



PARKER CONSERVATION
CONSERVATION, TRANSLOCATIONS, RESTORATION, RESEARCH, MANAGEMENT

Literature review of methods for estimating population size of burrowing petrels based on extrapolations from surveys

Graham C. Parker & Kalinka Rexer-Huber

Department of Conservation, Conservation Services Programme Objective 7, Project POP2014-02



Literature review of methods for estimating population size of burrowing petrels based on extrapolations from surveys

Department of Conservation, Conservation Services Programme
Objective 7, Project POP2014-02

Graham C. Parker^{1*} & Kalinka Rexer-Huber^{1,2}

¹ Parker Conservation, 126 Maryhill Tce, Dunedin New Zealand

² University of Otago, PO Box 56, Dunedin New Zealand

*Corresponding author: g.parker@parkerconservation.co.nz

Please cite as:

Parker, G.C. and Rexer-Huber, K. 2015. DRAFT Literature review of methods for estimating population size of burrowing petrels based on extrapolations from surveys. Unpublished DRAFT report to the Conservation Services Programme, Department of Conservation. Parker Conservation, Dunedin

Contents

| | |
|---|----|
| 1. Executive summary | 4 |
| 2. Introduction | 5 |
| Aims and scope | 5 |
| 3. Methods | 6 |
| 3.1 Literature search | 6 |
| 3.2 Focal definitions | 7 |
| 3.3 Literature analysis and summary | 8 |
| 4. Results and discussion | 9 |
| 4.1 Burrow contents | 9 |
| 4.1.1 Burrow occupancy | 9 |
| 4.1.2 Occupant detection | 10 |
| 4.2 Timing | 12 |
| 4.3 Availability bias | 15 |
| 4.3.1 Habitat availability | 15 |
| 4.3.2 Representativeness | 15 |
| 4.4 Burrow detection | 16 |
| 4.4.1 Burrow detection probability | 16 |
| 4.4.2 Extrapolation area | 18 |
| 4.5 Observer bias | 18 |
| 4.6 Summary | 19 |
| 5. Recommendations | 22 |
| 6. Acknowledgements | 23 |
| 7. References | 24 |

1. Executive summary

Robust population estimates are needed for conservation management of burrowing petrel populations. Estimates of population size for burrowing petrels are often obtained by extrapolation of burrow surveys to a population- or island-wide scale. However, extrapolation will also extrapolate bias or error, giving rise to potentially large error bounds reflecting imprecise estimates of population size. This hinders species risk assessment and limits the ability to detect trends in population size over time. We review methods for estimating the breeding population size of burrowing petrels by extrapolation from surveys, focusing in particular on the error associated with population estimates of the larger *Procellaria* petrels.

Sources of error in extrapolation of survey data are divided into five key areas: (1) uncertainty of burrow contents, (2) timing, (3) burrow detection probability, (4) availability bias and (5) observer bias. We reviewed 87 relevant studies. Of these, 45 published and unpublished studies deal specifically with quantitative surveys of burrowing petrels.

The review highlights that there is no single-best method for minimising error levels in population estimates. Rather, the most accurate and precise studies are those designed according to the specifics of the study resources, species and site, and we discuss a range of the factors that are important to consider.

To produce an accurate and precise population estimate from burrow counts, it is important to determine burrow contents and to distinguish between breeding and non-breeding birds. If a proportion of occupants are missed, further error is introduced to the population estimate, so it can be valuable to check occupant detection probability. The timing of burrow occupancy checks can help avoid assumptions about what proportion of breeding birds has not yet laid or has already failed. Extrapolation errors occur when the area sampled is not representative of the area that the samples are extrapolated to. If sampling sites are not representative, or if some part of a petrel's burrowing range is not accessible/available to sample, this availability bias can affect extrapolation. Burrow detection rates can also affect the accuracy of extrapolation, so the assumption that every burrow in the sampled area was detected should be checked. Whether planar map area or true surface area is used for extrapolation can be a further source of error. Observers may differ in their ability to detect burrows or burrow contents and this observer bias should be tested for.

Several key points are relevant to all studies: the need for a good pilot study to minimise error sources in the main survey; the need for sufficient time to cover enough ground while including contingency for weather; and the need to document burrowing petrel survey methods in enough depth to be repeatable.

Keywords: burrow breeding, seabird, petrel, Procellariiformes, population size, error

2. Introduction

Seabirds are a critical component of marine and many terrestrial ecosystems (Mulder et al. 2009), yet seabirds are one of the most-threatened bird groups in the world, with nearly half of all species known or suspected to be declining (Croxall et al. 2012). Burrow-nesting petrels are a broad group of seabird species spanning the globe. By nesting underground, burrow nesting petrels protect themselves from predators and achieve stable nest temperatures (Brooke 2004).

Some species of burrowing petrel suffer high rates of incidental mortality in fisheries (Bartle 1990; Waugh et al. 2008; Barbraud et al. 2009; Richard and Abraham 2013) and many have greatly reduced populations due to introduced mammals at their breeding grounds (Zotier 1990; Schulz et al. 2005; Martin et al. 2009; Bourgeois et al. 2013; Cuthbert et al. 2013).

For the conservation management of burrowing petrel populations, robust data is required to estimate demographic parameters and understand population dynamics. However, since burrowing petrels nest in underground burrows and cavities, frequently on remote islands that are hard to access, obtaining quality data is often much more difficult than for surface nesting seabirds (Brooke 2004; Schumann et al. 2013).

Historically, subjective population estimates were often made from brief visits to colonies (Bailey and Sorenson 1962; Taylor 1988), but this approach has largely been replaced with attempts to obtain quantitative data to estimate population size (Rayner et al. 2007a; Ryan et al. 2012; Bell et al. 2013a; Whitehead et al. 2014). Estimates of population size for burrowing petrels are often obtained by extrapolation of burrow surveys to a population- or island-wide scale. This typically involves surveying active nests, obtaining a representative sample of burrow density via transects or plots, and applying that density estimate, corrected for burrow occupancy, to the available area of nesting habitat (e.g. Burger and Lawrence 2001; Lawton et al. 2006; Baker et al. 2008; Lavers 2015).

However, extrapolation will also extrapolate bias or error, giving rise to potentially large error bounds and resulting in imprecise estimates of population size. Wide error margins also hinder species risk assessment and limit the ability to detect trends in population size over time (e.g. Oppel et al. 2014). To illustrate, trends may not be detected if population estimates have a precision of $\pm 25\%$ or more, unless the changes in breeding numbers are large, because the error margins of estimates will most likely overlap (R. Phillips pers. com.).

To reduce the error margins around population estimates, it is vital to consider potential sources of error and develop methodologies to lessen their effects on estimates. There are three general categories of error in estimates of population size; spatial variability, temporal variability and detection probability (Wolfaardt and Phillips 2011).

Aims and scope

The purpose of this project is to review methods used to estimate the breeding population size of burrowing petrels by extrapolation from surveys. Surveys are generally conducted along transects (e.g. Barbraud et al. 2009; Waugh et al. 2015), in quadrats (e.g. Rexer-Huber et al. 2014), and in plots or blocks (e.g. Burger and Lawrence 2001; Catty et al. 2003; Bell et al.

2013a; Whitehead et al. 2014). We will make recommendations, illustrated by examples, of methods by which future surveys can minimise error levels in population estimates.

Obtaining a population estimate is distinct from monitoring population trends. Methodologies need to be tailored according to whether a population estimate or population monitoring is the goal (Hunter et al. 2001; Rodway and Lemon 2011). Although these two questions share many of the same methods, this review focuses specifically on the methodology used in studies aiming to produce population estimates.

Within burrowing petrels, we focus on the error associated with population estimates of the larger *Procellaria* petrels. Four of the five species of *Procellaria* petrels breed in the New Zealand region and all are a high priority to New Zealand Department of Conservation's (DOC) Conservation Services Programme (CSP) since all are incidentally caught in commercial fisheries (Richard and Abraham 2013). Black petrels *Procellaira parkinsoni* breed only on Little Barrier / Hauturu and Great Barrier / Aotea Islands, and Westland petrel *P. westlandica* breed in a small area on the west coast of the South Island. In the NZ region, white-chinned petrel *P. aequinoctialis* breed on Antipodes and Auckland Islands, and grey petrel *P. cinerea* breed on Antipodes and Campbell Islands. The fifth *Procellaria* species is the spectacled petrel *P. conspicillata*, which breeds only on Inaccessible Island in the South Atlantic Ocean.

The literature on *Procellaria* petrels is not extensive. For many populations of *Procellaria* petrels only a single quantitative population estimation has been conducted and for others no quantitative population estimates exist. For this reason we also review research on other petrels and shearwaters where there is direct overlap in methodologies.

3. Methods

3.1 Literature search

To locate relevant literature we combined keywords and search terms as follows: `procellari* AND seabird OR petrel OR shearwater with population AND (status OR census OR estimate OR number OR abundance) AND (error OR variance OR precision) with burrowing`. We searched for published peer-reviewed literature in the database Web of Science, accessing BioOne, Biological Abstracts and Wiley Online. We searched for reference books, theses, governmental and non-governmental reports via Te Puna Search, which accesses Australian National Bibliographic Database and OCLC WorldCat along with New Zealand catalogues, and looked for relevant science publications with Ketu Search and the DOC Science online catalogue. We also searched for unpublished literature on population estimates of burrowing petrels via web-search engines Google and Google Scholar. Lastly, relevant experts were contacted for details on unpublished or pending work on burrowing seabird population estimation.

3.2 Focal definitions

Published and unpublished population estimates (**studies**) were reviewed taking into account the following known sources of error and assessing whether other sources were involved. To identify, and where possible quantify, the sources of error in extrapolation of survey data we divide the three general sources of error into five key areas: (1) uncertainty of burrow contents, (2) timing, (3) burrow detection probability, (4) availability bias and (5) observer bias. These and other key terms are defined below. Studies were scored in an attempt to prioritise the sources of error, from most to least influential on population estimates.

Uncertainty of burrow contents. The **burrow occupancy rate**, or proportion of burrows that contain the seabird species of interest, is a key element of a population estimate. To estimate the breeding population, burrow occupancy must also distinguish burrows occupied by breeding birds versus those occupied by failed or non-breeding birds. We ranked the proportion of burrow contents accessed accurately (Table 1). The rate of **false-negative errors** (that is, the number of burrows where an occupant was missed, expressed as a rate) can have a large influence on the population estimate (Tyre et al. 2003). Therefore studies were scored on whether or not they assessed **occupant detection probability**, the probability of correctly determining burrow contents.

Timing. Here, **timing** refers to the period the survey is undertaken relative to the breeding phenology of the target species. Burrow occupancy is particularly sensitive to survey timing (Schumann et al. 2013). Surveys after the early to mid-incubation stage may underestimate breeding attempts as breeding failures prior to surveys are missed (Hunter et al. 2001). We therefore scored each study based on when they assessed burrow occupancy (Table 1).

Availability bias. **Habitat availability**, or the proportion of petrel habitat available for sampling, can be limited by terrain and vegetation; for example, when species reside on cliff ledges or in fragile environments (Baker et al. 2010; Oppel et al. 2014). We scored the percentage of petrel habitat available for sampling based on descriptions in methods and discussion. Since **burrow density** (the number of burrows per unit area) and occupancy are rarely consistent among colonies or areas (Barbraud et al. 2009; Bell et al. 2009; Whitehead et al. 2014), it is important that sampling is representative. **Representativeness** refers to the extent to which sampled areas are representative of the habitat that a survey is extrapolated to. Studies were assessed on whether sampling was representative or not (Table 1).

Burrow detection probability. **Burrow detection probability**, the number of burrows counted as a proportion of the number available, is a measure of the ease with which burrows are found (**detectability**). Detectability varies with factors like vegetation, topography, and entrance size (Rayner et al. 2007b), but if burrows are missed during counts, extrapolated burrow numbers will be underestimated. We scored whether burrow detection probability was assessed (Table 1).

Observer bias. **Observer bias** from differences among observers can influence survey results (MacKenzie et al. 2006), so testing for differences among workers may help produce standardised and replicable results. Each study was scored by whether they tested for observer bias or not (Table 1).

Studies were assigned scores for each category of error, with the highest score in each category given when studies did not state if that source of error was accounted for (**not**

stated) (Table 1). To prevent sources of error with five categories being weighted more than those with three categories, scoring was standardised so that three categories were scored as 1, 3, 5 and five categories 1, 2, 3, 4, 5.

Table 1. Error classes and scores used to assess influence of error sources on breeding population estimates for burrowing seabirds.

| SOURCE OF ERROR | SCORE | DEFINITION | SECTION |
|---|-------|--|---------|
| Burrow contents: Accessibility | | <i>Proportion of burrow contents accessible</i> | 4.1.1 |
| | 1 | 90-100% | |
| | 2 | 70-90% | |
| | 3 | 50-70% | |
| | 4 | 0-50% | |
| Burrow contents: Occupant detection | 5 | not stated | 4.1.2 |
| | | <i>Probability of correctly determining burrow contents tested</i> | |
| | 1 | yes | |
| | 3 | no | |
| | 5 | not stated | |
| Timing | | <i>Study conducted during:</i> | 4.2 |
| | 1 | main laying period | |
| | 2 | early lay, before main laying period | |
| | 3 | after main laying period | |
| | 4 | non-breeding season | |
| Availability bias: Habitat availability | | <i>Proportion of petrel habitat accessed for sampling</i> | 4.3.1 |
| | 1 | 100% | |
| | 2 | 80-100% | |
| | 3 | 60- 80% | |
| | 4 | 40-60% | |
| Availability bias: Representativeness | 5 | not stated | 4.3.2 |
| | | <i>Sample representative, stratification used?</i> | |
| | 1 | representative and/or stratified | |
| | 3 | not representative and/or no stratification | |
| | 5 | not stated | |
| Burrow detection probability | | <i>Has the probability of encountering burrows been tested?</i> | 4.4 |
| | 1 | yes | |
| | 3 | no | |
| Observer bias | 5 | not stated | 4.5 |
| | | <i>Observer bias tested?</i> | |
| | 1 | yes | |
| | 3 | no | |
| | 5 | not stated | |

3.3 Literature analysis and summary

We summed scores (Table 1) for each population estimate (study); studies with low final scores were those that managed the most sources of error, whereas higher scores indicated studies where least error was accounted for. In order to relate sources of error to the precision of population estimates, precision must also be comparable among studies, so measures of precision (i.e. lower and upper 95% confidence intervals) were scaled for each estimate by dividing through the population estimate (Walther and Moore 2005).

4. Results and discussion

In total, 87 relevant studies were reviewed, 79 of which were published (peer-reviewed, reports, books) and eight unpublished (reports, theses, workshop proceedings, unpublished data). Of these, 45 published and unpublished studies deal specifically with quantitative surveys of burrowing petrels.

The literature on burrowing petrel survey extrapolation is discussed in the following sections, and summarised according to sources of error in Table 2.

4.1 Burrow contents

4.1.1 Burrow occupancy

To produce an accurate and precise population estimate from burrow counts, it is important to accurately determine the proportion of burrows that contain the seabird species of interest (burrow occupancy rate), distinguishing breeding birds from non-breeding or failed birds. Burrow occupancy is typically assessed from a sub-sample of burrows, with few exceptions (Barbraud et al. 2009; Rexer-Huber et al. 2014). The most widely used tools to determine burrow occupancy are the burrowscope (54% of studies reviewed e.g. Lavers 2015; Waugh et al. 2015), call-playback response (24%, e.g. Barbraud et al. 2009; Soanes et al. 2012) or feeling for an occupant by hand/with a probe (**grubbing**) (24%, e.g. Schulz et al. 2005). Some studies use inspection hatches (e.g. Waugh et al. 2006; Bell et al. 2009), and many use a combination of techniques like burrowscoping with excavation (Lawton et al. 2006), playback with grubbing (Burger and Lawrence 2001; Ryan and Ronconi 2011; Dilley et al. submitted) or grubbing and digital cameras (Baker et al. 2008). Most studies use direct inspection methods like these that rely on detection of bird(s) in a burrow, but indirect methods (e.g. using activity sign at burrow entrances) are still occasionally combined with other methods (Ryan et al. 2012; Schumann et al. 2013). Each method has limitations, and these limitations drive the choice of which is most appropriate for a particular species.

Response to playback of calls appears to be useful for some species, and minimally so for others. For example, a high proportion of white-chinned petrels responded to playback (Berrow 2000; Barbraud et al. 2008; Barbraud et al. 2009), but less than a quarter of Atlantic petrels *Pterodroma incerta* responded to playback (Rexer-Huber et al. 2014). Response rates can vary within species (18%–70%; Ratcliffe et al. 1998; Vaughan and Gibbons 1998), over a breeding season (Ryan et al. 2006), and even within individuals (Berrow 2000). The probability of a bird responding is influenced by factors including breeding condition, sex of bird in burrow, time of day, and playback features like the range of calls used, duration and volume (Berrow 2000; Burger and Lawrence 2001; Barbraud and Delord 2006; Soanes et al. 2012). However, even when playback reliably indicates occupancy, the presence of non-breeding birds can obscure true breeding numbers. Since population estimates typically hinge on knowing whether birds present in burrows are breeding, methods that visually inspect the contents of nesting chambers, like a burrowscope or inspection hatches, are critical.

Burrowscopes work particularly well for species with relatively large, straight burrows like those of Westland petrels (Waugh et al. 2003) and white-chinned petrels (Rexer-Huber et al. 2015). Burrows can be inspected in their entirety with confidence that no branches or

chambers were missed, and the small number not fully inspected reported as inaccessible and removed from calculations (e.g. Cuthbert 2001; Rexer-Huber et al. 2014). Complex burrows are more difficult to inspect fully by burrowscope; for example, some species have entrance moats, multiple entrances and/or chambers, collapsed chambers or entrances, and deviating tunnels (Hamilton 2000; McKechnie et al. 2007; Newman et al. 2008; Cuthbert et al. 2013; Dilley et al. submitted). Moats can be spanned by adding a firm but bendable splint to the burrowscope cable, but highly curved, tight or deviating burrows can be harder to inspect in full. For example, burrow occupancy was underestimated by up to 17% when a burrowscope was used to check the typically long, convoluted burrows of sooty shearwaters (McKechnie et al. 2007), and by 10–19% in Hutton's shearwater burrows (Cuthbert 2001; Cuthbert and Davis 2002).

For some species, burrows can be accessed by hand / with a flexible probe (burrow grubbing) (e.g. Schulz et al. 2005; Baker et al. 2008; Ryan and Ronconi 2011; Schumann et al. 2013), especially when progressive access holes are dug to allow full inspection of the nesting chamber (G. Taylor pers. com.). This technique is widely used for its ease and minimal equipment needs, and has no ambiguity about burrow contents when access is by hand (e.g. Barbraud and Delord 2006). However, grubbing with a probe is less accurate, particularly during incubation, and access to deeper or more convoluted burrows is often not possible by grubbing.

A range of other techniques are in use. Access hatches dug into or closer to the nesting chamber are sometimes used to assess burrow occupancy; typically, these are long-term study burrows established in successful, occupied burrows (e.g. Cuthbert 2001; Rayner et al. 2008; Bell et al. 2009), which may introduce positive bias in occupancy (Hunter et al. 2001). Digging access holes may be disruptive to an unknown extent (Hamilton 1998; Ryan et al. 2006), depending on the species and the timing. Some burrowing seabird species are more prone to deserting their nests when disturbed during early incubation (Warham 1990), although there are no records of desertion as a result of study during incubation for *Procellaria* species.

Photography of burrow contents is occasionally used (Baker et al. 2008). Dogs trained in detecting seabirds are sometimes used to great effect for both burrow surveys and detecting occupancy (Bell et al. 2013b; G.B. Baker pers. com.), but a well-trained dog and its handler are costly. Like people, the detection ability of dogs differs so the accuracy of a given dog must be calibrated (G.B. Baker pers. com.). Other burrow occupancy techniques being tested are the use of heat sensors (E. Bell pers. com.) and CO₂ sensors (S. Waugh pers. com.).

4.1.2 Occupant detection

Burrow occupancy may vary on both spatial and temporal scales (Berrow 2000; Lawton et al. 2006; McKechnie et al. 2009; Baker et al. 2010; Whitehead et al. 2014). Spatial variation can be minimised by using representative sites for sampling burrow occupancy (e.g. Sutherland and Dann 2012; Whitehead et al. 2014). Occupancy rates will best reflect the maximum number of breeding pairs early in the breeding season, just after laying has finished (Section 4.2 Timing). However, temporal variation in occupancy rates may become more of a factor when it is not possible to be present at that point (Sutherland and Dann 2012).

It is important to consider the assumption that occupants are always detected. False-negative errors, arising from the failure to record an occupant when it is present, can potentially occur

under any of the burrow occupancy methods discussed here. For some species like sooty shearwaters, inaccuracy from failure to detect some birds was greater than between-observers and between-sites sources of error (McKechnie et al. 2007). There is a large literature on response rate correction for call playback estimates of burrow occupancy (e.g. Burger and Lawrence 2001; Barbraud and Delord 2006; Ryan et al. 2006; Soanes et al. 2012; Rexer-Huber et al. 2014), and playback is rarely used without accounting for response rate. However, the probability of detecting burrow occupants (occupant detection probability) is less frequently checked in studies applying other methods; for example, 48% of studies using burrowscopes checked for occupant detection rates. Occupant detection is assessed in a smaller proportion of *Procellaria* studies (23%) than in those on other burrowing petrel species (59%) (Fig. 1). This may be due to the large, simple burrows dug by some *Procellaria* relative to those of smaller burrowing petrel species, which ease wall-to-wall inspection by burrowscope (Waugh et al. 2003). With smaller and more complex burrows, physical limitations are more readily apparent, which may explain why their impact on occupant detection is more frequently accounted for.

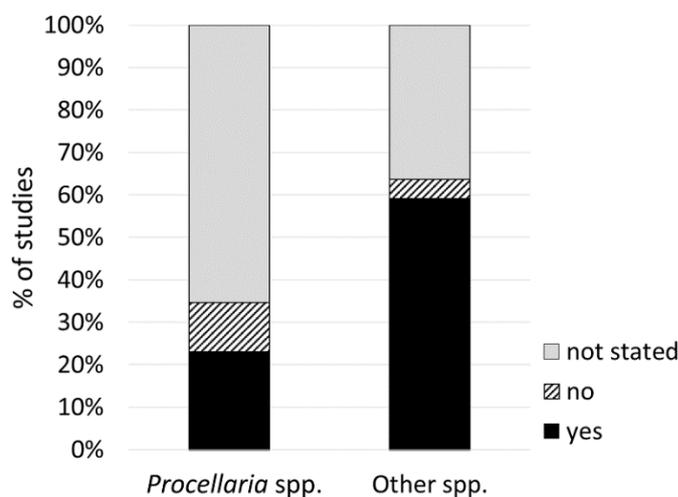


Figure 1. Occupant detection: proportion of studies on *Procellaria* (n=26 studies) and other burrowing petrel species (n=22) that account for the probability of detecting an occupant in burrows.

Occupant detection probability is most commonly assessed by variants of resampling methods: two observers checking the same sub-sample of burrows for occupants (e.g. Whitehead et al. 2014), or via repeated checks of the same burrows over a time period (MacKenzie et al. 2006; Ryan et al. 2006; Bailey et al. 2007; Baker et al. 2008; Rayner et al. 2009). Occupant or playback detection can also be assessed against a different inspection method, such as burrowscoping (Barbraud et al. 2008; Barbraud et al. 2009), grubbing (Ryan et al. 2012) or excavation / study hatches (Barbraud and Delord 2006; McKechnie et al. 2007; Baker et al. 2008; Baker et al. 2010). Detection is generally calculated simply as the probability of detecting an occupant, but may be estimated via a form of capture-mark-recapture (CMR) modelling where burrows represent ‘individuals’ and the presence of an occupant a ‘capture’ (Sutherland and Dann 2012; Whitehead et al. 2014).

On the other hand, **false-positive errors** arise when a burrow is incorrectly classed as occupied. This is minimised by the use of direct inspection tools rather than indirect activity-sign based methods (Rexer-Huber et al. 2014). False-positives might occur during direct inspection due to mistaken identity, i.e. when multiple species occur at a site (Rayner et al.

2008; Cuthbert et al. 2013; Whitehead et al. 2014), or when non-breeding birds are included in breeding bird totals.

False-positive errors are thought to be less prevalent than false-negative errors (Tyre et al. 2003), mainly because mistaken identity or breeding status should rapidly become rare with even small amounts of observer experience and training.

4.2 Timing

Strictly speaking, temporal variability is a process resulting from variation in demographic parameters, not an error source, and can be detected using estimates of population size at different times (C. Barbraud pers. com.). However, when a burrowing petrel population is infrequently surveyed, timing should be treated as an important source of error in estimating population size and accounted for in study design. Timing has particular relevance for determining burrow contents, while burrow entrance counts are less sensitive to this issue (Ryan et al. 2012). Burrow occupancy data collected at the most appropriate time in the breeding season can help avoid assumptions about what proportion of breeding birds has not yet laid or has already failed. For burrowing seabirds, the contents of burrows is ideally assessed immediately following the **main laying period**, when all birds except perhaps a few very late birds have laid and begun incubating (Lawton et al. 2006; Schumann et al. 2013; Defos du Rau et al. 2015). Surveying too early in the egg-laying period will miss a proportion of breeders that have not yet attempted to breed and thus underestimate the breeding population (Sutherland and Dann 2012). Monitoring burrows after early incubation may also underestimate the year's breeding attempts, since failures prior to survey may remain undetected (Hunter et al. 2001). There may also be fewer non-breeders occupying burrows after egg-laying is complete than before and during egg-laying (B. Dilley, pers. com.), helping to minimise false-positive errors.

Of the 48 studies reviewed, 31% collected data during the main laying period and 58% during incubation, after the main lay period (Table 2). When only studies of *Procellaria* petrels were examined, 42% collected data during main lay and a further 42% during incubation (Fig. 2), mainly in mid-incubation (Catry et al. 2003; Waugh et al. 2003; Barbraud et al. 2009; Rexer-Huber et al. 2014). Few *Procellaria* studies conduct surveys in the early part of the laying period (4%) but 21% of studies on other species survey prior to the main lay period (Fig. 2). A small proportion of studies collected data outside of the breeding season; 11.5% of *Procellaria* studies, but none of the studies on other species (Fig. 2). Some studies accounted for conducting burrow surveys outside of the breeding season. For example, Ryan et al. (2012) conducted an island-wide survey for white-chinned petrel burrow entrances at the end of the breeding season as more workers were available during this time, but collected occupancy data during the main laying period. Surveys in Wood and Otley (2013) were carried out over a four year period but occupancy was assessed during early incubation. Demographic data from previous work can help determine egg loss that occurred during incubation (e.g. Granadeiro et al. 2006).

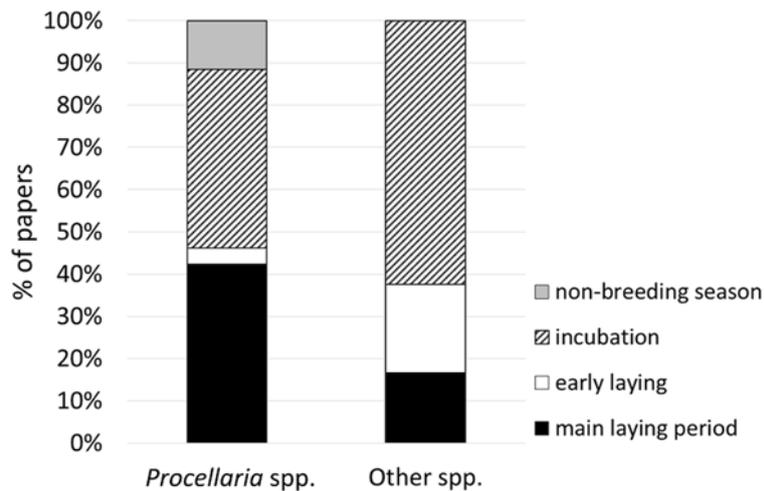


Figure 2. Timing in breeding season of research for population studies of *Procellaria* (n=26 studies) and other (n=24) burrowing petrel species.

Frequently, data on the timing of the laying period are not available. When demographic studies only exist for some populations, it may be possible to make inferences from other populations. For example, lay dates for white-chinned petrels from the well-studied population on Bird Island, South Georgia (Hall 1987; Berrow 2000) were used to guide survey planning for the Auckland Island population (Rexer-Huber et al. 2015). Laying data exist for most *Procellaria* petrels (Zotier 1990; Waugh et al. 2006; Barbraud et al. 2008; Bell et al. 2013b). Ideally a pilot study aimed at estimating the main-laying period should be conducted the year prior to burrow surveys. For species where this is not possible and there has not been previous demographic study, we recommend using all available information, for example anecdotal records and descriptions from workers with experience at the location, to predict the likely main laying period.

For some species there may not be a clear laying period, it may vary among populations, or it may change from year to year. For example, the lay dates of some white-chinned petrel populations can vary, driven by late-melting snow and ice cover in some years preventing birds from accessing burrows (Berrow 2000). Grey petrels on Marion Island start laying two weeks earlier than grey petrels on Gough Island (B. Dilley pers. com.). The lay period of black petrels on Great Barrier Island also varies from year to year, and they lay over an extended period of six weeks (E. Bell pers.com). To account for this, researchers collect population estimate data during three visits to the study colony (Bell et al. 2009). An alternative method for species that have asynchronous laying is a capture-mark-recapture (CMR) study on a sample of burrows, collecting information on occupancy, and using these CMR data to estimate the total number of burrows occupied in a given year (Williams et al. 2011). When comparison of population estimates to investigate population trends is an objective of the study, consistency in the timing of surveys between years is important (Bell et al. 2009; Rexer-Huber et al. 2014).

Table 2. Sources of error in burrowing petrel survey extrapolation. Number of studies in each error-type category expressed as percentage of total number of studies reviewed (parentheses in column header).

| | All spp. (48) | <i>Procellaria</i> spp. (26) |
|---|------------------|---------------------------------|
| Burrow contents: Accessibility | | |
| 90-100% | 35.4 | 34.6 |
| 70-90% | 6.3 | 7.7 |
| 70-50% | 2.1 | 3.8 |
| 0-50% | 2.1 | 3.8 |
| not stated | 54.2 | 50.0 |
| Burrow contents: Occupant detection | | |
| yes | 39.6 | 23.1 |
| no | 8.3 | 11.5 |
| not stated | 52.1 | 65.4 |
| Timing | | |
| main laying period | 31.3 | 42.3 |
| early lay, before main lay period | 4.2 | 3.8 |
| after main lay period | 58.3 | 42.3 |
| non-breeding season | 6.3 | 11.5 |
| Availability bias: Habitat availability | | |
| 100% | 27.1 | 34.6 |
| 80-100% | 29.2 | 26.9 |
| 60- 80% | 4.2 | 0.0 |
| 40-60% | 6.3 | 7.7 |
| not stated | 33.3 | 30.8 |
| Availability bias: Representativeness | | |
| representative, stratified | 69.4 | 50.0 |
| not representative &/or not stratified | 25.0 | 37.5 |
| not stated | 5.6 | 12.5 |
| 70-90% | 6.3 | 7.7 |
| 70-50% | 2.1 | 3.8 |
| 0-50% | 2.1 | 3.8 |
| not stated | 54.2 | 50.0 |
| Burrow detection probability | | |
| yes | 27.1 | 38.5 |
| no | 4.2 | 3.8 |
| not stated | 68.8 | 57.7 |
| Observer bias | | |
| yes | 25.0 | 38.5 |
| no | 18.8 | 34.6 |
| not stated | 56.3 | 26.9 |

Other aspects of timing are important to consider; for example, whether investigator disturbance is greater at a particular stage in the breeding season. This is often discussed in the context of shearwaters (Warham 1990), or focuses on the impacts of handling (Carey 2009), but to our knowledge the effects of burrow inspection that does not involve handling (burrowscope, access hatch, acoustic playback) have not been recorded for *Procellaria* petrels. Weather may also be important: the number of non-breeding birds can fluctuate greatly from one day to the next, depending on weather- or food-driven attendance cycles (R. Phillips pers. com.). Ryan et al. (2006) ascribed an increase in burrow occupancy to windy and misty conditions causing an influx of non-breeding and prospecting birds. This illustrates the advantage of collecting occupancy data over a number of days in order to identify and sample such phenomena.

4.3 Availability bias

4.3.1 Habitat availability

An availability bias results from some part of a petrel's burrowing range not being accessible, or available, to sample (Wolfaardt and Phillips 2011), creating problems for extrapolation (see section 4.3.2 below). Despite this, relatively few studies report the proportion of habitat that could not be surveyed (but see Catry et al. 2003; Barbraud and Delord 2006).

Burrowing petrels frequently inhabit remote, mountainous and steep terrain, areas with difficult vegetation or areas that are simply too fragile to access (Granadeiro et al. 2006; Baker et al. 2008; Oppel et al. 2014; Lavers 2015). However, if that terrain is not available for sampling, a sizeable proportion of the population may go unaccounted for. For example, an unknown proportion of grey petrels on Antipodes Island nest on cliffs, so Bell et al. (2013a) used dawn, day and dusk observations from vantage points to obtain presence / absence data. Spotlighting or acoustic surveys are also sometimes used to determine whether and/or which species are utilising inaccessible sites (Bell 2002; G. Taylor pers.com). It may be possible to calibrate vantage point, spotlighting or acoustic surveys against areas where quantitative data were collected (Oppel et al. 2014). If an important proportion of the population inhabits difficult terrain, presence / absence surveys would ideally be followed by access to at least some areas to quantify burrow numbers and estimate occupancy. In the case of cliffs, appropriate rope-access equipment could be used to survey the area (e.g. Defos du Rau et al. 2015), providing nests are not damaged or destroyed.

Fragile ground can also affect the ability to re-sample areas. For example Ryan et al. (2006) assessed the accuracy of single compared to multiple burrow occupancy checks, but limited repeat visits to colonies because of the risk of collapsing burrows during each visit to a colony. In very fragile areas, boards or plywood squares to distribute weight during surveys, or access paths developed while breeding birds are absent from an island, are possible solutions (Baker et al. 2010; G. Taylor pers. com.). However, this is clearly not always possible or the impact / effort justifiable, and other tools like acoustic monitoring may be necessary (Oppel et al. 2014).

4.3.2 Representativeness

Extrapolation errors occur when the area sampled was not representative of the area that the samples are extrapolated to (Hunter 2001; Baker et al. 2010). Ensuring that sampling areas

are fully representative of the extrapolation area can be a particular challenge when working on patchily-distributed species. A truly randomised sampling design may be representative but not encounter any burrows, even when present, making it necessary to target burrowed areas for sampling (e.g. Rayner et al. 2008; Parker et al. 2015; E. Bell pers. com.). However, a targeted sample is clearly not representative of adjacent areas, so simple extrapolation would not provide meaningful numbers.

A number of approaches have been used for representative sampling of patchy distributions. If every colony can be identified, non-colony areas could be excluded from extrapolation. For example, in an area with good prior knowledge of colony distribution, colonies can be sampled (or if small enough, counted in their entirety) and the density extrapolated specifically to the colony area (Burger and Lawrence 2001; Baker et al. 2008; Rayner et al. 2008; Baker et al. 2010; Sutherland and Dann 2012). If burrows occur according to a pattern – perhaps at greater density above 300 m, or on flatter terrain with drier soil, or in a particular forest type – then a stratified design may reduce extrapolation error (Cuthbert 2004; Lawton et al. 2006; Rayner et al. 2007a; Charleton et al. 2009; Francis and Bell 2010; Bell et al. 2011), but not always (Schumann et al. 2013). Stratification can also be important if dealing with very large or spatially discontinuous geographical areas; for example, on Kerguelen or South Georgia (Barbraud et al. 2009; Martin et al. 2009). This ensures that samples collected within a particular area, or stratum, are extrapolated only to the area of that stratum, and those summed for the overall population estimate. However, useful patterns are not always available to guide stratification, or produce strata with few burrows and result in large error bounds. A systematic study design can be useful to improve precision in these circumstances (Buckland et al. 2001; Fewster et al. 2009). For example, a small island known to contain a dense cluster of burrows along one edge and few elsewhere could be sampled with systematic transect lines (the sampling unit) spanning the low- to high-density gradient (Fewster et al. 2009). More patchily-distributed species may require more closely-spaced systematic transects, increasing survey effort (Lormée et al. 2012; Schumann et al. 2013). Alternatively, an adaptive cluster sampling design (Thompson 1991) can be useful to capture relatively rare species whose burrows are distributed in patches (C. Barbraud, pers.com.). This is not an exhaustive list of study design options (see Borchers et al. 2002; Morrison et al. 2008 for thorough review); rather, we review the more commonly-used approaches in the burrowing petrel literature.

4.4 Burrow detection

Extrapolation errors commonly arise when burrow counts under- or overestimate actual burrow numbers, and when extrapolation utilised planar map areas rather than true landscape surface areas.

4.4.1 Burrow detection probability

Accurate extrapolation depends on the assumption that every burrow in the sampled area was detected. If a number of burrows are missed during counts, extrapolated burrow numbers will be inaccurate to an unknown extent. Correcting for missed burrows may increase the variance around an estimate, with the advantage that this variance is more likely to overlap the true population size (Sutherland and Dann 2012).

The ease with which burrows are detected, or detectability, may vary with factors like vegetation, topography and burrow size (Lawton et al. 2006; Ryan et al. 2006; Rayner et al. 2007b; Barbraud et al. 2009). Therefore it follows that the likelihood of missing some burrows also varies (Francis and Bell 2010). Sampling design may also have an influence, since it is easier to miss some burrows in large plots or wider transects, for example, than in smaller plots. Even direction of travel can affect detectability: burrows are typically easier to detect when moving upslope than down in some habitats (Lawton et al. 2006; Rexer-Huber et al. 2015), a factor to consider if, for example, transects are oriented at random with respect to slope.

Even a moderate, realistic false-negative rate (the converse of detection rate) can have a large effect on study conclusions (Tyre et al. 2003; Archaux et al. 2012), particularly on the effect of habitat on site occupancy (e.g. whether burrows are detected at that site or not). While most burrowing petrel studies reviewed apparently did not test burrow detection probability (either not tested or not reported; 73%) (Table 2), those that did found that a number of burrows were missed, across a wide range of survey design, species, and physical variables. Detection probability was quantified in a larger proportion of *Procellaria* studies (39%) than in studies on other species (14%), with 82% of studies on non- *Procellaria* species failing to record whether detection probability was checked or not (Fig. 3).

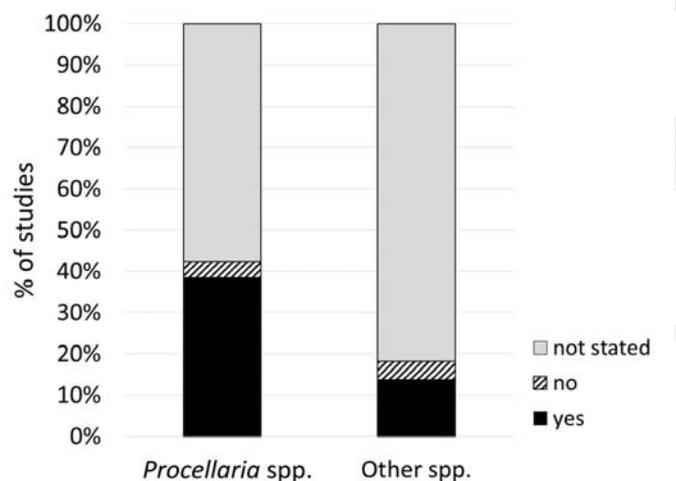


Figure 3. Burrow detection probability: proportion of studies on *Procellaria* (n=26 surveys) and other (n=22) burrowing petrels that account for burrow detection.

Detection probability is often derived from repeat surveys, with all sites surveyed several times over a visit, or season, or repeat surveys at a subset of sites (MacKenzie et al. 2006). A double-observer approach may be appropriate to determine burrow detection probability, where other observer/s revisit a subset of transects or plots (Lawton et al. 2006; Baker et al. 2008; Baker et al. 2010; Lormée et al. 2012; Defos du Rau et al. 2015). This assumes that the detection probability of the subset is the same at the other plots (MacKenzie et al. 2006), so care is required that this subsample is representative. The accuracy of sampling methods can be tested within small colonies or areas by exhaustive strip-search counts of a colony to determine the true burrow number (Barbraud et al. 2009; Ryan et al. 2012). In turn, exhaustive strip-search counts can be checked for detection probability via validation transects run in a different direction and angle to search strips (Parker et al. 2015). Some approaches to estimate burrow numbers explicitly take detection probability into account, like

capture-mark-recapture and distance sampling methods (Buckland et al. 2001; Morrison et al. 2008). Distance sampling is becoming more widely used (Lawton et al. 2006; Barbraud et al. 2009; Defos du Rau et al. 2015; Rexer-Huber et al. 2015). This uses the distance from a line to an object to account for objects becoming less detectable with distance, and incorporates this detection probability into estimates of burrow numbers (Buckland et al. 2001).

4.4.2 Extrapolation area

Another source of error when extrapolating samples of burrow counts is area; in particular, whether **planar area** (one-dimensional area as represented on a map) is used, rather than true **surface area** (three-dimensional surface area of a landscape). If counts within a given sampling area (plot diameter or transect width) on the landscape are extrapolated by the planar/map area of the study (Bell 2002; Sommer et al. 2010), true numbers may be underestimated (Jenness 2004). This is less of an issue when extrapolating counts based on planar sampling area to other planar areas; for example, if the sampled area is line length measured by GPS (Jamieson and Waugh 2015; Waugh et al. 2015; but see Baker et al. 2008). Calculating line length this way may however introduce measurement error from the GPS, which is rarely reported (but see Waugh et al. 2015).

The true landscape surface area was used for extrapolations in half (50%) of studies. A range of approaches were used to correct planar/map area to reflect true landscape surface area. Repeated slope measurements were used to estimate slope-corrected surface areas (Newman et al. 2008; Barbraud et al. 2009; Rexer-Huber et al. 2015); topographic contours used find the slope with which to correct planar area (Martin et al. 2009); and digitised contours, spot-height data or extensive GPS tracks are used to create digital elevation models in geographic information systems software (Catry et al. 2003; Lawton et al. 2006; Rayner et al. 2007b; Rayner et al. 2008; Defos du Rau et al. 2015).

4.5 Observer bias

Producing standardised and replicable results is a critical aspect of animal surveys, yet observers differ (MacKenzie et al. 2006). Observers may differ in their ability to detect burrows or burrow contents, to reach birds, or to estimate burrow numbers in a discrete area (Cuthbert 2001; Cuthbert and Davis 2002; Ryan and Ronconi 2011; Lormée et al. 2012), and this difference should be tested for (MacKenzie et al. 2006; Morrison et al. 2008). Observers with longer arms may have better bird detection rates in some burrow types, and observer experience can play a role in some methods to determine occupancy. Only 25% of all the studies we reviewed tested for differences in detection between workers (Table 2). More than half (56%) did not mention observer bias and 19% made it clear that observer bias was not assessed (Table 2). A higher proportion of *Procellaria* petrel studies tested for observer bias, with 39% testing, 27% not stated and 35% not testing for observer difference (Table 2).

The most common way to calibrate the reliability of observers is through the use of multiple counts (Morrison et al. 2008). For example, multiple counts were conducted along a proportion of transects for Westland petrels (Baker et al. 2008) and flesh-footed shearwaters *Puffinus carneipes* (Baker et al. 2010). Although neither study showed a difference among observers, bias should be tested for to increase confidence in the results. Other studies reviewed did show an effect of observer bias. Scan counts of the number of burrows in

discrete areas always showed differences among observers when checked against thorough ground counts (Ryan and Ronconi 2011). Counts of spectacled petrels differed by as much as 2–5% between observers in any given year (Ryan and Moloney 2000; Ryan et al. 2006; Ryan and Ronconi 2011), and observers of white-chinned petrels on Marion Island consistently under and over-estimated the number of burrows (81–123%) (Ryan et al. 2012).

Training is a logical approach to ensure that search effort is consistent across workers. It may also be important to provide measuring devices (e.g. tape, length of line, plastic pipe) to reduce measurement error (Morrison et al. 2008); for example, transect width or radius of a plot. When surveys are compared over years, having leaders with experience can help standardise methodologies between years (Ryan and Ronconi 2011; Bourgeois et al. 2013; Dilley et al. submitted). Furthermore, it may be useful for a single observer to do aspects of the work that may be more sensitive to observer error/experience, such as burrow inspection by burrowscope (Cuthbert 2001; Parker et al. 2015; S. Waugh pers. com.).

4.6 Summary

The degree to which studies accounted for sources of error in the study methods influenced the reported variance around population estimates (Fig. 4). Studies ranked according to total score, from most error sources accounted for to least, showed a tendency for wider confidence intervals (Fig. 4). Studies that did not report any indicator of variance around the estimate tended to have higher error scores, indicating that fewest error sources were accounted for (Fig. 4).

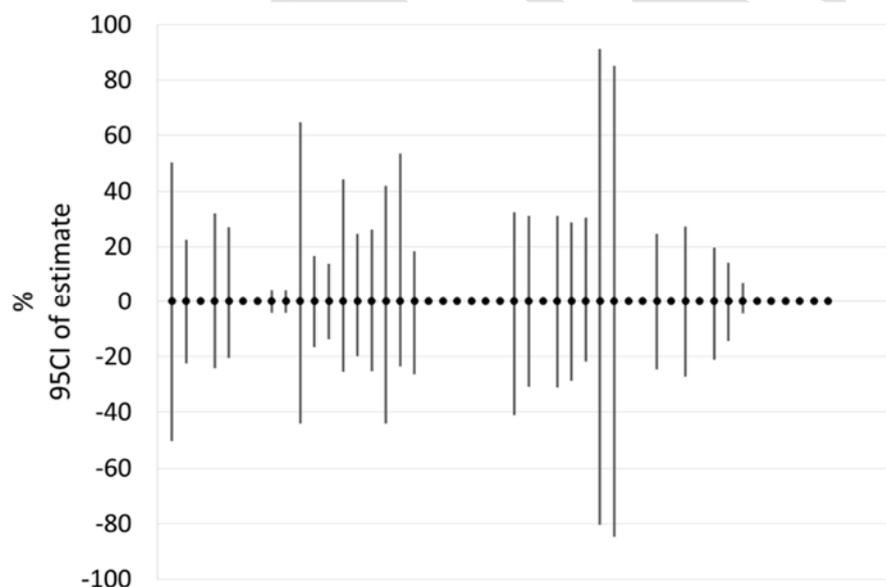


Figure 4. Impact of combined error sources on variance of burrowing petrel population estimates. Studies (points, $n=47$) are ranked along the x axis from lowest to highest; lowest score represents most sources of error accounted for; highest, least. Error bars show reported variance (lower and upper 95% confidence intervals scaled as percentage of estimate).

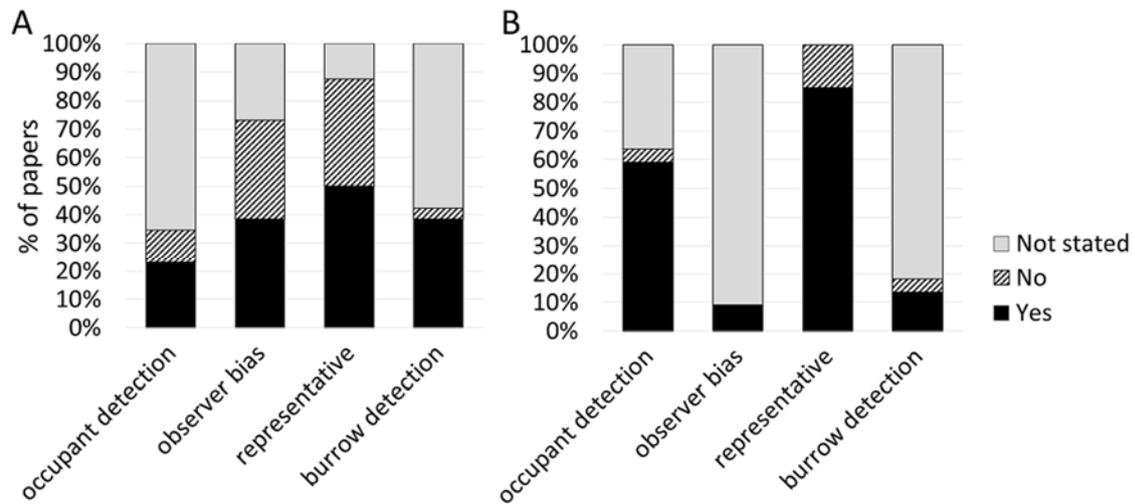


Figure 5. Source of error in burrowing petrel population estimates for A) *Procellaria* species (n=26 studies) and B) other large burrowing petrel species (n=22). Percentage of estimates by error category.

When error sources are viewed individually, some are more reliably accounted for than others. Overall, fewest *Procellaria* studies accounted for occupant detection (23%) and burrow detection (39%) probabilities (Fig 5). Burrow detection was also rarely addressed for other burrowing petrels (14%), while observer bias was not addressed in any of the studies on other species (Fig. 5). Surveys were more frequently timed around the peak lay period for *Procellaria* (42%) than for other burrowing petrels (18%) (Fig. 5).

Several recurring themes stand out in reviewing the literature for sources of error in population estimates: the need for a good pilot study to minimise error sources in the main survey (e.g. Defos du Rau et al. 2015); the need for sufficient time to cover enough ground while including contingency for weather (e.g. Baker et al. 2008), and the need to document burrowing petrel survey methods in enough depth to be repeatable (e.g. Waugh et al. 2013).

A surprisingly high proportion of studies do not state whether an error source was addressed or not (grey bars, Fig. 5), and few studies report not addressing an error source (striped bars, Fig. 5). This lack of survey methodology information hinders understanding of how population estimates were obtained, and makes it difficult for future surveys to repeat methods and analyse trends (Waugh et al. 2015). ‘Not stated’ was most frequent for occupant detection, observer bias and burrow detection, the three error sources that may have the largest impact on extrapolated population estimates. Clearly documenting methods and corrections used – and not used – is critical for interpretation and future repeatability of burrowing petrel surveys.

For reference, we summarise the key problems associated with extrapolation of burrowing petrel surveys, and the implications if these error sources are not addressed (Table 3).

Table 3. Error sources in extrapolation of burrowing petrel surveys to obtain population estimates: problems and implications if not addressed

| Factor | Problem | Implications if not addressed | Section discussed |
|-------------------------------------|---|--|--------------------------|
| Burrow contents | | | |
| <i>Accessibility</i> | Occupancy of burrows with unconfirmed contents may not be the same as burrows where the contents can be confirmed | Under or overestimate of burrow occupants in breeding population estimate | 4.1.1 |
| <i>Occupant detection</i> | Incorrectly assign occupied burrow as empty | Underestimate burrows occupied by breeding birds | 4.1.2 |
| | Non-breeding birds included in breeding bird totals | Include non-breeding burrow occupants in breeding population estimate | |
| | Mistaken identity | Include other burrowing seabird species in breeding population estimate | |
| Timing | Burrows not occupied by the majority of the season's breeding pairs | Under or over-estimate population size depending on breeding phenology | 4.2 |
| | Detection probability varies with time | Detection probability for some survey methods affected (response to taped playbacks, sign) | |
| | Some spp. may be more disturbance prone during incubation | Detrimental disturbance to study species | |
| | Inter-annual surveys conducted at different times of breeding season | Between year estimates not comparable | |
| Availability bias | | | |
| <i>Habitat availability</i> | Incorrectly assume birds are present or absent in the non-surveyed habitat | Decreased accuracy of population estimate | 4.3.1 |
| | Assume burrow density or occupancy consistent across habitat | Decreased accuracy and precision of estimate | |
| <i>Representativeness</i> | Sampled area not representative of the area samples are extrapolated to | Extrapolation error resulting in poor accuracy and precision | 4.3.2 |
| | Incorrect or no stratification of sampled areas when required | Estimate less precise | |
| Burrow detection probability | | | |
| | Burrows that are present are not detected | Underestimate burrow numbers | 4.4 |
| | | Increased variance due to correction factor applied for missed burrows | |
| Observer bias | | | |
| | Burrows not detected | Decreased precision due to difference in observers data | 4.5 |
| | Burrows of different species not accurately discerned | Decreased precision due to difference in observers data | |
| | Burrow occupants not accurately identified | Decreased precision due to difference in observers data | |
| | Distances (e.g. 1 m either side of a transect) estimated differently | Decreased precision due to difference in observers data | |

5. Recommendations

Burrowing petrel surveys are time-consuming, labour intensive and often temporally-limited by funding and logistics such as appropriate boating conditions (Schumann et al. 2013). It is of great interest to design surveys that will produce accurate and precise estimates while constrained by these limitations.

There is no single-best method for minimising error levels in population estimates; rather, the most accurate and precise studies are those designed according to the specifics of the study resources, species and site.

It is widely accepted that population estimates, usually expressed as numbers of breeding pairs, requires that counts of burrows be corrected by the proportion of burrows that contain a breeding pair (burrow occupancy rate). A range of tools are used to check burrow contents, but are not equally useful for all burrowing petrel species; that is, the accuracy and precision of a tool (e.g. burrowscope, or acoustic playback) varies according to species and breeding site. To reduce extrapolation errors, it is critical to test the assumption that every occupant will be found (occupant detection probability) and that breeders can be accurately discerned from non-breeders.

A good pilot study can be a valuable tool for reducing the variance around population estimates. A pilot study testing sampling techniques and sample sizes informed a Scopoli's shearwater *Calonectris diomedea* population estimate of good precision (Defos du Rau et al. 2015). Pilot studies which gauge the time required for good spatial coverage, including contingency for weather, can minimise the sample size and representativeness problems that most commonly result from having insufficient time. A pilot study may also allow the main laying period to be estimated for those species or populations for which it is unknown.

Planning the timing of surveys so that fewest assumptions and/or corrections are needed can also reduce error. This is illustrated by burrow occupancy: studies timed when the year's breeding pairs have all laid, but before many egg failures have occurred will best reflect the number of breeding pairs. A demographic study prior to survey can determine when laying occurs, but is also important when surveys around peak lay are not possible: the rate of failures prior to the survey can correct studies timed during incubation or the chick-rearing period.

Ensuring that sampled habitat is representative of the extrapolation area while minimising variance is one of the key challenges when addressing availability biases in population estimates. Careful study design is the best tool to address this bias. Depending on the species and site, representative sampling may require random plots or transects throughout an area (Burger and Lawrence 2001); random sampling within strata or subareas (stratified design, (Lawton et al. 2006; Rayner et al. 2007a); systematic sampling (Fewster et al. 2009; Whitehead et al. 2014), or adaptive cluster sampling (C. Barbraud pers. com.), to give a few examples. Habitat not available for sampling (cliffs, very fragile areas) introduces error if included in extrapolations, so it should be clearly documented how unavailable habitat was accounted for.

Burrow detection probability, testing the assumption that every burrow in an area is found, can have a large impact on how accurate a burrowing petrel estimate is, yet less than a third

of studies accounted for burrow detection. The appropriate method to determine burrow detection is, again, species and site dependent, but can involve simple repeat surveys, or checking counts against other sampling methods. Alternatively, methods that explicitly take detection probability into account (e.g. variants of capture-mark-recapture or distance sampling) may be best for some studies. Burrowing petrel estimates can also be influenced heavily by the area used for extrapolation. Seabird colonies with variable topography or elevation change may require an area correction to ensure that available surface area is not underestimated.

Observer bias may introduce unnecessary variance around a point estimate of population size. The assumption that well-trained and experienced observers will be consistent (Baker et al. 2008) does not always hold (Ryan and Ronconi 2011), so it is important to check burrow count and contents data for observer differences.

6. Acknowledgements

We thank Barry Baker, Biz Bell, Igor Debski, Jo Hiscock, Sarah Jamieson, Graeme Taylor, David Thompson and Susan Waugh for contributing expert knowledge to this document through participation in a workshop discussion. Christophe Barbraud, Ben Dilley, Richard Phillips, Graeme Taylor and Barbara Wienecke reviewed and improved this document and we thank them for their generous contribution of time and knowledge.

This project was funded by the Department of Conservation's Conservation Services Programme as Objective 7, Project POP2014-02.

7. References

- Archaux F, Henry PY, Gimenez O (2012) When can we ignore the problem of imperfect detection in comparative studies? *Methods in Ecology and Evolution* 3:188-194
- Bailey AM, Sorenson JH (1962) Subantarctic Campbell Island. Proceedings No. 10. Denver Museum of Natural History, Denver
- Bailey LL, Hines JE, Nichols JD, MacKenzie DI (2007) Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. *Ecology Applications* 17:281-290
- Baker GB, Cunningham R, Hedley G, King S (2008) Data collection of demographic, distributional and trophic information on the Westland petrel to allow estimation of effects of fishing on population viability. Report prepared for The Ministry of Fisheries PRO2006-01H. Latitude 42, Tasmania
- Baker GB, Hedley G, Cunningham R (2010) Data collection of demographic, distributional, and trophic information on the flesh-footed shearwater to allow estimation of effects of fishing on population viability: 2009-10 Field Season. Report prepared for The Ministry of Fisheries PRO2006-01 I. Latitude 42, Tasmania
- Barbraud C, Delord K (2006) Population census of blue petrels *Halobaena caerulea* at Mayes Island, Iles Kerguelen. *Antarct Sci* 18:199-204
- Barbraud C, Delord K, Marteau C, Weimerskirch H (2009) Estimates of population size of white-chinned petrels and grey petrels at Kerguelen Islands and sensitivity to fisheries. *Anim Conserv* 12:258-265
- Barbraud C, Marteau C, Ridoux V, Delord K, Weimerskirch H (2008) Demographic response of a population of white-chinned petrels *Procellaria aequinoctialis* to climate and longline fishery bycatch. *J Appl Ecol* 45:1460-1467
- Bartle JA (1990) Sexual segregation of foraging zones in Procellariiform birds: Implications of accidental capture on commercial fishery longlines of grey petrels *Procellaria cinerea*. *Notornis* 37:146-150
- Bell EA (2002) Grey petrels (*Procellaria cinerea*) on Antipodes Island, New Zealand: research feasibility, April to June 2001. DOC Science Internal Series 60. Department of Conservation, Wellington
- Bell EA, Bell BD, Sim JL, Imber MJ (2013a) Notes on the distribution, behaviour and status of grey petrel (*Procellaria cinerea*) on Antipodes Island, New Zealand. *Notornis* 60:269-278
- Bell EA, Sim JL, Scofield P (2009) Population parameters and distribution of the black petrel (*Procellaria parkinsoni*), 2005/06. DOC Res Dev Ser 307:1-47
- Bell EA, Sim JL, Scofield P (2011) Population parameters and distribution of the black petrel (*Procellaria parkinsoni*) on Great Barrier Island (Aotea Island), 2007/08. DOC Marine Conservation Services Series 8:1-37
- Bell EA, Sim JL, Scofield P, Francis C, Landers TJ (2013b) At-sea distribution and population parameters of the black petrels (*Procellaria parkinsoni*) on Great Barrier Island (Aotea Island), 2012/13. POP2012/13: Black petrels. Unpublished report by Wildlife Management International.
- Berrow SD (2000) The use of acoustics to monitor burrow-nesting white-chinned petrels *Procellaria aequinoctialis* at Bird Island, South Georgia. *Polar Biol* 23:575-579
- Borchers DL, Buckland ST, Zucchini W (2002) Estimating animal abundance: Closed populations. Springer-Verlag, London
- Bourgeois K, Ouni R, Pascal M, Dromzee S, Fourcy D, Abiadh A (2013) Dramatic increase in the Zembretta Yelkouan shearwater breeding population following ship rat

- eradication spurs interest in managing a 1500-year old invasion. *Biol Invasions* 15:475-482
- Brooke M (2004) *Albatrosses and Petrels Across the World*. Oxford University Press, Oxford
- Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2001) Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford
- Burger AE, Lawrence AD (2001) Census of wedge-tailed shearwaters *Puffinus pacificus* and Audubon's shearwaters *P. lherminieri* on Cousin Island, Seychelles using call-playback. *Mar Ornithol* 29:57-64
- Carey MJ (2009) The effects of investigator disturbance on procellariiform seabirds: a review. *N Z J Zool* 36:367-377
- Catry P, Campos A, Segurado P, Silva M, Strange I (2003) Population census and nesting habitat selection of thin-billed prion *Pachyptila belcheri* on New Island, Falkland Islands. *Polar Biol* 26:202-207
- Charleton K, Bragg C, Knight B, Fletcher D, Moller H, Newman J, Scott D (2009) Spatial variation in burrow entrance density of the sooty shearwater (*Puffinus griseus*). *Notornis* 56:1-10
- Croxall JP, Butchart SHM, Lascelles B, Stattersfield AJ, Sullivan B, Symes A, Taylor P (2012) Seabird conservation status, threats and priority actions: a global assessment. *Bird Conserv Int* 22:1-34
- Cuthbert R (2004) Breeding biology of the Atlantic Petrel, *Pterodroma incerta*, and a population estimate of this and other burrowing petrels on Gough Island, South Atlantic Ocean. *Emu* 104:221-228
- Cuthbert R, Davis LS (2002) Adult survival and productivity of Hutton's Shearwaters. *Ibis* 144:423-432
- Cuthbert RJ (2001) Conservation and ecology of Hutton's shearwater (*Puffinus huttoni*). Conservation Advisory Services Notes No. 335. Department of Conservation, Wellington
- Cuthbert RJ, Louw H, Lurling J, Parker G, Rexer-Huber K, Sommer E, Visser P, Ryan PG (2013) Low burrow occupancy and breeding success of burrowing petrels at Gough Island: a consequence of mouse predation. *Bird Conserv Int* 23:113-124
- Defos du Rau P, Bourgeois K, Thévenet M, Ruffino L, Dromzée S, Ouni R, Abiadh A, Estève R, Durand J, Anselme L, Faggio G, Mohamed Yahya J, Rguibi H, Renda M, Miladi B, Hamrouni H, Alilech S, Nefla A, Jaouadi W, Agrebi S, Renou S (2015) Reassessment of the size of the Scopoli's Shearwater population at its main breeding site resulted in a tenfold increase: implications for the species conservation. *J Ornithol* doi:10.1007/s10336-015-1187-4
- Dilley BJ, Schramm M, Ryan PG (submitted) Are mice suppressing the recovery of burrow-nesting petrels after the eradication of cats from Marion Island? *Biol Conserv* submitted
- Fewster RM, Buckland ST, Burnham KP, Borchers DL, Jupp PE, Laake JL, Thomas L (2009) Estimating the encounter rate variance in distance sampling. *Biometrics* 65:225-236
- Francis RICC, Bell EA (2010) Fisheries risks to the population viability of black petrel (*Procellaria parkinsoni*). New Zealand Aquatic Environment and Biodiversity Report No. 51. Ministry of Fisheries, Wellington
- Granadeiro JP, Dias MP, Rebelo R, Santos CD, Catry P (2006) Numbers and population trends of Cory's shearwater *Calonectris diomedea* at Selvagem Grande, Northeast Atlantic. *Waterbirds* 29:56-60

- Hall AJ (1987) The breeding biology of the white-chinned petrel *Procellaria aequinoctialis* at South Georgia. *J Zool (Lond)* 212:605-617
- Hamilton S (1998) A test of burrow occupancy of Sooty Shearwaters (*Puffinus griseus*) using chick response to sound. *Notornis* 45:64-66
- Hamilton S (2000) How precise and accurate are data obtained using an infra-red scope on burrow-nesting sooty shearwaters *Puffinus griseus*? *Mar Ornithol* 28:1-26
- Hunter CM (2001) Demography of Procellariids: model complexity, chick quality and harvesting. Thesis, University of Otago, Dunedin
- Hunter CM, Fletcher DJ, Scofield RP (2001) Preliminary modelling of black petrels (*Procellaria parkinsoni*) to assess population status. Department of Conservation, Wellington
- Jamieson SE, Waugh SM (2015) An assessment of recent population trends of flesh-footed shearwaters (*Puffinus carneipes*) breeding in New Zealand. *Notornis* 62:8-13
- Jenness JS (2004) Calculating landscape surface area from digital elevation models. *Wildl Soc Bull* 32:829-839
- Lavers JL (2015) Population status and threats to Flesh-footed Shearwaters (*Puffinus carneipes*) in South and Western Australia. *ICES J Mar Sci* 72:316-327
- Lawton K, Robertson G, Kirkwood R, Valencia J, Schlatter R, Smith D (2006) An estimate of population sizes of burrowing seabirds at the Diego Ramirez archipelago, Chile, using distance sampling and burrow-scoping. *Polar Biol* 29:229-238
- Lormée H, Delord K, Letournel B, Barbraud C (2012) Population survey of Leach's storm-petrels breeding at Grand Colombier Island, Saint-Pierre and Miquelon Archipelago. *Wilson J Ornithol* 124:245-252
- MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LL, Hines JE (2006) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, San Diego
- Martin AR, Poncet S, Barbraud C, Foster E, Fretwell P, Rothery P (2009) The white-chinned petrel (*Procellaria aequinoctialis*) on South Georgia: population size, distribution and global significance. *Polar Biol* 32:655-661
- McKechnie S, Bragg C, Newman J, Scott D, Fletcher D, Moller H (2009) Assessing the monitoring of sooty shearwater (*Puffinus griseus*) abundance in southern New Zealand. *Wildl Res* 36:541-552
- McKechnie S, Fletcher D, Moller H, Scott DS, Newman J, Bragg C (2007) Estimating and correcting for bias in population assessments of sooty shearwaters. *J Wildl Manage* 71:1325-1335
- Morrison ML, Block WM, Strickland MD, Collier BA, Peterson MJ (2008) Wildlife study design. Second edition. Springer, New York
- Mulder CPH, Grant-Hoffman MN, Towns DR, Bellingham PJ, Wardle DA, Durrett MS, Fukami T, Bonner KI (2009) Direct and indirect effects of rats: does rat eradication restore ecosystem functioning of New Zealand seabird islands? *Biol Invasions* 11:1671-1688
- Newman J, Scott D, Moller H, Fletcher D (2008) A population and harvest intensity estimate for sooty shearwater, *Puffinus griseus*, on Taukihepa (Big South Cape), New Zealand. *Pap Proc R Soc Tasman* 142:177-184
- Oppel S, Hervías S, Oliveira N, Pipa T, Silva C, Geraldés P, Goh M, Immler E, McKown M (2014) Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. *Nat Conserv* 7:1-13
- Parker GC, Rexer-Huber K, Thompson D (2015) An assessment of the breeding distribution of grey petrels *Procellaria cinerea*, Campbell Island, New Zealand. Report to the Department of Conservation. Parker Conservation, Dunedin

- Ratcliffe N, Vaughan D, Whyte C, Shepherd M (1998) Development of playback census methods for Storm Petrels *Hydrobates pelagicus*. *Bird Study* 45:302-312
- Rayner MJ, Clout MN, Stamp RK, Imber MJ, Brunton DH, Hauber ME (2007a) Predictive habitat modelling for the population census of a burrowing seabird: A study of the endangered Cook's petrel. *Biol Conserv* 138:235-247
- Rayner MJ, Dunphy BJ, Landers TJ (2009) Grey-faced petrel (*Pterodroma macroptera gouldi*) breeding on Little Barrier Island, New Zealand. *Notornis* 56:222-223
- Rayner MJ, Hauber ME, Clout MN (2007b) Breeding habitat of the Cook's Petrel (*Pterodroma cookii*) on Little Barrier Island (Hauturu): Implications for the conservation of a New Zealand endemic. *Emu* 107:59-68
- Rayner MJ, Parker KA, Imber MJ (2008) Population census of Cook's Petrel *Pterodroma cookii* breeding on Codfish Island (New Zealand) and the global conservation status of the species. *Bird Conserv Int* 18:211-218
- Rexer-Huber K, Parker GC, Ryan PG, Cuthbert R (2014) Burrow occupancy and population size in the Atlantic Petrel: a comparison of methods. *Mar Ornithol* 42:137-141
- Rexer-Huber K, Parker GC, Sagar PM, Thompson D (2015) White-chinned petrel population census, Disappointment Island (Auckland Islands). Report to the Agreement for the Conservation of Albatrosses and Petrels. Parker Conservation, Dunedin
- Richard Y, Abraham ER (2013) Risk of commercial fisheries to New Zealand seabird populations. New Zealand Aquatic Environment and Biodiversity Report No. 109. Ministry for Primary Industries, Wellington
- Rodway MS, Lemon MJF (2011) Use of permanent plots to monitor trends in burrow-nesting seabird populations in British Columbia. *Mar Ornithol* 39:243-253
- Ryan PG, Dilley BJ, Jones MGW (2012) The distribution and abundance of white-chinned petrels (*Procellaria aequinoctialis*) breeding at the sub-Antarctic Prince Edward Islands. *Polar Biol* 35:1851-1859
- Ryan PG, Dorse C, Hilton GM (2006) The conservation status of the spectacled petrel *Procellaria conspicillata*. *Biol Conserv* 131:575-583
- Ryan PG, Moloney CL (2000) The status of spectacled petrels *Procellaria conspicillata* and other seabirds at Inaccessible Island. *Mar Ornithol* 28:93-100
- Ryan PG, Ronconi RA (2011) Continued increase in numbers of spectacled petrels *Procellaria conspicillata*. *Antarct Sci* 23:332-336
- Schulz M, Robinson S, Gales R (2005) Breeding of the Grey Petrel (*Procellaria cinerea*) on Macquarie Island: population size and nesting habitat. *Emu* 105:323-329
- Schumann N, Dann P, Hoskins AJ, Arnould JPY (2013) Optimizing survey effort for burrow-nesting seabirds. *J Field Ornithol* 84:69-85
- Soanes LM, Thomas RJ, Bolton M (2012) Evaluation of field and analytical methods for estimating the population size of burrow-nesting seabirds from playback surveys. *Bird Study* 59:353-357
- Sommer E, Boyle D, Baer J, Fraser MJ, Palmer D, Sagar PM (2010) Antipodes Island white-chinned petrel and grey petrel field work report, 2009-10. Unpublished Final Research Report to Ministry of Fisheries. NIWA, Wellington.
- Sutherland DR, Dann P (2012) Improving the accuracy of population size estimates for burrow-nesting seabirds. *Ibis* 154:488-498
- Taylor GA (1988) Report to the Department of Conservation on an expedition to the Auckland Islands National Reserve, February 1988. Department of Conservation, Auckland
- Thompson S (1991) Adaptive cluster sampling: designs with primary and secondary units. *Biometrics* 47:1103-1115

- Tyre AJ, Tenhumberg B, Field SA, Niejalke D, Parris K, Possingham HP (2003) Improving precision and reducing bias in biological surveys: Estimating false-negative error rates. *Ecol Appl* 13:1790-1801
- Vaughan D, Gibbons DW (1998) The status of breeding Storm Petrels *Hydrobates pelagicus* on Skokholm Island in 1995. *Seabird* 20:12-21
- Walther BA, Moore JL (2005) The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography* 28:815-829
- Warham J (1990) *The Petrels: their Ecology and Breeding Systems*. Academic Press, London
- Waugh SM, Barbraud C, Adams L, Freeman AND, Wilson K-J, Wood G, Landers TJ, Baker GB (2015) Modeling the demography and population dynamics of subtropical seabird, and the influence of environmental factors. *Condor* 117:147-164
- Waugh SM, Cabrera H, Wood GC, Davis LS (2003) Burrow occupancy in Westland petrels (*Procellaria westlandica*). *Notornis* 50:123-127
- Waugh SM, Doherty PF, Freeman AND, Adams L, Woods GC, Bartle JA, Hedley GK (2006) Demography of Westland Petrels (*Procellaria westlandica*), 1995-2003. *Emu* 106:219-226
- Waugh SM, MacKenzie DI, Fletcher D (2008) Seabird bycatch in New Zealand trawl and longline fisheries, 1998-2004. *Pap Proc R Soc Tasman* 142:45-66
- Waugh SM, Tennyson AJD, Taylor GA, Wilson K-J (2013) Population sizes of shearwaters (*Puffinus* spp.) breeding in New Zealand, with recommendations for monitoring. *Tuhinga*:159-204
- Whitehead AL, Lyver POB, Jones CJ, Bellingham PJ, MacLeod CJ, Coleman M, Karl BJ, Drew K, Pairman D, Gormley AM, Duncan RP (2014) Establishing accurate baseline estimates of breeding populations of a burrowing seabird, the grey-faced petrel (*Pterodroma macroptera gouldi*) in New Zealand. *Biol Conserv* 169:109-116
- Williams KA, Frederick PC, Nichols JD (2011) Use of the superpopulation approach to estimate breeding population size: an example in asynchronously breeding birds. *Ecology* 92:821-828
- Wolfaardt A, Phillips R (2011) Guideline census methodologies for albatrosses and petrels. Agreement on the Conservation of Albatrosses and Petrels. Joint Fourth Meeting of Breeding Sites Working Group (BSWG4) and Sixth Meeting of Status and Trends WG (STWG6). ACAP, Guayaquil, Ecuador
- Wood GC, Otley HM (2013) An assessment of the breeding range, colony sizes and population of the Westland petrel (*Procellaria westlandica*). *N Z J Zool* 40:186-195
- Zotier R (1990) Breeding ecology of a Sub-Antarctic winter breeder - the grey petrel *Procellaria cinerea* on Kerguelen Islands. *Emu* 90:180-184