Estimating the abundance of burrow-nesting species through the statistical analysis of combined playback and visual surveys

Anne-Sophie Bonnet-Lebrun, Lícia Calabrese, Gerard Rocamora and Andrés López-Sepulcre

The conservation of elusive species relies on our ability to obtain unbiased estimates of their abundance trends. Many species live or breed in cavities, making it easy to define the search units (the cavity) yet hard to ascertain their occupancy. One such example is that of certain colonial seabirds like petrels and shearwaters, which occupy burrows to breed. In order to increase the chances of detection for these types of species, their sampling can be done using two independent methods to check for cavity occupancy: visual inspection, and acoustic response to a playback call.

This double-detection process allows us to estimate the probability of burrow occupancy by accounting for the probability of detection associated with each method. Here we provide a statistical framework to estimate the occupancy and population size of burrow-dwelling species. We show how to implement the method using both maximum likelihood and Bayesian approaches, and test its precision and bias using simulated datasets. We subsequently illustrate how to extend the method to situations where two different species may occupy the burrows, and apply it to a dataset on wedge-tailed shearwaters Puffinus pacificus and tropical shearwaters P. bailloni on Aride Island, Seychelles.

The simulations showed that the single-species model performed well in terms of error and bias except when detection probabilities and occupancies were very low. The two-species model applied to shearwaters showed that detection probabilities were highly heterogeneous. The population sizes of wedge-tailed and tropical shearwaters were estimated at 13,716 (95% CI: 12,909–15,874) and 25,550 (23,667–28,777) pairs respectively.

The advantages of formulating the call-playback sampling method statistically is that it provides a framework to calculate uncertainty in the estimates and model assumptions. This method is applicable to a variety of cavity-dwelling species where two methods can be used to detect cavity occupancy.
fact that these corrections assume constant detection probabilities over time and space. If that is not the case, this will strongly hinder spatiotemporal comparisons.

In species that inhabit dens or burrows, estimates of population densities rely on our ability to infer the proportion of those cavities that are inhabited by a breeding pair. This is the case for burrow nesting seabirds, which have a tradition of correcting for detectability through ad hoc application of correction factors (Warham 1996, Ratcliffe et al. 1998, Burger and Lawrence 2001, Scott et al. 2009). For example, some studies have used multiple visits to ascertain occupancy in a subsample of burrows, calculated an average occupancy rate and applied it to the total number of burrows of the sample area (Rayner et al. 2007, Whitehead et al. 2014). Other studies use the playback method, which combines visual and acoustic detection, and calculate a playback response rate which, applied to the number of burrows with undetected birds in it, gives an estimate of the proportion of apparently empty burrows that contained an unresponsive bird (James and Robertson 1985, Burger and Lawrence 2001). While some studies use both visual and acoustic detection methods only on a subset of data to later apply the calculated playback response rate to surveys using only acoustic sampling (James and Robertson 1985), others use both detection methods for all burrows (Burger and Lawrence 2001). In the first case, the number of occupied burrows is calculated as:

$$\hat{N} = n_v p_R$$  \hspace{1cm} (1)

where $n_v$ is the number of birds in which a bird responded to the playback call, and $p_R$ is the response rate calculated from the subsample as the proportion of visible birds that responded to the playback (James and Robertson 1985). When both methods are used for all burrows, the total number of occupied burrows $N$ is subsequently estimated using the following formula:

$$\hat{N} = n_v + n_o \frac{n_{ov}}{n_{ov}}$$  \hspace{1cm} (2)

where $n_v$ is the total number of birds seen, $n_o$ is the number of burrows with unresponsive birds (not seen nor heard), $n_{ov}$ is the number of birds that responded but were not seen, and $n_{ov}$ is the number of birds that were seen and responded.

While the above methods address the bias in our estimates of burrow occupancy by accounting for imperfect detection, they do not provide measures of uncertainty in the estimate and do not overcome the problem that population changes are confounded by changes in detection probability over time and space. Because uncertainty and the ability to make spatiotemporal comparisons in population estimates are key to the management of species, it is clearly desirable to formalize the joint estimation of detection and abundance probabilities in a formal statistical framework.

The aim of this article is to provide a statistical framework for the analysis of playback census. First, we derive the necessary likelihood functions. Second, we will implement them on simulated data using both maximum likelihood and Bayesian approaches, in order to assess the bias and precision of the occupancy estimates. Third, we extend the approach to cases where a burrow can be occupied by two different species. Finally, we implement the approach using data on two species of burrow-nesting seabirds, the wedge-tailed Puffinus pacificus and the tropical shearwater P. bailloni, breeding on Aride Island, an Important Bird Area of the Seychelles (Rocamora and Skerrett 2001). Although particularly useful for the sampling of burrow-nesting seabirds, the approach is applicable to any species where two independent methods of detection (e.g. passive sighting and active luring) can be applied to estimate the occupancy of an animal’s dwelling (e.g. crabs in burrows, fish in shelters, or woodpeckers in tree-holes).

**Material and methods**

**Statistical framework**

Consider a set of $m$ burrows that may be occupied by the species of interest. For each burrow, the detection procedure occurs by two independent methods with different detection probabilities: visual examination and playback. The data is arranged in a matrix $X$ of $m$ rows and two columns, whereby for each burrow $i$, elements $x_{i1}$ and $x_{i2}$ contain a 0 or 1 depending on whether an individual was detected or not with either the first (visual) or second (playback) method respectively. For example, an individual (or pair) that was not seen but heard on burrow $i$, will correspond to $x_{i} = [0,1]$. If we denote the probability of occupancy of the burrow $\Psi$, the probability of visual detection $p_V$, and the probability of playback response $p_R$, the likelihood of all four possible outcomes is given in Table 1. The total likelihood of the data $X$ is therefore:

$$L(X | \Psi, p_V, p_R) = \prod_{i=1}^{m} \left[ \psi_{i} (p_V x_{i1} + (1-p_V)(1-x_{i1})) \right]$$

$$= \prod_{i=1}^{m} \left[ p_R x_{i2} + (1-p_R)(1-x_{i2}) \right] + (1-\psi)(1-x_{i2} x_{i2})$$  \hspace{1cm} (3)

Note that for this to be true, we must make the following biological assumptions. First, we assume that each burrow can only be occupied by a single individual (or pair). Second, we assume independence of the two detection methods: in other words, visual inspection does not affect the probability of response nor vice-versa. In order to ensure this is fulfilled, it is recommendable to use the potentially more disturbing method last (e.g. playback should be done after visual inspection, to ensure birds do not change their behaviour in ways that affect their visibility). Third, it assumes that unoccupied burrows represent the absence of a breeding bird, rather than a temporary absence (e.g. to forage). To ensure this, it is important to time the sampling during the appropriate stage in the breeding cycle and during the animal’s inactive period of the day, which will depend on the target species (e.g. at night for $P. bailloni$ and $P. pacificus$). The latter assumption

| Scenario | Data $x_{i1}$ | Likelihood $L(x_{i1}, | \psi, p_V, p_R)$ |
|----------|---------------|-----------------------------------------|
| Yes | Yes | (1,1) | $\psi p_V p_R$ |
| Yes | No | (1,0) | $\psi p_V (1-p_R)$ |
| No | Yes | (0,1) | $\psi (1-p_V)p_R$ |
| No | No | (0,0) | $\psi (1-p_V)(1-p_R) + (1-\psi)$ |
can be relaxed by the use of repeated observations through time. Note that the illustrated estimate refers to the number of burrows that are occupied at the time of sampling. While burrow occupation of non-breeding individuals is rare, the tropical shearwater is known to breed throughout the year, and thus a proper breeding census would require including samples throughout the year, rather than a single time point as in our example.

The estimation of parameters $\psi$, $p_S$, and $p_R$ can proceed by maximum likelihood or Bayesian methods. For the latter, it will be necessary to define priors. Since all three parameters are probabilities we will use a uniform distribution bounded between 0 and 1 as recommended in Royle and Dorazio (2008). The total number of individuals in a sampling unit can be estimated as $\hat{N} = \psi m$.

Simulated datasets and performance comparison

In order to evaluate the performance of both maximum likelihood and Bayesian methods, we simulated a series of datasets. All datasets represented a plot with 30 burrows (representative of the number of burrows found in a typical plot for the shearwater case study described below). Data was generated using binomial trials for three-way combinations of the following parameter values. For both detection probabilities $p_V$ and $p_R$, we used $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. The occupancy probabilities $\psi$ used were $\{0.1, 0.3, 0.5, 0.7, 0.9\}$. For each combination, we performed a total of 20 simulations.

Maximum likelihood estimation was performed in program R 3.1.2 (R Core Team), using function mle2 in package bbmle (Bolker and R Development Core Team 2014). Bayesian estimation was implemented in JAGS 3.4.0 (Plummer 2003). As uninformative priors for all parameters ($\psi$, $p_V$, $p_R$) we used a uniform distribution from 0 to 1. For each estimation we ran 2000 chains with a burnin of 500.

We evaluated the performance of the estimation by calculating the average error and bias in estimating the total number of occupied burrows, the most likely measure of interest. As a measure of error we used the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{(\hat{N} - N)^2}{n}}$$

where $N$ is the known population size, $\hat{N}$ is its estimate, and $n$ is the expected number of occupied burrows (number of burrows $\times p_w$). The results are presented in standardized units by subtracting the mean RMSE for a given parameter combination and dividing by its standard deviation (Fig. 1).

We calculated relative bias as the proportional difference between estimated and known population size $\frac{(\hat{N} - N)}{N}$.

Models performed better when the probability of occupancy increased, both in terms of bias (Fig. 1) and error (Fig. 2). This is because higher occupancies imply larger sample sizes to estimate detection probabilities. Only plots with $\psi = 0.1$ (an average of 3 out of 30 occupied burrows) showed considerable levels of error and bias. Expectedly, the models also performed better when the probabilities of detection increased. Note that the two probabilities of detection (here visual detection and response to a playback call) are mathematically interchangeable in the simulation. The likelihood framework performed better under low occupancy rates and very different detection probabilities among both methods (e.g. the probability of visual detection being high and the probability of response being low).

Extension to two species

It is not uncommon for burrows and other types of refuge to be suitable for more than one species. In the example of a two-species scenario where the burrow can be occupied...
and is suspected to harbor the largest colony of *P. bailloni* in the world (Rocamora and Skerrett 2001). Both species breed in the same type of natural burrow on the hillsides of the island. While *P. pacificus* is a seasonal breeder found mainly from September to February, *P. bailloni* has no clear breeding season in Seychelles, and incubating birds may be found all year round (Skerrett et al. 2001).

To carry out the sampling analysed here, we followed the playback census protocol described in Betts (1998). In each plot, we noted all potentially suitable burrows (between 1 and 38 burrows per plot, 267 in total) and inspected them visually with a head torch for the presence of a nesting bird of either species. We played recorded male–female duet calls for both species (Rocamora et al. 2000) in the case of unknown content, and for the observed species when the bird was visible. We played the call at the opening of the burrow and noted whether the bird responded by the end of the recordings (1:24 min for *P. bailloni* and 1.58 for *P. pacificus*).

Although *P. bailloni* is a year-round breeder, in this illustrative application, we did only one round of sampling for both species, in November 2011. We surveyed 19 circular plots of 100 m² randomly selected in 1996 for a previous survey (Betts 1998). All surveys occurred at night between 20:00 h and 23:00 h once every sampling month.

---

**Case study: estimating shearwater densities on aride island**

We here present the application of the two-species model to the estimation of the size of a mixed colony on Aride Island (Seychelles) with two breeding shearwaters: the wedge-tailed *Puffinus pacificus* and tropical shearwater *P. bailloni*. Aride Island Nature Reserve comprises 73 ha and is suspected to harbor the largest colony of *P. bailloni* in the world (Rocamora and Skerrett 2001). Both species breed in the same type of natural burrow on the hillsides of the island. While *P. pacificus* is a seasonal breeder found mainly from September to February, *P. bailloni* has no clear breeding season in Seychelles, and incubating birds may be found all year round (Skerrett et al. 2001).

To carry out the sampling analysed here, we followed the playback census protocol described in Betts (1998). In each plot, we noted all potentially suitable burrows (between 1 and 38 burrows per plot, 267 in total) and inspected them visually with a head torch for the presence of a nesting bird of either species. We played recorded male–female duet calls for both species (Rocamora et al. 2000) in the case of unknown content, and for the observed species when the bird was visible. We played the call at the opening of the burrow and noted whether the bird responded by the end of the recordings (1:24 min for *P. bailloni* and 1.58 for *P. pacificus*).

Although *P. bailloni* is a year-round breeder, in this illustrative application, we did only one round of sampling for both species, in November 2011. We surveyed 19 circular plots of 100 m² randomly selected in 1996 for a previous survey (Betts 1998). All surveys occurred at night between 20:00 h and 23:00 h once every sampling month.

---

**Table 2. Likelihood for all possible scenarios in a 2 species case.**

<table>
<thead>
<tr>
<th>Sighting</th>
<th>Response</th>
<th>Species</th>
<th>Data $X_i$.</th>
<th>Likelihood $L(X_i\mid P_{\text{VA}}, P_{\text{VB}}, P_{\text{RA}}, P_{\text{RB}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>A</td>
<td>(1,1)</td>
<td>$\psi_A P_{\text{VA}} P_{\text{RA}}$</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>A</td>
<td>(1,0)</td>
<td>$\psi_A P_{\text{VA}} (1-P_{\text{RA}})$</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>A</td>
<td>(0,1)</td>
<td>$\psi_A (1-p_{\text{VA}}) P_{\text{RA}}$</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>B</td>
<td>(2,2)</td>
<td>$\psi_B P_{\text{VB}} P_{\text{RB}}$</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>B</td>
<td>(2,0)</td>
<td>$\psi_B P_{\text{VB}} (1-P_{\text{RB}})$</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>B</td>
<td>(0,2)</td>
<td>$\psi_B (1-P_{\text{VB}}) P_{\text{RB}}$</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td></td>
<td>(0,0)</td>
<td>$\psi_A (1-p_{\text{VA}})(1-p_{\text{RA}}) + \psi_B (1-p_{\text{VB}})(1-p_{\text{RB}}) + (1-\psi_A-\psi_B)$</td>
</tr>
</tbody>
</table>

---

**Figure 2.** Relative bias of estimates of population size for maximum likelihood and Bayesian estimation for the simulated plots under a variety of parameters.
Parameters were estimated using a Bayesian framework. We specified the following uninformative parameter priors for the random variables:

\[ N(0, 10) \]
\[ \text{Unif}(0, 10) \]

When probabilities were set as constant across plots, we used a uniform prior ranging from 0 to 1.

We ran three independent MCMC chains with 10,000 iterations each, a burnin of 5,000 and thinning of every 10 samples. Convergence was considered achieved when the Gelman–Rubin statistic (Gelman and Rubin 1992) for all parameters was lower than 1.1.

Results

Simulated datasets and performance comparison

Figure 1 and 2 show the standardized root mean squared error (RMSE) and bias, respectively, for the estimation of shearwater densities in the simulated plots. Models per-
formed better when the probability of occupancy increased, both in terms of error (Fig. 1) and bias (Fig. 2). This is because higher occupancies imply larger sample sizes to estimate detection probabilities. Only plots with \( \psi = 0.1 \) (and therefore an average of 3 out of 30 occupied burrows) showed considerable levels of error and bias. As expected, the models also performed better when the probabilities of detection increased. Note that the two probabilities of detection are interchangeable. The Bayesian framework performed worse under low occupancy rates and very different detection probabilities among both methods (i.e. one being high and the other low).

**Case study: estimating shearwater densities on aride island**

Supplementary material Appendix 1, Table A1 shows the performance of two-species shearwater models varying in assumptions of detection heterogeneity. The model with heterogeneity in all probabilities of detection was clearly superior (ΔDIC = 24.08). The parameter estimates using the best model are shown in Supplementary material Appendix 1, Table A2.

Total population sizes were estimated at 25 550 pairs (95% CI: 23 667–28 777) for *P. bailloni* and 13 716 pairs (12 909–15 874) for *P. pacificus*. This is considerably lower than previous estimates using a correction factor approach, particularly for the *P. bailloni* (estimated at 98 000 pairs in 2006–2007, Sampson and Sampson 2007). However, it is important to note that previous estimates were obtained using correction factors rather than the statistical method we here propose.

**Discussion**

We have described a simple statistical framework to estimate the abundance of cavity-dwelling species when cavities are easy to detect and count. In these species, abundance estimation reduces to a problem of occupancy, where the key quantity to estimate is the probability of an individual (or breeding pair) being present in the cavity. This is analogous to the estimation of species occurrence in occupancy models (MacKenzie et al. 2005).

While species occupancy models typically use multiple site visits and a single detection method to estimate detection probabilities, some studies have extended them to incorporate multiple detection methods (Coggins et al. 2014). Nichols et al. (2008) provide a general framework for analysing multiple-method data on species occurrence. The method we presented is closely related to the special case of single site visits, yet applied to the estimation of abundance, rather than presence, when the species inhabits discrete units such as burrows and can be extended to other cavity-dwelling species.

Our formulation bears some implicit assumptions to be considered when interpreting the density estimates. The first one is that all birds are present in their burrow at the time of the survey. If individuals are absent foraging at the time of sampling, or their breeding attempt has ended before sampling, they will not be accounted for. This problem is minimized during egg-laying, incubation, or chick brooding, when at least one bird (one of the adults or the chick) stays behind in the burrow at all times. If this issue, however, is deemed important for the species at hand, the design can be extended to performing multiple visits per burrow in an analogous way to Nichols et al. (2008). In the case of shearwaters, performing the surveys at night, when birds return to their burrows can minimize the risk of sampling when individuals are absent. Moreover, the subset of data we analyze in this article represents a single snapshot of the population in time. Estimation of the year-round population size would require sampling at different times of the year. Design considerations such as how often to sample in order to have a representative sample of breeding attempts will depend on the biological details of the species. In our case, this will be particularly important for *P. bailloni*, which breeds year round. A second important assumption is that both detection methods (visual and playback) are independent of each other (i.e. the use of one does not affect the other). In our case, we decided to always first inspect burrows before playing the call, in order to avoid the call affecting the activity level, and thus visibility, of the bird. Finally, our model also assumes an accurate count of burrows per unit area, which we believe is justifiable given the terrain and dimensions of our plots. However, surveys in areas of more difficult access might require the explicit incorporation of burrow count error through repeated measurements.

Extensions of the model may be developed to adapt to the idiosyncrasies of different species and surveys. For example, while uncommon in our study, burrows may be occupied by more than one individual, potentially requiring detection probabilities to be modeled as dependant on the number of individuals present.

Other possible extensions to the method include modeling the dependence of occupancy or detection probabilities on habitat characteristics (e.g. soil depth, aspect, etc.), as has been done in studies with a single detection method (Pearson et al. 2013, Oppel et al. 2014). This may be important to yield reliable predictions at larger scales that include a variety of environments.

Although we have illustrated an application to the study of shearwaters, the method we outline is applicable to any cavity-dwelling species that may be sampled non-invasively using two detection methods. This may include, not only a variety of burrow-nesting birds such as shearwaters, petrels, penguins or burrowing owls, for which playback detection is commonly used (Haug and Didiuk 1993, Jouventin and Aubin 2002, Conway et al. 2008); but also a variety of other animals like den-living mammals and burrowing crabs or spiders. For example, den occupancy might be estimated using a combination of camera traps and the presence of tracks. Populations of burrowing crabs or spiders might be estimated with a combination of visual burrow inspection and luring the animal, or checking whether a piece of lint placed in the entrance is destroyed, as illustrated by Pombo and Turra (2013).

One advantage of our proposed method is that it does not require destructive sampling to yield reliable estimates of burrow occupancy. Destructive methods such as burrow excavation have been used to calibrate imperfect methods of
Accurately representing uncertainty in our estimates of species population densities is of central importance to effective and sustainable management (Ludwig et al. 1993). This is the main advantage of our proposed method over other methods used to estimate densities of cavity-dwelling species. While these other methods aim to account for detection bias through application of correction factors, they do not calculate the uncertainty caused by imperfect detection (James and Robertson 1985, Warham 1996, Ratcliffe et al. 1998, Gusset and Burgener 2005, Scott et al. 2009, Oppel et al. 2014). Moreover, they do not account for possible spatial heterogeneity in the probabilities of detection.

Conway et al. (2008), for example, showed that the possible spatial heterogeneity in the probabilities of detecting species depends on factors such as ambient temperature. In our observation and occurrence in our estimation of animal abundances. While these other methods aim to account for detection bias through application of correction factors, they do not calculate the uncertainty caused by imperfect detection (James and Robertson 1985, Warham 1996, Ratcliffe et al. 1998, Gusset and Burgener 2005, Scott et al. 2009, Oppel et al. 2014). Moreover, they do not account for possible spatial heterogeneity in the probabilities of detection.

Accurately representing uncertainty in our estimates of species population densities is of central importance to effective and sustainable management (Ludwig et al. 1993). This is the main advantage of our proposed method over other methods used to estimate densities of cavity-dwelling species. While these other methods aim to account for detection bias through application of correction factors, they do not calculate the uncertainty caused by imperfect detection (James and Robertson 1985, Warham 1996, Ratcliffe et al. 1998, Gusset and Burgener 2005, Scott et al. 2009, Oppel et al. 2014). Moreover, they do not account for possible spatial heterogeneity in the probabilities of detection.

Acknowledgements – We would like to thank the staff and volunteers of Aride Island for assistance in data collection, and Vincent Bretagnolle and Swanne Gordon for comments and discussion. ALS would like to thank M. Betts for an inspiring introduction to call-playback census. This research was approved by the Seychelles Bureau of Standards (Research Permit no. A0157). Funding was provided by the Island Conservation Society, the CNRS, Fondation TOTAL (to LC and GR), the École Normale Supérieure de Paris (to ASBL) the Agence Nationale de la Recherche’s project EvoRANGE (to ALS), and the Direcció General de la Recerca de Catalunya (2009-SGR-481, to ALS).

References


Bolker, B. and R Development Core Team 2014. bbmle: tools for general maximum likelihood estimation. – R package.


Pombo, M. and Turra, A. 2013. Issues to be considered in counting burrows as a measure of Atlantic ghost crab populations, an important biindicator of sandy beaches. – PLoS One 8: e83792.


