A comparison of the analytical methods used in population studies of storm petrels and the implications for colony estimates

Michael C. Shewry[1*] and Stuart Murray[2]

* *Correspondence author. Email: mcshewry@gmail.com*
[1] *49 Oakbank Road, Perth, PH11HG, UK;*
[2] *Easter Craigie Dhu, Dunkeld, Perthshire PH80EY, UK.*

Abstract
The standard census method for burrow-nesting petrels involves using an estimated response rate to correct the number of responses to tape playback obtained across a study site. Response rates can vary significantly between sites and between surveys and are estimated by determining the level of response within a calibration plot for which the number of occupied sites must also be estimated. The field methods used for such surveys are well established. However, a range of analytical methods have been employed to determine the number of occupied sites within the calibration plot and hence the response rate. The choice of method can have a huge effect on the estimated population size. Data from surveys of Leach's *Oceanodroma leucorhoa* and European Storm-petrel *Hydrobates pelagicus* on North Rona in 2001, 2009 and 2015 have been used to compare the three methods commonly employed to estimate the number of occupied burrows within the calibration plot. This number is usually subject to uncertainty and is key to estimating the population size. Our analysis suggests that the widely used 'reciprocal transformation' method can give rise to implausibly high population estimates which are up to 70% higher than those provided by other methods. We conclude that whilst potentially useful in some situations, this method should not be used in isolation. No single method appears best in all situations, but for North Rona the alternative 'curve fitting' and du Feu methods consistently give more plausible population estimates.

Introduction
The field methods employed for conducting census counts of burrow-nesting species such as Leach's *Oceanodroma leucorhoa* and European Storm-petrel *Hydrobates pelagicus* are well established and based on Ratcliffe *et al*. (1998) and Gilbert *et al*. (1998). Tape-playback is employed to locate 'apparently occupied sites' (AOS) across a study site. The number of AOS is then corrected for non-response using an estimated response rate. A calibration plot is used to estimate the response rate by undertaking daily repeat visits to all burrows in the plot, in order to ascertain the proportion of visits that elicit a response. For more detail see, for example, Bolton *et al*. (2010) or Murray *et al*. (2010).

Whilst there is a standard method for the fieldwork, there is less standardisation in the analytical methods employed to arrive at a population estimate. In particular, the response rate relies on estimating the number of AOS in the calibration plot, some of which may not have responded by the last visit. Study site population estimates are very sensitive to this number and the available methods can give very different results. Ensuring that sufficient visits are made and then selecting the correct analytical method is critical to avoid reporting biased population estimates. Bolton *et al*. (2010) compare two of the available methods using data for European Storm-petrel from Mousa, Shetland Islands, and they encourage others with similar data to conduct analyses to compare with their own findings. Thus, this short paper uses equivalent data from surveys on North Rona to assess the methods currently employed and aims to arrive at some recommendations for future surveys.

The standard method requires repeat visits over at least seven days and whilst, in theory, more visits could be made to increase the likelihood of finding most AOS, this is usually impractical. Also, the assumption is made that the response rate stays reasonably constant but this may only apply during the incubation period or shortly after hatching. Thus, the total number of occupied burrows has to be estimated and at least three methods are in use for this purpose. Once the total has been estimated, the response rate is usually calculated by simply dividing the mean daily number of responses obtained by the number of AOS, although other methods may be applicable if, for example, a declining response rate over the study period is suspected. This was evident for the European Storm-petrel survey on North Rona in 2001 and there the response rate was estimated by only making use of the first response from each burrow (Murray *et al*. 2010, 2016).

Two of the methods for estimating the calibration plot population are asymptotic approaches. The first fits an exponential curve to the cumulative number of responding burrows by visit (Mayhew *et al*. 2000). This is fitted using a non-linear regression method and the result is illustrated in the first graph of Figure 1a, with the population size provided by the curve's asymptote (46 in this example). The full methodology can be found in Appendix 2 of Mitchell *et al*. (2004). The second approach fits a straight line to a reciprocal version of the same plot, illustrated in the second graph of Figure 1a, and is based on Fowler (2001). The population size is provided by the value at the intercept with the y-axis (in this example 1/0.0169 = 59). For convenience, these methods will be termed the 'curve fitting' and 'reciprocal transformation' methods respectively. The third approach is du Feu's mark-recapture method, which treats second and subsequent responses at each burrow as recaptures (du Feu *et al*. 1983). Assuming each capture is equally likely, du Feu *et al*. (1983) provides a formula, which can be solved iteratively to estimate the population size. Although this method does make more complete use of the available data, its assumption of equal capture probability of all individuals on all days will be invalidated when, for example, some nests are no longer attended by adults during daylight on later surveys, as chicks no longer require brooding.

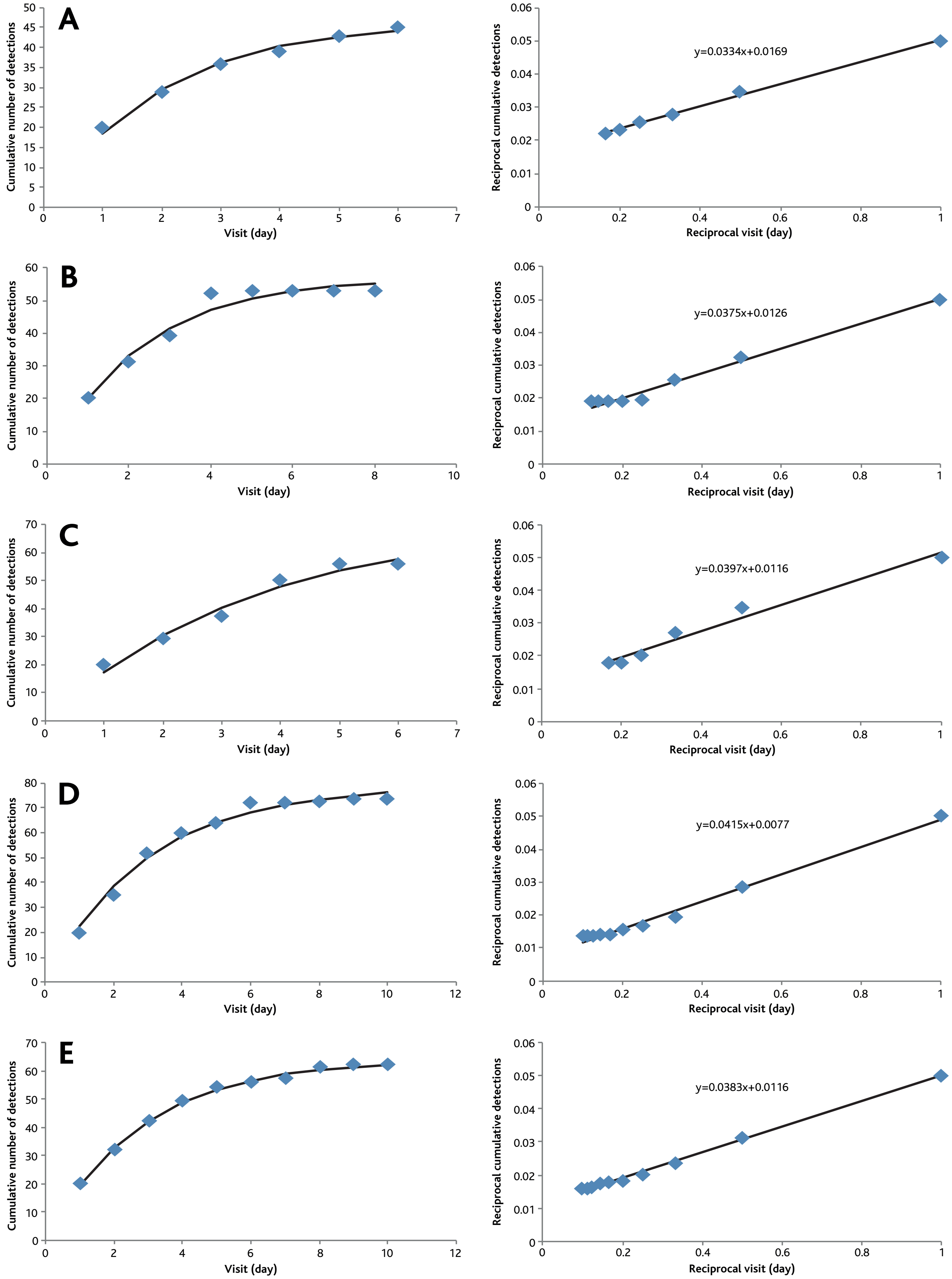

Figure 1. Cumulative and Reciprocal transformations from the North Rona surveys for European Storm-petrel *Hydrobates pelagicus* in a) 2001 and b) 2009 and Leach's Storm-petrel *Oceanodroma leucorhoa* in c) 2001, d) 2009 and e) 2015.

Several authors have compared the merits of the reciprocal transformation and du Feu methods in terms of their accuracy and the number of visits required for robust population estimation. Bolton *et al*. (2010) in presenting the results of European Storm-petrel surveys on Mousa, Shetland Islands, compared the accuracy of these methods. By choosing sites where the nest chamber was visible the true population size for the calibration plot could be ascertained, although the observers undertaking the calibration analysis were unaware of each chamber's occupancy status. They found that both methods gave similarly reliable estimates after four visits, with the du Feu method being slightly more accurate after seven visits. However, the method that was most accurate varied according to the number of visits made.

Soanes *et al*. (2012) compared the same methods using data for European Storm-petrel from two calibration plots on Skomer Island. In this case the true population sizes were unknown. The methods provided very different estimates, although the reciprocal transformation method was unable to provide an estimate for one plot even after nine visits. Du Feu was found to give the most statistically precise estimates (narrower confidence interval) but this does not mean it was more accurate.

Surveys on North Rona of European Storm-petrel in 2001 and 2009 and of Leach's Storm-petrel in 2001, 2009 and 2015 provide a further opportunity to compare these methods (Murray *et al*. 2010, 2016). These surveys were carried out using the same methods as those on Mousa and Skomer Island except that six daily visits were made to the calibration plots in 2001 and 8–10 visits in 2009 and 2015. The use of more than seven days in 2009 and 2015, together with higher response rates, allowed the true population size for the plots to be estimated with greater certainty.

Comparison of population methods

Population estimates for the calibration plots are shown in Table 1 by species and year. For the most part the curve fitting and du Feu's method give similar results, while those for the reciprocal transformation method are somewhat higher. The results for European Storm-petrel in 2001 are the exceptions. As there was evidence for a declining response rate over the study period this is likely to invalidate du Feu's method as the response probability will vary between occupied burrows. The curve fitting estimate also looks low given that the number of detections was still on the rise at day six (Murray *et al*. 2008). This was also true for Leach's Storm-petrel, although the level of recapture (second and subsequent response) was higher. On average there were 1.7 recaptures compared to 0.8 for European Storm-petrel. Partly as a result of this the curve fit and du Feu estimates look more plausible for Leach's Storm-petrel.

In 2009 and 2015, with more visits and higher recapture rates, the curve fit and du Feu estimates are very close. The reciprocal transformation estimates are out of step and look implausible, particularly as no new burrows were detected on the last visit to all three plots (Murray *et al*. 2010, 2016).

Table 1. Population size estimates for North Rona calibration plots.

	Year	Number of visits	Number of detections	Calibration plot population estimates Reciprocal transformation	Curve fit	du Feu
European	2001	6	45	59	46	62
	2009	8	53	79	57	55
Leach's	2001	6	56	86	71	61
	2009	10	74	130	79	76
	2015	10	62	86	64	63

Note: Estimates have been rounded to whole numbers.

Graphs showing the curve fit and the reciprocal transformation fitted straight line are shown in Figure 1. Focussing on the 2009 calibration plots sheds further light on why the reciprocal transformation estimates are suspect (Figures 1b and d). For both European and Leach's Storm-petrel the reciprocal transformation provides a relatively poor straight line fit because there were so few new responses in the last four visits. Thus, when response rates are high, the curve fitting approach is able to model this more effectively. It is true that the curve fit is relatively poor in this case for European Storm-petrel but the method is robust enough to still give a plausible population estimate.

Bolton *et al*. (2010) compare how quickly the methods converge to a final population estimate as the number of visits increases. For the North Rona plots this is shown in Figure 2, albeit just for the curve fitting and du Feu methods. The reciprocal transformation method has been left out for clarity and because of its implausible estimates for four out of the five plots.

For European Storm-petrel in 2001, du Feu's method must be viewed with suspicion because of a decline in the response rate over the study period. Estimates from the curve fitting approach are consistent from day three. In 2009, with no evidence of a changing response rate, the two estimates are close from day five and are converging to a value at or slightly above the total number of detections.

For Leach's Storm-petrel, the 2001 results are also erratic but with more consistent results being obtained using du Feu's method. By eye, the fit of the curve to the cumulative responses was relatively poor. In 2009 and 2015 the methods are consistent and stable from about day five. Response rates were relatively high and the curve provided a good fit to the cumulative responses.

Conclusions

Accurate estimation of the population size of the calibration plot is critical, as any errors will feed through proportionately to the estimated population of the study site. If the plot population is overestimated then so will be the study site population. Our analysis, and that conducted elsewhere, suggests that no single method for

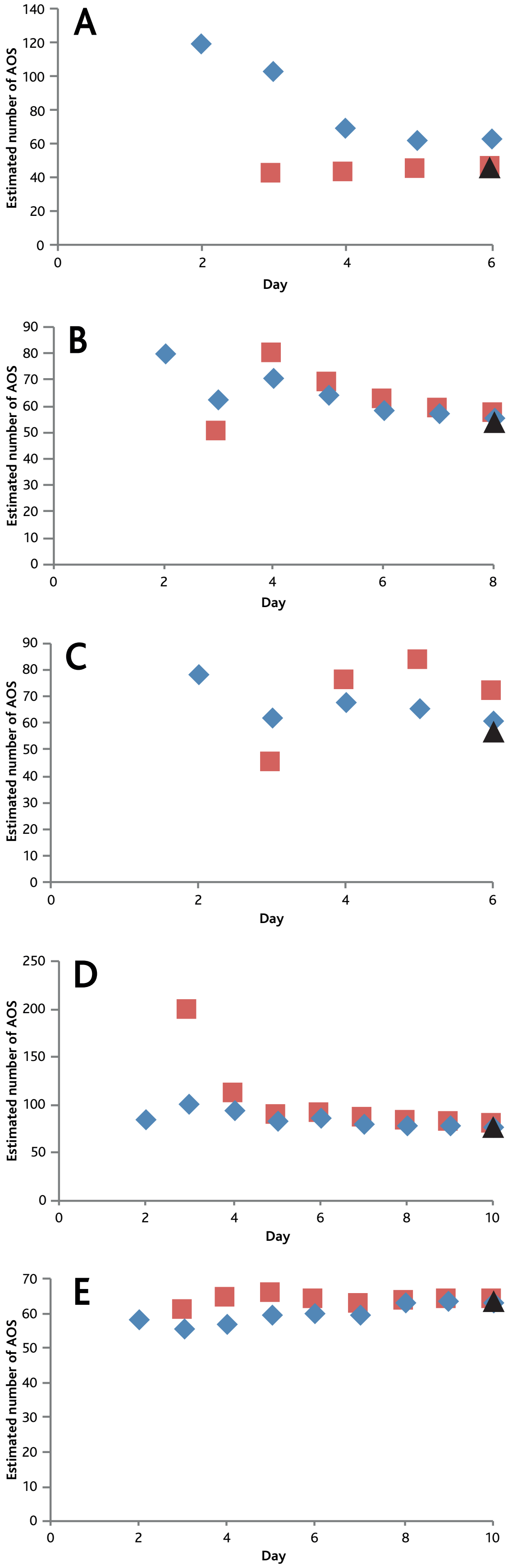

Figure 2. Effect of number of visits on population size estimates for European Storm-petrel *Hydrobates pelagicus* in a) 2001 and b) 2009 and Leach's Storm-petrel *Oceanodroma leucorhoa* in c) 2001, d) 2009 and e) 2015. ◆ = du Feu, ■ = Curve fitted, ▲ = Total AOS detected.

estimating the plot population is best in all situations: it will vary between sites and between years.

However, the reciprocal transformation method performed poorly with most of the North Rona data. For all but one calibration plot its estimate of the population size was significantly larger than that provided by the other methods. This matches the findings of Soanes *et al*. (2012), where the combined plot population estimate was 60% higher for the reciprocal transformation method compared to that using du Feu's method. However, in the absence of a plot of the cumulative number of responses by visit, it is not possible to judge which of these appears most realistic. Brown (2006) also uses the reciprocal transformation method and his Figure 2 does provide such plots for the 2003 and 2004 surveys of European Storm-petrel on Skomer Island, where 11 visits were made in 2003 and 15 in 2004. In both cases the estimated population sizes of 50.8 and 39.5 appear high compared to the number detected (34 and 35).

The chief disadvantages of the reciprocal transformation method are that:

- results are sensitive to the number of responses on the first visit. Regression lines are particularly sensitive to the data at the extremes and it can be seen from Figure 1b that visit one, far right, is highly influential;
- the method assumes a functional relationship between visit and the cumulative number of responding burrows which gives rise to implausible population estimates in many situations. The curve fitting method utilises just one of many possible functional forms, but seems to be more robust and has the distinct advantage of giving more equal weight to the data from each visit.

This second issue can be seen in both the Leach's and European Storm-petrel calibration plots from 2009 (Figure 1b and d). The cumulative number of detected AOS by visit increases more quickly than predicted by the reciprocal transformation's fitted line (the points are below the fitted line for several visits). This contrasts with the results from Mousa where response rates are much lower and the increase in the number of detected AOS corresponding slower (Figure 1 in Bolton *et al*. 2010). The reciprocal transformation method provides more sensible estimates in those circumstances.

It is possible that the reciprocal transformation method is seen as being equivalent to curve fitting. Appendix 2 of Mitchell *et al.* (2004) appears to suggest this when they say that the cumulative response curve "can be easily straightened using a reciprocal transformation of both axes". Straight line fitting is an attractive approach given that no specialist software is required but the methods are not the same: they imply different functional relationships between visit number and the cumulative number of responses.

In terms of the number of visits required, the results for European Storm-petrel (Figures 2a and b) suggest that there is merit in making at least seven visits where possible. In 2009, although the two methods are close from day five, they still exhibit further convergence towards the true population size on subsequent days. The results for 2001 are rather anomalous due to the calibration plot's declining response rate. Bolton *et al*. (2010) found that four or five visits were sufficient to get within 13% of the true population on Mousa, but our work suggests that will not necessarily be the case for other colonies.

For the Leach's Storm-petrel, ignoring the anomalous results for 2001, convergence was faster than for the European Storm-petrel: by day five both estimates had converged to within 15% of the final value in 2009 and within 5% in 2015. Nonetheless, some further convergence was evident on subsequent days.

In summary, we can conclude that:

- There is value in making seven or more visits where practical. More visits are likely to be needed if levels of response are low as the methods are likely to converge to the true population more slowly;
- The reciprocal transformation method produced suspiciously high estimates for North Rona and possibly on Skomer as well. It may well work best when levels of response are low and thus the cumulative number of AOS will converge more slowly to the true value. At the very least, a visual check of the straight line fit should be carried out;
- There is little to choose between the curve fitting and du Feu methods for the North Rona plots and both should be considered. Curve fitting does require the use of specialist software although R, which is freely available, does provide a Nonlinear Least Squares (nls) function which is reasonably straightforward to use;
- If feasible, plots such as those shown in Figure 2 should be produced to provide additional insight into the likely population size.

In general, selecting a low estimate for the plot's population will provide a conservative estimate of the study site's population. Whilst a biased estimate is undesirable, there is merit in choosing the lowest of several equally plausible plot populations. In addition, when estimating changes over time, care is needed to avoid the danger of over-estimating the population for one survey and under-estimating for another. Indeed, the results of earlier surveys may have to be revisited if a later survey indicates a change of approach. Thus, trends as well as population sizes are sensitive to the choice of method.

Perhaps above all, it is sensible to carry out a visual comparison of the population estimates with the plot of cumulative responses to help identify implausibly low or high estimates.

Acknowledgements

The 2009 and 2015 North Rona expeditions were partnership projects, made possible by funding provided by the Seabird Group, Scottish Natural Heritage, the Scottish Ornithologists' Club, the Gibson Estate, the Royal Society for the Protection of Birds, the Royal Commission on the Ancient and Historical Monuments of Scotland, Historic Scotland and the British Broadcasting Corporation, for which we are very grateful. The 2001 survey was carried out under contract to the Joint Nature Conservation Committee. Thanks are due to Mark Bolton, Mike Smith and an unnamed referee for extremely valuable comments on the draft manuscript.

References

Brown, J. 2006. Census of European Storm-petrels on Skomer Island. *Atlantic Seabirds* 8: 12–30.

Bolton, M., Brown, J. G., Moncrieff, H., Ratcliffe, N. & Okill, J. D. 2010. Playback re-survey and demographic modelling indicate a substantial increase in breeding European Storm-petrels *Hydrobates pelagicus* at the largest UK colony, Mousa, Shetland. *Seabird* 23: 14–24.

du Feu, C. R., Hounsome, M. V. & Spence, I. M. 1983. A single session mark/recapture method of population estimation. *Ringing & Migration* 4: 211–226.

Fowler, J. A. 2001. *Ecological studies in the maritime approaches to the Shetland oil terminal*. Report to Shetland Oil terminal Advisory Group & Scottish Natural Heritage. De Montefort University. Unpublished.

Gilbert, G., Gibbons, D. W. & Evans, J. 1998. *Bird monitoring methods, a manual of techniques for key UK species*. RSPB, Sandy.

Mayhew, P., Chisholm, K., Insley, H. & Ratcliffe, N. 2000. A survey of Storm Petrels on Priest Island in 1999. *Scottish Birds* 21: 78–84.

Mitchell, P. I., Newton, S. F., Ratcliffe, N. & Dunn, T. E. (eds.) 2004. *Seabird populations of Britain and Ireland*. Poyser, London.

Murray, S., Money, S., Griffin, A. & Mitchell, P. I. 2008. A survey of Leach's Storm-petrel *Oceanodroma leucorhoa* and European Storm-petrel *Hydrobates pelagicus* populations on North Rona and Sula Sgeir, Western Isles, Scotland. *Seabird* 21: 32–43.

Murray, S., Shewry, M. C., Harden, J., Jamie, K. & Parsons, M. 2010. A survey of Leach's and European Storm-petrel populations on North Rona and Sula Sgeir, Western Isles, Scotland, in 2009. *Seabird* 23: 25–40.

Murray, S., Shewry, M. C., Elliott, S., Jones, D. & Harden, J. 2016. A survey of Leach's Petrel *Oceanodroma leucorhoa* on North Rona, Western Isles, Scotland in 2015. *Seabird 26: 66–75*.

Ratcliffe, N., Vaughan, D., Whyte, C. & Shepherd, M. 1998. Development of playback census methods for Storm-petrels *Hydrobates pelagicus*. *Bird Study* 45: 302–312.

Soanes, L. M., Thomas, R. J. & 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.