REVIEW

Integrated population models: a novel analysis framework for deeper insights into population dynamics

Michael Schaub · Fitsum Abadi

Abstract Integrated population models (IPMs) represent the single, unified analysis of population count data and demographic data. This modelling framework is quite novel and can be implemented within the classical or the Bayesian mode of statistical inference. Here, we briefly show the basic steps that need to be taken when an integrated population model is adopted, and review existing integrated population models for birds and mammals. There are important advantages of integrated compared to conventional analyses that analyse each dataset separately and then try to make an inference about population dynamics. First, integrated population models allow the estimating of more demographic quantities, because there is information about all demographic processes operating in a population, and this information is exploited. Second, parameter estimates become more precise, and this enhances statistical power. Finally, all sources of uncertainty due to process variability and the sampling process(es) are adequately included. Core of the integrated models is the link of changes in the population size and the demographic rates via a demographic model (usually a Leslie matrix model) and the likelihoods of all existing datasets. We discuss some critical assumptions that are typically made in integrated population models and highlight fruitful areas of future research. Currently, we have found 25 studies that used integrated population models. Central to most studies was statistical development rather than their application to address an ecological question, which is not surprising given that integrated population models are still a new development. We predict that integrated population models will become a common and important tool in studies of population dynamics, both in ecology and its applications, such as conservation biology or wildlife management.

Keywords Bayesian · Demography · Leslie matrix · Population dynamics · Population growth rate · State-space model

M. Schaub (✉)
Swiss Ornithological Institute, 6204 Sempach, Switzerland
e-mail: michael.schaub@vogelwarte.ch

M. Schaub · F. Abadi
Institute of Ecology and Evolution, Division of Conservation Biology, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland

F. Abadi
South African National Biodiversity Institute, Kirstenbosch Research Centre, Claremont 7735, South Africa

F. Abadi
Animal Demography Unit, Department of Zoology, University of Cape Town, Rondebosch 7701, South Africa

Zusammenfassung Integrierte Populationmodelle (IPM) sind universelle Auswertungsmodelle mit denen jährliche Populationszählungen und demographische Daten simultan ausgewertet werden können. Diese Auswertungsmodelle sind relativ neu und können sowohl Bayesianisch wie auch frequentistisch analysiert werden. In diesem Artikel zeigen wir die wichtigsten Schritte, die es braucht, um ein IPM aufzustellen und geben eine Übersicht über die bisher auf Vögel- und Säugerdaten angewendeten IPM. Die Anwendung integrierter Populationsmodellen hat wichtige Vorteile gegenüber einer klassischen Auswertung, die die einzelnen Datensätze separat auswertet. Erstens, erlauben die IPM die Schätzung von demographischen Parametern, von denen keine spezifischen Daten vorliegen. Dies ist möglich, weil in den Populationszählungen Information

über alle demographischen Prozesse vorhanden ist, und diese Information wird in IPM explizit extrahiert. Zweitens werden alle Parameter präziser geschätzt, was Rückschlüsse und die weitere Modellierung erleichtern kann. Und drittens werden die gesamten Unsicherheiten, die auf Grund der Datensammlung bestehen, adequat berücksichtigt. Zentral für ein IPM ist eine Beziehung zwischen den demographischen Parametern und der Populationsgrösse (meist via einer Leslie Matrix) und Wahrscheinlichkeitsmodelle aller Datensätze. Wir diskutieren die kritischen Annahmen der IPM und zeigen mögliche zukünftige Forschungsfelder auf. Wir fanden 25 Studien, die IPM verwendet haben. Ein zentraler Punkt bei fast allen war die statistische Weiterentwicklung. Wir sind überzeugt, dass sich die IPM für viele Studien im Bereich der Populationsdynamik, aber auch von Naturschutz-und Wildbiologie, zu einem wichtigen Auswertungsinstrument entwickeln werden.

Introduction

A fundamental goal in population ecology is to understand how populations are regulated (Newton 1998). The dynamics of a population is the variation of the population size over time. Population dynamics is the result of losses and gains of individuals in the population and therefore of the four demographic key processes: recruitment, survival, emigration and immigration (Sibly and Hone 2002). Any change in the population growth rate requires a change of at least one of these underlying demographic rates. Knowledge about how the environment affects demographic rates and how this is translated into a change of population growth is therefore the key towards an understanding of how populations are regulated. While there is fundamental interest in this endeavour in ecological science, such a demographic assessment of populations is also an essential step in conservation, as it allows one to identify the demographic causes of population declines and to guide management decisions to halt or reverse a decline (Caughley 1994; Norris 2004; Baillie and Schaub 2009). Inversely, for the management of a pest species, it may permit identification of the most efficient way to control it.

Modern population dynamical research is necessarily quantitative, and many statistical methods exist to describe and explain demographic processes in populations. In this paper, we focus on the recently developed integrated population models with a special objective concerning their application to birds and mammals. We show how such a model is set up by highlighting the conceptual steps and avoiding much of the mathematical details. We then provide a short review of published integrated population models, discuss the current state of research and attempt to predict future developments.

Elements of integrated population models

An integrated analysis in general is a joint analysis of several datasets on different quantities. Inference is based on the joint likelihood, which is created by the multiplication of the likelihoods from the single datasets. A key element is that one or several parameters are common in several components of the likelihood. Examples are the joint analysis of capture–recapture and mark–recovery data (Burnham 1993; Lebreton et al. 1995) or of capture–recapture and carcass inspection data (Goodman 2004). We here define integrated population models as models that jointly analyse data on population size and data on demographic parameters. The focus of inference is the population growth as well as demographic rates. This approach helps to gain a deep insight into population dynamics, as it allows for a study of the links between impacting factors (environment and density), demography and population growth.

Analyses using population size and demographic data are not new developments. Traditionally, the available data have been analysed separately (e.g. Jenouvrier et al. 2003; Schaub et al. 2004). The population growth rate is estimated from the population count data and demographic parameters are estimated from demographic data, each with a different, separate likelihood. The estimated demographic parameters are then used as an input in population projection matrices and important inferences are obtained by comparisons of the results from the two separate analyses. First, the comparison between the population growth rate derived from projection matrices and the one obtained from the count data allows inferring whether all relevant demographic parameters are considered in the population model (Jenouvrier et al. 2003). Second, the correlations between the observed population growth rate and demographic rates may suggest which demographic rate was the most important driver of population growth (e.g. Thomson et al. 1997; Siriwardena et al. 1999; Peach et al. 1999; Freeman and Crick 2003; Robinson et al. 2004; Freeman et al. 2007). However, while there is nothing wrong in this approach, it has a number of drawbacks: (1) it is inefficient as it does not use all the available information, (2) there is no formal way to account for the uncertainty (variances and covariances) in demographic parameters and population growth, and (3) unless information about all demographic parameters is included, population growth rates derived from projection matrices are biased. We expand on these points below.

The recently developed integrated population models (Besbeas et al. 2002; Brooks et al. 2004) hold promise to

overcome these drawbacks. In these models, all available data are analysed within a single model; that is, the single data likelihoods are used to create the joint likelihood upon which inference is based. They make use of one or several state-space model components to model the true state separately from the observation process. For instance, as a description of the observed population counts the state equation describes the link between population sizes and demography, while the observation equation describes the link between population size and population counts. The combination of additional models that directly estimate demographic parameters from demographic data (such as ring recoveries) with the state-space model for the population count data forms the integrated population model. This modelling framework has several advantages: all uncertainty emerging from the fact that the data stem from a random sample of animals in a population of interest are accounted for, demographic parameters for which no explicit demographic data are available can often be estimated, and parameter estimates become more precise (Besbeas et al. 2002; Tavecchia et al. 2009; Abadi et al. 2010a). The latter two points are a direct consequence of the more efficient exploitation of all the available information: population growth is a function of demography, therefore the population count data contain information about all demographic processes in the population. This information is formally exploited in integrated population models.

How to set up an integrated population model

The construction of an integrated population model consists of three basic steps. First, a model that links population sizes and demographic rates needs to be specified; second, the component likelihoods of the available datasets have to be formulated; and third, the joint likelihood of all datasets combined must be created on which inference is based. By means of an example, we will now show how these steps are taken. We consider the case of a geographically open population of a short-lived bird species (imagine a passerine) from which we have annual counts of breeders, capture–recapture data and data on productivity. The resulting model is described in detail in Abadi et al. (2010b), where computer code to run the model as well as data are also available.

Step 1: define a population model

In the first step, a model that links the population sizes with the demographic rates is defined. Typically, this is a projection matrix model (Caswell 2001). Depending on the amount of data, the goal of the study and the assumptions one is ready to accept, this model may be more or less complicated. It may be a true age-structured model (i.e. Leslie matrix), a stage-structured model (i.e. Lefkovitch matrix) and it may or may not consider sexes. In our example, we consider a female-based model with a pre-breeding census and two stages. One stage refers to the number of individuals that are 1 year old and another to those that are older than 1 year (adult). A life cycle graph for this species is provided in Fig. 1. The population sizes of each stage (age class) can be written mathematically as a function of the population sizes in the previous year and the demographic rates:

$$N_{1,t+1} = N_{1,t}b_tS_{1,t}F_{1,t} + N_{a,t}b_tS_{1,t}F_{1,t} + (N_{1,t} + N_{a,t})\omega_t$$
$$N_{a,t+1} = N_{1,t}S_{a,t}F_{a,t} + N_{a,t}S_{a,t}F_{a,t}. \quad (1)$$

$N_{1,t}$ and $N_{a,t}$ are the number of females at time t in age classes 1-year and adult, respectively, b_t is the number of fledged females in year t that are produced per breeding female, $S_{1,t}$ and $S_{a,t}$ are the annual survival probabilities during the first year of life and later, respectively, $F_{1,t}$ and $F_{a,t}$ are the stage-specific probabilities that females are philopatric to the population, given that they have survived, and ω_t is the immigration rate of females in year t. The immigration rate ω_t is defined as the number of immigrated females in year $t + 1$ per female present in year t. We assume that immigrants are 1 year old. However, since the demographic parameters of individuals aged 1 year and older are identical, this assumption is not likely to be overly restrictive. An alternative way to write the same model is by considering vectors of population sizes and a transition matrix (here a Lefkovich matrix):

$$\begin{bmatrix} N_1 \\ N_a \end{bmatrix}_{t+1} = \begin{bmatrix} bS_1F_1 + \omega & bS_1F_1 + \omega \\ S_aF_a & S_aF_a \end{bmatrix}_t \begin{bmatrix} N_1 \\ N_a \end{bmatrix}_t. \quad (2)$$

The product of true survival and philopatry is apparent survival $(\phi_t = S_tF_t)$, which is the parameter estimated from capture–recapture data using the standard Cormack–Jolly–Seber model (Lebreton et al. 1992). Because the available data do not allow a decomposition into S and F, we will use ϕ hereafter.

Equation 1 describes the exact relationship between stage-specific population sizes and demographic parameters. However, we might want to include demographic stochasticity, because this is important in small populations (Lande 1993). A natural way to include demographic stochasticity is by considering appropriate statistical distributions. For the 1-year-old individuals, such a distribution should have the properties that it generates integer values (there are no half individuals), that the lowest possible number is 0 (if no recruitment occurs), that the highest possible number is ∞ (if recruitment is immense) and that the expected value satisfies Eq. 1 The Poisson distribution has these properties, and thus we may write
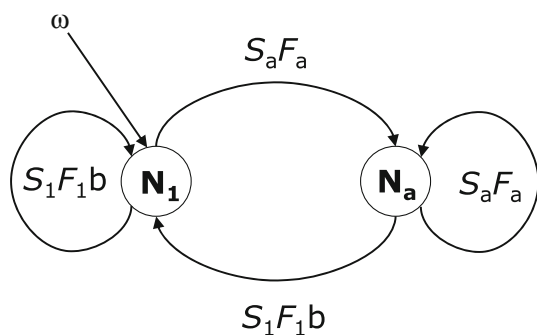
Fig. 1 Life cycle graph of a geographically open population of a short-lived passerine. The model is female-based and considers a pre-breeding census. The two stages are the 1-year-old individuals (N_1) and more than 1-year-old individuals (N_a). The demographic parameters are age-specific survival (S_1, S_a), site fidelity (F_1, F_a), productivity (b) and immigration (ω)

$$N_{1,t+1} \sim Pois(N_{1,t}b_t\phi_{1,t} + N_{a,t}b_t\phi_{1,t} + N_{1,t}\omega_t + N_{a,t}\omega_t)$$
$$= N_{1,t+1} \sim Pois((N_{1,t} + N_{a,t})(b_t\phi_{1,t} + \omega_t)). \quad (3)$$

For modelling the number of adult individuals, the appropriate distribution must have the properties that it generates integer values, that the lowest possible number is 0 (if no individual survives), that the highest possible number is the population size of the year before (if all survive), and that the expected value satisfies Eq. 1. Such a distribution is the binomial distribution, and we can write

$$N_{a,t+1} \sim Bin(N_{1,t} + N_{a,t}, \phi_{a,t}). \quad (4)$$

This concludes the description of our population model. To estimate its parameters for the population under study, we must now define the likelihoods.

Step 2: define the likelihoods for the available data

The next step is to define the likelihoods for the available dataset separately (here: population counts, capture–recapture data, productivity data). We start with the population counts data, for which we use the likelihood of a state-space model (De Valpine and Hastings 2002; Clark and Björnstad 2004). A state-space model is composed of a set of process equations describing the true but unknown development of the states across time and of observation equations linking the true states with the data (population counts). The process equations are already defined (Eqs. 3 and 4), they describe the population size as a function of demographic rates and the population size of the preceding year. The observations are conditional on the process, and we assume that the population counts in year t (y_t) are conducted shortly before reproduction starts, that counts consist of the sum of the number of breeding females which in our example is the sum of the number of females

in each stage class, and that these counts are related to the population sizes as

$$y_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{a,t} \end{bmatrix} + \varepsilon_t$$
$$\varepsilon_t \sim N\left(0, \sigma_y^2\right). \quad (5)$$

We assume that in each year t we make an error while counting the population (ε_t), in one year we may miss more individuals than we double-count, in another year this may be the other way round. We also assume that these counting errors can be described by a Normal distribution with mean zero and the variance (σ_y^2); the latter is denoted as observation error. This error is in fact a residual error and therefore incorporates not only observation errors of the counts but also lack of fit of the state equations to the true (but unobserved) dynamics of the population under study (Tavecchia et al. 2009). Other distributions for the residual error are also possible, like the Poisson or the log-Normal distribution. Sometimes it is also possible to get separate counts from different age classes, with the advantage of being able to extract more detailed information from the counts. The observation Eq. 5 does then need a slight adaptation (see, e.g., Tavecchia et al. 2009).

The likelihood of the state-space model (L_{SS}) for the count data is the product of the likelihood of the process (L_{SY}) and the observation equations (L_{OB}):

$$L_{SS}\left(\mathbf{y}|\mathbf{N}, \phi_1, \phi_\mathbf{a}, \mathbf{b}, \omega, \sigma_\mathbf{y}^2\right)$$
$$= L_{OB}\left(\mathbf{y}|\mathbf{N}, \sigma_\mathbf{y}^2\right) \times L_{SY}(\mathbf{N}|\phi_1, \phi_\mathbf{a}, \mathbf{b}, \omega). \quad (6)$$

Note that here and in the following the likelihoods are written using vectors and matrices (in bold), and thus the subscript referring to time are lost.

The state-space likelihood for the counts already contains all parameters that we would like to estimate. However, the parameters of the model are very difficult to estimate from the count data alone and usually require additional information such as specific constraints (King et al. 2010) or informative priors (Buckland et al. 2004, 2007; Thomas et al. 2005). In the integrated model, we bring in additional information which comes from the demographic data available.

The second dataset for which we create the likelihood are the capture–recapture data. We use the likelihood of the Cormack–Jolly–Seber model that is described in many papers (e.g. Lebreton et al. 1992; Williams et al. 2002; Abadi et al. 2010a). The capture–recapture data are summarised in the m-array format ($\mathbf{m}$; Lebreton et al. 1992) and a multinomial likelihood is used to relate the data with the apparent survival and recapture probability (p). The likelihood is $L_{CMR}(\mathbf{m}|\phi_1, \phi_\mathbf{a}, \mathbf{p})$.

The third dataset are the annual number of fledglings (J_t) in the surveyed broods (R_t). We use a simple Poisson regression model to estimate annual productivity (b_t) as $J_t \sim Pois(R_t b_t)$. The likelihood is defined by $L_{RS}(\mathbf{J}, \mathbf{R}|\mathbf{b})$. Typically, no state-space formulation is adopted for this model component, i.e. it is assumed that fecundity is measured without error.

Step 3: make inference from the joint likelihood

As the likelihoods of the three datasets are defined, we first have to create the joint likelihood, i.e. the likelihood of the integrated model. If we can assume independence among the component datasets of the integrated analysis, the joint likelihood (L_{IPM}) is just the product of the component likelihoods:

$$L_{IPM}\left(\mathbf{y}, \mathbf{m}, \mathbf{J}, \mathbf{R}|\mathbf{N}, \phi_1, \phi_a, \mathbf{b}, \omega, \mathbf{p}, \sigma_{\mathbf{y}}^2\right)$$
$$= L_{OB}(\mathbf{y}|\mathbf{N}, \sigma_{\mathbf{y}}^2) \times L_{SY}(\mathbf{N}|\phi_1, \phi_a, \mathbf{b}, \omega)$$
$$\times L_{CMR}(\mathbf{m}|\phi_1, \phi_a, \mathbf{p}) \times L_{RS}(\mathbf{J}, \mathbf{R}|\mathbf{b}). \quad (7)$$

The independence of the datasets is a fairly crucial assumption. Under one—restrictive—view, this means that no animal in the count data may occur in the capture–recapture data or in the productivity data. In practice, we often have different kinds of demographic data from a single and rather small population, indeed, it is for small populations where the combination of information in an integrated population model is particularly valuable (Schaub et al. 2007). Therefore, it is important to understand whether the violation of the independence assumption has a strong impact on the parameter estimates. From basic statistical intuition regarding non-independence in datasets, one would assume that dependent datasets would not result in biased parameter estimates, but in overestimated precision of the estimates. This is a classical result in statistics when non-independent data are analysed as if they were independent (Lebreton et al. 1992; Anderson et al. 1994). The information from one individual is used more than once and, hence, the genuine sample size is not as large as it seems. Abadi et al. (2010a) simulated data types similar to the ones we use here with different degrees of non-independence. They found that the violation of independence had almost no effect on the accuracy of the parameter estimates. However, this may not be the case if completely different types of data are combined, for instance ring recovery and count data (Besbeas et al. 2009).

The likelihoods of the three datasets have some parameters in common, as displayed graphically in Fig. 2. By combining these data sources into a single analysis using an integrated population model, more information can be used to estimate demographic parameters.

To make inference, the joint likelihood (Eq. 7) needs to be analysed. This can be done either in the frequentist or in the Bayesian framework. In the frequentist framework, one maximises the joint likelihood. This is possible under some additional assumptions (e.g. Normality and linearity) and if Kalman filtering is applied (Besbeas et al. 2002). An additional assumption refers to the distributions of the process equations (Eqs. 3 and 4) which need to be approximated by normal distributions (the Poisson and binomial distributions cannot be used; Besbeas et al. 2003). The Bayesian framework is more flexible, but requires the definition of prior distributions of all parameters. These are then combined with the joint likelihood to get the posterior distribution of all parameters. The posterior distribution can easily be approximated using simulation-based approaches such as Markov chain Monte Carlo (MCMC; Brooks 2003). The model just described has been applied to data of populations of Little Owls (*Athene noctua*; Abadi et al. 2010b).

A benefit of this particular integrated population model is that it allows to estimate immigration rate, which is otherwise very hard to estimate (see discussion in Abadi et al. 2010b). In addition, it is also possible to assess how immigration is impacted by environmental factors and to estimate the temporal variance of immigration. Modelling of covariates is performed via the usual generalised linear model framework (McCullagh and Nelder 1989). Abadi et al. (2010b) modelled the logarithm of the immigration rate as

$$\log(\omega_t) = \beta_0 + \beta_1 x_t \quad (8)$$

where x is the covariate (vole density) and β_0 and β_1 are the intercept and the slope, respectively, which are estimated. Clearly, other parameters of the integrated population model can be modelled in a similar way, and it is also
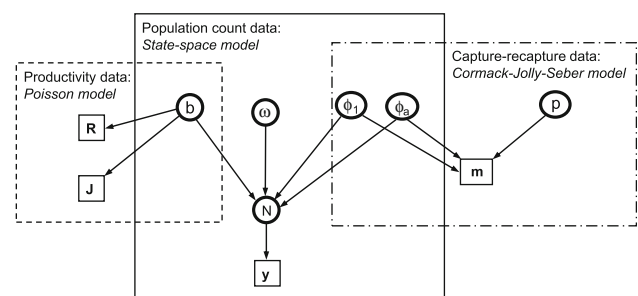

Fig. 2 Directed acyclic graph of the presented integrated population, showing that some demographic parameters are common to different sources of information. Estimated parameters are represented by *circles* and the data are represented by *rectangles*. Arrows represent dependences between nodes. Node notations: *m* capture–recapture data; *y* population count data; *J* number of newborns; *R* number of surveyed broods; *b* productivity; ϕ_1 juvenile apparent survival probability; ϕ_a adult apparent survival probability; ω immigration rate; *p* recapture probability; *N* population size. Note that the priors are not shown on this graph

possible to include random effects (e.g. Abadi et al., in review). Abadi et al. (2010b) have shown that the immigration into a Little Owl population was stronger in years when voles, the main prey, was abundant.

This integrated population model makes a number of assumptions which should be fulfilled, otherwise parameter estimates and/or the estimates of their precision may be biased. The assumptions are the same as those that would hold if the single datasets had been analysed separately. For instance, the Cormack–Jolly–Seber model requires that all individuals have the same survival and recapture probabilities, that the individuals are independent, and that no marks are lost or misread (Lebreton et al. 1992). Depending on the datasets that are used, different assumptions might need to be fulfilled. In addition, the used population model (i.e. the Leslie or Lefkovich matrix) must be appropriate for the population under study. Finally, there are assumptions due to the integration of all datasets, and these are more general, i.e. common to all integrated population models. As discussed above, the model needs the assumption of the independence of the different datasets. In addition, we have to assume that the demography of the individuals included in the different datasets is the same. The more the different datasets are independent, the larger is the risk that this last assumption is violated.

Review on existing integrated population analyses

In this section, we review publications that made use of an integrated population model, i.e. the joint modelling of counts and one or several types of demographic data. The goal is to provide a list of available studies using integrated population models to show their potential and circumstances of use. We focus mostly on bird studies, but include also studies on mammals, because similar datasets are typically collected. For a long time, integrated models have been applied repeatedly in fisheries (e.g. Schnute 1994; Millar and Meyer 2000; Maunder 2004), but we have not included these studies in our review, because the sampled data are usually quite different from those collected for bird studies. For the same reason, we also did not consider studies that integrated counts with harvest data (e.g. Trenkel et al. 2000; White and Lubow 2002). From each considered study, we extracted the following information: (1) types of data that were used, (2) mode of statistical analysis (frequentist or Bayesian), (3) whether or not the studied population was geographically open, (4) study duration, (5) included covariates, and (6) the main focus of the study (statistical development; model evaluation; ecology; conservation). Finally, we summarised for each study which additional parameter could be estimated, because an integrated population model was used. To find the published studies, we used the web of science (date of accession, September 2010).

In total, we found 25 studies (Table 1). Besbeas et al. (2002) was the first publication about an integrated population model that was applied to a bird. Since then, the annual number of publications has been increasing. The 25 studies used data on 18 different species; 2 studies were pure simulations and used no empirical data, but were set up to mimic birds. 12 studies used the frequentist approach while 13 applied Bayesian methods to make inference from the joint likelihood. The average duration of the studies was 29 years (range 9–72). Thus, generally, they were rather long. Twelve datasets refer to populations that were geographically closed, and 6 to geographically open populations. Geographically closed studies often include a very large area such as a complete country with the result that immigration and emigration are not important and do not need to be considered in the model. Small-scale studies on specific populations, which need an estimation of emigration and immigration, are a rather recent development in integrated population modelling (first study with empirical data: Reynolds et al. 2009). The main focus of most studies was the statistical development of methods and their evaluation. Some papers specifically deal with how census data are included (Besbeas and Freeman 2006), and Cave et al. (2010) explored the use of ringing data from constant effort sites as indices of population size. A recent and very promising development is the integration of multi-state capture–recapture with census data (Borysiewicz et al. 2009; Péron et al. 2010; McCrea et al. 2010), which allows the estimating of a rich set of biological parameters (e.g. natal and breeding dispersal, site-specific recruitment and productivity, survival). The benefit in terms of parameter precision of the integration of census data with multi-state capture–recapture data is much more pronounced than when census data are combined with single state capture–recapture data (Borysiewicz et al. 2009). There are only a relatively few and recent studies whose primary aim was ecology or conservation. This is not surprising given that integrated population models are still a new development. Moreover, since the models are very flexible and can be adapted to the specific sampling design and study question, we expect that some statistical development is likely to be included in most future studies whose focus is ecology or conservation. The benefit of the integrated population models over conventional analysis was often that they provided an estimate of productivity in the large-scale studies and estimates of movement parameters (site fidelity, immigration) in the small-scale studies. A further benefit due to the use of an integrated model was, in all studies, increased precision of the parameters estimates (some studies explicitly demonstrate that), but this is not shown in Table 1.

Table 1 Summary table of the reviewed 25 studies applying integrated population models in birds or mammals

Species	Data types[a]	Mode of analysis	Geographically open/closed	Study duration (years)	Included covariates	Focus[b]	Benefit of IPM	Reference
Lapwing *Vanellus vanellus* and Grey Heron *Ardea cinerea*	I, MR (Lapwing) C, MR (Grey Heron)	Frequentist	Closed	35 (Lapwing) 72 (Grey Heron)	Temporal trends Frost days	Statistics, evaluation	Productivity	Besbeas et al. (2002)
Lapwing *Vanellus vanellus* and Grey Heron *Ardea cinerea*	I, MR (Lapwing) C, MR (Grey Heron)	Frequentist	Closed	35 (Lapwing) 72 (Grey Heron)	Temporal trends Frost days	Statistics	Productivity	Besbeas et al. (2003)
Lapwing *Vanellus vanellus*	I, MR	Bayesian	Closed	33	Temporal trend Frost days	Statistics	Productivity	Brooks et al. (2004)
Spotted Dolphin *Stenella attenuata*	C, age-at-death, stage (colour)-at-death	Bayesian	Closed	40	Population size	Statistics, ecology	Productivity, survival	Hoyle and Maunder (2004)
American Black Duck *Anas rubipes*	C, MR	Bayesian	Closed	11	Competition (Mallards)	Statistics, harvest regulations	Productivity	Fonnesbeck and Conroy (2004)
Lapwing *Vanellus vanellus*	I, MR	Frequentist	Closed	35	Temporal trends Frost days Habitat	Statistics	Productivity	Besbeas et al. (2005)
Lapwing *Vanellus vanellus*	C, MR	Frequentist	Closed	35	Temporal trends Frost days	Statistics, evaluation	Productivity	Besbeas and Freeman (2006)
Greater Snow Goose *Chen caerulescens*	C, CMR, P	Frequentist	Closed	18	Harvest rate	Statistics, evaluation, ecology	Population size forecast	Gauthier et al. (2007)
Greater Horseshoe Bat *Rhinolophus ferrumequinum*	C, CMR (recaptures from 2 years only), P	Bayesian	Closed	15	Temporal trends	Ecology, conservation	Productivity	Schaub et al. (2007)
Black-footed Albatross *Phoebastria nigripes*	C, CMR, estimate of productivity was taken from literature	Frequentist	Closed	12	Fishing effort	Ecology, conservation	Immature survival	Véran and Lebreton (2008)
Lapwing *Vanellus vanellus*	I, MR	Bayesian	Closed	35	Temporal trends Frost days	Statistics, evaluation, conservation	Productivity	King et al. (2008)
Song Trush *Turdus pilomelos*	I, MR	Bayesian	Closed	32	Frost days Drought days	Ecology, conservation	Productivity	Baillie et al. (2009)

Table 1 continued

Species	Data types[a]	Mode of analysis	Geographically open/closed	Study duration (years)	Included covariates	Focus[b]	Benefit of IPM	Reference
Grey Heron *Ardea cinerea*	C, MR	Frequentist	Closed	70	Winter severity	Statistics	Productivity	Besbeas et al. (2009)
Soay Sheep *Ovis aries*	C, CMR, P	Frequentist	Closed	19	Temporal trend Population size NAO	Statistics, evaluation, ecology	Sex ratio at birth	Tavecchia et al. (2009)
Simulation study mimicking Great Cormorants *Phalacrocorax carbo*	C, MR, CMR all at multiple sites	Frequentist	Open	–	–	Statistics, evaluation	Productivity	Borysiewicz et al. (2009)
Common Guillemot *Uria aalge*	C, CMR, MR, P	Bayesian	Open	23	–	Statistics, ecology	Emigration, ring loss rate	Reynolds et al. (2009)
Sedge Warbler *Acrocephalus schoenobaenus*	CMR, MR, age ratio	Bayesian	Closed	19	–	Statistics	Site fidelity, Productivity	Cave et al. (2010)
Simulation study mimicking Hoopoes *Upupa epops*	C, CMR, (P)	Bayesian	Closed	–	–	Evaluation	(Productivity)	Abadi et al. (2010a)
Little Owl *Athene noctua*	C, CMR, P	Bayesian	Open	26	Prey abundance	Statistics	Immigration	Abadi et al. (2010b)
Eagle Owl *Bubo bubo*	C, P, age-at-death data, telemetry	Bayesian	Open	20	–	Statistics, ecology, conservation	Immigration, adult survival	Schaub et al. (2010)
Blackbird *Turdus merula*	C, P, MR	Bayesian	Closed	20	Weather Spatial variation	Ecology	Components of productivity	Robinson et al. (2010)
Black-headed Gull *Larus ridibundus*	C, MS	Frequentist	Open	20	–	Statistics, ecology	Site-specific productivity, dispersal at regional scale	Péron et al. (2010)
Great Cormorant *Phalacrocorax carbo*	C, MS	Frequentist	Closed	14	–	Statistics, ecology	Productivity	McCrea et al. (2010)
Black Bear *Ursus americanus*	CMR, age-at-harvest data, telemetry	Frequentist	Closed	28	Hunting effort Food index	Statistics, ecology	Abundance	Fieberg et al. (2010)
Bighorn Sheep *Ovis canadensis*	C, CMR, telemetry	Bayesian	Closed	28	Population density Snow depth Precipitation	Statistics, ecology	Productivity	Johnson et al. (2010)

The studies are ordered chronologically

[a] Data types: *C* population counts, *I* population indices, *CMR* capture–recapture data, *MR* mark-recovery data, *MS* multi-state capture–recapture data, *P* productivity data, *NS* nest survival

[b] Focus: *Statistics* development of new models; *Evaluation* model evaluation; *Ecology* understanding population dynamics; *Conservation* identifying causes of decline

Discussion and future research directions

The development and application of integrated population models is an active field of research. This is certainly a consequence of the many advantages (estimation of demographic parameters without explicit data; high modelling flexibility) that these models have and the new possibilities they provide. Other advantages that we have not addressed here at length are that the estimates of demographic parameters have improved precision (Besbeas et al. 2002; Gauthier et al. 2007; Abadi et al. 2010a), which increases the power to detect temporal variation between them and environmental covariates. Missing data and unequal temporal coverage of the different datasets can easily be dealt with (Brooks et al. 2004; Schaub et al. 2007). Moreover, it is straightforward to get predictions of future population sizes with full propagation of parameter uncertainty (Fonnesbeck and Conroy 2004; Gauthier et al. 2007; Tavecchia et al. 2009).

There is certainly a need to further develop integrated population models. We see several possible directions. First, there is currently no goodness-of-fit test available for these models, thus one does not know how well the models are fitting the data and whether basic model assumptions are fulfilled. A practical, but not fully satisfying, solution is the application of existing goodness-of-fit tests for some of the individual datasets. An informal test currently seems to consist of the plot of the observed and the predicted population trajectory and of the visual evaluation of their similarity. Second, more research is needed to develop models that include density dependence (Baillie et al. 2009). Conceptually, this should be easy to do and would allow the evaluating of whether or not there is density-dependent regulation of the population and through which demographic mechanism it is imposed. Third, more research is needed for model selection and to render model selection approachable for ecologists. In these complex models, there are many possible forms of uncertainty, and we should be able to explore this. While in the frequentist framework model ranking based on AIC is possible (Tavecchia et al. 2009), techniques like reversible jump MCMC have been in use in the Bayesian framework (King et al. 2008; Baillie et al. 2009). Methods that account for the uncertainty in the underlying population model (matrix projection model) should also be developed (see, e.g., Thomas et al. 2005). Fourth, we also expect to see in future integrated population models which use count data that allow explicit estimation of detection probability, i.e. that temporally and spatially repeated counts are modelled within the integrated population model. Fifth, parameter redundancy of integrated population models will be evaluated either using established methods (Catchpole and Morgan 1997; Gimenez et al. 2009b) or with newly developed tools. Sixth, the joint likelihood of integrated population models might be developed in such a way that it explicitly accounts for the dependence among datasets, such that the assumption of independence can be relaxed. Seventh, the integrated population models could be modified in such a way that integral (i.e. continuous) population models (Ellner and Rees 2006) instead of stage- or age-structured (i.e. categorical) population models are used to define the link between population size and demography. This would allow the quantifying of the impact of individual continuous traits on population dynamics. Finally, another avenue is to explore additional data types that contain demographic information and that may be linked to a population model.

The possibility to analyse demographic data with integrated population models may also have consequences for population monitoring. The possibility to estimate demographic data without explicit data using integrated population models is particularly appealing for monitoring because data collection for some demographic parameters can be very expensive. Therefore, it may be enough to collect data on only some parameters. However, there are limitations in this endeavour: generally, the precision of parameters for which no explicit data are sampled is low (Abadi et al. 2010a), and the estimators may not always perform well when time-dependent models are applied (Tavecchia et al. 2009). With the current knowledge, it is premature to give general advice about how an efficient monitoring can be designed. A fruitful area of future research is therefore to study how data can be sampled such that the uncertainty in the parameter estimates is reduced to a tolerable minimum while at the same time minimising the monitoring costs.

Our review has revealed that few studies whose primary aim was an ecological question or conservation have so far applied integrated population models. On the one hand, this is a bit disappointing given the suitability for these methods to address these questions, in particular for limited data (Schaub et al. 2007; Véran and Lebreton 2008). On the other hand, this is not surprising given that integrated population models are relatively new and that fitting them may pose a challenge for many ecologists. The development of specific computer software might be useful to facilitate the access for ecologists, but we believe that a relatively profound understanding of the specialist software used to fit these models (e.g. WinBUGS) will always remain necessary when integrated population models are applied. This is somehow the price that has to be paid for this very flexible tool. But with more and more quantitatively trained ecologists, good teaching materials (e.g. McCarthy 2007; Royle and Dorazio 2008; Gimenez et al. 2009a; Kéry 2010; King et al. 2010) and the availability of program code and example data (e.g. Brooks et al. 2004;

Schaub et al. 2007; Fieberg et al. 2010; Johnson et al. 2010; Abadi et al. 2010b), we believe that these powerful models will become the standard modelling framework for inference about population ecology in scientific and applied studies.

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 modeling. J Appl Ecol 47:393–400

Anderson DR, Burnham KP, White GC (1994) AIC model selection in overdispersed capture–recapture data. Ecology 75:1780–1793

Baillie SR, Schaub M (2009) Understanding changes in bird populations - the role of bird marking. Ringing Migr 24:189–198

Baillie SR, Brooks SP, King R, Thomas L (2009) Using a state-space model of the British song trush *Turdus pilomenos* population to diagnose the causes of a population decline. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 541–561

Besbeas P, Freeman SN (2006) Methods for joint inference from panel survey and demographic data. Ecology 87:1138–1145

Besbeas P, Freeman SN, Morgan BJT, Catchpole EA (2002) Integrating mark-recapture–recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58:540–547

Besbeas P, Lebreton JD, Morgan BJT (2003) The efficient integration of abundance and demographic data. Appl Stat 52:95–102

Besbeas P, Freeman SN, Morgan BJT (2005) The potential of integrated population modelling. Austr N Z J Stat 47:35–48

Besbeas P, Borysiewicz RS, Morgan BJT (2009) Completing the ecological jigsaw. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 513–539

Borysiewicz RS, Morgan BJT, Hénaux V, Bregnballe T, Lebreton JD, Gimenez O (2009) An integrated analysis of multisite recruitment, mark–recapture–recovery and multisite census data. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 579–591

Brooks SP (2003) Bayesian computation: a statistical revolution. Philos Trans R Soc Lond A 361:2681–2697

Brooks SP, King R, Morgan BJT (2004) A Bayesian approach to combining animal abundance and demographic data. Anim Biodiv Conserv 27.1:515–529

Buckland ST, Newman KB, Thomas L, Koesters NB (2004) State-space models for the dynamics of wild animal populations. Ecol Mod 171:157–175

Buckland ST, Newman KB, Fernandez C, Thomas L, Harwood J (2007) Embedding population dynamics models in inference. Stat Sci 22:44–58

Burnham KP (1993) A theory for combined analysis of ring recovery and recapture data. In: Lebreton JD (ed) Marked individuals in the study of bird populations. Birkhäuser, Basel, pp 199–213

Caswell H (2001) Matrix population models. Construction, analysis, and interpretation. Sinauer, Sunderland, MA

Catchpole EA, Morgan BJT (1997) Detecting parameter redundancy. Biometrika 84:187–196

Caughley G (1994) Directions in conservation biology. J Anim Ecol 63:215–244

Cave VM, King R, Freeman SN (2010) An integrated population model from constant effort bird-ringing data. J Agric Biol Environ Stat 15:119–137

Clark JS, Björnstad ON (2004) Population time series: process variability, observation errors, missing values, lags, and hidden states. Ecology 85:3140–3150

De Valpine P, Hastings A (2002) Fitting population models incorporating process noise and observation error. Ecol Monogr 72:57–76

Ellner SP, Rees M (2006) Integral projection models for species with complex demography. Am Nat 167:410–428

Fieberg JR, Shertzer KW, Conn PB, Noyce KV, Garshelis DL (2010) Integrated population modeling of black bears in Minnesota: implications for monitoring and management. PloS ONE 5:e12114

Fonnesbeck CJ, Conroy MJ (2004) Application of integrated Bayesian modeling and Markov chain Monte Carlo methods to the conservation of a harvested species. Anim Biodiv Conserv 27.1:267–281

Freeman SN, Crick HQP (2003) The decline of the spotted flycatcher *Muscicapa striata* in the UK: an integrated population model. Ibis 145:400–412

Freeman SN, Robinson RA, Clark JA, Griffin BM, Adams SY (2007) Changing demography and population decline in the common starling *Sturnus vulgaris*: a multisite approach to integrated population monitoring. Ibis 149:587–596

Gauthier G, Besbeas P, Lebreton JD, Morgan BJT (2007) Population growth in snow geese: a modeling approach integrating demographic and survey information. Ecology 88:1420–1429

Gimenez O, Bonner SJ, King R, Parker RA, Brooks SP, LE Jamieson, Grosbois V, Morgan BJT, Thomas L (2009a) WinBUGS for population ecologists: Bayesian modeling using Markov Chain Monte Carlo methods. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 883–915

Gimenez O, Morgan BJT, Brooks SP (2009b) Weak identifiability in models for mark-recapture–recovery data. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. Environmental and ecological statistics, vol 3. Springer, New York, pp 1055–1067

Goodman D (2004) Methods for joint inference from multiple data sources for improved estimates of population size and survival rates. Mar Mam Sci 20:401–423

Hoyle SD, Maunder MN (2004) A Bayesian integrated population dynamics model to analyze data for protected species. Anim Biodiv Conserv 27.1:247–266

Jenouvrier S, Barbraud C, Weimerskirch H (2003) Effects of climate variability on the temporal population dynamics of southern fulmars. J Anim Ecol 72:576–587

Johnson HE, Mills LS, Wehausen JD, Stephenson TR (2010) Combining ground count, telemetry, and mark-resight data to infer population dynamics in an endangered species. J Appl Ecol 47:1083–1093

Kéry M (2010) Introduction to WinBUGS for ecologists—a Bayesian approach to regression, ANOVA, mixed models and related analyses. Academic, Burlington

King R, Brooks SP, Mazzetta C, Freeman SN, Morgan BJT (2008) Identifying and diagnosing population declines: a Bayesian assessment of lapwings in the UK. Appl Stat 57:609–632

King R, Morgan BJT, Gimenez O, Brooks SP (2010) Bayesian analysis for population ecology. Chapman & Hall, Boca Raton

Lande R (1993) Risks of population extinction from demographic and enviornmental stochasticity and random catastrophes. Am Nat 142:911–927

Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypothesis using marked animals: a unified approach with case studies. Ecol Monogr 62:67–118

Lebreton JD, Morgan BJT, Pradel R, Freeman SN (1995) A simultaneous survival rate analysis of dead recovery and live recapture data. Biometrics 51:1418–1428

Maunder MN (2004) Population viability analysis based on combining Bayesian, integrated, and hierarchical analyses. Acta Oecol 26:85–94

McCarthy MA (2007) Bayesian methods for ecology. Cambridge University Press, Cambridge

McCrea RS, Morgan BJT, Gimenez O, Besbeas P, Lebreton J-D, Bregnballe T (2010) Multi-site integrated population modelling. J Agric Biol Environ Stat 15:539–561

McCullagh P, Nelder JA (1989) Generalized linear models. Chapman & Hall, London

Millar RB, Meyer R (2000) Bayesian state-space modeling of age-structured data: fitting a model is just the beginning. Can J Fish Aquat Sci 57:43–50

Newton I (1998) Population limitation in birds. Academic, London

Norris K (2004) Managing threatened species: the ecological toolbox, evolutionary theory and declining-population paradigm. J Appl Ecol 41:413–426

Peach WJ, Siriwardena GM, Gregory RD (1999) Long-term changes in over-winter survival rates explain the decline of reed buntings *Emberiza schoeniclus* in Britain. J Appl Ecol 36:798–811

Péron G, Crochet P-A, Doherty PF, Lebreton JD (2010) Studying dispersal at the landscape scale: efficient combination of population surveys and capture–recapture data. Ecology 91:3365–3375

Reynolds TJ, King R, Harwood J, Frederiksen M, Harris MP, Wanless S (2009) Integrated data analysis in the presence of emigration and mark loss. J Agric Biol Environ Stat 14:411–431

Robinson RA, Green RE, Baillie SR, Peach WJ, Thomson DL (2004) Demographic mechanisms of the population decline of the song trush *Turdus philomelos* in Britain. J Anim Ecol 73:670–682

Robinson RA, Baillie SR, King R (2010) Population processes in European blackbirds *Turdus merula*: a state–space approach. J Ornithol. doi:10.1007/s10336-010-0612-y

Royle JA, Dorazio RM (2008) Hierarchical modeling and inference in ecology. Academic, Amsterdam

Schaub M, Pradel R, Lebreton JD (2004) Is the reintroduced white stork (*Ciconia ciconia*) population in Switzerland self-sustainable? Biol Conserv 119:105–114

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 21:945–955

Schaub M, Aebischer A, Gimenez O, Berger S, Arlettaz R (2010) Massive immigration balances high human induced mortality in a stable eagle owl population. Biol Conserv 143:1911–1918

Schnute JT (1994) A general framework for developing sequential fisheries models. Can J Fish Aquat Sci 51:1676–1688

Sibly RM, Hone J (2002) Population growth rate and its determinants: an overview. Philos Trans R Soc Lond B 357:1153–1170

Siriwardena GM, Baillie SR, Wilson JD (1999) Temporal variation in the annual survival rates of six granivorous birds with contrasting population trends. Ibis 141:621–636

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

Thomas L, Buckland ST, Newman KB, Harwood J (2005) A unified framework for modelling wildlife population dynamics. Austr N Z J Stat 47:19–34

Thomson DL, Baillie SR, Peach WJ (1997) The demography and age-specific annual survival of song thrushes during periods of population stability and decline. J Anim Ecol 66:414–424

Trenkel VM, Elston DA, Buckland ST (2000) Fitting population dynamics models to count and cull data using sequential importance sampling. J Am Stat Ass 95:363–374

Véran S, Lebreton JD (2008) The potential of integrated modelling in conservation biology: a case study of the black-footed albatross (*Phoebastria nigripes*). Can J Stat 36:85–98

White GC, Lubow BC (2002) Fitting population models to multiple sources of observed data. J Wildl Manag 66:300–309

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic, San Diego