

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Besbeas, Panagiotis and Morgan, Byron J. T. (2017) Variance estimation for integrated population models. *Advances in Statistical Analysis*. ISSN 1863-8171.

### DOI

<https://doi.org/10.1007/s10182-017-0304-5>

### Link to record in KAR

<http://kar.kent.ac.uk/62889/>

### Document Version

Author's Accepted Manuscript

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

1 **SPECIAL ISSUE - ECOLOGICAL STATISTICS:**

2 **Variance estimation for integrated population models**

3 **Panagiotis Besbeas · Byron J.T. Morgan**

4  
5 Received: date / Accepted: date

6 **Abstract** State-space models are widely used in ecology. However it is well known that in practice  
7 it can be difficult to estimate both the process and observation variances that occur in such models.  
8 We consider this issue for integrated population models, which incorporate state-space models for  
9 population dynamics. To some extent the mechanism of integrated population models protects against  
10 this problem, but it can still arise, and two illustrations are provided, in each of which the observation  
11 variance is estimated as zero. In the context of an extended case study involving data on British Grey  
12 herons we consider alternative approaches for dealing with the problem when it occurs. In particular  
13 we consider penalised likelihood, a method based on fitting splines and a method of pseudo replication,  
14 which is undertaken via a simple bootstrap procedure. For the case study of the paper it is shown that  
15 when it occurs, an estimate of zero observation variance is unimportant for inference relating to the  
16 model parameters of primary interest. This unexpected finding is supported by a simulation study.

17 **Keywords** Bootstrap · Cross validation · Cubic splines · Grey heron · Mark-recovery-recapture  
18 data · Overfitting · Penalised likelihood · Plug-in method · Process/Observation Error estimation ·  
19 State-space models · Time-dependent parameters

---

P. Besbeas

Department of Statistics, Athens University of Economics and Statistics, Athens, Greece and National Centre for  
Statistical Ecology, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent,  
CT2 7NF, England

Tel.: +1227 816049

Fax: +1227 827932

E-mail: p.t.besbeas@kent.ac.uk

B.J.T.Morgan

National Centre for Statistical Ecology, School of Mathematics, Statistics and Actuarial Science, University of Kent,  
Canterbury, Kent, CT2 7NF, England E-mail: b.j.t.morgan@kent.ac.uk

## 1 Introduction to integrated population modelling using state-space models

Different surveys of wild animals can produce separate data sets, each with information on common parameters in population dynamics models. Surveys might be undertaken to estimate survival or productivity, or produce time-series of counts, which are the consequence of the interplay between survival and productivity, and often arise from national censuses.

Likelihoods describing productivity are frequently relatively straightforward, often based on binomial and related distributions, to describe count data on numbers of newborn individuals. Likelihoods describing survival typically arise in the analysis of capture-recapture data of different kinds, and are often based on multinomial distributions. More complex are likelihoods to describe census/count data, and an attractive framework for analysing such data is provided by state-space models. Here uncertainty is modelled in both a hidden stochastic process and in describing the observations made on the states of that process. This paper is devoted to investigating the estimation of these two types of variation.

Integrated population modelling combines the information in different independent surveys by forming joint likelihoods, which are products of component likelihoods, one for each survey. We follow the approach of Besbeas *et al.* (2002) which is now widely used; see eg McCrea *et al.* (2010). A Bayesian approach is described by Brooks *et al.* (2004) and Chapter 11 of Kéry and Schaub (2012) provides a comprehensive overview. A useful survey of applications is given by Schaub and Abadi (2011). Technical issues, such as how to provide initial population values for analysing time series of abundance data, how to perform model selection and how to gauge model fit are described by Besbeas and Morgan (2011), Besbeas *et al.* (2015) and Besbeas and Morgan (2014) respectively. In the case of model selection, standard use of the Akaike Information Criterion can over fit the data; see also Bengtsson and Cavanaugh (2006). Besbeas and Morgan (2014) use a method of calibrated simulation for judging goodness of fit.

Potential benefits of integrated population modelling (IPM) include improved precision of the estimates of common parameters, the estimation of parameters on which there is no direct information, and the coherent estimation of standard errors. Recent work is described by Chandler and Clark (2014) and Mazzetta *et al.* (2010), which includes relaxing the requirement that the different surveys need to be independent; see also Besbeas *et al.* (2009) and Abadi *et al.* (2010). We base the investigations of this paper on a single case study which has been used to illustrate a number of developments in the analysis of integrated data, and has a number of complex features that require appropriate description. However the work has wider implications in the estimation of population dynamics.

The plan of the paper is as follows: Section 2 describes the case study, involving British Grey herons, and which is used throughout the paper; this section provides detail regarding integrated population modelling and how it operates. Section 3 describes how integrated population modelling

proceeds, and gives the details of how component likelihoods are constructed. Section 4 explains how process and observation/measurement variance estimation might be difficult when simple state-space models are fitted to time-series data alone. Section 5 demonstrates that although this is less of a problem in integrated population modelling, there can still be difficulties with the correct estimation of observation variance in this context. Three possible solutions are proposed and evaluated in Section 6, and Section 7 provides a simulation study to investigate the wider implications of the findings. The paper ends with discussion and a recommendation in Section 8. An Appendix provides a short R program and additional figures are available at the Online Resource for the paper.

## 2 Grey heron case study

In the UK, survey information for estimating the demographic parameters of Grey herons, *Ardea cinerea*, is available at the British Trust for Ornithology. These birds nest in colonies, high up in trees, and as a result it is difficult to obtain access to nests and obtain reliable data on productivity. However there is a national heron census, which dates back to its inception in 1928, in which counts are made of what are judged to be active nests of breeding pairs, as opposed to individuals and there is also ring-recovery (MRR) data from birds ringed as chicks throughout the UK, which may be used to estimate annual survival probabilities. Guidance on taking the census is given at <http://www.bto.org/volunteer-surveys/heronries-census/taking-part>.

Ringling is of chicks, and takes place when they are still in the nest.

See <http://app.bto.org/birdfacts/results/bob1220.htm> for a summary of important features of the studies of grey herons in the UK. In this work we use the ring-recovery information from 1955-1997. Ring-recovery data are summarised by a table with each row corresponding to a year of ringing. Each year the number of birds ringed forms a multinomial index for the multinomial distribution which describes the numbers of herons reported dead that year from that cohort of ringed birds; see Chapter 4 of McCrea and Morgan (2014).

The Grey heron census from 1928 to 1998 inclusive is illustrated in Figure 1. Early values appear to be rounded to at least the nearest 50, an issue that we do not resolve here, but which is one indication of the presence of measurement error in the values. For continuity with earlier analyses, it is this data set that we analyse in the paper.

The ringed birds form a small fraction of the national population. Furthermore the ring-recovery information spans 43 of the 71 years of the census information in the study. Consequently the two data sets, of census and recovery information, may be regarded as independent, and we shall make that assumption throughout this work.

[Fig. 1 about here.]

### 3 Integrated population modelling

Models for the two data sets share common survival probabilities, and integrated population modelling exploits this feature. We use methods of classical statistical inference. Likelihoods are formed for the two component data sets and the product of these likelihoods is maximised to produce maximum-likelihood estimates for the complete set of model parameters. In practice therefore we set

$$\log(L_j) = \log(L_r) + \log(L_c), \quad (1)$$

where  $L_j$  denotes the joint likelihood,  $L_r$  denotes the likelihood for the ring-recovery data, given below in Equation 2, and  $L_c$  denotes the likelihood for the census data. Note that similar approaches are to be found in fisheries stock assessment models, where there is also consideration of differential weighting of the components in the joint log likelihood; see Francis (2011). The model for the census data includes a productivity parameter,  $p$ , and although there is not a data set providing direct information on productivity, as a consequence of integrated population modelling we can estimate this parameter, along with an estimate of its standard error. In fact for Grey herons, Besbeas *et al.* (2002) assumed that productivity is constant, resulting in an estimate of productivity given by  $\hat{p} = 0.96(0.07)$ . Here and throughout the paper we denote estimated standard errors by the terms in parentheses following the maximum-likelihood estimates. It is shown in Besbeas and Morgan (2012) how it is possible to formulate a complex, realistic structure for productivity, in which increases are triggered by drops in population size, relative to size thresholds, and we comment again on this feature later in the paper.

#### 3.1 Likelihood formation: ring-recovery data

Suppose, in a  $T$ -year study,  $d_{i,j}$  individuals are reported dead at time  $t_j$ , from a cohort of  $R_i$  individuals ringed as chicks at time  $t_i$ , and let  $u_i$  be the number of animals that are not recovered from the  $i^{\text{th}}$  cohort, so that  $u_i = R_i - \sum_j d_{i,j}$ . The probability corresponding to the  $d_{i,j}$  is denoted by  $p_{i,j}$ , and we write  $q_i = 1 - \sum_j p_{i,j}$ ,  $i = 1, \dots, T$ . Making use of the assumption of independence of individuals between cohorts, the MRR data can be modelled by a product of multinomials, and the log-likelihood is given by

$$\log(L_r) = \text{constant} + \sum_{i=1}^T \sum_{j=i+1}^T d_{i,j} \log(p_{i,j}) + \sum_{i=1}^T u_i \log(q_i). \quad (2)$$

The  $p_{i,j}$  are modelled in terms of annual survival and reporting probabilities. For illustration suppose that parameters do not vary with time, and let  $\phi_\ell$  denote the annual survival probability of individuals in the  $\ell^{\text{th}}$  year of life,  $\ell = 1, \dots, a-1, a^+$ , and  $\lambda$  denote the probability that an individual which dies is reported dead. Thus there are  $a$  age groups, and  $a^+$  refers to all individuals in the  $a^{\text{th}}$

116 and older years of life. For example, for  $j \leq a^+$ ,  $p_{i,j} = \prod_{k=1}^{j-1} \phi_k(1 - \phi_{j-i+1})\lambda$ . However, probabilities  
 117 will be time dependent in general (McCrea and Morgan 2014, p62).

118 3.2 Likelihood formation: census data

119 Census data form a time series that can be modelled by means of a state-space model (see also Pat-  
 120 terson *et al.* (2017), who discuss the relationship with hidden Markov models, and the Kalman filter,  
 121 and uses for modelling movement data). State-space models are based on two equations, a transition  
 122 equation and an observation equation. In our application the states are discrete. The transition equa-  
 123 tion describes how a population changes over time, through a Leslie matrix with elements which are  
 124 functions of parameters of survival and productivity. The census itself is then assumed to result from  
 125 observations made on the states of the underlying process. For further discussion, see King (2012),  
 126 King (2014) and Newman *et al.* (2014).

127 In the case of Grey herons, we assume that birds breed after the second year of life, and that the  
 128 transition equation of the model has the general form

$$\begin{pmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{a-1,t} \\ N_{a^+,t} \end{pmatrix} = \begin{pmatrix} 0 & p\phi_1 & \dots & p\phi_1 & p\phi_1 & p\phi_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \phi_2 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \phi_{a-1} & 0 & 0 \\ 0 & 0 & \dots & 0 & \phi_{a^+} & \phi_{a^+} \end{pmatrix} \begin{pmatrix} N_{1,t-1} \\ N_{2,t-1} \\ \vdots \\ N_{a-1,t-1} \\ N_{a^+,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{a-1,t} \\ \epsilon_{a^+,t} \end{pmatrix} \quad (3)$$

129 where  $N_{r,t}$ ,  $1 \leq r \leq a - 1$  and  $N_{a^+,t}$  denote, respectively, the unknown numbers of female birds aged  
 130  $r$ -years and greater than  $(a - 1)$ -years at time  $t$  and the  $\{\epsilon_{i,t}\}$  denote appropriate process errors,  
 131 which are discussed below. For Grey herons, based on analysis of the ring-recovery data in Besbeas  
 132 and Morgan (2012), we take  $a = 4$ .

133 An important feature in integrated population modelling is that the variances of the process errors  
 134 are not free parameters, to be estimated, which is frequently the case in modelling abundance data  
 135 alone. Rather they have a particular structure resulting from the modelling: in their simplest forms,  
 136 survival may correspond to binomial sampling, and recruitment due to productivity is assumed to  
 137 follow from a Poisson distribution; see Besbeas *et al.* (2002). Thus corresponding to Equation 3, when  
 138  $a = 4$  we have

$$\begin{aligned}
\text{Var}(\epsilon_{1,t}) &= p\phi_1(N_{2,t-1} + N_{3,t-1} + N_{4+,t-1}) \\
\text{Var}(\epsilon_{2,t}) &= \phi_2(1 - \phi_2)N_{1,t-1} \\
\text{Var}(\epsilon_{3,t}) &= \phi_3(1 - \phi_3)N_{2,t-1} \\
\text{Var}(\epsilon_{4+,t}) &= \phi_{4+}(1 - \phi_{4+})(N_{3,t-1} + N_{4+,t-1}).
\end{aligned}
\tag{4}$$

139 The first of these expressions corresponds to assuming a Poisson form for recruitment and the  
140 remaining result from the binomial model for survival.

141 Grey herons are not thought to skip breeding in general, and so it is acceptable to equate counting  
142 nests to counting breeding females. The situation is different for longer lived birds, which do skip  
143 breeding. The observed counts,  $\{y_t\}$ , are then given by the corresponding observation equation

$$y_t = (0, 1, 1, 1)(N_{1,t}, N_{2,t}, N_{3,t}, N_{4+,t})' + \eta_t, \tag{5}$$

144 where for all  $t$  we further assume that the observation error,  $\eta_t \sim N(0, \sigma^2)$ , with the observation  
145 variance  $\sigma^2$ , measuring the accuracy of the census, being a crucially important parameter to be  
146 estimated. Note that the normal distribution is justified as inactive nests might be counted in error,  
147 and in addition data might be rounded. As in the modelling of MRR data, parameters will be time-  
148 dependent in general.

149 It is shown in Besbeas *et al.* (2002) how the Kalman filter can be simply used to produce an ap-  
150 proximate likelihood for census data, and an additional simplification is given in Besbeas *et al.* (2003),  
151 which is valuable if a component likelihood is constructed using a stand-alone computer package, such  
152 as Program MARK; see <http://www.phidot.org/software/mark/docs/book/>. The theory of the  
153 Kalman filter is based upon the assumption of normally distributed random variables, and in Besbeas  
154 *et al.* (2002) the Poisson and binomial forms are taken to be approximately normal. Typically obser-  
155 vations are sufficiently large to justify this, which is certainly true of the Grey heron data; see Figure  
156 1. The work of Brooks *et al.* (2004) provides a Bayesian analysis of integrated population models,  
157 which does not make use of the Kalman filter and the normality assumptions, and results there show  
158 how robust these normal approximations are, even for small population sizes. This conclusion is also  
159 supported by simulation studies such as those of Section 6. For further details of the Kalman filter,  
160 see Durbin and Koopman (2001), and for more discussion of integrated population modelling, see  
161 Chapter 12 of McCrea and Morgan (2014) and Chapter 9 of Newman *et al.* (2014).

### 162 3.3 Parameter structure for the case study

163 We can expect time variation in survival probabilities, productivity and the reporting probability,  $\lambda$ , of  
164 dead ringed birds, which has been confirmed by various analyses such as Besbeas and Morgan (2012).  
165 Here we shall assume constant productivity, a reporting probability that may be fully time dependent

166 or regressed logistically on time, and survival probabilities that may be constant, fully time-dependent  
167 or are each regressed logistically on a measure of winter severity, the number of frost days recorded at  
168 a central England location,  $w_t$ ; see Besbeas and Morgan (2012). The reason for this is that the heron  
169 feeds on aquatic food, and cold winters can result in frozen ponds, thereby reducing food availability;  
170 this is a particular problem for younger birds. Additionally, the reporting probability of dead ringed  
171 birds has been decreasing over time in the UK, explaining the need for temporal variation in  $\lambda$ .

172 We give the notation for the model parameters in Table 1.

173 [Table 1 about here.]

### 174 3.4 Estimated correlation matrix for the case study

175 [Table 2 about here.]

176 We give in Table 2 the estimated correlation matrix after an integrated analysis of the heron  
177 data. A diffuse start has been used for the Kalman filter; see Besbeas and Morgan (2011). Although  
178 there is relevant discussion in Besbeas *et al.* (2002) from fitting a simpler model, we do not believe  
179 that such a table has been presented previously. We note the generally low values, in absolute terms,  
180 of the correlations, and in particular in the bottom row, where correlations relate the observation  
181 variance to the other model parameters. The largest absolute value in this row (0.2562) is indicated  
182 in bold face, and corresponds to the correlation with the slope of the first year survival regression.  
183 This is due to the fact that without regressing first-year survival upon the weather covariate, the  
184 model does not describe the data well, resulting in a larger estimate of  $\sigma$ . Also notable in the table  
185 are the high correlations, in absolute terms, shown in bold face, in the penultimate line of the table,  
186 corresponding to the productivity,  $p$ , being negatively correlated with the estimated intercepts of the  
187 survival probability regressions. This is due to the complete confounding of  $p$  and  $\phi_1$  in the census  
188 likelihood, seen in Equation 3, and the correlations between the estimated intercepts of the survival  
189 probability regressions. We shall comment further on the implications of the low correlations between  
190  $\log(\sigma^2)$  and the other parameters of Table 2 later in the paper.

## 191 4 Apportioning variance correctly when fitting state-space models

192 State-space models have been used for modelling abundance data alone, for simpler models than  
193 that presented in the last section. An illustration is provided by Dennis *et al.* (2006), who modelled  
194 abundance on the logarithmic scale. When only abundance data are being modelled, a difficulty may  
195 arise from the need to distinguish the two types of variances in the model: the process variances and  
196 the observation variances. See also Freckleton *et al.* (2006) for more discussion of this feature. Accurate  
197 determination of the observation variance requires replication of time series, corresponding to more



198 than a single independent observation at each time. This point has been made by Dennis *et al.* (2010),  
 199 and is investigated further by Knappe *et al.* (2013). Replication can arise from measurements being  
 200 made on different segments of the population, and an illustration of this is provided by Tavecchia  
 201 *et al.* (2009), where separate censuses are made of lambs and male and female adult Soay sheep,  
 202 *Ovis aries*. However typically in such a case different observation variances are associated with each  
 203 different type of census observation, so that this does not provide the desired replication. As pointed  
 204 out by Dennis *et al.* (2010), in practice replicated surveys of wild animals are generally hard to carry  
 205 out. See Chapter 11 of McCrea and Morgan (2014) and Newman *et al.* (2014) for more discussion of  
 206 state-space modelling.

207 For integrated population modelling there are two protections against the difficulty of correctly  
 208 allocating variance between process and observation. The first results from the nature of the modelling,  
 209 as the census data are not being analysed in isolation. In the heron illustration, for example, the  
 210 ring-recovery likelihood produces information on annual survival probabilities, which stabilises the  
 211 estimation and assists in the correct estimation of the two types of variance in integrated modelling.  
 212 The second protection arises from the structural specification of the process variances, illustrated in  
 213 Equation 4, which prevents them from being free parameters, as discussed above.

214 The motivation for this paper has been our experience that there can still be difficulties with  
 215 correctly estimating the state-space model variances in integrated population modelling. We now give  
 216 two illustrations in the next section, where the problem arises with the estimation of observation  
 217 variance.

## 218 **5 Two examples of complexity in integrated population models**

219 Integrated population modelling opens up possibilities for making models more realistic, and one  
 220 example of this has been mentioned already, when Besbeas and Morgan (2012) introduced a complex  
 221 structure for modelling heron productivity. We give here two instances of such elaboration giving rise  
 222 to unrealistic, zero estimates of observation variance. The context is again models for the heron data.

### 223 **5.1 Overdispersion**

224 We first consider an important case of adding overdispersion to both the binomial survival process of  
 225 the state-space model and also to the ring-recovery model, using the heron example as an illustration.  
 226 Extra-binomial variation might arise for several reasons, such as individual heterogeneity, the failure  
 227 to include relevant covariates, etc. We can do this in various ways, and for the ring-recovery model  
 228 we introduce a beta,  $\text{Beta}(\alpha, \beta)$ , distribution, with probability density function given by

$$f(\phi) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \phi^{\alpha-1}(1 - \phi)^{\beta-1}, \quad 0 \leq \phi \leq 1.$$

229 This is done for just the adult survival probability  $\phi_{4+}$ ; the motivation here arises from the fact  
 230 that the age class for the oldest individuals spans multiple ages and individuals may survive with  
 231 different survival probabilities, for example due to senescence; see for example Burnham and Rexstad  
 232 (1993) and Pollock and Raveling (1982). However overdispersion may be due to a range of features  
 233 not included in the simpler models, such as temporal and spatial variation in the environment . It  
 234 is useful to reparameterise the beta distribution in terms of its mean,  $\mu = \alpha/(\alpha + \beta)$ , and precision  
 235 parameter  $\theta = 1/(\alpha + \beta)$ , which allows  $\mu$  to be time-dependent or depend on covariates. The resulting  
 236 cell probabilities for the recovery matrix are then as shown in Table 3, for the illustrative case  $a = 2$   
 237 and time-varying reporting probability  $\lambda$ .

238 [Table 3 about here.]

239 The parameter  $\theta$  is an over-dispersion index, and setting  $\theta = 0$  removes the overdispersion. There  
 240 are several ways to add overdispersion to a state-space model binomial variance, including, for example,  
 241 a simple scaling of  $\text{Var}(\epsilon_{4+,t})$ . Here we use the expression of Equation 6, taken from Besbeas *et al.*  
 242 (2009):

$$\text{Var}(\epsilon_{4+,t}) = (N_{3,t-1} + N_{4+,t-1})\mu(1 - \mu) \left\{ 1 + \frac{\theta(N_{3,t-1} + N_{4+,t-1})}{1 + \theta} \right\}. \quad (6)$$

243 When  $\theta = 0$  here the parameter  $\mu$  reduces to  $\phi_{4+}$ , and Equation 6 reduces to Equation 4. Thus  
 244 the parameter  $\theta$  appears in both likelihood components of the integrated model in Equation 1. The  
 245 introduction of the new parameter  $\theta$  which enters this process variance prevents the variance from  
 246 being fully determined by the other model parameters, and can result in boundary estimation of the  
 247 observation variance. This is demonstrated in the profile log-likelihoods illustrated in Figure 2.

248 In general, for brevity we specify integrated models using forms such as *ccct/t/c*. Here the first four  
 249 letters correspond to how the survival probabilities are modelled, in order of age, the fifth corresponds  
 250 to describing the reporting probability and the sixth determines how productivity is described; *c* indi-  
 251 cates a constant parameter, *t* indicates full time dependence, with a separate parameter for each year,  
 252 and *v* indicates that the relevant parameter is regressed on a covariate. For the survival probabilities,  
 253 the covariate dependence we consider is logistic dependence on the winter weather covariate, while  
 254 in the case of  $\lambda$ , the only covariate dependence considered is logistic dependence on time. We can  
 255 see from Figure 2 that as model complexity increases, while the estimate of  $\theta$  steadily decreases, the  
 256 more dramatic effect is upon the estimate of  $\sigma$ , which ends up, for the *vvvv/v/c* model, on the  $\sigma = 0$   
 257 boundary. It is not necessary to have overdispersion in the ring-recovery model for this feature to  
 258 arise, and in fact performance is worse if  $\theta$  is estimated from census data alone; related results are  
 259 not shown here. We note also that in principle all four process variances in Equation 4, corresponding  
 260 to the four age classes in the model, could be over dispersed, which might result in more potential  
 261 boundary estimates.

[Fig. 2 about here.]

Due to the consistency of maximum-likelihood estimators, increasing the length of the time series will reduce the incidence of boundary estimates for the observation variance. We can see this from the simulated results of Table 4, however it is noticeable that for the complex model case very long time series would be necessary in order to reduce appreciably the incidence of boundary estimates.

[Table 4 about here.]

## 5.2 Time-variation

Zero estimates of observation variance can also arise without overdispersion in the model, but with process model flexibility. This is demonstrated in Table 5 for a range of models involving time-dependent parameters. Shown in the table is the maximum-likelihood estimate of  $\sigma^2$ , and we can see that several models result in the estimate  $\hat{\sigma} = 0$ .

[Table 5 about here.]

In terms of AIC, the best model for the data has logistic weather covariate dependence of each of the survival probabilities and time dependence for the recovery and productivity parameters,  $vvvv/t/t$ , and also results in  $\hat{\sigma} = 0$ ; results not provided here. In practical terms this boundary estimate is unrealistic and may suggest that the data are being over-fitted. As already observed above, it is tempting to add time-dependence to model parameters such as productivity, as has been done in a particular way by Besbeas and Morgan (2012), and this runs the risk of boundary estimation of the observation variance. In addition, Besbeas and Morgan (2012) demonstrate good fit of their model to the data.

## 6 Alternative approaches to dealing with zero estimates of observation variance in integrated population models

We present and investigate three alternative approaches.

### 6.1 Plug in value for $\sigma$

In the absence of an external data-based value, we examine alternative possibilities, based on fitting a cubic spline to the census data. In order to produce a simple check of their results, in Besbeas *et al.* (2002) the authors fitted a spline to the time series data and formed the standard deviation of the jackknife residuals, obtaining a value of 288, which they regarded as the same order of magnitude as  $\hat{\sigma} = 465(43)$ , resulting from their integrated population modelling. This approach could always be

291 applied, and the resulting estimate of  $\sigma$  used in subsequent analysis. However it is over simplistic, as  
 292 the spline simply treats the time series of census values as a sequence without structure, in particular  
 293 without underlying process variability. Furthermore, the plug-in approach might result in conservative  
 294 error estimation for the other model parameters, though simulation results later in Section 6 suggest  
 295 that this is not an important issue. See also Francis (2011).

## 296 6.2 Penalised likelihood

297 Another approach that might be adopted is to penalise the log likelihood for the census data, in order  
 298 for the likelihood maximisation to avoid the boundary  $\sigma = 0$ . Here one simply adds  $\alpha h(\sigma)$  to the log  
 299 likelihood for suitable penalty parameter  $\alpha > 0$  and penalty function  $h(\sigma)$ ; see for example Wang and  
 300 Lindsay (2005). This is equivalent to using an informed prior distribution on  $\sigma$  in a Bayesian analysis  
 301 of the problem. We demonstrate the use of this approach in Figures 3 and 4, respectively for the two  
 302 cases where we assume that  $\theta = 0$  and  $\theta > 0$ , where the model is  $vvvv/v/c$ . Penalising the likelihood  
 303 works well, but suffers from the need to decide on a suitable function for  $h(\sigma)$ , and how to choose  $\alpha$ :  
 304 we simply used the logarithmic function for  $h(\sigma)$ .

305 Thus instead of the joint likelihood  $L_j$  we maximised the penalised log likelihood given by

$$\log(L_j^p) = \log(L_r) + \log(L_c) + \alpha \log(\sigma). \quad (7)$$

306 Other penalty functions were also investigated, with similar conclusions resulting.

307 [Fig. 3 about here.]

308 [Fig. 4 about here.]

309 From Figure 3 we can see that when there is no assumption of overdispersion the effect of increas-  
 310 ing  $\alpha$  is primarily to increase the estimate of  $\sigma$ , and the remaining parameter estimates are largely  
 311 unchanged. For this model there is no boundary estimation, and the effect on estimating  $\sigma$  is to be  
 312 expected, from Equation 7. It is interesting to observe the behaviour in Figure 4 when  $\alpha$  increases.  
 313 Increasing the penalty is seen to increase the influence of the ring-recovery data. An important gen-  
 314 eral conclusion which might be drawn is that the parameters of interest are stable with respect to the  
 315 choice of  $\alpha$ . A natural approach to choosing  $\alpha$  is to use cross validation (Green and Silverman 1994,  
 316 p30). However this approach is time consuming, and in this application it gives  $\alpha = 0$ , as demonstrated  
 317 by Figure 1 in the Online Resource for the paper. Therefore we shall not investigate this approach  
 318 further in the simulation experiments reported in Section 6.

319 In the next section we revisit the observation that what is needed in order to estimate  $\sigma$  well, even  
 320 in the complex cases that we have considered, is replicated census data. So that ideally what we seek  
 321 is to augment Equation 5 with

$$\tilde{y}_t = (0, 1, 1, 1)(N_{1,t}, N_{2,t}, N_{3t,t}, N_{4+,t})' + \tilde{\eta}_t, \quad (8)$$

where  $\tilde{\eta}_t$  is independent of and with the same distribution as  $\eta_t$ .

### 6.3 Pseudo replication

We have found, for the heron data, and also for other similar ecological time series that after first-order differencing there remains little structure in the series. This is an interesting finding, and one which suggests that the first-order Markov modelling that we employ is appropriate for these data. It also supports the use of pseudo replication.

As the census data are clearly non-stationary, standard bootstrap approaches for time series are inappropriate; a review of such methods is provided by Gonçalves and Politis (2011). Shown in Figure 1 are four alternative pseudo replicates of the census data. These are obtained in an ad hoc fashion, by at each appropriate time, ie., apart from the first and last, selecting a value observed at either the current, previous or next time, each with probability 1/3; see also eg, Rice (1984). This approach can be extended to include the observed data at other neighbouring time points; see for example the results of Figure 2 in the Online Resource for the paper, where neighbouring points which are two time points before and after the current time point are also included. For stationary time series this approach would not change expectations. However in general it induces a correlation between the real data and the pseudo replicate at each time point. It is easy to show that this correlation decreases with increasing observation variance, in accordance with intuition, and also increases with the serial covariances of the time series. We can see, in comparison with Figure 1, the extra variation that results from sampling from 5 alternatives as compared with just sampling from 3. We have found that after first-order differencing the simple pseudo replicated series, based on just three values at each time point, has the appearance of a first-order moving average series, as one might expect. The pseudo replicates are used to augment Equation 5, as in Equation 8.

We start by testing the use of pseudo replication when there is no boundary estimate of the observation variance,  $\sigma^2$ . The model fitted in Table 6 is  $vvvv/v/c$ , without overdispersion. We see from Table 6 that the main change as a result of replication is the reduction in the estimate of the standard error of  $\hat{\sigma}$ . As a consequence of the low correlations between  $\hat{\sigma}$  and all the other model parameter estimators, seen in Table 2, this is to be expected.

[Table 6 about here.]

In Table 7 we present results for when there is heterogeneity present in the model  $vvvv/v/c$ . Adding overdispersion in fact improves the fit of the model to the data. We compare the use of 3 and 5 neighbouring values in the replication. The results from using pseudo replicates are very good. We can see from Table 7 that the primary effect of using more points for the pseudo replication lies in increasing the estimate of  $\sigma^2$ , which is to be expected, and also in reducing its estimate of standard

error, which is also as one might expect. As with using penalised likelihood, we find that the main effect of pseudo replication is on the estimate of  $\sigma$  and its standard error.

[Table 7 about here.]

A referee has observed that other possibilities for pseudo replication could be considered. For instance, the model could just be fitted to all data replicated for each year prior and after the actual year. We have not considered the properties of such a scheme.

## 7 Simulation

In order to compare the methods, and explore the wider relevance of the results from the observed data, we analyse simulated data based on the fitted heron models. We shall vary the amount of overdispersion ( $\theta = 0, 0.03, 0.0553, 0.07$ ), the size of the measurement error ( $\log(\sigma^2) = 10, 11, 12.258$ ) and the amount of ring-recovery data in the joint analysis (complete, 50%, 25% and 10%). Remaining parameter values are given in Table 1. There were 200 replications for each parameter combination. The reduction of the MRR data by 90% results in a major loss of MRR information, and the resulting table of recorded deaths is very sparse. For brevity we only present a small, representative sample of extensive results.

First we present in Table 8 root mean square error (RMSE) results from one of the cases for when there is no overdispersion, and no boundary estimates for  $\sigma$ .

[Table 8 about here.]

What we see from Table 8 is an improvement in estimation from moving from MRR only to IPM, and this is especially marked for the reduced MRR case, as expected; we see that estimating  $\sigma$  using pseudo replication increases some RMSEs slightly in (a), due to the fact that the census data no longer follow the model exactly, though that is not generally the case in (b), and that plugging in estimates for  $\sigma$  has little effect. We consider this last feature to be a consequence of the low correlations between the estimator of  $\sigma$  and the other estimators in Table 2. However it is also a matter of using an appropriate value, and using very large values for  $\sigma$  has the effect of diminishing the contribution of the census data (results not shown).

In Table 9 we present the results from one of the cases when there is overdispersion:  $\theta = 0.07$ , focussing on the 38 instances out of the 200 simulations taken which resulted in boundary estimates in each of cases (a) and (b). Note that the RMSE for  $\sigma$  from integrated population modelling is necessarily relatively small, due to the fact that the observation variance is estimated as zero in these cases and there is no variance contribution.

Again, the table shows the improvement that arises from integrated modelling, and how that varies, in a sensible way, with respect to the thinning of the ring-recovery data, despite the boundary

388 estimation. The performance of pseudo replication for estimating  $\sigma$  is seen to depend on the amount  
389 of ring-recovery data present, performance diminishing as that information is reduced. However re-  
390 maining parameter estimates are remarkably stable.

391 Regarding estimating observation variance when there is no boundary estimate, relevant histograms  
392 are provided in Figures 3 and 4 in the Online Resource for the paper. We can see that using the spline  
393 approach can result in large overestimation, due to the fact that there is no modelling of the census  
394 data, and thus the variance estimated using the spline accounts for both types of variation, process  
395 and observation variability. In fact in results not shown here we see this overestimation increased when  
396 we add overdispersion to the simulations ( $\theta > 0$ ). The overestimation of  $\sigma$  is a function of the value  
397 of  $\sigma^2$  used in the simulations, with increased overestimation corresponding to smaller  $\sigma^2$  used. This is  
398 due to the proportionally greater effect of the process variance when the observation error is smaller.  
399 For the particular cases investigated we found that plugging in half of the  $\sigma^2$  value estimated using  
400 the spline approach performed well. We see also that the joint analysis provides good estimates of the  
401 observation variance, and that the addition of a single pseudo replicate results in a more positively  
402 skewed distribution of estimates of observation variance.

403 [Table 9 about here.]

## 404 8 Conclusion

405 Naturally, the conclusions we draw are based on just the one case study considered, and the simulations  
406 based upon that. One might expect results to depend upon the relative strength of information in the  
407 two data sets being analysed. As one check of this we have repeated analyses using 50%, 75% and  
408 90% thinning of the ring-recovery information and obtained broadly similar conclusions. Additionally,  
409 for pseudo replication we have only presented information for where there was only a single pseudo  
410 replicate, but the effect of taking multiple replicates has also been considered. Results are as expected,  
411 and it appears that only single replication is needed, to ensure that  $\hat{\sigma} > 0$ .

412 The primary finding of this investigation of integrated population modelling is that the estimates of  
413 parameters of primary interest which arise when the observation variance is estimated as zero appear  
414 to be reliable. This is counter intuitive, and in particular is in contrast to the findings of de Valpine and  
415 Hastings (2002), de Valpine and Hilborn (2005), Knappe and Korner-Nieverage (2015) and Maunder  
416 *et al.* (2015) in the context of working just with time series data of population abundances.

417 Should an estimate of observation variance be required when we obtain  $\sigma = 0$ , then it is simple  
418 to use what results from use of splines, perhaps in conjunction with the estimate from using pseudo  
419 replication. However the value obtained might be regarded as an upper bound.

420 **Acknowledgements** We thank the Associate Editor, Roland Langrock, two anonymous referees, Stephen Freeman,  
421 Mark Maunder, Leo Polanski and Martin Ridout for their very helpful comments.

## 422 9 Appendix

423 Here we give the short R program for fitting a cubic spline, using cross validation, and evaluating the  
424 standard deviation of the jackknife residuals. In the program  $x$  is a vector of the census years and  $y$  is  
425 a vector of the census values.

426

```
427 fit <- smooth.spline(x,y,CV=TRUE)    # smooth spline fit  
428 res <- (fit$yin - fit$y)/(1-fit$lev) # jackknife residuals  
429 sigma <- sqrt(var(res))             # estimate standard deviation
```

430

431

## 432 References

- 433 Abadi, F., Gimenez, O., Arlettaz, R. and Schaub, M. “An assessment of integrated population models: bias, accuracy,  
434 and the violation of the assumption of independence.” *Ecology* **91** 7–14 (2010)
- 435 Barry, S. C., Brooks, S. P., Catchpole, E. A. and Morgan, B. J. T. The analysis of ring-recovery data using random  
436 effects. *Biometrics*, **59**, 54–65, (2003)
- 437 Bengtsson, T. and Cavanaugh, J. E. An improved Akaike information criterion for state-space model selection. *Com-*  
438 *putational Statistics & Data Analysis*, **50**, 2635–2654, (2006)
- 439 Besbeas, P., Borysiewicz, R. S. and Morgan, B. J. T. Completing the ecological jigsaw. In *Modelling demographic pro-*  
440 *cesses in marked populations*. D. L. Thomson, E.G. Cooch, and M. J. Conroy (Eds.) Springer series: *Environmental*  
441 *and Ecological Statistics*, **3**, 513–540, (2009)
- 442 Besbeas, P., Freeman, S. N., Morgan, B. J. T. and Catchpole, E. A. Integrating mark-recapture-recovery and census  
443 data to estimate animal abundance and demographic parameters. *Biometrics* **58**, 540–547, (2002)
- 444 Besbeas, P., Lebreton, J.-D. and Morgan, B. J. T. The efficient integration of abundance and demographic data. *Applied*  
445 *Statistics*, **52**, 95–102, (2003)
- 446 Besbeas, P., McCrea, R. S. and Morgan, B. J. T. Integrated population model selection in ecology. *University of Kent*  
447 *Technical Report*, <https://kar.kent.ac.uk/id/eprint/48039>.
- 448 Besbeas, P. and Morgan, B. J. T. Kalman filter initialisation for integrated population modelling. *Applied Statistics*,  
449 **61**, 151–162, (2011)
- 450 Besbeas, P. and Morgan, B. J. T. A threshold model for heron productivity. *Journal of Biological, Agricultural and*  
451 *Environmental Statistics*, **17**, 128–141, (2012)
- 452 Besbeas, P. and Morgan, B. J. T. Goodness of fit of integrated population models using calibrated simulation. *Methods*  
453 *in Ecology and Evolution*, **5**, 1373–1382, (2014)
- 454 Brooks, S. P., King, R. and Morgan, B. J. T. A Bayesian approach to combining animal abundance and demographic  
455 data. *Animal Biodiversity and Conservation* **27**, 515–529, (2004)
- 456 Burnham, K. P. and Rexstad, E. A. Modeling heterogeneity in survival rates of banded waterfowl. *Biometrics*, **49**,  
457 1194–1208, (1993)
- 458 Chandler, R and Clark, J Spatially explicit integrated population models. *Methods in Ecology and Evolution*, **5**, 1351–  
459 1360, (2014)
- 460 Dennis, B. , Ponciano, J. M., Lele, S. R., Taper, M. L. and Staples, D. F. Estimating density dependence, process  
461 noise and observation error. *Ecological Monographs*, **76**, 323–341, (2006)

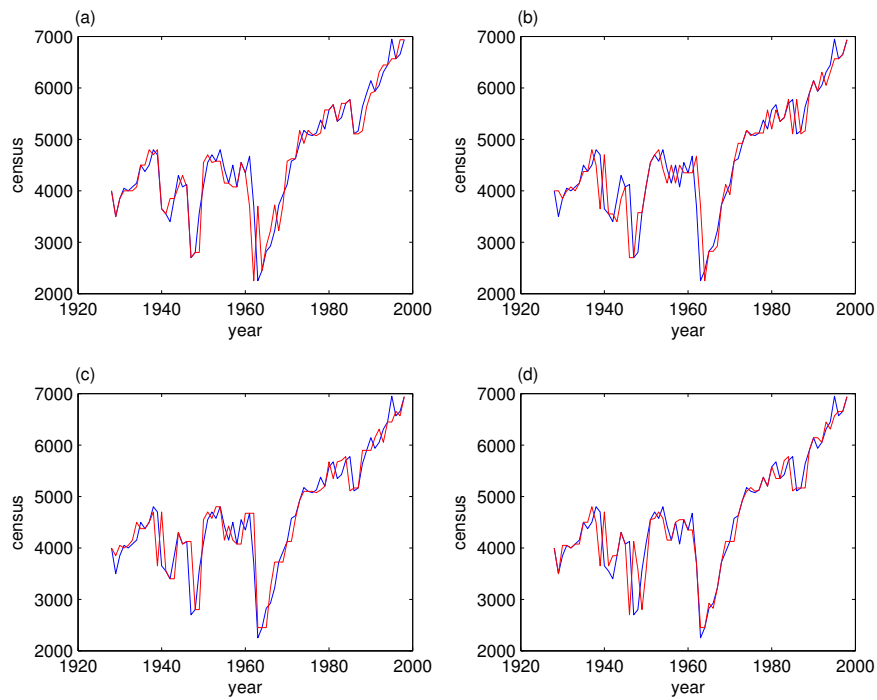


- 462 Dennis, B., Ponciano, J. M. and Taper, M. L. Replicated sampling increases efficiency in monitoring biological popula-  
463 tions. *Ecology*, **91**, 610–620, (2010)
- 464 de Valpine, P. and Hastings, A. Fitting population models incorporating process noise and observation error. *Ecological*  
465 *Monographs*, **72**, 57–76, (2002)
- 466 de Valpine, P. and Hilborn, R. State-space likelihoods for nonlinear fisheries time-series. *Can. J. Fish. Aquat. Sci.*, **62**,  
467 1937–1952, (2005)
- 468 Durbin, J. and Koopman, S. J. *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford, (2001)
- 469 Francis, R. I. C. C. Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci.*, **68**,  
470 1124–1138, (2011)
- 471 Freckleton, R. P., Watkinson, A. R., Green, R. E. and Sutherland, W. J. Census error and the detection of density  
472 dependence. *Journal of Animal Ecology*, **75**, 837–851, (2006)
- 473 Gonçalves, S. and Politis, D. Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean*  
474 *Statistical Society*, **40**, 383–386, (2011)
- 475 Green, P. and Silverman, B. *Nonparametric Regression and Generalized Linear Models: A roughness penalty approach*  
476 Chapman & Hall/CRC Press, Boca Raton, (1994)
- 477 Kéry, M. and Schaub, M. *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*, Academic Press,  
478 MA, USA, (2012)
- 479 King, R. A review of Bayesian state-space modelling of capture-recapture-recovery data. *Interface Focus*, **2**, 190–204,  
480 (2012)
- 481 King, R. Statistical Ecology. *Annual Review of Statistics and its Application*, **1**, 401–426, (2014)
- 482 Knape, J. Estimability of density dependence in models of time series data. *Ecology*, **89**, 2994–3000, (2008)
- 483 Knape, J., Besbeas, P. and de Valpine, P. Using uncertainty estimates in analyses of population time series. *Ecology*,  
484 **94**, 2097–2107, (2013)
- 485 Knape, J. and Korner-Nievergelt, F. Estimates from non-replicated population surveys rely on critical assumptions.  
486 *Methods in Ecology and Evolution*, doi:10.1111/2041-210X.12329, (2015)
- 487 McCrea, R. S. and Morgan, B. J. T. *Analysis of Capture-recapture Data*, CRC Chapman & Hall, Boca Raton, (2014)
- 488 McCrea, R. S., Morgan, B. J. T., Gimenez, O., Besbeas, P., Bregnballe, T. and Lebreton, J.-D. Multi-site integrated  
489 population modelling. *Journal of Biological, Agricultural and Environmental Statistics*, **15**, 539–561, (2010)
- 490 Maunder, M. N., Deriso, R. B. and Hanson, C. H. Use of state-space population dynamics models in hypothesis testing:  
491 advantages over simple log-linear regressions for modeling survival, illustrated with application to longfin smelt.  
492 (*Spirinchus thaleichthys*). *Fisheries Research*, **164**, 102–111, (2015)
- 493 Mazzettta, C., Morgan, B. J. T. and Coulson, T. A state-space modelling approach to population size estimation.  
494 Technical report, University of Kent Technical Report: UKC/SMSAS/10/025, (2010)
- 495 Newman, K. B., Buckland, S. T., Morgan, B. J. T., King, R., Borchers, D. L., Cole, D. J., Besbeas, P. T., Gimenez, O.  
496 and Thomas, L. *Modelling Population Dynamics: Model Formulation, Fitting and Assessment using State-space*  
497 *Methods*. Springer, New York, (2014)
- 498 Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L. and King, R. Statistical modelling of individual  
499 animal movement: an overview of key methods and a discussion of practical challenges. arXiv:1603.07511v3 [stat.AP],  
500 (2017)
- 501 Pollock, K. H. and Raveling, D. G. Assumptions of modern band-recovery models, with emphasis on heterogeneous  
502 survival rates. *Journal of Wildlife Management*, **46**, 88–98, (1982)
- 503 Rice, J. Bandwidth choice for nonparametric regression. *The Annals of Statistics* **12**, 1215–1230, (1984)
- 504 Schaub, M. and Abadi, F. Integrated population models: a novel analysis framework for deeper insights into population  
505 dynamics. *Journal of Ornithology* **152**, 227–237, (2011)
- 506 Searle, S. R. *Matrix Algebra Useful for Statistics*, Wiley, New York, (1982)

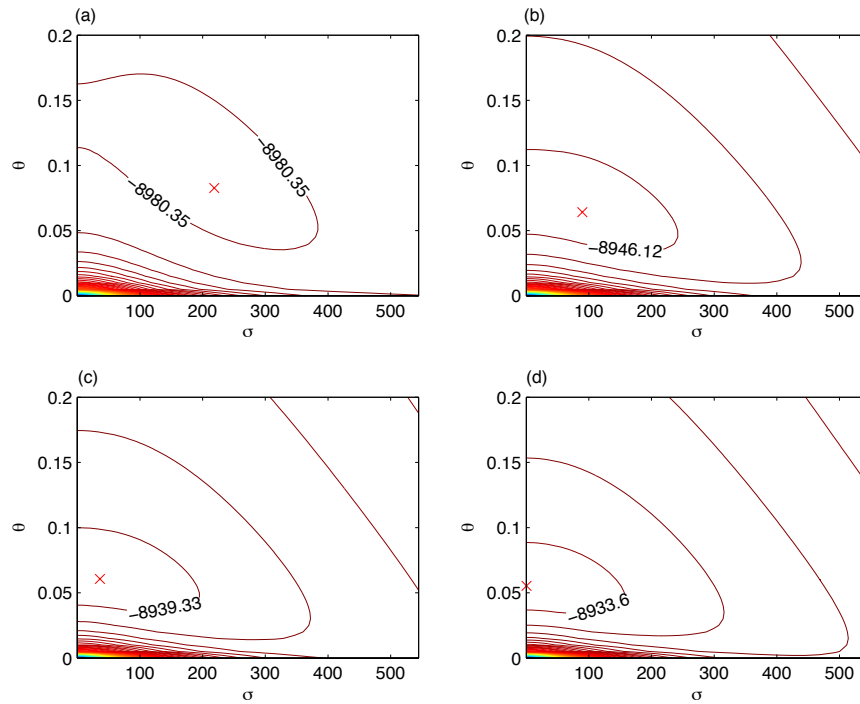
- 
- 507 Tavecchia, G., Besbeas, P., Coulson, T., Morgan, B. J. T. and Clutton-Brock, T. H. Estimating population size and  
508 hidden demographic parameters with state-space modelling. *The American Naturalist* **173**, 722–733, (2009)
- 509 Wang, J.-P. and Lindsay, B. G. A penalized nonparametric maximum likelihood approach to species richness estimation.  
510 *Journal of the American Statistical Association* **100**, 942–959, (2005)

511 **List of Figures**

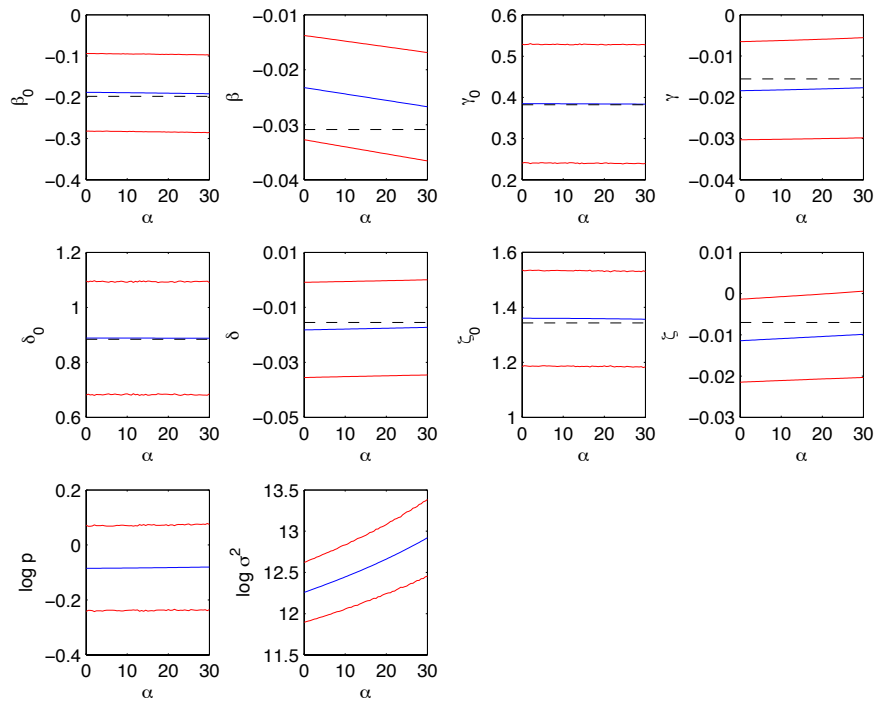
|     |   |   |    |
|-----|---|---|----|
| 512 | 1 | Examples of pseudo replication of the heron census data. The real data are shown in blue, in common in each of the four panels, (a)–(d), and four pseudo replicates are given in red. The pseudo replicates, $\tilde{y}_t, t = 2, \dots, T - 1$ , are obtained by selecting at random a value at each time that is the observed value at that time, $y_t$ , or the preceding time, $y_{t-1}$ , or the following time, $y_{t+1}$ , each with equal probability. Each panel presents a different pseudo-replicate of the census data. . . . . | 19 |
| 513 | 2 | Boundary estimation when modelling the heron data using overdispersion: profile log likelihoods with respect to $\sigma$ and $\theta$ . The indicated contours of the log-likelihood are $\chi^2_2(5\%)/2$ below the maximum value. The models fitted are: (a) $cccc/v/c$ ; (b) $vvcc/v/c$ ; (c) $vvvc/v/c$ ; (d) $vvvv/v/c$ . The maximum-likelihood estimate of $(\theta, \sigma)$ is marked with ‘ $\times$ ’. The best model in terms of AIC is $vvvv/v/c$ , with the boundary estimate for $\sigma$ as shown. . . . .                  | 20 |
| 514 | 3 | The effect of the value of the penalty scaling parameter $\alpha$ in penalised likelihood estimation; using all of the MRR data. The blue lines provide the maximum-likelihood estimates of the parameters, and the red lines provide 95% confidence intervals. The black dotted lines are the maximum-likelihood estimates using the recovery data alone, where these are available. In this case the model is $vvvv/v/c$ , and there is no overdispersion. Estimates tend to the MRR values as $\alpha$ increases. . . . .                | 21 |
| 515 | 4 | The effect of the value of the penalty scaling parameter $\alpha$ in penalised likelihood estimation; using all of the MRR data. The blue lines provide the maximum-likelihood estimates of the parameters, and the red lines provide 95% confidence intervals. The black dotted lines are the maximum-likelihood estimates using the recovery data alone, where available. In this case the model is $vvvv/v/c$ , and there is overdispersion. . . . .   | 22 |
| 516 |   |   |    |
| 517 |   |   |    |
| 518 |   |   |    |
| 519 |   |   |    |
| 520 |   |   |    |
| 521 |   |   |    |
| 522 |   |   |    |
| 523 |   |   |    |
| 524 |   |   |    |
| 525 |   |   |    |
| 526 |   |   |    |
| 527 |   |   |    |
| 528 |   |   |    |
| 529 |   |   |    |
| 530 |   |   |    |
| 531 |   |   |    |
| 532 |   |   |    |
| 533 |   |   |    |
| 534 |   |   |    |



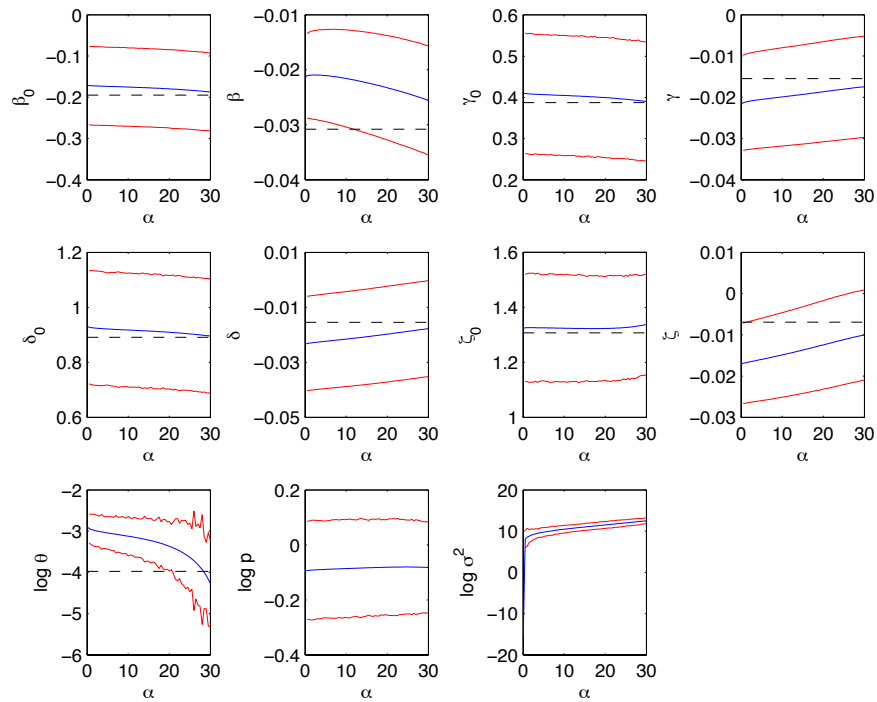
**Fig. 1** Examples of pseudo replication of the heron census data. The real data are shown in blue, in common in each of the four panels, (a)–(d), and four pseudo replicates are given in red. The pseudo replicates,  $\tilde{y}_t, t = 2, \dots, T - 1$ , are obtained by selecting at random a value at each time that is the observed value at that time,  $y_t$ , or the preceding time,  $y_{t-1}$ , or the following time,  $y_{t+1}$ , each with equal probability. Each panel presents a different pseudo-replicate of the census data.



**Fig. 2** Boundary estimation when modelling the heron data using overdispersion: profile log likelihoods with respect to  $\sigma$  and  $\theta$ . The indicated contours of the log-likelihood are  $\chi_2^2(5\%)/2$  below the maximum value. The models fitted are: (a)  $cccc/v/c$ ; (b)  $vvcc/v/c$ ; (c)  $vvvc/v/c$ ; (d)  $vvvv/v/c$ . The maximum-likelihood estimate of  $(\theta, \sigma)$  is marked with ‘ $\times$ ’. The best model in terms of AIC is  $vvvv/v/c$ , with the boundary estimate for  $\sigma$  as shown.



**Fig. 3** The effect of the value of the penalty scaling parameter  $\alpha$  in penalised likelihood estimation; using all of the MRR data. The blue lines provide the maximum-likelihood estimates of the parameters, and the red lines provide 95% confidence intervals. The black dotted lines are the maximum-likelihood estimates using the recovery data alone, where these are available. In this case the model is  $vvvv/v/c$ , and there is no overdispersion. Estimates tend to the MRR values as  $\alpha$  increases.



**Fig. 4** The effect of the value of the penalty scaling parameter  $\alpha$  in penalised likelihood estimation; using all of the MRR data. The blue lines provide the maximum-likelihood estimates of the parameters, and the red lines provide 95% confidence intervals. The black dotted lines are the maximum-likelihood estimates using the recovery data alone, where available. In this case the model is  $vvvv/v/c$ , and there is overdispersion.

535 **List of Tables**

536 1 Parameter notation and values used to generate the simulated data in Section 6 of the  
 537 paper. In an obvious notation to indicate time dependence of the parameters, we have  
 538  $\text{logit}(\phi_{1,t}) = \beta_0 + \beta w_t$ ,  $\text{logit}(\phi_{2,t}) = \gamma_0 + \gamma w_t$ ,  $\text{logit}(\phi_{3,t}) = \delta_0 + \delta w_t$ ,  $\text{logit}\phi_{4+,t} = \zeta_0 + \zeta w_t$ ,  
 539 and  $\text{logit}(\lambda_t) = \nu_0 + \nu t$ . Here  $w_t$  is a measure of winter temperature: see text. Three  
 540 alternatives for  $\sigma^2$  and four for  $\theta$  are considered. Note that all of the parameter values  
 541 are the estimates obtained from fitting the real data; see Table 6. . . . . 25

542 2 Estimated correlation matrix for integrated population modelling of Grey heron data.  
 543 Those values shown in bold face are discussed in the text. The parameter notation used  
 544 is given in Table 1. . . . . 26

545 3 Multinomial cell probabilities,  $p_{i,j}, i = 1, \dots, 3, j = 1, \dots, 4$ , for ring-recovery data  
 546 assuming overdispersion where, for illustration,  $a = 2$ . . . . . 27

547 4 A demonstration of the effect of increasing time-series sample size,  $n$ , on the frequency  
 548 of boundary estimates of observation variance. The simple model here corresponds to  
 549 model  $cccc/v/c$ , and the complex model has the form  $vvvv/v/c$ . Shown is the number  
 550 of times the zero-variance boundary was encountered for the measurement variance out  
 551 of 500 replications in each case. The true value of observation standard deviation was  
 552 taken to be  $\sigma = e^{8.9/2} = 85.63$ . . . . . 28

553 5 A demonstration of how time-variation in parameters can result in an estimated zero  
 554 observation error. In all cases there is time variation of reporting probability. The first  
 555 five models have one other instance of time variation and for the remaining models  
 556 there is one further instance of time variation. . . . . 29

557 6 A comparison between the analysis of the Grey heron data with and without the use  
 558 of pseudo replication, when there is no heterogeneity in the model, and no boundary  
 559 estimate of observation variance. The labelling of the parameters is the same as in Table  
 560 2. In the pseudo replication case we present averages of 20 separate simulations. . . . 30

561 7 A comparison between the analysis of the Grey heron data using model  $vvvv/v/c$  with  
 562 no replication and with a single replicate based on 3 or 5 neighbouring values when  
 563 there is heterogeneity in the model. In both replicate cases we present averages of 20  
 564 separate simulations. Variation between simulations was small. Standard errors for the  
 565 case of no replication, when there is a boundary estimate, are obtained for the other  
 566 parameters using an appropriate singular-value decomposition approach; (Searle 1982,  
 567 p. 318). This is the case illustrated in panel (d) of Figure 2 . The resulting estimated  
 568 standard errors can be expected to be conservative. . . . . 31

569 8 Root mean square error results (RMSEs), all multiplied by 1000, from fitting model  
 570  $vvvv/v/c$  with  $\sigma = e^{12.258/2} = 458.97$  and no overdispersion, so that  $\theta = 0$ . There were  
 571 no boundary estimates for  $\sigma$  by the IPM method in the 200 simulations undertaken.  
 572 We compare the results of just analysing the recovery data alone (MRR only), of using  
 573 integrated population modelling (IPM), of using integrated population modelling with  
 574 a single pseudo replication (IPM with pseudorep), and of plugging in various multiples  
 575 of the estimated value of  $\sigma$  obtained from fitting a spline,  $\tilde{\sigma}$ , for selected parameters.  
 576 This is done twice, (a) for when we use the full MRR data and (b) for when we use a  
 577 random 10% of the MRR data. Note that (a) and (b) result from different simulation  
 578 runs. . . . . 32



|     |  |    |
|-----|--|----|
| 579 | <p>9 Root mean square error results (RMSEs), all multiplied by 1000, from fitting model <math>vvvv/v/c</math> with <math>\sigma = e^{10/2} = 148.41</math> and overdispersion, with <math>\theta = 0.07</math>. Separately for (a) and (b), which arise from different simulation runs, results are only for the 38 cases out of 200 simulations undertaken when there were boundary estimates for IPM. We compare the results of just analysing the recovery data alone (MRR only), of using integrated population modelling (IPM), of using integrated population modelling with a single pseudo replication (IPM with pseudorep), and of plugging in various multiples of the estimated value of <math>\sigma</math> obtained from fitting a spline, <math>\tilde{\sigma}</math>, and also for the true value, <math>\sigma = 148.41</math>, for selected parameters. This is done twice, (a) for when we use the full MRR data and (b) for when we use a random 10% of the MRR data. . . . .</p> | 33 |
| 580 |  |    |
| 581 |  |    |
| 582 |  |    |
| 583 |  |    |
| 584 |  |    |
| 585 |  |    |
| 586 |  |    |
| 587 |  |    |
| 588 |  |    |

**Table 1** Parameter notation and values used to generate the simulated data in Section 6 of the paper. In an obvious notation to indicate time dependence of the parameters, we have  $\text{logit}(\phi_{1,t}) = \beta_0 + \beta w_t$ ,  $\text{logit}(\phi_{2,t}) = \gamma_0 + \gamma w_t$ ,  $\text{logit}(\phi_{3,t}) = \delta_0 + \delta w_t$ ,  $\text{logit}\phi_{4+,t} = \zeta_0 + \zeta w_t$ , and  $\text{logit}(\lambda_t) = \nu_0 + \nu t$ . Here  $w_t$  is a measure of winter temperature: see text. Three alternatives for  $\sigma^2$  and four for  $\theta$  are considered. Note that all of the parameter values are the estimates obtained from fitting the real data; see Table 6.

| parameter       | value                   |
|-----------------|-------------------------|
| $\beta_0$       | -0.188                  |
| $\beta$         | -0.023                  |
| $\gamma_0$      | 0.385                   |
| $\gamma$        | -0.018                  |
| $\delta_0$      | 0.889                   |
| $\delta$        | -0.018                  |
| $\zeta_0$       | 1.360                   |
| $\zeta$         | -0.011                  |
| $\nu_0$         | -2.027                  |
| $\nu$           | -0.832                  |
| $\log p$        | -0.085                  |
| $\log \sigma^2$ | {10,11,12.258}          |
| $\theta$        | {0, 0.03, 0.0553, 0.07} |

|                 |                |                |                |          |                |          |                |         |         |        |          |                 |
|-----------------|----------------|----------------|----------------|----------|----------------|----------|----------------|---------|---------|--------|----------|-----------------|
| $\beta_0$       | 1              |                |                |          |                |          |                |         |         |        |          |                 |
| $\beta$         | -0.0398        | 1              |                |          |                |          |                |         |         |        |          |                 |
| $\gamma_0$      | 0.1067         | -0.0233        | 1              |          |                |          |                |         |         |        |          |                 |
| $\gamma$        | -0.0165        | -0.0888        | -0.0418        | 1        |                |          |                |         |         |        |          |                 |
| $\delta_0$      | 0.0818         | -0.0160        | 0.0974         | 0.0297   | 1              |          |                |         |         |        |          |                 |
| $\delta$        | -0.0249        | -0.1192        | -0.0315        | -0.1558  | -0.2022        | 1        |                |         |         |        |          |                 |
| $\zeta_0$       | 0.1349         | -0.1476        | 0.1432         | -0.0530  | 0.1604         | -0.0501  | 1              |         |         |        |          |                 |
| $\zeta$         | -0.0430        | -0.2308        | -0.0571        | -0.3714  | -0.0495        | -0.2846  | -0.1567        | 1       |         |        |          |                 |
| $\nu_0$         | 0.1227         | -0.0275        | 0.0993         | -0.0105  | 0.0884         | -0.0182  | 0.1460         | -0.0276 | 1       |        |          |                 |
| $\nu$           | -0.0658        | -0.0174        | -0.0505        | 0.0073   | -0.0424        | 0.0067   | -0.0393        | 0.0002  | -0.1668 | 1      |          |                 |
| $\log p$        | <b>-0.4830</b> | 0.1176         | <b>-0.5313</b> | 0.0434   | <b>-0.4561</b> | 0.0867   | <b>-0.7779</b> | 0.0990  | -0.1964 | 0.0803 | 1        |                 |
| $\log \sigma^2$ | -0.0181        | <b>-0.2562</b> | -0.0049        | 0.0319   | 0.0093         | 0.0334   | 0.0061         | 0.1044  | 0.0003  | 0.0058 | 0.0026   | 1               |
|                 | $\beta_0$      | $\beta$        | $\gamma_0$     | $\gamma$ | $\delta_0$     | $\delta$ | $\zeta_0$      | $\zeta$ | $\nu_0$ | $\nu$  | $\log p$ | $\log \sigma^2$ |

**Table 2** Estimated correlation matrix for integrated population modelling of Grey heron data. Those values shown in bold face are discussed in the text. The parameter notation used is given in Table 1.

|   | Year of recovery        |                            |   |  |
|---|-------------------------|----------------------------|---|--|
|   | 1                       | 2                          | 3   | 4  |
| 1 | $(1 - \phi_1)\lambda_1$ | $\phi_1(1 - \mu)\lambda_2$ | $\phi_1 \frac{\mu}{1+\theta}(1 - \mu)\lambda_3$ | $\phi_1 \frac{\mu(\mu+\theta)}{(1+\theta)(1+2\theta)}(1 - \mu)\lambda_4$ |
| 2 |                         | $(1 - \phi_1)\lambda_2$    | $\phi_1(1 - \mu)\lambda_3$                      | $\phi_1 \frac{\mu}{1+\theta}(1 - \mu)\lambda_4$                          |
| 3 |                         |                            | $(1 - \phi_1)\lambda_3$                         | $\phi_1(1 - \mu)\lambda_4$   |

**Table 3** Multinomial cell probabilities,  $p_{i,j}$ ,  $i = 1, \dots, 3$ ,  $j = 1, \dots, 4$ , for ring-recovery data assuming overdispersion where, for illustration,  $a = 2$ .

| $n$ | Model $cccc/v/c$<br>$\theta$ |      |      | Model $vvvv/v/c$<br>$\theta$ |      |      |
|-----|------------------------------|------|------|------------------------------|------|------|
|     | 0.01                         | 0.03 | 0.04 | 0.01                         | 0.03 | 0.04 |
| 51  | 2                            | 11   | 22   | 40                           | 118  | 144  |
| 71  | 2                            | 10   | 28   | 31                           | 114  | 145  |
| 91  | 1                            | 4    | 23   | 29                           | 103  | 137  |
| 111 | 1                            | 4    | 10   | 13                           | 97   | 133  |

**Table 4** A demonstration of the effect of increasing time-series sample size,  $n$ , on the frequency of boundary estimates of observation variance. The simple model here corresponds to model  $cccc/v/c$ , and the complex model has the form  $vvvv/v/c$ . Shown is the number of times the zero-variance boundary was encountered for the measurement variance out of 500 replications in each case. The true value of observation standard deviation was taken to be  $\sigma = e^{8.9/2} = 85.63$ .

---

| Model           | $\hat{\sigma}^2$ |
|-----------------|------------------|
| <i>tccc/t/c</i> | 0.0              |
| <i>ctcc/t/c</i> | 0.0              |
| <i>cctc/t/c</i> | 238.2            |
| <i>ccct/t/c</i> | 39.5             |
| <i>cccc/t/t</i> | 17.0             |
| <i>ttcc/t/c</i> | 0.0              |
| <i>tctc/t/c</i> | 0.0              |
| <i>tcct/t/c</i> | 0.0              |
| <i>tccc/t/t</i> | 16.5             |
| <i>cttc/t/c</i> | 0.0              |
| <i>ctct/t/c</i> | 0.0              |
| <i>ctcc/t/t</i> | 17.4             |
| <i>cctt/t/c</i> | 0.0              |
| <i>cctc/t/t</i> | 0.0              |
| <i>ccct/t/t</i> | 0.0              |

---

**Table 5** A demonstration of how time-variation in parameters can result in an estimated zero observation error. In all cases there is time variation of reporting probability. The first five models have one other instance of time variation and for the remaining models there is one further instance of time variation.

| parameter       | No replication |        | With replication |        |
|-----------------|----------------|--------|------------------|--------|
|                 | estimate       | se     | estimate         | se     |
| $\beta_0$       | -0.1880        | 0.0478 | -0.1854          | 0.0480 |
| $\beta$         | -0.0232        | 0.0048 | -0.0188          | 0.0043 |
| $\gamma_0$      | 0.3851         | 0.0731 | 0.3878           | 0.0736 |
| $\gamma$        | -0.0184        | 0.0061 | -0.0177          | 0.0059 |
| $\delta_0$      | 0.8888         | 0.1036 | 0.8907           | 0.1055 |
| $\delta$        | -0.0182        | 0.0088 | -0.0182          | 0.0089 |
| $\zeta_0$       | 1.3604         | 0.0877 | 1.3566           | 0.0894 |
| $\zeta$         | -0.0114        | 0.0051 | -0.0120          | 0.0050 |
| $\nu_0$         | -2.0275        | 0.0256 | -2.0275          | 0.0256 |
| $\nu$           | -0.8321        | 0.0461 | -0.8339          | 0.0462 |
| $\log p$        | -0.0850        | 0.0781 | -0.0906          | 0.0789 |
| $\log \sigma^2$ | 12.2583        | 0.1884 | 12.2221          | 0.1353 |

**Table 6** A comparison between the analysis of the Grey heron data with and without the use of pseudo replication, when there is no heterogeneity in the model, and no boundary estimate of observation variance. The labelling of the parameters is the same as in Table 2. In the pseudo replication case we present averages of 20 separate simulations.

| parameter       | No Replication |        | Replication with 3 values |        | Replication with 5 values |        |
|-----------------|----------------|--------|---------------------------|--------|---------------------------|--------|
|                 | estimate       | se     | estimate                  | se     | estimate                  | se     |
| $\beta_0$       | -0.1711        | 0.0486 | -0.1742                   | 0.0486 | -0.1738                   | 0.0488 |
| $\beta$         | -0.0213        | 0.0038 | -0.0228                   | 0.0043 | -0.0217                   | 0.0047 |
| $\gamma_0$      | 0.4106         | 0.0743 | 0.4099                    | 0.0745 | 0.4094                    | 0.0751 |
| $\gamma$        | -0.0216        | 0.0058 | -0.0177                   | 0.0060 | -0.0163                   | 0.0061 |
| $\delta_0$      | 0.9306         | 0.1054 | 0.9243                    | 0.1061 | 0.9240                    | 0.1070 |
| $\delta$        | -0.0233        | 0.0087 | -0.0194                   | 0.0088 | -0.0172                   | 0.0088 |
| $\zeta_0$       | 1.3251         | 0.1002 | 1.3413                    | 0.1004 | 1.3420                    | 0.1018 |
| $\zeta$         | -0.0171        | 0.0049 | -0.0126                   | 0.0051 | -0.0085                   | 0.0054 |
| $\nu_0$         | -2.8954        | 0.1722 | -3.1561                   | 0.2155 | -3.1760                   | 0.2588 |
| $\nu$           | -0.8312        | 0.0462 | -0.8330                   | 0.0463 | -0.8350                   | 0.0464 |
| $\log \theta$   | -2.0194        | 0.0261 | -2.0206                   | 0.0260 | -2.0206                   | 0.0260 |
| $\log p$        | -0.0942        | 0.0917 | -0.1104                   | 0.0925 | -0.1178                   | 0.0951 |
| $\log \sigma^2$ | -13.3313       | –      | 10.7247                   | 0.1540 | 11.4949                   | 0.1487 |

**Table 7** A comparison between the analysis of the Grey heron data using model  $vvvv/v/c$  with no replication and with a single replicate based on 3 or 5 neighbouring values when there is heterogeneity in the model. In both replicate cases we present averages of 20 separate simulations. Variation between simulations was small. Standard errors for the case of no replication, when there is a boundary estimate, are obtained for the other parameters using an appropriate singular-value decomposition approach; (Searle 1982, p. 318). This is the case illustrated in panel (d) of Figure 2. The resulting estimated standard errors can be expected to be conservative.



(a): using all the MRR data

| method                     | RMSE $\times$ 1000 |         |            |          |          |          |
|----------------------------|--------------------|---------|------------|----------|----------|----------|
|                            | $\beta_0$          | $\beta$ | $\gamma_0$ | $\gamma$ | $\log p$ | $\sigma$ |
| MRR only                   | 492                | 50      | 802        | 73       | -        | -        |
| IPM                        | 492                | 43      | 798        | 62       | 787      | 40520    |
| IPM with pseudorep         | 495                | 42      | 798        | 62       | 811      | 43876    |
| Plug $\tilde{\sigma}^2$    | 492                | 43      | 799        | 62       | 791      | -        |
| Plug $\tilde{\sigma}^2/2$  | 493                | 43      | 798        | 62       | 787      | -        |
| Plug $3\tilde{\sigma}^2/2$ | 492                | 44      | 800        | 63       | 794      | -        |

(b): using 10 % of the MRR data

| method                     | RMSE $\times$ 1000 |         |            |          |          |          |
|----------------------------|--------------------|---------|------------|----------|----------|----------|
|                            | $\beta_0$          | $\beta$ | $\gamma_0$ | $\gamma$ | $\log p$ | $\sigma$ |
| MRR only                   | 4539               | 162     | 5772       | 259      | -        | -        |
| IPM                        | 2727               | 98      | 3831       | 197      | 5201     | 43901    |
| IPM with pseudorep         | 2429               | 95      | 3431       | 213      | 4675     | 42752    |
| Plug $\tilde{\sigma}^2$    | 2774               | 98      | 3890       | 197      | 5341     | -        |
| Plug $\tilde{\sigma}^2/2$  | 2462               | 103     | 3496       | 202      | 4718     | -        |
| Plug $3\tilde{\sigma}^2/2$ | 3042               | 100     | 4189       | 201      | 5835     | -        |

**Table 8** Root mean square error results (RMSEs), all multiplied by 1000, from fitting model  $vvvv/v/c$  with  $\sigma = e^{12.258/2} = 458.97$  and no overdispersion, so that  $\theta = 0$ . There were no boundary estimates for  $\sigma$  by the IPM method in the 200 simulations undertaken. We compare the results of just analysing the recovery data alone (MRR only), of using integrated population modelling (IPM), of using integrated population modelling with a single pseudo replication (IPM with pseudorep), and of plugging in various multiples of the estimated value of  $\sigma$  obtained from fitting a spline,  $\tilde{\sigma}$ , for selected parameters. This is done twice, (a) for when we use the full MRR data and (b) for when we use a random 10% of the MRR data. Note that (a) and (b) result from different simulation runs.

| (a): using all the MRR data     |                    |         |            |          |          |          |          |
|---------------------------------|--------------------|---------|------------|----------|----------|----------|----------|
| method                          | RMSE $\times$ 1000 |         |            |          |          |          |          |
|                                 | $\beta_0$          | $\beta$ | $\gamma_0$ | $\gamma$ | $\theta$ | $\log p$ | $\sigma$ |
| MRR only                        | 698                | 54      | 842        | 75       | 858      | -        | -        |
| IPM                             | 605                | 44      | 820        | 69       | 220      | 1060     | 148412   |
| IPM with pseudorep              | 600                | 48      | 813        | 68       | 269      | 967      | 242942   |
| Plug $\tilde{\sigma}^2$         | 614                | 50      | 842        | 69       | 335      | 967      | -        |
| Plug $\tilde{\sigma}^2/2$       | 610                | 48      | 836        | 68       | 269      | 978      | -        |
| Plug $3\tilde{\sigma}^2/2$      | 616                | 51      | 845        | 70       | 362      | 962      | -        |
| Plug 148.41 <sup>2</sup>        | 606                | 45      | 825        | 68       | 173      | 1023     | -        |
| (b): using 10 % of the MRR data |                    |         |            |          |          |          |          |
| method                          | RMSE $\times$ 1000 |         |            |          |          |          |          |
|                                 | $\beta_0$          | $\beta$ | $\gamma_0$ | $\gamma$ | $\theta$ | $\log p$ | $\sigma$ |
| MRR only                        | 1690               | 161     | 3158       | 225      | 1572     | -        | -        |
| IPM                             | 1619               | 91      | 2767       | 197      | 1564     | 2748     | 148412   |
| IPM with pseudorep              | 1652               | 116     | 2867       | 192      | 1570     | 3793     | 201596   |
| Plug $\tilde{\sigma}^2$         | 1581               | 96      | 2742       | 198      | 1557     | 2835     | -        |
| Plug $\tilde{\sigma}^2/2$       | 1591               | 97      | 2760       | 195      | 1561     | 2915     | -        |
| Plug $3\tilde{\sigma}^2/2$      | 1567               | 98      | 2738       | 197      | 1556     | 2729     | -        |
| Plug 148.41                     | 1612               | 95      | 2756       | 193      | 1562     | 2811     | -        |

**Table 9** Root mean square error results (RMSEs), all multiplied by 1000, from fitting model  $vvv/v/c$  with  $\sigma = e^{10/2} = 148.41$  and overdispersion, with  $\theta = 0.07$ . Separately for (a) and (b), which arise from different simulation runs, results are only for the 38 cases out of 200 simulations undertaken when there were boundary estimates for IPM. We compare the results of just analysing the recovery data alone (MRR only), of using integrated population modelling (IPM), of using integrated population modelling with a single pseudo replication (IPM with pseudorep), and of plugging in various multiples of the estimated value of  $\sigma$  obtained from fitting a spline,  $\tilde{\sigma}$ , and also for the true value,  $\sigma = 148.41$ , for selected parameters. This is done twice, (a) for when we use the full MRR data and (b) for when we use a random 10% of the MRR data.