error (García et al., 2021). These approaches can be categorized into four general metrics. The first two approaches involve counts of either all (hereafter referred to as counts), or only the reproductive (counts_R) individuals within permanent sampling units (> 70 % of MUs). The third and fourth approaches score either the presence/absence (occupancy) or percent cover over multiple quadrats (plant cover). To facilitate the comparison between estimates of abundance between populations using different metrics, the total number of presences and raw plant cover data were used in subsequent analyses rather than their relative proportion over the total number of possible presences or surveyed area.

2.2. Estimation of observation error

All participants were asked to do at least one double survey during one of the monitoring years. These were always conducted by the same team responsible for that MU and consisted of dismantling and relocating all the sampling plots and repeating the survey on the same day or a few days after. Resurveys occur in all or part of the sampling units for a MU, and are required to take the same amount of time.

The mean of the squared difference between the natural logarithms of the two abundance estimates taken across years and populations of a species was used as an estimate of the observation error for that species. Use of the log of abundance to make an estimate of variance is based on an assumption of constant proportional errors in abundance, relative to the mean, such that variance on the log scale will be scale invariant, which was confirmed by examination of abundance *vs* squared differences of logged abundances (Appendix A, Fig. A1). If a species lacked a double survey in a particular population but had an estimate of observation error in others (19.6 % of the cases), we used the double surveys from those populations with estimates as an approximation for the error. A total of 157 populations of 93 different plant species were double surveyed, and constitute the dataset for this analysis.

2.3. Stochastic exponential model

To estimate the mean and variance of the annual stochastic population growth rate of each plant population while accounting for OE, we fit a density-independent Bayesian stochastic exponential growth model using JAGS 4.3.0 (Plummer, 2003), *rjags* (Plummer, 2022) and *runjags* (Denwood, 2016). The model is built assuming that the logarithm of observed population size at time T, $\log(M_T)$, has a normal distribution with mean and variance equal to $\log(N_T)$ and ε^2_{OE}, respectively, and where $\log(N_T)$ is the logarithm of true population size at time T. In addition, we assume stochastic exponential growth, with $N_T = \lambda_T{*}N_{T-1}$, and where $log(\lambda_T)$ is normally distributed with mean $\overline{log(\lambda)}$ and variance, $\sigma^2_{Process}$. We simultaneously fit this population growth process and estimated the variance from OE, ε^2_{OE}. Thus, the model uses both the data on observed numbers across time for a population as well as pairs of observations for the years and population plots with resurveys. Use of this model is based on the assumption that sampling is sufficient in each survey so that log growth rates are normally distributed due to the central limit theorem. This model assumes that proportional observation error is constant through time. This is clearly a simplifying assumption, and is likely to be particularly poor for mobile organisms, annual plants, or monitoring studies with non-permanent plot designs. However, all plants in the Adopt a Plant program are long-lived or multi-annual species, and thus easily detectable when using a permanent plot sampling design. In addition, the training efforts made to ensure that volunteers are able to identify their assigned plant species, to keep effort similar across surveys, and to always visit the MUs during the same time of the year should contribute to keep detectability high and relatively constant through time.

The ε^2_{OE} and the population sizes from one year to the next advance according to a model defined by the following equations:

$$log(N_T) = log(\lambda_{T-1}) + log(N_{T-1}) \tag{1}$$

$$log(M_T) + \varepsilon_T = log(\lambda_{T-1}) + log(M_{T-1}) + \varepsilon_{T-1} \tag{2}$$

$$log(\lambda_T) \sim Normal\left(\overline{log(\lambda)}, \sigma^2_{Process}\right) \tag{3}$$

$$\varepsilon_T \sim Normal\left(0, \varepsilon^2_{OE}\right) \tag{4}$$

$$(log(M_{t1}) - log(M_{t2})\ )^2 \sim Gamma\left(1/2, 1/2{*}1/\varepsilon^2_{OE}\right) \tag{5}$$

Where Eq. (1) shows the basic underlying process model, Eq. (2) expresses the process using observed population sizes and observation errors (ε_T values), and Eqs. (3)–(5) show the distributional relationships of growth rates and observation errors, where ε^2_{OE} is the estimated observation error variance and M_{t1} and M_{t2} are the values of each repeated survey within one year. Each pair of double survey values provide one estimate of ε^2_{OE} and these estimates are assumed to follow a chi-square distribution (Cochran, 1934), fit in our models as a gamma distribution using the relationship between chi-square and gamma variants (Hogg et al., 1978). The gamma distribution in Eq. (5) is shown with the shape and rate parametrization used in JAGS. Note that in JAGS the normal distribution is parameterized with precision, the inverse of the variance, but to explain the model we express the normal with variance as is more common in the general literature. Finally, in the actual model we also include a random effect of *Plot* for the distribution of the mean $\log(\lambda_T)$ to account for variation between sampling units within a population, and in JAGS the equations shown above are rearranged to facilitate parameter estimation (see code in Appendix E).

Prior to applying this model to our data, we tested if the stochastic exponential model performed well on the different abundance metrics used in the Adopt a plant program by comparing its estimates with three different approaches commonly used to analyze resurvey data: the N-mixture model of Royle (2004) for counts, the dynamic site occupancy model of MacKenzie et al. (2002) for presence/absence data and the zero-augmented beta regression proposed by Wright et al. (2017) for plant cover (Appendix D). We simulated data with varying growth rates, plant detectability, process variances and proportion of observation error relative to the process variance, as well as different number of plots. While we simulated data using parameter values that were broadly comparable to situations seen with the Adopt a Plant program, we also simulated data with a broader range of parameters to test the applicability of our approach to other studies. In most instances the stochastic exponential model performed comparably well to the rest of the models and correctly estimated the true population mean growth rate and process variance (Appendix D). The only exception occurred when analyzing presence/absence data with a high proportion of observation error or low detectability. In this case, our model was able to accurately estimate the true growth rate of the population but tended to underestimate process variance. Thanks to all the efforts put into ensuring that participants in the Adopt a Plant program provide robust estimates of plant abundance, situations in which observation error is high and plant detectability is low are unlikely. Thus, we considered our exponential stochastic model a suitable approach for analyzing data from all plant populations in our dataset, ensuring the comparability of the estimates between abundance metrics coming from different populations and species.

To fit each population model on the Adopt a Plant data, we ran 4 MCMC chains for 1,000,000 iterations with a burn-in period of 100,000 samples and a thinning interval of 10 samples. We used uninformative priors for all model parameters (Table 1). Convergence of the MCMC chains was assessed visually and with the Gelman-Rubin statistic, with values <1.05 being considered a good indicator of chain convergence. The models were fitted separately with and without OE (henceforth referred to as models with and without OE) to compare results between both methods. In addition, we fitted a model to estimate only the OE of each population, using only the double count information, without the multi-year count data, as a baseline to compare the estimates of OE in the full model.

2.4. *Probability of quasi-extinction*

For each population we calculated the probability of quasi-extinction according to its estimated average growth rate and process variance. To take into account the uncertainty estimated by the Bayesian approach, we extracted the last 1000 samples from each MCMC, for a total of 4000 samples, to obtain samples of the average growth rate and process variance of each population. Then, we computed the probability of quasi-extinction as the cumulative probability for that population's size to fall below 10 % of the initial population size in the next 50 years using the *extCDF* function in the *popbio* package (Stubben and Milligan, 2007), which applies the diffusion approximation method of Lande and Orzack (1988). We set an arbitrary initial size of 5000 for all populations, as the main goal was to showcase the effect of including OE in those estimates rather than calculating the extinction risk itself. Finally, we averaged the results over all 4000 samples to get an estimated overall probability of quasi-extinction (p_{qe}) along with its variance (σ^2_{qe}). This process was

Table 1
Prior distributions for the parameters used in each Bayesian model. Normal distributions are shown with their mean (μ) and precision (τ), and gamma distributions with their shape (r) and rate (s) following the parameterization used in JAGS.

Model	Parameter	Definition
Stochastic exponential model	Growth rate	$log(\lambda_T) \sim Normal(\mu = 0, \tau = 0.001)$
Stochastic exponential model	Process precision	$1/\sigma^2_{Process} \sim Gamma(r = 0.001, s = 0.001)$
Stochastic exponential model	Plot precision	$1/\sigma^2_{Plot} \sim Gamma(r = 0.001, s = 0.001)$
Stochastic exponential model	Precision of the observation error	$1/\varepsilon^2_{OE} \sim Gamma(r = 0.001, s = 0.001)$
Meta-analysis	Growth rate	$log(\lambda) \sim Normal(\mu = 0, \tau = 10^{-6})$
Meta-analysis	Monitoring Unit precision	$1/\sigma^2_{MU} \sim Gamma(r = 0.001, s = 0.001)$
Meta-analysis	Species precision	$1/\sigma^2_{TAXON} \sim Gamma(r = 0.001, s = 0.001)$

repeated separately for each population and for the growth rates and variances estimated by the models with and without OE.

2.5. *Analysis of growth rates*

To characterize estimates of population dynamics across all populations, and to estimate the importance of considering OE and the magnitude of OE across the populations in our study, we followed the approach by Haase et al. (2023) and Pilotto et al. (2020) and conducted a set of Bayesian meta-analyses, using the output from our individual population model fits as input. In the first of these models, we used log *(λ)* and $\sigma_{Process}$ estimated by models with observation error for each plant population. For these analyses we only used data from those populations whose estimates for *log(λ)* and $\sigma_{Process}$ had a Gelman-Rubin statistic between 0.95 and 1.05, indicating good chain convergence and model fit (100 % of populations). Using this approach, we modeled the average trend across populations while taking into account the standard deviation of their posteriors as a measure of their uncertainty. The meta-analysis model assumed Gaussian distributions for *log(λ)* and $\sigma_{Process}$, considered no fixed effects and included two random effects, one for MU and another for plant species nested within that MU to account for differences between monitoring units and species. We ran 4 MCMC chains for 1000,000 iterations with a burn-in period of 100,000 samples and a thinning interval of 10. To summarize the quasi-extinction probabilities, we followed a similar approach but assuming that those probabilities followed a Beta distribution with a *logit* link function and whose shape parameters were estimated *via* moment matching using the following equations: $\alpha = \mu_{qe} {}^{*} log\left(\frac{1}{\sigma^2_{qe}}\right)$ and $\beta = \mu_{qe}\left(1-\mu_{qe}\right){}^{*} log\left(\frac{1}{\sigma^2_{qe}}\right)$ were μ_{qe} and σ^2_{qe} are, respectively, the average quasi-extinction probability of a population in 50 years (p_{qe}) and its variance calculated over the 4000 values of p_{qe} obtained from the posteriors of our initial model. To test if including an estimate of OE affected our estimates of growth rates and their variance as well as quasi-extinction probabilities, we followed the same meta-analysis approach just described, but with the posterior distributions estimated by the model without observation error and compared the resulting posteriors from both meta-analyses. Finally, we tested the differences between different abundance metrics with a meta-regression model, which follows a similar approach as the previous meta-analysis models but also includes a fixed effect for the abundance metric used.

3. Results

The stochastic exponential model performed well on most simulated data (Appendix B, Figs. B1, B2 & B3), correctly estimating the average growth rates and stochastic variances of the simulated populations and showing good convergence for a wide range of parameters. When fit to our real data, including the observation error led to narrower credible intervals in both parameters (Fig. 1A & 1B). Longer data series improved the fit of the model and led to narrower credible intervals of the estimated parameters (Fig. 1C & 1D). Including an estimate of OE in the model did not have a significant effect on the estimated long-term stochastic growth rates of most populations (Fig. 2A), however it substantially reduced the estimated process variance in all of them (Fig. 2B). Although the proportion of variation associated with OE differed greatly between populations, on average it accounted for 31.8 % (SD = 32.1) of the total variation in observed numbers, and ranged up to a maximum of 98.7 % (Fig. 3A). In most cases, OE estimated directly within the exponential model was similar or lower than that estimated on its own (Fig. 3B), although its 95 % credible intervals tended to be narrower (Appendix A, Fig. A2).

At the individual population level, models without OE estimated that only 3 populations (1.9 %) had average growth rates that deviated from 0 (CI_{95} not overlapping with 0), with one showing an increase in population size over time whereas two showed signs of decline. On the other hand, models with OE estimated that six populations (3.8 %) had growth rates that deviated from stability, of which four had negative growth rates and two positive ones.

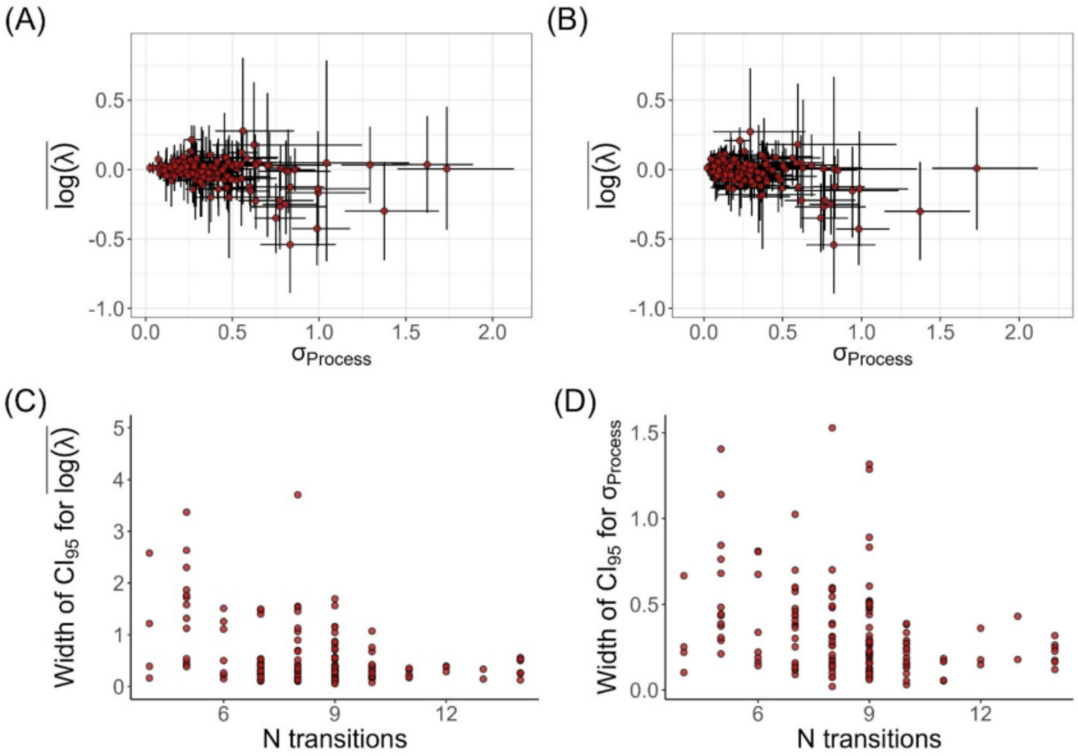

Fig. 1. Relationship between the median estimated *log(λ)* and $\sigma_{Process}$ for the models without (A) and with observation error (B). Panels (C) and (D) show the width of the 95 % credible interval of *log(λ)* and $\sigma_{Process}$ in respect to the number of transitions used in the model with OE.

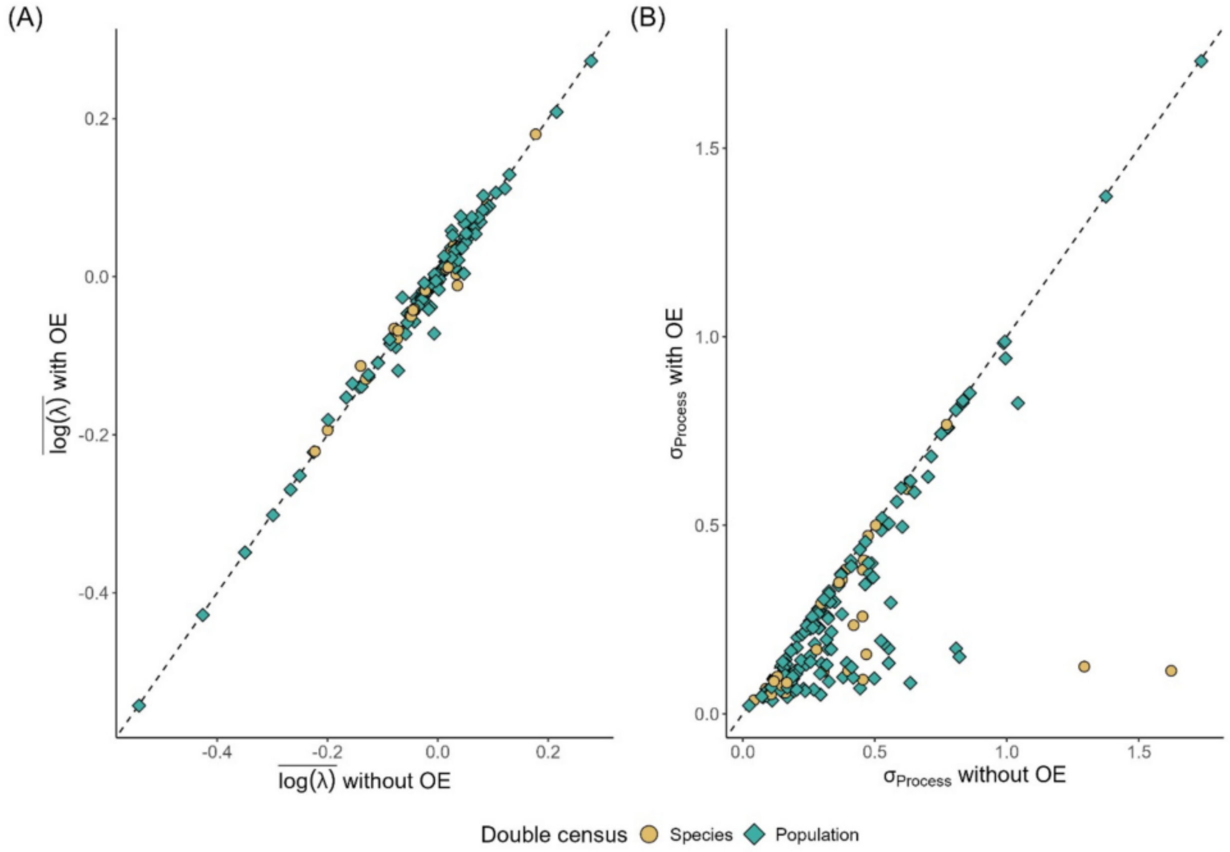

Fig. 2. Relationship between estimates of *log(λ)* (A) and $\sigma_{Process}$ (B) obtained from the models with and without observation error.

The Bayesian meta-analysis of the models with observation error estimated an overall population growth rate of −0.0024 with a 95 % credible interval (CI_{95}) of −0.017 and 0.014 (Fig. 4A), and an average $\sigma_{process}$ of 0.25 (CI_{95} = 0.19, 0.31) (Fig. 4B). The average quasi-extinction probability in 50 years was 0.28 (SD = 0.086) (Fig. 4C). In contrast, the average growth rate estimated by models that did not account for OE was −0.0039 (CI_{95} = −0.021, 0.014) with an average $\sigma_{process}$ of 0.37 (CI_{95} = 0.3, 0.44) and an average quasi-extinction probability in 50 years of 0.42 (SD = 0.081). We also found that the magnitude of OEs were similar across all four survey methods, although occupancy had the lowest OE variance (Median = 0.075, CI_{95} = 0.043, 0.11), followed by plant cover (Median = 0.09, CI_{95} = 0.043, 0.14), counts (Median = 0.098, CI_{95} = 0.078, 0.12) and $counts_R$ (Median = 0.15, CI_{95} = 0.062, 0.23).

4. Discussion

In this study we developed a Bayesian approach to population viability analysis that takes into account the variance estimated with easily collected data on observation error. We applied it to data from 157 plant populations living across a wide range of environmental conditions and for which we had measures of observation error taken during the sampling process. We found that including such estimates directly into the population model improved the precision of parameter estimates. Whereas this had little to no impact on estimates of mean growth, it greatly reduced both the estimated stochastic variance of growth rates and the estimated quasi-extinction probabilities. The relative contribution of OE to the estimated variation in growth rates differed between populations, but on average it accounted for almost a third of the total variation. We also compared the observation error estimated by different abundance sampling metrics. Population growth estimates and their credible intervals were similar between metrics, although results based on populations surveyed with occupancy-based methods tended to be more accurate, followed by plant cover and counts of individuals, with counts on only reproductive individuals being the least accurate. However, as only one metric was used for each population and we intentionally chose the metric to suit the studied species, this result must be interpreted cautiously. Most of the plant species that we studied show stable population growth rates over time, with only a very small fraction of them having significantly decreasing or increasing population sizes.

4.1. *Accounting for observation error improves PVA estimates*

The importance of observation error as a source of bias in population viability analysis is widely acknowledged in population viability literature, and different methods have been proposed to account for it (Dennis et al., 2006, 2010; Humbert et al., 2009; Kéry et al., 2009; Staples et al., 2004). However, very few studies on plant population viability have measured observation error directly during the sampling process or incorporated it into their models, although specific efforts in this regard have been made for vegetation cover (Pardo et al., 2015; van Strien et al., 2024; Sykes et al., 1983; Wright et al., 2017). Here, we estimated observation error by following a simple protocol: surveying twice all or part of the population of interest, and parameterizing the variance of those double surveys directly in our models. By considering the OE during the sampling process we were able to decompose the variance of growth rates into its process and observation components, improving our estimates of both. Including an error term slightly increased the estimated growth rates and greatly reduced the estimated variance in comparison with model fits that ignored OE. These results are consistent with theoretical expectations, since observed variance should be the sum of process variance and observation error, and higher process variance will negatively influence population viability

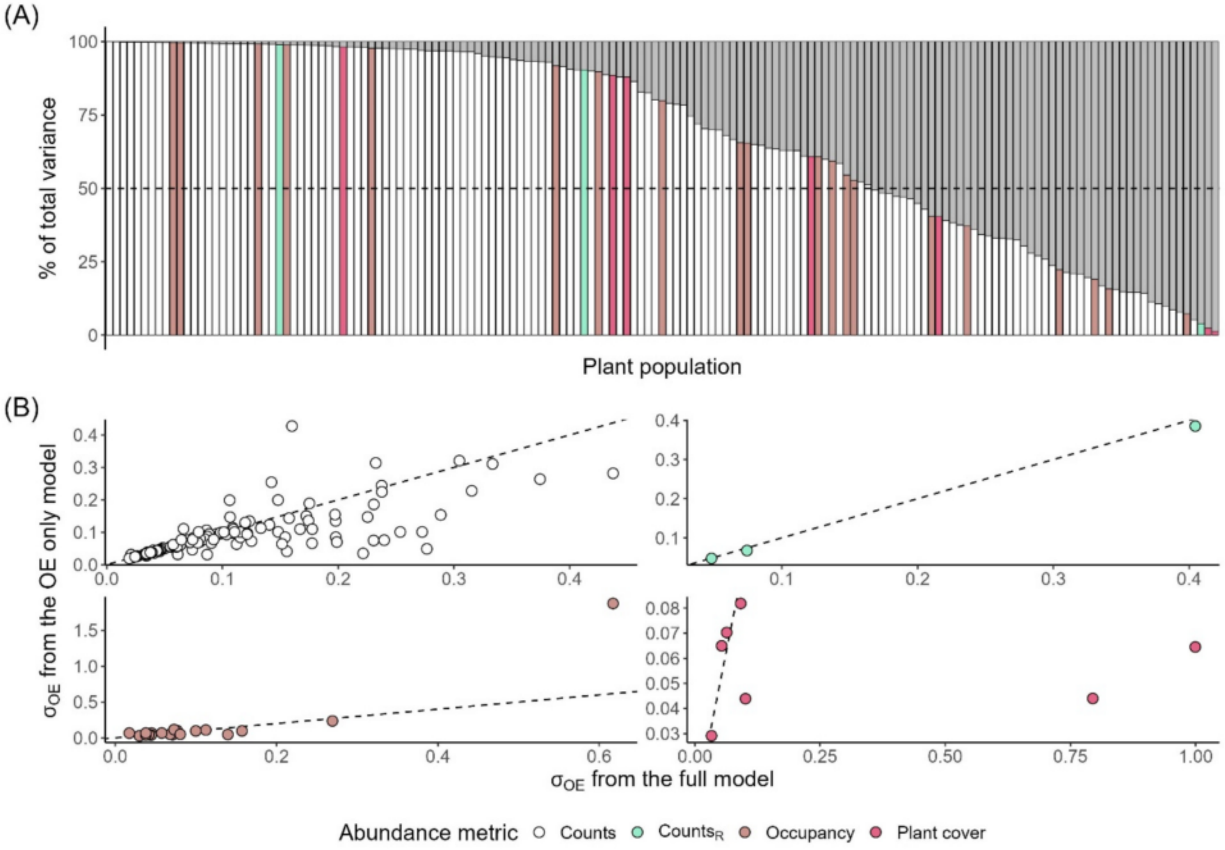

Fig. 3. (A) Proportion of the total variance estimated for each population that corresponds to process variance (color) and observation error (gray). (B) Relationship between the OE variation estimated by the full model and the OE only model. Dashed lines in panel (B) indicate the one-to-one ratio.

(Gillespie, 1977; Lande, 1993; Lande and Orzack, 1988). Accordingly, this reduction in the estimated environmental stochasticity (process variance) and the slight increase in average growth rates should have an important positive effect on the predicted quasi-extinction probabilities of all populations, which were 14 % lower when estimated using the growth rates and variances obtained from the model with observation error.

While the direction of effects we find is entirely as predicted, another important finding was that OE accounted for a meaningful portion of the variation in population sizes and growth rates for most of our species and populations, suggesting that ignoring it when assessing the viability of populations might significantly bias our estimates of growth and extinction (McNamara and Harding, 2004; See and Holmes, 2015).

We also found that longer time series resulted in better estimates for all population parameters, evidenced by narrower credible intervals in growth rates and their variance. Although the use of repeated measures through time is a well-known method for improving the estimates of different PVA parameters (Dennis et al., 2010; Knape et al., 2013; See and Holmes, 2015), to our knowledge this is the first study to directly estimate the error associated with the observation process for such a large and varied set of plant populations and years. A simple procedure like this could be implemented in other monitoring programmes to improve estimates of extinction risk and thus contribute to better population management (Lindenmayer and Likens, 2010).

While taking some double count data is clearly valuable, these can be used in different ways to estimate OE. In particular, an alternative approach to ours would be making a separate estimate of OE using only the double count data, then incorporating these estimates into a model to estimate the mean and variance in growth rates. However, we found that incorporating OE estimation directly into the model led not only to better estimates of the demographic parameters, but also to more accurate estimates of observation error itself. While this means using a more involved analysis method, it is still not difficult and could also be easily modified to, for example, include models in which annual growth rates depend upon environmental drivers such as climate variables, or is density-dependent.

4.2. *The importance of the sampling method*

When it comes to sampling plant populations and its changes in abundance, many metrics can be used, although the choice between them often boils down to the detectability of the focal plant (Elzinga et al., 2015; Morrison, 2016). This, in turn, is determined by the size of the plant, its density in the study area, how easy it is to tell apart from other plants or the density of the surrounding vegetation (Morrison, 2016; Perret et al., 2023). The goal of choosing the appropriate abundance metric and sampling method is to reduce any biases during the observation process and produce the most accurate estimates of abundance possible (Bonham, 2013). The Adopt a Plant program follows the workflow proposed by García et al. (2021) to choose between four abundance metrics widely used in plant ecology and conservation (Bonham, 2013; Elzinga et al., 2015). Each metric is chosen according to the characteristics of the focal plant within its local context, in order to minimize observation error. Thus, the abundance of plants that are difficult to count, such as those forming mats, was assessed using presence/absence or abundance-cover methods in quadrats. For larger species with clearer distinction between individuals we used counts of all individuals when possible or only reproductive individuals when distinction between those in vegetative states was difficult. Plots and quadrats are placed in a stratified manner across the population, including different environmental conditions and plant densities while avoiding extensive sampling of areas with extremely high or low

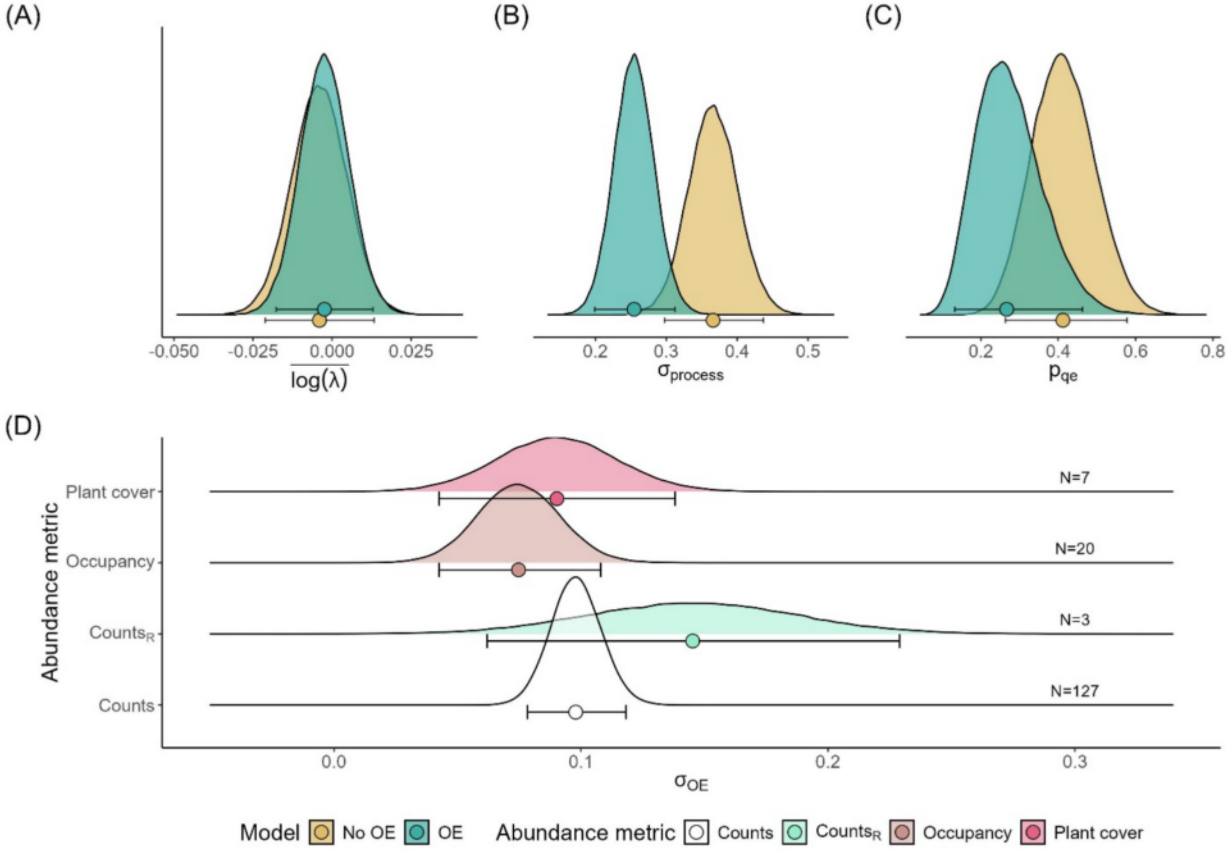

Fig. 4. Posterior distributions of the average long-term stochastic growth rate (A) of all populations, their variance (B) and the average quasi-extinction probability (C) estimated by the models with (blue) and without observation error (yellow). (D) Posterior distributions of the estimated effect of sampling method on observation error along with the number of populations where that method is used. In all panels points and whiskers indicate the median value of the corresponding distribution and its 95 % credible interval. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

density. Many authors have discussed the potential consequences of preferential sampling based on abundance of the focal organism, which may lead to spurious positive or negative estimates of abundance change when plots are placed in very dense or sparse areas (Aubry et al., 2024; Fournier et al., 2019; Mentges et al., 2021). However, by sampling a wide number of conditions and densities within each population the design used in the Adopt a Plant program may be less affected by these biases. In addition, plant detectability should not be affected by preferential sampling and thus estimates of observation error obtained with the scheme used in the Adopt a Plant program are likely to be robust to potential biases in this regard.

The differences in estimated sampling error between metrics used in our study were low. Initially, we expected quadrat-based methods, such as occupancy or plant-cover, to have lower observation error due to their reduced sensitivity to changes in their respective metrics in comparison to counts of individuals (Vittoz et al., 2010). However, the variance associated with observation error in counts was comparable to other methods, suggesting that incorporating measures to increase plant detectability (*i.e.,* observer training, permanent plots, monitoring of multi-annual plant species) can reduce observation error to similar levels across measurement metrics while keeping the higher resolution of counting individuals. On the other hand, counting reproductive individuals led to the highest error rates, which was unexpected since colorful flowers and fruit tend to be more conspicuous and thus easily detected than plants alone (Perret et al., 2023). However, counts based on flowering plants can vary considerably in a few days, and only three of the populations in our study used that abundance metric, so this pattern might be a consequence of the small sample size for this method. With this exception, the patterns of observation error are fairly consistent with what could be intuitively expected, although a direct comparison between different abundance metrics in the same population would be necessary to better understand the relative importance of the sampling method *versus* the local features of the plant populations in determining detectability and observation error. Nevertheless, our results highlight the importance of choosing the appropriate sampling method during the design process of any population under monitoring to reduce any potential biases, putting special attention on plant detectability as well as the context of the focal population (Morrison, 2016). It is also interesting to note that despite our efforts to reduce observation error, all metrics had substantial estimated OE, stressing the importance of using the adequate statistical procedures to control for OE in studies of population growth and viability.

The sampling method of choice also determines how best to analyze the resulting data to estimate plant abundance or occupancy and its changes through time. Many analytical approaches have been developed to deal with the specific needs and assumptions of different metrics, especially when using repeated surveys with imperfect detection. The ideal practice may be to use the method that best accommodates all the assumptions and limitations of a particular data set (*i.e.,* occupancy and plant cover data are inherently bounded between 0 and the number of sites or total area surveyed). N-mixture models, for example, are better suited for repeated counts over several sampling units (Royle, 2004). Similarly, occupancy models work best to estimate the abundance and detection probability of a species based on presence/absence data in discrete sites (MacKenzie et al., 2002), and zero-augmented beta regression approaches are the best choice when working with plant cover data (Wright et al., 2017). When trying to estimate and compare overall abundance changes across populations monitored using different abundance metrics, it may also be useful to have a common method able to give reasonable trend estimates even if some assumptions of the data are not met, and even then, the model can be easily modified to better accommodate these nuances. Our model also allows the estimation of other parameters important for population viability analysis such as the variance of mean growth rates in a straightforward way. This can be easily decomposed into its process and observation error components and compared between populations monitored using different metrics. In addition, these could be used to estimate other common metrics of PVA such as quasi-extinction probabilities or mean time to extinction. Our simulation tests showed that our single approach can be a good option in such circumstances, given that certain assumptions about the population and the sampling methods are met. These include having a relatively constant detection probability, surveying plant species with multi-annual life cycles, using permanent sampling units which are representative of the population as a whole or that few sampling unit is empty or entirely covered by the plant. This good result is certainly caused by the central limit theorem, which guarantees that when data are aggregated, assumptions of normality (on the proper scale) are approximately correct, and thus a single relatively simple approach to population growth estimation can work quite well.

5. Conclusions

Population viability analysis is a fundamental tool in conservation biology, but estimates of extinction risk are heavily influenced by the data that are used to parameterize population models, including inaccuracies and biases in even the simplest metrics of population size or change. These inaccuracies can in turn bias our estimates of important population parameters like growth rates and their variability, namely by giving the impression that populations may fluctuate more than they actually do. Our study demonstrated that by using a simple method such as repeated surveys in all or part of the sampling area of a population, observation error can be taken easily into account, which considerably improves the estimates of viability models. In the current crisis of biodiversity loss, this is a crucial procedure for the correct assessment of the state of populations, which can help our decision-making regarding species conservation.

CRediT authorship contribution statement

Héctor Miranda-Cebrián: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Daniel F. Doak:** Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Formal analysis. **María Begoña García:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization.

Open research statement

Data will be made available on request.

Funding

This paper has been made possible by the support of many projects and >400 collaborators over time. In recent years it has been supported by projects VUBI (PROY_E35_24), BIOTREND (TED2021-131513B-I00), NATURAL PATRIMONY (RECEUSGL2021), and REFUGIA (PID2021-129056OB-I00). HMC was supported by a grant from the Spanish Ministry of Science (grant no. PRE2018-083868). DFD was supported by a grant from the US National Science Foundation no. 1753954.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

The authors would like to thank all the volunteers and personnel involved in the Adopt a Plant program, whose hard work and altruistic dedication have made this work possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2025.111141.

Data availability

Data will be made available on request.

References

Aubry, P., Francesiaz, C., Guillemain, M., 2024. On the impact of preferential sampling on ecological status and trend assessment. Ecol. Model. 492, 110707. https://doi.org/10.1016/j.ecolmodel.2024.110707.

Bonham, C.D., 2013. Measurements for Terrestrial Vegetation. John Wiley & Sons.

Boyce, M., Haridas, C., Lee, C., The NCEAS Stochastic Demography Working Group, 2006. Demography in an increasingly variable world. Trends Ecol. Evol. 21 (3), 141–148. https://doi.org/10.1016/j.tree.2005.11.018.

Buonaccorsi, J.P., Staudenmayer, J., 2009. Statistical methods to correct for observation error in a density-independent population model. Ecol. Monogr. 79 (2), 299–324. https://doi.org/10.1890/07-1299.1.

Carpenter, S.R., Cottingham, K.L., Stow, C.A., 1994. Fitting Predator-Prey Models to Time Series with Observation Errors. Ecology 75 (5), 1254–1264. https://doi.org/10.2307/1937451.

Cochran, W.G., 1934. The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Math. Proc. Camb. Philos. Soc. 30 (2), 178–191. https://doi.org/10.1017/S0305004100016595.

Cowie, R.H., Bouchet, P., Fontaine, B., 2022. The Sixth Mass Extinction: Fact, fiction or speculation? Biol. Rev. 97 (2), 640–663. https://doi.org/10.1111/brv.12816.

Dail, D., Madsen, L., 2011. Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. Biometrics 67 (2), 577–587. https://doi.org/10.1111/j.1541-0420.2010.01465.x.

Dennis, B., Munholland, P.L., Scott, J.M., 1991. Estimation of growth and extinction parameters for endangered species. Ecol. Monogr. 61 (2), 115–143. https://doi.org/10.2307/1943004.

Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L., Staples, D.F., 2006. Estimating density dependence, process noise, and observation error. Ecol. Monogr. 76 (3), 323–341. https://doi.org/10.1890/0012-9615(2006)76[323:EDDPNA]2.0.CO;2.

Dennis, B., Ponciano, J.M., Taper, M.L., 2010. Replicated sampling increases efficiency in monitoring biological populations. Ecology 91 (2), 610–620. https://doi.org/10.1890/08-1095.1.

Denwood, M.J., 2016. Runjags: an R package providing Interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. J. Stat. Softw. 71 (9), 1–25. https://doi.org/10.18637/jss.v071.i09.

Elzinga, C.L., Salzer, D.W., Willoughby, J.W., 2015. Measuring and Monitoring Plant Populations. Createspace Independent Publishing Platform.

Fournier, A.M.V., White, E.R., Heard, S.B., 2019. Site-selection bias and apparent population declines in long-term studies. Conserv. Biol. 33 (6), 1370–1379. https://doi.org/10.1111/cobi.13371.

García, M.B., Silva, J.L., Tejero, P., Pardo, I., 2021. Detecting early-warning signals of concern in plant populations with a Citizen Science network. Are threatened and other priority species for conservation performing worse? J. Appl. Ecol. 58 (7), 1388–1398. https://doi.org/10.1111/1365-2664.13890.

Gillespie, J.H., 1977. Natural selection for variances in offspring numbers: a new evolutionary principle. Am. Nat. 111 (981), 1010–1014. https://doi.org/10.1086/283230.

Haase, P., Bowler, D.E., Baker, N.J., Bonada, N., Domisch, S., Garcia Marquez, J.R., Heino, J., Hering, D., Jähnig, S.C., Schmidt-Kloiber, A., Stubbington, R., Altermatt, F., Álvarez-Cabria, M., Amatulli, G., Angeler, D.G., Archambaud-Suard, G., Jorrín, I.A., Aspin, T., Azpiroz, I., Welti, E.A.R., 2023. The recovery of European freshwater biodiversity has come to a halt. Nature 620 (7974), 582–588. https://doi.org/10.1038/s41586-023-06400-1.

Hogg, R.V., Craig, A.T., Craig, A.T., 1978. Introduction to Mathematical Statistics, 4. ed. Macmillan.

Holmes, E.E., 2001. Estimating risks in declining populations with poor data. Proc. Natl. Acad. Sci. 98 (9), 5072–5077. https://doi.org/10.1073/pnas.081055898.

Humbert, J.-Y., Scott Mills, L., Horne, J.S., Dennis, B., 2009. A better way to estimate population trends. Oikos 118 (12), 1940–1946. https://doi.org/10.1111/j.1600-0706.2009.17839.x.

Kellner, K.F., Swihart, R.K., 2014. Accounting for imperfect detection in ecology: a quantitative review. PLoS One 9 (10), e111436. https://doi.org/10.1371/journal.pone. 0111436.

Kéry, M., 2004. Extinction rate estimates for plant populations in revisitation studies: importance of detectability. Conserv. Biol. 18, 570–574. https://doi.org/10.1111/j.1523-1739.2004.00105.x.

Kéry, M., Dorazio, R.M., Soldaat, L., Van Strien, A., Zuiderwijk, A., Royle, J.A., 2009. Trend estimation in populations with imperfect detection. J. Appl. Ecol. 46 (6), 1163–1172. https://doi.org/10.1111/j.1365-2664.2009.01724.x.

Knape, J., Besbeas, P., De Valpine, P., 2013. Using uncertainty estimates in analyses of population time series. Ecology 94 (9), 2097–2107. https://doi.org/10.1890/12-0712.1.

Lande, R., 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am. Nat. 142 (6), 911–927 (JSTOR).

Lande, R., Orzack, S.H., 1988. Extinction dynamics of age-structured populations in a fluctuating environment. Proc. Natl. Acad. Sci. 85 (19), 7418–7421. https://doi.org/10.1073/pnas.85.19.7418.

Lindenmayer, D.B., Likens, G.E., 2010. The science and application of ecological monitoring. Biol. Conserv. 143 (6), 1317–1328. https://doi.org/10.1016/j.biocon.2010.02.013.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83 (8), 2248–2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2.

McLoughlin, P.D., Messier, F., 2004. Relative Contributions of Sampling Error in Initial Population Size and Vital Rates to Outcomes of Population Viability Analysis: *Sampling Error in PVA*. Conserv. Biol. 18 (6), 1665–1669. https://doi.org/10.1111/j.1523-1739.2004.00500.x.

McNamara, J.M., Harding, K.C., 2004. Measurement error and estimates of population extinction risk. Ecol. Lett. 7 (1), 16–20. https://doi.org/10.1046/j.1461-0248.2003.00550.x.

Mentges, A., Blowes, S.A., Hodapp, D., Hillebrand, H., Chase, J.M., 2021. Effects of site-selection bias on estimates of biodiversity change. Conserv. Biol. 35 (2), 688–698. https://doi.org/10.1111/cobi.13610.

Morris, W.F., Doak, D.F., 2002. *Quantitative conservation biology: theory and practice of population viability analysis*. Sinauer Associates.

Morrison, L.W., 2016. Observer error in vegetation surveys: a review. J. Plant Ecol. 9 (4), 367–379. https://doi.org/10.1093/jpe/rtv077.

Ono, K., Langangen, Ø., Stenseth, N. Chr., 2019. Improving risk assessments in conservation ecology. Nat. Commun. 10 (1). https://doi.org/10.1038/s41467-019-10700-4.

Pardo, I., Doak, D.F., García-González, R., Gómez, D., García, M.B., 2015. Long-term response of plant communities to herbivore exclusion at high elevation grasslands. Biodivers. Conserv. 24 (12), 3033–3047. https://doi.org/10.1007/s10531-015-0996-3.

Perret, J., Besnard, A., Charpentier, A., Papuga, G., 2023. Plants stand still but hide: imperfect and heterogeneous detection is the rule when counting plants. J. Ecol. 111 (7), 1483–1496. https://doi.org/10.1111/1365-2745.14110.

Pilotto, F., Kühn, I., Adrian, R., Alber, R., Alignier, A., Andrews, C., Bäck, J., Barbaro, L., Beaumont, D., Beenaerts, N., Benham, S., Boukal, D.S., Bretagnolle, V., Camatti, E., Canullo, R., Cardoso, P.G., Ens, B.J., Everaert, G., Evtimova, V., Haase, P., 2020. Meta-analysis of multidecadal biodiversity trends in Europe. Nat. Commun. 11 (1), 3486. https://doi.org/10.1038/s41467-020-17171-y.

Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling.

Plummer, M., 2022. rjags: Bayesian graphical models using MCMC. https://CRAN.R-project.org/package=rjags.

Prendergast, J., Bazeley-White, E., Smith, O., Lawton, J., Inchausti, P., Kidd, D., Knight, S., 2010. The Global Population Dynamics Database [Text/xml]. KNB Data Repository. https://doi.org/10.5063/F1BZ63Z8.

Royle, J.A., 2004. *N*-Mixture models for estimating population size from spatially replicated counts. Biometrics 60 (1), 108–115. https://doi.org/10.1111/j.0006-341X.2004.00142.x.

See, K.E., Holmes, E.E., 2015. Reducing bias and improving precision in species extinction forecasts. Ecol. Appl. 25 (4), 1157–1165. https://doi.org/10.1890/14-2003.1.

Shaffer, M.L., 1981. Minimum Population Sizes for Species Conservation. BioScience 31 (2), 131–134. https://doi.org/10.2307/1308256.

Staples, D.F., Taper, M.L., Dennis, B., 2004. Estimating population trend and process variation for PVA in the presence of sampling error. Ecology 85 (4), 923–929. https://doi.org/10.1890/03-3101.

Stubben, C., Milligan, B., 2007. Estimating and analyzing demographic models using the popbio package in *R*. J. Stat. Softw. 22 (11). https://doi.org/10.18637/jss.v022.i11.

Sykes, J.M., Horrill, A.D., Mountford, M.D., 1983. Use of Visual Cover Assessments as Quantitative Estimators of Some British Woodland Taxa. J. Ecol. 71 (2), 437. https://doi.org/10.2307/2259726.

Tuljapurkar, S., 1990. Population Dynamics in Variable Environments, vol. 85. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-642-51652-8.

Van Strien, A.J., Irvine, K.M., Retel, C., 2024. Trends in plant cover derived from vegetation plot data using ordinal zero-augmented beta regression. J. Veg. Sci. 35 (4), e13295. https://doi.org/10.1111/jvs.13295.

Vittoz, P., Bayfield, N., Brooker, R., Elston, D.A., Duff, E.I., Theurillat, J.-P., Guisan, A., 2010. Reproducibility of species lists, visual cover estimates and frequency methods for recording high-mountain vegetation: reproducibility of methods for recording high-mountain vegetation. J. Veg. Sci. 21 (6), 1035–1047. https://doi.org/10.1111/j.1654-1103.2010.01216.x.

Walker, K.J., Stroh, P.A., Ellis, R.W., 2017. *Threatened plants in Britain and Ireland*. Botanical Society of Britain and Ireland.

Wilson, H.B., Kendall, B.E., Possingham, H.P., 2011. Variability in Population Abundance and the Classification of Extinction Risk. Conserv. Biol. 25 (4), 747–757. https://doi.org/10.1111/j.1523-1739.2011.01671.x.

Wright, W.J., Irvine, K.M., Warren, J.M., Barnett, J.K., 2017. Statistical design and analysis for plant cover studies with multiple sources of observation errors. Methods Ecol. Evol. 8 (12), 1832–1841. https://doi.org/10.1111/2041-210X.12825.