



# Analysing Mark–Recapture–Recovery Data in the Presence of Missing Covariate Data Via Multiple Imputation

Hannah WORTHINGTON, Ruth KING, and Stephen T. BUCKLAND

We consider mark–recapture–recovery data with additional individual time-varying continuous covariate data. For such data it is common to specify the model parameters, and in particular the survival probabilities, as a function of these covariates to incorporate individual heterogeneity. However, an issue arises in relation to missing covariate values, for (at least) the times when an individual is not observed, leading to an analytically intractable likelihood. We propose a two-step multiple imputation approach to obtain estimates of the demographic parameters. Firstly, a model is fitted to only the observed covariate values. Conditional on the fitted covariate model, multiple “complete” datasets are generated (i.e. all missing covariate values are imputed). Secondly, for each complete dataset, a closed form complete data likelihood can be maximised to obtain estimates of the model parameters which are subsequently combined to obtain an overall estimate of the parameters. Associated standard errors and 95% confidence intervals are obtained using a non-parametric bootstrap. A simulation study is undertaken to assess the performance of the proposed two-step approach. We apply the method to data collected on a well-studied population of Soay sheep and compare the results with a Bayesian data augmentation approach.

Supplementary materials accompanying this paper appear on-line.

**Key Words:** Individual time-varying continuous covariates; Mark–recapture–recovery data; Missing values; Multiple imputation; Two-step algorithm.

## 1. INTRODUCTION

The collection and analysis of mark-recapture data have a long history. The traditional model fitted to such data is typically referred to as the Cormack–Jolly–Seber (CJS) model (Cormack 1964; Jolly 1965; Seber 1965; for a review see for example McCrea and Morgan 2014; King 2013, 2014 and references therein). Initially only live recaptures were modelled, but the term CJS model is now more generally used for mark–recapture–recovery (MRR)

---

Hannah Worthington (✉) (E-mail: [hw233@st-andrews.ac.uk](mailto:hw233@st-andrews.ac.uk)) is a PhD Student, Ruth King is a Reader and Stephen T. Buckland is a Professor are in School of Mathematics and Statistics and Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Fife KY16 9LZ, Scotland, UK.

© 2014 The Author(s)  
*Journal of Agricultural, Biological, and Environmental Statistics*  
DOI: [10.1007/s13253-014-0184-z](https://doi.org/10.1007/s13253-014-0184-z)

data (i.e. allowing additional dead recoveries). MRR data are typically displayed in the form of the encounter histories for each animal observed in the study, recording whether each individual is observed (alive or dead) or not at each capture event. The model parameters of interest are typically the survival probabilities of the individuals under study, with the recapture probabilities (for live resightings) and recovery probabilities (of dead recoveries) nuisance parameters. Heterogeneity can be modelled by the inclusion of covariate information. For discrete valued individual time-varying covariate(s), the Arnason–Schwarz (AS) model is often used where the transition probabilities between discrete states are typically assumed to be first-order Markovian (Brownie et al. 1993; Schwarz et al. 1993). We consider the more complex case where the individual time-varying covariates are continuous valued.

The main issue that arises for time-varying individual covariates is that of missing data. If an individual is unobserved at any point within the study, its corresponding covariate value is also unobserved. Even if an individual is observed, the corresponding individual covariate value may still not be recorded. For example, for the Soay sheep data that we will consider, observations of individuals following their initial capture are a combination of physical recaptures and resightings from a distance. If an individual is only resighted at a distance, their corresponding weight is unknown. Note that we use the term “recaptures” to include both physical recaptures and resightings. In the presence of unknown continuous covariate values, the corresponding likelihood is analytically intractable, and expressed as an integral over all the missing covariate values (see for example, Bonner and Schwarz 2006; King et al. 2009; Langrock and King 2013 for further discussion). Several approaches have been proposed for dealing with this issue. An early approach was simply to coarsely discretise the continuous covariates considered (Nichols et al. 1992), thus reducing the model to the AS model (the integrals are essentially replaced by summations). However, the discretisation performed is arbitrary and generally discards large amounts of information. Langrock and King (2013) refine this approach by considering a very fine discretisation for the set of possible values for the unobserved continuous covariates. The corresponding probability transition matrix for the covariate process is expressed as a function of an underlying covariate model specified on continuous space. Alternatively, a Bayesian data augmentation (or complete data likelihood) approach has been proposed, defining the joint posterior distribution over the model parameters and unobserved covariate values (Bonner and Schwarz 2006; King et al. 2008; Schofield et al. 2009). Finally, the so-called “trinomial” approach of Catchpole et al. (2008) considers a conditional likelihood approach, using only the observed covariate values, discarding a potentially significant proportion of data. Bonner et al. (2010) compare the performance of the trinomial, Bayesian data augmentation and a single (deterministic) imputation method. A comparison of a fine discretization method to the trinomial can be found in Langrock and King (2013).

We propose a new approach for dealing with the missing covariate data. In particular, we consider a two-step process: (1) fit a parametric model to the observed covariate data and impute the missing covariate values given the fitted covariate process model; (2) conditional on the covariate values (observed and imputed) maximise the likelihood of the observed encounter histories to obtain corresponding parameter estimates. Given the covariate values, the corresponding maximum likelihood estimates (MLEs) of the model parameters for the observed encounter histories can be obtained using standard software, such as MARK (White

and Burnham 1999) or the associated R interface, RMark (Laake 2013). Repeatedly imputing the missing covariate values leads to a multiple imputation approach, where the overall estimates of the parameters are obtained by combining the estimates of the different imputed datasets. Associated confidence intervals are obtained using a non-parametric bootstrap in order to incorporate the additional uncertainty relating to the covariate model parameters.

In Sect. 2 we introduce the data motivating the approach and the underlying model to be fitted. In Sect. 3 we describe how the model is fitted to MRR data using multiple imputation. We conduct a simulation study in Sect. 4 before considering a real dataset relating to Soay sheep in Sect. 5. Finally, in Sect. 6 we conclude with a discussion.

## 2. DATA AND MODEL

We briefly describe the data before introducing the notation and discussing the general form of the likelihood and model in terms of the covariate dependence on the survival probabilities.

### 2.1. DATA

We consider MRR data of Soay sheep on the island of Hirta in the St Kilda archipelago. This is a well-studied population (Clutton-Brock and Pemberton 2004), resulting in a rich source of data, including the collection of individual time-varying covariates in the form of the weight of individual sheep. Individuals are marked as lambs with uniquely numbered ear tags allowing for identification at subsequent capture occasions. We consider the data from 1990 to 2009, from the annual autumn census, and restrict the dataset to females captured and marked as lambs between 1990 and 2008, providing a total of 1,745 individuals. An individual history is most easily expressed in the form of the combination of an encounter history (whether an individual is observed at each capture event) and covariate history (the weight of the individual at each capture event). For example, consider the following history of an individual first observed as a lamb in 1991 and recovered dead in 2000:

1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	(Year)
1	1	0	0	1	0	1	0	0	2	(Encounter history)
12.6	?	?	?	20.3	?	?	?	?	–	(Covariate history)

The second line corresponds to the encounter history, where “0” denotes the individual is not observed, “1” the individual is observed, and “2” the individual is recovered dead. Thus this individual was initially captured and marked in 1991, recaptured in 1992, 1995 and 1997 before being recovered dead in 2000. The third line corresponds to the observed weight at each corresponding capture event, where ? denotes that the weight is unknown. In this example the weight of the individual is only observed in 1991 and 1995, and is unknown in all other years, including 1992 and 1997 when it was also recaptured. In total, for the dataset that we consider, only 54.1 % of live recaptures have an associated observed

covariate value. For further discussion of the study see for example, [Clutton-Brock and Pemberton \(2004\)](#) and [Bonner et al. \(2010\)](#) (and references therein).

## 2.2. NOTATION

Let there be a total of  $T$  capture events labelled  $t = 1, \dots, T$ . At each capture event all observed individuals are recorded. The first time that an individual is observed, it is uniquely marked and released back into the population, so that the individual can be uniquely identified at each subsequent recapture (we assume there is no mark loss). Let  $N$  denote the total number of individuals observed within the study period. For each individual  $i = 1, \dots, N$ , the associated encounter history is denoted by  $\mathbf{h}_i = \{h_{i,t} : t = 1, \dots, T\}$ , such that,

$$h_{i,t} = \begin{cases} 0 & \text{if individual } i \text{ is not observed at time } t; \\ 1 & \text{if individual } i \text{ is observed at time } t; \\ 2 & \text{if individual } i \text{ is recovered dead at time } t. \end{cases}$$

We let  $\mathbf{h} = \{\mathbf{h}_i : i = 1, \dots, N\}$  denote the set of observed encounter histories in the study.

The corresponding likelihood of the data is a function of survival, recapture and recovery probabilities. In particular, we set,

$$\phi_{i,t} - \mathbb{P}(\text{individual } i \text{ survives until time } t + 1 \mid \text{individual is alive at time } t)$$

$$p_{i,t} - \mathbb{P}(\text{individual } i \text{ is observed at time } t \mid \text{individual is alive at time } t)$$

$$\lambda_{i,t} - \mathbb{P}(\text{individual } i \text{ is recovered dead at time } t \mid \text{individual dies in the interval } [t-1, t)).$$

We let  $\boldsymbol{\phi} = \{\phi_{i,t} : i = 1, \dots, N; t = 1, \dots, T-1\}$  and similarly for  $\mathbf{p}$  and  $\boldsymbol{\lambda}$  (for times  $t = 2, \dots, T$ ). The set of model parameters is  $\boldsymbol{\theta} = \{\boldsymbol{\phi}, \mathbf{p}, \boldsymbol{\lambda}\}$ . The likelihood of the observed encounter histories is given by,

$$f(\mathbf{h}|\boldsymbol{\theta}) = \prod_{i=1}^N H(i), \tag{1}$$

where  $H(i)$  denotes the probability of the encounter history for individual  $i$ . For simplicity, let  $\chi_{i,t}$  denote the probability that individual  $i$  is not observed after time  $t$ , with

$$\chi_{i,t} = 1 - (1 - \phi_{i,t})\lambda_{i,t+1} - \phi_{i,t}[1 - (1 - p_{i,t+1})\chi_{i,t+1}]$$

such that  $\chi_{i,T} = 1$ , for all  $i = 1, \dots, N$ . This term considers all possible outcomes after a final live encounter (death and not recovered or survival but not seen again).

We let  $f_i$  denote the first time that individual  $i$  is observed and  $l_i$  the corresponding final time that the individual is observed (alive or dead). The observed population can be divided into two groups: individuals seen only once ( $f_i = l_i$ ) and those seen (alive or dead) more than once ( $f_i < l_i$ ). For individuals seen more than once, then they must be alive from time  $f_i$  to (at least) time  $l_i - 1$  if they are recovered dead at time  $l_i$  or to time  $l_i$  if they are observed alive at time  $l_i$ . The encounter history probability between these occasions is a product of Bernoulli trials. For the final encounter there are two options to consider: (i) a

live recapture after which survival of the animal is unknown (so that all subsequent possible outcomes must be considered); and (ii) a dead recovery. This leads to the probability of an encounter history of the form:

$$H(i) = \begin{cases} \chi_{i, f_i} & f_i = l_i \\ \prod_{t=f_i}^{l_i-2} \phi_{i,t} \{I(h_{i,t+1} = 1)p_{i,t+1} + I(h_{i,t+1} = 0)(1 - p_{i,t+1})\} \\ \quad \times \{I(h_{i,l_i} = 1)\phi_{i,l_i-1}p_{i,l_i}\chi_{i,l_i} + I(h_{i,l_i} = 2)(1 - \phi_{i,l_i-1})\lambda_{i,l_i}\} & f_i < l_i, \end{cases}$$

where  $I(\cdot)$  denotes the indicator function; and we take the convention that the product term is unity if  $l_i = f_i + 1$ . For further discussion on this likelihood see for example [Catchpole et al. \(1998\)](#), [King and Brooks \(2002\)](#), [King et al. \(2009\)](#) and [McCrea and Morgan \(2014\)](#).

### 2.3. MODEL

We assume no individual heterogeneity for the recapture and recovery probabilities (removing the subscript  $i$  notation). To incorporate additional individual heterogeneity in the survival probabilities,  $\phi$ , we specify the survival probabilities as a function of individual time-varying continuous covariates ([Catchpole et al. 2000, 2004](#); [Bonner and Schwarz 2006](#); [King et al. 2006, 2008, 2009](#)). We consider a single covariate, for simplicity, but the method can be immediately extended to multiple covariates (see Sect. 6). Let  $w_{i,t}$  denote the value of the individual covariate for individual  $i = 1, \dots, N$  at time  $t = f_i, \dots, T$ . The survival probability for individual  $i$  at time  $t$  is typically specified in logistic form:

$$\text{logit } \phi_{i,t} = \alpha + \beta w_{i,t},$$

where  $\alpha$  and  $\beta$  are the regression parameters (see for example [North and Morgan 1979](#); [King et al. 2009](#)). The set of model parameters is  $\theta = \{\mathbf{p}, \boldsymbol{\lambda}, \alpha, \beta\}$ , since the survival probabilities are a deterministic function of  $\alpha$  and  $\beta$ . However, a complexity arises due to the presence of unknown covariate values (if an individual is not observed at a given capture event, the corresponding covariate value is also unobserved, and hence unknown).

Let  $x_{i,t}$  denote the recorded covariate value for individual  $i = 1, \dots, N$  at time  $t = f_i, \dots, T$ , such that  $x_{i,t} = NA$  if the covariate value is unknown (including following death) and set  $\mathbf{x}_i = \{x_{i,t} : t = f_i, \dots, T\}$ . Let  $\mathbf{x} = \{x_{i,t} : i = 1, \dots, N; t = f_i, \dots, T\}$ . We let  $w_{i,t}$  denote the true covariate value for individual  $i$  at time  $t = f_i, \dots, T$  and set  $\mathbf{w} = \{w_{i,t} : i = 1, \dots, N, t = f_i, \dots, T\}$  (so that  $w_{i,t} = x_{i,t}$  when  $x_{i,t} \neq NA$ ). Finally, we let  $\mathbf{z}$  denote the set of unobserved covariate values, so that  $\mathbf{z} = \{w_{i,t} : x_{i,t} = NA \text{ for } i = 1, \dots, N; t = f_i, \dots, T\}$ .

We assume an underlying model for the covariate values over time, with model parameters  $\boldsymbol{\eta}$ . The likelihood for the observed encounter histories can be expressed in the form,

$$f(\mathbf{h}, \mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\eta}) = \int f(\mathbf{h}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta}, \boldsymbol{\eta})d\mathbf{z} = \int f(\mathbf{h}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{z})f(\mathbf{y}, \mathbf{z}|\boldsymbol{\eta})d\mathbf{z},$$

where  $f(\mathbf{y}, \mathbf{z}|\boldsymbol{\eta})$  denotes the joint probability density function of the covariate values for the underlying covariate model and  $f(\mathbf{h}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{z})$  the likelihood of the observed encounter

histories, conditional on the demographic parameters and all covariate values (both observed and unobserved), as given in Eq. (1). However, this integral is analytically intractable.

### 3. MODEL FITTING

Previous classical approaches for obtaining parameter estimates in the presence of individual covariates have included removing individuals for which there are any missing covariate values (Catchpole et al. 2000, for only time-invariant individual covariates); the trinomial approach using a conditional likelihood (Catchpole et al. 2008); and discretisation methods, either coarsely (Nichols et al. 1992), which discards significant information, or finely (Langrock and King 2013). For further discussion of these (and additional approaches) see for example Catchpole et al. (2000) and Langrock and King (2013). Alternatively, a Bayesian data augmentation approach has been proposed (Bonner and Schwarz 2006; King et al. 2006, 2008) forming the joint posterior distribution over the model parameters and unobserved covariate values. A Markov chain Monte Carlo algorithm is used to explore the posterior distribution, updating the missing covariate values at each iteration of the chain, and integrated out by considering the posterior marginal distribution of the parameters of interest. See Bonner et al. (2010) and Langrock and King (2013) for further discussion and comparisons between some of these approaches.

Within this paper, we consider a two-step multiple imputation process for fitting the MRR models. This involves initially fitting a model to the observed covariate data only, implicitly assuming that the unobserved values are *missing at random* (Heitjan and Basu 1996). Given the fitted covariate model, we obtain a set of complete datasets by repeatedly imputing the missing covariate values from the predictive distribution of the covariate values. For each imputed covariate dataset we maximise the likelihood of the observed encounter histories, conditional on the observed and imputed covariate values, to obtain a set of estimates for the demographic parameters. The estimates obtained for each imputed dataset are combined to obtain an overall multiple imputation estimate of the model parameters. For a complete review of imputation techniques see for example Little and Rubin (2002).

We note that the fitted covariate model does not use any direct information from the encounter histories, for example, in terms of whether the individual is known to be alive or not at given times (although we only record covariate values when an individual is observed). Thus, there is no “feedback” loop from the information contained within the encounter histories into the covariate model. However, unless an individual is recovered dead or observed at the last capture event, there is further uncertainty as to the status of the individual following final sighting. Repeated non-sightings of individuals and corresponding unobserved covariate values following their final sighting may arise because the animal had died and was not recovered. Thus, it is clear that the covariate values are not *missing at random*, so that the implicit independence assumption of the two-step multiple imputation process is violated. However, the amount of information within the encounter histories regarding the covariate values is likely to be relatively small (compared to that for the observed covariate values). This is because the associated likelihood component of the encounter histories contains information on the covariate values only indirectly through the survival probabilities,

with additional uncertainty resulting from recapture and recovery probabilities that are less than one. We assess the impact of making this missing-at-random assumption within the imputation step of the missing covariate values in Sect. 4 via a simulation study.

### 3.1. DEMOGRAPHIC PARAMETER ESTIMATION

We initially fit an underlying model to the observed covariate values (ignoring the observed encounter histories). Given this fitted model, we propose a multiple imputation approach (imputing the unobserved covariate values) to obtain estimates of the demographic parameters. To illustrate the associated modelling issues we consider two possible covariate models.

#### 3.1.1. Covariate Process

We consider the covariate data *independently* of the encounter histories,  $\mathbf{h}$ . The covariate process describes the changing covariate values for each individual over time. We assume an underlying normal model for the covariate process of the form,

$$w_{i,t} | \mu_{i,t}, \sigma^2 \sim N(\mu_{i,t}, \sigma^2),$$

for  $i = 1, \dots, N$  and  $t = f_i, \dots, T$ , independently of each other. Clearly, knowledge of the underlying biological system should be used in the specification of the covariate model. We consider in further detail two possible models for the underlying mean  $\mu_{i,t}$ : (1) a simple additive model with fixed individual, temporal and age factors; and (2) a random walk model corresponding to a first-order Markovian structure for the covariate values. In both cases we provide the predictive distributions for the unobserved covariate values  $z$ , given the fitted covariate model, that will be used in the following imputation step.

##### *Model 1*

We consider a simple additive model for the mean of the form  $\mu_{i,t} = \delta_i + \kappa_t + \gamma_j$ , for  $i = 1, \dots, N$  and  $t = f_i, \dots, T$ . The covariate model parameters are  $\boldsymbol{\eta} = \{\boldsymbol{\delta}, \boldsymbol{\kappa}, \boldsymbol{\gamma}, \sigma^2\}$  with the individual component  $\boldsymbol{\delta} = \{\delta_i : i = 1, \dots, N\}$ , temporal component  $\boldsymbol{\kappa} = \{\kappa_t : t = 1, \dots, T\}$  and age component  $\boldsymbol{\gamma} = \{\gamma_j : j = 2, \dots, J\}$  where  $j$  is the age of individual  $i$  at time  $t$  and  $J$  the maximum observed age. The corresponding likelihood of the observed covariate data is simply a product over independent normal distributions, which can be easily maximised to obtain the MLEs,  $\widehat{\boldsymbol{\eta}}$ . We note that this assumes that at least one recorded covariate value is obtained for each individual within the study (else there is no associated MLE for  $\delta_i$ ).

The corresponding predictive distribution of an unknown covariate value for individual  $i$  at time  $t$ , denoted  $z_{i,t}$ , for the fitted covariate model is simply of the form,

$$z_{i,t} | \widehat{\boldsymbol{\eta}} \sim N(\widehat{\delta}_i + \widehat{\kappa}_t + \widehat{\gamma}_j, \widehat{\sigma}^2).$$

Thus, each of the  $z_{i,t}$  values can be simulated (independently) from their predictive covariate distribution (given the underlying model and parameters  $\widehat{\boldsymbol{\eta}}$ ). However, we note that in

general, this model may be not biologically realistic, since we would generally expect some dependence of covariate values over time.

*Model 2*

We remove the previous independence assumption of the additive model and consider a first-order Markov model for the covariate process, with temporal and age components, to allow for temporal heterogeneity (for example, due to environmental conditions) and the effect of age (for example, growth in young sheep). The covariate process can be described as follows. For the initial covariate value (i.e. initial state), for each individual  $i = 1, \dots, N$ , we set

$$w_{i,f_i} \sim N(v_{f_i}, \tau^2), \quad (2)$$

and for  $t = f_i + 1, \dots, T$ , we set,

$$w_{i,t} \sim N(w_{i,t-1} + \kappa_t + \gamma_j, \sigma^2). \quad (3)$$

We note that the interpretation of  $\kappa_t$  and  $\gamma_j$  corresponds to the year and age effects on weight relative to the weight in the previous year. Notationally, we let  $\eta = \{\kappa, \boldsymbol{\gamma}, \mathbf{v}, \sigma^2, \tau^2\}$  where  $\kappa = \{\kappa_t : t = 2, \dots, T\}$ ,  $\boldsymbol{\gamma} = \{\gamma_j : j = 2, \dots, J\}$  and  $\mathbf{v} = \{v_t : t = 1, \dots, T - 1\}$ . The likelihood of the observed covariate values can be explicitly written down, using standard normal distributional results and the MLEs of the parameters obtained, once more denoted by  $\hat{\eta}$ . We note that for the “null” case, where no covariate value is recorded for an individual observed within the study, the underlying model parameters  $\eta$  are still estimable (due to the initial state distribution being specified).

Due to the Markovian structure, in order to derive the predictive distribution for the missing covariate values,  $z$ , given the underlying fitted model, we need to consider a set of possible cases. These correspond to (i) after the final recorded covariate value; (ii) between successively recorded covariate values; and (iii) before the first recorded covariate value for each individual. In order to consider these separate cases, we let  $b_i$  and  $c_i$  denote the first and final time the covariate value for individual  $i$  is observed, respectively. The three cases we consider correspond to (i)  $t > c_i$ ; (ii)  $b_i < t < c_i$  and (iii)  $t < b_i$ .

**Case (i)  $t > c_i$ :** The predictive distribution for the unknown covariate value for individual  $i$  at time  $t$ , denoted  $z_{i,t}$  (assuming the individual is alive) given the fitted underlying covariate model is

$$z_{i,t} | \hat{\eta}, w_{i,t-1} \sim N \left( w_{i,t-1} + \hat{\kappa}_t + \hat{\gamma}_j, \hat{\sigma}^2 \right). \quad (4)$$

Here we note that

$$w_{i,t-1} = \begin{cases} y_{i,t-1} & t = c_i + 1; \\ z_{i,t-1} & t = c_i + 2, \dots, T. \end{cases}$$

In other words,  $w_{i,t-1}$  corresponds to the last observed value when  $t = c_i + 1$  and the previously imputed value when  $t = c_i + 2, \dots, T$ .

**Case (ii)**  $b_i < t < c_i$ :

Let  $y_{i,t+k}$  denote the next observed covariate value for individual  $i$  after time  $t$  (for  $k \geq 1$ ). The predictive distribution for  $z_{i,t}$  given the fitted covariate model, next observed covariate value and covariate value at time  $t - 1$ , denoted  $w_{i,t-1}$ , has the form,

$$z_{i,t} | \widehat{\boldsymbol{\eta}}, w_{i,t-1}, y_{i,t+k} \sim N \left( \frac{k(w_{i,t-1} + \widehat{\kappa}_t + \widehat{\gamma}_j) + y_{i,t+k} - \sum_{g=1}^k (\widehat{\kappa}_{t+g} + \widehat{\gamma}_{j+g})}{k+1}, \frac{k\widehat{\sigma}^2}{k+1} \right), \quad (5)$$

where  $w_{i,t-1}$  is as above.

**Case (iii)**  $t < b_i$ : We initially consider the missing covariate value for individual  $i$  at time  $f_i$ , the first time the individual is observed. Using standard normal distributional theory, it is straightforward to show that, given the distribution specified on the initial covariate values,

$$z_{i,f_i} | \widehat{\boldsymbol{\eta}}, y_{i,b_i} \sim N \left( \frac{\widehat{\tau}^2 \left[ y_{i,b_i} - \sum_{g=1}^k (\widehat{\kappa}_{f_i+g} + \widehat{\gamma}_{1+g}) \right] + k\widehat{\sigma}^2 \widehat{v}_{f_i}}{k\widehat{\sigma}^2 + \widehat{\tau}^2}, \frac{k\widehat{\sigma}^2 \widehat{\tau}^2}{k\widehat{\sigma}^2 + \widehat{\tau}^2} \right),$$

where  $k = b_i - f_i$ . The predictive distribution for the remaining missing covariate values  $z_{i,t}$  for  $t = f_i + 1, \dots, b_i - 1$  follow analogously to case (ii), given the parameters  $\widehat{\boldsymbol{\eta}}$ , imputed covariate value  $w_{i,t-1}$  and observed covariate value  $y_{i,b_i}$ , and is of the same form as given in Eq. (5). Derivations of the predictive distributions given in cases (ii) and (iii) are provided in the Online Supplementary Material Appendix A.

Finally, we note that for the “null” case, where no covariate value is recorded for an individual observed within the study, the corresponding predictive distribution for the covariate values is simply the fitted covariate process model (i.e. the covariate model defined at the MLEs of the parameters,  $\widehat{\boldsymbol{\eta}}$ ). Thus, to simulate the corresponding missing covariate values, an initial covariate value is simulated from the fitted initial state distribution, and each subsequent covariate value simulated from the fitted covariate model conditional on the previous simulated value (i.e. case (i)).

Clearly other covariate models can be used (see for example [Langrock and King 2013](#)) and in general a model selection procedure can be performed within the analysis of the covariate data, using only the observed covariate values, via likelihood ratio tests or AIC statistic, for example. Extending the model to allow survival to depend upon multiple covariates is straightforward to achieve through the definition of an appropriate (multivariate normal) covariate model.

### 3.1.2. Multiple Imputation

Given the fitted underlying covariate process model, we consider a multiple imputation approach for obtaining parameter estimates of the demographic parameters of interest. In

particular, we repeatedly impute a “complete” set of covariate values a total of  $M$  times from the predictive distribution of the covariate values, given the fitted covariate process model (see Sect. 3.1.1 for particular examples). These imputed covariate values are denoted by  $z^1, \dots, z^M$ , where  $z^m = \{z_{i,t}^m : x_{i,t} = NA \text{ for } i = 1, \dots, N; t = f_i, \dots, T\}$  for  $m = 1, \dots, M$ . For the  $m$ th imputed dataset, the corresponding survival probabilities for individual  $i = 1, \dots, N$  at time  $t = f_i, \dots, T$  can be written in the form, logit  $\phi_{i,t} = \alpha + \beta w_{i,t}^m$ , where

$$w_{i,t}^m = \begin{cases} y_{i,t} & x_{i,t} \neq NA; \\ z_{i,t}^m & x_{i,t} = NA. \end{cases}$$

Thus, the survival probabilities for the likelihood of the observed encounter histories are all explicitly known, since they are no longer a function of any unknown covariate values (the associated missing covariate values have been imputed). For each imputed dataset,  $m = 1, \dots, M$ , the likelihood of the encounter histories, conditional on the set of complete covariate values (i.e. observed and imputed),  $f(\mathbf{h}|\boldsymbol{\theta}, \mathbf{y}, \mathbf{z}^m)$ , can be efficiently maximised using the standard computer package MARK/RMark to obtain an estimate of the demographic parameters, denoted by  $\hat{\boldsymbol{\theta}}^m$ . The overall estimate of the demographic parameters is:

$$\hat{\boldsymbol{\theta}} = \frac{1}{M} \sum_{m=1}^M \hat{\boldsymbol{\theta}}^m.$$

In other words, we take the sample mean of the estimated demographic parameters to be the multiple imputation estimate of the demographic parameters. Typically only a small number of complete datasets need to be imputed to obtain reliable estimates (Little and Rubin 2002). Unsurprisingly, the lower the proportion of missing covariate values, the fewer imputations are necessary (see Sect. 4 for further discussion). For the Soay sheep dataset considered in Sect. 5, we note that the parameter estimates obtained appeared to stabilise around 10 imputations. To obtain an associated standard error and/or 95 % confidence interval (CI), we wish to take into account the parameter uncertainty with regard to the underlying covariate model. Thus, to incorporate this additional uncertainty we consider a non-parametric bootstrap.

### 3.1.3. Standard Errors/Confidence Intervals

The standard errors (and hence CIs) of the demographic parameters can be estimated from the results of the multiple imputation conducted on the covariate values. However, these multiple imputations are conditional on the fitted covariate model (and MLEs of the covariate model parameters). Thus, to fully incorporate the uncertainty of the covariate model within the CIs of the demographic parameters we consider a non-parametric bootstrap. We use a non-parametric bootstrap, sampling individual encounter histories each time, as this is more robust than the associated parametric bootstrap since it does not assume that the underlying model is correct. In particular, we sample with replacement a total of  $N$  individuals from the set of  $N$  observed individuals. Each bootstrap replicate is drawn such

that the number of individuals captured for the first time each year remains the same as for the original dataset. We let  $\mathbf{h}_b$  denote the set of  $N$  sampled encounter histories with associated observed covariate values  $\mathbf{y}_b$ . The subscript  $b$  is used to denote that this is a bootstrap replicate. For each bootstrap replicate, the previous two-step multiple imputation algorithm is then implemented. Namely, for a given bootstrap replicate,  $b$ , this involves fitting the underlying covariate model to the corresponding observed covariate values,  $\mathbf{y}_b$ , to obtain the corresponding MLEs, denoted by  $\hat{\boldsymbol{\eta}}_b$ . A multiple imputation approach is then used, simulating a set of missing covariate values from their predictive covariate distributions, denoted by  $z_b^1, \dots, z_b^M$ , for the given bootstrapped capture histories. Secondly, for each imputed dataset, the complete data likelihood of the bootstrapped encounter histories is maximised, conditional on the observed and imputed covariate values and combined to obtain a bootstrap estimate  $\hat{\boldsymbol{\theta}}_b$ . These bootstrapped multiple imputation estimates can be used to obtain standard errors for the parameters (by taking the sample standard deviation) or 95% CI (by taking the lower and upper 2.5% quantiles of the bootstrapped estimates). Using this approach we can identify a “hierarchy of imputations”, with a single bootstrap corresponding to imputing the set of capture histories (to allow for uncertainty with regard to the fitted covariate model) followed by the imputation of the missing covariate values (conditional on the model parameters). See, for example, [Buckland \(1984\)](#) and [Buckland and Garthwaite \(1991\)](#) for further discussion of the bootstrap algorithm, with particular reference to ecological applications.

#### 4. SIMULATION STUDY

We assess the proposed multiple imputation approach using a simulation study. In particular we consider the impact of assuming the data are *missing at random*, when ignoring the information contained in the encounter histories on the missing covariate values. We consider a study with  $T = 10$  capture events, where 100 individuals are first observed at each capture event  $t = 1, \dots, 9$ , so that there are  $N = 900$  individuals in total. We assume time-invariant capture and recovery probabilities and consider four scenarios: (a)  $p = 0.9$ ,  $\lambda = 0.9$ ; (b)  $p = 0.9$ ,  $\lambda = 0.3$ ; (c)  $p = 0.3$ ,  $\lambda = 0.9$ ; and (d)  $p = 0.3$ ,  $\lambda = 0.3$ . For the covariate model, we use the first-order Markov model described above including temporal and age effects and simulate  $v_t \sim N(-1.2, 0.4)$  for  $t = 1, \dots, 9$ ,  $\tau = 0.4$ ,  $\kappa_t \sim N(0, 0.15)$  for  $t = 2, \dots, 10$ ,  $\boldsymbol{\gamma} = \{0.75, 0.65, 0.3, 0.1, 0, -0.05, -0.1, -0.2, -0.2\}$  and  $\sigma = 0.3$ . For the logistic regression on survival, we divide the animals into three age groups: lambs, yearlings and adults, corresponding to the first, second and remaining years of life respectively. We use subscripts on the regression parameters to denote the different age groups, where the subscript denotes the year of life. We set the intercept terms to be  $\alpha_1 = 2$ ,  $\alpha_2 = 2.1$  and  $\alpha_{3+} = 1.7$ , and gradient terms  $\beta_1 = 1.9$ ,  $\beta_2 = 1.7$  and  $\beta_{3+} = 1.1$  (the resulting survival functions are given in Fig. 1). In addition, we consider 2 scenarios for the probability that a covariate value is observed ( $p_w$ ), given the individual is observed:  $p_w = 1$  (when an individual is observed at any given time, the corresponding covariate value is also observed); and  $p_w = 0.6$  (when an individual is observed, with probability 0.4 the associated covariate value is unobserved/missing). Note that for the second scenario,  $p_w = 0.6$  the same data is

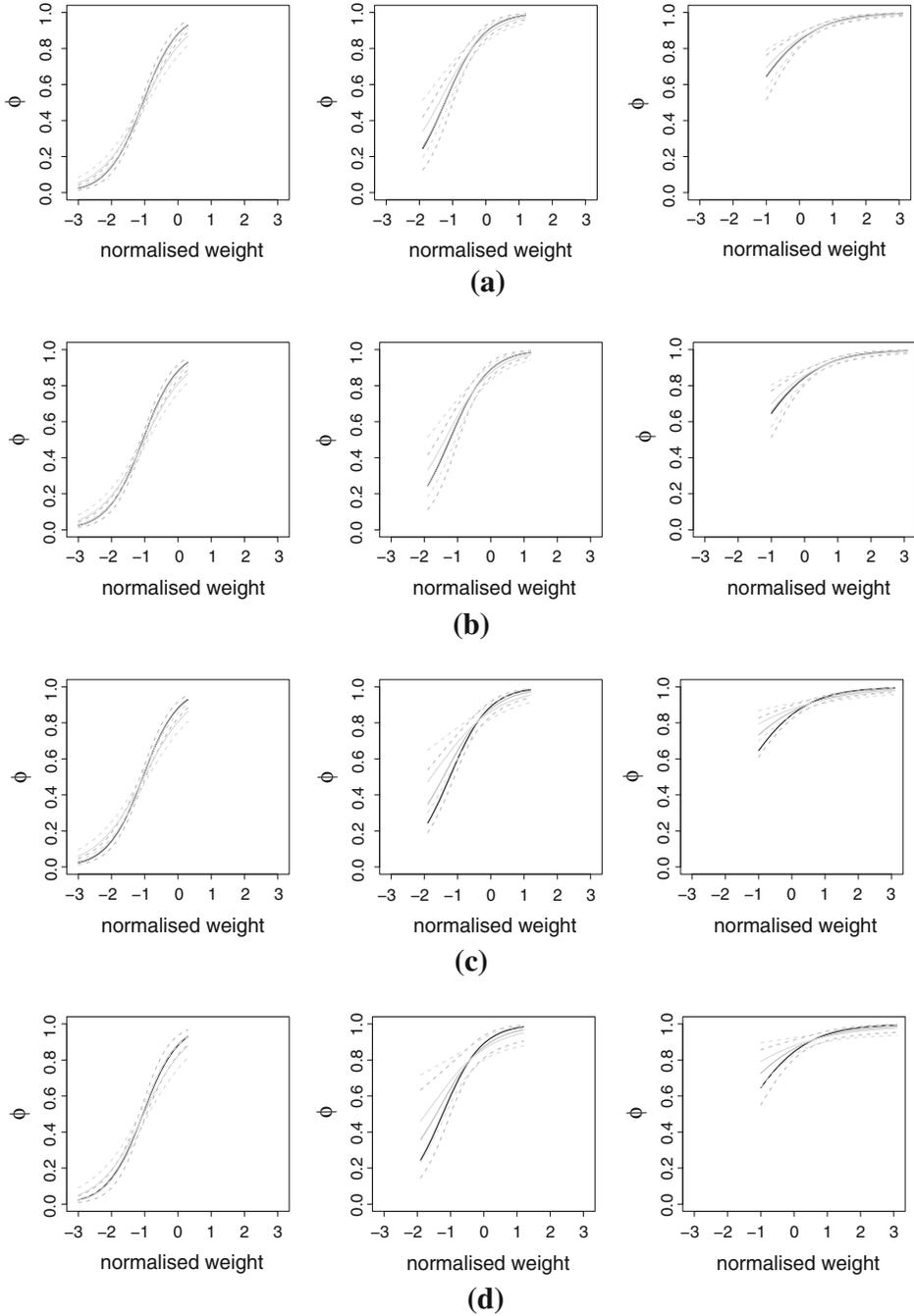


Figure 1. True (in black) and mean estimated survival function, with associated 95% CI for the three age groups (lambdas (left), yearlings (centre) and adults (right)) for  $p_w = 1$  (in mid-grey) and  $p_w = 0.6$  (in light-grey) for each scenario **a**  $p = 0.9, \lambda = 0.9$ ; **b**  $p = 0.9, \lambda = 0.3$ ; **c**  $p = 0.3, \lambda = 0.9$  and; **d**  $p = 0.3, \lambda = 0.3$ .

used as for  $p_w = 1$  but with the observed covariate values “thinned”. In total 100 datasets are simulated for each of the 8 scenarios corresponding to the combination of recapture, recovery and covariate observation probabilities.

For the simulation study we initially consider the number of imputations needed in order for the corresponding parameter estimates to stabilise (plots are given in Online Supplementary Material Appendix B). From the simulations conducted, it is clear that fewer imputations are needed for the parameter estimates to converge when the recapture probability is large and/or  $p_w = 1$ . This is to be expected as these scenarios lead to fewer missing covariate values. An increased recovery probability, however, appears to have limited impact on the number of imputations needed. For the remainder of the simulation study we use 20 imputations to calculate the parameter estimates for each simulated dataset, with 10 imputations used for each bootstrapped dataset. The mean estimated survival function with associated 95% CI for each age group for each combination of recapture and recovery probabilities are presented in Fig. 1. Boxplots of the capture and recovery probabilities (nuisance parameters) for each scenario are presented in Online Supplementary Material Appendix C.

From Fig. 1, the most accurate results are achieved when the covariate is observed with certainty when observed (i.e.  $p_w = 1$ ). This is to be expected as this maximises the amount of covariate information available. This suggests that, in this case, there is little additional information contained in the encounter history relating to the (unobserved) covariate values, in comparison to the observed covariate values, and that the missing-at-random assumption is a reasonable approximation. However, the survival estimates for lambs (and possibly for yearlings) tend to be a weaker function of weight when additional covariate values are removed for observed individuals (i.e. when  $p_w = 0.6$ ). Further investigation suggests that this may be related to a larger proportion of missing covariate values for the younger age group(s), and in particular individuals in the study who do not have any observed covariate values. In particular, it is well known that completely omitting individuals with no observed covariate values from MRR analyses leads to positively biased survival estimates (Kendall et al. 2003; King 2014), since individuals that die early are less likely to have observed covariate values. This same mechanism appears to lead to the poorer estimates for the youngest individuals (which will have the highest proportion of individuals with no observed covariates within their encounter history). This is supported by additional simulations (results omitted) where the covariate value is always recorded at the initial capture, but at each subsequent recapture, the corresponding covariate value is only reported with probability 0.5. For these simulated datasets, the survival estimates are similar to the case  $p_w = 1$ . Finally, as expected higher recapture/recovery probabilities leads to increased precision of the estimates. Overall, these results suggest that the two-step approach performs well when covariate values are always observed (or observed at least once for each individual), but that caution is required when a significant proportion of individuals are observed with no recorded covariate values.

## 5. SOAY SHEEP ANALYSIS

We consider data from the annual autumn census between 1990 and 2009 ( $T = 20$  with  $t = 1$  referring to 1990). 1,745 individuals are marked as lambs and released during this period, with a total of 5,306 capture/recaptures and 1,203 recoveries. In addition, there are 2,869 observed covariate values (45.9% of captures do not have an associated weight;

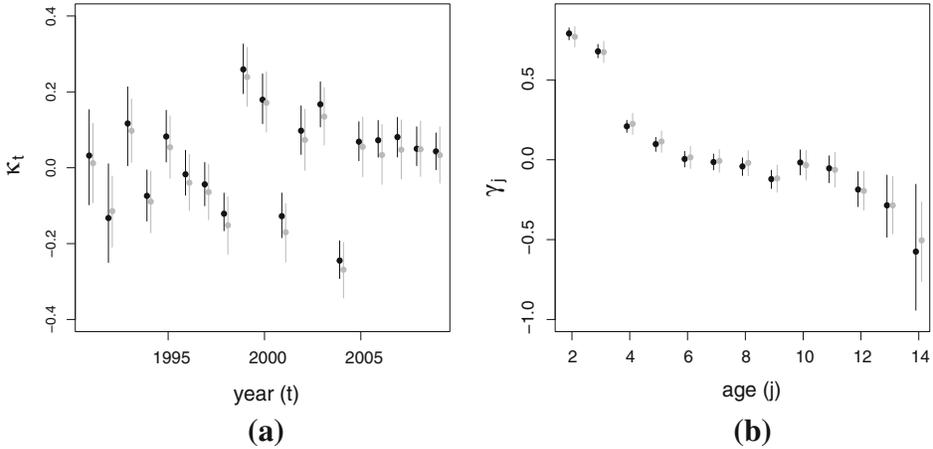


Figure 2. MLEs and 95% (bootstrap) confidence intervals (in *black*) and posterior mean and 95% symmetric credible intervals (in *grey*) for: **a** the time effects  $\kappa_t$  for years  $t = 1991, \dots, 2009$  and; **b** the age effects  $\gamma_j$  for ages  $j = 2, \dots, 14$  for the Soay sheep data. Note that since the oldest observed sheep (and corresponding covariate value) is age 14 we set  $\gamma_j = 0$  for all  $j \geq 15$ .

34% of individuals do not have any observed weights). For comparison we also conduct a Bayesian data augmentation approach, assuming vague priors on the parameters (see Online Supplementary Material Appendix D for details).

### 5.1. MODEL

We assume individual homogeneity with regard to the recapture and recovery probabilities but assume that these parameters are time-varying. For the survival probabilities, we include an age-dependence for the logistic regression coefficients (following the age-dependent model considered by [Bonner et al. 2010](#)) with four distinct age-classes corresponding to lambs (year 1), yearlings (year 2), adults (years 3–7) and seniors (years 8+). Individuals typically increase weight in early years of life so that we also allow the covariate model to be dependent on the age of the individual and permit additional annual variation. Thus we consider a first-order Markov model for the weight of an individual with additive temporal and age dependent terms i.e. model 2 in Sect. 3.1.1 and described by Eqs. (2) and (3).

The parameter estimates and posterior means with associated 95% confidence/credible intervals for the age effects and the time effects are provided in Fig. 2. Clearly, the time effects vary with respect to year, most likely as a function of (unmodelled) environmental conditions. The age effects are unsurprising with lambs and immature sheep increasing in weight before stabilising around age 5–6 as mature adults with a further possible decrease in weight from age 12–13 as they enter “old-age” (with a significant increase in uncertainty for these latter age effects most likely due to the very limited number of observations). In general, a model selection procedure can be performed to select a suitable model. For these data we obtain  $\Delta\text{AIC}$  values of 1978 and 253 for the only time effect model (excluding age effects) and only age effect model (excluding time effects) against the time and age effect

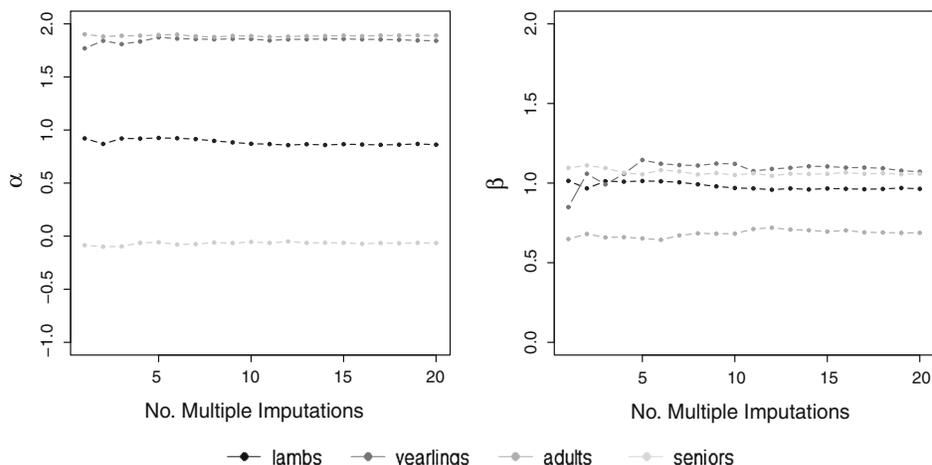


Figure 3. MLEs of the survival regression parameters for each age group plotted against the number of imputed datasets used for the Soay sheep dataset.

model, respectively, providing further evidence that both age and time effects are important. We note that further models could be considered, such as age *groups*, a parametric/semi-parametric functional form for the age effects, or inclusion of environmental covariates for the time effects (though this is beyond the current focus of this analysis).

## 5.2. DEMOGRAPHIC PARAMETERS

Conditional on the imputed covariate values, we fit the model in the computer package RMark. We initially investigate the number of imputed datasets necessary in order for the corresponding MLEs of the parameters to stabilise. Figure 3 provides the MLEs of the survival parameters (i.e. the parameters of primary interest) against increasing numbers of imputed datasets. The estimates stabilise with very few imputed datasets, so for conservative reasons we use 20 imputed datasets for obtaining the MLEs and 10 imputed datasets within each bootstrap iteration to obtain associated standard errors.

### 5.2.1. Recapture/Recovery Probabilities

The corresponding parameter estimates for the recapture and recovery probabilities are provided in Fig. 4. The recovery probabilities are more variable than the recapture probabilities with generally appreciably larger confidence intervals. The high recapture probabilities (generally  $>0.9$ ) explains the need for relatively few multiple imputations to obtain stabilised parameter estimates.

### 5.2.2. Survival Probabilities

The estimates of the survival probabilities for the different age classes are provided in Fig. 5. Clearly, there is a positive dependence between the survival probability and weight for all ages, as expected. In particular, the survival probabilities are very strongly related to

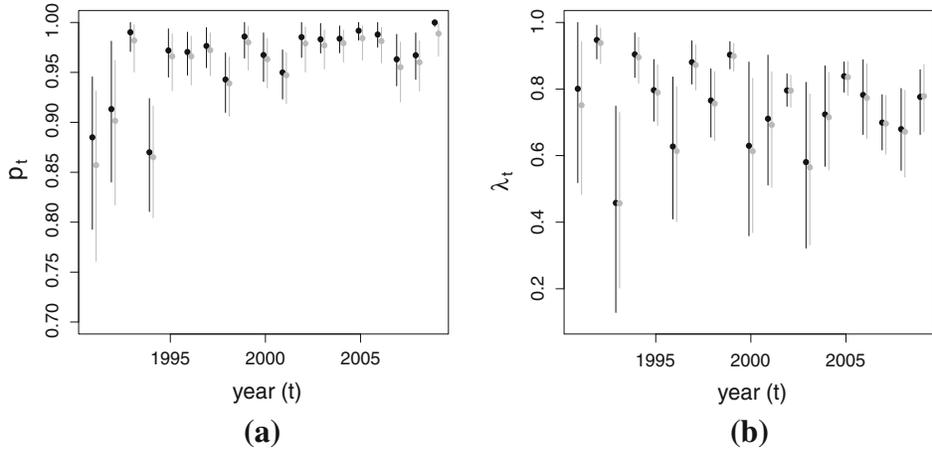


Figure 4. The parameter estimates and 95% (bootstrap) confidence interval for the multiple imputation approach (in *black*) and posterior means and 95% symmetric credible interval for the Bayesian data augmentation approach (in *grey*) for **a** recapture probabilities  $p_t$  and; **b** recovery probabilities  $\lambda_t$  for years  $t = 1991, \dots, 2009$  for the Soay sheep data. Note that the scale of the y axis differs between plots.

weight for lambs and yearlings, with the strength of the dependence decreasing in adulthood (years 3–7) and the level of uncertainty in the relationship between weight and survival increasing in the oldest age class (years 8+). As suggested by the simulation study in Sect. 4, in the case where not all individuals have at least one observed covariate value, the survival probabilities for the multiple imputation approach may lead to an underestimate in the gradient of the survival probabilities (decreasing as individuals age)—this can be seen by comparison with the Bayesian analysis). A positive dependence between survival and weight for the different age classes has been identified in previous analyses using a variety of techniques and slightly different datasets (e.g. varying years of the study) and models (see for example King et al. 2008; Bonner et al. 2010; Langrock and King 2013). However, we note that Bonner et al. 2010 and Langrock and King 2013 only consider individuals with at least one observed weight, which as mentioned in Sect. 4 leads to a positive bias in the survival estimates, most noticeably for lambs.

## 6. DISCUSSION

We present a new likelihood-based approach for analysing complex MRR data in the presence of individual time-varying continuous covariates. The approach builds on previous work by combining the idea of a complete data likelihood with a multiple imputation approach (with the compromise being the removal of the feedback loop from the encounter histories on the covariate process). The approach is straightforward to implement, using a two-step process: initially fit a model to the covariate data; given the model fitted generate a series of complete covariate datasets and combine these with the observed encounter histories to obtain the MLEs of the model parameters. The proposed multiple imputation approach is computationally fast (in the order of seconds) to obtain the MLEs of the model

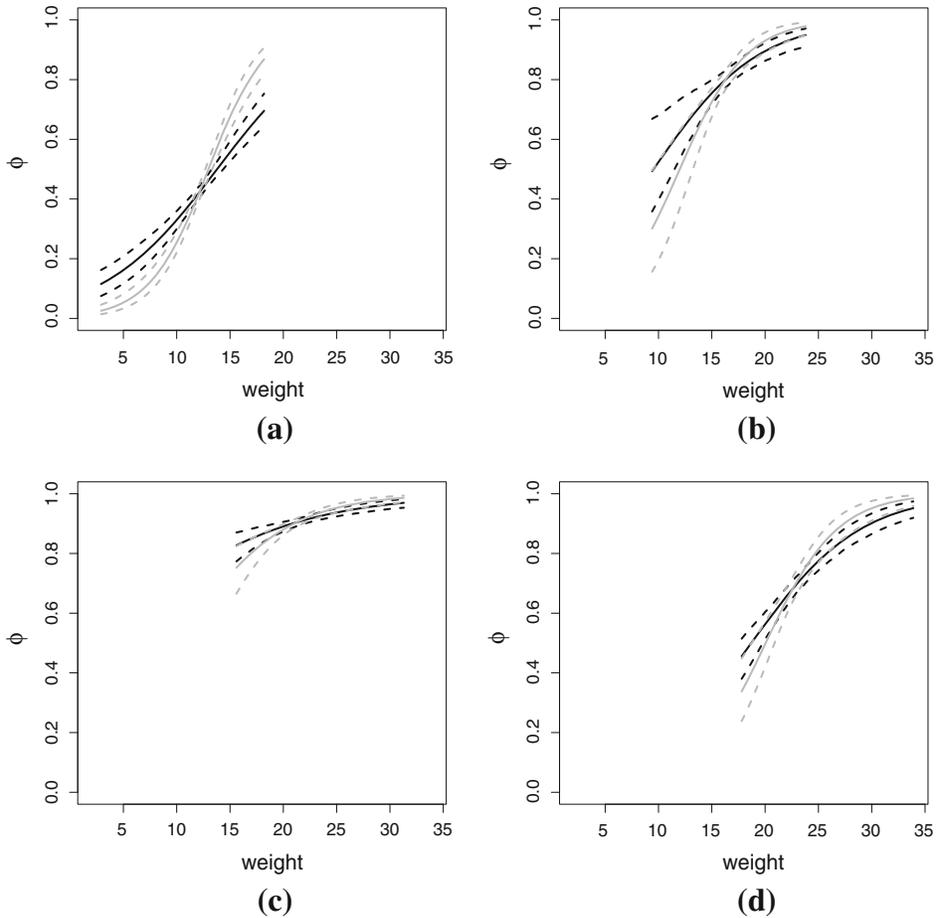


Figure 5. The parameter estimates for the multiple imputation approach with 95 % (bootstrap) confidence interval (in *black*) and posterior means and 95% symmetric credible interval (in *grey*) for the survival probability as a function of weight (across the observed range of weights) for **a** lambs; **b** yearlings; **c** years 3–7 and; **d** senior sheep in years 8+.

parameters conditional on the observed and imputed covariate values, using standard software, such as RMark. This is in contrast to the alternative Bayesian data augmentation and numerical integration using an approximate hidden Markov model. The trinomial approach is computationally fast and can also be fitted within the computer package MARK (Bonner 2013), but discards a significant amount of information.

The trade-off of this two-step multiple imputation approach is that it assumes that the covariate values are missing at random with the missing covariate values imputed independently of the capture encounter histories. The simulation study suggests that if the covariate values are always observed when a capture occurs (or at least for the initial capture occasion) then the two-step approach produces reliable estimates of the survival probabilities. However, introducing individuals with no observed covariate values leads to a decrease in performance.

The extension to multiple covariates is immediate for the multiple imputation approach, with limited additional computational expense. In particular, this primarily impacts on step

1 of the algorithm—fitting the observed covariate values to a specified covariate model. The covariate model may be specified independently for each covariate or jointly over covariates if they may interact with each other. The estimation of the MLEs of the demographic parameters, conditional on the covariate values, follows as before, simply with additional regression parameters to be estimated. Alternatively, for a numerical integration approach (such as that applied by [Langrock and King 2013](#)) multiple covariates lead to significant computational expense which quickly becomes infeasible, for example, beyond 2 or 3 covariates due to the “curse of dimensionality”. The trinomial approach can be applied as before, where the conditioning on covariate values now refers to all covariate values of a single individual at a given capture event (if one or more covariate values are missing at a given time point, the corresponding likelihood contribution is discarded).

The two-step multiple imputation algorithm to impute missing values, given some underlying model, can be applied more generally to obtain parameter estimates and associated confidence intervals, for example, in the future prediction of systems when the model parameters are expressed as a function of environmental (time-varying) covariates, so that future covariate values are unknown. See for example [Little and Rubin \(2002\)](#) for a summary of many multiple imputation methods. The application and performance of the bootstrap algorithm to other missing data problems is the focus of current research, particularly in relation to the analysis of capture–recapture data.

## ACKNOWLEDGEMENTS

We would like to thank all the volunteers and members of the Soay sheep project who collected the data from the sheep at St. Kilda, and allowing us access to the data, and the National Trust for Scotland and the Scottish Natural Heritage who have supported this project. In addition, we would like to thank Tim Coulson for providing the data in the necessary form and for discussions with regard to the data and Bill Browne, Matthew Schofield, Ruth Dundas and Roland Langrock for useful discussions relating to this work. Financial support was provided by a Carnegie Scholarship awarded to Hannah Worthington by the Carnegie Trust for the Universities of Scotland and EPSRC/NERC Grant EP/1000917/1.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

[Received September 2013. Accepted September 2014.]

## REFERENCES

- Bonner, S. J. (2013), Implementing the trinomial mark–recapture–recovery model in program MARK. *Methods in Ecology and Evolution* **4**, 95–98
- Bonner, S. J., Morgan, B. J. T. and King, R. (2010), Continuous covariates in mark–recapture–recovery analysis: A comparison of methods. *Biometrics* **66**, 1256–1265
- Bonner, S. J. and Schwarz, C. J. (2006), An extension of the Cormack–Jolly–Seber model for continuous covariates with application to *Microtus Pennsylvanicus*. *Biometrics* **62**, 142–149
- Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H. and Hestbeck, J. B. (1993), Capture–recapture studies for multiple strata including non-Markovian transitions. *Biometrics* **49**, 1173–1187

- Buckland, S. T. (1984), Monte Carlo confidence intervals. *Biometrics* **40**, 811–817
- Buckland, S. T. and Garthwaite, P. H. (1991), Quantifying precision of mark–recapture estimates using the bootstrap and related methods. *Biometrics* **47**, 255–268
- Catchpole, E. A., Fan, Y., Morgan, B. J. T., Clutton-Brock, T. H. and Coulson, T. (2004), Sexual dimorphism, survival and dispersal in red deer. *Journal of Agricultural, Biological, and Environmental Statistics* **9**, 1–26
- Catchpole, E. A., Freeman, S. N., Morgan, B. J. T. and Harris, M. P. (1998), Integrated recovery/recapture data analysis. *Biometrics* **54**, 33–46
- Catchpole, E. A., Morgan, B. J. T., Coulson, T. N., Freeman, S. N. and Albon, S. D. (2000), Factors influencing Soay sheep survival. *Journal of the Royal Statistical Society, Series C* **49**, 453–472
- Catchpole, E. A., Morgan, B. J. T. and Tavecchia, G. (2008), A new method for analysing discrete life history data with missing covariate values. *Journal of the Royal Statistical Society, Series B* **70**, 445–460
- Clutton-Brock, T. H. and Pemberton, J. M. (Eds) (2004). *Soay Sheep: Dynamics and Selection in an Island Population*. Cambridge University Press, Cambridge.
- Cormack, R. M. (1964), Estimates of survival from the sighting of marked animals. *Biometrika* **51**, 429–438
- Heitjan, D. F. and Basu, S. (1996), Distinguishing “Missing at Random” and “Missing Completely at Random”. *The American Statistician* **50**, 207–213
- Jolly, G. M. (1965), Explicit estimates from capture–recapture data with both death and immigration stochastic model. *Biometrika* **52**, 225–247
- Kendall, W. L., Hines, J. E. and Nichols J. D. (2003) Adjusting multistate capture–recapture models for misclassification bias: manatee breeding proportions. *Ecology* **84**, 1058–1066
- King, R. (2013), A review of Bayesian state-space modelling of capture–recapture–recovery data. *Interface Focus* **2**, 190–204
- King, R. (2014), Statistical ecology. *Annual Review of Statistics and its Application* **1**, 401–426
- King, R. and Brooks, S. P. (2002), Model selection for integrated recovery/recapture data. *Biometrics* **58**, 841–851
- King, R., Brooks, S. P. and Coulson, T. (2008), Analysing complex capture–recapture data in the presence of individual and temporal covariates and model uncertainty. *Biometrics* **64**, 1187–1195
- King, R., Brooks, S. P., Morgan, B. J. T. and Coulson, T. (2006), Bayesian analysis of Soay sheep survival. *Biometrics* **62**, 211–220
- King, R., Morgan, B. J. T., Gimenez, O. and Brooks, S. P. (2009), *Bayesian Analysis for Population Ecology*. CRC Press, Boca Raton
- Laake, J. L. (2013), *RMark: An R Interface for Analysis of Capture–Recapture Data with MARK*. AFSC Processed Rep 2013–2001, 25p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115
- Langrock, R. and King, R. (2013), Hidden Markov-type maximum likelihood estimation of mark–recapture–recovery models in the presence of continuous covariates. *Annals of Applied Statistics* **7**, 1709–1732
- Little, R. J. A. and Rubin, D. B. (2002), *Statistical Analysis with Missing Data*, 2nd edition. Wiley, Hoboken
- McCrea, R. S. and Morgan, B. J. T. (2014), *Analysis of Capture–Recapture Data*. CRC Press, Boca Raton
- Nichols, J. D., Sauer, J. R., Pollock, K. H. and Hestbeck, J. B. (1992), Estimating transition probabilities for stage-based population projection matrices using capture–recapture data. *Ecology* **73**, 306–312
- North, P. M. and Morgan, B. J. T. (1979), Modeling heron survival using weather data. *Biometrics* **35**, 667–681
- Schwarz, C. G., Schweigert, J. F. and Arnason, A. N. (1993), Estimating migration rates using tag-recovery data. *Biometrics* **49**, 177–193
- Schofield M. R., Barker, R. J. and MacKenzie, D. I. (2009), Flexible hierarchical mark–recapture modeling for open populations using WinBUGS. *Environmental and Ecological Statistics* **16**, 253–268
- Seber, G. A. F. (1965), A note on the multiple-recapture census. *Biometrika* **52**, 249–259
- White, G. C. and Burnham, K. P. (1999), Program MARK: Survival estimation from populations of marked animals. *Bird Study* **46** Supplement, 120–138