Fay, Rémi, Authier, Matthieu, Hamel, Sandra, Jenouvrier, Stéphanie, van de Pol, Martijn, Cam, Emmanuelle, Gaillard, Jean-Michel, Yoccoz, Nigel G., Acker, Paul, Allen, Andrew, Aubry, Lise M., Bonenfant, Christophe, Caswell, Hal, Coste, Christophe F.D., Larue, Benjamin, Le Cœur, Christie, Gamelon, Marlène, Macdonald, Kaitlin R., Moiron, Maria, Nicol-Harper, Alex, Pelletier, Fanie, Rotella, Jay, Teplitsky, Celine, Touzot, Laura, Wells, Caitlin, and Sæther, Bernt-Erik (2022) ***Quantifying fixed individual heterogeneity in demographic parameters: performance of correlated random effects for Bernoulli variables.*** **Methods in Ecology and Evolution, 13 (1) pp. 91-104.**

https://doi.org/10.1111/2041%2D210X.13728

Methods in Ecology and Evolution

DR RÉMI FAY (Orcid ID : 0000-0002-7202-367X)

DR STEPHANIE JENOUVRIER (Orcid ID : 0000-0003-3324-2383)

MRS. PAUL ACKER (Orcid ID : 0000-0002-3815-772X)

DR ANDREW M. ALLEN (Orcid ID : 0000-0002-0119-2425)

DR LISE MYRIAM AUBRY (Orcid ID : 0000-0003-3318-7329)

DR CHRISTOPHE BONENFANT (Orcid ID : 0000-0002-9924-419X)

MR BENJAMIN LARUE (Orcid ID : 0000-0002-4608-9288)

DR MARLÈNE GAMELON (Orcid ID : 0000-0002-9433-2369)

KAITLIN R MACDONALD (Orcid ID : 0000-0002-6376-1486)

PROFESSOR BERNT-ERIK SÆTHER (Orcid ID : 0000-0002-0049-9767)

Article type : Research Article

Quantifying fixed individual heterogeneity in demographic parameters: Performance of correlated random effects for Bernoulli variables

Rémi Fay[1*], Matthieu Authier[2], Sandra Hamel[3], Stéphanie Jenouvrier[4,5], Martijn van de Pol[6], Emmanuelle Cam[7], Jean-Michel Gaillard[8], Nigel G. Yoccoz[9], Paul Acker[1], Andrew Allen[6], Lise M. Aubry[10], Christophe Bonenfant[8], Hal Caswell[11], Christophe F. D. Coste[1], Benjamin Larue[16], Christie Le Cœur[12], Marlène Gamelon[1,8], Kaitlin R. Macdonald[13], Maria Moiron[14], Alex Nicol-Harper[5,15], Fanie Pelletier[16], Jay Rotella[13], Celine Teplitsky[14], Laura Touzot[8], Caitlin Wells[10], Bernt-Erik Sæther[1]

1 Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
2 Observatoire PELAGIS, UMS-CNRS 3462, Université de la Rochelle, La Rochelle, France
3 Département de biologie, Université Laval, Québec City, QC, Canada
4 Centre d'Etudes Biologiques de Chizé, UMR 7372, Centre National de la Recherche Scientifique, 79360 Villiers en Bois, France
5 Biology Department, MS-50, Woods Hole Oceanographic Institution, Woods Hole, MA, USA
6 Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, the Netherlands
7 LEMAR, CNRS, IRD, Ifremer, Université Brest, Plouzané, 29280, France
8 Laboratoire de Biométrie et Biologie Évolutive, CNRS, Unité Mixte de Recherche (UMR) 5558, Université Lyon 1, Université de Lyon, Villeurbanne, France
9 Department of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
10 Fish, Wildlife and Conservation Biology Department, Colorado State University, Fort Collins, CO 80523-1474, USA
11 Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
12 Department of Biosciences, Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo, Norway
13 Department of Ecology, Montana State University, 310 Lewis Hall, Bozeman, MT 59718, USA
14 Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 Campus CNRS, Montpellier, France
15 School of Ocean and Earth Science, National Oceanography Centre, University of Southampton Waterfront Campus, European Way, Southampton, SO14 3ZH, UK
16 Département de Biologie, Université de Sherbrooke, Sherbrooke, Québec, Canada

* corresponding author: fay.remi@gmail.com

running headline: correlated random effects for Bernoulli variables

Abstract

1. An increasing number of empirical studies aim to quantify individual variation in demographic parameters because these patterns are key for evolutionary and ecological processes. Advanced approaches to estimate individual heterogeneity are now using a multivariate normal distribution with correlated individual random effects to account for the latent correlations among different demographic parameters occurring within individuals. Despite the frequent use of multivariate mixed models, we lack an assessment of their reliability when applied to Bernoulli variables.

2. Using simulations, we estimated the reliability of multivariate mixed effect models for estimating correlated fixed individual heterogeneity in demographic parameters modelled with a Bernoulli distribution. We evaluated both bias and precision of the estimates across a range of scenarios that investigate the effects of life-history strategy, levels of individual heterogeneity, and

presence of temporal variation and state-dependence. We also compared estimates across different sampling designs to assess the importance of study duration, number of individuals monitored, and detection probability.

3. In many simulated scenarios, the estimates for the correlated random effects were biased and imprecise, which highlight the challenge in estimating correlated random effects for Bernoulli variables. The amount of fixed among-individual heterogeneity was frequently overestimated, and the absolute value of the correlation between random effects was almost always underestimated. Simulations also showed contrasting performances of mixed models depending on the scenario considered. Generally, estimation bias decreases and precision increases with slower pace of life, large fixed individual heterogeneity, and large sample size.

4. We provide guidelines for the empirical investigation of individual heterogeneity using correlated random effects according to the life-history strategy of the species, as well as, the volume and structure of the data available to the researcher. Caution is warranted when interpreting results regarding correlated individual random effects in demographic parameters modelled with a Bernoulli distribution. Because bias varies with sampling design and life history, comparisons of individual heterogeneity among species is challenging. The issue addressed here is not specific to demography, making this warning relevant for all research areas, including behavioral and evolutionary studies.

Key-words: accuracy, among-individual variation, capture-recapture, generalized linear mixed models, individual quality, joint mixed models, multivariate normal distribution, precision

Introduction

Populations are composed of individuals that differ in their attributes, both at the phenotypic and genetic level, which influences their fitness. This among-individual heterogeneity is ubiquitous across populations and is a fundamental topic in ecology and evolution (Bolnick et al., 2011; Hamel et al., 2018). Among-individual heterogeneity profoundly affects population responses as the average performance of all individuals in a population is typically different from the performance of a population of average individuals (van de Pol & Verhulst, 2006; Vaupel & Yashin, 1985). More generally, individual heterogeneity affects the estimation of critical parameters such as vital rates, population growth rate, and components of demographic variance (i.e., demographic stochasticity, environmental stochasticity, and density dependence), with profound implications for population dynamics, phenotypic selection, and the evolution of life history strategies (Lomnicki, 1978 for a pioneer study; Snyder & Ellner, 2018; Vindenes et al., 2008; Vindenes & Langangen, 2015 for recent developments).

Various definitions of individual heterogeneity have been formulated (Cam et al., 2016; Gimenez et al., 2018; Wilson & Nussey, 2010). In evolutionary and behavioural studies, individual heterogeneity often refers to the among-individual variance observed in a phenotypic trait. In this context, individual heterogeneity is generally trait specific and may vary within individuals over time (e.g., Jolles et al., 2020). Here, we define individual heterogeneity more restrictively as the among-individual variance in demographic parameters. Some of the factors generating individual heterogeneity can be easily observed (e.g., sex, age, size), but some are typically not observed by biologists (e.g., those due to dominance, personality, or genetic make-up). Here, individual heterogeneity refers to this unobserved heterogeneity in demographic parameters that persists after accounting for observed differences such as age, state and sex. Our demographic definition of individual heterogeneity is identical to what has been called *fixed heterogeneity* (Tuljapurkar et al., 2009; van Daalen & Caswell, 2020) or *demographic* *heterogeneity* (Stover et al., 2012), and align with the concept of *frailty*, although the latter is specific to individual variation in survival (Vaupel & Yashin, 1985).

Because quantifying all aspect of phenotypic variation that cause among-individual variation in demographic parameters is impossible, individual heterogeneity is frequently defined as an unmeasured latent variable (Cam et al., 2016). Many recent studies aimed to specifically quantify the amount of among-individual heterogeneity in demographic parameters to evaluate its biological importance and determine its drivers. Two main modelling approaches have been used to estimate individual heterogeneity as a latent variable: the finite mixture models characterizing the presence of unobserved groups (Hamel et al., 2017; Pledger et al., 2003) and the mixed effect models quantifying random individual effects (van de Pol & Verhulst, 2006; Gimenez & Choquet, 2010; Hamel et al., 2018). Although both approaches have been employed to account for and quantify unobserved individual heterogeneity (Gimenez et al., 2018), mixed effect models are most widely used for two reasons. First, they are convenient because they allow for straightforward quantification, interpretation and comparison of heterogeneity across traits and populations. Second, they are easier to implement. Indeed, mixture models often suffer from convergence problems, especially when Bernoulli-distributed traits are included, and defining the appropriate number of groups is not an easy task (Cubaynes et al., 2012; Hamel et al., 2017).

Although many studies have focused on measuring individual heterogeneity for a single demographic parameter, quantifying individual heterogeneity in multiple demographic parameters with their covariation is critical. Indeed, covariation in individual heterogeneity in multiple demographic parameters may reveal biological processes driving individual heterogeneity such as allocation trade-offs or among-individual variation in resource acquisition (van Noordwijk & de Jong, 1986). For instance, some studies found that individuals with a high survival probability also tend to have a high probability of breeding (Cam et al., 2002; McElligott et al., 2002; Pigeon et al., 2017) supporting the hypothesis that the overall covariations shaping individual heterogeneity may correspond to a continuum of low to high quality individuals (Wilson & Nussey 2010). Excitingly, important advances have been made to expand statistical techniques to not only quantify the amount of unobserved heterogeneity in demographic rates, but also look at whether patterns exist in how different demographic rates covary within a population. Recent approaches have quantified individual heterogeneity in multiple demographic parameters based on mixed effect models using a multivariate normal distribution (e.g., Browne et al., 2007; Cam et al., 2002; Knape et al., 2011; Paterson et al., 2018). In these models, normally distributed individual random effects and their covariation are jointly estimated in several demographic parameters accounting explicitly for the non-independence in among-individual heterogeneity occurring in demographic parameters.

No studies, however, have assessed the statistical reliability of multivariate mixed effects models in estimating correlated individual random effects for traits modelled with a Bernoulli distribution (hereafter Bernoulli-distributed traits). Previous studies have investigated the performance of multivariate mixed effects models (also referred to as joint mixed effects models) for normally distributed traits (Martin et al., 2011; van de Pol, 2012). Based on simulations, they found that reliable estimates and statistical inferences could be reached with sample sizes of a few hundred individuals. However, the difficulty in estimating individual heterogeneity could vary with the type of trait. For Bernoulli-distributed traits, accurately estimating individual random effects could be more challenging (Hamel et al., 2012; Kain et al., 2015). Previous studies have shown that in situations where individual variation in continuous traits is accurately estimated, all else being equal, estimates of individual heterogeneity in Bernoulli-distributed traits can be biased (Bonnet & Postma, 2016). Bernoulli-distributed data contain less information than continuous response data (i.e., presence or absence vs. presence, absence and magnitude of the response). Furthermore, data available to estimate individual variability in demographic parameters are generally scarce (Browne et al., 2007). In longitudinal studies of wild populations, individuals are often observed only once or a few times (<5) throughout their lifetime due to imperfect detection and a short lifespan. Thus, the reliability of multivariate mixed effects models to estimate correlated individual random effects for Bernoulli-distributed demographic parameters remains an unresolved issue.

To fill this knowledge gap, we performed simulations to evaluate the reliability of multivariate mixed effects models in estimating correlated among-individual heterogeneity in demographic parameters that follow a Bernoulli distribution. Previous studies suggested that the amount of among-individual heterogeneity and the number of observations for each individual are critical to estimate individual random effects (Kain et al., 2015). Because lifespan affects the number of occasions when an individual can be observed, and thus the amount of information potentially available to estimate demographic parameters, we may expect model performance to vary according to lifespan, and thereby with the life-history strategy of the species considered. We first investigated the effects of life-history strategy and the amount of among-individual heterogeneity on the bias and precision of estimated correlated individual random effects in survival and reproduction. Furthermore, temporal variation and state-dependent variation (i.e., the probability that a given event for individual *i* at time *t* depends on the state of that individual at time *t-1*) which are both pervasive in the wild, can be mistakenly attributed to fixed individual heterogeneity if ignored (Authier et al., 2017; Cam et al., 2016). Positive state-dependence can be particularly problematic because the variation it generates in individual life-history trajectories can mimic that induced by fixed individual heterogeneity (Cam et al., 2016). For instance, if the probability of reproducing successfully is higher after a successful reproductive attempt, state-dependence will generate state-persistence in life histories with some individuals accumulating successes and others accumulating failures, in the same way fixed individual heterogeneity in reproductive success acts. Thus, empirical studies investigating individual heterogeneity have to estimate both individual heterogeneity and state-dependence simultaneously; otherwise, the estimates might be biased because state-dependence and fixed individual heterogeneity could be confounded (Authier et al., 2017; Cam et al., 2016). Thus, in a second step, we assessed the reliability of mixed effects models to estimate correlated individual random effects including temporal variation and positive state-dependence in our simulations. Finally, because the sample size (i.e., the number of individuals monitored) and the design of long-term studies show large variation, we also considered the effect of the number of individuals and the sampling design (i.e., detection probability and duration of the monitoring) on the bias and precision of the correlated individual random effects. We compared results across scenarios to provide guidelines for quantifying individual heterogeneity according to the life-history strategy, the structure and the volume of data available to the researcher.

Material and Methods

1. Data simulation

1.1 Baseline model

We simulated individual life-history trajectories considering two demographic parameters, annual survival probability and reproductive success probability (i.e., the probability of successfully raising at least one offspring to independence), that are each modelled with a

Bernoulli distribution. Each individual's trajectory starts when the individual is recruited as a first-time breeder in the population, and we did not simulate any age effect. The survival process was modelled as follows:

$\mathrm{SURVIVAL}_{it} \sim \mathrm{Bernoulli}\ (\mathrm{logit}^{-1}(\mu_{\Phi} + \alpha_{i,\Phi}))$

Where $\mathrm{SURVIVAL}_{it}$ is the survival of individual *i* from year *t-1* to year *t* and μ_{Φ} is the logit-transform of Φ, which is the average survival probability. Conditional on its survival, individual *i* may breed successfully in year *t* following an additional Bernoulli process where:

$\mathrm{SUCCESS}_{it} \mid (\mathrm{SURVIVAL}_{it}=1) \sim \mathrm{Bernoulli}\ (\mathrm{logit}^{-1}(\mu_{\psi} + \alpha_{i,\psi}))$,

where μ_{ψ} is the logit-transform of ψ, which is the average reproductive success. $\alpha_{i,\Phi}$ and $\alpha_{i,\psi}$ are individual random effects that determine the fate of each individual and follow a multivariate normal distribution:

$$\begin{pmatrix} \alpha_{i,\Phi} \\ \alpha_{i,\psi} \end{pmatrix} \sim \mathrm{MVN} \left(\begin{matrix} 0, \\ 0, \end{matrix} \quad \begin{bmatrix} \sigma^2_{\Phi} & \mathrm{cov}_{\Phi\psi} \\ \mathrm{cov}_{\Phi\psi} & \sigma^2_{\psi} \end{bmatrix} \right),$$

where σ^2_X is the variance of trait x (x refers to either survival Φ or reproductive success ψ) and $\mathrm{cov}_{\Phi\psi}$ is the covariance between the two demographic parameters. The correlation between the two demographic parameter is calculated as $\mathrm{cor}_{\Phi\psi} = \frac{\mathrm{cov}_{\Phi\psi}}{\sigma_{\Phi} \times \sigma_{\psi}}$.

1.2 Full model

Individual variation in demographic parameters may originate from processes other than individual heterogeneity such as temporal variation due to changing environmental conditions and state-dependence, i.e., the probability that a given survival or reproductive event for individual *i* at time *t* depends on the state of that individual at time *t-1*. If not accounted for, positive state-dependency would increase the estimated individual heterogeneity. Inversely, negative state-dependency would lead to an underestimation of individual heterogeneity. To account for these two additional processes, we modified the baseline model by including temporal variation and state-dependence. The model then becomes:

$\mathrm{SURVIVAL}_{it} \sim \mathrm{Bernoulli}\ (\mathrm{logit}^{-1}(\mu_{\Phi} + \alpha_{i,\Phi} + \gamma_{\Phi} \times \mathrm{success}_{i(t-1)} + \varepsilon_{t,\Phi}))$,

and

$\text{SUCCESS}_{it} \mid (\text{SURVIVAL}_{it}=1) \sim \text{Bernoulli}\ (\text{logit}^{-1}(\mu_\psi + \alpha_{i,\psi} + \gamma_\psi \times \text{SUCCESS}_{i(t-1)} + \varepsilon_{t,\psi}))$,

where γ_Φ and γ_ψ are the parameters quantifying the intensity of state-dependence in survival and reproductive success probabilities, respectively. $\varepsilon_{t,\Phi}$ and $\varepsilon_{t,\psi}$ are the temporal random effects simulating the environmental effects following normal distributions of mean 0 and variance $\sigma^2_{\varepsilon_\Phi}$ and $\sigma^2_{\varepsilon_\psi}$, respectively (Authier et al., 2017).

1.3 Parameterization

Based on the models described above, we simulated datasets with different parameter values corresponding to various scenarios (Table 1). For the means μ_Φ and μ_ψ, we considered two sets of values corresponding to a fast and a slow life-history strategy. These values were chosen to reflect the pace of life of a small passerine ($\mu_\Phi = 0.5$ and $\mu_\psi = 0.7$, generation time of two years assuming recruitment at 1 year old) and a long-lived seabird ($\mu_\Phi = 0.9$ and $\mu_\psi = 0.8$, generation time of 19 years assuming recruitment at 10 years). We simulated small and large amounts of individual heterogeneity in survival and reproductive success probability. Because the variance of a Bernoulli process is maximized at a mean probability of 0.5 and is constrained towards 0 as the mean approaches 0 or 1, we slightly adjusted the value representing a small and large amount of individual heterogeneity according to the life-history strategies as traits' means markedly differed between these strategies (Table 1, Fig. 1). The correlation between the random effects was set to be 0.6, based on previous studies reporting positive covariations between demographic parameters (Cam et al., 2002; Fay et al., 2018; McLean et al., 2019). We simulated the absence or the presence of both temporal variation in demographic parameters and state-dependence (Table 1). Specifically, we included positive state-dependence, i.e., higher survival and reproductive success following a successful reproductive event the previous year. Although negative state-dependence is predicted by life-history trade-offs (Bell, 1980), empirical studies on natural populations have frequently reported positive state-dependence (McElligott et al., 2002; Smith, 1981), which may persist even when individual heterogeneity is accounted for (Cam et al., 2013; Zhang et al., 2015). After a failure, we used values for which survival value was Φ-0.1 and reproductive success was ψ-0.1 regardless of the life-history strategy. These values reflect effect sizes reported in empirical studies (e.g., McElligott et al., 2002; Lescroël et al., 2009; Fay et al., 2018). For simplicity, individuals perform as if they were previously unsuccessful at the first occasion. Finally, we also simulated different sampling designs by using data sets consisting of 10, 20, or 40 years of monitoring with 25 or 100 new individuals, recruited as first-time breeders, added each year, which produce six combinations of sample size, from 250 to 4000 individuals. These designs allowed us to disentangle the effect of the number of individuals from the number of years of monitoring. For instance, to investigate fixed individual heterogeneity, one may ask whether it is better to have a sample size of 1000 individuals that come from 20 years of monitoring with 100 new individuals recruited as first breeders per year, or from 40 years of monitoring with 25 new individuals recruited as first breeders per year. Lastly, because animal monitoring in the wild is usually akin to imperfect detection, we simulated datasets with either perfect ($p = 1$) or imperfect detection ($p = 0.5$) (Table 1). Ultimately, our simulations captured two distinct life-history strategies, two levels of individual heterogeneity, the presence or absence of temporal variability and state-dependence, three levels of monitoring duration, two marking effort schemes and two levels of detection probability, thereby leading to 192 scenarios. The parameter space explored was a trade-off between the number of factors investigated and the number of resulting scenarios and computation time. Although the parameter space investigated remained relatively limited, contrasting two or three levels for each factor allowed describing the relevant patterns regarding model performance.

1.4. Mixture of binomial and continuous traits

Reliably estimating individual random effects is particularly challenging for Bernoulli variables, but is easier for non-binary traits (Bonnet & Postma, 2016). One may thus suggest that the inclusion of additional demographic parameters following a Poisson or Normal distribution for instance, would improve the reliability of estimates of other individual random effects on survival and reproduction. Assuming that all random effects are correlated, accurately estimating one may improve the estimation of the others. To evaluate this possibility, we ran six additional scenarios to test how bias and precision of correlated random effects in Bernoulli-distributed traits change when we include a third Poisson-distributed demographic parameter correlated to the previous two. See Appendix S1 for details.

Table 1: Parameter values used to simulate the datasets. For convenience, parameter values are alternatively given on the probability scale (PS) or logit scale (LS).

Biological process			
Parameters	Meaning	Scenarios	Values
Φ	Mean survival (PS)	fast – slow	0.5 or 0.9
ψ	Mean reproductive success (PS)	fast – slow	0.7 or 0.8
σ_{Φ}	Standard deviation of the individual heterogeneity in survival (LS)	low – high	0.2/0.3 or 0.6/0.8
σ_{ψ}	Standard deviation of the individual heterogeneity in reproductive success (LS)	low – high	0.2/0.3 or 0.6/0.8
$cor_{\Phi\psi}$	Correlation between individual random effects for survival and reproductive success (LS)	Quality	0.6
$\sigma_{\varepsilon_\Phi}$	Standard deviation of the temporal variation in survival (LS)	absent - present	0 or 0.5
$\sigma_{\varepsilon_\Phi}$	Standard deviation of the temporal variation in reproductive success (LS)	absent - present	0 or 0.5
γ_Φ	State-dependence in survival (PS)	absent - present	0 or 0.1
γ_ψ	State-dependence in reproduction (PS)	absent - present	0 or 0.1
Sampling process			
Parameters	Meaning	Scenarios	Values
n_{year}	Study duration	low - medium - high	10 or 20 or 40
n_{ind}	Number of individuals marked per year	low – high	25 or 100
p	Detection probability (PS)	imperfect - perfect	0.5 or 1

2. Analyzing the simulated data

To assess the quality of the estimates provided by the multivariate mixed models, we simulated 100 datasets for each scenario, which led to n=19,800 simulated datasets in total. For each dataset, we ran a Bayesian multi-state capture-recapture model identical to the model used to simulate the data. We computed the bias, both non-scaled and scaled, and the precision for all estimates for each scenario using the set of 100 simulated datasets/fitted models. The bias was the difference between the average estimate over the 100 fitted models and the simulated value. The scaled bias was the bias divided by the simulated value. The precision was the average coefficient of variation of the estimate. Ninety five percent credible interval (CRI) coverages were computed over 300 fitted models. To reduce computation time, we computed CRI for a subset of 44 scenarios (over 192) including two distinct life-history strategies, two levels of individual heterogeneity, the simultaneous presence or absence of both temporal variability and state-dependence, three levels of monitoring duration, two levels of detection probability and the low marking effort level (that is 25 new individuals, recruited as first-time breeders, added each year). When simulated datasets had perfect detection, we fixed the detection probability to 1 in the model analyzing the data rather than estimating its value. This corresponds to the choice made in practice when detection is equal or close to 1 in real datasets (e.g., Cam et al., 2002; Knape et al., 2011). This means that for scenarios with perfect detection, we were not using capture-mark-recapture (CMR) models but classical generalized linear mixed models (GLMMs) with correlated random effects. We simulated data using R 3.5.1 (R Core Team, 2018) and conducted all analyses in JAGS (Plummer, 2003) using the 'jagsUI' R package (Kellner, 2016). R and JAGS codes used are provided in Appendix S2. We used a modified Cholesky decomposition (Chen & Dunson, 2003) to specify the prior of the covariance matrix. In order to improve mixing of chains, we used parameter expansion as in (Dunson, 2008), a technique to improve computational efficiency by reducing dependence among MCMC draws (Browne, 2004). Details including the description of the prior used for the covariance matrix is given in Appendix S3. Given the large amount of computation required to fit Bayesian models with individual random effects, the analyses were run on two super computers located in Canada.

Results

Model performance across life histories and amount of individual heterogeneity

Simulations based on the baseline model showed that both life-history strategy and amount of individual heterogeneity have a critical effect on the ability of the model to accurately estimate correlated random effects (Fig. 2). When we simulated datasets corresponding to a fast life-history strategy, individual heterogeneity in both survival and reproductive success tended to be overestimated (relative bias of the two demographic parameters ranging from -6 to +157% according to the level of heterogeneity, sample size and detection probability; Appendix S4, Fig. 2a,b). Correlation and covariation among random effects were strongly underestimated (bias -4 to -96% and -3 to -94% respectively; Fig. 2a,b). Additional simulations with different correlation values showed that it was the absolute value of the correlation that was underestimated (Fig. S1). By contrast, the amount of individual heterogeneity was estimated with smaller bias (bias -11 to +124%) and higher precision for the slow strategy. The correlation and covariation between random effects were also underestimated for the slow strategy and bias was potentially large, but slightly smaller than for a fast life-history strategy (bias -0.5 to -94% and -1 to -89% respectively; Fig. 2a,b). Finally, the amount of simulated individual heterogeneity had a strong effect on the reliability of estimates. When simulated individual heterogeneity was high, bias decreased for individual heterogeneity estimates (bias -8 to +16% compared with -4 to +157% for low individual heterogeneity) and for estimates of correlations and covariation between random effects (bias -0.5 to -76% compared with -21 to -96% and -1 to -65% compared with -19 to -94% respectively), but estimates generally became less precise (Fig. 2a,b). Coverages of 95% CRI were generally high (>80%) showing that despite frequent bias, CRIs were large and included the true parameter value most of the time.

The effect of temporal variation and state-dependence

The full model included two additional processes: temporal variation and positive state-dependence. Generally, the inclusion of these processes made the estimation of the variance of individual random effects and correlation and covariation between random effects more challenging (Fig. 2c,d). The inclusion of these processes accentuated the bias and decreased the precision compared with the estimates obtained from the baseline model. These effects were independent of the simulated amount of individual heterogeneity, but it was more detrimental for fast life-history strategies. For this latter, relative bias in the estimated individual heterogeneity ranged from -6 to +157% for the baseline model compared with a range of -14 to +225% when including temporal variation and state-dependence (Appendix S4, Fig. 2). In contrast, for the slow life-history strategy, bias ranged from -11 to +124% for the baseline model and from -7 to +129% for the full model (Fig. 2). Although both temporal variation and state-dependence tended to decrease the quality of the estimates, they did not contribute equally to this deterioration. An increased bias was mainly observed when state-dependence was included (Appendix S4). While we obtained unbiased estimation of temporal variation on average for most scenarios, state-dependence estimates were frequently biased negatively. When state-dependence was underestimated, the individual random effects were strongly positively biased, which suggests that random effects captured part of the individual heterogeneity in demographic parameters generated by the positive state-dependence.

Influence of sampling design

As expected, sampling design had a strong effect on the performance of the estimates. All else being equal, increasing the study duration, the number of marked individuals, and the detection probability reduced bias and increased precision (Figs 3 & 4). However, increasing sample size may reduce 95% CRI coverage when estimates are biased (Appendix S4). For large sample sizes (≥1,000 recruited individuals), individual heterogeneity estimates were fairly accurate for most scenarios for the slow life-history strategy (relative bias -6 to +6%) (Figs 3e, f & 4b-f). However, the clear underestimation of the correlation and covariation between random effects persisted when we simulated low individual heterogeneity (relative bias -30 to -60% and -31% to -0.56% respectively; Figs 3e & 4a,c,e). For the fast life-history strategy, the bias persisted for a sample size of 1000 individuals, especially when the model included state-dependence (Appendix S4). With very large sample sizes (4,000 recruited individuals), a slight bias of the individual heterogeneity in reproductive success persisted (relative bias -11 to +12%) and the correlation and covariation between random effects were still strongly underestimated (relative bias -15 to -93% and -12 to -91% respectively; Fig. 4e,f). Doubling the study duration had more impact on improving the reliability of estimates than doubling the number of marked individuals per year for the slow life-history strategy (Figs 3c,d vs. 3e,f & 4a,b), but not for the fast life-history strategy. Imperfect detection generally increased the bias and lowered the precision, but these effects were stronger for the estimated amount of individual heterogeneity. When detection probability decreased from 1 to 0.5, the bias in individual heterogeneity increased from 0 to 232%, whereas the bias in the correlation and covariation between the random effects increased from 0 to 47% and from 7 to 832% respectively (Fig. S2).

Mixture of Bernoulli- and Poisson-distributed demographic parameters

Additional simulations showed that including Poisson-distributed traits - such as the number of offspring produced per successful breeding attempt - had only a weak effect on the ability of the model to accurately estimate correlated random effects for the Bernoulli-distributed traits. Although estimated individual heterogeneity of the demographic parameter following a Poisson distribution was unbiased and precise under all simulated scenarios, bias in correlated individual random effects for Bernoulli-distributed traits were almost unchanged (Fig. S3). Individual heterogeneity estimates for Bernoulli variables were the same with and without the Poisson variable and the bias of the correlation between random effects was slightly decreased only for high individual heterogeneity. For high individual heterogeneity, relative bias in the estimated correlation ranged from -0.22 to -0.64% when modelling only Bernoulli traits compared with -0.12 to -0.60% when including a Poisson-distributed trait, whereas for low individual heterogeneity, the relative bias remained the same with or without the inclusion of a Poisson-distributed trait, ranging from -0.93 to -0.96%.

Discussion

Although ecologists have shown increasing interest in estimating individual heterogeneity by modelling correlated random effects in multivariate mixed models (Bonnet & Postma, 2016; Cam et al., 2013; Knape et al., 2011; Paterson et al., 2018), an assessment of how reliably these models quantify individual heterogeneity in demographic parameters was lacking. Our simulations fill this gap and reveal that estimating correlated random effects for Bernoulli variables is challenging because estimations of fixed individual heterogeneity in survival and reproductive success and their correlation could be strongly biased and imprecise for most of the scenarios investigated in our study. Simulations also indicated that bias in estimates strongly depends on the life-history strategy of the species, which we measured by the species pace of life (generation times spanning over an order of magnitude), as well as the amount of individual heterogeneity and the sample size, both covering the ranges commonly reported in empirical studies. Generally, estimates become less biased and more precise when a large sample size was obtained from a focal population that had a slow pace of life, and higher individual heterogeneity. Although our study raises concerns regarding the biological interpretation of previously published empirical estimates of correlated individual random effects for Bernoulli-distributed demographic parameters, it also provides useful guidelines for future empirical studies determining under which conditions reliable estimates could be obtained depending on the type of life-history strategy of the study species considered. Most importantly, although we address this issue in a demographic framework, the problem treated here is not specific to demography and these results are relevant for all research areas using correlated random effects for Bernoulli-distributed traits.

Bias in the variance and correlation estimates

In many scenarios, estimates from the variance-covariance matrix were biased and imprecise. Although individual heterogeneity was frequently overestimated, the absolute value of the correlation between random effects was almost always underestimated. Bias tended to be stronger for estimates of the correlation between random effects than for estimates of the amount of individual heterogeneity. Here, we used the same model for generating and analysing the data, meaning that we describe model performances under the best-case scenario. These results confirm the concerns raised by Knape et al. (2011) regarding the large uncertainty associated with empirical estimates of correlated random effects.

Correlation and standard deviation of the individual heterogeneity in survival and reproduction are directly related since $cor_{\Phi\psi} = cov(\Phi,\psi)/(\sigma_{\Phi} \times \sigma_{\psi})$. Therefore, for a given covariance level, an overestimation of individual heterogeneity also results in an underestimation of the correlation. The systematic underestimation of the correlation among random effects is in line with results from previous simulations investigating the reliability of temporal correlation estimates among demographic components (Riecke et al., 2019). Based on datasets including temporal random effects simulated with a multivariate normal distribution, these authors found that the temporal correlation could be underestimated whichever the sign of the true correlation. The systematic underestimation of covariation could be compared with the well-known problem of regression dilution in linear models. When fitting a linear model, the random measurement error in the explanatory variable systematically biases the estimate of the regression slope toward zero (Spearman, 1904). In our case, the estimation error in individual heterogeneity causes a systematic underestimation of the correlation between random effects.

The priors we used for the covariance matrix may affect the observed bias and precision. To assess the sensitivity of the results to prior choice, we re-ran simulations with two different priors (Appendix S5). Results show that this choice of prior distributions has a small effect on the magnitude of the bias in the estimated amount of individual heterogeneity and correlation between random effects (Fig. S4). Overestimation of the individual heterogeneity and strong underestimation of the correlation estimate were observed irrespective of the prior used. The prior used for our simulation study tends to shrink the correlation estimates toward 0. This effect was expected since we used a slightly informative prior favouring a null value for the correlation (Appendix S3). Ensuring a marginal uninformative prior is straightforward in the case of a 2-by-2 covariance matrix but not for matrices of higher dimensions. The advantage of the prior used for this simulation study is that it can be used for more than two traits (e.g., Cam et al., 2013; Appendix S1). Finding priors with marginal uniform correlations for multivariate covariance matrices is an active area of research (Huang & Wand, 2013).

Effect of the pace of life and state-dependence on estimates

We found contrasting model performance depending on the pace of life of the species. Although the two life-history strategies simulated may not be representative of the whole slow-fast continuum, they clearly suggest patterns according to the species' pace of life and reveal key aspects affecting model performances. Estimates were substantially less biased and more precise for the slow life-history strategy, especially for individual heterogeneity in reproductive success. This contrasting performance according to the life-history is likely due to variation in the number of reproductive attempts per individual caused by differences in the average lifespan within contrasting life-history strategies. In our simulations, individuals bred once a year, meaning individuals with a fast strategy (mean survival = 0.5 leading to an adult life expectancy of 1 year) reproduced twice on average (i.e., at recruitment and the year after), whereas individuals with a slow strategy (mean survival = 0.9 leading to an adult life expectancy of 9 years) reproduced ten times on average. Thus, the information available to estimate individual-specific performance in reproduction was larger for individuals with a slow life-history strategy. In contrast, the information available to estimate individual heterogeneity in survival probability was less affected by the pace of life, likely because mortality occurs only once per individual regardless of the pace of life.

Although temporal variation in demographic parameters has weak effects on model performance, the simulation results showed the detrimental effect of positive state-dependence on the estimation of the individual random effects. This reveals the difficulty in disentangling individual heterogeneity in life-history trajectory due to positive state-dependence, from that of inherent individual differences in survival and reproductive ability. In many scenarios, individual heterogeneity generated by state-dependence was partly captured by the inflated variances of the individual random effects. Because both processes can replace each other in accounting for state-persistence over time, they have to be estimated simultaneously when both are present to get unbiased estimates (Authier et al., 2017). In accordance with previous studies, our results show that simultaneously estimating state-dependence and fixed individual heterogeneity is challenging (Hamel et al., 2012; Nerlove, 2014). Still, our simulations show that disentangling these processes is possible when sample size is large enough. For the slow life-history strategy, relatively reliable estimates of state-dependence were obtained from sample sizes of 1,000 or more recruited individuals. For the fast life-history strategy, unbiased estimation of state-dependence seemed possible from 4,000 individuals.

Importance of the sampling design

The reliability of estimates of the variance-covariance matrix depended strongly on the study design. Most of the difficulties described above vanished with large sample sizes, that is >1,000 individuals monitored. This demonstrates that observed bias is not due to the unidentifiability of the parameters but rather related to a lack of information in the data. Clearly, reliable estimation of correlated random effects for Bernoulli variables requires very large sample sizes. According to the simulation results, the order of magnitude for an adequate sample size should be >1,000 individuals. Although this is larger than most sample sizes available from individual-based long-term studies in the wild, some datasets meet this requirement (e.g., Cam et al., 2013; Gillespie et al., 2013; Paterson et al., 2018).

For the sampling designs we investigated, study duration seemed to be more influential than total number of individuals monitored for the slow life-history strategy, but not for the fast life-history strategy. This difference according to the pace of life is, again, likely due to the average lifespan associated with each life-history strategy. Since average lifespan of an individual with a slow life-history strategy is longer than that of an individual with a fast life-history strategy, increasing the study duration is more likely to increase the number of observations per individual for the former. As a general rule, it seems more efficient to increase the number of observations per individual than to increase the number of individuals (see Fig. S6 for an illustration of the effect of the number of observations per individuals).

Implications for future research

Results from studies estimating correlated individual random effects among demographic parameters modelled with Bernoulli distribution should be interpreted cautiously because bias is likely to be pervasive and strong. We also found that it can vary according to demographic parameters, pace of life, and true amount of fixed individual heterogeneity. These difficulties make comparative studies very challenging to perform. Although multi-species comparisons play a key role in life-history research, differences in estimates of individual heterogeneity could be affected, or even driven, by biases that change according to the species' pace of life, the amount of individual heterogeneity and sample size. Variable bias according to the true amount of individual heterogeneity is particularly problematic because we cannot know the true amount of fixed individual heterogeneity in any specific demographic parameter *a priori*.

Although our results raise concerns about the biological interpretation of individual random effects for Bernoulli-distributed variables, investigating individual heterogeneity with multivariate mixed models is not a hopeless cause. Studies interested in estimating and interpreting individual heterogeneity from correlated individual random effects should favour the investigation of continuous traits. When Bernoulli-distributed traits are involved, estimating individual heterogeneity reliably is possible if studies are based on very large sample sizes (i.e., thousands of individuals or more for fast life-history species), and include individuals with near complete life histories. Although such datasets are not common, some monitoring on long-lived sea birds, marine mammals, small passerines or humans meet this requirement. Ideally, such studies should simulate data to assess precision of estimates prior to drawing inference on estimated correlated random effects from their data. We provided R code (see Appendix S2) that will help researchers perform customized simulations for their specific study system and sampling design.

Acknowledgments

The idea for this paper was developed during a workshop on "Individual heterogeneity in animal's life histories – more than meets the eye" and we acknowledge the Lorentz Centre of Leiden University for their support and facilitating this meeting. The analyses presented in this paper were run on two super computers located in Canada (Beluga & Cedar). We thank the support provided by Compute Canada (www.computecanada.ca). We acknowledge the support of NSF OPP

1640481 and 1840058 to JR and SJ, respectively. Finally, we thank Julien Martin and two anonymous reviewers for providing useful comments on this manuscript.

Conflicts of Interest: No conflict of interest to declare.

Authors' contributions: All authors discussed the ideas that lead to the study. RF, MA, SA, SJ, MP, EC, JMG, NGY designed the simulation plan; RF and MA developed the R and JAGS codes. RF ran the analyses and led the writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

Data Availability: Code from this study have been deposited in Dryad Digital Repository: https://doi.org/10.5061/dryad.kd51c5b6p Fay et al. (2021).

References

Authier, M., Aubry, L. M., & Cam, E. (2017). Wolf in sheep's clothing: Model misspecification undermines tests of the neutral theory for life histories. *Ecology and Evolution*, *7*(10), 3348–3361.

Bell, G. (1980). The costs of reproduction and their consequences. *The American Naturalist*, *116*(1), 45–76.

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, *26*(4), 183–192.

Bonnet, T., Morrissey, M. B., Morris, A., Morris, S., Clutton-Brock, T. H., Pemberton, J. M., & Kruuk, L. E. (2019). The role of selection and evolution in changing parturition date in a red deer population. *PLoS Biology*, *17*(11), e3000493.

Bonnet, T., & Postma, E. (2016). Successful by chance? The power of mixed models and neutral simulations for the detection of individual fixed heterogeneity in fitness components. *The American Naturalist*, *187*(1), 60–74.

Browne, W. J. (2004). An illustration of the use of reparameterisation methods for improving MCMC efficiency in crossed random effect models. *Multilevel Modelling Newsletter*, *16*(1), 13–25.

Browne, W. J., McCleery, R. H., Sheldon, B. C., & Pettifor, R. A. (2007). Using cross-classified multivariate mixed response models with application to life history traits in great tits (Parus major). *Statistical Modelling*, *7*(3), 217–238.

Cam, E., Aubry, L. M., & Authier, M. (2016). The conundrum of heterogeneities in life history studies. *Trends in Ecology & Evolution*, *31*(11), 872–886.

Cam, E., Gimenez, O., Alpizar-Jara, R., Aubry, L. M., Authier, M., Cooch, E. G., Koons, D. N., Link, W. A., Monnat, J.-Y., & Nichols, J. D. (2013). Looking for a needle in a haystack: Inference about individual fitness components in a heterogeneous population. *Oikos*, *122*(5), 739–753.

Cam, E., Link, W. A., Cooch, E. G., Monnat, J.-Y., & Danchin, E. (2002). Individual covariation in life-history traits: Seeing the trees despite the forest. *The American Naturalist*, *159*(1), 96–105.

Chen, Z., & Dunson, D. B. (2003). Random effects selection in linear mixed models. *Biometrics*, *59*(4), 762–769.

Cubaynes, S., Lavergne, C., Marboutin, E., & Gimenez, O. (2012). Assessing individual heterogeneity using model selection criteria: How many mixture components in capture–recapture models? *Methods in Ecology and Evolution*, *3*(3), 564–573.

Dunson, D. B. (2008). *Random effect and latent variable model selection*. Springer.

Fay, R., Barbraud, C., Delord, K., & Weimerskirch, H. (2018). From early life to senescence: Individual heterogeneity in a long-lived seabird. *Ecological Monographs*, *88*(1), 60–73.

Fay, R., Authier, M., Hamel, S., Jenouvrier, S., van de Pol, M., Cam, E., Gaillard, J. M., Yoccoz, N. G., … Sæther B. E. (2021) Code from: Quantifying fixed individual heterogeneity in demographic parameters: Performance of correlated random effects for Bernoulli variables. Dryad Digital Repository. https://doi.org/10.5061/dryad.kd51c5b6p.

Gimenez, O., Cam, E., & Gaillard, J.-M. (2018). Individual heterogeneity and capture–recapture models: What, why and how? *Oikos*, *127*(5), 664–686.

Gimenez, O., & Choquet, R. (2010). Individual heterogeneity in studies on marked animals using numerical integration: Capture–recapture mixed models. *Ecology*, *91*(4), 951–957.

Hamel, S., Gaillard, J.-M., Douhard, M., Festa-Bianchet, M., Pelletier, F., & Yoccoz, N. G. (2018). Quantifying individual heterogeneity and its influence on life-history trajectories: Different methods for different questions and contexts. *Oikos*, *127*(5), 687–704.

Hamel, S., Yoccoz, N. G., & Gaillard, J.-M. (2012). Statistical evaluation of parameters estimating autocorrelation and individual heterogeneity in longitudinal studies. *Methods in Ecology and Evolution*, *3*(4), 731–742.

Hamel, S., Yoccoz, N. G., & Gaillard, J.-M. (2017). Assessing variation in life-history tactics within a population using mixture regression models: A practical guide for evolutionary ecologists. *Biological Reviews*, *92*(2), 754–775.

Huang, A., & Wand, M. P. (2013). Simple marginally noninformative prior distributions for covariance matrices. *Bayesian Analysis*, *8*(2), 439–452.

Jolles, J. W., King, A. J., & Killen, S. S. (2020). The role of individual heterogeneity in collective animal behaviour. *Trends in Ecology & Evolution*, *35*(3), 278–291.

Kain, M. P., Bolker, B. M., & McCoy, M. W. (2015). A practical guide and power analysis for GLMMs: Detecting among treatment variation in random effects. *PeerJ*, *3*, e1226.

Kellner, K. (2016). *JagsUI: A Wrapper Around "rjags" to Streamline "JAGS" Analyses. R package version 1.4.4. Https://CRAN.R-project.org/package=jagsUI.*

Knape, J., Jonzén, N., Sköld, M., Kikkawa, J., & McCallum, H. (2011). Individual heterogeneity and senescence in Silvereyes on Heron Island. *Ecology*, *92*(4), 813–820.

Lescroël, A., Dugger, K. M., Ballard, G., & Ainley, D. G. (2009). Effects of individual quality, reproductive success and environmental variability on survival of a long-lived seabird. *Journal of Animal Ecology*, 798–806.

Lomnicki, A. (1978). Individual differences between animals and the natural regulation of their numbers. *The Journal of Animal Ecology*, *47*(2), 461–475.

Martin, J. G., Nussey, D. H., Wilson, A. J., & Réale, D. (2011). Measuring individual differences in reaction norms in field and experimental studies: A power analysis of random regression models. *Methods in Ecology and Evolution*, *2*(4), 362–374.

McElligott, A. G., Altwegg, R., & Hayden, T. J. (2002). Age-specific survival and reproductive probabilities: Evidence for senescence in male fallow deer (Dama dama). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1496), 1129–1137.

McLean, E. M., Archie, E. A., & Alberts, S. C. (2019). Lifetime fitness in wild female baboons: Trade-offs and individual heterogeneity in quality. *The American Naturalist*, *194*(6), 745–759.

Nerlove, M. (2014). Individual heterogeneity and state dependence: From George Biddell Airy to James Joseph Heckman. *ØEconomia. History, Methodology, Philosophy*, *4–3*, 281–320.

Paterson, J. T., Rotella, J. J., Link, W. A., & Garrott, R. (2018). Variation in the vital rates of an Antarctic marine predator: The role of individual heterogeneity. *Ecology*, *99*(10), 2385–2396.

Pigeon, G., Festa-Bianchet, M., & Pelletier, F. (2017). Long-term fitness consequences of early environment in a long-lived ungulate. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1853), 20170222.

Pledger, S., Pollock, K. H., & Norris, J. L. (2003). Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics*, *59*(4), 786–794.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, *2*, 1–10.

Riecke, T. V., Sedinger, B. S., Williams, P. J., Leach, A. G., & Sedinger, J. S. (2019). Estimating correlations among demographic parameters in population models. *Ecology and Evolution*, *9*(23), 13521–13531.

Smith, J. N. (1981). Does high fecundity reduce survival in Song Sparrows? *Evolution*, *35*(6), 1142–1148.

Snyder, R. E., & Ellner, S. P. (2018). Pluck or luck: Does trait variation or chance drive variation in lifetime reproductive success? *The American Naturalist*, *191*(4), E90–E107.

Spearman, C. (1904). The proof and measurement of association between two things. *International Journal of Epidemiology*, *39*(5), 1137–1150.

Steiner, U. K., & Tuljapurkar, S. (2012). Neutral theory for life histories and individual variability in fitness components. *Proceedings of the National Academy of Sciences*, *109*(12), 4684–4689.

Stover, J. P., Kendall, B. E., & Fox, G. A. (2012). Demographic heterogeneity impacts density-dependent population dynamics. *Theoretical Ecology*, *5*(2), 297–309.

Tuljapurkar, S., Steiner, U. K., & Orzack, S. H. (2009). Dynamic heterogeneity in life histories. *Ecology Letters*, *12*(1), 93–106.

van Daalen, S., & Caswell, H. (2020). Variance as a life history outcome: Sensitivity analysis of the contributions of stochasticity and heterogeneity. *Ecological Modelling*, *417*(108856), 108856.

van de Pol, M. (2012). Quantifying individual variation in reaction norms: How study design affects the accuracy, precision and power of random regression models. *Methods in Ecology and Evolution*, *3*(2), 268–280.

van de Pol, M., & Verhulst, S. (2006). Age-dependent traits: A new statistical model to separate within-and between-individual effects. *The American Naturalist*, *167*(5), 766–773.

van Noordwijk, A. J., & de Jong, G. (1986). Acquisition and allocation of resources: Their influence on variation in life history tactics. *The American Naturalist*, *128*(1), 137–142.

Vaupel, J. W., & Yashin, A. I. (1985). Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *The American Statistician*, *39*(3), 176–185.

Vindenes, Y., Engen, S., & Sæther, B.-E. (2008). Individual heterogeneity in vital parameters and demographic stochasticity. *The American Naturalist*, *171*(4), 455–467.

Vindenes, Y., & Langangen, Ø. (2015). Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology Letters*, *18*(5), 417–432.

Wilson, A. J., & Nussey, D. H. (2010). What is individual quality? An evolutionary perspective. *Trends in Ecology & Evolution*, *25*(4), 207–214.

Zhang, H., Vedder, O., Becker, P. H., & Bouwhuis, S. (2015). Contrasting between-and within-individual trait effects on mortality risk in a long-lived seabird. *Ecology*, *96*(1), 71–79.

Figure legends:

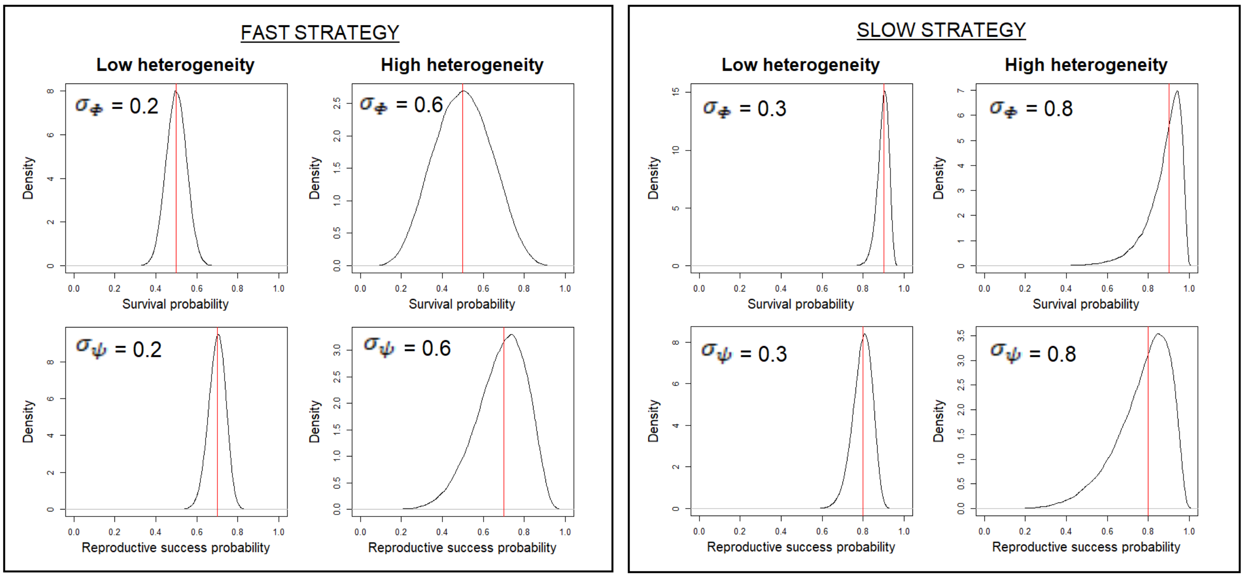
Figure 1: Distribution of the simulated individual heterogeneity (σ) in survival (Φ) and in reproductive success (ψ) probabilities according to the life-history strategy and the amount of individual heterogeneity simulated. The red lines display the means. In each case, individual heterogeneity included in the trajectories corresponds to a continuum of increasing individual performance along the x-axis that can be interpreted as frailty for survival probability and as a measure of reproductive ability for reproductive success probability. The correlation between these axes of performance is positive and thereby corresponds to a continuum of individual quality.

Figure 2: Bias and precision in the estimates of individual random effects and their correlation for Bernoulli variables. Parameters include the standard deviation of the individual heterogeneity in survival (σ_{Φ}) and reproductive success (σ_{ψ}) and their correlation (cor) and covariance (cov). Each plot displays the estimates for a scenario with a specific amount of individual heterogeneity (i.e., low vs. high), and for scenarios that include or exclude temporal variation in demographic parameters and positive state-dependence (i.e., baseline vs. full model). The study duration was 20 years and the number of new individuals recruiting as first-time breeders per year was 25, leading to a sample size of 500 individuals. Diamonds in darker color give the values used to simulate the datasets and points in lighter color give the average estimates over the 100 models fitted to the 100 simulated datasets. Error bars give the range including 95% of the estimated values.

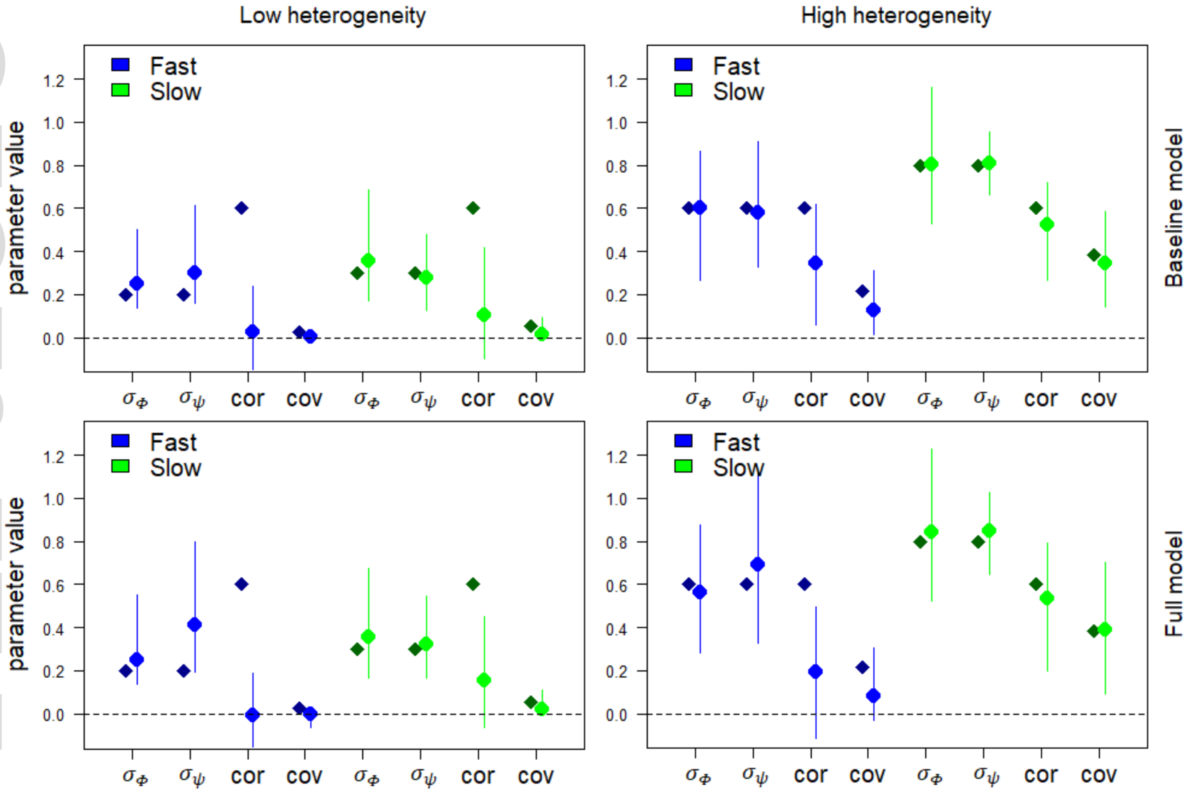
Figure 3: Bias and precision in the estimates of individual random effects and their correlation for Bernoulli variables with the full model, i.e., including temporal variation in demographic parameters and state-dependence. Parameters include the standard deviation of individual heterogeneity in survival (σ_{Φ}) and reproductive success (σ_{ψ}) and their correlation (cor) and covariance (cov). Each plot displays the estimates for a scenario with a specific amount of individual heterogeneity (i.e. low vs. high) and study duration (10, 20 or 40 years). The number of new individuals recruited as first-time breeders per year was 25, leading to the sample sizes of 250, 500 and 1,000 individuals according to the study duration. Diamonds in darker color give the values used to simulate the datasets and points in lighter color give the average estimates over the 100 models fitted to the 100 simulated datasets. Error bars give the range including 95% of the estimated values.

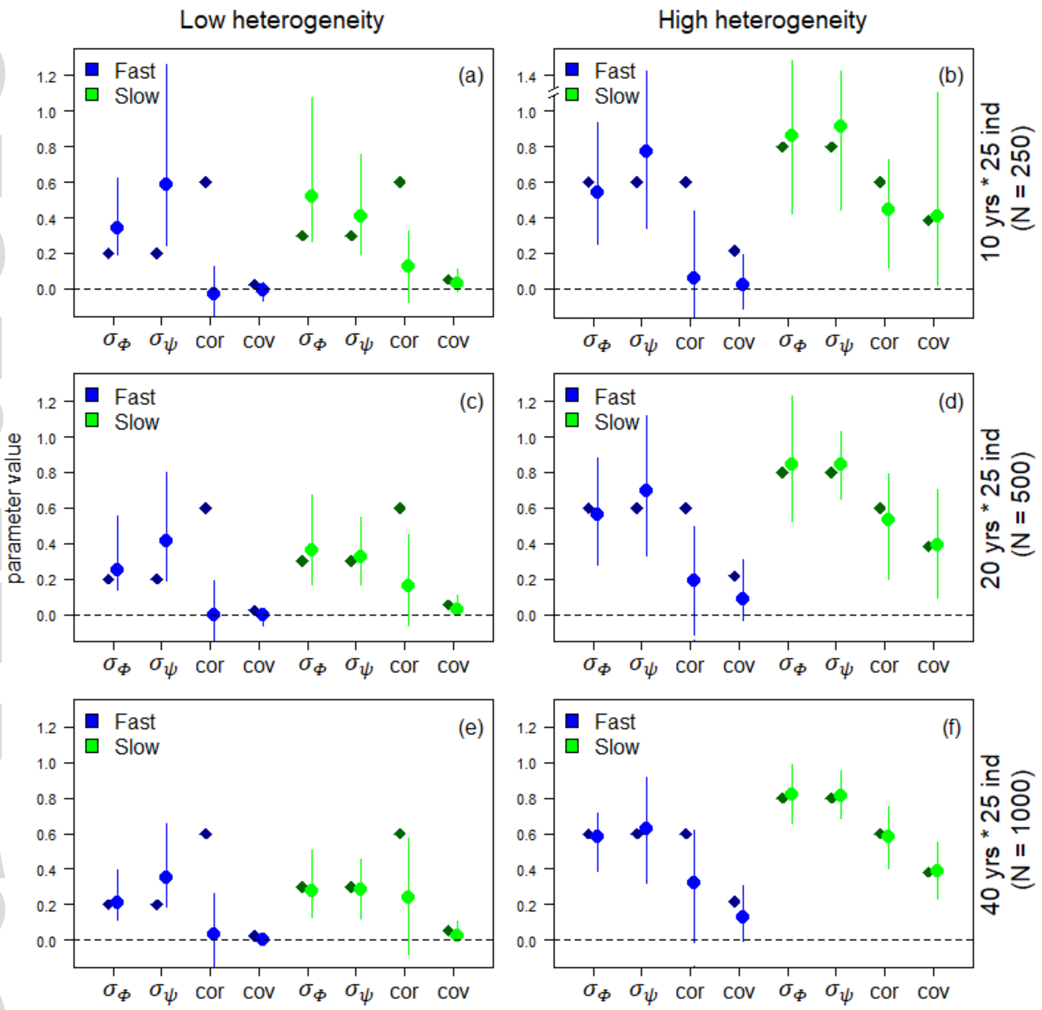
Figure 4: Bias and precision in the estimates of individual random effects and their correlation for Bernoulli variables with the full model, i.e., including temporal variation in demographic parameters and state-dependence. Parameters include the standard deviation of the individual heterogeneity in survival (σ_{Φ}) and reproductive success (σ_{ψ}) and their correlation (cor) and covariance (cov). Each plot displays the estimates for a scenario with a specific amount of individual heterogeneity (i.e., low vs. high) and study duration (10, 20 or 40 years). The number of new individuals recruited as first-time breeders per year was 100, leading to the sample sizes of

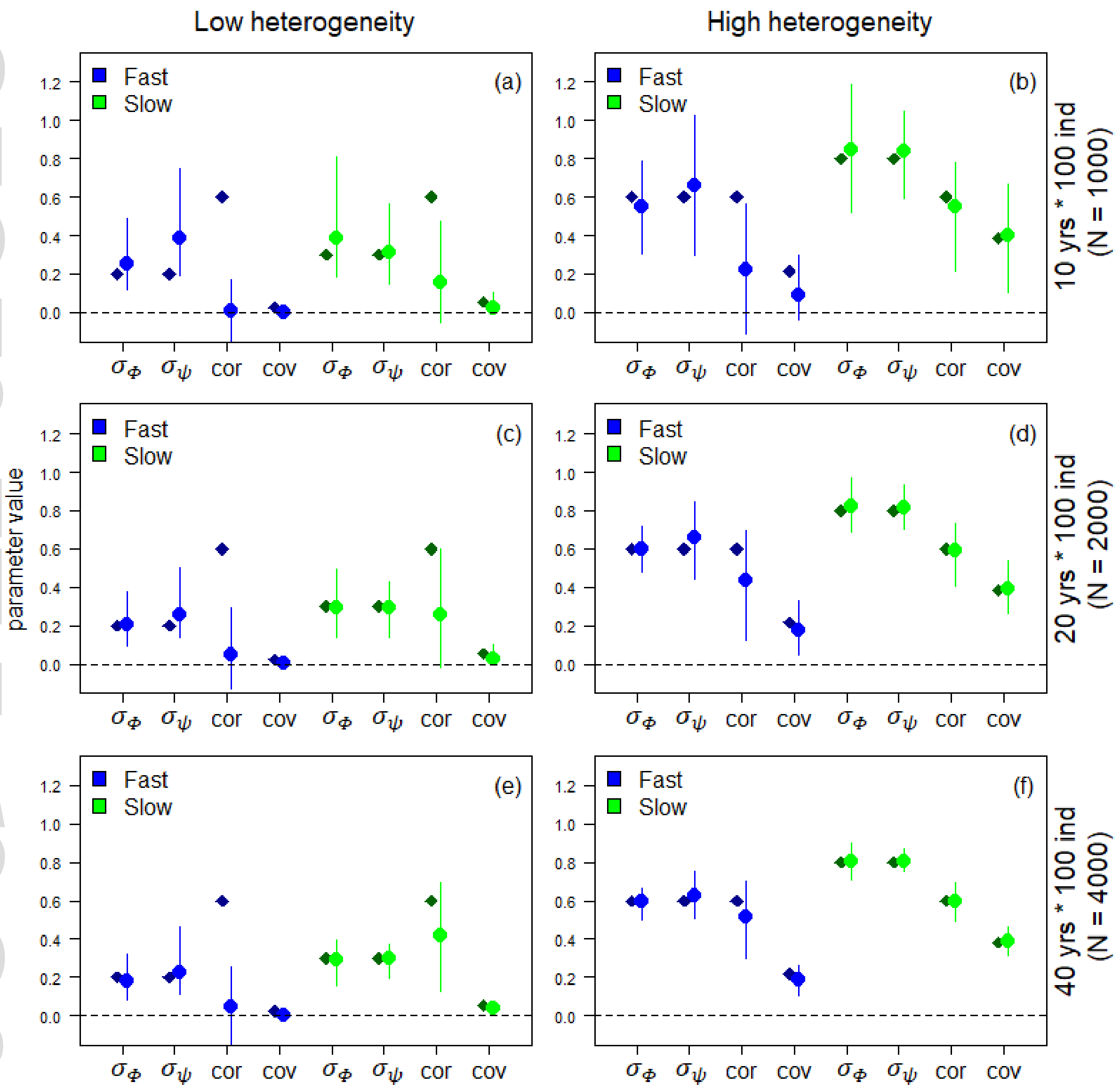
1,000, 2,000 and 4,000 individuals according to the study duration. Diamonds in darker color give the values used to simulate the datasets and points in lighter color give the average estimates over the 100 models fitted to the 100 simulated datasets. Error bars give the range including 95% of the estimated values.


FAST STRATEGY
Low heterogeneity
High heterogeneity
σ_ϕ = 0.2
σ_ϕ = 0.6
σ_ψ = 0.2
σ_ψ = 0.6
SLOW STRATEGY
Low heterogeneity
High heterogeneity
σ_ϕ = 0.3
σ_ϕ = 0.8
σ_ψ = 0.3
σ_ψ = 0.8
Density
Survival probability
Reproductive success probability
0.0
0.2
0.4
0.6
0.8
1.0

mee3_13728_f2.docx




Low heterogeneity
High heterogeneity
parameter value
Fast
Slow
(a)
(b)
(c)
(d)
(e)
(f)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
σ_Φ
σ_ψ
cor
cov
10 yrs * 25 ind (N = 250)
20 yrs * 25 ind (N = 500)
40 yrs * 25 ind (N = 1000)


Low heterogeneity
High heterogeneity
Fast
Slow
(a)
(b)
(c)
(d)
(e)
(f)
parameter value
1.2
1.0
0.8
0.6
0.4
0.2
0.0
σ_{Φ}
σ_{ψ}
cor
cov
10 yrs * 100 ind (N = 1000)
20 yrs * 100 ind (N = 2000)
40 yrs * 100 ind (N = 4000)