CSIRO PUBLISHING
*Wildlife Research*
http://dx.doi.org/10.1071/WR16226

**Review**

# Sampling effort determination in bird surveys: do current norms meet best-practice recommendations?

*David M. Watson*

Institute for Land, Water and Society, Charles Sturt University, PO Box 789, Albury Wodonga, NSW 2640, Australia. Email: dwatson@csu.edu.au

**Abstract.** A critical design component of studies measuring diversity is sampling effort. Allocation of sampling effort dictates how many sites can be sampled within a particular time-frame or budget, as well as sample duration, frequency and intensity, thereby determining the resolution and reliability of emergent inferences. Conventional survey techniques use fixed-effort methods that assume invariant detectabilities among sites and species. Several approaches have been developed in the past decade that account for variable detectability by using alternative sampling methods or by adjusting standard counts before analysis, but it is unclear how widely adopted these techniques have been or how current bird surveying norms compare with best-practice recommendations. I conducted a systematic search of the primary literature to ascertain how sampling effort is determined, how much effort is devoted to sampling each site and how variation in detectability is dealt with. Of 225 empirical studies of bird diversity published between 2004 and 2016, five used results-based stopping rules (each derived independently), 54 used proportional sampling, and 159 (71%) used implicit effort-based stopping rules (fixed effort). Effort varied widely, but 61% of studies used samples of 10 min or less and 62% of studies expended total effort per datum of 2 h or less, with 78% providing no justification for sampling efforts used and just 15% explicitly accounting for estimated detectability. Given known variation in detectability, relying on short-duration fixed-effort approaches without validation or *post hoc* correction means that most bird diversity studies necessarily under-sample some sites and/or species. Having identified current bird surveying norms and highlighted their shortcomings, I provide five practical solutions to improve sampling effort determination, urging contributors and consumers of empirical ecological literature to consider survey data in terms of sample completeness.

**Additional keywords:** completeness, detectability, diversity, stopping rule.

Received 13 December 2016, accepted 17 April 2017, published online 26 May 2017

## Introduction

A key design component of empirical studies of diversity is sampling effort. The way in which sampling effort is allocated dictates the number of sites sampled within a particular time-frame or budget, as well as sample duration, frequency and intensity, thereby determining the strength, breadth and reliability of emergent inferences (Quinn and Keough 2002; Beck and Schwanghart 2010). Whereas overall sampling effort is ultimately constrained by available time and/or funding, effort expended at individual sites is generally determined by balancing quantitative rigour and desired resolution against perceived sources of variation in sample completeness (the proportion of species or individuals detected). Although most studies are designed by considering data resolution and logistical constraints, sources of variation in sample completeness are poorly-known and rarely acknowledged.

Much of the discussion associated with sampling effort and completeness is framed in terms of detectability i.e. the likelihood of recording a species when it is present (Gu and Swihart 2004; de Solla *et al.* 2005; Zhang *et al.* 2014). Detectability is variously considered in study design, namely when framing questions and selecting a sampling approach (*a priori*; see MacKenzie and Royle 2005) or once the data have been collected (*a posteriori*; see Kéry *et al.* 2010). When estimating species richness is the priority, richness estimators allow the number of species detected to be considered as a function of the total number of species predicted to occur (Peterson and Slade 1998). Although routinely used for invertebrates (Longino *et al.* 2002; Ellison *et al.* 2007) and plants (Archaux *et al.* 2006; Colwell *et al.* 2012), these predictive approaches are rarely used with other groups (Watson 2010), with priority being given to identifying species rather than estimating the proportion of species detected (Watson 2003). Emphasis in the recent ornithological literature has been on *a posteriori* approaches i.e. building models to estimate detection probability and correct for it (Marques *et al.* 2007; Murray *et al.* 2011; Hutto 2016). Although they can be applied at the community scale, these removal models are typically species-specific and are used primarily to generate standardised density or occupancy estimates (Farnsworth *et al.* 2002; Kéry and Schmid 2004; Royle and Link 2006), but they are unable to account for species present but not detected.

 

There is growing unease among ecologists and conservation biologists that correcting for variation in detectability *a posteriori* may introduce more bias into the dataset than the use of uncorrected estimates or, as Welsh *et al*. (2013) (p. 15) phrased it, 'ignoring non-detection can actually be better than trying to adjust for it'(but see Guillera-Arroita *et al*. 2014). Rather than collecting inconsistent data, there has been renewed interest in developing and using sampling approaches that explicitly account for variation in detectability, enabling resultant data to be analysed with confidence that they are both comparable and representative of the actual variation in occurrence. Thus, 'given a well-thought out and balanced sampling design, we suggest that unadjusted estimates of single- and multiple-species responses to ecological gradients can be just as robust as estimates that were *a posteriori* controlled for covariates of detection probability' (Banks-Leite *et al*. 2014, p. 857).

The confounding effects of detectability – specifically, variation in sample completeness among species and/or sites – are closely related to sampling effort (Gotelli and Colwell 2011) and of critical importance for conservation and management. Although the proportion of individuals and species present but not detected necessarily decreases with increasing effort per site, greater sampling efforts may not be practicable given design/budgetary constraints. Regardless of sampling effort, sample completeness varies among sites (Bebber *et al*. 2007) – in all but the most homogeneous of systems, the same quantum of sampling effort will under-sample some sites and over-sample others (Beck and Schwanghart 2010; Watson 2010), leading to potentially spurious inferences and misguided management recommendations. One solution to this conundrum is to scale effort to desired completeness, using results-based stopping rules to determine when a site is adequately sampled (Cam *et al.* 2002; Watson 2003). By minimising variation in the proportion of species or individuals missed, comparability of resultant estimates is maximised, regardless of the effort required (Peterson and Slade 1998; Chao *et al*. 2009). Although there is an extensive literature on the issues of detectability and how best to ensure that it does not confound estimates of density and species richness, it is unclear whether these approaches have become widely adopted by practitioners and empirical researchers.

In this quantitative review, I address the following questions: (1) how is sampling effort determined; (2) what approaches are used to minimise or otherwise account for variation in detectability; and (3) how do current methodological norms compare with best-practice recommendations? Having established current norms using a systematic literature search, I consider relationships between study design and survey effort, and distinguish determinants of variation in sample duration, stopping rules and consideration of detectability (in terms of the location and scale of study and the journal in which it was published). Finally, I present five recommendations for best-practice sampling effort determination, urging research practitioners and consumers alike to consider sample completeness explicitly, and to normalise estimations of variation in detectability. Although the present review is based on studies of birds – a group widely studied as surrogates for terrestrial diversity (Larsen *et al*. 2012) – the issues explored herein relate to diversity estimation generally, making this contribution relevant to anyone collecting or consuming data on species richness or occurrence.

## Literature search and analysis

To gain an overview of recent bird survey practices and identify current norms used to estimate bird richness and abundance, I conducted a systematic literature search. An initial search for publications containing the words 'landscape', 'bird' and 'diversity' in the abstract published between March 2004 and March 2016 (using Scopus) yielded 730 results. Once studies written in languages other than English, relating to wetland or coastal birds, using informal or purely qualitative methods (typically single-site studies), reviews, methodological comparisons and multiple papers using the same dataset were excluded, 225 studies presenting empirical data on determinants of bird diversity in terrestrial landscapes were identified for consideration (full dataset is summarised in Table S1, available as supplementary material to this paper). Rather than exhaustive or necessarily representative, this restricted search provided an indicative sample of the range of approaches that have been used to survey terrestrial bird communities and the associated sampling efforts and design protocols employed. For each study meeting the above criteria, the following data were retrieved: sample duration, total sampling effort used for each datum (if variable effort, mean value used), justification given for sampling effort and stopping rule applied. Total effort per datum was calculated on the basis of the finest scale of comparisons used in that particular study, whether at the treatment, habitat patch or landscape scale. Thus, if the study estimated richnesses for forest patches based on quarterly 10-min point counts for 4 years, the effort per datum would be 160 min, whereas if comparisons were conducted separately for each season, the effort per datum would be 40 min. Thirty-one studies considered effort in terms of area rather than time, so the number of studies used for the first two metrics is 194. Studies were also classified by location (North America, Europe, Africa and Middle East, Asia, Australasia, Latin America), habitat type (forest, woodland, grassland, farmland, plantation, urban), scale of inquiry (patch, vegetation type or landscape) and the impact factor of the journal in which they were published (2009 values used; see Table S1 for further details). Most results in the review are qualitative; however, some quantitative analyses were conducted, using either product–moment correlations (Pearson) between various measures of sampling effort, or generalised linear models to compare measures of sampling effort among different groups of studies defined by sampling design, with bivariate comparisons of means reported as Student's *t*-tests. All analyses were conducted in R (R Core Team 2013), using log-transformed data to normalise leptokurtic distributions, with untransformed values reported as means to facilitate interpretation.

## Results

### *Sampling effort*

Sample duration varied widely, but most studies (169 of the 194 studies that provided sampling effort in minutes; 87%) used samples of 20 min or shorter duration (Fig. 1), with 10 min

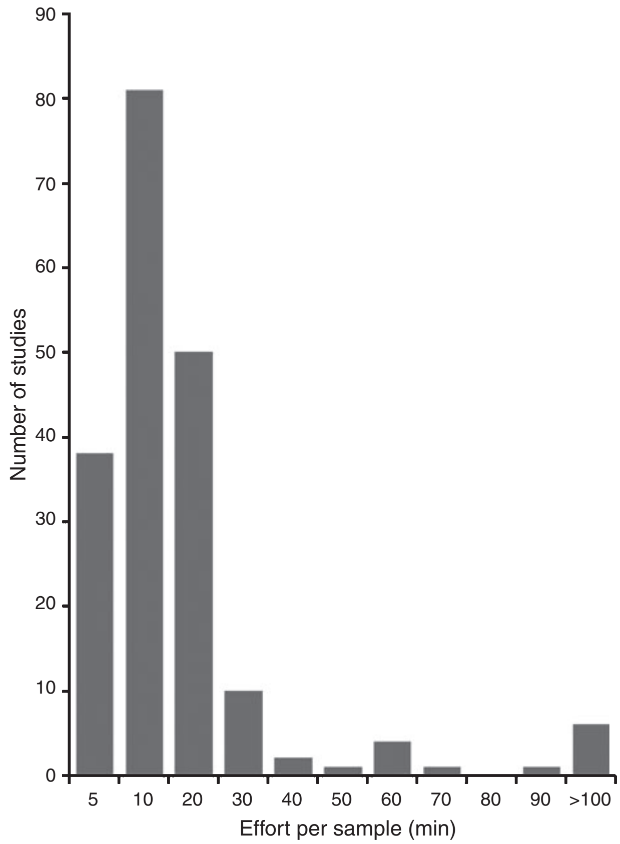


**Fig. 1.** Sample duration used by the 194 studies that provided sampling effort in minutes. For those studies using variable effort, mean values were used.

being the most frequently used sample duration (69 studies; 36%) followed by 5 min (36 studies; 19%), 20 min (28 studies; 14%) and 15 min (21 studies; 11%). Sampling effort per datum varied more widely, owing to initial differences in sample duration and variation in study design, sampling frequency and study duration (Fig. 2). Most studies (121; 62%) devoted 120 min or less to each datum; the most frequently employed effort was 40 min (16 studies), with 15 studies using 20 min, 14 studies 10 min and 13 studies 30 min.

Sample duration is just one component of study design and, although there was a significant positive relationship between the duration of a single sample and the total effort per datum ($r^2 = 0.399$, $P < 0.001$), total effort per datum was more closely related to frequency of sampling (number of occasions a study unit was sampled; $r^2 = 0.595$, $P = 0.003$) and sampling intensity (number of samples per site or study unit on each sampling occasion; $r^2 = 0.475$, $P < 0.001$).

Of the 225 studies considered, 175 (78%) did not explain how sampling effort was determined, with 169 providing no justification for the sampling effort expended and six citing reference(s) to a standard method (Table 1). A second set of 10 studies (4%) claimed that their sampling effort was sufficient; six of these found the effort to be sufficient in prior (unpublished) work, whereas four performed repeat visits to a subset of sites to ensure that no species were missed. The final set of 40 studies (18%) presented data demonstrating sufficiency of sampling efforts; 15 of these used rarefaction to adjust estimates to harmonise completeness, five studies used a results-based stopping rule to standardise sample completeness, five presented species accumulation data to demonstrate sufficiency, and 15 used richness estimates for analysis to control for differential completeness (Table 1).

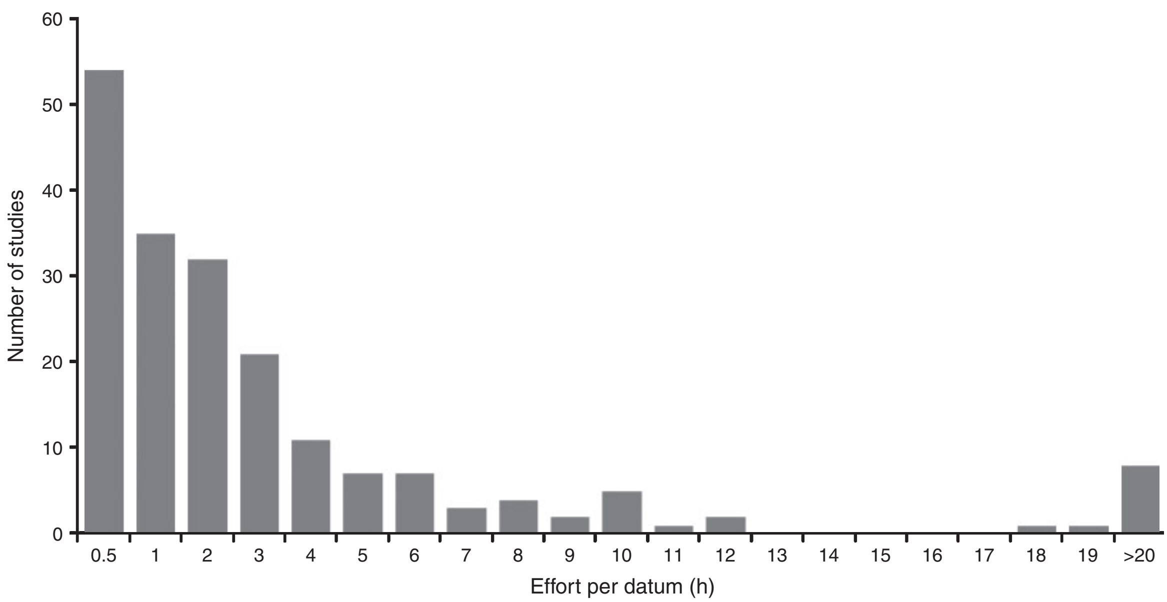


**Fig. 2.** Total effort per datum for the 194 studies that provided sampling effort in minutes (rescaled to hours). In addition to relating to sample duration, sampling frequency and sampling intensity (the number of samples per site on each sampling occasion), total effort relates to the question being asked. Thus, if monthly 20-min point counts are used for seasonal comparisons, effort per datum is 1 h, and 4 h if used for inter-annual comparisons.

These three groups of studies (sufficiency ignored ($n = 175$), sufficiency claimed ($n = 10$) and sufficiency demonstrated ($n = 40$)) were compared and no significant differences were found among them in terms of sample frequency, sample duration or impact factor of the journal in which they were published. Two significant differences were detected; both effort per datum and sampling intensity were greater in those studies where sufficiency was demonstrated than in those where sufficiency was ignored ($t = 2.76$, $P = 0.006$ and $t = 2.83$, $P = 0.005$ respectively, Fig. 3). Those studies that did not explain how sampling effort was determined expended an average of 290 min per datum, compared with 902 min for those studies where sampling sufficiency was demonstrated. To evaluate whether approaches to determine sampling effort have changed over time, proportions of studies in these three groups were compared for three time periods (2004–08, 2009–12, 2013–16), with the proportion of studies ignoring sufficiency increasing, and the proportion of studies claiming or demonstrating sufficiency decreasing (Table 1).

Few of the variables examined were significant predictors of variation in sampling effort, but the 20 studies from Asia were most divergent in terms of total effort per datum (mean of 945 min compared with 119 min), sample duration (46 min vs 15 min) and frequency (mean of 11 vs 4; $t = 3.26$, 4.05, 3.58 respectively, $P < 0.001$) compared with those from Europe. Likewise, several significant differences were noted among studies of different habitats; the 10 studies undertaken in plantations used samples with a mean duration of 68 min, compared with a mean of 10 min for grasslands ($t = 3.11$, $P = 0.002$). Survey method was related to sampling effort, with point counts being significantly lower than transects in terms of sample duration (12 min vs 58 min; $t = 6.23$, $P < 0.001$) and total effort per datum (321 min vs 873 min; $t = 2.36$, $P = 0.02$). There were significant differences between studies comparing birds

**Table 1. Determination of sampling effort**

Arrows (up for increasing, down for decreasing) denote consistent trends across the three subsets of the 12-year dataset

| Justification for sampling effort | *N* in 2004–08 (% of 62) | *N* in 2009–12 (% of 73) | *N* in 2013–16 (% of 90) | *N* in 2004–16 (% of 225) |
|---|---|---|---|---|
| No justification provided | 40 | 57 | 72 | 169 (75%) |
| Deemed sufficient by referenced paper | 3 | 1 | 2 | 6 (3%) |
| Sufficiency ignored | 43 (69%) | 58 (79%) | 74 (82%) | 175 (78%)↑ |
| Found sufficient in prior study | 3 | 2 | 1 | 6 (3%) |
| Consistent composition after repeat visits | 2 | 1 | 1 | 4 (2%) |
| Sufficiency claimed | 5 (8%) | 3 (4%) | 2 (2%) | 10 (4%)↓ |
| Tested using rarefaction | 7 | 4 | 4 | 15 (7%) |
| Results-based stopping rule | 1 | 1 | 3 | 5 (2%) |
| Rate of species accumulation | 3 | 1 | 1 | 5 (2%) |
| Calculated estimated richness | 3 | 6 | 6 | 15 (7%) |
| Sufficiency demonstrated | 14 (23%) | 12 (16%) | 14 (16%) | 40 (18%)↓ |

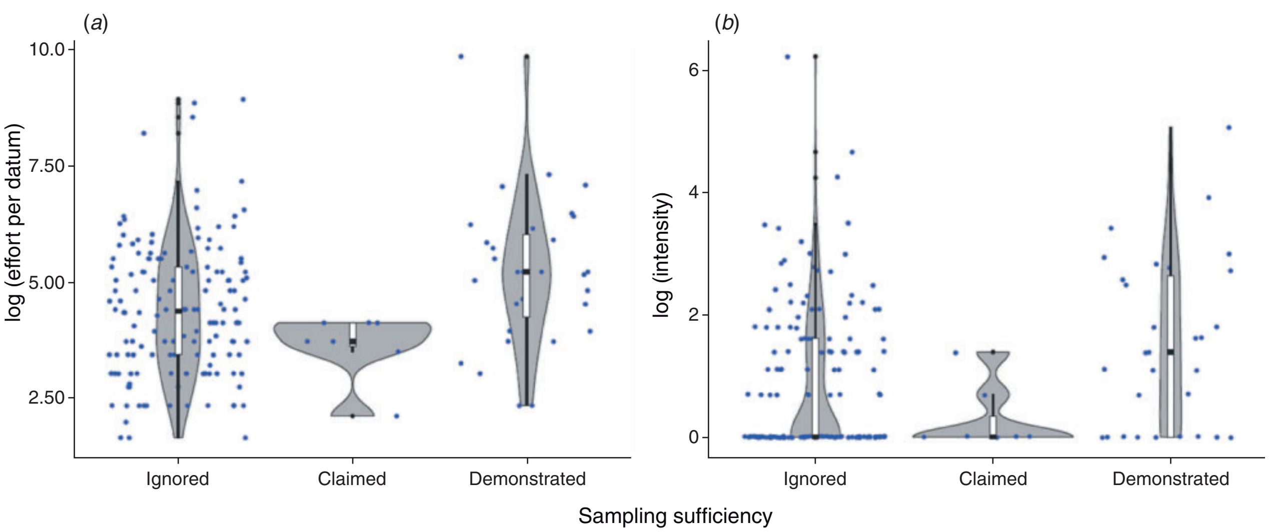


**Fig. 3.** Differences in (*a*) effort per datum (number of samples collected during each sampling occasion and (*b*) sampling intensity among studies based on how sufficiency of sampling effort was determined (log-transformed data plotted). Those studies that provided no information about how sampling effort was determined expended significantly less effort and lower sampling intensity than did those studies where sufficiency was demonstrated, reflecting a greater than three-fold difference in the mean sampling effort per datum (290 min vs 902 min).

associated with different vegetation types and patch-scale studies in terms of sample intensity (28 vs 4; $t = 4.43$, $P < 0.001$) and total effort per datum (50 min vs 805 min; $t = 3.22$, $P = 0.002$). No significant relationships were detected between journal impact factor and sample duration, sampling intensity, frequency or total effort per datum.

### *Stopping rules*

Of 225 studies considered, 54 (24%) used proportional sampling, with 53 using greater efforts at larger sites and one study using greater effort at sites with more complex vegetation. In total, 159 studies (71%) applied fixed-effort sampling – either sampling fixed areas, using samples of fixed duration or both; 12 studies (5%) used variable effort, of which seven used territory-mapping or area-search methods applied at the whole-of-patch scale and five used results-based stopping rules (Table 2).

Significant differences were discerned on the basis of how sampling effort was allocated. Fixed-effort studies employed a mean total effort per datum of 250 min, compared with 377 min for studies employing proportional effort and 472 min for studies using variable effort ($t = 4.03$, $P < 0.001$ and $t = 2.33$, $P = 0.02$ respectively). Sampling intensity also differed; fixed-effort studies employed a mean sampling intensity of 7, compared with 11 for proportional-effort and 25 for variable-effort designs ($t = 3.321$, $P = 0.001$; $t = 3.75$, $P < 0.001$). Sampling frequency also differed, with studies that used fixed-effort regimes sampling on an average of four occasions compared with three occasions for studies that used variable-effort designs ($t = 2.09$, $P = 0.037$). No significant differences were noted in terms of the impact factor of the journal publishing these studies. In terms of temporal variation, the proportion of studies using fixed-effort designs increased during the 12 years considered, whereas the proportion using proportional designs consistently decreased, with no clear pattern for the 12 studies using variable effort (Table 2).

### *Detectability*

In terms of detectability, 166 studies (74%) did not account for detectability; of those, 147 studies did not consider detectability at all, 10 studies claimed detectabilities were comparable (among both sites and species) but presented no corroborative data, and nine studies measured detection distance but did not consider detectability further nor did they adjust count data before analysis (Table 3). A second group of 26 studies (12%) considered detectability explicitly, but did not adjust their data before analysis, including 23 studies that minimised variation in detectability via design, and three studies that found detectabilities to be comparable in their system during previous work. The third and final group comprised 33 studies (15%) that explicitly accounted for detectability, including 13 studies that quantified detectability (four using removal models, four using occupancy models), 15 studies that used distance (Buckland *et al.* 1993) to adjust density estimates (five of which found density estimates comparable, so no adjustments were warranted), and five studies that calculated

**Table 2. Procedure used to allocate sampling effort**

Arrows (up for increasing, down for decreasing) denote consistent trends across the three subsets of the 12-year dataset

| Allocation of sampling effort | *N* in 2004–08 (% of 62) | *N* in 2009–12 (% of 73) | *N* in 2013–16 (% of 90) | *N* in 2004–16 (% of 225) |
|---|---|---|---|---|
| Fixed | 38 (61%) | 54 (74%) | 67 (74%) | 159 (71%)↑ |
| Variable (proportional[A]) | 19 (31%) | 18 (25%) | 17 (19%) | 54 (24%)↓ |
| Variable (patch-scale area search) | 4 (6%) | 0 (1%) | 3 (3%) | 7 (3%) |
| Variable (results-based stopping rule) | 1 (2%) | 1 (1%) | 3 (3%) | 5 (2%) |

[A]Of those 47 studies that used a proportional design, all but one study (where an additional sample was collected at sites with more complex vegetation) scaled effort to the site area.

**Table 3. Procedure used to estimate variation in detectability (either among species or sites)**

| Consideration of detectability | *N* in 2004–08 (% of 62) | *N* in 2009–12 (% of 73) | *N* in 2013–16 (% of 90) | *N* in 2004–16 (% of 225) |
|---|---|---|---|---|
| Not considered | 37 | 51 | 59 | 147 |
| Presumed similar | 4 | 3 | 3 | 10 (4%) |
| Measured detection distance (counts unadjusted) | 3 | 1 | 5 | 9 (4%) |
| Ignored detectability | 44 (71%) | 55 (75%) | 67 (74%) | 166 (74%) |
| Minimised via design | 8 | 6 | 9 | 23 (10%) |
| Found to be comparable in previous work | 2 | 0 | 1 | 3 (1%) |
| Considered detectability | 10 (16%) | 8 (11%) | 10 (11%) | 26 (12%) |
| Modelled detectability explicitly | 3 | 5 | 5 | 13 (6%) |
| Used distance | 2 | 6 | 7 | 15 (7%) |
| Used adjusted richness estimates | 3 | 1 | 1 | 5 (2%) |
| Accounted for detectability | 8 (13%) | 12 (16%) | 13 (14%) | 33 (15%) |

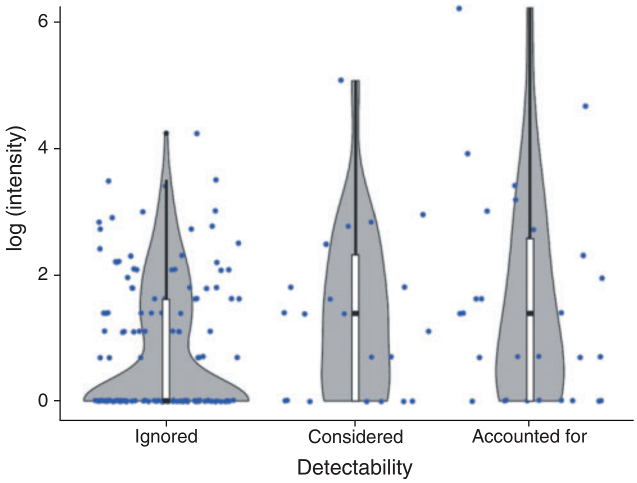


**Fig. 4.** Differences in sampling intensity (number of samples collected during each sampling occasion, log-transformed data plotted) among studies on the basis of how detectability was dealt with; those studies that made no mention of detectability (ignored) using fewer samples per sampling occasion (mean of 4) than did studies that considered detectability but made no adjustments (mean of 13), and those that accounted for variation in detectability explicitly (mean of 26).

estimated richness values, thereby accounting for species present but not detected (Table 3).

There were no significant differences among these three groups of studies in terms of either sampling frequency, sample duration, effort per datum or impact factor; however, those studies that ignored detectability used fewer samples per sampling occasion (mean of 4) than did either those studies that considered detectability but did not make any adjustments, or those studies that adjusted data before analysis (means of 13 and 26 respectively); differences deemed significant ($t = 2.15$, $P = 0.03$ and $t = 3.1$, $P = 0.002$ respectively, Fig. 4). No clear trends were detected in terms of temporal variation; more recently published studies were not more likely to consider or account for detectability than were those published earlier (Table 3).

## Discussion

The lack of conformity in all aspects of terrestrial bird surveys was striking, especially considering the restrictive criteria used to select studies for inclusion. Although journal impact factors have many shortcomings as measures of journal quality, the lack of any association between impact factor and sampling approach suggested that more highly-cited research does not necessarily employ more rigorous study design. Even considering studies conducted in similar ecosystems at comparable scales, there was no uniformity in sample duration, survey frequency, sampling intensity nor the overall effort per datum. Hence, for those 81 studies of forest bird communities published from 2004 to 2016, sample duration varied from 5 min to 3 h, sites were visited 1–40 times, with 1–500 samples being recorded on each occasion, resulting in a total sample effort per datum ranging across five orders of magnitude. This variability notwithstanding, the distribution of sampling effort (in terms of both single samples and total effort per datum) was strongly left-skewed (Figs 1, 2). Thus, for the same set of forest bird community studies, more than half employed total efforts per datum of 150 min or less. As discussed below, a longer sampling duration need not yield more consistent or comparable diversity estimates but, for a given study design, the potential confounding effects of both differential detectability and completeness on diversity estimates increase with diminishing sampling efforts (Watson 2004).

Also noteworthy was the redundancy in the literature surveyed, exemplified by multiple, independent derivations of results-based stopping rules to determine sampling effort. Of those five studies employing a results-based stopping rule, Aerts *et al*. (2008) re-invented the standardised search (Watson 2003), whereas Wyshynski and Nudds (2009) independently developed a results-based stopping rule to determine how many point counts were required. Werling *et al*. (2014) and McMahon *et al*. (2013) used 'whole-field searches', i.e. informal results-based stopping rules whereby sampling continued until the observer was satisfied that all individuals present were detected, with speed of walking and overall effort being determined by the number of birds present. Davis *et al*. (2013) (p. 429) used a similar informal approach to survey birds in urban woodland remnants and parks, where 'observers were instructed to take as much time as was necessary to survey all habitat types'. Although the approach used by this last study was clearly inspired by the standardised search (citing relevant papers in the associated reports that justify selection of methods in detail), not one of these five empirical studies cited any of the relevant primary literature, nor described the approach they used in terms of stopping rules. Indeed, none of the 225 studies considered here used the term 'stopping rule', demonstrating a disconnect between field-based empiricists applying methods and the comparative literature evaluating and refining them (e.g. Biaduń and Zmihorski 2011).

### *Detectability*

The very low proportion of studies that accounted for detectability (15%) is at odds with the current emphasis on adjusting for detectability, both in ornithology (Farnsworth *et al*. 2002; Marques *et al*. 2007; Watson 2010) and in the wider ecological literature (de Solla *et al*. 2005; Gotelli and Colwell 2011; Banks-Leite *et al*. 2014). Indeed, the majority (65%) of studies considered made no mention whatsoever of detectability. Whether or not this is problematic depends partly on one's perspective, but also relates to the objectives of the individual study. Thus, if a study aims to compare occurrence patterns in adjacent vegetation types (e.g. MacGregor-Fors and Schondube 2011), apparent differences could be confounded by different likelihoods of detection, potentially invalidating any inferences based on unadjusted counts (Gu and Swihart 2004). Conversely, if a study aims to identify which subset of an assemblage inhabiting native vegetation occurs within adjacent plantations (e.g. Azhar *et al*. 2013), differences in detectability may result in overlooked species in one or both habitats, but would be unlikely to diminish the overall validity of the findings (Colwell *et al*. 2012). Distance-based approaches to estimate individuals present but not recorded accounted for 7% of the studies evaluated; 9 of the 15 studies were conducted in North America or Europe. Several authors have identified shortcomings with distance-based methods to adjust density estimates (Elphick 2008; Welsh *et al*. 2013; Banks-Leite *et al*. 2014), and this approach is more frequently used for studies where estimating changes in abundance are a greater priority than determining variation in species richness (Elphick 2008; Hutto 2016). Of the 13 studies that modelled detectability explicitly, four used occupancy analysis, a modelling approach designed specifically for field-based studies recording species with variable detection probabilities (MacKenzie *et al*. 2003; MacKenzie and Royle 2005). Although developed in 2003 and becoming widely used in other disciplines, the fact that only four studies of the 225 considered here adopted this approach (three of which were conducted in North America) indicates that occupancy analysis is either unfamiliar to avian ecologists or ill-suited to community-scale studies of occurrence patterns (see Welsh *et al*. 2013).

An alternative approach that is gaining popularity (11 of 15 studies published in the past 5 years) is using predictive methods to estimate richness on the basis of the ratio of species seen repeatedly (duplicates) to those species seen only once (uniques), and either using these predicted values for analysis or adjusting recorded values or sampling effort on the basis of the estimated number of missed species (Peterson and Slade 1998). By scaling sampling effort to survey completeness, this approach is a robust means of generating comparable richness estimates in a wide range of habitats with very different likelihoods of detectability. Rather than numbers of individuals, species can be considered in terms of the proportion of samples in which they were detected (Watson 2003), with the resultant incidence values being far less susceptible to many of the species- and habitat-based factors known to confound density estimates (summarised by Verner 1985). Several studies employing fixed-effort designs used accumulation-based methods to check the overall completeness of surveys (e.g. Suarez-Rubio and Thomlinson 2009; Guldemond and van Aarde 2010; Pineda-Diez de Bonilla *et al*. 2012) or determine the number of samples (e.g. Fontana *et al*. 2011), but not to evaluate sufficiency of single samples. Although reassuring the reader that the overall community has been adequately sampled, this *post hoc* validation does not allow completeness of individual inventories to be estimated, nor consistency to be checked to ensure comparability. In addition to affecting inferences by the author(s), the ramifications of confounded comparisons extend to future application of reported results, which are of particular concern when management applications interpret reported absences as actual absences (Gu and Swihart 2004).

### *Sampling effort: how much is enough?*

More than 90% of the studies evaluated relied on samples of 30 min or shorter duration, with 72% using samples of 15 min or shorter duration. Most of the studies that were evaluated (175, 78%) did not explain how sampling effort was determined, and used less than a third the effort per datum of those 40 studies (18%) where sampling sufficiency was demonstrated (Fig. 3). Interestingly, several studies discussed explicitly the balance between sample duration and logistics, exemplified by Loss *et al*. (2009) (p. 2579) as follows 'Counts of longer duration were not used in order to allow for sampling of all survey points (185 total points in 2005; 145 points in 2006) twice during each breeding season.'. What constitutes a longer duration also varied: Linden *et al*. (2012) considered 30 min to be longer than a typical bird survey, whereas Sánchez-Oliver *et al*. (2014) (p. 137) considered their 10 min point counts (conducted twice each season) to be 'considerably longer than that used in previous studies recording species richness in woodland islands' (see also Matsuoka *et al*. 2014).

Many studies adhered to a fixed sampling protocol in the belief (expressed or implied) that this standardised protocol ensures comparability; for example, 'a standardised sampling protocol is most important and a potential underestimation of local diversity would enter the statistical analysis merely as a systematical error' (Flohre *et al*. 2011; see also Hutto 2016). Fixed-effort approaches (which accounted for 71% of the studies evaluated) generate comparable diversity estimates only if sites are sampled to the same degree (i.e. equivalent proportions of individuals/species are missed, yielding comparable completeness estimates); however, in the absence of completeness estimates, comparability is unknown. This concern is exacerbated by the small sample durations typically associated with fixed-effort approaches (5 min and 10 min being the two most commonly used survey durations) and the lack of corroborating data demonstrating that these sample durations are sufficient. It is worth noting that these short-duration fixed-effort counts were originally developed to estimate bird densities in continuous, homogeneous forests (Merikallio 1958); transect and point count methods are poorly suited for species richness estimation in patchy landscapes (Verner 1985; Watson 2004).

At the other extreme, greater sampling efforts can also be problematic, in terms of both reduced efficiency and inflated richness estimates. Although the studies employing results-based stopping rules used comparable sampling efforts, studies employing proportional sampling involved a significantly greater effort per datum. Of those 66 studies applying variable effort to different sites, most (60 studies; 91%) scaled sampling effort to site area. This approach pre-supposes a positive relationship between diversity and patch area, effectively forcing area sensitivity (Haila 1986; Horn *et al*. 2000). Although some studies recognised this issue explicitly (e.g. Shake *et al*. 2012), others (e.g. Loss *et al*. 2009) used more samples for larger sites, but then considered overall richness per sampling point, confounding underlying species–area relationships.

### *Dependent-variable selection: how informative is species richness?*

Although studies considered here necessarily measured species richness ('diversity' was one of three keywords used for the literature search), some ecologists consider abundance to be far more ecologically important (Šizling *et al*. 2009; Gaston 2010), so sampling approaches that overlook rarely encountered species need not be discarded. Even for those studies solely concerned with density or rank abundance, ensuring consistency of estimates across multiple sites and/or time periods remains critical. Although playing a more minor role in food webs and energy transfer than common species (Evans *et al*. 2005), rarely encountered species are often of greater management concern and more likely to be sensitive to changes in habitat extent or quality (see Matthews *et al*. 2014). Indeed, by explicitly focusing sampling methods on common species, it begs the question – how rare does an event need to be in order to justify exclusion? Numerous ecological interactions (e.g. pollination and seed dispersal) are disproportionately influenced by infrequent events – even one-off events can have profound and far-reaching consequences. Rather than simply relating to the duration of the sampling period, overall study design – especially the number of sites and sampling frequency – determines the likelihood of detecting infrequent events, and must be considered explicitly when selecting dependent variables.

### *Five practical solutions to improve bird survey methods*

In addition to researchers, consultants and other practitioners undertaking empirical studies, referees, editors and research managers share the responsibility to increase adoption of best-practice approaches to quantify bird occurrence patterns. To assist in this regard, I offer five practical solutions to be considered when planning or evaluating studies relying on estimates of bird diversity. Rather than advocating any particular approach, these solutions focus on maximising alignment among the methods selected, the study system and the question being addressed. By maximising efficiency and accuracy of diversity estimates, the inferential basis of studies can be improved and comparability among them enhanced.

#### *(1) Consider completeness*

Recognising that bird surveys (and biodiversity inventories generally) are necessarily incomplete, richness estimates should be considered relative to the number of species that were actually present. Dividing the former (recorded) integer by the latter (predicted) value yields a single number; namely, the estimated completeness of recorded species richness. By considering survey data in terms of sampling completeness during the design stage, the appropriateness of different sampling regimes can be objectively assessed relative to the question(s) being addressed. Hence, fixed-effort sampling pre-supposes minimal variation in sample completeness, whereas proportional sampling pre-empts the findings by scaling sampling effort to the area being sampled. By calculating completeness estimates, the validity of these assumptions can be iteratively evaluated, enabling researchers to gain a real-time appreciation for the appropriateness of their approach, identifying and correcting incipient issues in the field. Subsequently, presenting completeness data as a component of reporting allows research consumers to realise the inferential limitations of resultant richness estimates, contextualising uncertainty associated with the reported data and emergent interpretations.

#### *(2) Estimate detectability*

Given the very small proportion of ornithological studies presenting information on detectability, a comprehensive understanding of the relative influence of various sources of variation in detectability is not possible, leading to a pervasive nagging concern that detectability must be a confounding factor. The best answers to the question 'How was variation in detectability dealt with?', which is invariably asked of any ecological study, start with 'On the basis of estimated variation in detectability. . .'. Rather than invoking assumptions based on previous work, widely used methods or personal experience, data on the nature and magnitude of variation in detectability should be presented (e.g. Crates *et al*. 2017). Estimated variation in detectability may be minimal, or may vary in unanticipated ways. Adjustments to density or richness estimates may not be needed, or particular analyses may need to be regarded as provisional without undermining the rigour of the overall study. This information does not need to be onerous to collect or present, and may include a preliminary dataset comparing time till first detection for several key species in different patch sizes/habitat types, an evaluation of density estimates from surveys at different times of the day, or detection distances in different times of the day or seasons (e.g. La and Nudds 2016). By normalising the inclusion of such data, not only will strength of inference improve but a more complete understanding of the determinants of detectability will emerge.

#### *(3) Before using any method, test it*

Although so obvious it should go without saying, the present review highlighted the lack of objectivity associated with bird survey selection – only 40 (~18%) of the published papers considered herein demonstrated the sufficiency of sampling efforts employed. Just as research questions vary, so do their budgets, time-frames, study species, study systems and surveyors; so, every study represents a unique combination of factors to be reconciled when deciding on a sampling design and survey method(s). As with detectability estimation, testing a method need not require exhaustive additional work, but should be an integral component of initial fieldwork. Having selected a particular survey method, determining the sampling effort required can easily be achieved by sampling a small subset of sites exhaustively and examining the relationship between sampling effort and estimated completeness. In addition to resolving the most efficient way to apportion sampling effort, these initial trials can be used to demonstrate the sufficiency of sampling effort in resultant reports and publications.

#### *(4) Refine the question*

As part of methodological trials, consider closely the question you are addressing and ensure the time-frame, spatial scale, grain size and variables are all tailored to the question. If you are asking questions about the temporal dynamics of rainforest bird assemblages, do you need to include raptors, owls and swifts? If so, understand that this may require a doubling or tripling of the survey effort involved (Watson 2010). Conversely, if asking a question about woodland bird responses to a competitively dominant species, is species richness the best variable to use, or is rank abundance more ecologically meaningful? By alternating between aspects of study design and the exact scope of the question being asked, not only will a mismatch between approach and question be avoided, but efficiency of data collection will be maximised.

### *(5) Optimise sampling frequency to the question being addressed*

Although just one aspect of study design, frequency of sampling is one of the least considered aspects, despite having great influence over the total survey effort required in a particular study. To revisit the last two examples – for a study of rainforest residents, is it essential to conduct surveys every month or would paired surveys in the wet season and dry season suffice? If so, the same effort would yield 6 years of seasonal data instead of 1 year of monthly dynamics. Likewise, for the woodland bird competition study, why conduct four seasonal surveys over 2 years when surveys over the peak breeding season could be undertaken over twice the time-frame using half the overall effort? Repeat sampling is often combined with low sampling effort (e.g. Cunningham *et al*. 2014; Paker *et al*. 2014) when one single longer survey at the most ecologically relevant time may be far more instructive (and more cost effective).

## Acknowledgements

I am grateful to Dale Nimmo for assistance with data analysis and visualisation. Discussions with students in the Graduate Certificate of Ornithology helped clarify many points summarised here. Maggie Watson and Chris Elphick provided considered input on a previous version of this manuscript.

## References

Aerts, R., Lerouge, F., November, E., Lens, L., Hermy, M., and Muys, B. (2008). Land rehabilitation and the conservation of birds in a degraded Afromontane landscape in northern Ethiopia. *Biodiversity and Conservation* **1**, 53–69. doi:10.1007/s10531-007-9230-2

Archaux, F., Gosselin, F., Bergès, L., and Chevalier, R. (2006). Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* **17**, 299–306. doi:10.1111/j.1654-1103.2006.tb02449.x

Azhar, B., Lindenmayer, D. B., Wood, J., Fischer, J., Manning, A., McElhinny, C., and Zakaria, M. (2013). The influence of agricultural system, stand structural complexity and landscape context on foraging birds in oil palm landscapes. *The Ibis* **155**, 297–312. doi:10.1111/ibi.12025

Banks-Leite, C., Pardini, R., Boscolo, D., Cassano, C. R., Püttker, T., Barros, C. S., and Barlow, J. (2014). Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. *Journal of Applied Ecology* **51**, 849–859. doi:10.1111/1365-2664.12272

Bebber, D. P., Marriott, F. H. C., Gaston, K. J., Harris, S. A., and Scotland, R. W. (2007). Predicting unknown species numbers using discovery curves. *Proceedings of the Royal Society B* **274**, 1651–1658. doi:10.1098/rspb.2007.0464

Beck, J., and Schwanghart, W. (2010). Comparing measures of species diversity from incomplete inventories: an update. *Methods in Ecology and Evolution* **1**, 38–44. doi:10.1111/j.2041-210X.2009.00003.x

Biaduń, W., and Zmihorski, M. (2011). Factors shaping a breeding bird community along an urbanization gradient: 26-year study in medium size city (Lublin, SE Poland). *Polish Journal of Ecology* **59**, 381–389.

Buckland, S. T., Anderson, D. R., Burnham, K. P., and Laake, J. L. (1993). 'Distance Sampling: Estimating Abundance of Biological Populations.' (Chapman & Hall: London.)

Cam, E., Nichols, J. D., Sauer, J. R., and Hines, J. E. (2002). On the estimation of species richness based on the accumulation of previously unrecorded species. *Ecography* **25**, 102–108. doi:10.1034/j.1600-0587.2002.250112.x

Chao, A., Colwell, R. K., Lin, C.-W., and Gotelli, N. J. (2009). Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* **90**, 1125–1133. doi:10.1890/07-2147.1

Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S.-Y., Mao, C. X., Chazdon, R. L., and Longino, J. T. (2012). Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* **5**, 3–21. doi:10.1093/jpe/rtr044

Crates, R., Terauds, A., Rayner, L., Stejanovic, D., Ingwersen, D., and Webb, M. (2017). An occupancy approach to monitoring regent honeyeaters. *The Journal of Wildlife Management* **81**, 699–377. doi:10.1002/jwmg.21222

Cunningham, R. B., Lindenmayer, D. B., Crane, M., Michael, D. R., Barton, P. S., Gibbons, P., Okada, S., Ikin, K., and Stein, J. A. R. (2014). The law of diminishing returns: woodland birds respond to native vegetation cover at multiple spatial scales and over time. *Diversity & Distributions* **20**, 59–71. doi:10.1111/ddi.12145

Davis, R. A., Gole, C., and Roberts, J. D. (2013). Impacts of urbanisation on the native avifauna of Perth, Western Australia. *Urban Ecosystems* **16**, 427–452. doi:10.1007/s11252-012-0275-y

de Solla, S. R., Shirose, L. J., Fernie, K. J., Barrett, G. C., Brousseau, C. S., and Bishop, C. A. (2005). Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. *Biological Conservation* **121**, 585–594. doi:10.1016/j.biocon.2004.06.018

Ellison, A. M., Record, S., Arguello, A., and Gotelli, N. J. (2007). Rapid inventory of the ant assemblage in a temperate hardwood forest: species composition and assessment of sampling methods. *Environmental Entomology* **36**, 766–775. doi:10.1093/ee/36.4.766

Elphick, C. S. (2008). How you count counts: the importance of methods research in applied ecology. *Journal of Applied Ecology* **45**, 1313–1320. doi:10.1111/j.1365-2664.2008.01545.x

Evans, K. L., Greenwood, J. J. D., and Gaston, K. J. (2005). Relative contribution of abundant and rare species to species-energy relationships. *Biology Letters* **1**, 87–90. doi:10.1098/rsbl.2004.0251

Farnsworth, G. L., Pollock, K. H., Nichols, J. D., Simons, T. R., Hines, J. E., and Sauer, J. R. (2002). A removal model for estimating detection probabilities from point-count surveys. *The Auk* **119**, 414–425. doi:10.1642/0004-8038(2002)119[0414:ARMFED]2.0.CO;2

Flohre, A., Fischer, C., Aavik, T., Bengtsson, J., Berendse, F., Bommarco, R., Ceryngier, P., Clement, L. W., Dennis, C., Eggers, S., Emmerson, M., Geiger, F., Guerrero, I., Hawro, V., Inchausti, P., Liira, J., Morales, M. B., Oñate, J. J., Pärt, T., Weisser, W. W., Winqvist, C., Thies, C., and Tscharntke, T. (2011). Agricultural intensification and biodiversity partitioning in European landscapes comparing plants, carabids, and birds. *Ecological Applications* **21**, 1772–1781. doi:10.1890/10-0645.1

Fontana, C. S., Burger, M. I., and Magnusson, W. E. (2011). Bird diversity in a subtropical South-American city: effects of noise levels, arborisation and human population density. *Urban Ecosystems* **14**, 341–360. doi:10.1007/s11252-011-0156-9

Gaston, K. J. (2010). Valuing common species. *Science* **327**, 154–155. doi:10.1126/science.1182818

Gotelli, N., and Colwell, R. (2011). Estimating species richness. In 'Biological Diversity: Frontiers in Measurement and Assessment'. (Eds A. E. Magurran and B. J. McGill.) pp. 39–54. (Oxford University Press: Oxford, UK.)

Gu, W., and Swihart, R. K. (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation* **116**, 195–203. doi:10.1016/S0006-3207(03)00190-3

Guillera-Arroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., and McCarthy, M. A. (2014). Ignoring imperfect detection in biological surveys is dangerous: a response to 'fitting and interpreting occupancy models'. *PLoS ONE* **9**, e99571.

Guldemond, R. A. R., and van Aarde, R. J. (2010). Forest patch size and isolation as drivers of bird species richness in Maputaland, Mozambique. *Journal of Biogeography* **37**, 1884–1893.

Haila, Y. (1986). North European land birds in forest fragments: evidence for area effects? In 'Wildlife 2000. Modelling Habitat Relationships of Terrestrial Vertebrates'. (Eds J. Verner, M. L. Morrison and C. J. Ralph.) pp. 315–319. (The University of Wisconsin Press: Madisson, WI.)

Horn, D., Fletcher, R. Jr, and Koford, R. K. (2000). Detecting area sensitivity: a comment on previous studies. *American Midland Naturalist* **144**, 28–35. doi:10.1674/0003-0031(2000)144[0028:DASACO]2.0.CO;2

Hutto, R. L. (2016). Should scientists be required to use a model-based solution to adjust for possible distance-based detectability bias? *Ecological Applications* **26**, 1287–1294. doi:10.1002/eap.1385

Kéry, M., and Schmid, H. (2004). Monitoring programs need to take into account imperfect species detectability. *Basic and Applied Ecology* **5**, 65–73. doi:10.1078/1439-1791-00194

Kéry, M., Royle, J. A., Schmid, H., Schaub, M., Volet, B., Häfliger, G., and Zbinden, N. (2010). Site-occupancy distribution modelling to correct population-trend estimates derived from opportunistic observations. *Conservation Biology* **24**, 1388–1397. doi:10.1111/j.1523-1739.2010.01479.x

La, V. T., and Nudds, T. D. (2016). Effect of revisitation surveys on detection of wetland birds with different diel vocalization patterns. *Journal of Fish and Wildlife Management* **7**, 509–519. doi:10.3996/082015-JFWM-077

Larsen, F. W., Bladt, J., Balmford, A., and Rahbek, C. (2012). Birds as biodiversity surrogates: will supplementing birds with other taxa improve effectiveness? *Journal of Applied Ecology* **49**, 349–356. doi:10.1111/j.1365-2664.2011.02094.x

Linden, D. W., Roloff, G. J., and Kroll, A. J. (2012). Conserving avian richness through structure retention in managed forests of the Pacific Northwest, USA. *Forest Ecology and Management* **284**, 174–184. doi:10.1016/j.foreco.2012.08.002

Longino, J. T., Coddington, J. A., and Colwell, R. K. (2002). The ant fauna of a tropical rain forest: estimating species richness three different ways. *Ecology* **83**, 689–702. doi:10.1890/0012-9658(2002)083[0689:TAFOAT]2.0.CO;2

Loss, S. R., Ruiz, M. O., and Brawn, J. D. (2009). Relationships between avian diversity, neighborhood age, income, and environmental characteristics of an urban landscape. *Biological Conservation* **142**, 2578–2585. doi:10.1016/j.biocon.2009.06.004

MacGregor-Fors, I., and Schondube, J. E. (2011). Use of tropical dry forests and agricultural areas by neotropical bird communities. *Biotropica* **43**, 365–370. doi:10.1111/j.1744-7429.2010.00709.x

MacKenzie, D. I., and Royle, J. A. (2005). Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology* **42**, 1105–1114. doi:10.1111/j.1365-2664.2005.01098.x

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., and Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* **84**, 2200–2207. doi:10.1890/02-3090

Marques, T. A., Thomas, L., Fancy, S. G., and Buckland, S. T. (2007). Improving estimates of bird density using multiple-covariate distance sampling. *The Auk* **124**, 1229–1243. doi:10.1642/0004-8038(2007)124[1229:IEOBDU]2.0.CO;2

Matsuoka, S. M., Mahon, C. L., Handel, C. M., Sólymos, P., Bayne, E. M., Fontaine, P. C., and Ralph, C. J. (2014). Reviving common standards in point-count surveys for broad inference across studies. *The Condor* **116**, 599–608. doi:10.1650/CONDOR-14-108.1

Matthews, T. J., Cottee-Jones, H. E., and Whittaker, R. J. (2014). Habitat fragmentation and the species–area relationship: a focus on total species richness obscures the impact of habitat loss on habitat specialists. *Diversity & Distributions* **20**, 1136–1146. doi:10.1111/ddi.12227

McMahon, B. J., Carnus, T., and Whelan, J. (2013). A comparison of winter bird communities in agricultural grassland and cereal habitats in Ireland: implications for Common Agricultural Policy reform. *Bird Study* **60**, 176–184. doi:10.1080/00063657.2012.758227

Merikallio, E. (1958). Finnish birds: their distribution and numbers. *Fauna Fennica* **5**, 1–81.

Murray, L. D., Gates, R. G., and Spinola, R. M. (2011). Evaluation of three methods to estimate density and detectability from roadside point counts. *The Journal of Wildlife Management* **75**, 1072–1081. doi:10.1002/jwmg.131

Paker, Y., Yom-Tov, Y., Alon-Mozes, T., and Barnea, A. (2014). The effect of plant richness and urban garden structure on bird species richness, diversity and community structure. *Landscape and Urban Planning* **122**, 186–195. doi:10.1016/j.landurbplan.2013.10.005

Peterson, A. T., and Slade, N. A. (1998). Extrapolating inventory results into biodiversity estimates and the importance of stopping rules. *Diversity & Distributions* **4**, 95–105. doi:10.1046/j.1365-2699.1998.00021.x

Pineda-Diez de Bonilla, E., León-Cortés, J. L., and Rangel-Salazar, J. L. (2012). Diversity of bird feeding guilds in relation to habitat heterogeneity and land-use cover in a human-modified landscape in southern Mexico. *Journal of Tropical Ecology* **28**, 369–376. doi:10.1017/S026646741200034X

Quinn, G. P., and Keough, M. J. (2002). 'Experimental Design and Data Analysis for Biologists.' (Cambridge University Press: Cambridge, UK.)

R Core Team (2013). 'R: a Language and Environment for Statistical Computing.' (R Foundation for Statistical Computing: Vienna, Austria.) Available at http://www.R-project.org/ [verified 2 May 2017].

Royle, J. A., and Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* **87**, 835–841. doi:10.1890/0012-9658(2006)87[835:GSOMAF]2.0.CO;2

Sánchez-Oliver, J. S., Rey Benayas, J. M., and Carrascal, L. M. (2014). Differential effects of local habitat and landscape characteristics on bird communities in Mediterranean afforestations motivated by the EU Common Agrarian Policy. *European Journal of Wildlife Research* **60**, 135–143. doi:10.1007/s10344-013-0759-y

Shake, C. S., Moorman, C. E., Ridd Le, J. D., and Burchell, M. R. (2012). Influence of patch size and shape on occupancy by shrubland birds. *The Condor* **114**, 268–278. doi:10.1525/cond.2012.110107

Šizling, A. L., Šizlingová, E., Storch, D., Reif, J., and Gaston, K. J. (2009). Rarity, commonness and the contribution of individual species to species richness patterns. *American Naturalist* **174**, 82–93. doi:10.1086/599305

Suarez-Rubio, M., and Thomlinson, J. R. (2009). Landscape and patch-level factors influence bird communities in an urbanized tropical island. *Biological Conservation* **142**, 1311–1321. doi:10.1016/j.biocon.2008.12.035

Verner, J. (1985). Assessment of counting techniques. *Current Ornithology* **2**, 247–302. doi:10.1007/978-1-4613-2385-3_8

Watson, D. M. (2003). The 'standardized search': an improved way to conduct bird surveys. *Austral Ecology* **28**, 515–525. doi:10.1046/j.1442-9993.2003.01308.x

Watson, D. M. (2004). Comparative evaluation of new approaches to survey birds. *Wildlife Research* **31**, 1–11. doi:10.1071/WR03022

Watson, D. M. (2010). Optimizing inventories of diverse sites: insights from Barro Colorado Island birds. *Methods in Ecology and Evolution* **1**, 280–291.

Welsh, A. H., Lindenmayer, D. B., and Donnelly, C. F. (2013). Fitting and interpreting occupancy models. *PLoS One* **8**, e52015. doi:10.1371/journal.pone.0052015

Werling, B. P., Dickson, T. L., Isaacs, R., Gaines, H., Gratton, C., Gross, K. L., Liere, H., Malmstrom, C. M., Meehan, T. D., Ruan, L., Robertson, B. A., Robertson, G. P., Schmidt, T. M., Schrotenboer, A. C., Teal, T. K., Wilson, J. K., and Landis, D. A. (2014). Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes. *Proceedings of the National Academy of Sciences, USA* **111**, 1652–1657. doi:10.1073/pnas.1309492111

Wyshynski, S. A., and Nudds, T. D. (2009). Pattern and process in forest bird communities on boreal landscapes originating from wildfire and timber harvest. *Forestry Chronicle* **85**, 218–226. doi:10.5558/tfc85218-2

Zhang, J., Nielsen, S. E., Grainger, T. N., Kohler, M., Chipchar, T., and Farr, D. R. (2014). Sampling plant diversity and rarity at landscape scales: importance of sampling time in species detectability. *PLoS One* **9**, e95334. doi:10.1371/journal.pone.0095334