Chapter 9
Integrated Population Modelling

9.1 Introduction

In Chap. 5, we recommended that formulation of population dynamics models should be guided by aims to answer specific scientific questions or assess or predict the effects of management actions. Management actions might target a specific life stage. For example, we might ask "How does removing wetland plants (such as bulrush or cattail) that have started to cover ponds and reduce the amount of open water in a waterfowl breeding area affect reproductive success?" The consequences of actions, however, typically ripple throughout the entire population life history and effective management requires more detailed ecological study. This in turn requires information about demographic processes and abundances for *multiple* life stages to characterize the population dynamics.

Information at the population and individual levels is often simultaneously available in monitoring programmes of wildlife populations. For example one survey might be designed to provide information about survival for a specific life stage, another for reproductive success, and a third for total population abundance. Often the information from each survey is analysed in isolation, and the separate results are used to fill the elements of a Leslie or Lefkovitch population projection matrix (Chap. 2). However, the surveys may provide overlapping information about demographic processes or abundances for multiple life stages and analyses that utilize that overlap are likely to be more powerful and provide more information than multiple piecemeal analyses.

In this chapter, we discuss approaches that do utilize the overlapping information, namely integrated population modelling (IPM). We define IPM to be the fitting of a population dynamics model to two or more sources of data where (i) the fitting is done in a *single or simultaneous* stage, and (ii) each source provides information at either the population or individual level. A third common feature of IPM, but not necessary to our definition, is (iii) at least two sources provide overlapping information about one or more population processes. A review of applications of integrated population modelling is provided by Schaub and Abadi (2011).

K.B. Newman et al., *Modelling Population Dynamics: Model Formulation, Fitting and Assessment using State-Space Methods*, Methods in Statistical Ecology,
DOI 10.1007/978-1-4939-0977-3_9,

We emphasise *single or simultaneous* to distinguish IPM from multi-stage or sequential fitting procedures. An example of a multi-stage fitting procedure is to use one data set to estimate survival probabilities for a sequence of years, another to estimate reproductive success, and then a third to estimate annual abundance estimates, where the three sets of estimates are calculated independently of one another, and the three sets of estimates are then used to fit a population dynamics model. In contrast, an IPM analysis uses all three data sets in a single combined analysis to fit the population dynamics model with simultaneous estimation of survival, reproduction and abundance. Maunder (1998) discusses several advantages of IPMs over multi-stage (non-simultaneous) analyses and we highlight some of these advantages and others below.

An IPM analysis can account for correlations between survival, reproduction and abundances. For example, years of higher survival are often years of higher reproductive success. Parameters shared by multiple data sets can be estimated with greater precision, or at least the uncertainty in an estimate based on combined data sets can be more coherently estimated, and the sharing of parameters can overcome parameter redundancy problems (Sect. 5.2). The gray whale analysis in Sect. 6.4.2 is an example of an IPM where non-identifiability problems were mitigated to a degree; abundance estimates alone make separation of survival and reproduction rates difficult at best, but the addition of harvest data provided additional information about both sets of parameters and improved estimability.

The availability of multiple surveys can also influence model formulation (Sect. 5.1), allowing for finer temporal and spatial modelling of population dynamics when for example different surveys are taken at different times of the year.

In keeping with the theme of the book, we first show how state-space models can provide a unifying framework for readily integrating data from multiple sources. However, such a completely unified SSM approach to analysing multiple data sources can be technically intricate and daunting. Most of this chapter focuses on an alternative IPM approach, an approach we call a connected likelihood approach. This latter approach can be less technically demanding, where a SSM is combined with other non SSM models for different data sets. The connected likelihood approach builds naturally on methods for estimating survival probabilities (Chap. 7) and abundances (Chap. 8) over time. In Chap. 8, we focussed on methods for estimating animal abundance from mark-recapture data. However, population abundance measures or estimates can be provided by other kinds of data and surveys, such as line transect surveys or aerial counts for randomly selected plots. Such surveys in general are aimed at detecting trends, such as growth or decline, in total population numbers, often at large scale, e.g. abundances at a national level. We shall use the term *census* in a generic sense in this chapter for any procedure used for estimating the size of a population or of a predefined part of it from field data. General procedures for estimating the size of a population are discussed in Chap. 6.

The structure of this chapter is as follows. Sect. 9.2 discusses the single SSM approach to IPM mentioned previously. Section 9.3 describes the connected likelihood approach and includes a worked example. An approximation that greatly assists the connected likelihood model is given and illustrated in Sect. 9.4 and

technical issues that arise in forming the Kalman filter likelihood are considered in Sect. 9.5. Extending the connected likelihood approach to multi-site and multi-state models is the topic of Sect. 9.6. The chapter ends (Sect. 9.7) with a variety of extensions and issues particular to IPMs, including conditional Gaussian (i.e. normal) modelling (Sect. 9.7.1), Bayesian methods (Sect. 9.7.3), and goodness-of-fit (Sect. 9.7.2).

9.2 Integrated Modelling within an SSM Framework

As discussed in Chap. 5, the temporal, spatial and biological resolution of the available data constrains formulation of the state process model in terms of what state vector components and parameters can be estimated. When more than one survey or data set is available, flexibility in the formulation of the SSM may be greater than what is feasible with a single data set. Multiple surveys can provide information on population abundances distinguished by age, sex, maturity or life history stage. Multiple surveys can also provide information about different population dynamic processes such as survival and reproduction. Data from multiple surveys can be incorporated in the SSM framework in two different ways: either expand the number of components of the observation vector or increase the number of observation vectors.

When multiple surveys make measurements on the population at the same point in time, the observation vector is enlarged to include data from each survey. If the different surveys are measuring the same components of the state vector, the data are simply replicate measurements that may or may not be independent of each other (conditional on the state component) and in general have different variances. For example, a mark-recapture survey and a line transect survey might be carried out within a week of each other and the population is assumed to be relatively static during that week. Estimates of abundance for week t from the two surveys are calculated ($y_{mr,t}$ and $y_{lt,t}$ for mark-recapture and line transect respectively) with corresponding standard errors ($s_{mr,t}$ and $s_{lt,t}$). Assuming that both are unbiased estimates of total abundance N_t, and are independent of each other, a general expression for the observation model is the following.

$$\begin{bmatrix} y_{mr,t} \\ y_{lt,t} \end{bmatrix} \sim D\left(\boldsymbol{\mu}_t = \begin{bmatrix} N_t \\ N_t \end{bmatrix}, \Sigma = \begin{bmatrix} s^2_{mr,t} & 0 \\ 0 & s^2_{lt,t} \end{bmatrix}\right)$$

where D is an arbitrary bivariate distribution with expected value vector $\boldsymbol{\mu}_t$, and variance-covariance matrix Σ. In this example, the observations are taken to be derived quantities rather than the raw data (e.g. numbers marked and recaptured in the mark-recapture survey, numbers counted and distances to animals in the line transect survey). The raw data could be the observations, but this often requires considerably more complex distributional structures; e.g. the likelihood models

underlying distance sampling (Buckland et al. 2001). Knape et al. (2013) examined the impact of using derived quantities instead of raw data and found that the loss of information was slight in the cases considered (also see Sect. 5.1). Another scenario is that different management agencies carry out separate surveys for distinct, non-overlapping land units. For example, duck surveys are made at three different wildlife refuges at nearly the same point in time. The population total might be spatially partitioned by each refuge, where $N_{A,t}$, $N_{B,t}$ and $N_{C,t}$ are the abundances for refuges A, B and C, respectively. The state vector is expanded accordingly and the components of the observation vector are independent abundance estimates, denoted $y_{A,t}$, $y_{B,t}$, and $y_{C,t}$, which are matched with each refuge's population abundance. For example:

$$\begin{bmatrix} y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{bmatrix} \sim D\left(\boldsymbol{\mu}_t = \begin{bmatrix} n_{A,t} \\ n_{B,t} \\ n_{C,t} \end{bmatrix}, \Sigma = \begin{bmatrix} s^2_{A,t} & 0 & 0 \\ 0 & s^2_{B,t} & 0 \\ 0 & 0 & s^2_{C,t} \end{bmatrix}\right)$$

In the case of multiple surveys making measurements at different points in time and perhaps focussing on different population sub-processes, additional observation vectors could be inserted and paired with different state vectors. For a concrete example, we revisit the BRS model of Sects. 2.3 and 3.2.2. The BRS model had two states (immature and mature animals) and three sub-processes, survival (S), growth (i.e. maturation, R) and birth (B), occurring in that sequence. We assume that three separate surveys were carried out independently to estimate the parameters corresponding to these processes, namely ϕ_1, ϕ_2, π and ρ (see Eqs. (2.2), (2.5) and (2.6)). We also assume that a fourth survey was carried out just after the breeding season, giving an estimate of total abundance $(n_{1,t}, n_{2,t})$; thus these are *census* data as defined previously. For simplicity, assume that the animals sampled to estimate the survival probabilities are then followed throughout the subsequent growth and reproduction processes with perfect detectability, i.e. their growth and reproduction numbers are known without error. The observation sub-processes are modelled as follows.

$$\begin{aligned}
(y_{1,t-1}, y_{2,t-1}) &\sim D\left((n_{1,t-1}, n_{2,t-1}), \theta\right) \\
y_{1(s),1,t} &\sim \text{binomial}\left(n_{1,t-1}, p\phi_1\right) \\
y_{1(s),2,t} &\sim \text{binomial}\left(n_{2,t-1}, p\phi_2\right) \\
y_{2(r),2,t} &\sim \text{binomial}\left(y_{1(s),1,t}, \pi\right) + y_{1(s),2,t} \\
y_{3(b),1,t} &\sim \text{binomial}\left(y_{2(r),2,t}, \rho\right)
\end{aligned}$$

The terms $y_{1,t-1}$ and $y_{2,t-1}$ are the census data with D an arbitrary distribution reflecting the uncertainty in the census data with corresponding parameter θ, while p is the probability of capture for the survival study. If different animals were used for sampling the three sub-processes, then the population abundances at each point

in time would be substituted for $y_{1(s),1,t}$ and $y_{2(r),2,t}$ and two additional capture probabilities are added. For example, using the notation from Eqs. (3.25) and (3.26):

$$\begin{aligned}
(y_{1,t-1}, y_{2,t-1}) &\sim D\left((n_{1,t-1}, n_{2,t-1}), \theta\right) \\
y_{1(s),1,t} &\sim \text{binomial}\left(n_{1,t-1}, p_s\phi_1\right) \\
y_{1(s),2,t} &\sim \text{binomial}\left(n_{2,t-1}, p_s\phi_2\right) \\
y_{2'(r),2,t} &\sim \text{binomial}\left(u_{1(s),1,t}, p_r\pi\right) \\
y_{3(b),1,t} &\sim \text{binomial}\left(u_{2(r),2,t}, p_b\rho\right)
\end{aligned}$$

where p_s, p_r, and p_b denote the probability of sampling animals for the three subprocesses and $y_{2'(r),2,t}$ denotes the immature animals that just matured.

Section 9.3 presents an alternative approach to using data from multiple surveys in which the state-space model formulation is just used for a subset of one or more surveys and alternative formulations characterize data from other surveys. In particular Sect. 9.3 constructs different likelihoods for different surveys but two or more likelihoods have one or more parameters in common.

9.3 Integrated Modelling with Connected Likelihoods

Census data are naturally dependent upon animal survival and fecundity, and this observation motivates this section and much of the remainder of this chapter. When information from several sources is available for a particular species, it is natural to consider the extent to which the available types of survey data are compatible and how they can complement each other. In early work, the emphasis was on the former, and matching the different types of analyses was done in an *ad hoc* way, by comparing model-based and census-based population growth rates (Coulson et al. 2001) or by checking visually the similarity of model-based and census-based population trajectories (Kanyamibwa and Lebreton 1992). In this section, we provide a formal methodology for the simultaneous analysis of mark-recapture-recovery data and population information such as census data.

9.3.1 Data, Models and Integrated Modelling

We introduce integrated population modelling through an extensive example in which ring-recovery data from marked birds are combined with abundance data on the same species. However, the approach is quite general, as we shall see. Our example is described in Besbeas et al. (2002) and involves observations on lapwings *Vanellus vanellus* breeding in Britain. This is a species of conservation concern in the UK due to its dramatic decline in recent years and it has been placed onto

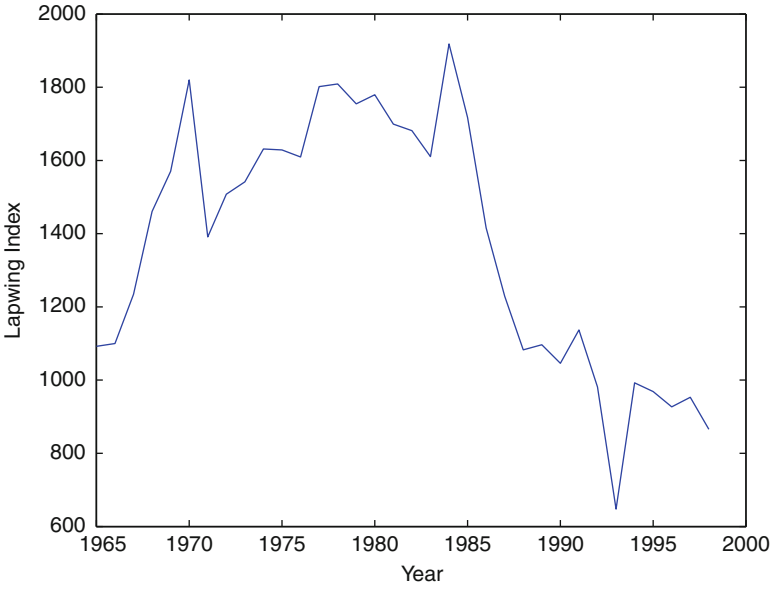

Fig. 9.1 The lapwing CBC index. The CBC index is sometimes given relative to a baseline year. As shown here, it is an estimate of the total population of lapwing territories for the set of sites included in the analysis

the Red List of species of highest conservation concern (Eaton et al. 2009). The census data we use are an index, derived from the common birds census (CBC) (Marchant et al. 1990). The CBC data are collected from specific survey sites by volunteers, and contain a large number of missing values. An index is constructed using a generalized linear model (ter Braak et al. 1994) which estimates site effects s_i (for site $i = 1, \ldots, S$) and year effects u_t (for year $t = 1, \ldots, T$) subject to an arbitrary constraint, e.g. $s_S = 0$. Annual index values y_t are then calculated as

$$y_t = \sum_{i=1}^{S} \exp(s_i + u_t), \quad t = 1, \ldots, T. \tag{9.1}$$

The resulting index estimates the relative abundance of the national breeding population from 1965 to 1998 inclusive, and it is plotted in Fig. 9.1.

The corresponding ring-recovery data provide the numbers of birds recovered dead in successive years after being ringed as chicks from 1963 to 1997. Note that these are national figures and are unlikely to share common individuals with the CBC data. The raw data are given in Besbeas et al. (2002), and as an illustration a subset of the recovery data is presented in Table 9.1.

We model the ring-recovery data using annual survival probabilities $\boldsymbol{\phi}$, with components which describe age-dependence, and a recovery probability κ, which

Table 9.1 Illustrative recovery sub-table for lapwings, taken from Besbeas et al. (2002)

Year of ringing	Number ringed	Year of recovery								No. of days below freezing (w)
		1991	1992	1993	1994	1995	1996	1997	1998	
1990	4170	12	3	3	2	1	0	2	0	12
1991	4314		9	4	6	1	0	1	0	12
1992	3480			18	3	1	2	0	1	9
1993	3689				6	5	2	2	1	6
1994	3922					12	4	6	0	3
1995	3591						7	5	1	18
1996	4488							7	0	10
1997	4339								5	0

The data show the numbers of British lapwings recovered dead in successive years after being ringed as nestlings, for birds ringed between 1990 and 1997 inclusive. Reporting probabilities of dead birds have been declining over time. By the 1990s, recovery rate had roughly halved relative to the 1960s. Also shown in the table is the number (w) of days in the year that a measure of Central England temperature fell below freezing. Note that year i ($1990 \leq i \leq 1997$) indicates the period of time from 1 April of year i to 31 March of year $(i + 1)$

Table 9.2 Cell probabilities $\{p_{ij}\}$ for a simple model for ring-recovery data, with no over-dispersion or time-variation.

Year of ringing	Year of recovery			
	1	2	3	4
1	$(1-\phi_1)\kappa$	$\phi_1(1-\phi_a)\kappa$	$\phi_1\phi_a(1-\phi_a)\kappa$	$\phi_1\phi_a^2(1-\phi_a)\kappa$
2		$(1-\phi_1)\kappa$	$\phi_1(1-\phi_a)\kappa$	$\phi_1\phi_a(1-\phi_a)\kappa$
3			$(1-\phi_1)\kappa$	$\phi_1(1-\phi_a)\kappa$

In this illustrative example, there are 3 years of ringing, 4 years of recovery, and three model parameters

is the probability of recovery and reporting of marked dead birds, and which varies over time. For the lapwings, there are two age classes of survival, corresponding to birds in their first year of life and older birds. Both of the annual survival probabilities are regressed on a single measure of winter severity w, using logistic regression. Thus we have $\text{logit}(\phi_1) = \beta_0 + \beta w$, for birds in their first year of life, with first-year annual survival probability of ϕ_1, and $\text{logit}(\phi_a) = \delta_0 + \delta w$, which applies to birds aged ≥ 1, with annual survival probability ϕ_a. In addition, the reporting probabilities for dead wild birds in Britain are generally found to be declining over time (Baillie and Green 1987), and so we set $\text{logit}(\kappa) = \nu_0 + \nu t$, where t measures year. See also McCrea et al. (2012b). We do not here consider over-dispersion, but if necessary that may be easily incorporated, for example by means of suitable additive random effects, as described by Barry et al. (2003), or through the use of beta distributions.

Morgan and Freeman (1989) and Freeman and Morgan (1992) describe more general models for ring-recovery data, involving additional age-dependence in survival, and/or time dependence in all parameters. See also Chap. 7. In order to write down the likelihood function, we introduce additional notation as follows. Let the number of birds ringed in year i be R_i, let the number recovered in year j, having been ringed in year i, be m_{ij}, and let the number which are not recovered from the year i cohort be $u_i = R_i - \sum_{j=i}^{c} m_{ij}$. Let the probability of recovery in year j given a bird was ringed in year i, corresponding to m_{ij}, be p_{ij} and let $q_i = 1 - \sum_{j=i}^{c} p_{ij}$ be the probability of non-recovery from the i^{th} cohort. The recovery probability p_{ij} is a product of annual survival probabilities (from year i to year $j-1$, mortality in year j, and recovery in year j (e.g. Table 9.2). A particular model for the data consists of a specification of the probabilities $p_{ij} \equiv p_{ij}(\boldsymbol{\phi}, \kappa)$ in terms of the model parameters. In order to display appropriate multinomial cell probabilities, we shall take a ring-recovery study with birds ringed as nestlings for $r = 3$ successive years, and recoveries recorded for the $c = 4$ years following the initial ringing. In the simplest case, these parameters are constant, and the recovery probabilities are given in Table 9.2. For each cohort, the probabilities of non-recovery are (1− the corresponding row totals). Then, provided the birds suffer independent fates, both within and between cohorts, the likelihood for the ring-recovery data is product-multinomial in form, with log-likelihood given below. Here and later in this chapter we shall for convenience suppress the dependence of likelihoods on data.

$$\log L_r(\boldsymbol{\phi}, \kappa) = \sum_{i=1}^{r} \sum_{j=i}^{c} m_{ij} \log p_{ij} + \sum_{i=1}^{r} u_i \log q_i, \tag{9.2}$$

where terms not depending on the parameters have been omitted. We shall refer to Eq. (9.2) as the ring-recovery log-likelihood below.

The census data are described by means of a state-space model involving a productivity measure p and measurement error variance σ^2, in addition to the survival probabilities. For the lapwings we set $\log(p) = \gamma_0 + \gamma t$, to allow for a decline in productivity; Besbeas et al. (2005) consider instead a change-point in productivity, corresponding to the start of the decline in numbers. The elements involved in forming the census likelihood are as follows.

Let N_{1t} and N_{at} denote, respectively, the underlying numbers of one-year-old female birds and female birds aged ≥ 2 years at time t, and define $\mathbf{N}_t = (N_{1t}, N_{at})'$ and $N_t = N_{1t} + N_{at}$. If we assume no sex effect on survival and that breeding starts at age 2, then a natural process model for the underlying population sizes would be

$$N_{1,t+1} \mid \mathbf{N}_t \sim \text{Poisson}(N_{at}\, p\phi_1)$$

and

$$N_{a,t+1} \mid \mathbf{N}_t \sim \text{binomial}(N_t, \phi_a)$$

where the parameter p now denotes the annual productivity of females per female. The random variables N_{1t} and N_{at} can be approximated by appropriate independent normal variables, resulting in the following model

$$\begin{bmatrix} N_{1,t+1} \\ N_{a,t+1} \end{bmatrix} = \begin{bmatrix} 0 & p\phi_1 \\ \phi_a & \phi_a \end{bmatrix} \begin{bmatrix} N_{1t} \\ N_{at} \end{bmatrix} + \begin{bmatrix} \delta_{1t} \\ \delta_{at} \end{bmatrix}, \tag{9.3}$$

where the δ terms have zero means, and variances which are given by suitable Poisson and binomial expressions:

$$\begin{aligned} \text{Var}(\delta_{1t}) &= E(N_{a,t})\, p\phi_1 \\ \text{Var}(\delta_{at}) &= E(N_t)\phi_a(1-\phi_a). \end{aligned}$$

If one is using classical inference, then it is necessary to use expectations in the variance expressions in order to comply with the assumptions of the Kalman filter; see Sullivan (1992) as well as the gray whale example in Sect. 6.4.2. The matrix above is a familiar Leslie matrix. Buckland et al. (2004, 2007) provide a general framework for deriving population projection matrices by considering intermediate sub-processes (Chap. 2). If we assume that only breeding birds are censused, then what we observe, y_t, which we take here as the CBC index, is given by the measurement equation,

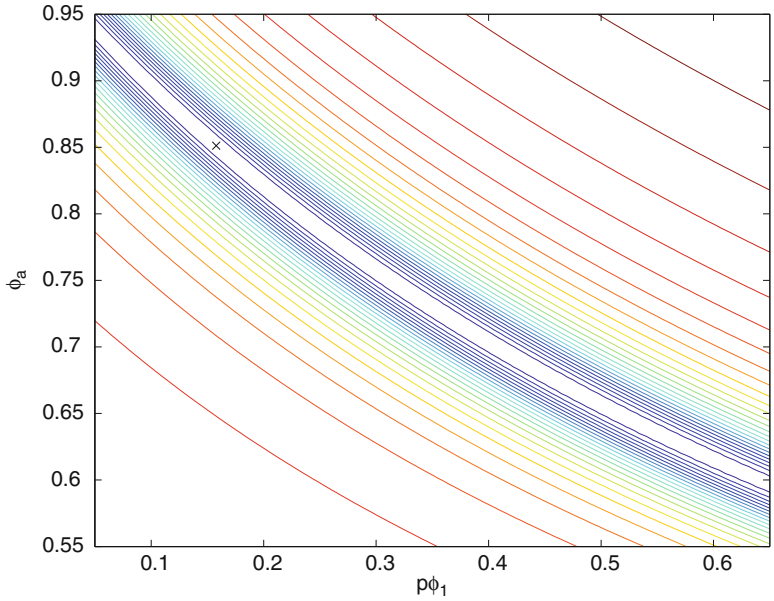

Fig. 9.2 Profile log-likelihood contours for $\theta = p\phi_1$ and ϕ_a from the lapwing census data, obtained by maximizing $\log L_c(p, \phi_1, \phi_a, \sigma)$ with respect to σ. The location of the maximum is shown by $\times$

$$y_t = (0 \;\; 1)\begin{pmatrix} N_{1t} \\ N_{at} \end{pmatrix} + \epsilon_t. \tag{9.4}$$

We assume that ϵ_t is normally distributed with constant variance σ^2, so that $\epsilon_t \sim \text{normal}(0, \sigma^2)$. (In some cases it may be sensible for σ^2 to vary with time.)

Equations (9.3) and (9.4) collectively form a normal dynamic linear model (see Sect. 4.4). Thus the likelihood for the census data, $L_c(p, \phi_1, \phi_a, \sigma)$, can be derived using the Kalman filter. However, the parameters p and ϕ_1 in this likelihood only appear as a product. Furthermore likelihood functions for census data generally provide limited information on the underlying demographic mechanisms individually. Figure 9.2 plots the two-dimensional profile log-likelihood contours of L_c for parameters $\theta = p\phi_1$ and ϕ_a for the lapwing data and illustrates how θ and ϕ_a are negatively correlated.

However ϕ_1 also occurs in L_r, and so we may obtain a full-rank model (see Sect. 5.2.2), in which all parameters may in principle be estimated, by maximizing the joint likelihood,

$$L_j(\phi_1, \phi_a, \kappa, p, \sigma) = L_r(\phi_1, \phi_a, \kappa) \times L_c(p, \phi_1, \phi_a, \sigma). \tag{9.5}$$

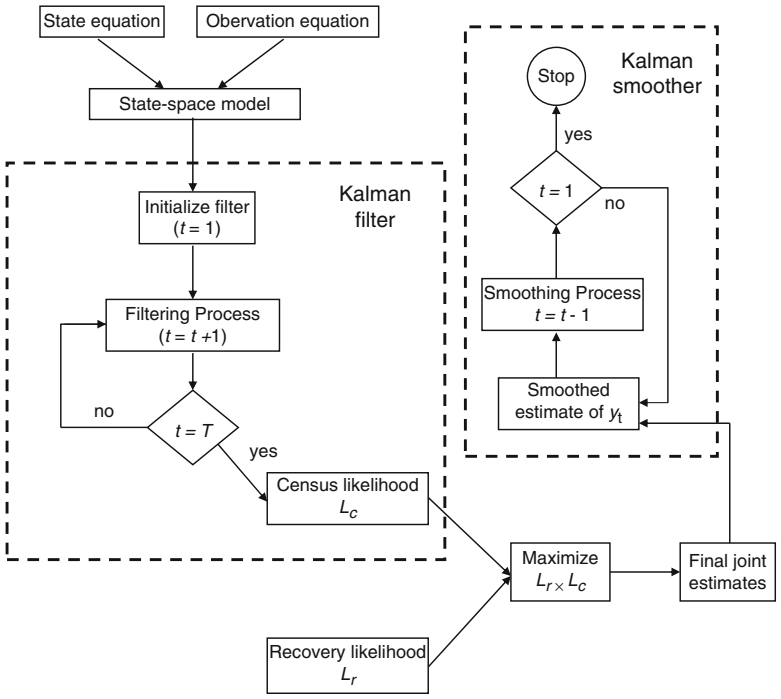

Fig. 9.3 Flow chart summarising the steps in fitting an integrated population model using classical inference incorporating the Kalman filter

(To reflect the various logistic/logarithmic regressions in the model, we can also write the joint likelihood as $L_j(\beta_0, \beta, \delta_0, \delta, \nu_0, \nu, \gamma_0, \gamma, \sigma)$, but we use the notation in Eq. (9.5) in a generic fashion, for brevity.) This is very useful, as the productivity of a species in decline is a parameter that we particularly want to investigate. The assumption of independence made in multiplying the two likelihoods together in Eq. (9.5) is not likely to be violated. The approach extends flexibly to include additional likelihood components, corresponding to data on say productivity or movement. This methodology has been termed *integrated population modelling* in the literature. The flow chart in Fig. 9.3 summarizes the steps of integrated population modelling, when classical inference is being used incorporating the Kalman filter. We shall now illustrate the performance of the integrated approach by application to the lapwing data before discussing a number of technical issues and extensions.

Table 9.3 Maximum likelihood parameter estimates from fitting the model $\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$ to the lapwing data, (i) using ring-recovery data alone, and (ii) using integrated population modelling, incorporating both the ring-recovery data and the census data

		Parameter estimates		Estimated standard errors	
		(i)	(ii)	(i)	(ii)
ϕ_1	Intercept (β_0)	0.5158	0.5231	0.0675	0.0672
	Slope (β)	−0.0241	−0.0228	0.0072	0.0070
ϕ_a	Intercept (δ_0)	1.5011	1.5210	0.0683	0.0690
	Slope (δ)	−0.0360	−0.0279	0.0051	0.0045
κ	Intercept (ν_0)	−4.5668	−4.5632	0.0350	0.0350
	Slope (ν)	−0.5729	−0.5841	0.0641	0.0636
p	Intercept (γ_0)		−1.1513		0.0880
	Slope (γ)		−0.4323		0.0743
σ			159.4691		21.8712

The estimated standard errors result from inverting a numerical approximation to the Hessian matrix at the maximum likelihood estimate. Reprinted with permission from *Biometrics*

9.3.2 Results for Lapwing Example

The decline of lapwings may be described through inclusion of time in survival and/or productivity parameters. Here we follow Besbeas et al. (2002) and allow the productivity parameter p to be a function of time: $p(year)$. Thus we have fitted to the lapwing data the model,

$$\phi_1(w), \phi_a(w)/\kappa(year)/p(year),$$

which indicates that both ϕ_1 and ϕ_a are logistic functions of the single covariate, w, which measures the number of days in the year when the temperature at a location in central England was below freezing. We concentrate on logistic regressions, as described in the last section, but alternative link functions might also be used. Additionally, we let $\kappa(year)$, $p(year)$ denote respectively logistic and logarithmic regressions of κ and p on year. Note that p is not a probability, and so is not bounded above by unity. The maximum likelihood point estimates from the joint data, and also from the recovery data only, are given in Table 9.3.

We can see that the joint analysis changes slightly the maximum likelihood estimates from the ring-recovery data, as now the estimates describe both the data sets. There is little difference between the standard errors of the common parameters in the two analyses, due to the dominance of the ring-recovery data in this illustration. The same conclusion arises from a Bayesian analysis; see Brooks et al. (2004). When we repeated the analysis with a subset of the recovery data, then we found that the joint analysis produced substantially reduced estimates of

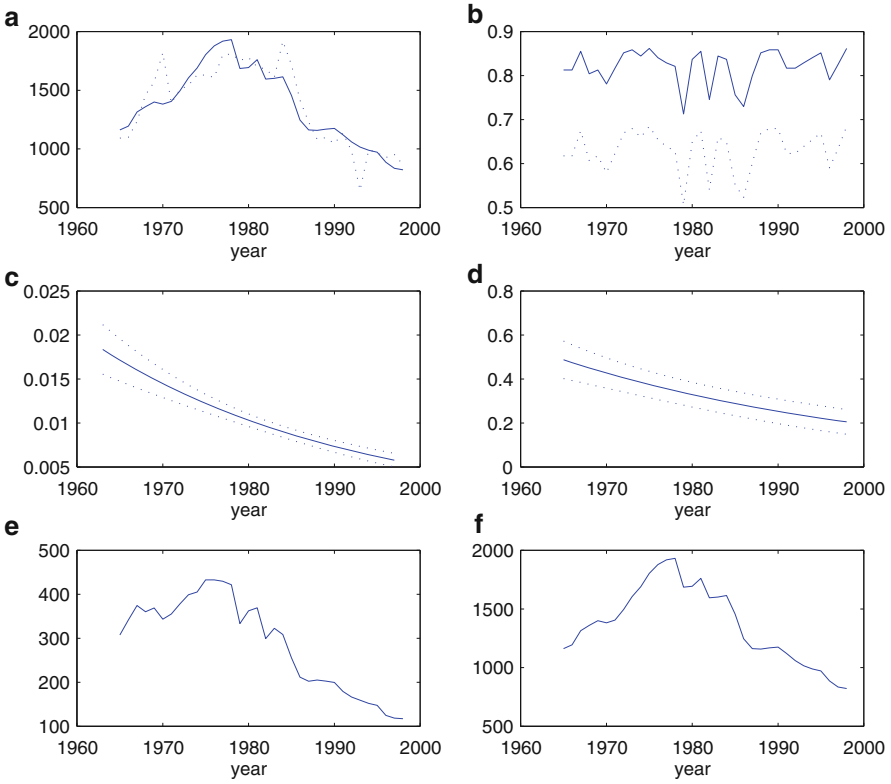

Fig. 9.4 The result of fitting model $\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$ to the lapwing data. (**a**) Observed census data (*dotted line*); fitted curve (*solid line*). (**b**) $\hat{\phi}_a$ (*solid line*); $\hat{\phi}_1$ (*dotted line*). (**c**) $\hat{\kappa}(year)$, with estimated 95 % confidence bands. (**d**) $\hat{p}(year)$, with estimated 95 % confidence bands. (**e**) Graph of $\hat{N}_{1t}$; (**f**) Graph of $\hat{N}_{at}$, equivalent to the fitted curve in (**a**). Reprinted with permission from *Biometrics*

standard error compared with the ring-recovery ones; see Besbeas et al. (2002). For the ring-recovery data alone, due to the length of the study, correlations between parameter estimators are typically quite small, ranging in magnitude from -0.20 to 0.29. The addition of the census data has little effect on those correlations. However there are now correlations of -0.49 between $\hat{\gamma}_0$ and $\hat{\beta}_0$, and -0.91 between $\hat{\gamma}_0$ and $\hat{\delta}_0$, where γ_0, β_0 and δ_0 are defined in Table 9.3. These are sensible findings, since increasing the productivity p requires a decrease in survival in order to match the data. We would also expect a stronger correlation with the intercept estimate of the adult survival, $\hat{\delta}_0$, as is seen above.

We show in Fig. 9.4 the results from maximizing the combined likelihood L_j. The confidence bands shown result from applying the δ-method. The decline of κ with time agrees with Catchpole et al. (1999). However of greater interest to us here is the time-varying behaviour of the parameters ϕ_1, ϕ_a and p. We can see

Table 9.4 The results from fitting by maximum likelihood a range of models to the lapwing data

(i)

Model	$-\ell$	No. of parameters	AIC
$\phi_1(w), \phi_a(w)/\kappa(year)$	7156.33	6	14325
$\phi_1(w, year), \phi_a(w, year)/\kappa(year)$	7155.16	8	14326

(ii)

Model	$-\ell$	No. of parameters	AIC	ΔAIC
$\phi_1(year), \phi_a(year)/\kappa(year)/p$	7409.54	8	14835	52
$\phi_1(w), \phi_a(w)/\kappa(year)/p$	7398.40	8	14813	30
$\phi_1(w), \phi_a(w, year)/\kappa(year)/p$	7382.37	9	**14783**	0
$\phi_1(w, year), \phi_a(year)/\kappa(year)/p$	7401.14	9	14820	37
$\phi_1(w, year), \phi_a(w, year)/\kappa(year)/p$	7381.45	10	**14783**	0
$\phi_1(year), \phi_a(year)/\kappa(year)/p(year)$	7409.29	9	14837	54
$\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$	7383.38	9	**14785**	2
$\phi_1(w), \phi_a(w, year)/\kappa(year)/p(year)$	7381.99	10	**14784**	1
$\phi_1(w, year), \phi_a(w)/\kappa(year)/p(year)$	7383.36	10	14787	3
$\phi_1(w, year), \phi_a(w, year)/\kappa(year)/p(year)$	7381.45	11	14786	3

In (i) we fit only the ring-recovery data; in (ii) we use integrated population modelling to fit a range of alternative models to both ring-recovery and census data. Here ℓ denotes the value of the log-likelihood evaluated at the maximum likelihood estimates of the parameters. AIC denotes the Akaike Information Criterion, and ΔAIC indicates the difference between the model AIC value and the smallest AIC value for the set of models considered. In (ii) we show in bold type the AIC values corresponding to alternative acceptable models for the combined data set, with $\Delta\text{AIC} \leq 2$. Reprinted with permission from *Biometrics*

that the decline in lapwing numbers since 1980 is compatible with a major drop in the productivity parameter p. An alternative explanation of the recent decline in lapwing numbers is that there is a decline in survival probability over time. There was no evidence for this in Catchpole et al. (1999), or for the more extensive ring-recovery data set analysed here (see Table 9.4(i)).

However we can see from Table 9.4(ii) that when we analyse the combined data using integrated population modelling, then in terms of AIC, several models provide comparable best fit to the data, in particular with constant p and declining probabilities of survival over time. Detailed studies of breeding lapwings have shown a decrease in chicks produced over the period. This is usually attributed to the switch from spring to autumn sowing of cereals and intensification of pasture management (see Wilson et al. 2001 and references therein). Thus in order to demonstrate model performance, we shall here only consider the model $\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$, as in Besbeas et al. (2002). For further discussion, see King et al. (2008).

We show in Fig. 9.5 the regressions of $\hat{\phi}_1$ and $\hat{\phi}_a$ on w, combined with plots of $\hat{\phi}_{1,t}$ and $\hat{\phi}_{a,t}$ resulting from a model with separate ϕ_1 and ϕ_a parameters for each year, and denoted by $\{\phi_{1,t}\}$ and $\{\phi_{a,t}\}$ respectively. These regressions are seen to provide a fair description of the relationship between annual survival and w. Note that the

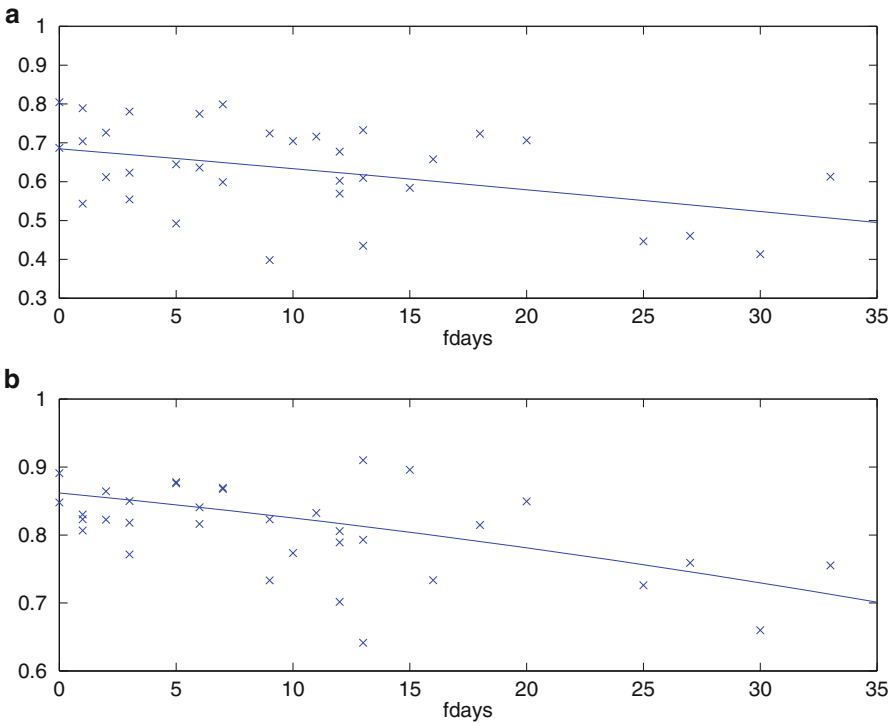

Fig. 9.5 A graphical demonstration of the logistic regressions of ϕ_1 and ϕ_a in the model, $\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$, for the lapwing data. Here the covariate w denotes the number of days in a year when the temperature in central England was below freezing. In (**a**) we graph $\hat{\phi}_1$ against w and in (**b**) we graph $\hat{\phi}_a$ against w. Also plotted are the values $\{\hat{\phi}_{1t}\}$ and $\{\hat{\phi}_{at}\}$ respectively, corresponding to having a separate value for ϕ_1 and ϕ_a for each year. Reprinted with permission from *Biometrics*

estimates of β and δ in Table 9.3 are approximately equal, and a model in which ϕ_1 and ϕ_a share a common slope parameter produces virtually no change to the likelihood.

Shown in Fig. 9.4e is the smoothed estimate of N_{1t}. The figure demonstrates the large decline in N_{1t} in recent years, in line with the predicted decline in p over time in this model.

We do not provide a formal test of goodness-of-fit of the selected model, but in Fig. 9.6 we plot the observed numbers of recoveries against the fitted numbers, and also provide a Q-Q plot of the prediction errors from the Kalman filter, which are expected to have a normal distribution. In the latter case, the single large prediction error is due to the initial census value, and is to be expected because of the way in which the Kalman filter analysis is initiated. Overall there is no serious indication of systematic lack of fit.

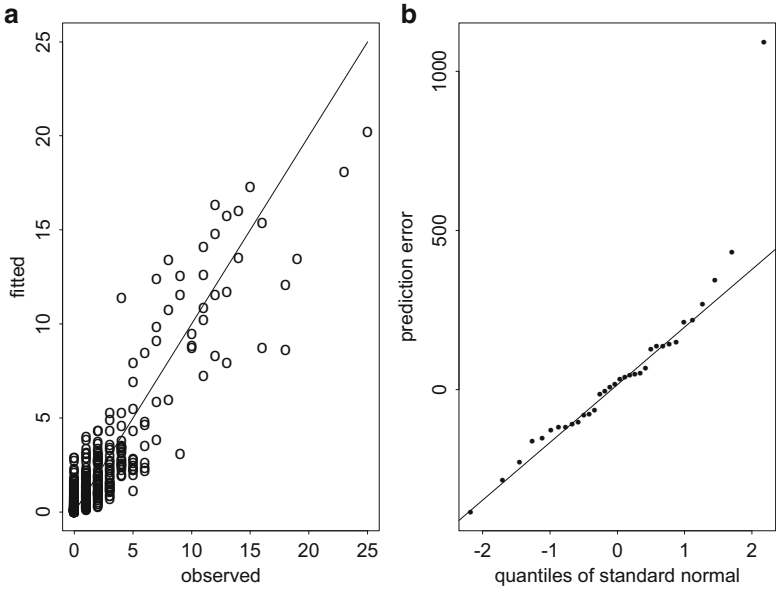

Fig. 9.6 Graphical checks of goodness-of-fit of the model $\phi_1(w), \phi_a(w)/\kappa(year)/p(year)$, for the lapwing data: (**a**) observed numbers plotted vs expected numbers for the recovery data; (**b**) a Q-Q plot of the prediction errors from the Kalman filter. Reprinted with permission from *Biometrics*

9.4 Facilitating Connected Likelihood Modelling

9.4.1 Normal Approximation

The integrated approach described above requires specialized computer code, not only for the Kalman filter component, but also for the additional likelihood components, some of which may be complex, and may even have been derived using specialist computer packages. Quite often, for example, survival data would be analysed by Program MARK (White and Burnham 1999). This then poses obvious difficulties for combined analysis and is likely to preclude the use of the integrated approach in practice. A solution to this problem is provided by an approximation, suggested and evaluated by Besbeas et al. (2003). Here, a multivariate normal approximation is adopted for the form of the likelihood of the ring-recovery data, making use of the parameter estimates, and their corresponding estimates of dispersion obtained from analysing the ring-recovery data alone. In particular, we make a multivariate normal approximation to $L_r(\boldsymbol{\phi}, \kappa)$:

$$\begin{aligned} \log L_r(\boldsymbol{\phi}, \kappa) &\approx \text{constant} - \frac{1}{2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' \hat{\boldsymbol{\Sigma}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ &\equiv \log \hat{L}_r(\boldsymbol{\phi}, \kappa) \end{aligned} \tag{9.6}$$

where we write $\boldsymbol{\theta}$ to denote the model parameters on the logistic scale, and where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ are respectively the maximum likelihood estimates of $\boldsymbol{\theta}$ and the dispersion matrix of $\hat{\boldsymbol{\theta}}$, both obtained from a separate model-fitting exercise for the ring-recovery data alone. The approximation is motivated naturally by the asymptotically multivariate normal distribution of maximum likelihood estimators, and has been used to good effect in Lebreton et al. (1992). The integrated population modelling then proceeds by replacing the exact likelihood $L_r(\boldsymbol{\phi}, \kappa)$ in (9.5) with expression (9.6):

$$L_j(\phi_1, \phi_a, \kappa, p, \sigma) = \hat{L}_r(\phi_1, \phi_a, \kappa) \times L_c(p, \phi_1, \phi_a, \sigma).$$

The approach has great potential to approximate likelihood components for abundance data obtained from complex sampling schemes (Knape et al. 2013). This means that particular programs or packages for survival and also for fecundity or other data, if appropriate, need only be run once, to obtain the relevant maximum likelihood estimates of relevant parameters and their estimates of standard error and correlation. This clearly greatly simplifies both the resulting form for L_j and its maximization.

9.4.2 Results

For the example of the last section, we evaluate the approximation in Table 9.5; the standard errors are obtained from the observed information matrices as in Table 9.3. The agreement between exact and approximate results is seen to be very good. An additional benefit from using the approximate approach is that in comparison with the exact analysis, it is far less sensitive to starting values for the maximum likelihood iteration. Additionally the approximate analysis was found to be about 2.5 times faster than the exact method.

We can see from Table 9.5 that the change in the value of $\hat{\boldsymbol{\theta}}$ from the recovery analysis alone to the joint analysis is not large. It is both the overall magnitude of this change and the size of the data sets which determine the effectiveness of the multivariate normal approximation made in the paper, as we require the multivariate normal approximation to L_r to be good for the value of $\boldsymbol{\theta}$ which maximizes L_j, and not just for the value that maximizes L_r. We show in Fig. 9.7 the good agreement between exact and approximate $\log L_r(\boldsymbol{\phi}, \kappa)$ for the small illustrative data set of Table 7.1, by means of profile log-likelihoods, for the parameter β_0. We have found the good agreement observed in Fig. 9.7 also for profiles with respect to other parameters. Besbeas et al. (2003) provide further details. The good performance of the approximation extends to the case of multi-site data (McCrea et al. 2010).

Table 9.5 An evaluation of the multivariate normal approximation for the ring-recovery likelihood for the lapwing data

Parameter	Ring-recovery alone	Exact combination	Approximate combination
β_0	0.5158 (0.0679)	0.5231 (0.0679)	0.5226 (0.0678)
β	−0.0241 (0.0072)	−0.0228 (0.0070)	−0.0227 (0.0070)
δ_0	1.5011 (0.0685)	1.5210 (0.0693)	1.5191 (0.0686)
δ	−0.0360 (0.0051)	−0.0279 (0.0045)	−0.0280 (0.0045)
ν_0	−4.5668 (0.0351)	−4.5632 (0.0352)	−4.5634 (0.0351)
ν	−0.5729 (0.0640)	−0.5841 (0.0637)	−0.5837 (0.0638)
γ_0		−1.1513 (0.0886)	−1.1489 (0.0876)
γ		−0.4323 (0.0743)	−0.4314 (0.0740)
σ		159.469 (22.062)	159.613 (21.875)

In each case we show the maximum likelihood parameter estimate and the corresponding standard error. The estimates of error are obtained as in Table 9.3. Reproduced with permission from *Applied Statistics*

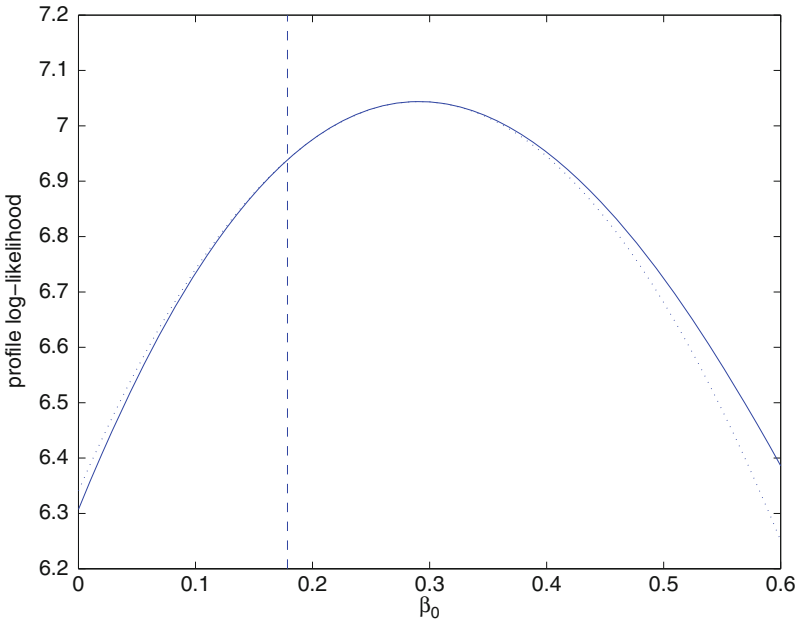

Fig. 9.7 Agreement between $\log L_r(\boldsymbol{\phi}, \kappa)$ and the multivariate normal approximation for the data of Table 7.1. The two curves are profile log-likelihoods taken with respect to the parameter β_0. The *solid curve* is the exact likelihood, and the *dotted curve* approximate. Also shown is the location of $\hat{\beta}_0 = 0.1786$, the value which maximizes the combined exact likelihood, making use of the entire run of the census data for lapwings. Reprinted with permission from *Applied Statistics*

9.5 Technical Issues for Classical Analysis

The connected likelihood approach depends on making a number of assumptions and adopting certain procedures. For example, we assume that we can approximate discrete distributions by normal distributions; that we can suitably start the Kalman filter iterations; that the different surveys are independent; and that the state-space model adopted correctly partitions variation between its transition and measurement processes. Brooks et al. (2004) demonstrated the robustness of the normality assumptions. In this section, we explore the other issues listed above, starting with the problem of initializing the Kalman filter.

9.5.1 Kalman Filter Initialization

Computational algorithms in state space analyses are mainly based on recursions in which we calculate values at time t from earlier values for $t-1,\ldots,1$. The question of how these recursions are started up at the beginning of the series is called initialization. The initialization problem of the Kalman filter when $\boldsymbol{\mu}_0$ and $\mathbf{Q}_0$ are unknown is an issue requiring attention in general, in areas such as economics and engineering, but the problem may be more important in population ecology, where there are often small samples, models may involve a large number of states and unknown parameters, and models are usually non-stationary. We shall now describe an approach to initialize the Kalman filter which is designed for integrated population modelling in ecology.

9.5.1.1 Stable-Age Initialization

In population ecology, the state vector $\mathbf{n}_t$ is typically a vector denoting the numbers of individuals in the population in a number of classes at time t. Typical elements of $\mathbf{A}_t$ include age- or stage-specific survival and productivity parameters and rates at which individuals in one state make the transition to another state, for example through immigration or emigration. The matrix $\mathbf{A}_t$ is referred to as a Leslie or Lefkovitch matrix, depending on whether the population is age- or stage-classified, respectively (Chap. 2). We encountered a Leslie matrix in Sect. 9.3.1.

The Perron-Frobenius theorem states that, for appropriate *constant* transition matrices $\mathbf{A}$, there exists a real positive eigenvalue κ that is greater in absolute value (or in modulus, if some of the other eigenvalues are complex) than all of the other eigenvalues. The implications are that the dominant eigenvalue κ represents the asymptotic growth rate of the population, and the normalized right eigenvector associated with κ represents the asymptotic proportion of every age or stage class in the total population. We call the eigenvalue κ *the asymptotic growth rate* and its corresponding right eigenvector, $\mathbf{v}$, is called the *stable age* (or *stable stage*) distribution.

We propose starting the Kalman filter by taking the initial mean vector $\boldsymbol{\mu}_0$ to be proportional to the stable age (stage) distribution of a Leslie (Lefkovitch) matrix $\mathbf{A}$, with the proportions scaled by the total size of the first observation, $\mathbf{n}_0$. To choose $\mathbf{Q}_0$, we adopt the conservative approach of requiring the lower end of an appropriate confidence interval for each element of $\mathbf{n}_0$ to be non-negative, and that elements are independent. Thus for example for univariate observations and 95 % confidence intervals we take

$$\boldsymbol{\mu}_0 = \mathbf{v}y_1/(\mathbf{Z}_1\mathbf{v}), \qquad \mathbf{P}_1 = \mathrm{diag}((\boldsymbol{\mu}_0/1.96)^2)$$

where y_1 and $\mathbf{Z}_1$ are the first observation and measurement vector respectively. In practice, in order to derive the stable age distribution, we need to know $\mathbf{A}$, which may contain unknown parameters. We select the values for the parameters in $\mathbf{A}$ that are in common with the demographic analyses by using their maximum likelihood estimates obtained from analysing the demographic data alone. Any remaining parameter(s) can then be obtained by iteration; for details see Besbeas and Morgan (2012b). For instance if productivity p is unknown, then we take an arbitrary value for p, and obtain a maximum likelihood fit, which provides an estimate, $\hat{p}^{(1)}$. We then use this value to start the Kalman filter, resulting in an estimate $\hat{p}^{(2)}$, and so on. When the matrix $\mathbf{A}_t$ is time-dependent, then we obtain $\mathbf{A}$ by an appropriate time-average of the $\mathbf{A}_t$. The good performance of this approach compared with alternatives in which the elements of the initial state vector are diffuse, that is, treated as random variables with infinite variance, or are treated as unknown constants to be estimated by maximum likelihood, is shown in simulation studies presented by Besbeas and Morgan (2012b).

9.5.2 Lack of Independence

It may be that census data and demographic data are not completely independent. Besbeas et al. (2009) consider the effect of dependence between a ring-recovery data set and census data. This was done for a model with two age-classes for survival, as in Sect. 9.3.1, and with constant parameters, ϕ_1, ϕ_a and κ. Life histories spanning eight years were constructed for a 20-year period, with $\phi_1 = 0.5$, $\phi_a = 0.7$, for a range of values of κ, and probability of recapture of a live animal. Observation error was added to the life histories, resulting in census data. It was shown that, in some circumstances, combining dependent data sets but treating them as independent can actually reduce estimator precision. While this was only a single study, the message is that one should take care conducting combined analyses for dependent data sets.

9.5.3 Heterogeneity in the State-Space Model

In Sect. 9.3.1, we used binomial error variances for when a number n of individuals survive or die in any particular year, and Poisson error variances for recruitment arising from reproduction. In both of these cases, we can allow for heterogeneity. A study with an application to data on grey herons *Ardea cinerea* is provided by Besbeas et al. (2009) in which census data are combined with ring-recovery data. It was found that when overdispersion is included in the integrated modelling, it is possible to obtain $\hat{\sigma} = 0$. We return to this finding in the next section. Thus while it is in principle straightforward to include overdispersion in integrated population modelling, there may be an interaction between the roles of the overdispersion parameter and the parameter denoting measurement error.

9.6 From Modelling an Index to Multi-Sites: Additional Complexity

In this section we consider how the connected likelihood approach can be extended to multi-site and multi-state models.

9.6.1 Accounting for Different Habitats

The CBC sites on which the index y_t was based in Sect. 9.3.1 may be classified as arable, grazing, mixed (i.e. arable and grazing) or "other" (which are not farmland and could include estuaries, for example). The study by Besbeas et al. (2002) did not make use of this information but it is interesting to consider how it might be used to give a breakdown with respect to habitat. For illustration, we outline the integrated analysis of the grazing and arable sites, which account for 31 % of the sites; Besbeas et al. (2005) give detailed results and also discuss the complete set of results from all four habitats.

The ring-recovery component estimates overall survival, and does not produce a breakdown with respect to habitat. However we use the state-space model given below, using superscripts A and G to indicate arable and grazing, respectively. Let $\mathbf{n}_t = (n_{1t}^A, n_{at}^A, n_{1t}^G, n_{at}^G)'$ and $\mathbf{y}_t = (y_t^A, y_t^G)'$. Then, with a general notation for respective error terms $\boldsymbol{\delta}_t$ and $\boldsymbol{\epsilon}_t$,

$$\mathbf{n}_t = \mathbf{A}\mathbf{n}_{t-1} + \boldsymbol{\delta}_t, \qquad \mathbf{y}_t = \mathbf{B}\mathbf{n}_t + \boldsymbol{\epsilon}_t,$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & p^A\phi_1 & 0 & 0 \\ \phi_a & \phi_a & 0 & 0 \\ 0 & 0 & 0 & p^G\phi_1 \\ 0 & 0 & \phi_a & \phi_a \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{bmatrix}.$$

Thus n^A_{1t}, for example, is the number of one-year-old female lapwings fledged, and assumed subsequently to live, on arable land at time t, and y^A_t denotes the CBC index value at time t, derived from the arable sites only. Thus the model now assumes different productivities for arable sites, p^A, and grazing sites, p^G, and we also assume that there are different observation error standard deviations, σ^A and σ^G for arable and grazing respectively. Note that the model analyses the arable and grazing populations separately and assumes no movement of individuals between the two habitats. We discuss extensions to multi-site modelling incorporating movement probabilities in Sect. 9.6.3. The likelihood function from this multivariate model is equivalent to the product of likelihood functions for univariate data

$$L_c(\boldsymbol{\phi}, p^A, \sigma^A; \mathbf{y}^A) \times L_c(\boldsymbol{\phi}, p^G, \sigma^G; \mathbf{y}^G)$$

where $\boldsymbol{\phi} = (\phi_1, \phi_a)$, and $\mathbf{y}^A$ ($\mathbf{y}^G$) denote all the observations on arable (grazing) sites over time. These likelihood functions are constructed for the relevant index alone, as in Sect. 9.3.1. The likelihood can readily be extended to deal with more than two indices as well as individual site data. The analysis then follows the same lines as above, based on the model

$$\phi_1(w), \phi_a(w), \kappa(year), p^A(year), p^G(year).$$

The results indicate that productivity varies with habitat; see Besbeas et al. (2005) for details.

9.6.2 Modelling Individual Site Data

Besbeas and Freeman (2006) provide an alternative approach to dealing with individual site data which sidesteps the intermediate process of deriving an index of abundance. The approach is single-stage and has several advantages, including no loss of information in summarizing the raw data using an index. This approach fits the survey data directly by incorporating the population model into the generalized linear model used to derive the index in the existing approach. For example, from Eq. (9.1) and the structural part of the state-space model of Eqs. (9.3) and (9.4), we obtain the recursive relationship for the year effects

$$u_t = \phi_a u_{t-1} + p\phi_1\phi_a u_{t-2}, \quad t > 2, \tag{9.7}$$

which gives rise to the likelihood function, $L_{glm}(\phi_1, \phi_a, p, u_{1:2}, \mathbf{s})$. The form of L_{glm} depends on the distributional assumptions made for the individual site data, c_{it}. For example, when the c_{it} are assumed to follow independent Poisson distributions, then the log-likelihood function is given by

$$\log L_{glm}(\phi_1, \phi_a, p, u_{1:2}, \mathbf{s}) = \text{constant} + \sum_{i=1}^{S} \sum_{t=1}^{T} c_{it}(s_i + u_t) + e^{s_i + u_t}$$

where $u_{1:2} = (u_1, u_2)$ and u_t is given by Eq. (9.7) for $t > 2$. The integrated population modelling procedure then replaces the component corresponding to the abundance index, $L_c(\phi_1, \phi_a, p, \sigma)$, in Eq. (9.5), by $L_{glm}(\phi_1, \phi_a, p, u_{1:2}, \mathbf{s})$. The procedure is shown to perform well relative to analysing an index; see also Maunder (2001) for related work. Freeman and Besbeas (2012) provide further development with respect to analysing presence-absence survey data. The ease with which the connected likelihood approach can integrate different types of information, assuming independence between the sources, makes it very appealing in practice.

9.6.3 Multi-Site and Multi-State Modelling

The work described so far assumes no movement of individuals. How the general theory of integrated population modelling extends to the multi-site case is described by McCrea et al. (2010). The illustrative example of that paper involved great cormorants *Phalacrocorax carbo sinensis*, moving between three sites in Denmark, with the additional complication that birds could move between breeding and non-breeding states, as well as between sites. Model selection was based on a step-up procedure applied to the recapture data alone, in order to avoid possible problems with overfitting the census data, as discussed in Sect. 9.5.3. The final integrated population model included complex state- and site-dependent transitions. There was a pronounced improvement in the precision of estimators from combining the two data sets in a single analysis.

9.7 Additional Aspects of IPM

9.7.1 Conditional Gaussian Modelling

The main advantage of analysing census data using normal dynamic linear models lies in the use of the Kalman filter, which greatly facilitates the estimation process. Despite the conceptual simplicity of this model, it is both flexible and general, and extends to a range of other models, such as linear matrix population models which have wide application in ecology and demography (Caswell 2001). The non-linear

case however is important in ecological settings, for example in accommodating density dependence in the population model or allowing observation variance to depend on population size; see Sect. 4.4.2. In certain important cases, non-linear models are still amenable to analysis by the Kalman filter. These models are known as *conditional Gaussian models* and are described in detail in Harvey (1989, Sect. 3.7.1). Here system matrices $\mathbf{A}_t$, $\mathbf{B}_t$, $\mathbf{Q}_t$ and $\mathbf{R}_t$ may depend upon previous observations, up to and including $\mathbf{n}_t$ in the notation of Eq. (4.9). The essence of these models is that even though the matrices may depend on observations up to time t, they may be regarded as fixed once time t has been reached, and thus the derivation of the Kalman filter still applies.

Besbeas et al. (2009), Tavecchia et al. (2009) and Besbeas and Morgan (2012a) provide illustrations of this approach on three different species. The illustration below is taken from this last paper, based on the grey heron. This example also involves ring-recovery and census data but uses a more elaborate survival age structure than the lapwing illustration, involving four age classes. The heron census data have been published widely and a notable characteristic of this species, other than its marked population crashes, is that the population rebounds quickly after a crash.

Besbeas and Morgan (2012a) considered several models for heron productivity p_t in year t, including a direct density-dependent model

$$\log p_t = \gamma_0 + \gamma y_t,$$

and a threshold model, where for a suitable threshold, τ,

$$\log p_t = \begin{cases} \nu_0 + \nu_1 & \text{if } y_t < \tau, \\ \nu_0 & \text{if } y_t \geq \tau. \end{cases}$$

The motivating assumption of this model is that as the birds nest in heronries, then when numbers are low, there might be less competition for space and resources, resulting in higher productivity than when the numbers are high. We would thus expect $\nu_1 > 0$. The threshold model is found to perform well relative to alternative models; see Besbeas and Morgan (2012a) for details, including extensions to more than one threshold. This paper also compares the results of a conditional Gaussian model with those from a Bayesian approach.

9.7.2 Goodness of Fit

A global goodness-of-fit test does not exist for integrated population models. For IPMs using a single SSM, the methods discussed in Sect. 5.6, e.g. recursive residuals, are a possibility.

For connected likelihood models, it seems intuitively reasonable to examine goodness of fit for each component of the model separately, which enables us to use off-the-shelf goodness-of-fit techniques in each of MRR modelling and time-series modelling, for example. Thus, in the case of mark-recovery and census data, as in Sect. 9.3.1, it is possible to examine plots of observed versus fitted numbers of recoveries which will indicate the fit of the mark-recovery component of the model, whilst for the census component, the prediction errors $\mathbf{v}_t$ are expected to be normally distributed and thus it might be appropriate to examine a Q-Q plot of these errors. Figure 9.6 provides an example of this approach for the lapwing data. The per-component goodness of fit can also be examined using standard tests, such as chi-square tests or numerical tests for normality. The extension to other types of mark-recapture-recovery data is obvious. However in all cases, it is important to note that integrated population modelling estimates are joint estimates, describing several components simultaneously, and may thus inherently manifest some lack-of-fit to each component individually.

The time-series plot of the observed and corresponding smoothed population estimates is an additional, natural diagnostic tool for the census component to examine whether the fitted model exhibits the stylised characteristics concerning the series. This plot is illustrated in Fig. 9.4a for the lapwing data. The time-series plot of the smoothed states against time is also a valuable diagnostic tool to check if the components extracted provide a suitable representation of these characteristics, and these are illustrated in Fig. 9.4e and f. The estimated observation variance provides a quantitative indication of the overall fit, with smaller values indicating a better model fit. However as discussed in the last section, it is important to consider this approach in parallel with model selection and heterogeneity in the state-space model, and the approach may need to be informed by additional considerations, which might for instance provide guidance on the appropriate magnitude of observation error. The potential to over-fit census data exists as shown in Besbeas et al. (2009), and any integrated population modelling should conclude with a particular assessment of the estimated measurement error, and consideration of whether or not it is appropriate.

In recent work, Besbeas and Morgan (2014) illustrate how Monte Carlo simulation can be used in the evaluation of goodness of fit of integrated models. We expect goodness-of-fit Monte Carlo simulation techniques to be become increasingly adopted both in integrated modelling and in capture-recapture in the future.

9.7.3 Bayesian Methods

The emphasis in this chapter has been on the connected likelihood approach using methods of classical analysis. Bayesian methods for state-space modelling have been described in Sect. 4.3. See also, for example, Meyer and Millar (1999) and Millar and Meyer (2000a,b). These papers involve fisheries applications, where there have been numerous other applications— see for example Rivot et al. (2001,

2004), Rivot and Prévost (2002) and Rifflart et al. (2006). The Bayesian approach to integrated population modelling was initially described by Brooks et al. (2004) and for more information, as well as simple `WinBUGS` programs to perform the necessary Markov chain Monte Carlo iterations, see Schaub et al. (2007), Gimenez et al. (2009b) and King et al. (2009). A clear advantage of the Bayesian approach is that it removes the need to use the Kalman filter, and hence the need to make normal approximations, as well as to specify the variances of the δ terms of Eq. (9.3) in terms of expectations. However the work of Brooks et al. (2004) showed that classical methods are robust to the normal approximation. We can also model density dependence directly with a Bayesian approach, and avoid the conditionally Gaussian modelling of the last section. In addition, goodness of fit can be assessed using Bayesian p-values; see Sect. 5.6.4. However King (2011) discusses how MCMC can be slow for fitting SSMs. The interplay between integrated modelling and Bayesian analysis is discussed in Maunder (2003).

9.7.4 Model Selection

The general procedures discussed in Sect. 5.6 still apply but some specifics to the connected likelihood method are worth noting. Besbeas et al. (2002) combined ring-recovery data with census data. The number of age classes, and other model aspects, were those obtained from analysing the ring-recovery data alone. An alternative approach would be to base model selection on the connected likelihood. However for that kind of data combination, when selecting the number of age classes using AIC, Besbeas et al. (2014) found that it can be better to use recovery data alone.

9.7.5 Integrated Modelling in Fisheries Research

In this chapter, we have seen how state-space models and integrated population modelling allow the inclusion of information at the population and individual levels in a single framework. The state-space model for the population survey data opens the way to integrating different types of information and we have shown how data from multiple surveys can be simultaneously analysed in two different ways: by adding additional model structure in the model, as in Sect. 9.2, or by combining different likelihoods for different surveys with shared parameters for common processes, as in Sect. 9.3. The focus in the chapter has been on the latter and its application in population ecology, but integrated analysis has a long history in fisheries, dating back to Fournier and Archibald (1982), with developments in e.g. Maunder (1998, 2001, 2003). A recent review of integrated analysis in fisheries research is provided by Maunder and Punt (2013). Hoyle and Maunder (2004) provide a non-fisheries example, northeastern offshore spotted dolphin *Stenella attenuata*. Maunder (2004) uses IPM to carry out Population Viability Analysis (see Sect. 4.2).

One of the highly generalized integrated analysis models in fisheries is termed stock synthesis (Methot and Wetzel 2013), in which different data sets often contain contradictory information, and one of the main difficulties is determining the appropriate weighting factor between the different data types. In such cases, sensitivity analyses that investigate the influence of the weighting factors are an important part of the assessment. There are thus interesting analogies but also differences between integrated analysis in population ecology in general and particular applications in fisheries research and management. These differences stem from the different types of data, processes such as catchability, selectivity and aging imprecision, but also modelling purpose and complexity. However the underlying logic is the same, and research in both areas would benefit from better connection.