Estimation of immigration rate using integrated population models

Fitsum Abadi[1]*, Olivier Gimenez[2], Bruno Ullrich[3], Raphaël Arlettaz[1,4] and Michael Schaub[1,4]

[1]*Institute of Ecology and Evolution – Conservation Biology, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland;* [2]*Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, 1919 Route de Mende, 34293 Montpellier Cedex 5, France;* [3]*Zeller Strasse 15, 73110 Hattenhofen, Germany;* [4]*Swiss Ornithological Institute, CH-6024 Sempach, Switzerland*

Summary

1. The dynamics of many populations is strongly affected by immigrants. However, estimating and modelling immigration is a real challenge. In the past, several methods have been developed to estimate immigration rate but they either require strong assumptions or combine in a piecewise manner the results from separate analyses. In most methods the effects of covariates cannot be modelled formally.

2. We developed a Bayesian integrated population model which combines capture–recapture data, population counts and information on reproductive success into a single model that estimates and models immigration rate, while directly assessing the impact of environmental covariates.

3. We assessed parameter identifiability by comparing posterior distributions of immigration rates under varying priors, and illustrated the application of the model with long term demographic data of a little owl *Athene noctua* population from Southern Germany. We further assessed the impact of environmental covariates on immigration.

4. The resulting posterior distributions were insensitive to different prior distributions and dominated by the observed data, indicating that the immigration rate was identifiable. Average yearly immigration into the little owl population was 0·293 (95% credible interval 0·183–0·418), which means that ca 0·3 female per resident female entered the population every year. Immigration rate tended to increase with increasing abundance of voles, the main prey of little owls.

5. *Synthesis and applications*. The means to estimate and model immigration is an important step towards a better understanding of the dynamics of geographically open populations. The demographic estimates obtained from the developed integrated population model facilitate population diagnoses and can be used to assess population viability. The structural flexibility of the model should constitute a useful tool for wildlife managers and conservation ecologists.

Key-words: *Athene noctua*, Bayesian, capture–recapture, identifiability, population counts, reproductive success, survival, state-space model

Introduction

Immigration and emigration are important demographic processes which can have a strong impact on population dynamics (Ward 2005; Lampila *et al.* 2006; Schaub *et al.* 2006; Grøtan *et al.* 2009). In an attempt to fully understand the dynamics of a population, it is therefore important to be able to estimate and include these parameters into a population dynamical model. Knowledge of immigration and emigration is also essential for identifying whether a local population is a source or a sink (Pulliam 1988; Peery, Becker & Beissinger 2006) which may be of interest in conservation and other management applications.

Existing methods can be used to estimate emigration rate by combining capture–recapture and dead recovery data (Burnham 1993, Reynolds *et al.* 2009). In most population studies where emigration cannot be directly estimated, it is accounted for in the estimate of apparent survival (function of mortality and permanent emigration), which is obtained from capture–recapture data (Lebreton *et al.* 1992). Unlike emigration, the

*Correspondence author. E-mail: fitsum.gebreselassie@iee.unibe.ch

rate of immigration is notoriously difficult to estimate. Consequently, relatively little is known about its role in population dynamics and about factors affecting its variability across space and time. The problem of estimating immigration stems from the difficulty of making direct observations: immigrants can only be distinguished from established individuals if all members of one or both groups are marked. Only in very intense studies, in which all individuals in the population are marked each year, can immigration be readily inferred from the number of unmarked individuals. However, such an intensive capture effort is very rarely applied because in natural conditions complete capture is usually not possible (Møller 2002; Grøtan *et al.* 2009). To date, several approaches have been used to deal with this problem. Key to all of them is the availability of capture–recapture data. However, some methods require additional demographic data which then need to be combined with capture–recapture data to estimate immigration rate.

Initially, immigration was inferred from direct observations of the movement of marked individuals between populations using multistate capture–recapture models (Spendelow *et al.* 1995). These methods solve the problem that not all individuals in the population are captured each year. However, immigrants can originate from very different populations that may be far away, and in practice it is almost impossible to mark individuals at all potential source populations. Therefore, this method usually results in an underestimation of the immigration rate.

Secondly, if all locally born individuals are marked each year and if a fraction of adults is captured each year, immigration can be estimated (O'Hara, Lampila & Orell 2009). Under this sampling protocol, it is clear that any unmarked individual in the population must be an immigrant, but owing to the imperfect capture of adults, it is not clear when they have immigrated. This latter problem can be solved by the application of the Jolly-Seber model (O'Hara *et al.* 2009). The critical requirement is that all young born in the population must be marked every year. This can only be achieved in rare situations.

Thirdly, immigration can be estimated if capture–recapture data are sampled under the robust design (Nichols & Pollock 1990). Here the population is sampled at least twice within short time each year, allowing the estimation of population sizes of young and adult individuals as well as their survival probabilities. By rearranging these quantities, the total number of immigrants can be estimated. The disadvantage of this approach is that this special sampling design is rarely applied due to the high sampling effort it requires.

Fourthly, immigration can be estimated if ordinary capture–recapture data of young and adults and data on the reproductive success are available (Cooch, Rockwell & Brault 2001; Peery *et al.* 2006; Schaub *et al.* 2006). Pradel's (1996) temporal symmetry model applied to the capture–recapture data of the adults allows estimating population growth rate (λ) or total recruitment (f, immigration plus local recruitment). From the capture–recapture data, apparent juvenile (ϕ_{ju}) and adult (ϕ_{ad}) survival probabilities are estimated (Lebreton *et al.* 1992). Apparent survival probability is the joint probability to

survive and to remain in the study population (i.e. a function of both mortality and permanent emigration). For species that start to reproduce in their first year of life, immigration rate (*imm*) can then be estimated either as $imm = f - b\phi_{ju}$ (Peery *et al.* 2006) or as $imm = \lambda - b\phi_{ju} - \phi_{ad}$ (Schaub *et al.* 2006), where b is the reproductive output. Similar calculations are possible for species with delayed start of reproduction (Cooch *et al.* 2001). Standard errors of the estimates are obtained using the delta method or by bootstrapping. Disadvantages of this method are that covariances between the parameters stemming from different data sets remain unknown and immigration rate is not treated as a model parameter but is derived. Therefore, this method is not flexible enough to test whether immigration is changing as a function of environmental covariates. Other drawbacks relate to the assumptions of the temporal symmetry model. If immediate trap response or transients occur, or if the initial capture probability is different than the recapture probability (typical in studies using colour bands), the estimated λ and f are strongly biased (Hines & Nichols 2002), and consequently immigration would be as well. Trap response and transients are frequent in capture–recapture studies (Perret *et al.* 2003; Schmidt, Feldmann & Schaub 2005; Schaub *et al.* 2009).

Fifthly, Grøtan *et al.* (2009) recently modelled the immigration process using a beta-binomial model based on population count data (i.e. the number of breeding birds). To estimate immigration they assumed that the number of available nest boxes in the study area is equal to the maximum number of immigrants. However, this assumption is only reasonable if very few birds breed in natural cavities. Further, they assumed that all adults unmarked or marked outside the study area are immigrants, which may not always be true.

From this overview, it is evident that a coherent method to estimate immigration is needed, which is applicable to the typical data that are sampled in geographically open populations. In this paper, we propose a new method to estimate immigration using integrated population models (Besbeas *et al.* 2002) that does not suffer from these problems. Integrated population models need population counts or indices of population size which are combined with other sources of data (typically capture–recapture data, data on fecundity) in one coherent model, with the advantage that otherwise non-estimable parameters can become estimable (Besbeas *et al.* 2002). This is possible because the population size contains information about all demographic processes in the population, and this information is extracted. The rationale behind our novel approach is similar. We show how to estimate immigration, and how to model immigration as a function of covariates. We illustrate the method with a case study on little owls *Athene noctua* Scopoli. Immigration in this population was previously estimated using the combination of capture–recapture and fecundity data (Schaub *et al.* 2006). Keeping the best structure for survival and recapture probabilities, as well as fecundity from the previous analysis, we tested whether immigration showed trends across time and whether immigration is affected by the density of voles *Microtus arvalis* Pallas, the main prey of the little owl.

Materials and methods

THE INTEGRATED POPULATION MODEL

The core of the integrated population model is a population projection model describing the development of the population sizes of each age class from 1 year to another. The projection matrix is parameterized with the demographic rates. Owing to the life history of little owl, we used a simple age-structured model (Caswell 2001) with two age classes (1 and 2 years or older) and a pre-breeding census. Let $N_{1,t}$ be the number of 1-year-old females at time t, $N_{2+,t}$ be the number of females older than 1 year at time t, $\phi_{ju,t}$ and $\phi_{ad,t}$ are the juvenile and adult (1 year or more) apparent survival probabilities of a female between time t and $t + 1$ respectively, b_t is fecundity at time t, and imm_t is the immigration rate, defined as the number of female immigrants in year $t + 1$ per female in the population in year t. Following Buckland *et al.* (2007), the expected number of females present in year $t + 1$ are given by

$$E(\boldsymbol{N}_{t+1}|\boldsymbol{N}_t) = E\left(\begin{bmatrix} N_{1,t+1} \\ N_{2,t+1} \end{bmatrix}\middle| \begin{bmatrix} N_{1,t} \\ N_{2+,t} \end{bmatrix}\right)$$
$$= \begin{bmatrix} \phi_{ju}\frac{b}{2} & \phi_{ju}\frac{b}{2} \\ \phi_{ad} + imm & \phi_{ad} + imm \end{bmatrix}_t \begin{bmatrix} N_{1,t} \\ N_{2+,t} \end{bmatrix}\middle| \begin{bmatrix} N_{1,t-1} \\ N_{2+,t-1} \end{bmatrix} \quad \text{eqn 1}$$

The model is female based, but fecundity (*b*) refers to the complete reproductive output. We assumed an even sex ratio of the fledglings and thus divide *b* by 2. Further, we assumed that immigrants add to the individuals that are at least 2 years old. However, results would be identical if immigrants were assumed to add to the 1-year-old individuals.

The capture–recapture data contribute to the estimation of survival, the data on reproductive success to the estimation of fecundity and the population count data to the estimation of the population sizes and all demographic rates. For each data set, a likelihood can be formulated and finally the complete integrated model is the joint likelihood of these different parts. In the following we describe briefly the different likelihoods (for details see Brooks, King & Morgan 2004; Schaub *et al.* 2007; Abadi *et al.* 2010).

We used the Cormack–Jolly–Seber (CJS) model to analyse the capture–recapture data (Lebreton *et al.* 1992). The frequency of individual encounter histories (*m*) follows a multinomial distribution with cell probabilities that are function of age-specific apparent survival (ϕ) and recapture probabilities (*p*). The formulation of the likelihood of this model ($L_{cr}(\boldsymbol{m} \mid \boldsymbol{\phi}, \boldsymbol{p})$) is straightforward and described in many papers (e.g. Lebreton *et al.* 1992).

We estimated fecundity (*b*) from the number of recorded reproducing females (*R*) and total number of offspring (*J*) produced by them in year *t*. We assumed that J_t follows a Poisson distribution with parameter written as a product of R_t and b_t, hence, $J_t \sim \text{Po}(R_t b_t)$. The likelihood of this model is denoted as $L_{rp}(\boldsymbol{J}, \boldsymbol{R} \mid \boldsymbol{b})$.

The population count data were analysed with a state-space model (Besbeas *et al.* 2002; de Valpine & Hastings 2002; Buckland *et al.* 2004). The state-space model is defined by two processes: a state process that describes the evolution of the true population sizes across time and an observation process which describes the observation of the true process. The state process is already defined by our population model (eqn 1), but here we introduce demographic stochasticity using Poisson and binomial distributions as

$$N_{1,t+1} \sim Po\big((N_{1,t} + N_{2+,t})(b_t/2)\phi_{ju,t}\big) \quad \text{eqn 2}$$

and

$$N_{2+,t+1} \sim \big(\text{Bin}((N_{1,t} + N_{2+,t}), \phi_{ad,t}) + \text{Po}((N_{1,t} + N_{2+,t})imm_t)\big)$$
$$\text{eqn 3}$$

The observation process is conditional on the state process. Here, only breeding females are counted, and we assume the counts to follow a Poisson distribution. That is, $y_t \sim \text{Po}(N_{1,t} + N_{2+,t})$, where y_t is the number of observed breeding females in year t. The likelihood of the state process is $L_{sy}(\boldsymbol{N} \mid \boldsymbol{\phi}, \boldsymbol{b}, \boldsymbol{imm})$, the likelihood of the observation process is $L_{ob}(\boldsymbol{y} \mid \boldsymbol{N})$ and the likelihood of the state-space model is the product of the two likelihoods.

LIKELIHOOD OF THE INTEGRATED MODEL

The likelihoods of the three data sets have some parameters in common, as displayed graphically by the directed acyclic graph (Fig. 1). Thus, by combining these data sources into a single analysis using an integrated population model, more information can be used to estimate demographic parameters. Under the assumption of independence between the three data sources, the joint likelihood for the combined data, which is a basis for inference, is obtained as the product of the likelihoods of the three data sources (Besbeas *et al.* 2002; Besbeas, Lebreton & Morgan 2003; Brooks *et al.* 2004; Abadi *et al.* 2010), thus

$$L_{joint}(\boldsymbol{m}, \boldsymbol{J}, \boldsymbol{R}, \boldsymbol{y}|\boldsymbol{N}, \boldsymbol{\phi}, \boldsymbol{b}, \boldsymbol{imm}, \boldsymbol{p}) = L_{cr}(\boldsymbol{m}|\boldsymbol{\phi}, \boldsymbol{p}) \times L_{rp}(\boldsymbol{J}, \boldsymbol{R}|\boldsymbol{b})$$
$$\times L_{ob}(\boldsymbol{y}|\boldsymbol{N}) \times L_{sy}(\boldsymbol{N}|\boldsymbol{\phi}, \boldsymbol{b}, \boldsymbol{imm}) \quad \text{eqn 4}$$

The calculation of the joint likelihood (eqn 4) relies on the assumption that the data are independent which is frequently violated in practice. However, a simulation study, which combined capture–recapture, population count and reproductive success data, showed that the violation of this assumption has only minimal impact on the accuracy of parameter estimates (Abadi *et al.* 2010). Since our model to estimate immigration uses the same kind of data as the simulation study by Abadi *et al.* (2010), the potential violation of the independence assumption is likely to have little impact on the estimators.
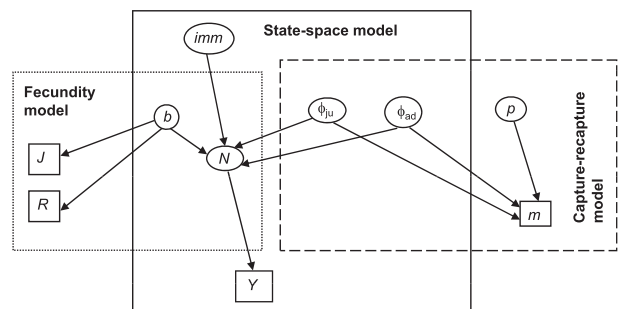

Fig. 1. Structure of the integrated population model: Directed acyclic graph of the integrated population model for the little owl population in Göppingen, showing that some demographic parameters are common to different sources of information. Estimated parameters are represented by circles and the 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 recorded reproducing females; *b*, fecundity; ϕ_{ju}, juvenile survival probability; ϕ_{ad}, adult survival probability; *imm*, immigration rate; *p*, recapture probability; *N*, population size.

CASE STUDY

Data

We used demographic data on little owls to illustrate the estimation and modelling of immigration rate using the integrated population model. The little owl is a small-sized owl preying mostly on voles in Central Europe and inhabiting semi-open landscapes. It is a cavity breeder and readily accepts nest boxes (Cramp 1985). We collected the demographic data near Göppingen (Southern Germany; 48°40′N, 9°38′E), where nest boxes were installed and the population was monitored from 1978 to 2003. Each year we checked all nest boxes several times, recorded reproductive success ($n = 353$ breeding attempts), marked nestlings with rings ($n = 659$) and captured and marked adults that were in the nest boxes at the time of the checks ($n = 73$). Thus, capture–recapture data, data on reproductive success and on population size (number of occupied nest boxes) were available. Further, we recorded a two-level index of vole abundance (high, low) each spring based on the number of vole hills and holes. For further details see Schaub *et al.* (2006).

PARAMETER ESTIMATION AND SENSITIVITY ANALYSIS

The joint likelihood of the model (eqn 4) is constructed based on data of females only. However, we had also capture–recapture data of males. These data were also included and modelled, but they contributed to the joint likelihood only, if at least one parameter was identical (e.g. if survival was not sex-specific), improving the precision of the estimates common in both sexes.

We performed a Bayesian analysis of the integrated population models, which gives us more flexibility regarding model assumptions to form the likelihood of the population count data (Brooks *et al.* 2004; Schaub *et al.* 2007; King *et al.* 2008). To estimate demographic rates, we mimic as closely as possible the Schaub *et al.* (2006) analysis. We used constant immigration and those structures of the other demographic rates that turned out to be best in Schaub *et al.* (2006). We assessed whether the integrated analysis provides an identifiable estimate of immigration by comparing the prior and posterior distributions for immigration rate. In the Bayesian context, a parameter is considered to be identifiable if the posterior distribution differs from the prior distribution (Garrett & Zeger 2000; Gimenez, Morgan & Brooks 2009). If this is so the data do contain information about the parameter under scrutiny. We specified four prior distributions for immigration (three with uniform distributions between 0 and 5 ($U(0, 5)$), 0 and 2 ($U(0, 2)$), and 0 and 1 ($U(0, 1)$); one with a normal distribution with mean 0 and variance 0·25 truncated to the values between 0 and 2 ($N(0,0·25)I(0,2)$)). The first three prior distributions assumed that immigration is equally likely within the given ranges, whereas the normal prior assumed that low immigration is much more likely than high immigration. We regard the $U(0,5)$ prior as non-informative and the normal prior as the most informative. If immigration is identifiable, we expect the choice of the prior to have only limited impact on the posterior distribution. Thirdly, we estimated the demographic rates while using only capture–recapture and population count data, but not the data on fecundity. This allows us to check whether estimation is still possible in the case where explicit data for two different parameter types (immigration and fecundity) are lacking.

In all cases we specified the following prior distributions for the other model parameters than immigration. Since survival and recapture probabilities are modelled as linear functions of time and sex on a logit scale, we assigned a $N(0,10^4)$ prior to the regression coefficients. Further we used a $U(0, 5)$ prior for fecundity and a $N(0,10^4)$ truncated to positive values on the age-specific population sizes.

Unfortunately, no measure of goodness-of-fit (GOF) is available yet for integrated population models. Therefore, the best thing to do is to assess the GOF of single data sets. Schaub *et al.* (2006) assessed the GOF of the CJS model and found no lack of fit.

MODELLING IMMIGRATION AS A FUNCTION OF COVARIATES

In this section, we show how the integrated model can be used to model immigration rate as a function of environmental covariates and to make inference. In fact, the inclusion of a covariate is easy and straightforward. We modelled the log of immigration rate as a linear function of covariates using the linear relationship

$$\log(imm_t) = \beta_0 + \sum_{j=1}^{v} \beta_j X_{jt} \qquad \text{eqn 5}$$

where X_{jt} is the value of the jth covariate at the tth sampling occasion, the β's are regression coefficients. We used a $N(0,10^4)$ prior for the unknown regression parameters (β's).

Either classical hypothesis tests (i.e. evaluating whether β is different from 0) or model selection (e.g. using the deviance information criterion DIC (Spiegelhalter *et al.* 2002; van der Linde 2005) or performing reversible jump MCMC (Green 1995)) can be used to make inference about immigration in relation to covariates. Here, we compared several models based on the DIC. We used the DIC as a preliminary tool for comparing different models. The models that we compared had all different structures for immigration, but the same structure for all other model parameters.

An extensive model selection by Schaub *et al.* (2006) revealed that survival was best estimated using the same linear time trend for both age classes with an additional sex effect, that recapture probability was best modelled time-specific with an additive sex effect and that reproductive success was different in each year (model denoted as $\phi_{age2 + T + sex}$, $p_{t + sex}$, b_t). We defined six different models for immigration. The first model considers that immigration is time-dependent (imm_t). Secondly, we assumed that immigration had systematically changed with time, either linearly (imm_T) or following a quadratic function ($imm_{T + T^2}$). The fourth model assumes the immigration to depend on the abundance of voles (imm_v). We expect high vole abundance to have a positive impact on immigration. The next model assumes that immigration is constant across time ($imm_.$). Finally, to test whether immigration was important at all, we considered a model where we assumed no immigration (imm_0).

MODEL IMPLEMENTATION

All analyses were carried out using WinBUGS (Lunn *et al.* 2000) called from R 2·7 ·1 (R Development Core Team 2008) with the R2WinBUGS package (Sturtz, Ligges & Gelman 2005). WinBUGS performs Markov chain Monte Carlo (MCMC) techniques to sample from the posterior distribution of each parameter. We used the Brooks and Gelman diagnostic ($\hat{R}$) to assess the convergence of the MCMC simulations (Brooks & Gelman 1998). Initial trials with three independent chains showed that convergence ($\hat{R} < 1·01$) was obtained after 20 000 iterations with a burn-in period of 10 000 iterations. For all analyses we therefore ran single chains of 200 000 iterations, of which the first 100 000 were removed as a burn-in period and thinned such that every 100th observation was retained. The R and

WinBUGS code for the model which included a covariate are provided in Appendix S1 (Supporting Information).

Results

PARAMETER ESTIMATION AND SENSITIVITY ANALYSIS

The estimated immigration rate from our simplest constant model was 0·293 with 95% credible interval (CRI) 0·183–0·418 (under prior $U(0, 2)$). This means that for every breeding female in year t, about 0·3 female had immigrated in year $t + 1$. As can be seen in Fig. 2, the prior distributions had little effect on the posterior distribution of immigration rate, indicating that the posterior distribution was highly dominated by the observed data. The posterior means and 95% CRIs for the immigration rate were almost the same under different priors [posterior mean (95% CRI): under $U(0, 5)$: 0·290 (0·173–0·417); under $U(0, 1)$: 0·292 (0·183–0·415); under $(N(0,0·25)I(0,2))$: 0·296 (0·175–0·425)]. Consequently, immigration is identifiable using the integrated population model. The estimated immigration rate obtained from the analysis without
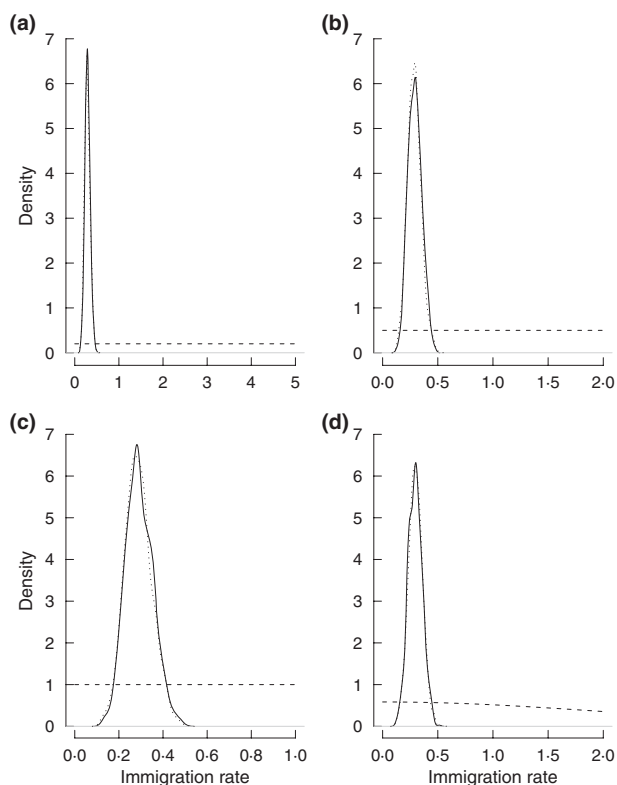
fecundity data was 0·288 with 95% credible interval 0·170–0·425, which was virtually identical to the one obtained from the analysis which did include fecundity data. Also in this case was the posterior distribution of immigration insensitive to the prior distributions, indicating again that immigration was identifiable (Fig. 2). However, the wide credible intervals indicated uncertainty in the estimates of immigration in all cases.

The annual point estimates of all demographic rates were very similar regardless of whether they stem from the integrated population models or from the 'two steps' approach (Schaub *et al.* 2006; see Appendix S2, Supporting Information).

MODELLING IMMIGRATION WITH COVARIATES AND MODEL SELECTION

Model selection of immigration rate showed that the model where immigration was a function of vole abundance was favoured by DIC (Table 1). However, the difference in DIC to the next best model with constant immigration rate was small. Thus, there is uncertainty whether immigration really depended on vole abundance. This is also reflected in the posterior distribution of the vole abundance effect which includes zero (Fig. 3). The models where immigration changes deterministically across time and where immigration rate was different in each year were clearly lower ranked. In particular, the model where no immigration occurred was ranked as the worst, indicating that immigration was important in this little owl population. The estimates of all demographic rates with associated 95% credible intervals based on the best model are shown in Fig. 4.

Discussion

Immigration to geographically open population is often an important demographic component to population growth (Lampila *et al.* 2006; Peery *et al.* 2006; Schaub *et al.* 2006; Grøtan *et al.* 2009), but immigration rate is one of the most difficult

Fig. 2. Parameter identifiability: Prior (dashed line) and posterior (solid and dotted lines for models with and without fecundity data, respectively) distributions of the immigration rate, estimated using the integrated population model, where immigration rate was constrained to be constant over time (model denoted as $\phi_{age2 + T + sex}$, $p_{t + sex}$, b_t, $imm_.$, Table 1). A $U(0,5)$ (Panel a), a $U(0,2)$ (Panel b), a $U(0,1)$ (Panel c) and a $N(0,0·25)I(0,2)$ (Panel d) prior are used. Note that the posterior for the model without fecundity is hardly visible, because it matches almost completely the posterior of the model including fecundity.

Table 1. Model selection results for different patterns of immigration rate (*imm*) of little owls using integrated population models

Immigration model	Deviance	pD	ΔDIC
imm_v	583·44	63·95	0·00
$imm_.$	573·99	73·70	0·31
imm_t	579·71	72·34	4·66
imm_T	584·51	67·86	4·98
$imm_{T + T^2}$	584·65	66·77	5·03
imm_0	588·18	76·82	17·61

The model deviance, the effective number of parameters (pD), and the DIC difference (ΔDIC) between the current model and the best model are provided. The model subscript t stands for year-specific rates, T and T^2 denote a linear and quadratic trend over time, respectively, 0 denotes no immigration, '.' denotes constancy over time, and v denotes a vole abundance effect. For survival ($\phi_{age2 + T + sex}$), recapture ($p_{t + sex}$) and fecundity (b_t) the same structure was used for all models.
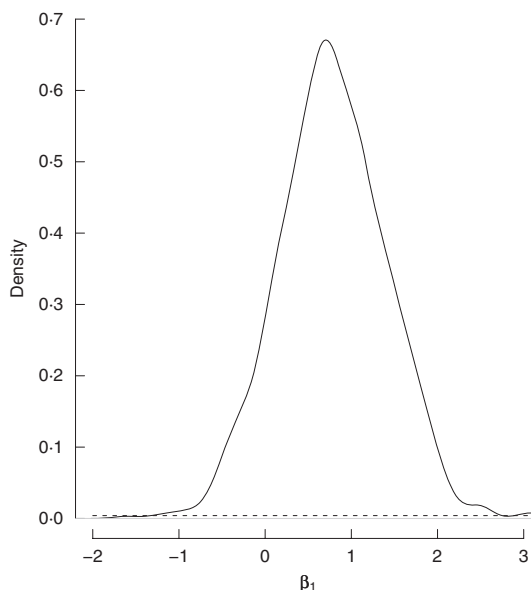
Fig. 3. Resource-dependence of immigration rate: Prior (dashed) and posterior (solid line) density of the vole abundance effect (β_1) on immigration rate. A $N(0,10000)$ prior for β_1 was used. The probability that β_1 was positive is 0·900.
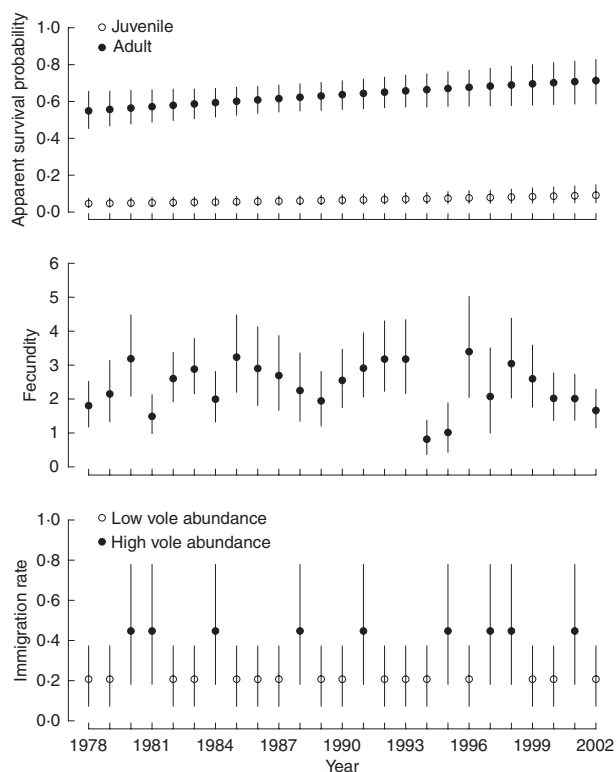

Fig. 4. Parameter estimates: Annual apparent survival probabilities, fecundity, and immigration rates of little owls females based on the best integrated population model (model $\phi_{age2+T+sex}$, p_{t+sex}, b_t, imm_v, Table 1). Points are the posterior means, vertical lines indicate the 95% credible intervals.

demographic rates to estimate. Here we propose an extension of an integrated population model to estimate immigration rate using population count, capture–recapture and fecundity

data. We show that immigration is estimable and can furthermore be modelled as a function of environmental covariates. This is a step forward to model the dynamics of geographically open populations.

A key advantage of integrated population models in general is that they allow estimating demographic parameters for which no explicit data are sampled (Besbeas *et al.* 2002, Abadi *et al.* 2010). This is possible because the population count data contain information about all demographic processes in the population and because this information is extracted with the inclusion of independent data about some of these processes. Previously developed integrated population models showed that fecundity can be estimated if count data are combined either with ring-recovery data (Besbeas *et al.* 2002; Brooks *et al.* 2004) or with capture–recapture data (Abadi *et al.* 2010). Here, we show that immigration can be estimated, if count data are combined with capture–recapture data. The inclusion of data on fecundity contributes little to the estimation of immigration rate. This might be different in a situation where survival from birth to recruitment as breeder is high. However, it is very encouraging that the integrated population model has the power to estimate two demographic parameters (i.e. immigration rate and fecundity) for which no explicit data are available.

Our analyses revealed that the posterior distributions of immigration were not very sensitive to different priors. Because the posterior distribution generally becomes independent of the prior distribution when the data are highly informative, our result indicated that the immigration rate could be estimated based on the information from the observed data.

Other approaches to estimate immigration have been proposed and we reviewed them briefly in the introduction. The integrated population model has several advantages over existing methods. The first advantage is that the model allows immigration to be modelled as a function of covariates. Often, the interest is not just in a point estimate, but on testing biological hypotheses about immigration. This is only possible if covariates can be included. Secondly, our approach does not rely on strong assumptions, as other approaches do. In particular there is no requirement for all unmarked individuals that are born in the population to be marked, that the population is counted without error or that all unmarked individuals are assumed to be immigrants. In our view, relaxing these assumptions provides much more realism to the data than has been possible previously. Yet, the integrated population model is also based on some assumptions. The different data must be independent from each other. A simulation study in a slightly different context has shown that the violation of this assumption has limited impact only on parameter accuracy (Abadi *et al.* 2010). Furthermore, the assumptions of the capture–recapture model (Lebreton *et al.* 1992) including identity of estimates of survival and recapture among individuals must hold. However, some frequent reasons for the violation of these assumptions such as occurrence of transients (Pradel 1996), trap-dependence (Pradel 1993) or non-random temporary emigration (Schaub *et al.* 2004) can be relaxed by specifying a different capture–recapture model.

This can be included into the integrated population model. In contrast, for methods that rely on the reverse-time modelling of capture–recapture data (Peery *et al.* 2006; Schaub *et al.* 2006) this is not possible.

The integrated population model yielded less precise estimates of the immigration rate in comparison to the 'two steps' approach (Schaub *et al.* 2006; see Appendix S2 in Supporting Information). However, the precision from the 'two steps' approach is based on the assumption that the covariances between all parameter estimates are zero. Presumably this was violated resulting in an overly optimistic precision. In contrast, the integrated model provides a more realistic picture of the uncertainty in all the estimates. In integrated population models, the precision of fecundity when no explicit fecundity data were included was much lower than when data on fecundity were included (see Appendix S2, Supporting Information, and Abadi *et al.* 2010). This is likely to be generally true. Thus, as there are no explicit data on immigration included, the low precision of the estimates could be expected. To increase precision, explicit data on immigration must be included. This could be achieved if capture–recapture data are sampled using the robust design, allowing estimates of immigration rates to be obtained (Nichols & Pollock 1990). The robust design model could then be included into the integrated population model. In the Bayesian context, one could also consider more informative priors derived from another population or from a closely related species to improve the precision of the estimates of immigration rate.

We foresee several areas for further developing the novel analytical framework presented here. First, the model could be modified to include random effects. This is of importance when the temporal variance of demographic rates and the temporal covariances between them must be estimated. Such estimates are crucial for assessing the contribution of demographic rates to the variation of the population growth rate (Horvitz, Schemske & Caswell 1997; Burnham & White 2002; Loison *et al.* 2002). Moreover, a hierarchical formulation is useful to model the effect of covariates on demographic rates, relaxing the assumption that the complete variation in the demographic rates over time is determined by the covariates alone. Also, a random-effects formulation for year-specific parameters is more parsimonious than explicitly estimating one independent effect for every year. Secondly, the state-space model could be extended to account for overdispersion by specifying beta-binomial or negative binomial distributions for the state process (King *et al.* 2008). Thirdly, the model could be adapted to different life histories, such as a life history with delayed maturity or with explicit inclusion of both sexes. It may then possible to estimate age and/or sex specific immigration rates.

In conclusion, our newly extended integrated population model will be a useful tool to estimate immigration rate in a geographically open population permitting a deeper understanding of the dynamics of such populations and to help conservation biologists conducting population diagnoses (e.g. population viability analysis) and proposing efficient management actions.

References

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

Besbeas, P., Freeman, S.N., Morgan, B.J.T. & Catchpole, E.A. (2002) Integrating mark–recapture–recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*, **58**, 540–547.

Besbeas, P., Lebreton, J.D. & Morgan, B.J.T. (2003) The efficient integration of abundance and demographic data. *Applied Statistics*, **52**, 95–102.

Brooks, S.P. & Gelman, A. (1998) General methods for monitoring convergence of iterative simulation. *Journal of Computational and Graphical Statistics*, **7**, 434–455.

Brooks, S.P., King, R. & Morgan, B.J.T. (2004) A Bayesian approach to combining animal abundance and demographic data. *Animal Biodiversity and Conservation*, **27**, 515–529.

Buckland, S.T., Newman, K.B., Fernandez, C., Thomas, L. & Harwood, J. (2007) Embedding populaton dynamics models in reference. *Statistical Science*, **22**, 44–58.

Buckland, S.T., Newman, K.B., Thomas, L. & Koesters, N.B. (2004) State-space models for the dynamics of wild animal populations. *Ecological Modelling*, **171**, 157–175.

Burnham, K.P. & White, G.C. (2002) Evaluation of some random effects methodology applicable to bird ringing data. *Journal of Applied Statistics*, **29**, 245–264.

Burnham, K.P. (1993) A theory for combined analysis of ring recovery and recapture data. *Marked Individuals in the Study of Bird Population* (eds J.-D. Lebreton & P.M. North), pp. 199–213. Birkhauser Verlag, Basel, Switzerland.

Caswell, H. (2001) *Matrix population Models: Construction, Analysis, and Interpretation*, 2nd edn. Sinauer Associates, Inc., Sunderland, MA, USA.

Cooch, E., Rockwell, R.F. & Brault, S. (2001) Retrospective analysis of demographic responses to environmental change: a lesser snow goose example. *Ecological Monographs*, **71**, 377–400.

Cramp, S. (1985) *Handbook of the Birds of Europe, the Middle East and North Africa*. Oxford University Press, New York.

de Valpine, P. & Hastings, A. (2002) Fitting population models incorporating process noise and observation error. *Ecological Monographs*, **72**, 57–76.

Garrett, E.S. & Zeger, S.L. (2000) Latent class model diagnosis. *Biometrics*, **56**, 1055–1067.

Green, P.J. (1995) Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, **82**, 711–732.

Gimenez, O., Morgan, B.J.T. & Brooks, S.P. (2009) Weak identifiability in models for mark–recapture–recovery data. *Modeling Demographic Processes in Marked Populations* (eds D.L. Thomson, E.G. Cooch & M.J. Conroy), pp. 1055–1067. Springer Series: Environmental and Ecological Statistics, New York.

Grøtan, V., Sæther, B.-E., Engen, S., van Balen, J.H., Perdeck, A.C. & Visser, M.E. (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. *Journal of Animal Ecology*, **78**, 447–459.

Hines, J.E. & Nichols, J.D. (2002) Investigations of potential bias in the estimation of lambda using Pradel's (1996) model for capture–recapture data. *Journal of Applied Statistics*, **29**, 573–587.

Horvitz, C., Schemske, D.W. & Caswell, H. (1997) The relative "importance" of life-history stages to population growth rate: prospective and retrospective analyses. *Structured Population Models in Marine, Terrestrial, and Freshwater Systems* (eds S. Tuljapurkar & H. Caswell). Champan and Hall, New York.

King, R., Brooks, S.P., Mazzetta, C., Freeman, S.N. & Morgan, B.J.T. (2008) Identifying and diagnosing population declines: a Bayesian assessment of lapwings in the UK. *Applied Statistics*, **57**, 609–632.

Lampila, S., Orell, M., Belda, E. & Koivula, K. (2006) Importance of adult survival, local recruitment and immigration in a declining boreal forest passerine, the willow tit *Parus montanus*. *Oecologia*, **148**, 405–413.

Lebreton, J.-D., Burnham, K.P., Clobert, J. & Anderson, D.R. (1992) Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs*, **62**, 67–118.

Loison, A., Saether, B.E., Jerstad, K. & Rostad, O.W. (2002) Disentangling the sources of variation in the survival of the European dipper. *Journal of Applied Statistics*, **29**, 289–304.

Lunn, D.J., Thomas, A., Best, N. & Spiegelhalter, D. (2000) WinBUGS – A Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, **10**, 325–337.

Møller, P. (2002) North Atlantic Oscillation (NAO) effects of climate on the relative importance of first and second clutches in a migratory passerine bird. *Journal of Animal Ecology*, **71**, 201–210.

Nichols, J.D. & Pollock, K.H. (1990) Estimation of recruitment from immigration versus in situ reproduction using Pollock's robust design. *Ecology*, **71**, 21–26.

O'Hara, R.B., Lampila, S. & Orell, M. (2009) Estimation of rates of births, deaths, and immigration from mark–recapture data. *Biometrics*, **65**, 275–281.

Peery, M.Z., Becker, B.H. & Beissinger, S.R. (2006) Combining demographic and count-based approaches to identify source-sink dynamics of a threatened seabird. *Ecological Applications*, **16**, 1516–1528.

Perret, N., Pradel, R., Miaud, C., Grolet, O. & Joly, P. (2003) Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology*, **72**, 567–575.

Pradel, R. (1993) Flexibility in survival anaysis from recapture data: Handling trap-dependence. *Marked Individuals in the Study of Bird Population* (eds J.D. Lebreton & P.M. North), pp. 29–37. Birkhaüser Verlag, Basel, Switzerland.

Pradel, R. (1996) Utilization of capture–mark–recapture for the study of recruitment and population growth rate. *Biometrics*, **52**, 703–709.

Pulliam, H.R. (1988) Sources, sinks, and population regulation. *American Naturalist*, **132**, 652–661.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, T.J., King, R., Harwood, J., Frederiksen, M., Harris, M.P. & Wanless, S. (2009) Integrated data analysis in the presence of emigration and mark loss. *Journal of Agricultural Biological and Environmental Statistics*, **14**, 411–431.

Schaub, M., Gimenez, O., Schmidt, B.R. & Pradel, R. (2004) Estimating survival and temporary emigration in the multistate capture–recapture framework. *Ecology*, **85**, 2107–2113.

Schaub, M., Gimenez, O., Sierro, A. & Arlettaz, R. (2007) Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology*, **21**, 945–955.

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., Zink, R., Beissmann, H., Sarrazin, F. & Arlettaz, R. (2009) When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. *Journal of Applied Ecology*, **46**, 92–100.

Schmidt, B.R., Feldmann, R. & Schaub, M. (2005) Demographic processes underlying population growth and decline in *Salamandra salamandra*. *Conservation Biology*, **19**, 1149–1156.

Spendelow, J.A., Nichols, J.D., Nisbet, I.C.T., Hays, H., Cormons, G.D., Burger, J., Safina, C., Hines, J.E. & Gochfeld, M. (1995) Estimating annual survival and movement rates of adults within a metapopulation of Roseate Terns. *Ecology*, **76**, 2415–2428.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R. & van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **64**, 583–616.

Sturtz, S., Ligges, U. & Gelman, A. (2005) R2WinBUGS: A pacakge for running WinBUGS from R. *Journal of Statistical Software*, **12**, 1–16.

van der Linde, A. (2005) DIC in variable selection. *Statistica Neerlandica*, **59**, 45–56.

Ward, M.P. (2005) The role of immigration in the decline of an isolated migratory bird population. *Conservation Biology*, **19**, 1528–1536.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Data and codes (R and WinBUGS) used in this article.

Appendix S2. Detailed summaries of the annual estimates of demographic rates of little owl females from Göppingen using different data and models

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.