

Revisiting the Effect of Capture Heterogeneity on Survival Estimates in Capture-Mark-Recapture Studies: Does It Matter?

Fitsum Abadi^{1,2}, Andre Botha³, Res Altwegg^{1,2*}

1 South African National Biodiversity Institute, Claremont, South Africa, **2** Animal Demography Unit, Department of Biological Sciences and Department of Statistical Sciences, University of Cape Town, Rondebosch, South Africa, **3** Endangered Wildlife Trust, Modderfontein, South Africa

Abstract

Recently developed capture-mark-recapture methods allow us to account for capture heterogeneity among individuals in the form of discrete mixtures and continuous individual random effects. In this article, we used simulations and two case studies to evaluate the effectiveness of continuously distributed individual random effects at removing potential bias due to capture heterogeneity, and to evaluate in what situation the added complexity of these models is justified. Simulations and case studies showed that ignoring individual capture heterogeneity generally led to a small negative bias in survival estimates and that individual random effects effectively removed this bias. As expected, accounting for capture heterogeneity also led to slightly less precise survival estimates. Our case studies also showed that accounting for capture heterogeneity increased in importance towards the end of study. Though ignoring capture heterogeneity led to a small bias in survival estimates, such bias may greatly impact management decisions. We advocate reducing potential heterogeneity at the sampling design stage. Where this is insufficient, we recommend modelling individual capture heterogeneity in situations such as when a large proportion of the individuals has a low detection probability (e.g. in the presence of floaters) and situations where the most recent survival estimates are of great interest (e.g. in applied conservation).

Citation: Abadi F, Botha A, Altwegg R (2013) Revisiting the Effect of Capture Heterogeneity on Survival Estimates in Capture-Mark-Recapture Studies: Does It Matter? PLoS ONE 8(4): e62636. doi:10.1371/journal.pone.0062636

Editor: Matt Hayward, Australian Wildlife Conservancy, Australia

Received: November 7, 2012; **Accepted:** March 23, 2013; **Published:** April 30, 2013

Copyright: © 2013 Abadi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work supported by grants from the South African National Research Foundation (no. 68561), the Endangered Wildlife Trust, BirdLife South Africa, the Claude Leon Foundation, and the University of Cape Town. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: res.altweg@gmail.com

Introduction

Survival of animals in the wild is an important fitness component, and unbiased survival estimates are critical for understanding, among other things, the patterns of life histories [1], evolutionary pressures in the wild (e.g. [2]), and for the conservation of populations [3]. Development of sophisticated open capture-mark-recapture models [4] has revolutionized our knowledge of survival in populations of wild animals (reviewed by [5]). In theory, these methods give unbiased survival estimates by incorporating an estimate of the detection probability (i.e. the probability of recapture (or resighting) an individual that is alive and in the population at the time of a survey) into the estimation of survival probability. The detection probability is often regarded as a nuisance parameter, and is usually of little biological interest (but see [6]). However, as methods are now used extensively (the key publication, [4], has been cited 2109 times, according to ISI Web of Science, accessed on 14 March 2013), the importance of accounting for the detection probability is becoming clear: almost all studies find detection probabilities < 1 (implying that perfect detection is hardly ever achieved). Further, most studies find the detection rate to vary among groups of individuals (e.g. age classes and sex), and over time and space. This suggests modelling variation in detection probabilities is critical for obtaining

unbiased survival estimates from capture-mark-recapture experiments on wild populations [4,7,8].

Conventional capture-mark-recapture methods make the critical assumption of equal detection probability among individuals within a group. This assumption is generally tested using standard tests (Test 2 in RELEASE, ([9]; also available within program MARK: [10]), and U-CARE, ([11,12])), and often found to be violated. However, since [13] found that the departure from the assumption of homogeneity in detection causes negligible bias in survival estimates, the resulting lack of fit (to which unmodelled individual heterogeneity contributes) is normally dealt with by multiplying the variance-covariance matrix by a constant variance inflation (overdispersion) factor, \hat{c} [4,10]. This inflates the confidence intervals, but does not attempt to correct any potential bias in the mean estimate.

More recently, however, concerns about the assumption that capture heterogeneity in the estimation of survival and/or population size can safely be ignored have been voiced [14–19]. At the same time, the development in hierarchical models for analysing capture-mark-recapture data has made it possible to address this issue as capture heterogeneity can now be modelled in various ways [15,20–22]. While appealing, these methods add complexity to the analyses, and are more difficult to fit and assess for general users of capture-mark-recapture methodology.

Therefore, the objectives of this study are to reassess effects of different forms of capture heterogeneity on survival estimates, and to identify situations in which the assumption of homogenous detection probability can safely be applied to empirical data sets. We use simulations and two case studies where we expected capture heterogeneity to be strong, to illustrate pros and cons of modelling capture heterogeneity in capture-mark-recapture studies.

Methods

The family of Cormack-Jolly-Seber (CJS) models is widely used to estimate survival probabilities from capture-recapture data. This model can be implemented either using the multinomial likelihood [4,9] or the state-space formulation [20,23,24]. Here, we used the latter approach as it provides a flexible framework for incorporating individual heterogeneity and easily fitting other modified models [20,23,25–28]. Before we present the model, we define the notations. Let T be the number of sampling occasions in year, ϕ be the survival probability, which is assumed constant over time and identical for all individuals, and p_i be the detection probability of individual i and constant over time. The model is expressed by the state process, which describes the true biological process (e.g. survival), and the observation process, which describes the error associated with the data (e.g. imperfect detection). Following [20,23], the state process (i.e. $z = 1$ (alive) or 0 (dead) for individual i at time t) is described by Bernoulli trials as

$$z_{i,t+1}|z_{i,t} \sim \text{Bernoulli}(z_{i,t}\phi) \tag{1}$$

The state of an individual at the time of first capture is known with probability 1 (i.e. it is alive with certainty). The observation process (i.e. $y = 1$ (seen alive), 0 (not seen) for individual i at time t) is modelled by Bernoulli trials as

$$y_{i,t}|z_{i,t} \sim \text{Bernoulli}(z_{i,t}p_i) \tag{2}$$

To assess the effect of ignoring capture heterogeneity, we considered a model that allowed for capture heterogeneity ($\{\phi, p_i\}$) vs. a model that assumed constant detection probability across individuals ($\{\phi, p\}$). In the former, we used a random effects model to account for capture heterogeneity, and individual random terms (ε_i) are the deviations from the global mean μ_p , which are distributed normally with a mean of 0 and standard deviation σ_p . That is,

$$\text{logit}(p_i) = \mu_p + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_p^2) \tag{3}$$

1. Simulation Study

We carried out a simulation study to assess the potential bias in estimates of survival probability when capture heterogeneity is ignored. In conducting simulations, we first needed to specify the parameters. The number of sampling occasions in year (T) set at 15 years, the survival probability (ϕ) was set at 0.7 and assumed to be constant over time. We considered four different scenarios of capture heterogeneity (p_i), which could frequently arise in empirical capture-recapture studies: 1) symmetric heterogeneity around a mean detection probability (Figure 1a); 2) right-skewed distribution of detection probabilities (e.g. most individuals have relatively low detection probabilities but a few individuals are

being caught repeatedly; a situation that could arise if the study area is relatively small in relation to the movement patterns of the individuals) (Figure 1b); 3) left-skewed distribution (e.g. most individuals have relatively high detection probabilities but a few individuals are unlikely to be detected because they have their home ranges along the periphery of the study area) (Figure 1c); and 4) two-group heterogeneity, a situation if the studied population consists of two groups (e.g. females and males, non-breeders and breeders, social status) that cannot be distinguished in the field but differ in their propensity to be trapped (Figure 1d). The first three scenarios of detection probabilities were generated from a beta distribution with means 0.4, 0.2, and 0.8 and standard deviations 0.148, 0.163, and 0.163, respectively. In the fourth scenario, we considered a detection probability of 0.2 for one group and 0.8 for the other group. For each scenario, the annual number of newly marked individuals was chosen to be 50. We then simulated 100 replicate data sets each under the assumption that survival probability is constant over time and identical for all individuals, and under the assumption that detection probability varies only across individuals.

To each simulated data set, we fitted a model that allows for heterogeneity in detection probability among individuals ($\{\phi, p_i\}$), and a model that assumes constant detection probability across individuals ($\{\phi, p\}$). The former model is equivalent to the generating model for the first scenario, with symmetric heterogeneity. We used this model ($\{\phi, p_i\}$) to analyse all data sets, including the ones generated under the other scenarios for heterogeneity, to see how well this approach works in different situations.

All the analyses were performed within the Bayesian framework, specifying non-informative priors to reflect little *a priori* knowledge about the parameters. We used uniform prior distributions between 0 and 1 (U(0,1)) for the survival and mean detection probabilities, and a U(0,10) prior distribution for the standard deviation of the detection probabilities. We first assessed the convergence of the Markov Chain Monte Carlo (MCMC) algorithm to the targeted posterior distribution by running 3 chains of 10000 iterations with a burn-in of 7000. The \hat{R} [29] values were below 1.01 for all parameters, indicating convergence. We then ran a single chain of 30000 iterations, a burn-in of 20000, and retained every 10th observation for each simulated data. Thus, the posterior summary statistics were computed based on 1000 MCMC samples. We then calculated the relative bias in estimates of survival probability as

$$\text{Bias}(\mathcal{G}) = \frac{\left(\frac{1}{ns} \sum_{k=1}^{ns} \mathcal{G}_k - \mathcal{G}\right)}{\mathcal{G}} \tag{4}$$

where ϕ_k is the estimated survival probability for the k^{th} simulated data, ϕ is the survival probability used to generate the data, and ns is the number of simulated data sets. Further, we assessed the effect of ignoring detection heterogeneity on the precision of survival estimates. The data simulation was carried out using R [30] and the analysis was performed using WinBUGS calling from R using R2WinBUGS [31].

Simulation results: Bias and precision of survival estimates. Ignoring heterogeneity in detection probability led to negative bias in estimates of survival probability (Figure 2a). Though the bias was small in all scenarios, our analyses revealed the largest bias in the cases of the right-skewed scenario (i.e., most individuals had relatively low detection probabilities) and two groups with different detection probabilities. A model accounting for heterogeneity essentially produced unbiased estimates of

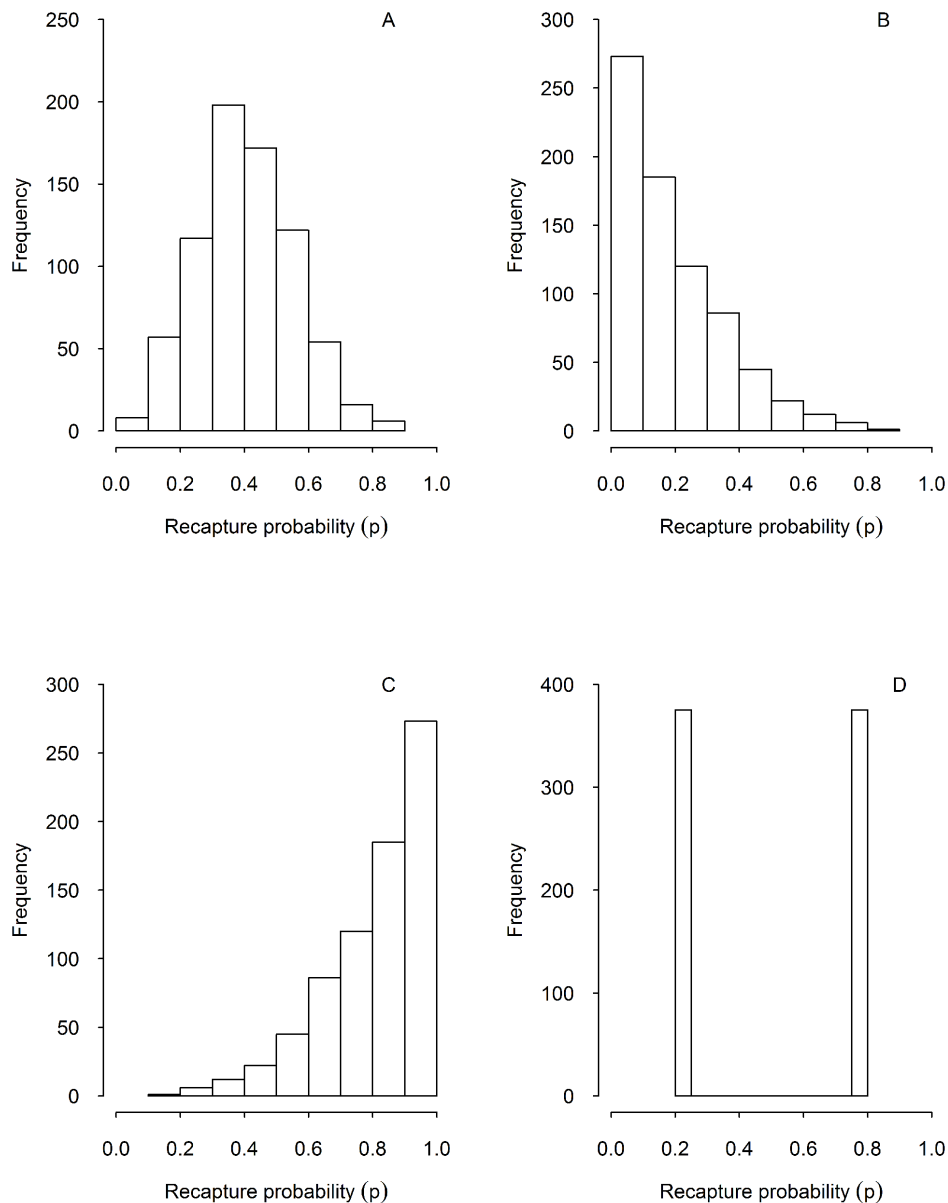


Figure 1. Different scenarios of heterogeneity in detection probabilities. a) Symmetric individual detection probabilities, b) right-skewed (most individuals had a lower detection probability), c) left-skewed (most individuals had a higher detection probability), and d) two-group heterogeneity (individuals with low and high detection probabilities). doi:10.1371/journal.pone.0062636.g001

survival probability in all cases (Figure 2a). As expected, the model that allows for heterogeneity in detection probability provided slightly less precise (i.e. large standard deviation) estimates of survival probability compared to the one ignoring it (Figure 2b). The individual random effects model slightly underestimated the standard deviations of the detection probabilities (with relative bias ranging from -0.19 to -0.05).

Goodness-of-fit (GOF) tests. To see whether commonly used diagnostics would flag the detection heterogeneity in our simulations, we randomly chose five simulated data sets per scenario and tested for detection heterogeneity in program U-CARE [11,12]. The overall GOF test (i.e., TEST3.SR+TEST3.SM+TEST2.CT+TEST2.CL) for the CJS model showed little evidence for lack of fit at the 5% significance level except for the scenario with two groups differing in detection

probabilities (Appendix S1). For this scenario, the directional tests for both transience and trap dependence were significant (all p-values < 0.05). For all selected data sets, the estimated overdispersion parameter (\hat{c}) varied between 1.000 and 2.162, suggesting evidence of overdispersion (Appendix S1). Interestingly, for the second scenario (Figure 1b), the TEST2.CT test showed signs of transience whereas for the third scenario (Figure 1c), the TEST3.SR test for trap dependence was statistically significant (Appendix S1).

2. Case Studies

Our simulation study clearly showed that ignoring capture heterogeneity led to a small bias in survival estimates and a slight overestimation of its precision (i.e. small standard deviation). We also examined the issue of capture heterogeneity in two case

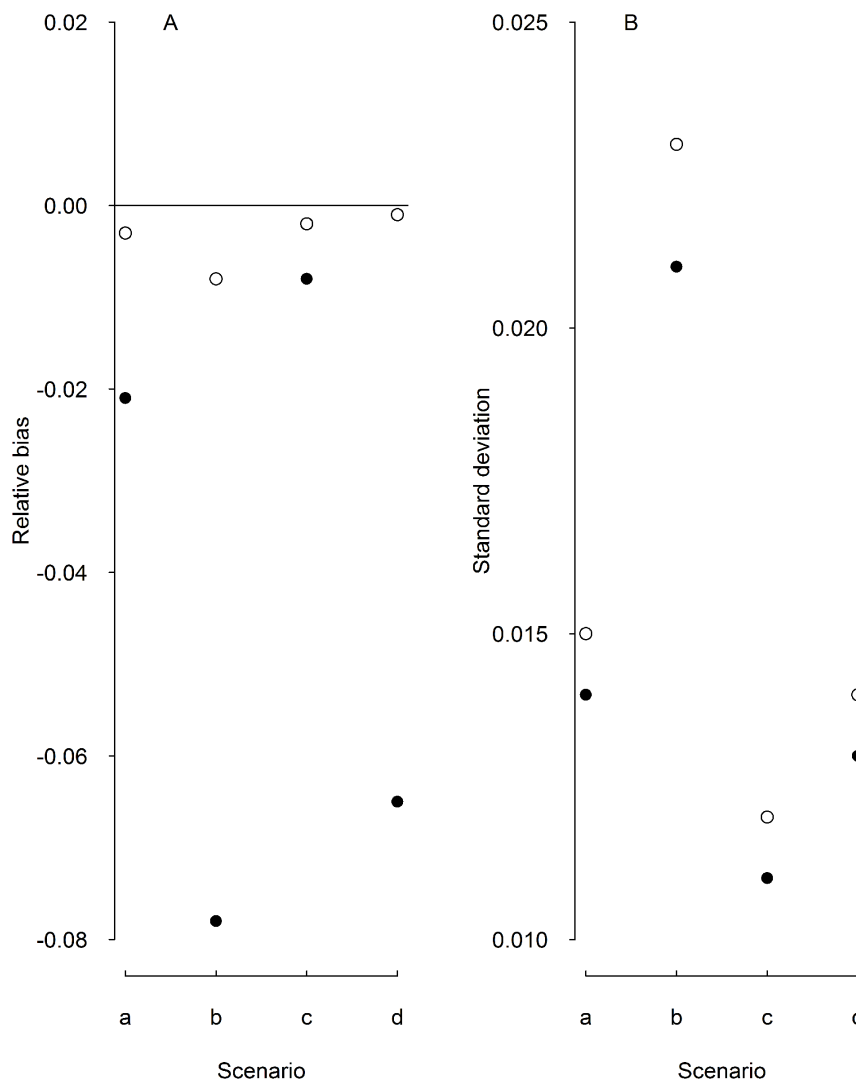


Figure 2. Relative bias and precision in the estimate of survival probability. The relative bias (panel 1) in the estimate of survival probability and precision in terms of standard deviation (panel 2) from a model that ignores (solid circle) and accounts for heterogeneity (open circle) in detection probability under different scenarios: a) symmetric individual heterogeneity, b) right skewed, c) left skewed, and d) two-group heterogeneity.

doi:10.1371/journal.pone.0062636.g002

studies involving data on African White-backed Vultures (*Gyps africanus*) and on African Penguins (*Spheniscus demersus*). While we advocate reducing capture heterogeneity by choosing a sampling design that minimizes the problem, this is not always possible and our case studies are examples of the latter. Both species are of conservation concern and subject to tagging programs that can be used to assess effective management strategies. In both data sets, we expected strong detection heterogeneity for reasons detailed below, and we assessed the extent to which this may lead to biased survival.

Ethics statement. Capturing and tagging of the birds was done under permits from SAFRING, CapeNature, and the Department of Environmental Affairs (DEA). Under South African laws, when in possession of a research permit allowing banding of penguins, no additional ethics clearance is required for these birds. They were captured by hand and stainless steel flipper bands were applied according to the guidelines approved by the DEA [32], who was also responsible for ethical oversight. Stainless steel bands are tear-drop shaped and the ends overlap, allowing

each band to be individually fitted to the penguin using custom-made banding pliers. The penguins were banded by officials of the DEA and the South African Bird Ringing Unit (SAFRING) with permits issued under the Sea Birds and Seals Protection Act No. 46 of 1973, the Marine Living Resources Act No. 18 of 1998, and the National Environmental Management Biodiversity Act No. 10 of 2004. Both institutions agreed on the use and publication of these data. The vultures were fitted with patagial tags, for which ethics clearance was provided by the Endangered Wildlife Trust Ethics Committee and the ethics committee at South African National Parks. The effect of tagging was minimised by adopting the standard protocol adopted for this practice in southern Africa [33].

The African White-backed Vulture Study. Ninety-three vultures were captured using carcass baited walk in traps between November 2005 and January 2007 at the Moholoholo Wildlife Centre near Kampersrus in Mpumalanga, South Africa. Captured birds were fitted with unique alphanumeric coded patagial tags and standard metal leg rings. Individuals were then resighted

monthly between December 2005 and June 2010 near the capture site where vultures were being fed and also within the Greater Kruger National Park where the birds are known to forage. Resightings away from the capture site were reported by members of the public visiting or staff working within the Kruger National Park and adjacent nature reserves. While some vultures visited this site regularly, others were only found there occasionally, thus creating strong individual heterogeneity in resighting probabilities (Figure 3). See [34] for more details.

The African Penguin study. We analysed a capture-mark-recapture data set consisting of 5558 adult African Penguins banded and resighted on Robben Island, South Africa, between 2002 and 2009. Birds returning to their breeding colony were observed with a spotting scope and their flipper bands read. The birds use different paths to walk from the beach to the colony and even though flipper bands were read in all areas that penguins used, most of the resighting effort was concentrated on two main paths used by most penguins. We expected the spatial heterogeneity in effort to translate into individual resighting heterogeneity if penguins consistently used the same areas in their colony, as is the norm for seabirds. There is also the possibility of unidentified groups with different resighting probabilities as we could not tell apart the sexes in the field and a range of band types of different quality were used [35], even though the latter effect could have been incorporated into the model structure. More information on the penguin study can be found in [36].

For both data sets, we fitted a model allowing for individual heterogeneity in resighting probabilities and one assuming constant resighting probabilities, using the same methods as for analysing the simulated data above. The models further allowed for a time effect on survival and resighting probabilities, treating time as a random effect in the longer vulture study and as a fixed effect in the shorter penguin study. We computed the posterior summary statistics using a single chain of 200000 iterations after discarding the initial 100000 iterations as a burn-in period and thinned by using every 10th observation. The R and WinBUGS codes used for fitting all models are available on request.

Case studies results. For the African White-backed Vulture data, the estimated monthly survival probabilities were close to 1 except for a few months (Figure 4a). The posterior distribution of the standard deviation of monthly individual resighting probability (on the logit scale) had a mean of 1.78 (95% CRI: 1.50–2.10),

```

100000000001000010000000000000...
100000000000011010000000001100...
101100101100011110111000111100...
111111111111111111111110111111...
1111101111111111110111001111001...
100110110011110110011101101100...
101000100000000000000000100000...
10000000000100000001000000000...
1010111101111111111111001001011...
111100000000000001000000100000...
10000001111111111111110111110...
    
```

Figure 3. Example of parts of the capture-resighting histories for a subset of the African White-backed Vultures. The rows correspond to individuals and the column to months. If a particular bird was seen in a given month, its sighting history contains a '1' in the corresponding column, and '0' otherwise. The capture histories suggest strong heterogeneity in resighting probabilities, probably due to individual differences in movement patterns. doi:10.1371/journal.pone.0062636.g003

indicating evidence of heterogeneity in resighting probabilities. However, both models, the one ignoring heterogeneity and the one allowing for it, provided similar estimates of survival probabilities for most months of the study (Figure 4a). The overall GOF test from U-CARE indicated a serious lack of fit ($\chi^2 = 627.903$, $df = 142$, $P < 0.001$) with the largest contribution coming from TEST2.CT ($\chi^2 = 487.068$, $df = 53$, $P < 0.001$). The directional test for transience was statistically significant and the \hat{c} obtained from the GOF test showed substantial overdispersion ($\hat{c} = 4.422$). In general, the model ignoring heterogeneity yielded survival estimates with shorter confidence intervals than the one accounting for it (Figure 4a). For African Penguins, the mean annual survival probabilities varied between 0.530 and 0.817, and the model ignoring heterogeneity tended to underestimate the survival probability in some years (Figure 4b). Our analysis revealed that the mean of the posterior distribution of the standard deviation of the resighting probability was 1.239 (95% CRI: 0.883–1.637), suggesting evidence of heterogeneity in resighting individuals. The overall GOF test result from U-CARE showed overdispersion ($\chi^2 = 83.924$, $df = 28$, $P < 0.001$), and TEST2.CT was highly significant ($\chi^2 = 37.292$, $df = 5$, $P < 0.001$). The directional test for transience was statistically significant and the \hat{c} obtained from the GOF test showed overdispersion ($\hat{c} = 2.997$). Yet, ignoring this variation only had a small effect on the survival estimates and the precision of the survival estimates was comparable for both models (Figure 4b).

Discussion

The traditional capture-mark-recapture modelling framework assumes constant survival and detection probabilities across individuals [4,37–39]. This is of course never strictly true in real situations, but early studies showed that heterogeneity among individuals results in only slight bias in survival estimates [13]. More recently, doubts have been raised whether individual heterogeneity can safely be ignored [14,19,40] and methods to account for such heterogeneity have been developed [15,21]. Individual heterogeneity can conveniently be modelled as individual random effects when formulating the model as a state-space process [20,23] but popular software packages also offer individual random effects within the classical capture-mark-recapture modelling framework (recent versions of MARK [10] and E-SURGE [21]). We used simulations and two case studies to examine in what situations the added complexity of individual random effects is necessary, and what the costs and gains may be.

In our simulations even large individual heterogeneity in detection probabilities caused little bias in survival estimates. The largest negative biases occurred in the cases of right-skewed heterogeneity (i.e., most individuals had relatively low detection probabilities) and with two groups that differ in detection probabilities. Intuitively, the negative bias arises because the detection estimate is dominated by individuals that are detected more frequently because their detection probability is high. This leads to positive bias in detection estimates and negative bias in the survival estimates. If it was known or suspected that detection heterogeneity was due to a small number of unrecognized groups that differ in detection probabilities, this structure could be modelled using discrete mixtures [40]. In all cases considered, however, modelling heterogeneity as continuous individual random effects essentially eliminated bias. The results of this simulation are thus in agreement with [13] who found small negative biases of unmodelled capture heterogeneity on survival estimates. [16] also found small negative biases in a situation with unmodelled two-group heterogeneity. However, such small

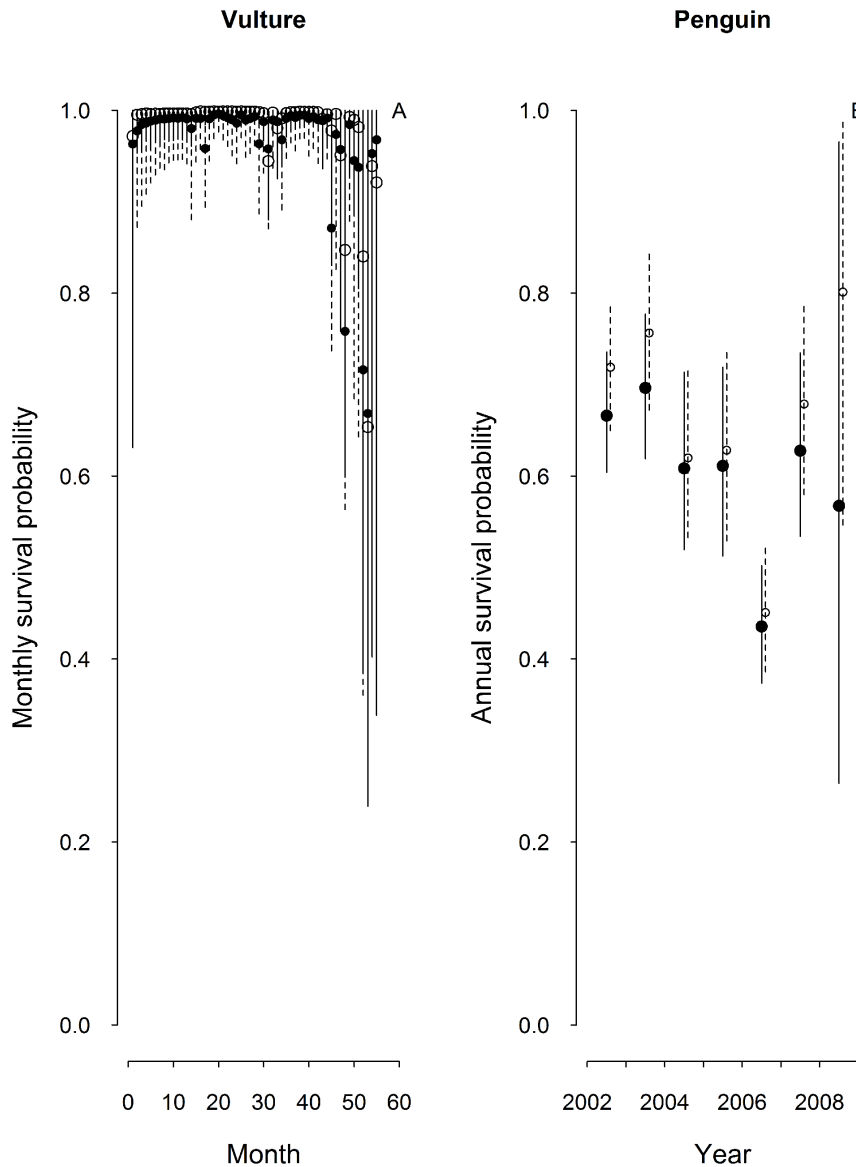


Figure 4. Estimates of survival probability. Mean survival estimates along with the 95% credible interval obtained from a model that ignores heterogeneity (solid symbols and lines) and a model that allows for heterogeneity (open symbols and broken lines) for (a) the African White-backed Vultures data, and (b) the African Penguins data. doi:10.1371/journal.pone.0062636.g004

negative bias in survival estimates may still have a strong impact on the ability to select the correct management measures, in particular for long-lived species where population dynamics is very sensitive to survival [19].

Our case studies supported the results observed in our simulations. Survival estimates slightly increased when we included individual random effects on detection with greater influence towards the end of the time series. This may be due to individuals with low detection probability ‘disappearing’ towards the end of the study, which would make the estimated detection probability increasingly influenced by the individuals that are easy to detect, and thus lead to survival estimates that are increasingly biased low towards the end of the study [16]. In applied conservation the most recent survival estimates are often the most interesting ones, because they are needed to gauge the effectiveness of conservation interventions or to predict future population

declines. We recommend individual random effects to be explored in such situations, particularly if detection probabilities are low. Other strategies for reducing detection heterogeneity (see below) should also reduce the problem of apparently declining survival estimates towards the end of the study.

Precision of survival estimates is also of concern when capture heterogeneity is unmodelled. Such heterogeneity could lead to confidence intervals that are too narrow due to failure to account for uncertainty in detection probabilities. In our simulations, we found that the survival estimates became slightly less precise when we used individual random effects to account for detection heterogeneity, demonstrating that ignoring detection heterogeneity can lead to optimistic confidence intervals. In the vulture case study, the model without individual random effects led to increased precision of survival estimates with confidence intervals that appear to be overly optimistic. However, in the penguin case

study, adding individual random effects had little influence on precision of survival estimates.

We advocate a three-stage strategy to dealing with potential capture heterogeneity. The first stage is to reduce detection heterogeneity by study design, e.g. by standardising field effort across time and space. However, even constant effort is not likely to yield constant detection probabilities as factors beyond the control of the observer can still vary. Furthermore, constant effort may not be possible or practical. As a second stage, we therefore recommend incorporating suspected sources of heterogeneity into the model, e.g. by using covariates that are thought to be related to detection probability. Indirect information on possible hidden heterogeneity (e.g. sex determined by uncertain cues [41]) can be used in multi-event models [42]. As a third stage, we recommend modelling individual heterogeneity where there is evidence for such heterogeneity to remain, for example from a high estimate of \hat{c} or significant goodness-of-fit results. Our simulations showed that heterogeneity can lead to apparent trap effects and transience. Using continuously distributed individual random effects may yield better survival estimates, in terms of bias and precision than applying an overdispersion factor.

In summary, our results suggest that individual detection heterogeneity only has small effects on survival estimates. In practice, the situation that is most likely to bias survival is if a considerable proportion of the individuals have a low detection probability, like our right skewed and two-group scenarios. Transients or the presence of floaters would have such an effect.

References

1. Stearns SC (1992) The evolution of life histories. Oxford: Oxford University Press.
2. Roulin A, Altwegg R, Jensen H, Steinsland I, Schaub M (2010) Sex-dependent selection on an autosomal melanic female ornament promotes the evolution of sex ratio bias. *Ecology Letters* 13: 616–626.
3. Morris WF, Doak DF (2002) Quantitative conservation biology. Sunderland, MA: Sinauer Associates Inc.
4. Lebreton J-D, Burnham KP, Clobert J, Anderson DR (1992) Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* 62: 67–118.
5. Sandercock BK (2006) Estimation of demographic parameters from live-encounter data: a summary review. *Journal of Wildlife Management* 70: 1504–1520.
6. Crespin L, Choquet R, Lima M, Merritt J, Pradel R (2008) Is heterogeneity of catchability in capture–recapture studies a mere sampling artifact or a biologically relevant feature of the population? *Population Ecology* 50: 247–256.
7. Williams BK, Nichols JD, Conroy MJ (2002) Analysis and Management of Animal Populations. San Diego, USA: Academic Press.
8. Gimenez O, Viallefont A, Charmantier A, Pradel R, Cam E, et al. (2008) The risk of flawed inference in evolutionary studies when detectability is less than one. *American Naturalist* 172: 441–448.
9. Burnham KP, Anderson DR, White GC, Brownie C, Pollock KH (1987) Design and Analysis Methods for Fish Survival Experiments Based on Release–Recapture. *American Fisheries Society Monograph* 5: 1–437.
10. White GC, Burnham KP (1999) Program MARK: Survival estimation from populations of marked animals. *Bird Study* 46: S120–S139.
11. 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.
12. Choquet R, Lebreton JD, Gimenez O, Reboulet AM, Pradel R (2009) U-CARE: Utilities for performing goodness of fit tests and manipulating CAapture–REcapture data. *Ecography* 32: 1071–1074.
13. Carothers AD (1979) Quantifying unequal catchability and its effect on survival estimates in an actual population. *Journal of Animal Ecology* 48: 863–869.
14. Pledger S, Efford M (1998) Correction of bias due to heterogeneous capture probability in capture–recapture studies of open populations. *Biometrics* 54: 888–898.
15. Clark JS, Ferraz GA, Oguge N, Hays H, DiCostanzo J (2005) Hierarchical Bayes for structured, variable populations: From recapture data to life-history prediction. *Ecology* 86: 2232–2244.
16. Devineau O, Choquet R, Lebreton JD (2006) Planning capture–recapture studies: Straightforward precision, bias, and power calculations. *Wildlife Society Bulletin* 34: 1028–1035.
17. Cubaynes S, Pradel R, Choquet R, Duchamp C, Gaillard J-M, et al. (2010) Importance of Accounting for Detection Heterogeneity When Estimating Abundance: the Case of French Wolves. *Conservation Biology* 24: 621–626.
18. Péron G, Crochet PA, Doherty PF, Lebreton JD (2010) Studying dispersal at the landscape scale: efficient combination of population surveys and capture–recapture data. *Ecology* 91: 3365–3375.
19. Fletcher D, Lebreton J-D, Marescot L, Schaub M, Gimenez O, et al. (2012) Bias in estimation of adult survival and asymptotic population growth rate caused by undetected capture heterogeneity. *Methods in Ecology and Evolution* 3: 206–216.
20. Royle JA (2008) Modeling individual effects in the Cormack–Jolly–Seber model: A state-space formulation. *Biometrics* 64: 364–370.
21. Gimenez O, Choquet R (2010) Individual heterogeneity in studies on marked animals using numerical integration: capture–recapture mixed models. *Ecology* 91: 951–957.
22. Smout S, King R, Pomeroy P (2011) Estimating demographic parameters for capture–recapture data in the presence of multiple mark types. *Environmental and Ecological Statistics* 18: 331–347.
23. Gimenez O, Rossi V, Choquet R, Dehais C, Doris B, et al. (2007) State-space modelling of data on marked individuals. *Ecological Modelling* 206: 431–438.
24. Dupuis JA (1995) Bayesian estimation of movement and survival probabilities from capture–recapture data. *Biometrika* 82: 761–772.
25. Gimenez O, Covas R, Brown CR, Anderson MD, Brown MB, et al. (2006) Nonparametric estimation of natural selection on a quantitative trait using mark–recapture data. *Evolution* 60: 460–466.
26. Gimenez O, Gregoire A, Lenormand T (2009) Estimating and visualizing fitness surfaces using mark–recapture data. *Evolution* 63: 3097–3105.
27. Marzolin G, Charmantier A, Gimenez O (2011) Frailty in state-space models: application to actuarial senescence in the dipper. *Ecology* 92: 562–567.
28. Kéry M, Schaub M (2012) Bayesian population analysis using WinBUGS – a hierarchical perspective. Burlington: Academic Press.
29. Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulation. *Journal of Computational and Graphical Statistics* 7: 434–455.
30. R Development Core Team (2009) R: A Language and Environment for Statistical Computing.
31. Sturtz S, Ligges U, Gelman A (2005) R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software* 12: 1–16.
32. Petersen SL, Branch GM, Crawford RJM, Cooper J, Underhill LG (2005) The future of flipper banding African penguins: discussion, recommendations and guidelines. *Marine Ornithology* 33: E1–E4.
33. Botha A (2007) A review of colour-marking techniques used on vultures in southern Africa. *Vulture News* 56: 52–63.

The effect of transients is well recognised and accounted for by assuming that these individuals are never recaptured [43,44]. Even though we did not explicitly examine a scenario where some individuals have zero recapture or resighting probability, our results suggest that continuously distributed individual random effects effectively eliminate the bias due to various types of individual heterogeneity. Our results also suggest that individual random effects can improve survival estimates towards the end of the study if detection heterogeneity is present.

Supporting Information

Appendix S1 Summary of the GOF test results from Program U-CARE [12].
(DOCX)

Acknowledgments

Thanks to Barbara Barham and SAFRING for collating and curating the penguin data. We thank all people who have helped to collect the African White-backed Vultures data and the African Penguin data. We thank Olivier Gimenez, Michael Schaub, and two referees for their helpful comments on an earlier version of this manuscript.

Author Contributions

Conceived and designed the experiments: FA RA AB. Analyzed the data: FA RA. Wrote the paper: FA RA AB.

34. Monadjem A, Botha A, Mum C (2012) Survival of the African white-backed vulture *Gyps africanus* in north-eastern South Africa. *African Journal of Ecology* 51: 87–93.
35. Klages NTW, Spencer KD (1996) Flipper bands on penguins: why newer is not always better. *Safring News*: 9–12.
36. Crawford RJM, Altwegg R, Barham BJ, Barham PJ, Durant JM, et al. (2011) Collapse of South Africa's penguins in the early 21st century. *African Journal of Marine Science* 33: 139–156.
37. Cormack RM (1964) Estimates of survival from sighting of marked animals. *Biometrika* 51: 429–438.
38. Jolly GM (1965) Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* 52: 225–247.
39. Seber GAF (1965) A note on multiple-recapture census. *Biometrika* 52: 249–259.
40. Pledger S, Schwarz CJ (2002) Modelling heterogeneity of survival in band-recovery data using mixtures. *Journal of Applied Statistics* 29: 315–327.
41. Genovart M, Pradel R, Oro D (2012) Exploiting uncertain ecological fieldwork data with multi-event capture – recapture modelling: an example with bird sex assignment. *Journal of Animal Ecology* 81: 970–977.
42. Pradel R (2005) Multievent: An extension of multistate capture – recapture models to uncertain States. *Biometrics* 61: 442–447.
43. Johnston JP, Peach WJ, Gregory RD, White SA (1997) Survival rates of tropical and temperate passerines: A Trinidadian perspective. *American Naturalist* 150: 771–789.
44. Pradel R, Hines JE, Lebreton JD, Nichols JD (1997) Capture-recapture survival models taking account of transients. *Biometrics* 53: 60–72.