Check for
updates

Consequences of violating assumptions of integrated population models on parameter estimates

Floriane Plard[1,2] · **Daniel Turek[3]** · **Michael Schaub[2]**

Abstract

While ecologists know that models require assumptions, the consequences of their violation become vague as model complexity increases. Integrated population models (IPMs) combine several datasets to inform a population model and to estimate survival and reproduction parameters jointly with higher precision than is possible using independent models. However, accuracy actually depends on an adequate fit of the model to datasets. We first investigated bias of parameters obtained from integrated population models when specific assumptions are violated. For instance, a model may assume that all females reproduce although there are non-breeding females in the population. Our second goal was to identify which diagnostic tests are sensitive to detect violations of the assumptions of IPMs. We simulated data mimicking a short- and a long-lived species under five scenarios in which a specific assumption is violated. For each simulated scenario, we fitted an IPM that violates the assumption (simple IPM) and an IPM that does not violate each specific assumption. We estimated bias and uncertainty of parameters and performed seven diagnostic tests to assess the fit of the models to the data. Our results show that the simple IPM was quite robust to violation of many assumptions and only resulted in small bias of the parameter estimates. Yet, the applied diagnostic tests were not sensitive to detect such small bias. The violation of some assumptions such as the absence of immigrants resulted in larger bias to which diagnostic tests were more sensitive. The parameters informed by the least amount of data were the most biased in all scenarios. We provide guidelines to identify misspecified models and to diagnose the assumption being violated. Simple models should often be sufficient to describe simple population dynamics, and when data are

✉ Floriane Plard
    floriane.plard@ens-lyon.fr

1   Department of Aquaculture and Fish Biology, Hólar University, Háeyri 1, 550, Sauðrkrókur, Iceland

2   Swiss Ornithological Institute, 6204 Sempach, Switzerland

3   Department of Mathematics and Statistics, Williams College, Williamstown, USA

abundant, complex models accounting for specific processes will be able to shed light on specific biological questions.

1 Introduction

Integrated population models (IPMs) describe population dynamics based on the population model assumed by the modeller and the usefulness of the available data (Besbeas et al. 2002; Schaub and Abadi 2011; Schaub and Kéry in press). Thus, any incongruence between the population ecology (e.g. life cycle, mating system, individual homogeneity and independence, etc.) and the model or inadequate modelling of the protocol of data collection has the potential to bias the estimates of population parameters. This will potentially lead to wrong conclusions or management/conservation decisions (Besbeas and Morgan 2014). While ecologists know that models come with assumptions that must be reasonably satisfied, consequences of their violation become vague as the model complexity increases including multiple datasets, many parameters and dependencies between them.

The essence of IPMs is to combine several datasets to inform a population model (Besbeas et al. 2002; Schaub and Kéry in press). The datasets originate from the population level (e.g., population count data) and from the individuals level (e.g., capture-recapture data, productivity data). IPMs are widely applied in population ecology (Schaub and Abadi 2011; Zipkin et al. 2019) because they describe transient dynamics and make possible to determine the influence of different environmental factors and of the contribution of each demographic rate and of population structure to population dynamics (Koons et al. 2016, 2017). This combination of data allows the estimation of demographic parameters with higher precision. It also provides the possibility to estimate parameters that would not be estimable by the separate analysis of the given datasets (Besbeas et al. 2002; Abadi et al. 2010a). IPMs have been used to investigate density-dependent effects (Abadi et al. 2012; Gamelon et al. 2016), and to estimate parameters for which no data were available (Tavecchia et al. 2009) such as immigration (Abadi et al. 2010b). However, the biases in these hidden parameter estimates are not systematically checked (Gamelon et al. 2016), and their accuracy and interpretation are more and more questioned (Riecke et al. 2019).

IPMs are increasingly used in conservation and management studies because they result in more accurate population viability analysis (Tempel et al. 2014; Arnold et al. 2018; Saunders et al. 2018; Zipkin et al. 2018; Plard et al. 2019a; Schaub and Ullrich in press). The combination of multiple dataset allow integrating different sources of information in populations where data are rare and sparse over time and space (Schaub et al. 2007). Data deficiency is a major challenge in conservation studies and IPMs have managed to provide insightful knowledge about population demography and status (Rhodes et al. 2011; Duarte et al. 2017). However, data deficiency is also one of the first sources of uncertainty and bias. Thus, we need to better understand how

different amount of data combined in IPMs influence estimates of demographic rates (Fletcher et al. 2019; Saunders et al. 2019).

The joint likelihood of the IPM is often created as the product of the likelihood of each dataset (Besbeas et al. 2002), assuming that the different datasets are independent. In reality, the independence assumption is almost always violated. Reproductive, survival and count data often partly involve the same individuals from one population which can result in the violation of the independence assumption. Nevertheless, previous simulations (Abadi et al. 2010a; Plard et al. 2019c; Weegman et al. 2021) suggested that IPMs are robust to the overlap of individuals in different data sets. Yet, the consequences of the violations of other assumptions on which IPMs rely on have only partially been investigated (Riecke et al. 2019; Schaub and Kéry in press).

When one builds an IPM, one often makes three different series of assumptions. A first one refers to the use of parametric distributions for demographic and observation parameters (Besbeas and Morgan 2014). For instance, the distributions of clutch or litter sizes can be underdispersed as compared to the modelled one using a Poisson distribution (Kendall and Wittmann 2010). Another example is the use of the a specific distribution for modeling the sampling error (e.g., Gaussian vs. Poisson) that may bias estimates of demographic parameters and population abundance when it is not appropriate (Maunder and Piner 2017). A second series of assumptions are necessary about various forms of heterogeneity of demographic and observation parameters (Besbeas and Morgan 2014; Maunder and Piner 2017). For a given population model, demographic parameters may be assumed to be homogeneous while in reality they vary spatially or temporally. Conflicts among datasets may emerge if data have been collected at different times or locations but also by different people. Within one population, individuals may be heterogeneous, consistently showing higher or lower reproduction. When disregarded, such heterogeneity has the potential to bias predictions of population dynamics (Sæther et al. 2004; Vindenes et al. 2008; Kendall et al. 2011). Both series of assumptions are common to many statistical and demographic models and need to be carefully checked.

A third series of assumptions is more specific to IPMs and concerns the population model (Carvalho et al. 2017). In the present paper, we will study the robustness of IPM to violation of this last series of assumptions. Because an IPM combines different data sets, we need to assume that the population model corresponds to the true underlying data generating process of each data set. For instance, we may assume that all females reproduce or the absence of density-dependence while in reality there are non-breeders or density-dependent processes. Riecke et al. (2019) has already warned about consequences for parameter estimates when some assumptions are violated. They showed that estimates of hidden parameters (such as immigration or breeding probability) are biased when processes, such as mark loss, were not accounted for. However, when we build a model, how do we know that we are not missing demographic processes? How do we know that the model we used is correct and produces relatively unbiased results?

Diagnostic tests can evaluate the fit of a model to the data and identify possible conflicts between datasets (Besbeas and Morgan 2014; Carvalho et al. 2017; Schaub and Kéry in press). A first possibility is to perform goodness of fit tests on each dataset (Besbeas and Morgan 2014). They can indicate heterogeneity in reproduction, survival

and recapture, but they can be challenging to apply to complex models. Another useful check has been proposed in fisheries research to assess stock abundance: the comparisons of parameter estimates from integrated and single data models (Carvalho et al. 2017).

In this paper we pursued three objectives. Our first goal was to understand how much parameter estimates are biased when an IPM is misspecified. Suppose that we assume all females reproduce but there are non-breeding females in the population. Or, suppose that some females are double brooding and we do not include this process. What are the consequences on estimates of survival and productivity? Our second goal was to identify if some diagnostic tests have the power to identify violated assumptions. Our third goal was to study the influence of data deficiency on parameter bias and uncertainty. We investigated the impact of the violation of five assumptions linked to the population model in IPMs using simulated data mimicking a short and a long-lived species.

2 Methods

As a basis for the simulation, we considered two hypothetical populations of species with a short-lived and a long-lived life-history. The short-lived species can be seen as a passerine bird, the long-lived species as a large mammal. For both species we distinguish between two age classes: juvenile (newborn fledgling or weaning) and adult individuals. Juvenile survival (from fledgling or weaning to first-year) differs from adult survival (after first-year) in both species. First breeding occurs at age one and two years in the short- and the long-lived species, respectively. We assume that reproductive performance is invariant with age from age at first reproduction in both species.

We considered six different simulated scenarios: a null scenario and five scenarios each with a different violation of a model assumption. We first present the simulation study of the null scenario and then specify the assumption that is violated in each of the five other scenarios. Second, we present the two different models used to analyze the data for each scenario: IPM_0: simple IPM; and one of IPM_1-IPM_5: an IPM that is adjusted to the specific assumption of each scenario, hence a model that fits better than IPM_0. Third, we describe the diagnostic tests that we used to assess the fit of each model.

2.1 Generating the data with the null scenario

2.1.1 Simulation of each population

For each scenario, we simulated data from our hypothetical short-lived (parameters given with lower case letters) and long-lived (parameters given with capital letters) species. We described the female part of the population and adopted a model for a post-breeding census. All scenarios shared the following steps: in year 0, each population was composed of 300 females with ages distributed according to the stable

age distribution of the scenario. The annual number of juvenile ($\phi^j_{t+1}/\Phi^J_{t+1}$) and adult ($\phi^a_{t+1}/\Phi^A_{t+1}$) survivors from t to $t+1$ was simulated with Binomial distributions with probabilities $s^j = 0.1$ and $S^J = 0.5$ for the juvenile survival of the short- and long-lived species, respectively and $s^a = 0.4$ and $S^a = 0.9$ for the adult survival of the short- and long-lived species, respectively. For the short-lived species, for instance:

$$
\begin{aligned}
\phi^j_{t+1} &= Bin(s^j, n^j_t) \\
\phi^a_{t+1} &= Bin(s^a, n^a_t)
\end{aligned}
\tag{1}
$$

where n^j_t and n^a_t are the annual number of juvenile and adult females in year t.

Surviving females reproduce and their annual reproductive success (number of female fledglings or weaned newborn females per mother i, $w_{i,t}/W_{i,t}$) was modelled with a Poisson distribution with mean $r = 6$, and a Binomial distribution of probability $R = 0.22$, for the short- and long-lived species, respectively.

$$
\begin{aligned}
w_{i,t} &= Pois(r) \\
W_{i,t} &= Bin(R, 1)
\end{aligned}
\tag{2}
$$

The populations were simulated for 15 years and the resulting population growth rates were 1 for both species. In all scenarios and for both populations we assumed absence of emigration. Moreover, in the null scenario, we assumed absence of immigration. Because population growth rates were 1 in most scenarios, annual population sizes remained around 300 individuals with $n^a_t = 42$ adults and $n^j_t = 258$ juveniles in the short lived species and $N^a_t = 249$ adults and $N^j_t = 51$ juveniles in the long-lived species. These proportions corresponded to the stable age distributions of each population. In our main analysis, we chose to simulate a relative large number of individuals that allows accurate estimation of the demographic rates. Then, we performed a second analysis using a lower number of individuals monitored to study the influence of data deficiency on bias and uncertainty in estimates (see Sect. 2.5).

2.1.2 Simulation of the data collection in each population

Data collection was simulated to mimic realistic field sampling processes. Annually, we sampled some females from the population to obtain the capture-mark-recapture, the reproductive and the population count datasets. Binomial trials were used to select the females that belonged annually to each dataset.

The capture-mark-recapture dataset was obtained by the annual capture of newborn and adult females. The recapture probability was set to $p_s = 0.5$, corresponding to a moderate effort of capture of about 150 females annually in both species. On average, this resulted in the long-lived species in a total of 742 (SD=40) females that were captured at least once (with an average of 3 captures per female). About 15 females were captured between 8 and 13 times. For the short-lived species, an average of 2086 (SD=776) females were captured at least once in total (with an average of 1.1 capture per female) and about 34 females were captured between 3 and 6 times.

Independently of the capture of females for the capture-mark-recapture data, reproductive successes data were sampled for a proportion of 50% ($p_r = 0.5$) of females, resulting on average in 110 (SD=5) and 22 (SD=8) mothers monitored annually in the long-lived and short-lived species, respectively. The reproductive dataset included the annual numbers of monitored mothers ($m_t = Bin(p_r, n_t^a)$ and $M_t = Bin(p_r, N_t^A)$) and of their female newborns (f_t and F_t; obtained by summing the numbers of female newborns of the mothers monitored).

The population count dataset was the annual total (juvenile + adult) number of females counted at the same time when the reproductive success data were recorded. Independently of the two other datasets, a proportion of 50% ($p_c = 0.5$, about 150 females annually for both species) of the females were counted annually: $c_t = Bin(p_c, n_t^a + n_t^j)$ and $C_t = Bin(p_c, N_t^A + N_t^j)$.

2.2 Generating the data with five different scenarios

Five different scenarios commonly occurring in wild populations were simulated. Each of them violated a specific assumption of the null scenario. In scenarios 1 and 2, the manner of data collection differed from the null model. In scenarios 3, 4 and 5, the population model differed from the null model. Here, we describe how each scenario differs from the null scenario.

1. *Heterogeneity in recapture* In many monitoring programs of bird species, capture and recapture of individuals occur at or close to nests and target reproducing individuals. Consequently, only adult females that successfully reproduce (produced at least one fledgling or weaned juvenile) are captured as unsuccessful females leave breeding locations before they could be captured. This sampling protocol violates the assumption of recapture homogeneity, i.e. that all females have the same recapture probability. For adult females we used $p_s = 0.5$ if $\omega_{i,t} > 0$ and $p_s = 0$ if $\omega_{i,t} = 0$. For all newborn females, the probability of first capture at independence was kept equaled to 0.5 in both species.

2. *Heterogeneity in timing of data collection* Counts and recaptures can occur in different seasons. For instance, recaptures in mammals often occur before females lactate to limit the disruption of reproduction and counts are often performed when vegetation is less dense to maximize detection probability. In this scenario, data were collected during different seasons. Count of females occurred six months after the annual capture sessions (Fig. S9). Survival was assumed to be homogeneous within the year. The equations used to simulate the data correspond to equations used in the statistical model to estimate the demographic rates, see Eqs. (9) and (10).

3. *Immigration* influences the dynamics of most wild populations (Ward 2005; Schaub et al. 2006; Grotan et al. 2009; Millon et al. 2019). In this scenario, the populations were assumed to be open, and annually a proportion of 20% (immigration rate: $\omega = 0.2$) females enters the population as adults. Thus, as survival and reproductive success were kept at the same values the population increased in this scenario. The numbers of immigrants was: $n_t^i = Pois(\omega n_t^a)$ and $N_t^I = Pois(\omega N_t^A)$ in the short- and long-lived species, respectively.

4. *Non-breeders* Non-breeders occur in many populations. If a population model does not include existing non-breeders, the modelled dynamics can be strongly biased (Lee et al. 2017). However, non-breeders are rarely included in the reproductive dataset because these individuals are often absent from the breeding areas. In this scenario, we assume that some females do not reproduce. The number of breeding females was simulated using a Binomial distribution with the breeding probability $\psi = 0.8$. The annual number of reproductive females was $n_t^{ar} = Bin(\psi, n_t^a)$ and $N_t^{AR} = Bin(\psi, N_t^P)$ in the short- and long-lived species, respectively. N_t^P is the annual number of mature (older than 1 year old) females in the long lived species. Annual individual reproductive successes $w_{i,t}/W_{i,t}$ were simulated only for these reproductive females with a Poisson distribution of mean $r = 7.5$, and a Binomial distribution of probability $R = 0.275$, in the short- and the long-lived species, respectively, to maintain a population growth rate of 1.

5. *Density-dependence* Density-dependent effects are common in most populations and density-dependence is one of the main drivers of population dynamics (Turchin 1995). In this scenario, reproductive success is negatively influenced by the size of the total population (n or N). We used the linear model $log(r) = \alpha + \beta * n$ ($\alpha = log(6)$, $\beta = -0.5$) for the short-lived species and $logit(R) = A + B * N$ ($A = logit(0.22)$, $B = -0.5$) for the long-lived species.

2.3 Statistical models

The simulated datasets corresponding to the null scenario without any violation of assumptions were analysed only with IPM_0, while the simulated datasets corresponding to one of the five scenarios with the violation of a specific assumption were each analysed with IPM_0 and IPM_i (i = 1 to 5, corresponding to the 5 scenarios), hence with a model that does not violate the assumptions (IPM_i) and with one model that violates a specific assumption (IPM_0).

Simple IPM (IPM_0) The likelihood of the IPM is the joint likelihood of the three datasets given values of the demographic parameters. We estimated the unknown demographic rates $\tilde{s}^j$; $\tilde{S}^J$; $\tilde{s}^a$; $\tilde{S}^a$, and $\tilde{r}$; $\tilde{R}$, the unknown recapture probability $\tilde{p}_s$, and an index of population size (i.e., the latent variable $\tilde{n} \approx p_c * n$ and $\tilde{N} \approx p_c * N$ in the short- and the long-lived species, respectively). Note that in our IPMs, we cannot estimate true population size as p_c and n/N are not identifiable. The population model of this simple IPM is defined as a state-space model and corresponded to the null scenario we presented above.

– *The state process model* is identical to the following annual matrix projection model for the short-lived species:

$$\begin{bmatrix} \tilde{n}_{t+1}^j \\ \tilde{n}_{t+1}^a \end{bmatrix} = \begin{pmatrix} \tilde{s}^j\tilde{r} & \tilde{s}^a\tilde{r} \\ \tilde{s}^j & \tilde{s}^a \end{pmatrix} \begin{bmatrix} \tilde{n}_t^j \\ \tilde{n}_t^a \end{bmatrix}$$

where in year t the number of newborns (weaning/fledgling) is $\tilde{n}_t^j$, and the number of adults is $\tilde{n}_t^a$. In the annual matrix projection model for the long-lived species

females do not reproduce in their first year:

$$\begin{bmatrix} \tilde{N}^J_{t+1} \\ \tilde{N}^A_{t+1} \end{bmatrix} = \begin{pmatrix} 0 & \tilde{S}^A \tilde{R} \\ \tilde{S}^J & \tilde{S}^A \end{pmatrix} \begin{bmatrix} \tilde{N}^J_t \\ \tilde{N}^A_t \end{bmatrix}$$

All demographic parameters are assumed to be constant. Including demographic stochasticity, the population projection for the short-lived species gets:

$$\tilde{n}^a_{t+1} \sim Bin(\tilde{s}^a, \tilde{n}^a_t) + Bin(\tilde{s}^j, \tilde{n}^j_t)$$
$$\tilde{n}^j_{t+1} \sim Pois(\tilde{r} * \tilde{n}^a_{t+1}) \tag{3}$$

For the long-lived species, $\tilde{N}^Y$ is the annual number of yearling females that do not reproduce and $\tilde{N}^P$ is the annual number of mature females older than 1 year old ($\tilde{N}^A_t = \tilde{N}^P_t + \tilde{N}^Y_t$):

$$\tilde{N}^Y_{t+1} \sim Bin(\tilde{S}^J, \tilde{N}^J_t)$$
$$\tilde{N}^P_{t+1} \sim Bin(\tilde{S}^A, \tilde{N}^P_t + \tilde{N}^Y_t)$$
$$\tilde{N}^J_{t+1} \sim Bin(\tilde{R}, \tilde{N}^P_{t+1}) \tag{4}$$

– *The state-space observation model* links the true population size with the annual number of observed females (count data, c_t and C_t) using a Gaussian distribution with a standard deviation $\tilde{\sigma}$ corresponding to the observation (residual) error. For the short-lived species:

$$c_t \sim Norm(\tilde{n}^j_t + \tilde{n}^a_t, \tilde{\sigma}) \tag{5}$$

For the long-lived species:

$$C_t \sim Norm(\tilde{N}^J_t + \tilde{N}^Y_t + \tilde{N}^P_t, \tilde{\sigma}) \tag{6}$$

We also performed all analyses using the Poisson distribution (rather than the Gaussian) to investigate whether a less flexible residual error modifies the results. Because results were similar, we report here only the results for the Gaussian distribution. Results using a Poisson distribution can be found in the supplementary material (Fig. S5-S9).

We used the likelihood of a generalized linear model for the reproductive dataset. The reproductive dataset included the annual numbers of female newborns (f_t and F_t) and of monitored mothers (m_t and M_t). A Poisson and a Binomial distribution linking $\tilde{r}$ and $\tilde{R}$ to the observed annual reproductive success were used to estimate reproductive success:

$$f_t \sim Pois(\tilde{r}, m_t) \tag{7}$$
$$F_t \sim Bin(\tilde{R}, M_t) \tag{8}$$

We transformed the capture-mark-recapture dataset into the m-array format and used a multinomial likelihood to estimate survival and recapture probabilities. For a detailed explanation of the transformation of capture-mark-recapture data into the m-array format and the use of the multinomial likelihood, see e.g. Kéry and Schaub (2012) (Chap. 7.10). This model assumes a survival probability that differs between the two age classes and a constant recapture probability.

IPM$_i$ A different IPM was created for each scenario. These models were similar to the IPM_0, but not misspecified with respect to each particular scenario. We now describe how each of them differed from IPM_0.

– IPM_1: *Heterogeneity in recapture* Only females that successfully reproduced were available to be captured with probability $\tilde{p}_s$. Because recapture was restricted to reproducing females, the recapture probability was equal to $\tilde{p}_s * (1 - exp(-\tilde{r}))$ and $\tilde{p}_s * \tilde{R}$ for the short- and the long-lived species, respectively. For the short-lived species, the probability of successful reproduction is $1 - exp(-\tilde{r})$ because $exp(-\tilde{r})$ is the probability that a female had no reproductive success. For the long-lived species, $\tilde{R}$ is the probability of successful reproduction.

– IPM_2: *Heterogeneity in the timing of data collection* We assume that survival is homogenous within the year and the population size is calculated after 6 months when count data are collected. Thus, in this IPM, population projections for the short-lived species are given by:

$$
\begin{aligned}
\tilde{n}_{t+1}^a &\sim Bin(\sqrt{\tilde{s}^a}, \tilde{n}_t^{\sqrt{a}}) + Bin(\sqrt{\tilde{s}^j}, \tilde{n}_t^{\sqrt{j}}) \\
\tilde{n}_{t+1}^j &\sim Pois(\tilde{r} * \tilde{n}_{t+1}^a)) \\
\tilde{n}_{t+1}^{\sqrt{a}} &\sim Bin(\sqrt{\tilde{s}^a}, \tilde{n}_{t+1}^a) \\
\tilde{n}_{t+1}^{\sqrt{j}} &\sim Bin(\sqrt{\tilde{s}^j}, \tilde{n}_{t+1}^j)) \\
c_t &\sim Norm(\tilde{n}_t^{\sqrt{j}} + \tilde{n}_t^{\sqrt{a}}, \tilde{\sigma})
\end{aligned}
\tag{9}
$$

where $\tilde{n}^{\sqrt{j}}$ and $\tilde{n}^{\sqrt{a}}$ are the annual numbers of newborn and adult females after 6 months when the count data are collected. For the long-lived species:

$$
\begin{aligned}
\tilde{N}_{t+1}^Y &\sim Bin(\sqrt{\tilde{S}^J}, \tilde{N}_t^{\sqrt{J}}) \\
\tilde{N}_{t+1}^P &\sim Bin(\sqrt{\tilde{S}^A}, \tilde{N}_t^{\sqrt{P}} + \tilde{N}_t^{\sqrt{Y}}) \\
\tilde{N}_{t+1}^J &\sim Bin(\tilde{R}, \tilde{N}_{t+1}^A) \\
\tilde{N}_{t+1}^{\sqrt{Y}} &\sim Bin(\sqrt{\tilde{S}^A}, \tilde{N}_{t+1}^Y) \\
\tilde{N}_{t+1}^{\sqrt{P}} &\sim Bin(\sqrt{\tilde{S}^A}, \tilde{N}_{t+1}^P) \\
\tilde{N}_{t+1}^{\sqrt{J}} &\sim Bin(\sqrt{\tilde{S}^J}, \tilde{N}_{t+1}^J) \\
C_t &\sim Norm(\tilde{N}_t^{\sqrt{J}} + \tilde{N}_t^{\sqrt{Y}} + \tilde{N}_t^{\sqrt{P}}, \tilde{\sigma})
\end{aligned}
\tag{10}
$$

where $\tilde{N}^{\sqrt{Y}}$ and $\tilde{N}^{\sqrt{P}}$ are the annual numbers of yearling and mature females, respectively, when the count data are collected.

– IPM_3: *Immigration* The observation model that links the count to the true state of the population (Eqs. (5) and (6)) includes the number of immigrants ($\tilde{n}^i$; $\tilde{N}^I$) which is estimated from the hidden immigration rate $\tilde{\omega}$ (i.e., without an explicit dataset that is informative about that parameter). For the short-lived species:

$$c_t \sim Norm(\tilde{n}_t^j + \tilde{n}_t^a + \tilde{n}_t^i, \tilde{\sigma})$$
$$\tilde{n}_t^i \sim Pois(\tilde{\omega}\tilde{n}_t^a) \tag{11}$$

For the long-lived species:

$$C_t \sim Norm(\tilde{N}_t^J + \tilde{N}_t^Y + \tilde{N}_t^P + \tilde{N}_t^I, \tilde{\sigma})$$
$$\tilde{N}_t^I \sim Pois(\tilde{\omega}(\tilde{N}_t^Y + \tilde{N}_t^P)) \tag{12}$$

– IPM_4: *Non-breeders* This IPM includes a supplemental step to estimate the annual number of breeding females ($\tilde{n}^{ar}$) from the annual breeding probability $\tilde{\psi}$. $\tilde{\psi}$ is again estimated as a hidden (or latent) parameter. For the short-lived species:

$$\tilde{n}_{t+1}^{ar} \sim Bin(\tilde{\psi}, \tilde{n}_{t+1}^a)$$
$$\tilde{n}_{t+1}^j \sim Pois(\tilde{r}\tilde{n}_{t+1}^{ar}) \tag{13}$$

For the long-lived species:

$$\tilde{N}_{t+1}^{AR} \sim Bin(\tilde{\psi}, \tilde{N}_{t+1}^P)$$
$$\tilde{N}_{t+1}^J \sim Bin(\tilde{R}, \tilde{N}_{t+1}^{AR}) \tag{14}$$

– IPM_5: *Density-dependence* Reproductive success is influenced by the size of the population and we use a linear regression model for this relationship. For the short-lived species:

$$log(\tilde{r}_t) = \tilde{\alpha} + \tilde{\beta}(\tilde{n}_{t-1}^j + \tilde{n}_{t-1}^a) \tag{15}$$

For the long-lived species:

$$logit(\tilde{R}_t) = \tilde{A} + \tilde{B}(\tilde{N}_{t-1}^J + \tilde{N}_{t-1}^Y + \tilde{N}_{t-1}^P) \tag{16}$$

2.4 Diagnostic tests

For each IPM, we performed seven diagnostic tests. We first conducted five commonly applied goodness of fit tests (GOF) using Bayesian p values (Brooks et al. 2000). Second, we compared estimates of survival (juvenile and adult survival) and reproductive success from the IPMs with estimates of these parameters from single data models,

i.e. Cormack–Jolly Seber models for survival and regression models for reproductive success. To shorten the results, we only present the diagnostic tests for survival, as estimates of adult or juvenile survival were always more biased than reproductive success.

– *GOF* Following Besbeas and Morgan (2014), we used posterior predictive checks (Gelman et al. 1996) to test the fit of the models to the different datasets. The basic idea for a posterior predictive check is that for each combination of estimated parameters (demographic rates and recapture probability), a dataset is simulated [see Besbeas and Morgan (2014), Kéry and Schaub (2012), Schaub and Kéry (in press) for detailed explanation]. For each combination of estimated parameters, we can estimate a discrepancy measure between the expected values (EX_i) and observed data (OB) and a discrepancy measure between the expected values (EX_i) and the simulated dataset (OB_i). When a model fits the data, one can expect that half of the points of a scatter plot of the two discrepancy measures to lie above the line of unit slope. The proportion of points above the line of unit slope is the Bayesian p value.
Following Besbeas and Morgan (2014) we used the Freeman–Tukey statistic ($D_{FT} = \sum(\sqrt{OB_i} - \sqrt{EX_i})^2$) for the capture-mark-recapture and the reproductive datasets as discrepancy measures. We also used an alternative Pearson chi-squared test for the reproductive dataset. The Pearson chi-squared could not be used for the capture-mark-recapture datasets because they include a large number of zeros when no individuals were recaptured. For the population count data, we used two discrepancy measures: the mean absolute percentage error ($D_{MAPE} = \frac{100}{T} \sum_{t=1}^{T} \left| \frac{OB_{i,t} - EX_{i,t}}{OB_{i,t}} \right|$, where T is the number of non-missing predictions), and the maximum percentage error ($D_{MPE} = 100 max \frac{OB_{i,t} - EX_{i,t}}{OB_{i,t}}$) (Besbeas and Morgan 2014). The population count dataset (c; C) corresponded to the observed data OB, while $\tilde{n}/\tilde{N}$ were the expected values EX. Then, we used the distribution of demographic rates to simulate a new $\tilde{N}_i$ using the state process model of each IPM and then we simulated OB_i with the observation model. For these five goodness of fit tests we treated a model as misspecified when the Bayesian p-value was above 0.975 or below 0.025.

– *Single data models* Conflicts among different data are inferred by comparing the demographic parameters estimated from an IPM to those estimated from a model analysing the capture-mark-recapture dataset and the reproductive dataset separately. In other words, we compared the estimated parameters obtained from the IPMs to those obtained from models excluding the count data. We transformed the capture-mark-recapture data into an m-array and used a multinomial likelihood that had the same structure as the corresponding model component in IPM_0. The reproductive datasets were analysed with generalized linear models identical to the model components for reproduction in IPM_0. Because the largest bias was always observed in the survival parameters, we performed two diagnostic tests using juvenile and adult survival parameters. We considered a model as misspecified when less than 75% (increasing this value did not change our conclusions) of the posterior distribution of each survival parameter estimated from the IPM

overlapped with the posterior distribution of each survival parameter estimated from the single data model.

2.5 Analysis

We performed 1000 simulations over 15 years under each of these six scenarios for both species. We sampled individuals to yield capture-mark-recapture, reproduction, and count datasets. To assess the performances of IPM_0 and IPM_{1-5}, we first compared the bias ($posterior\ mean - truth$) and precision (mean squared errors $MSE = bias^2 + variance$) of comparable parameters: mean recapture probability, mean juvenile and adult survival, and mean reproductive success. For the scenario 5 including density-dependent effect, we could not compare annually variable reproductive successes estimated from IPM_5 with the constant reproductive success estimated from IPM_0. Second, to investigate the influence of small datasets, we resampled 20% of the simulated sampled females of the reproductive and capture-mark-recapture datasets in each population. These data were analyzed again with the two corresponding IPMs as for the original dataset using the following combinations. (i) We included 20% of the capture-mark-recapture data and all the count and reproductive data. This would correspond to set $p_s = 0.1$ while $p_r = p_c = 0.5$. (ii), we included 20% of the reproductive data and all count and capture-recapture data. This would correspond to set $p_r = 0.1$ while $p_s = p_c = 0.5$. Third, we performed the 7 diagnostic tests on all models. In order to understand how the different scenarios influenced the diagnostic tests, we report the number of simulations when a model was evaluated as misspecified for each scenario.

We fit all models in the Bayesian framework and defined the following prior distributions. We used uniform distributions over the interval [0,1] as priors for all probabilities and uniform distributions over the interval [0,10] as priors for reproductive success of the short-lived species. We used uniform distributions over the interval $[0,C_1]$ where C_1 is the number of females counted in the first year as priors for the number of females of each age class in the first year ($\tilde{n}_1^a$, $\tilde{n}_1^a$, $\tilde{N}_1^J$, $\tilde{N}_1^Y$, $\tilde{N}_1^P$). For the scenario including a density-dependent effect, we used the Gaussian distribution with mean 0 and standard deviation 100 as priors for α, β, A, and B (Kéry and Schaub 2012). We generated three chains of length 25,000, discarding the first 12,500 samples as burn-in. Convergence of all chains was assessed using the Gelman and Rubin convergence diagnostic (R<1.1, Gelman and Rubin 1992). We fit the models using NIMBLE (de Valpine et al. 2017) run from R (R Core Team 2019).

3 Results

3.1 Accuracy and precision of the IPM under different scenarios

In the null scenario, the simple IPM_0 gave unbiased demographic parameters (Figs. 1, 2, left panels) with higher precision (Figs. 3, 4, left panels) compared to single data models. The higher precision of parameters obtained from IPMs compared to
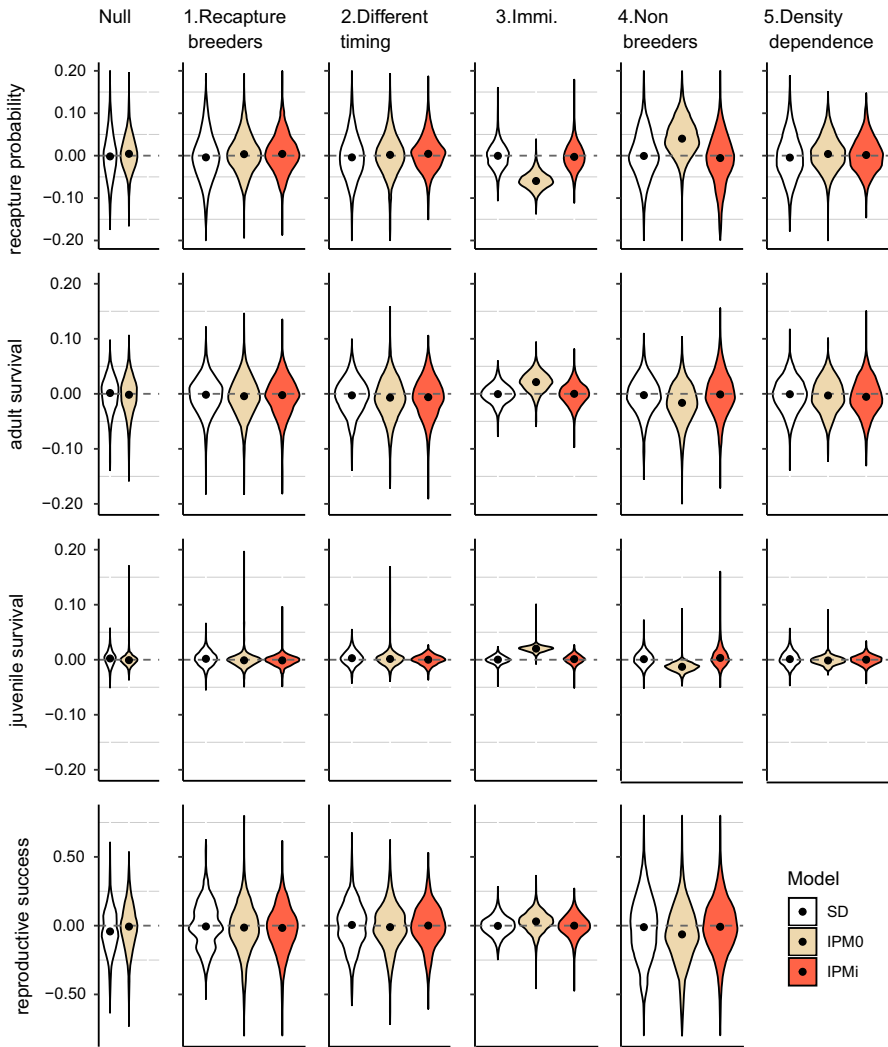
Fig. 1 Comparison of bias for estimates of recapture probabilities, juvenile and adult survival, and reproductive success in a short-lived species. 6 scenarios are shown: null scenario, 1: Heterogeneity in recapture when only breeders are recaptured 2: Heterogeneity in timing of the collection of the three datasets. 3: Immigration influences population dynamics. 4: Non-breeders influence population dynamics: breeding probability < 1. 5: Density-dependent effect on reproductive success. Parameters were estimated under single data CMR and reproduction models (SD, white), simple IPM_0 (beige) and IPM_i (red). Violin plots show the distributions of mean bias over 1000 simulations. The median of each distribution is shown with a black point
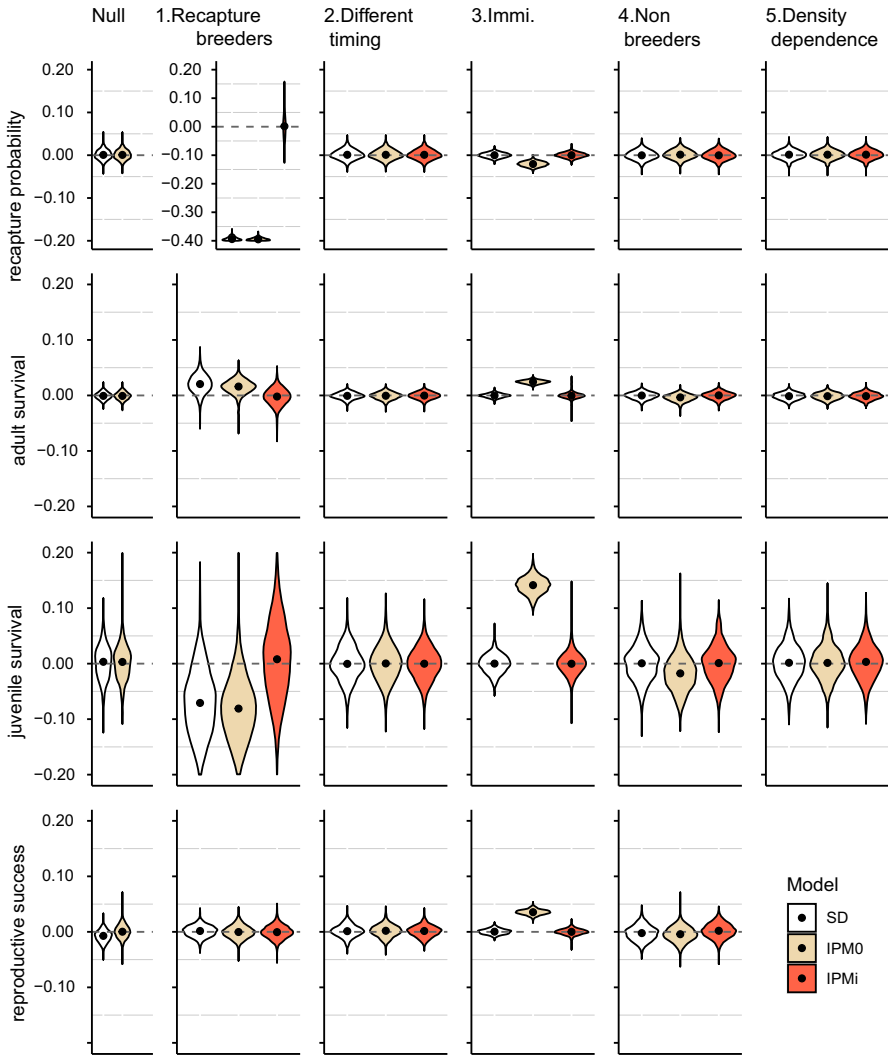
Fig. 2 Comparison of bias for estimates of recapture probabilities, juvenile and adult survival, and reproductive success in a long-lived species. 6 scenarios are shown: null scenario, 1: Heterogeneity in recapture when only breeders are recaptured 2: Heterogeneity in timing of the collection of the three datasets. 3: Immigration influences population dynamics. 4: Non-breeders influence population dynamics: breeding probability < 1. 5: Density-dependent effect on reproductive success. Parameters were estimated under single data CMR and reproduction models (SD, white), simple IPM_0 (beige) and IPM_i (red). Violin plots show the distributions of mean bias over 1000 simulations. The median of each distribution is shown with a black point
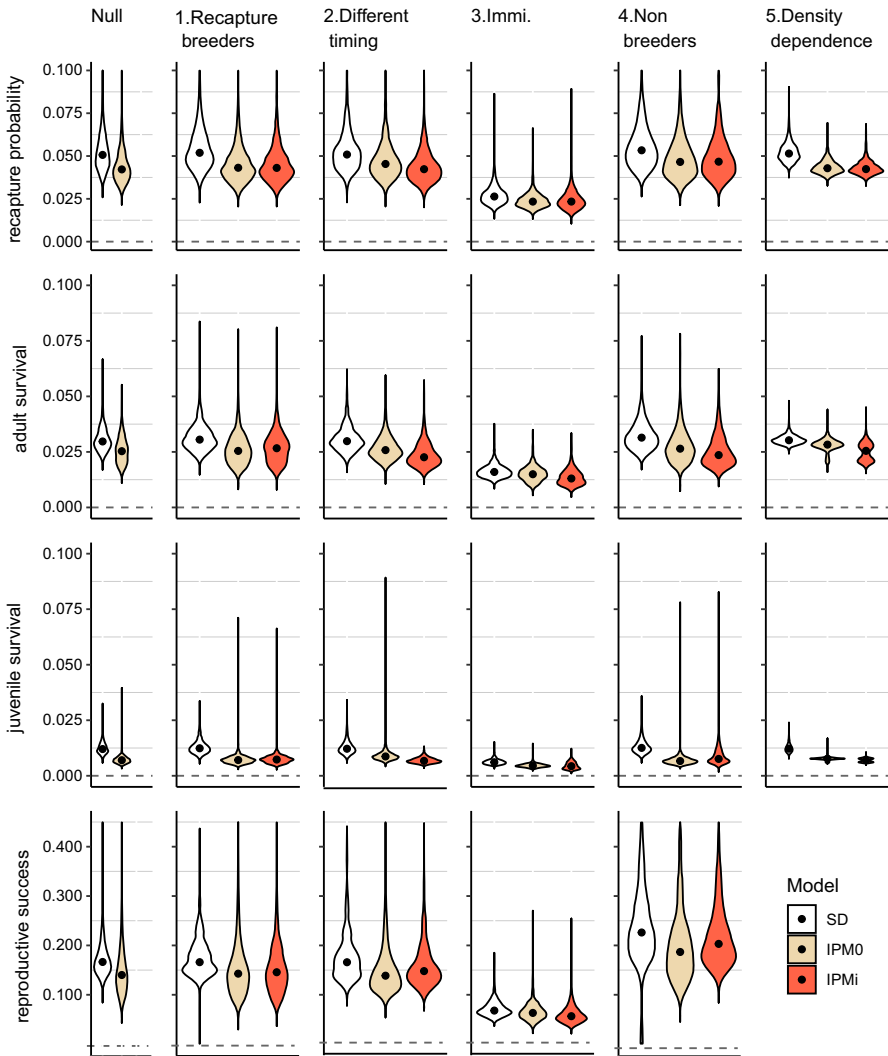
Fig. 3 Comparison of mean square errors (MSE) for estimates of recapture probabilities, juvenile and adult survival, and reproductive success in a short-lived species. 6 scenarios are shown: null scenario, 1: Heterogeneity in recapture when only breeders are recaptured 2: Heterogeneity in timing of the collection of the three datasets. 3: Immigration influences population dynamics. 4: Non-breeders influence population dynamics: breeding probability < 1. 5: Density-dependent effect on reproductive success. Parameters were estimated under single data CMR and reproduction models (SD, white), simple IPM_0 (beige) and IPM_i (red). Violin plots show the distributions of mean MSE over 1000 simulations. The median of each distribution is shown with a black point
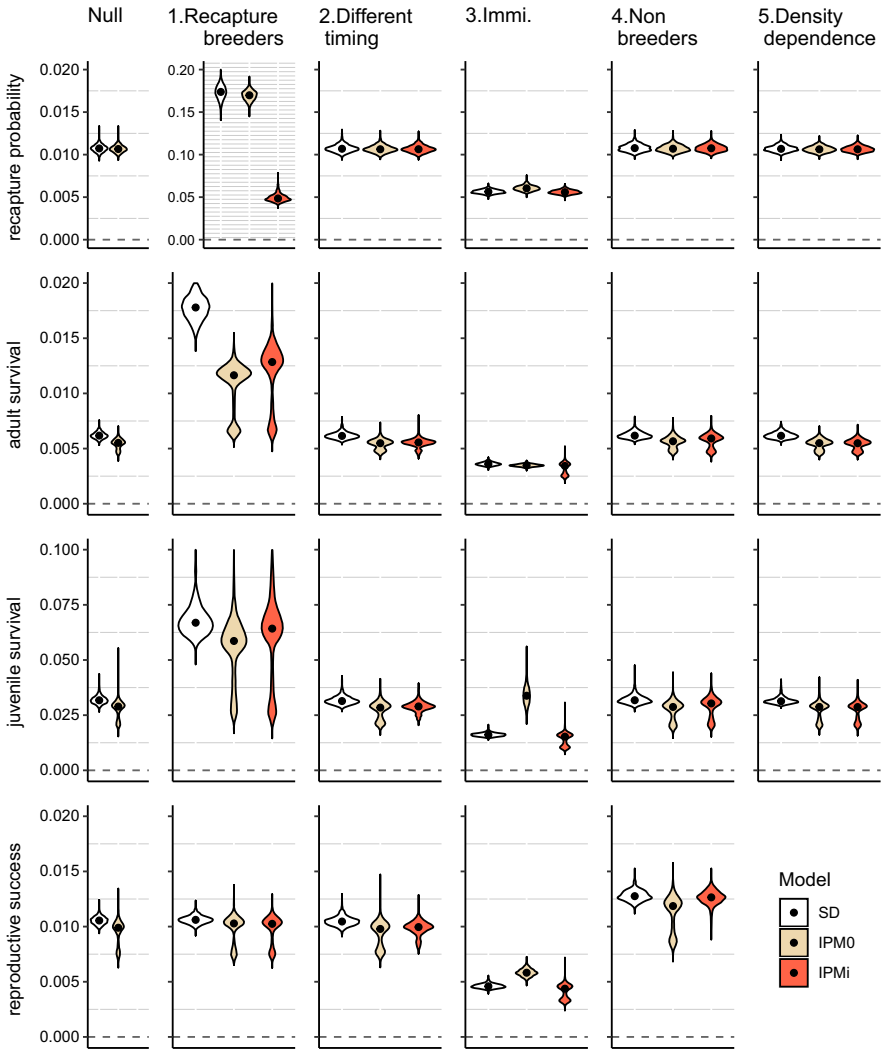
Fig. 4 Comparison of mean square errors (MSE) for estimates of recapture probabilities, juvenile and adult survival, and reproductive success in a long-lived species. 6 scenarios are shown: null scenario, 1: Heterogeneity in recapture when only breeders are recaptured 2: Heterogeneity in timing of the collection of the three datasets. 3: Immigration influences population dynamics. 4: Non-breeders influence population dynamics: breeding probability < 1. 5: Density-dependent effect on reproductive success. Parameters were estimated under single data CMR and reproduction models (SD, white), simple IPM_0 (beige) and IPM_i (red). Violin plots show the distributions of mean MSE over 1000 simulations. The median of each distribution is shown with a black point

single data models was verified for all scenarios. Recapture and adult survival were estimated with higher accuracy in the long-lived compared to the short-lived species while juvenile survival was estimated with higher accuracy and precision in the short-lived than in the long-lived species.

1. *Heterogeneity in recapture* In the scenario where only breeders were recaptured, the simple IPM_0 resulted in unbiased estimates of demographic parameters in the short-lived species (Fig. 1, second column). Indeed the assumption of recapture homogeneity was very weakly violated in this species. The recapture probability was only slightly different from 0.5 for all females ($p_S \approx 0.49875$) because, most females successfully reproduced (Probability of successful reproduction was $1 - exp(-6) \approx 0.9975$). By contrast, the probability of successful reproduction for females older than 1 year was 0.22 in the long-lived species. Thus, the recapture probability was highly heterogeneous in this species as 22% of the mature females had a recapture probability of 0.5 while 78% had a recapture probability of 0 (mean recapture probability ≈ 0.11). Yearling females also had a recapture probability of 0 because they did not reproduce. A direct consequence was that both the single data model and the simple IPM_0 under-estimated the recapture probability (Fig. 2, second column, $- 0.40[- 0.42:- 0.38]$ and $-0.40[-0.42:-0.39]$, respectively) (here and below mean and 95% interval of absolute bias over 1000 simulations are given). The violation of the homogeneous recapture assumption in the single data model and in the IPM_0 also resulted in biased estimates of survival parameters: juvenile survival was under-estimated ($-0.07[-0.17:0.06]$ and $-0.08[-0.19:0.07]$) and adult survival was over-estimated ($0.02[-0.01:0.05]$ and $0.02[-0.01:0.04]$). The IPM_1 resulted in unbiased estimates of adult survival parameters (bias lower than 0.01, on average). However, IPM_1 erroneously assumed that the recapture probability of yearling was $\tilde{p}_S * \tilde{R} \approx 0.11$ while in reality it was 0. The assumption of homogeneity in recapture was thus weakly violated in IPM_1 which may explain why the distribution of bias in juvenile survival was very wide compared to other scenarios (Fig. 2, second column). The estimates of the three main demographic parameters from IPM_1 were not more precise than the estimates from the simple IPM_0 (Figs. 3, 4, second column).
2. *Heterogeneity in timing of data collection* Estimates from the single data model, IPM_0, and IPM_2 were similarly accurate regardless of whether count data were collected at the same time as the two other datasets or at a different time (all biases lower than 0.002, Figs. 1, 2, third column). Precision was slightly better from IPM_2 than from IPM_0 (Figs. 3, 4, third column).
3. *Immigration* Ignoring immigration led to the strongest bias in parameter estimates of all scenarios considered here (Figs. 1, 2, fourth column). IPM_0 translated the assumed absence of immigration into a lower recapture probability and higher survival and reproductive success. For the short-lived species, the recapture probability ($-0.06[-0.09:-0.02]$) was the most biased parameter, followed by adult survival ($0.02[-0.01:0.05]$), juvenile survival ($0.02[0.01:0.03]$) and reproductive success ($0.03[-0.12:0.16]$). For the long-lived species, juvenile survival showed the highest bias ($0.1 [0.11:0.16]$). Adult survival ($0.02[0.02:0.03]$), reproductive success ($0.04[0.03:0.05]$) and recapture probability ($-0.02 [-0.03:-0.01]$) were

all similarly biased. Both the single data model and the IPM_3 including immigration resulted in accurate (bias lower than 0.002 for all parameters) and more precise (Figs. 3, 4, fourth column) parameter estimates than IPM_0. The single-data models showed no bias because their underlying assumptions were met. Single data model used only the capture-mark-recapture data or the reproductive data and did not use count data. Thus, contrary to IPM_0, the single data model did not use the information of the count data that the population was increasing at a higher rate than expected by the survival and reproductive rates. Using IPM_3, the estimate of the immigration rate had large uncertainty, and was slightly biased, the latter being larger in the short- than the long-lived species (bias in $\tilde{\omega}$: $-0.01[-0.12:0.15]$ and $0.001[-0.002:0.002]$ for the short- and the long-lived species, respectively).

4. *Non breeders* In the presence of non-breeders IPM_0 under-estimated all demographic parameters with the highest bias appearing in adult survival ($-0.02[-0.10:0.04]$) and reproductive success ($-0.07[-0.60:0.34]$) in the short-lived species and in juvenile survival ($-0.02[-0.07:0.05]$) in the long-lived species (Figs. 1, 2, fifth column). Bias occurred because of a conflict among datasets. Information about reproductive success $\tilde{r}$ corresponded to r when originating from the reproductive dataset because only reproducing individuals were included, while it corresponded to $\psi * r$ when the information originated from the count data set. The resulting estimate was somewhere between these two values. To compensate the bias in reproductive success $\tilde{r}$ and to achieve a close fit of the estimated population size with the observed counts, $\tilde{s}^j$, $\tilde{s}^a$, and $\tilde{p}_S$ were also biased in IPM_0. By contrast, both the single data models and the IPM_4 provided accurate estimates of demographic parameters (bias lower than 0.002). Breeding probability estimated from IPM_4 was more accurate for the long- than the short-lived species (bias in $\tilde{\psi}$: $-0.06[-0.35:0.14]$ and $-0.003[-0.12:0.14]$ for the short- and the long-lived species, respectively). IPM_4 estimated the parameters with higher precision than the single data models but not than IPM_0 (Figs. 3, 4, fifth column).

5. *Density-dependence* The simple IPM_0 estimated constant reproductive success and therefore was unable to properly estimate the density-dependent effect on reproductive success but it resulted in unbiased estimates of average survival parameters for both species (Figs. 1, 2, last column). The IPM_5 estimated the regression parameters of annual population density on reproductive success with relatively high bias ($\tilde{\beta}$: $0.03[-0.08 : 0.32]$ and $\tilde{B}$: $-0.38[-1.40 : 0.30]$) and uncertainty ($\tilde{\beta}$: $0.07[0.04 : 0.16]$ and $\tilde{B}$: $0.89[0.53 : 2.71]$), particularly for the long-lived species. Indeed, because the population size varied more in the short-lived than in the long-lived species due to demographic stochasticity (n = 350, temporal variability: SD = 14, and N = 298, SD = 2 in this scenario), the density-dependence parameter was more difficult to estimate in the more stable long-lived species.

3.2 Sparse datasets

If only 20% of the capture-recapture data were included, the precision in all demographic parameters for all models IPM_0 and IPM_{1-5} was reduced (Figs. S1–S4) and the magnitude of bias changed for some parameters. The bias from IPM_0 in

Table 1 Percentage of models (from 1000 simulations) identified as misspecified by each diagnostic test under different scenarios. See the main text for a full description of each scenario

| | | short-lived species | | | | | | | long-lived species | | | | | | |
| | | GOF | | | | | Compar. | | GOF | | | | | Compar. | |
		CMR_FT	Rep-FT	Rep-Chi²	MAPE	MPE	ad.surv	juv.surv	CMR_FT	Rep-FT	Rep-Chi²	MAPE	MPE	ad.surv	juv.surv
	Null	4	4	4	5	5	2	2	4	3	3	4	4	1	2
	1. Recapture breeders	4	4	5	5	5	3	3	20	3	3	4	5	5	2
IPM 0	2. Different timing	6	7	7	22	24	3	5	4	3	3	6	6	1	1
	3. Immigration	14	3	3	1	1	52	93	100	100	100	100	0	100	100
	4. Non-breeders	4	4	5	5	5	4	27	5	3	3	5	6	4	5
	5. Density-dependence	4	99	99	1	1	0	1	4	5	6	6	6	1	2
IPM 1	1. Recapture breeders	4	4	4	4	4	2	2	100	3	3	4	4	41	42
IPM 2	2. Different timing	4	3	3	23	41	4	2	4	3	3	2	83	0	0
IPM 3	3. Immigration	4	3	3	4	4	4	5	6	4	4	10	10	6	7
IPM 4	4. Non-breeders	4	3	4	6	28	4	5	5	3	2	8	8	3	2
IPM 5	5. Density-dependence	4	99	99	5	50	2	1	5	10	10	4	100	0	1

survival and recapture parameters declined in scenario 1 (heterogeneity in recapture). For the scenarios 3 and 4 (immigration or non-breeders), bias from IPM_0 increased for juvenile and adult survival but decreased for reproductive success. Surprisingly, we found that fitting the "true" IPM_{1-5} resulted in higher bias than the simple IPM_0 in some scenarios because the datasets were not large enough to inform all parameters estimated. This was true for the estimates of juvenile survival in the long-lived species in scenario 1 (heterogeneity in recapture; $-0.038[-0.17:0.12]$ and $0.023[-0.13:0.27]$ in IPM_1 and IPM_0, respectively), scenario 4 (non-breeders; $0.066[-0.13:0.28]$ and $-0.011[-0.07:0.05]$ in IPM_4 and IPM_0, respectively), and scenario 5 (density-dependence; $0.013[-0.12:0.19]$ and $0.008[-0.12:0.19]$ in IPM_5 and IPM_0, respectively).

Including only 20% of the reproductive data resulted in lower precision in reproductive parameters for all models IPM_0 and IPM_{1-5} (Figs. S1–S4). The bias was larger in estimates of reproductive success in scenarios 3 and 4 (immigration and non-breeders) for both species.

3.3 Diagnostic tests

We classified a test to be useful if it correctly recognized a model as misspecified in more than 95% of the simulations. Generally it appeared that the applied tests were little sensitive to the evaluated model violations (Table 1).

GOF tests The maximum percentage error (MPE) tests were very sensitive to uncertainty in demographic parameters. This test did not recognize IPM_0 as misspecified in any scenario but targeted IPM_5 as misspecified when it was not. Other GOF tests recognized IPM_0 as misspecified only for the immigration scenario for the long-lived species. The GOF test for capture-mark-recapture data identified IPM_1 as misspecified when it was not for the long-lived species. The GOF tests for the reproductive data

recognized both IPM_5 and IPM_0 as misspecified in scenario 5 for the short-lived species.

Comparison tests with single data models recognized IPM_0 as misspecified only for the immigration scenario of the long-lived species. They did not erroneously recognize any model as misspecified when it was not.

4 Discussion

Our results show that simple IPMs were quite robust to the violation of most but not of all assumptions that we evaluated. The use of an IPM that corresponded exactly to the data generating model improved the estimates compared to the simple (wrong) IPM often little, the notable exception being when immigration occurred. Unfortunately, the evaluated diagnostic tests performed similarly and were not sensitive to detect small bias and thus could not identify misspecified IPMs that produced only small bias in parameter estimates. Nevertheless, violation of assumptions resulting in large bias such as when an IPM wrongly assumes absence of immigration were correctly identified by most diagnostic tests.

Among demographic parameters, the parameter with the largest bias was always the parameter that was informed by the least amount of data, regardless of which assumption was violated. However, the magnitude of bias depended both on the type of assumption being violated and on the life-history of the studied species. For a long-lived species, the scenario including immigration and a dependency between recapture and reproductive success resulted in the largest bias when analysed with a simple IPM_0. For the short-lived species the scenarios including immigration or non-breeders resulted in the largest bias when analysed with a simple IPM_0.

Last but not least, our results show that complex models, even if correctly specified can result in biased parameters if data are sparse (when only 20% of the CMR data are used).

4.1 Generality and limits of our results

To maximize the generality of our results, we included in our simulations two different life-histories and simulated data for 15 years, corresponding to a typical duration of IPM studies, which is often between 10 and years (20 years: Tenan et al. 2017, 16 years: Plard et al. 2020, 15 years: Lieury et al. 2015; Hatter et al. 2017; Fay et al. 2019, 14 years: Duarte et al. 2016, 12 years: Brommer et al. 2017, 11 years: Cleasby et al. 2017), even if some studies last longer (22 years: Tempel et al. 2014, 30 years: Margalida et al. 2020) or shorter (7 years, Duarte et al. 2017). We chose to simulate a relative large number of individuals (300 individuals) compared to population sizes of empirical IPMs which was often between 20 and 300 individuals (in the articles cited above). The simulated sample sizes were large enough to estimate correctly most demographic rates, as shown by our results. Nevertheless, we performed a second analysis with a lower number of individuals to study the influence of sample size on

bias and uncertainty in estimates. We show that bias and uncertainty always increase with declining amount of data, even in correctly specified models.

The scenario including heterogeneity in recapture showed that when the assumption of homogeneity of recapture probability is violated, survival estimates are biased. This is in accordance with results of previous studies (Carothers 1973; Devineau et al. 2006; Fletcher et al. 2012; Abadi et al. 2013). These studies have shown that the bias increases with increasing heterogeneity in recapture probability and when the average recapture probability decreases (Devineau et al. 2006; Fletcher et al. 2012).

In the scenario where the count dataset were collected six months later than the reproductive and the capture-mark-recapture datasets, IPM_0 produced accurate estimates of reproductive success and of survival probabilities. Because we assumed constant survival probabilities before and after the collection of the count data, the variation in the number of counted females was proportional to the variation in the true population size at the time of recapture, which is why there was no bias in the parameter estimates. However, bias in demographic parameters is expected under this scenario if survival varies within the year (Gauthier et al. 2001; Rockwell et al. 2017; Robinson et al. 2020).

The scenario including immigration showed that immigration will result in overestimation of all demographic rates if not accounted for (Abadi et al. 2010b; Schaub and Fletcher 2015). This bias will increase as immigration rate increases. Bias in demographic parameters were higher in the long-lived species characterized by a higher proportion of adults and thus of immigrants, than in the short-lived species characterized by a higher proportion of juveniles. Moreover, the bias in parameter estimates is expected to change if immigrants have different survival and reproductive success than residents (Grist et al. 2017; Rolandsen et al. 2017; Barbraud and Delord 2021).

In the scenario including non-breeders, IPM_0 estimates of survival and reproductive success were only weakly biased because the breeding probability was high (0.8) in our simulations. If individuals have a low breeding probability researchers generally know that there are non-breeders in their study populations and therefore it is unlikely that an IPM is misspecified with respect to non-breeders. We have chosen a relatively high breeding probability for our simulations, hence considered a realistic scenario where a researcher is not aware of non-breeders and therefore is likely to misspecify an IPM. However, the higher the proportion of non-breeders in the population is, the higher the resulting bias in estimated demographic rates becomes if non-breeders are not explicitly included in the population model (Lee et al. 2017). The omission of non-breeders biases all demographic rates as shown by our results using IPM_0 because the presence of non-breeders creates a conflict between the predicted population size based on the wrong population model and the count data.

In the scenario including density-dependence, survival parameters were only weakly biased. Indeed, IPM_0 produced accurate estimates of *average* reproductive success and therefore *average* survival probabilities remained unbiased. Nevertheless, density-dependence has a higher potential to bias the estimated parameters if the population size fluctuates more than in our simulations. Then, however, density-dependence will also be easier to detect and to estimate.

4.2 Bias of simple *IPM*$_0$ and model complexity

IPMs are now widely and increasingly used in ecology and conservation to understand the mechanisms that drive population dynamics (Schaub and Abadi 2011; Abadi et al. 2017; Bled et al. 2017; Koons et al. 2017; Arnold et al. 2018). They have been used to make predictions about the future dynamics of populations (Schaub and Abadi 2011; Zipkin et al. 2019). Such predictions from IPMs are particularly valuable as they correctly include the uncertainties due to parameter estimation (Saunders et al. 2018; Schaub and Kèry in press). Yet, it is crucial to know when an IPM gives biased estimates to avoid wrong or misleading inference.

Our simulation results show that absolute bias in survival parameters obtained from the simple IPM_0 was below 0.05 on average for most scenarios. Nevertheless, our scenarios are constructed to reflect the likely magnitudes of violation. If they were larger (e.g., more immigrants or higher fraction of non-breeders), biases are expected to increase. For many demographic analyses, the simple IPM_0 is expected to produce accurate and robust results and hence the uncertainty is properly accounted for when used for population projections, for example. The inaccuracy of these estimates increases the fewer data are available, but this is true for all statistical models. However, when IPM_0 is fitted, one cannot study specific demographic processes such as density-dependence that may drive population dynamics. If such a process occurs and is not included in an IPM, the projection of *transient* population dynamics may be wrong (Hixon et al. 2002; Turchin 1995). More complex models are thus needed and useful to answer specific questions such as how strong is the influence of individual, environmental and population factors on population dynamics (Benton et al. 2006; Evans et al. 2013; Barraquand and Gimenez 2019; Plard et al. 2019b).

Nevertheless, more complex models do not provide estimates of demographic parameters with systematically higher accuracy. Moreover, when datasets are sparse, our results showed that for some scenarios such as the density-dependent scenario, complex models even if correctly specified gave estimates with larger bias than the simple IPM. For this particular scenario, the estimates of density-dependent parameters were highly imprecise, particularly for the long-lived species. Thus, increasing the complexity of IPMs needs substantial amount of data.

Finally, a complex model will never, by itself, estimate with accuracy processes for which data are unavailable. When a demographic parameter is estimated that is not directly informed by a dataset (hidden parameter), it may soak up inconsistencies from other parts of the model. For example immigration can be biased when mark loss in the capture-recapture data is not modelled adequately (Riecke et al. 2019). Therefore, estimates of hidden parameters need to be interpreted with care (Schaub and Kéry in press). Robinson et al. (2014) included a hidden parameter in studies of several bird species and declared them as correction parameter that includes any important demographic processes that are not captured by other parts of the model. They assumed that unmeasured processes were mostly related to productivity and specifically is due to the proportion of breeder for which no data were available. Thus, a hidden parameter is useful to get a better fit of the model, but the labelling of this parameter is speculative to some degree.

4.3 Differences between long-lived and short-lived species

The demographic parameter that is the most biased depends on the data collected but also directly on the life-history of the species studied. In populations of long-lived species few offspring are produced, and hence there is naturally little information about juvenile survival (Gaillard et al. 2000). In populations of short-lived species, juvenile survival tended to be estimated with lower bias and higher precision because a lot of newborns are produced and inform this parameter. At the opposite, sample size for estimating adult survival is larger in long-lived species compared to that of short-lived species. If an assumption is violated, the IPM tries to achieve a compromise. The estimates are weighted averages based on the information that each dataset contributes to that parameter. Thus, the parameters that are least informed by data can more easily be pulled. As a consequence, juvenile survival and adult survival will respectively be the parameter with the highest biased in long- and short-lived species, respectively. Yet, as stated above, this differs when the model contains a hidden parameter. Then the hidden parameter is the one that is least informed by data and more easily pulled and thus potentially biased.

Our results also show that the same violated assumption can have different impact on short and long-lived species, in accordance with another study (Earl 2019). In the scenario including recapture heterogeneity for instance, IPM_0 produced biased estimates mainly for the long-lived species. In the scenario including non-breeders, IPM_0 gave biased estimates mainly for the short-lived species. This is particularly apparent when only 20% of the survival or the reproductive data are included. The population growth rate is differentially sensitive to changes in demographic rates and the sensitivities depend on the life-history of the species. Generally, population growth rates of long-lived species are more sensitive to changes in adult survival and those of short-lived species to changes in recruitment related parameters (juvenile survival, reproductive success) (Saether and Bakke 2000). Therefore, the violation of an assumption that results primarily in a bias of adult survival is expected to be worse for the predicted dynamics of a long-lived than of a short-lived species, while the inverse is true for an assumption that affects mostly juvenile survival or reproductive success.

4.4 How can we know that a model is wrong?

The application of generic goodness of fit tests is the first step to identify whether a model is wrong (Gelman et al. 1996; Brooks et al. 2000; Johnson and Omland 2004; Kéry and Schaub 2012; Besbease and Morgan 2014). However, the various tests we have evaluated were not very sensitive to small bias and the violating assumption(s) could not be identified by using them. In most cases the parameter with the largest uncertainty or that is least informed by the data will be the most biased. However, this will not necessarily tell us which assumption is violated.

If a model is suspected as being wrong, the comparison between estimates from single data models and IPM can get us started to identify the problem (our results and Riecke et al. 2019; Schaub and Kéry in press). Any difference between these

estimates might indicate model misspecification. When a demographic process adding individuals to the population is missing, such as immigration, reproductive or/and survival parameters will be higher compared to the estimates we get from the single data models. Conversely, when a demographic process removing individuals from the population is missing, such as breeding probability lower than 1, reproductive and survival parameters will be lower compared to the estimates originating from the single data models. Second, comparing estimates from models using either all or a subset dataset can also highlight the conflicting parameter(s) (Carvalho et al. 2017). If one assumption is violated or if there is a conflict between datasets, changing the size of a dataset should result in different parameter estimates because by reducing the size of a dataset, we change the weight of each dataset (Fletcher et al. 2019; Schaub and Kéry in press). If the datasets are too small, another possibility to change the weight of the datasets is to artificially increase the size of one of them by cloning (Lele et al. 2007). If the assumptions are met and there is no conflict between datasets, changing the size of a dataset should not result in different parameter estimates though the uncertainty would be affected. As an illustration, the scenario including recapture heterogeneity, estimated recapture probability increased when a subset of the capture-recapture data was used.

Finally, comparisons of estimates with prior knowledge about the demography of a species can help to identify misspecified models. If no prior knowledge is available, knowledge from a related species or even allometric relationships with demographic rates can be used. A reasonable distrust is healthy. To hypothesize which assumptions really are violated, profound knowledge of the sampling design and of the species is required. Once we have a hypothesis about which assumption might be violated, a targeted GOF test can be performed. GOF tests must be specific to the particular assumption tested (Gelman et al. 1996; Choquet et al. 2009; Kéry and Schaub 2012; McCrea et al. 2016; Gimenez et al. 2018). Multiple specific tests should be performed to test the fit of a general demographic model (Gelman et al. 1996; Kéry and Schaub 2012; Besbeas and Morgan 2014). Besides the tests that we have evaluated, cross-validation has a large potential to compare the predictions of the population model to supplemental data, put aside to fit the model (Conn et al. 2018; Hooten and Hobbs 2015). However with long-term datasets, data are often too sparse to be held out. Moreover, in wild populations, the influence of spatial or temporal heterogeneity on demographic parameters can be overlooked if some data are left aside.

4.5 Conclusions

As correctly remembered by Conn et al. (2018) the goal of goodness of fit testing is not to find a perfectly fitting model, but one that does not violate assumptions which result in systematic errors. Simple models are often useful for many purposes because they are robust when the amount of data is large enough (Stephens et al. 2002). Nevertheless it should become a routine to test the fit of IPMs. Multiple comparisons of parameters estimated with single data models and using reduced or enlarged datasets help to identify lack of fit and conflicts between datasets.

For more complex models that allow addressing more specific questions, GOF tests should be used targeting specific assumptions. Consistency between estimates and expected knowledge are paramount to assessing model fit—unexpected and strange values may be a warning signal. Many demographic processes or other mechanisms may influence survival, reproduction or directly population growth rate in an additive, multiplicative, linear or non-linear way, and expert knowledge on the population being studied can set a complex model on the right path.

Declarations

References

Abadi F, Gimenez O, Arlettaz R, Schaub M (2010a) An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. Ecology 91:7–14

Abadi F, Gimenez O, Ullrich B, Arlettaz R, Schaub M (2010b) Estimation of immigration rate using integrated population models. J Appl Ecol 47(2):393–400

Abadi F, Gimenez O, Jakober H, Stauber W, Arlettaz R, Schaub M (2012) Estimating the strength of density dependence in the presence of observation errors using integrated population models. Ecol Model 242:1–9

Abadi F, Botha A, Altwegg R (2013) Revisiting the effect of capture heterogeneity on survival estimates in capture-mark-recapture studies: does it matter? PLoS ONE 8(4):e62636

Abadi F, Barbraud C, Gimenez O (2017) Integrated population modeling reveals the impact of climate on the survival of juvenile emperor penguins. Glob Change Biol 23:1353–1359

Arnold TW, Clark RG, Koons DN, Schaub M (2018) Integrated population models facilitate ecological understanding and improved management decisions. J Wildl Manag 82(2):266–274

Barbraud C, Delord K (2021) Selection against immigrants in wild seabird populations. Ecol Lett 24(1):84–93

Barraquand F, Gimenez O (2019) Integrating multiple data sources to fit matrix population models for interacting species. Ecol Model 411:108713

Benton TG, Plaistow SJ, Coulson TN (2006) Complex population dynamics and complex causation: devils, details and demography. Proc R Soc B 273:1173–1181

Besbeas P, Morgan BJT (2014) Goodness-of-fit of integrated population models using calibrated simulation. Methods Ecol Evol 5:1373–1382

Besbeas P, Freeman SN, Morgan BJT, Catchpole EA (2002) Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. Biometrics 3:540–547

Bled F, Belant JL, Van Daele LJ, Svoboda N, Gustine D, Hilderbrand G, Barnes VG Jr (2017) Using multiple data types and integrated population models to improve our knowledge of apex predator population dynamics. Ecol Evol 7:9531–9543

Brommer JE, Wistbacka R, Selonen V (2017) Immigration ensures population survival in the s iberian flying squirrel. Ecol Evol 7(6):1858–1868

Brooks SP, Catchpole EA, Morgan BJ et al (2000) Bayesian animal survival estimation. Stat Sci 15(4):357–376

Carothers A (1973) The effects of unequal catchability on jolly-seber estimates. Biometrics 29:79–100

Carvalho F, Punt AE, Chang YJ, Maunder MN, Piner KR (2017) Can diagnostic tests help identify model misspecification in integrated stock assessments? Fish Res 192:28–40

Choquet R, Lebreton J, Gimenez O, Reboulet A, Pradel R (2009) U-CARE: utilities for performing goodness of fit tests and manipulating CApture-REcapture data. Ecography 32(6):1071–1074. https://doi.org/10.1111/j.1600-0587.2009.05968.x

Cleasby IR, Bodey TW, Vigfusdottir F, McDonald JL, McElwaine G, Mackie K, Colhoun K, Bearhop S (2017) Climatic conditions produce contrasting influences on demographic traits in a long-distance arctic migrant. J Anim Ecol 86(2):285–295

Conn PB, Johnson DS, Williams PJ, Melin SR, Hooten MB (2018) A guide to bayesian model checking for ecologists. Ecol Monogr 88(4):526–542

de Valpine P, Turek D, Paciorek CJ, Anderson-Bergman C, Temple Lang D, Bodik R (2017) Programming with models: writing statistical algorithms for general model structures with NIMBLE. J Comput Graph Stat 26:403–417. https://doi.org/10.1080/10618600.2016.1172487

Devineau O, Choquet R, Lebreton JD (2006) Planning capture-recapture studies: straightforward precision, bias, and power calculations. Wildl Soc Bull 34(4):1028–1035

Duarte A, Weckerly FW, Schaub M, Hatfield JS (2016) Estimating golden-cheeked warbler immigration: implications for the spatial scale of conservation. Anim Conserv 19(1):65–74

Duarte A, Pearl CA, Adams MJ, Peterson JT (2017) A new parameterization for integrated population models to document amphibian reintroductions. Ecol Appl 27(6):1761–1775

Earl JE (2019) Evaluating the assumptions of population projection models used for conservation. Biol Conserv 237:145–154

Evans MR, Grimm V, Johst K, Knuuttila T, De Langhe R, Lessells CM, Merz M, O'Malley MA, Orzack SH, Weisberg M et al (2013) Do simple models lead to generality in ecology? Trends Ecol Evol 28(10):578–583

Fay R, Michler S, Laesser J, Schaub M (2019) Integrated population model reveals that kestrels breeding in nest boxes operate as a source population. Ecography 42(12):2122–2131

Fletcher D, Lebreton JD, Marescot L, Schaub M, Gimenez O, Dawson S, Slooten E (2012) Bias in estimation of adult survival and asymptotic population growth rate caused by undetected capture heterogeneity. Methods Ecol Evol 3(1):206–216

Fletcher RJ, Hefley T, Robertson EP, Zuckerberg B, McCleery RA, Dorazio RM (2019) A practical guide for combining data to model species distribution. Ecology 100:e02710

Gaillard JM, Festa-Bianchet M, Yoccoz NG, Loison A, Toïgo C (2000) Temporal variation in fitness components and population dynamics of large herbivores. Annu Rev Ecol Syst 31:367–393

Gamelon M, Grotan V, Sand EE, Bjokvoll VME, Saether BE (2016) Density dependence in an age-structured population of great tits: identifying the critical age classes. Ecology 97:2479–2490

Gauthier G, Pradel R, Menu S, Lebreton JD (2001) Seasonal survival of greater snow geese and effect of hunting under dependence in sighting probability. Ecology 82(11):3105–3119

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–511

Gelman A, Meng X, Stern H (1996) Posterior predictive assessment of model fitness via realised discrepancies. Stat Sinica 6:733–807

Gimenez O, Lebreton JD, Choquet R, Pradel R (2018) R2ucare: an r package to perform goodness-of-fit tests for capture-recapture models. Methods Ecol Evol 9(7):1749–1754

Grist H, Daunt F, Wanless S, Burthe SJ, Newell MA, Harris MP, Reid JM (2017) Reproductive performance of resident and migrant males, females and pairs in a partially migratory bird. J Anim Ecol 86(5):1010–1021

Grotan V, Saether BE, Engen S, van Balen JH, Perdeck AC, Visser ME (2009) Spatial and temporal variation in the relative contribution of density dependence, climate variation and migration to fluctuations in the size of great tit populations. J Anim Ecol 78:447–459

Hatter IW, Dielman P, Kuzyk GW (2017) An integrated modeling approach for assessing management objectives for mule deer in central british columbia. Wildl Soc Bull 41(3):508–515

Hixon MA, Pacala SW, Sandin SA (2002) Population regulation: historical context and contemporary challenges of open vs. closed systems. Ecology 83(6):1490–1508

Hooten MB, Hobbs NT (2015) A guide to Bayesian model selection for ecologists. Ecol Monogr 85:3–28

Johnson JB, Omland KS (2004) Model selection in ecology and evolution. Trends Ecol Evol 19:101–108. https://doi.org/10.1016/j.tree.2003.10.013

Kendall B, Wittmann M (2010) A stochastic model for annual reproductive success. Am Nat 175(4):461–468

Kendall BE, Fox GA, Fujiwara M, Nogeire TM (2011) Demographic heterogeneity, cohort selection, and population growth. Ecology 92:1985–1993

Kéry M, Schaub M (2012) Bayesian population analysis using WinBUGS: a hierarchical perspective. Academic Press, Burlington

Koons DN, Iles DT, Schaub M, Caswell H (2016) A life-history perspective on the demographic drivers of structured population dynamics in changing environments. Ecol Lett 19(9):1023–1031

Koons DN, Arnold TW, Schaub M (2017) Estimating the demographic drivers of realized population dynamics. Ecol Appl 27:2102–2115

Lee AM, Reid JM, Beissinger SR (2017) Modelling effects of nonbreeders on population growth estimates. J Anim Ecol 86(1):75–87

Lele SR, Dennis B, Lutscher F (2007) Data cloning: easy maximum likelihood estimation for complex ecological models using bayesian markov chain monte carlo methods. Ecol Lett 10(7):551–563

Lieury N, Gallardo M, Ponchon C, Besnard A, Millon A (2015) Relative contribution of local demography and immigration in the recovery of a geographically-isolated population of the endangered egyptian vulture. Biol Conserv 191:349–356

Margalida A, Jiménez J, Martínez JM, Sesé JA, García-Ferré D, Llamas A, Razin M, Colomer M, Arroyo B (2020) An assessment of population size and demographic drivers of the bearded vulture using integrated population models. Ecol Monogr 90(3):e01414

Maunder MN, Piner KR (2017) Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fish Res 192:16–27

McCrea RS, Morgan BJ, Gimenez O (2016) A new strategy for diagnostic model assessment in capture-recapture. J R Stat Soc 66(4):815–831

Millon A, Lambin X, Devillard S, Schaub M (2019) Quantifying the contribution of immigration to population dynamics: a review of methods, evidence and perspectives in birds and mammals. Biol Rev 94(6):2049–2067

Plard F, Bruns H, Cimiotti D, Helmecke A, Hötker H, Jeromin H, Roodbergen M, Schekkerman H, Teunissen W, van der Jeugd H et al (2019a) Low productivity and unsuitable management drive the decline of central european lapwing populations. Anim Conserv 23:286–296

Plard F, Fay R, Kéry M, Cohas A, Schaub M (2019b) Integrated population models: powerful methods to embed individual processes in population dynamics models. Ecology 100:e02715

Plard F, Turek D, Grüebler MU, Schaub M (2019c) IPM^2: toward better understanding and forecasting of population dynamics. Ecol Monogr 89(3):e01364

Plard F, Arlettaz R, Jacot A, Schaub M (2020) Disentangling the spatial and temporal causes of decline in a bird population. Ecol Evol 10(14):6906–6918

R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rhodes R, Fey Ng CF, de Villiers DL, Preece HJ, McAlpine CA, Possingham HP (2011) Using integrated population modelling to quantify the implications of multiple threatening processes for a rapidly declining population. Biol Conserv 144:1081–1088

Riecke TV, Williams PJ, Behmke TL, Gibson D, Leach AG, Sedinger BS, Street PA, Sedinger JS (2019) Integrated population models: model assumptions and inference. Methods Ecol Evol 10:1072–1082

Robinson RA, Morrison CA, Baillie SR (2014) Integrating demographic data: towards a framework for monitoring wildlife populations at large spatial scales. Methods Ecol Evol 5:1361–1372

Robinson RA, Meier CM, Witvliet W, Kéry M, Schaub M (2020) Survival varies seasonally in a migratory bird: Linkages between breeding and non-breeding periods. J Anim Ecol 89(9):2111–2121

Rockwell SM, Wunderle JM, Sillett TS, Bocetti CI, Ewert DN, Currie D, White JD, Marra PP (2017) Seasonal survival estimation for a long-distance migratory bird and the influence of winter precipitation. Oecologia 183(3):715–726

Rolandsen CM, Solberg EJ, Sæther BE, Moorter BV, Herfindal I, Bjørneraas K (2017) On fitness and partial migration in a large herbivore-migratory moose have higher reproductive performance than residents. Oikos 126(4):547–555

Saether B, Bakke O (2000) Avian life history variation and contribution of demographic traits to the population growth rate. Ecology 81:642–653

Sæther BE, Engen S, Pape Møller A, Weimerskirch H, Visser ME, Fiedler W, Matthysen E, Lambrechts MM, Badyaev A, Becker PH et al (2004) Life-history variation predicts the effects of demographic stochasticity on avian population dynamics. Am Nat 164(6):793–802

Saunders SP, Cuthbert FJ, Zipkin EF (2018) Evaluating population viability and efficacy of conservation management using integrated population models. J Appl Ecol 55(3):1380–1392

Saunders SP, Farr MT, Wright AD, Bahlai CA, Ribeiro JW Jr, Rossman S, Sussman AL, Arnold TW, Zipkin EF (2019) Disentangling data discrepancies with integrated population models. Ecology 100(6):e02714

Schaub M, Abadi F (2011) Integrated population models: a novel analysis framework for deeper insights into population dynamics. J Ornithol 152:S227–S237

Schaub M, Fletcher D (2015) Estimating immigration using a Bayesian integrated population model: choice of parametrization and priors. Environ Ecol Stat 22(3):535–549

Schaub M, Kéry M (in press) Integrated population models. Theory and ecological applications with R and JAGS. Academic Press, Burlington. ISBN 978–0128205648

Schaub M, Ullrich B (in press) A drop in immigration results in the extinction of a local woodchat shrike population. Anim Conserv. https://doi.org/10.1111/acv.12639

Schaub M, Ullrich B, Knötzsch G, Albrecht P, Meisser C (2006) Local population dynamics and the impact of scale and isolation: a study on different little owl populations. Oikos 115:389–400

Schaub M, Gimenez O, Sierro A, Arlettaz R (2007) Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. Conserv Biol 21(4):945–955

Stephens PA, Frey-Roos F, Arnold W, Sutherland WJ (2002) Sustainable exploitation of social species: a test and comparison of models. J Appl Ecol 39(4):629–642

Tavecchia G, Besbeas P, Coulson T, Morgan BJT, Clutton-Brock TH (2009) Estimating population size and hidden demographic parameters with state-space modeling. Am Nat 173:722–733

Tempel DJ, Peery MZ, Gutierrez RJ (2014) Using integrated population models to improve conservation monitoring: California spotted owls as a case study. Ecol Modell 289:86–95

Tenan S, Fasola M, Volponi S, Tavecchia G (2017) Conspecific and not performance-based attraction on immigrants drives colony growth in a waterbird. J Anim Ecol 86(5):1074–1081

Turchin P (1995) Population regulation: old arguments and a new synthesis. In: Cappuccino N, Price PW (eds) Population dynamics : new approaches and synthesis. Academic Press, San Diego, pp 19–40

Vindenes Y, Engen S, Sæther BE (2008) Individual heterogeneity in vital parameters and demographic stochasticity. Am Nat 171(4):455–467

Ward MP (2005) The role of immigration in the decline of an isolated migratory bird population. Conserv Biol 19:1528–1536

Weegman MD, Arnold TW, Clark RG, Schaub M (2021) Partial and complete dependency among data sets has minimal consequence on estimates from integrated population models. Ecol Appl e225831(3): e02258a

Zipkin EF, Saunders SP (2018) Synthesizing multiple data types for biological conservation using integrated population models. Biol Conserv 217:240–250

Zipkin EF, Inouye BD, Beissinger SR (2019) Innovations in data integration for modeling populations. Ecology 100:e02713

Floriane Plard got her PhD in 2014 in The University of Lyon 1, France. She worked as a postdoctoral fellow at Stanford University, USA during the following year. Then, she started this work during her four

year post-doctoral position at the Swiss Ornithological Institute, Sempach, Switzerland. Since then, she has been a year an Assistant Professor in statistics at the University Lyon 2, France and is now an invited associate professor at Holar University in Iceland. Her research is at the interface between the theory of life-histories, population ecology and quantitative ecology. She developed theoretical and statistical models to provide a better understanding and prediction of population dynamics using individual mechanisms.

Daniel Turek received his PhD in Statistics at the University of Otago, New Zealand. After completing his PhD, Daniel began a postdoctoral research position at the University of California, Berkeley, where he helped develop the NIMBLE software project. He is now an Assistant Professor of Statistics at Williams College, Massachusetts. Daniel's research interests are in computational statistics, Bayesian analyses of hierarchical models, and MCMC sampling algorithms.

Michael Schaub got his Ph.D. in Ecology from the University of Zurich, Switzerland. Thereafter he worked as a Postdoctoral research fellow in the CNRS Montpellier, France, and later at University of Bern, Switzerland. Now he is the head of the ecological research department at the Swiss Ornithological Institute, Sempach, Switzerland. Michael's research interests are the demography, population dynamics and conservation of birds and the development of statistical tools for their study.