Importance of accounting for phylogenetic dependence in multi-species mark-recapture studies

Fitsum Abadi[1,2*], Christophe Barbraud[3], Dominique Besson[3], Joël Bried[4], Pierre-André Crochet[1], Karine Delord[3], Jaume Forcada[5], Vladimir Grosbois[6], Richard A. Phillips[5], Paul Sagar[7], Paul Thompson[8], Susan Waugh[9], Henri Weimerskirch[3], Andrew G. Wood[5] , Olivier Gimenez[1]

[1]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, campus CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, France

[2] Animal Demography Unit, Department of Biological Sciences and Department of Statistical Sciences, University of Cape Town, Rondebosch 7701, South Africa

[3]Centre d'Etudes Biologiques de Chizé, UPR 1934 CNRS, 79360 Villiers en Bois, France

[4]Departamento de Oceanografia e Pescas, Centro do IMAR da Universidade dos Açores, 9901-862 Horta, Açores, Portugal

[5]British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

[6]Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), Département ES, UR22, TA C22/E, Campus international de Baillarguet, 34398 Montpellier Cedex 5, France

[7]National Institute of Water & Atmospheric Research Ltd, P O Box 8602, Christchurch, New Zealand.

[8]University of Aberdeen, Institute of Biological & Environmental Sciences, Lighthouse Field Station, Cromarty, IV11 8YL, UK.

[9]Museum of New Zealand Te Papa Tongarewa, PO Box 467, Wellington, New Zealand

[*]Corresponding author E-mail**:** fitaba@gmail.com

Abstract

Species in comparative demography studies often have a common phylogenetic or evolutionary ancestry and hence, they cannot fully be treated as independent samples in the statistical analysis. Although the serious implication of ignoring phylogeny has long been recognized, no attempt has been made so far to account for the lack of statistical independence due to phylogeny in multi-species mark-recapture comparative demography studies. In this paper, we propose a Bayesian hierarchical model that explicitly accounts for phylogenetic dependence among species, and to correct for imperfect detection, which is a common phenomenon in free-ranging species. We illustrate the method using individual mark-recapture data collected from 16 seabird species of the order Procellariiformes. Data on body mass and phylogeny of these species are compiled from literature. We investigate the relationship between adult survival and body mass with and without accounting for phylogeny. If we ignore phylogeny, we obtain a positive survival-body mass relationship. However, this relationship is no longer statistically significant once phylogenetic dependence is taken into account, implying that survival may actually depend on an unmeasured variable that is correlated with body mass due to a shared dependence on phylogeny. The proposed model allows the integration of multi-species mark-recapture data and phylogenetic information, and it is therefore a valuable tool in ecological and evolutionary biology.

Key-words: Bayesian hierarchical model; comparative demography; mark-recapture; phylogeny; Procellariiformes; survival probability

1. Introduction

Reliable estimates of demographic parameters are crucial to understand population dynamics of wildlife populations (Lebreton et al., 1992; Williams et al., 2002). Over the past few decades, important methodological advancements have improved the estimation of demographic parameters from mark-recapture data whilst accounting for imperfect detection. Mark-recapture methods have successfully been used, for instance, to estimate age-specific survival probabilities (Lebreton et al., 1992), recruitment and dispersal (Pradel, 1996; Lebreton et al., 2003; Péron et al., 2010), and population abundance (Pollock et al., 1990; Cubaynes et al., 2010). In addition, these methods are widely used for exploring how demographic parameters are affected by environmental covariates (e.g., Lebreton et al., 1992; Altwegg et al., 2003; Gimenez et al., 2006). However, these models have all focused on studying the demography of a single population or species.

Traditionally, multi-species demographic studies have been performed in two steps: 1) demographic parameters of each species are estimated independently, and 2) comparisons of the estimates of each species are conducted (Johnston et al., 1997; Peach et al., 2001). However, this *ad hoc* approach is inefficient as it ignores the uncertainty associated with the estimates and also fails to take account of inter-species variation (Papadatou et al., 2012). More recently, hierarchical (random effects) mark-recapture models have been proposed to account for inter-species variation and to overcome the constraints of this *ad hoc* approach (Lahoz-Monfort et al., 2011; Papadatou et al., 2012; Péron and Koons, 2012).

Despite these advances, multi-species comparative demography studies have ignored that some species have a common phylogenetic or evolutionary ancestry (Harvey and Pagel, 1991; Freckleton et al., 2002; Paradis and Claude, 2002; Bried et al., 2003; Freckleton, 2009). In the strict sense, this phylogenetic dependence implies that species cannot be treated as independent units, hampering the application of conventional statistical procedures (e.g., generalized linear models) in a general non-mark-recapture context (Freckleton et al., 2002; Paradis and Claude, 2002). Furthermore, other studies have highlighted that ignoring phylogenetic correlation among species can lead to an overestimation of the precision of parameter estimates and hence flawed inferences (Harvey and Pagel, 1991; Halsey et al., 2006; Ives and Zhu, 2006; Freckleton, 2009). Some papers have suggested using alternative methods such as generalized least squares (GLS) and generalized estimating equations (GEE) to account for correlations among species due to phylogeny (Garland and Ives, 2000; Paradis and Claude, 2002; Ives and Zhu, 2006). These methods are applied to study the relationship between dispersal and population synchrony (Paradis and Claude, 2002), to investigate the relationship between adult life expectancy and body mass (Bried et al., 2003), and to determine the relationship between home-range area and body mass with simulated data (Ives and Zhu, 2006). Despite the serious implications of ignoring phylogeny, there appear to be no published works that account for phylogeny in the context of multi-species mark-recapture studies.

The main goal of this paper is therefore to develop a Bayesian hierarchical model for comparative demography analysis, which explicitly integrates phylogeny in the estimation of survival probability whilst accounting for imperfect detection. Here we focus on the estimation of survival but, in principle, our method is valid for other demographic parameters. In addition, our method is quite flexible, and can assess the impact of environmental covariates or allometric relationships on demographic parameters. We illustrate the application of our model using individual mark-recapture data for 16 seabird species in the order Procellariiformes (i.e., albatrosses, petrels and shearwaters). Bried et al. (2003) studied the relationship between adult life expectancy (i.e., derived from the survival estimates) and body mass of Procellariiformes accounting for phylogeny using the GLS method. However, they ignored the uncertainty in the adult life expectancy and treated it as known quantities in the GLS regression analysis. Such a two-step approach of ‘doing statistics on statistics’ has been repeatedly criticized in the literature (e.g., Link, 1999; Grosbois et al., 2008). Consequently, we used our new statistical approach that utilizes the mark-recapture data and phylogenetic information simultaneously to examine the relationship between survival and body mass of the order Procellariiformes. We compared the parameter estimates obtained from the models with and without phylogeny. We provide the R and WinBUGS codes for implementing the models in the appendix.

2. Multi-species mark-recapture model accounting for phylogeny

In this section, we first build the likelihood for mark-recapture data based on the standard Cormack-Jolly-Seber (CJS) model (Lebreton et al., 1992). Next, we introduce a hierarchical model that takes into account phylogeny in the estimation of survival probability. We then briefly discuss how this model can be implemented within a Bayesian framework.

2.1.Likelihood for mark-recapture data

The standard mark-recapture protocol yields capture histories with a sequence of ones (mark or recaptured/resighted) and zeroes (not captured or seen) for individuals in the study population. We assume that such individual mark-recapture histories collected from several related species are available. For each species, we considered the general Cormack-Jolly-Seber (CJS) model that provides estimates of time-varying survival (ϕ_t) and recapture (p_t) probabilities (Lebreton et al., 1992). For computational purpose, we used minimal sufficient statistics for the CJS model under the form of an m-array (Burnham et al., 1987; Lebreton et al., 1992). We denote the m-array entries by $m_{i,j}, i = 1, 2, ..., T-1, j = i+1, 2, ..., T+1$ and T is the number of sampling occasions, which may vary from species to species. The m-array, $m_{i,j}$

, represent the number of individuals released at occasion i and recaptured for the first time at occasion j for $j \leq T$, and $m_{i,T+1}$ denotes the number of individuals never recaptured following release at occasion i. The CJS model likelihood is then constructed for each species based on a product of multinomial distributions assuming each row of the m-array is independent and for which the cell probabilities are functions of both survival and recapture probabilities (Lebreton et al., 1992).

2.2.Incorporating phylogeny

To account for phylogenetic dependence among species, we proposed the following hierarchical model with a logit link function for the survival probability.

$$\operatorname{logit}(\phi_{t,s}) = \mu + \eta_s + \varepsilon_{t,s} \qquad (1)$$

$$\eta_s \sim MVN(\mathbf{0}, \delta^2 \Sigma), \quad \varepsilon_{t,s} \sim N(0, \sigma_s^2)$$

where $\phi_{t,s}$ is the survival probability of species s between years $t-1$ and t, μ is the overall mean survival probability, $\varepsilon_{t,s}$ is a normally distributed random term with species-specific temporal variance (σ_s^2), and η_s is a random term that depends on the species and is distributed as a multivariate normal with variance-covariance matrix $\delta^2 \Sigma$. Here Σ is derived from the phylogenetic tree and treated as fixed known quantity in our model (Ives and Zhu, 2006; Revell, 2010; Blomberg et al., 2012; Hansen and Bartoszek, 2012). By scaling Σ to a height of one, we can interpret δ^2 as the residual variance (de Villemereuill et al., 2012). We used hypothetical data to illustrate the computation of Σ (Fig. 1). The off-diagonal values in Σ (i.e., covariance) always increase as the phylogenetic distance decreases (Freckleton et al., 2002). That is, the greater the shared history between the species, the higher the values in Σ. As shown in Fig. 1, the main diagonal (i.e., the variance) in Σ is computed as the distance from the root to the tip (e.g., S_{11}=6 (4+2), S_{55}=6(2+3+1)) and the off-diagonal elements (i.e., the covariance) are the total shared path lengths between each pair of species (e.g., S_{12}=4, S_{35}= 5= (2+3), S_{15}=0 (share no path lengths)). By setting Σ to an identity matrix (i.e. ones on the diagonal and zeroes elsewhere), the effect is reduced to a species random effect that assumes no phylogenetic dependence among species (e.g., Papadatou et al., 2012). Note that Pagel's λ can be incorporated into Σ to measure the strength of phylogenetic signal (e.g., Revell, 2010; de Villemereuill et al., 2012).

Our model can also be easily extended to assess the relationship between survival and covariates (e.g., body mass). This is of particular interest when one is investigating allometric relationships (e.g., Bried et al., 2003). Thus, we can modify Eq. (1) as

$$\text{logit}(\phi_{t,s}) = \mu_s + \eta_s + \varepsilon_{t,s} \qquad (2)$$

$$\mu_s = \beta_0 + \beta_1 X_s$$

where μ_s is the mean survival of species s, X_s is the covariate related to species s, and β_0 and β_1 are the regression coefficients. This model assumes that only a portion of the inter-species variation in survival is explained by the covariate and the error terms (η_s) are correlated as described in Eq. (1).

2.3.Model implementation

Fitting hierarchical models is fairly straightforward in a Bayesian framework (Link and Barker, 2004; Millar, 2009). We therefore adopt a Bayesian inference using the Markov Chain Monte Carlo (MCMC) algorithm to obtain marginal posterior estimates of the parameters in the model. Bayesian inference involves specification of priors and we used vague prior distributions (see below in the application section) that reflect our little knowledge about the parameters. We assessed convergence of the MCMC output using the standard Brooks-Gelman-Rubin statistic (Brooks and Gelman, 1998) and inspecting the diagnostic plots (e.g., trace, autocorrelation, and posterior density plots). We used the software WinBUGS (Spiegelhalter et al., 2003) calling it from program R with the package R2WinBUGS (Sturtz et al., 2005) to implement the models with and without phylogeny. The R and WinBUGS code used for fitting the models are provided in Appendix C.

3. Application to Procellariiformes data

We used long-term mark-recapture data of 16 seabird species from the order Procellariiformes (see Table A1, Appendix A) to illustrate the method. The duration of the mark-recapture studies were not the same for all species and varied between 9 and 41 years. We considered adult individuals only, hence focusing on adult survival, which is the parameter to which the population growth rate of long-lived species is most sensitive. We used published information on body mass and phylogenetic data for these species (Bried et al., 2003; Monteiro et al., 1996), and more details are available in Table A1, Appendix A and in Bried et al. (2003). Procellariiformes are long-lived birds and show the highest inter-specific body mass variation in any avian taxonomic order (from 19 g in the Least Storm-petrel *Halocyptena microsoma* to almost 10 kg in the Wandering Albatross *Diomedea exulans*) and share similar traits through common ancestry (Figs. A1 and A2, Appendix A). It is generally believed that survival depends on body mass with larger species having higher survival rates (Warham, 1990; Covas et al., 2002; Bried et al., 2003). In this study, we used our method to examine the allometry of adult survival (i.e. variation in adult survival with body mass) within the order Procellariiformes and, specifically, to compare the effect of (1) ignoring the phylogenetic information, and (2) incorporating this information in the model.

3.1.Goodness-of-fit (GOF) and model selection

Before fitting a multi-species hierarchical model that accounts for phylogeny, we performed preliminary analyses for each species. We first assessed the fit of the CJS model for each species independently (Pradel et al., 2005) using program U-CARE (Choquet et al., 2009). The overall GOF test of the CJS model for all species except *Pterodroma lessonii* was significant (P-value < 0.005). Further, either or both components Test 2 and Test 3 were significant at the 5% level for the 15 species, suggesting a lack of fit of the CJS model due to transience (an excess of individuals that are marked and never seen again) and/or trap-dependence (lack of independence between successive capture events). The results of the goodness-of-fit tests are provided in Appendix B. We used an *ad hoc* approach and discarded up to four capture occasions, depending on the species, to effectively remove the transience effect. We accounted for the trap dependence effect detected for the 14 species using two distinct recapture probabilities depending on whether a capture occurred or not the occasion before (Pradel, 1993). Once these effects were accounted for, no further lack of fit was detected.

We then conducted model selection considering different structures (e.g., time dependent, constant) for survival and recapture probabilities and the best model structure for each species was selected using the AIC criterion (Burnham and Anderson, 2002). For most of the species, the top-ranked model (i.e. the model with lowest AIC value) contained constant or time dependent survival and trap dependence effect in recapture probabilities. The summary of model selection results are provided in Appendix B. Preliminary analyses for each species were performed using program MARK (White and Burnham, 1999) calling it from R (R Development Core Team, 2012) with the RMark package (Laake, 2012). Next, we used each species-specific model structure selected previously and we treated time as random if the best model for a species was the one with time-dependent survival. We then performed the full analysis of multi-species hierarchical models using the software WinBUGS calling it from program R with the package R2WinBUGS (See Appendix C for code).

3.2.Multi-species models for the Procellariiformes

We fitted two different models for long-term data of the Procellariiformes. The first model assumes that only part of the inter-species variation is explained by body mass and there is a phylogenetic dependence among the species. This model is similar to Eq. 2 above where X is replaced by body mass. To derive Σ from the phylogenetic tree, we used the built-in function vcv.phylo from the ape package (Paradis et al., 2004) in program R and the output is displayed in Fig. A2, Appendix A. Our second model assumes that no phylogenetic dependence among the species and thus Σ is replaced by the identity matrix. We used normal prior distributions with mean 0 and variance 100 for β_0 and β_1, uniform prior distributions between 0 and 5 for δ and σ_s (see Eqs. (1) and (2)).

We ran two parallel chains of length 1,000,000 with a burn-in of 500,000 and thinned by taking every 100th value. The standard Brooks-Gelman-Rubin statistic (Brooks and Gelman, 1998) values were below 1.02 for all parameters and visual inspection of the diagnostic plots (e.g., trace, autocorrelation, and posterior density plots) showed no evidence of lack of convergence of the MCMC outputs. Note that we used the log transformed body mass to improve convergence. We then computed the posterior summary statistics based on the MCMC simulated values to make inference about the regression parameters with and without phylogeny. We used the deviance information criterion (DIC; Spiegelhalter et al., 2002) as a preliminary tool for comparing the two models. We noted that the use of DIC for comparing hierarchical models is controversial (e.g., Millar, 2009).

4. Results

Table 1 shows the regression coefficients estimates obtained from the model with and without phylogeny. The two models led to different conclusions about the relationship between survival and body mass. The model that accounted for phylogenetic information provided a wider 95% posterior confidence interval for the slope and overlapped with zero, suggesting no statistically significant biological signal (i.e., survival-body mass relationship), whereas there was a significant positive effect of body mass on survival when phylogeny was ignored (Table 1; Fig. 2). The model with phylogeny (DIC=29467) had also a lower DIC value than the one without phylogeny (DIC=29484), and hence it was better supported by the data. Fig. 3 shows the posterior mean survival probabilities with 95% credible intervals of the 16 species, which were derived from models where survival is a function of body mass. Models that ignored and accounted for phylogeny both yielded high estimates of survival probability, which are in agreement with the estimates reported by Bried et al. (2003). However, accounting for phylogeny provided less precise survival estimates compared to the model that ignored it (Fig. 3). For those species we considered time-dependent survival, the standard deviations (i.e., on the logit scale) of the time random effect were ranging from 0.30 to 1.20. The estimated residual standard deviations from the model with and without phylogeny were 1.29 (sd = 0.32) and 0.61(sd = 0.14), respectively.

5. Discussion

We present a hierarchical model for comparative demography analysis of free-ranging species using individual mark-recapture data. Our model is a generalization of the multilevel model of Papadatou et al. (2012) as it explicitly accounts for the dependence among species due to phylogeny. Using this framework the strength of dependence among species (i.e., derived from common ancestry) is translated into the variance-covariance matrix and incorporated as fixed known quantity in the model.

In the case study using the Procellariiformes data, the model that accounted for phylogeny yielded less precise (i.e., larger standard deviation) estimate of the body mass effect on survival in comparison to the model without phylogeny, which in turn led to slightly less precise estimates of the mean adult survival probabilities for each species. Interestingly, these two models have exactly the same number of parameters as the variance-covariance matrix (Σ) is known from the phylogenetic tree. However, the model without phylogeny assumes independence among species (i.e., more data points), resulting in an overly optimistic precision of the estimates of the regression coefficients and indicated a statistically significant relationship between survival and body mass. In our data set, the assumption of independence is clearly invalid as larger species tended to be closely related with each other and phylogenetically distinct from the smaller species (see Appendix A). In other words, some of the observations in the data are related to each other due to phylogeny and ignoring this correlation structure may, potentially, lead to invalid inferences.

Our findings are in line with previous studies that used the Generalized Least Squares (GLS) method to account for phylogeny. For instance, Bried et al. (2003) found no significant relationship between adult life expectancy and body mass of Procellariiformes. Ives and Zhu (2006) analyzed simulated data of home-range area and body mass and found no significant relationship after accounting for phylogeny. Therefore, our findings as well as previous studies clearly showed that ignoring dependence of species due to phylogeny may lead to spurious conclusions due to high Type I error rate. Now, we are not suggesting that there is really no relationship between body mass and survival in Procellariiformes. Firstly, there are mechanistic biological arguments suggesting that small species suffer from increased predation compared to large species (e.g., Mougeot and Bretagnolle, 2000; Keitt et al., 2004;

Oro et al., 2005; Bonnaud et al., 2009 for predation on small to medium sized Procellariiformes; studies documenting predation on adult of large-sized species proved much harder to find). Secondly, the survival-body mass relationships in birds is supported by several other studies (e.g., Gaillard et al., 1989; Saether, 1989). We have thus no doubt that Procellariiformes follow this general pattern and that larger species have higher adult survival. We want to stress, however, that this pattern in Procellariiformes is supported by essentially three "data points": the small-sized clade of low survival, the medium-sized clade of intermediate survival and the large-sized clade of high survival. Without support of external information, it would be impossible to conclude on the reality of the survival - body mass relationship in Procellariiformes. Our result could also imply that survival may actually depend on other behavioral, morphological and ecological species attributes that are correlated with body mass due to a shared dependence on phylogeny, and thus the direct effect of body mass was weak.

Although we showed how to directly incorporate phylogenetic information in comparative demography studies, we here highlighted the alternative views that accounting for phylogeny may not always be needed (Westoby et al., 1998). In the case of cross-species correlations where the interest is in the relationship between traits in the existing pool of species, treating each species as an independent sample (e.g., ignoring phylogeny) may be reasonable (Westoby et al., 1998). However, if the interest is in the evolution of these traits and then closely related species represent non-independent samples, then our method is valuable in properly accounting this lack of independence in the analysis.

In this article, we focused on adult survival, but our model can be applied to study the allometry of other parameters such as age dependence of survival and fecundity with appropriate link functions. The linearity assumption between survival (on the logit scale) and body mass can also be relaxed using semi-parametric or non-parametric models (Gimenez et al., 2006) whilst accounting for phylogeny. Here we paid little attention to the impact of different degrees of phylogenetic signal on biological inferences. This could be further evaluated using simulations. In this article, we also assumed that the phylogeny tree is known without error. However, some studies have shown that misspecification of the variance-covariance matrix $(\sum)$ can be problematic for comparative studies (Blomberg et al., 2012). Therefore, a logical development of this approach would be to evaluate the robustness of our analyses considering alternative phylogenies, and to extend the model to take account of phylogenetic uncertainty. In conclusion, our new statistical method integrates multi-species mark-recapture and phylogenetic data in a unified framework and hence, it is a valuable tool for comparative demography analysis in the field of ecology and evolutionary biology.

Acknowledgments

We thank two anonymous referees and Jasper Slingsby for helpful comments on the manuscript. We acknowledge the financial support of the Swiss National Science Foundation (PBBEP3_133501 to FA).We thank all people involved in the collection of the data.

References

Altwegg, R., Roulin, A., Kestenholz, M., Jenni, L., 2003. Variation and covariation in survival, dispersal, and population size in barn owls Tyto alba. Journal of Animal Ecology 72, 391–399.

Blomberg, S.P., Lefevre, J.G., Wells, J.A., Waterhouse, M., 2012. Independent Contrasts and PGLS Regression Estimators Are Equivalent. Systematic Biology 61, 382–391.

Bonnaud, E., Bourgeois, K., Vidal, E., Legrand, J., Le Corre, M., 2009. How can the Yelkouan shearwater survive feral cat predation? A meta-population structure as a solution? Population Ecology 51, 261–270.

Bried, J., Pontier, D., Jouventin, P., 2003. Mate fidelity in monogamous birds: a re-examination of the Procellariiformes. Animal Behaviour 65, 235–246.

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

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York.

Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., Pollock, K.H., 1987. Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture. American Fisheries Society Monograph 5, 1–437.

Choquet, R., Lebreton, J.-D., Gimenez, O., Reboulet, A.M., Pradel, R., 2009. U-CARE: Utilities for performing goodness of fit tests and manipulating CApture-REcapture data. Ecography 32, 1071–1074.

Covas, R., Brown, C.R., Anderson, M.D., Brown, M.B., 2002. Stabilizing selection on body mass in the sociable weaver Philetairus socius. Proceedings of the Royal Society B-Biological Sciences 269, 1905–1909.

Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J.-M., Lebreton, J.-D., Marboutin, E., Miquel, C., Reboulet, A.-M., Poillot, C., Taberlet, P., Gimenez, O., 2010.

Importance of Accounting for Detection Heterogeneity When Estimating Abundance: the Case of French Wolves. Conservation Biology 24, 621–626.

de Villemereuil, P., Wells, J.A., Edwards, R.D., Blomberg, S.P. 2012. Bayesian models for comparative analysis integrating phylogenetic uncertainty. BMC Evolutionary Biology 12:102.

Freckleton, R.P., 2009. The seven deadly sins of comparative analysis. Journal of Evolutionary Biology 22, 1367–1375.

Freckleton, R.P., Harvey, P.H., Pagel, M., 2002. Phylogenetic analysis and comparative data: A test and review of evidence. American Naturalist 160, 712–726.

Gaillard, J.-M., Pontier, D., Allainé, D., Lebreton, J.-D., Trouvilliez, J., Clobert, J., 1989. An analysis of demographic tactics in birds and mammals. Oikos 56, 59–76.

Garland, T., Ives, A.R., 2000. Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. American Naturalist 155, 346–364.

Gimenez, O., Crainiceanu, C., Barbraud, C., Jenouvrier, S., Morgan, B.J.T., 2006. Semiparametric regression in capture-recapture modelling. Biometrics 62, 691–698.

Grosbois, V., Gimenez, O., Gaillard, J.-M., Pradel, R., Barbraud, C., Clobert, J., Møller, A.P., Weimerskirch, H., 2008. Assessing the impact of climate variation on survival in vertebrate populations. Biological reviews 83, 357–99.

Halsey, L.G., Butler, P.J., Blackburn, T.M., 2006. A phylogenetic analysis of the allometry of diving. American Naturalist 167, 276–287.

Hansen, T.F., Bartoszek, K., 2012. Interpreting the Evolutionary Regression: The Interplay Between Observational and Biological Errors in Phylogenetic Comparative Studies. Systematic Biology 61, 413–425.

Harvey, P.H., Pagel, M.D., 1991. The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford.

Ives, A.R., Zhu, J., 2006. Statistics for correlated data: Phylogenies, space, and time. Ecological Applications 16, 20–32.

Johnston, J.P., Peach, W.J., Gregory, R.D., White, S.A., 1997. Survival rates of tropical and temperate passerines: A Trinidadian perspective. American Naturalist 150, 771–789.

Keitt, B.S., Tershy, B.R., Croll, D.A., 2004. Nocturnal behavior reduces predation pressure on black-vented shearwaters *puffinus opisthomelas*. Marine Ornithology 32, 173–178.

Laake, J., 2012. RMark: R Code for MARK Analysis.

Lahoz-Monfort, J.J., Morgan, B.J.T., Harris, M.P., Wanless, S., Freeman, S.N., 2011. A capture-recapture model for exploring multi-species synchrony in survival. Methods in Ecology and Evolution 2, 116–124.

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

Lebreton, J.-D., Hines, J.E., Pradel, R., Nichols, J.D., Spendelow, J.A., 2003. Estimation by capture-recapture of recruitment and dispersal over several sites. Oikos 101, 253–264.

Link, W.A., 1999. Modeling pattern in collections of parameters. The Journal of Wildlife Management 63, 1017–1027.

Link, W.A., Barker, R.J., 2004. Hierarchical mark-recapture models: a framework for inference about demographic processes. Animal Biodiversity and Conservation 27, 441–449.

Millar, R.B., 2009. Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors. Biometrics 65, 962–969.

Monteiro, L.R., Ramos, J.A., Furness, R.W., Del Nevo, J., 1996. Movements, morphology, breeding, molt, diet and feeding of seabirds in the Azores. Colonial Waterbirds 19, 82–97.

Mougeot, F., Bretagnolle, V., 2000. Predation risk and moonlight avoidance in nocturnal seabirds. Journal of Avian Biology 31, 376–386.

Oro, D., de León, A., Minguez, E., Furness, R.W., 2005. Estimating predation on breeding European storm-petrels (*Hydrobates pelagicus*) by yellow-legged gulls (*Larus michahellis*). Journal of Zoology 265, 421–429.

Papadatou, E., Pradel, R., Schaub, M., Dolch, D., Geiger, H., Ibanez, C., Kerth, G., Popa-Lisseanu, A., Schorcht, W., Teubner, J., Gimenez, O., 2012. Comparing survival among species with imperfect detection using multilevel analysis of markurecapture data: a case study on bats. Ecography 35, 153–161.

Paradis, E., Claude, J., 2002. Analysis of comparative data using generalized estimating equations. Journal of Theoretical Biology 218, 175–185.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.

Peach, W.J., Hanmer, D.B., Oatley, T.B., 2001. Do southern African songbirds live longer than their European counterparts? Oikos 93, 235–249.

Pollock, K.H., Nichols, J.D., Brownie, C., Hines, J.E., 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs 1–97.

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

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

Pradel, R., Gimenez, O., Lebreton, J.-D., 2005. Principles and interest of GOF tests for multistate capture-recapture models. Animal Biodiversity and Conservation 28, 189–204.

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

Péron, G., Koons, D.N., 2012. Integrated modeling of communities: parasitism, competition, and demographic synchrony in sympatric ducks. Ecology 93, 2456–2464.

R Development Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Revell, L.J., 2010. Phylogenetic signal and linear regression on species data. Methods in Ecology and Evolution 1, 319–329.

Saether, B.-E., 1989. Survival rates in relation to body weight in European birds. Ornis Scandinavica 20, 13–21.

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

Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D., 2003. WinBUGS user manual version 1.4.

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

Warham, J., 1990. The Petrels. Their Ecology and Breeding Systems. Academic Press, London.

Westoby, M., Cunningham, S.A., Fonseca, C.R., Overton, J.M., Wright, I.J., 1998. Phylogeny and variation in light capture area deployed per unit investment in leaves: designs for selecting study species with a view to generalising. Inherent variation in plant growth. Physiological mechanisms and ecological consequences (ed. by H. Lambers, H. Poorter and M.M.I. Van Vuuren). Blackhuys, Leiden, The Netherlands.

White, G.C., Burnham, K.P., 1999. Program MARK: Survival estimation from populations of marked animals. Bird Study 46, S120–139.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic Press, San Diego, USA.

Appendix A: Body mass, phylogenetic tree and the corresponding covariance matrix of the 16 species of Procellariiformes.

Appendix B: Supplementary Data - Preliminary analyses (goodness-of-fit (GOF) tests and model selection) for each species in program U-CARE and MARK.

Appendix C: Supplementary Data - The R and WinBUGS code used for the analysis in the paper.

Figures captions

Fig.1. A hypothetical phylogenetic tree of five species and the corresponding covariance matrix, $\sum$. Numbers on the tree are branch lengths that are used to calculate the elements of $\sum$.

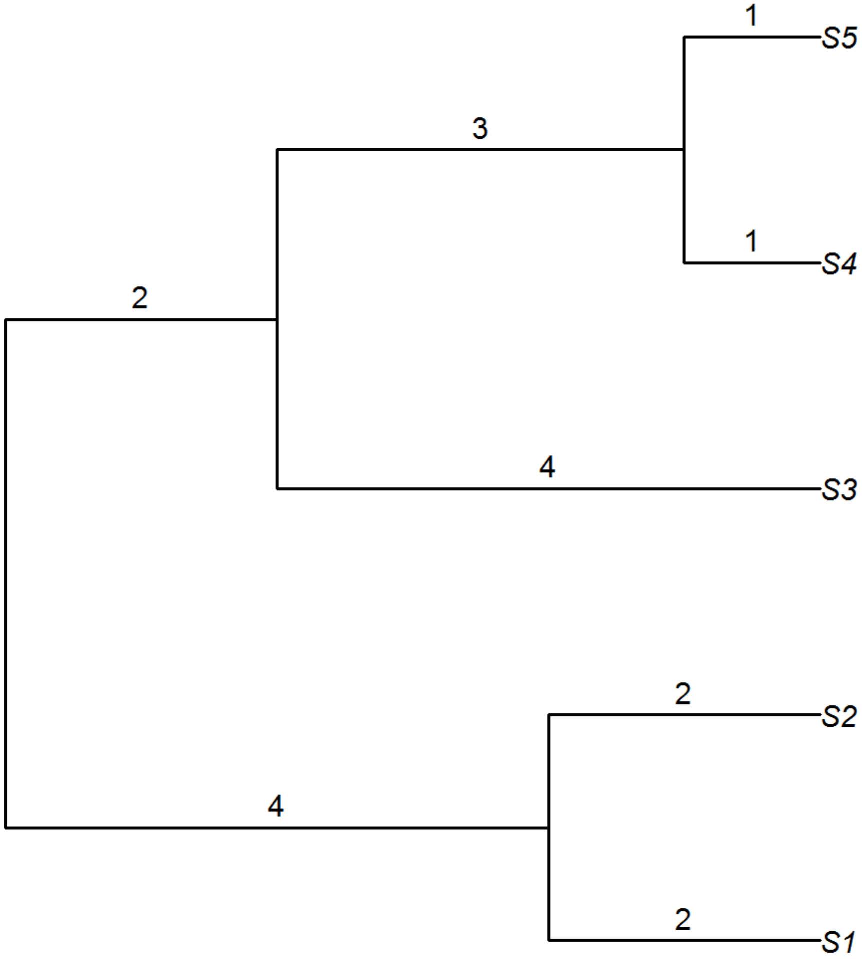
Fig. 2. Estimated relationship between mean survival of seabird species and body mass (on log scale) obtained from the models with phylogeny (solid line) and without phylogeny (dashed lines). Points are the mean survival estimates obtained from species-specific independent analyses. Note that different points (square, circle, and triangle) are used to highlight phylogenetic closeness (see Fig. A1, Appendix A).

Fig. 3. Estimated posterior means (and 95% credible intervals; vertical bars) of adult survival derived from models that accounted for phylogeny (Model 1; solid circle) and ignored phylogeny (Model 2; open circle). Species names are listed in the x-axis.

Table 1. Estimated posterior means, standard deviations and 95% credible intervals (CRI) of the regression coefficients obtained from the two models (Model 1: Accounting for phylogeny; Model 2: Ignoring phylogeny) fitted to the Procellariiformes data (see Eq. (2)).

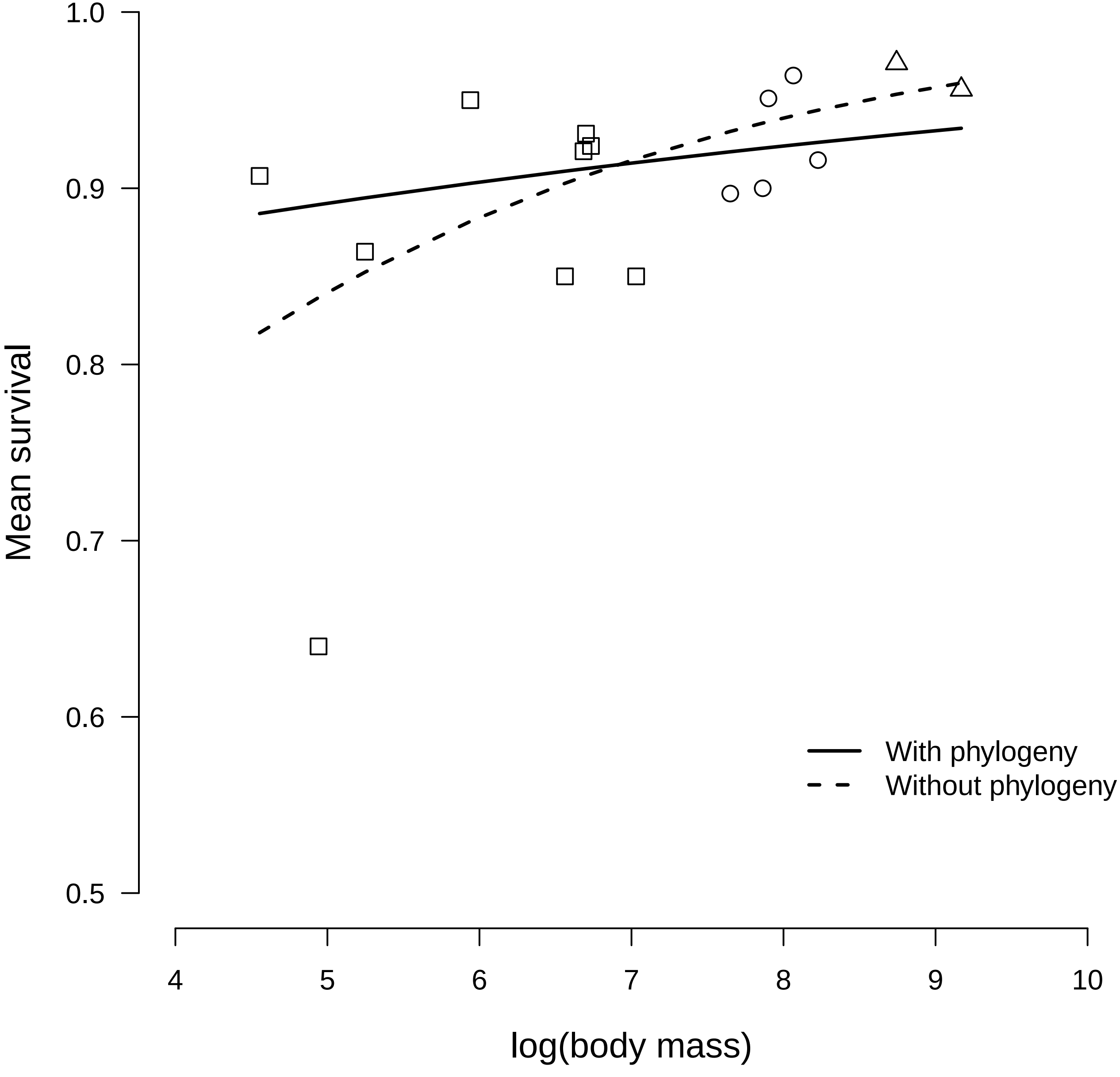
	Parameter	Mean	Std. dev.	95% CRI
Model 1	intercept (β_0)	1.452	1.703	[-1.693; 5.077]
	log(body mass) (β_1)	0.131	0.221	[-0.325; 0.541]
Model 2	intercept (β_0)	-0.143	0.931	[-1.967; 1.665]
	log(body mass) (β_1)	0.361	0.129	[0.110; 0.619]

Figure

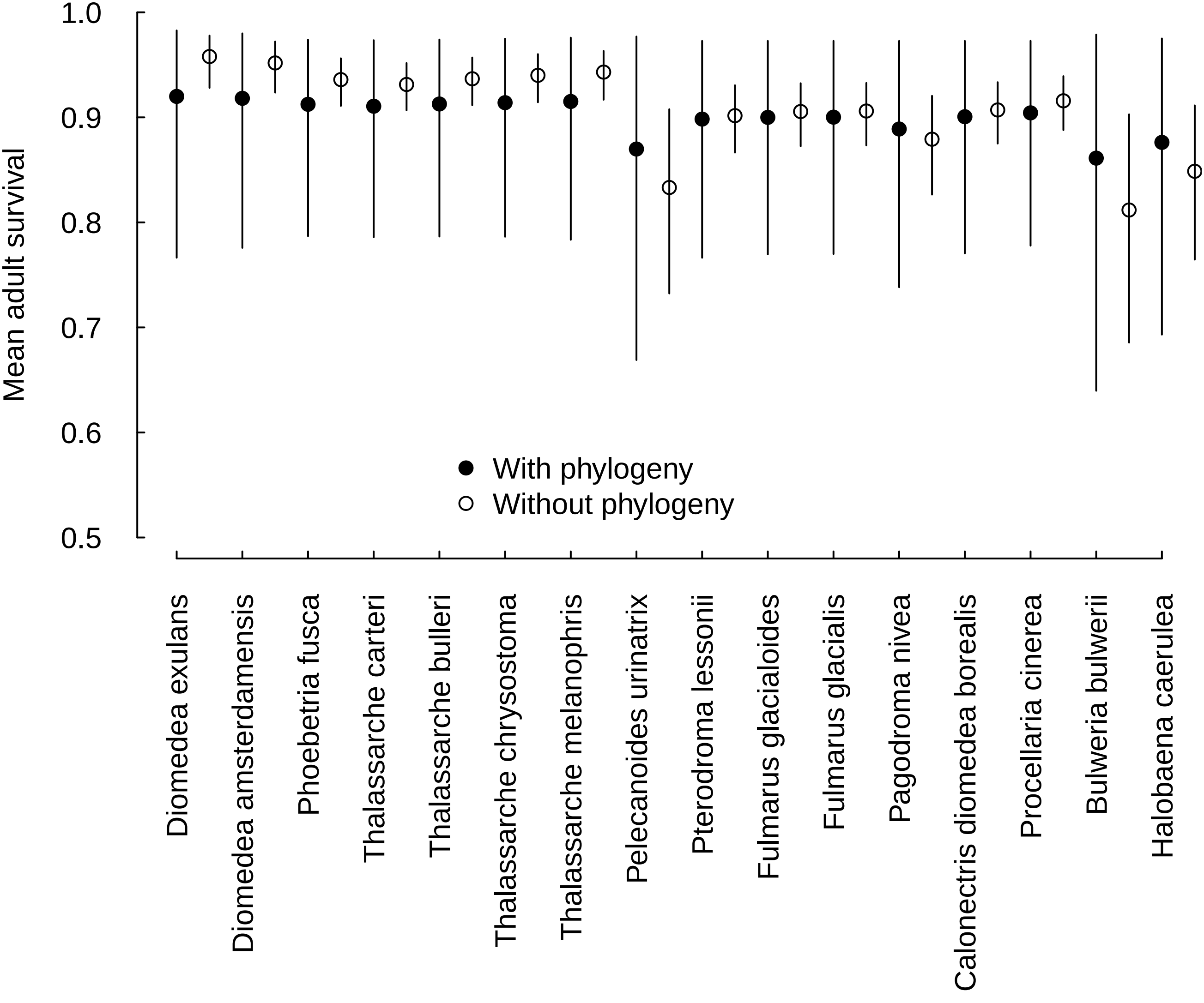


	S1	S2	S3	S4	S5
S1	6	4	0	0	0
S2	4	6	0	0	0
S3	0	0	6	2	2
S4	0	0	2	6	5
S5	0	0	2	5	6

Figure


1.0
0.9
0.8
0.7
0.6
0.5
Mean survival
4
5
6
7
8
9
10
log(body mass)
With phylogeny
Without phylogeny

Figure


1.0
0.9
0.8
0.7
0.6
0.5
Mean adult survival
● With phylogeny
○ Without phylogeny
Diomedea exulans
Diomedea amsterdamensis
Phoebetria fusca
Thalassarche carteri
Thalassarche bulleri
Thalassarche chrysostoma
Thalassarche melanophris
Pelecanoides urinatrix
Pterodroma lessonii
Fulmarus glacialoides
Fulmarus glacialis
Pagodroma nivea
Calonectris diomedea borealis
Procellaria cinerea
Bulweria bulwerii
Halobaena caerulea