Austral Ecology (2003) **28**, 515–525

The 'standardized search': An improved way to conduct bird surveys

DAVID M. WATSON*
Applied Ornithology Group, The Johnstone Centre, Charles Sturt University, Albury, New South Wales, Australia (Email: dwatson@csu.edu.au)

Abstract Bird surveys are among the most widely used biodiversity inventories and serve as the basis for an increasing proportion of pure and applied ecological research. It is rarely possible to conduct exhaustive censuses of all individuals present at a particular site, so stopping rules are routinely used to determine when sampling should finish. Most bird survey methods use (implicit) effort-based stopping rules, either fixed times, fixed sampling areas (quadrats) or both, to standardize samples of different sites. If between-site variation is high, however, a fixed sampling effort will generate samples of variable completeness with samples from smaller, less complex sites being more representative and complete than samples from larger, more complex sites. More importantly, quadrat-based methods shift the scope of the overall study from bird occurrence in sites to bird occurrence in quadrats within sites, diminishing the impact of the research given that results cannot be extrapolated to relevant biological and management scales. Here I advocate an alternative means of conducting bird surveys, whereby the entire site is sampled and a results-based stopping rule is used to ensure sample completeness is uniform across all sites. For example, a researcher may decide to continue sampling each site until two or fewer previously unencountered species are recorded in a 40-min period. Samples of different sites will vary in both area and duration but will all be equivalently accurate estimates of species richness. This approach allows the avifauna of entire sites (whether territories, woodland remnants or catchments) to be sampled and compared directly, generating results and implications at the appropriate scale. In addition to yielding reliable measures of species richness, data collected this way can be used to calculate estimates of sample completeness and species incidence, two valuable metrics for ecological studies. This paper includes detailed worked examples of how to conduct a 'standardized search' and calculate sample completeness and species incidence estimates. I encourage further research on bird survey methods, and suggest that most current methods are insufficient, inconsistent and unreliable.

Key words: birds, inventory, results-based stopping rule, sample completeness, sampling effort, sampling.

INTRODUCTION

An increasing amount of ecological research relies on measures of bird species richness to address a range of questions (Rosenstock *et al.* 2002). Whether comparing revegetation plots of various ages, different habitats or remnants of varying area, bird surveys are a relatively straightforward and efficient means of estimating biodiversity. In addition to research-based projects, these methods are gaining widespread use among agencies and extension groups to measure baseline patterns of diversity and gauge the effectiveness of management practices (Freudenberger 2001). Despite this popularity, attention is rarely given to the adequacy of sampling methods used and the best ways to apply them. Many studies use inappropriate methods, which may lead to large systematic errors in resultant data and, subsequently, erroneous conclusions and misguided management (Verner 1985). Rather than continue applying methods of questionable utility, Verner (1985) highlighted the need for research into alternative approaches. Several new methods or modifications to existing methods have recently been proposed (Mac Nally & Horrocks 2002; Thompson 2002), but many are not applicable to community-based richness estimates (Tarvin *et al.* 1998; Anthony *et al.* 1999) or apply solely to territorial species (Nichols *et al.* 2000; Rosenstock *et al.* 2002), a relatively small subset of Australian terrestrial birds (Recher 1988).

In this contribution, I examine several methodological issues that have been neglected in the ornithological and ecological literature and identify some shortcomings of standard bird surveying methods. In particular, I discuss the frequent mismatch between the scale of sampling and the question/hypothesis being addressed, and critically evaluate the criteria used to determine when sampling should cease (stopping rules). Having explored these issues, I introduce an

---

*Postal address: School of Environmental and Information Sciences, Charles Sturt University, PO Box 789, Albury, New South Wales 2640, Australia.

Accepted for publication March 2003.

alternative approach. Rather than constituting an entirely new surveying technique, it offers a new way to apply existing methods to ensure consistency and maximize the comparability of results.

Throughout this paper, I use bird richness in habitat fragments as examples; this is a topic with which I have some experience (Watson *et al.* 2000; Watson 2002a,b, 2003) and which is currently motivating considerable research. The issues discussed here, however, pertain to field ornithology and ecological sampling more generally, and apply to any comparisons of richness. Although not appropriate for all inventory-based studies, the approach advocated here is a straightforward, objective and repeatable means of estimating species richness of birds (and other organisms), making it a useful addition to the field ecologists' 'tool-box'.

SPATIAL SCALE AND SAMPLING DESIGN

Tailoring the scale and frequency of sampling to the question being addressed is the foundation of effective sampling (Schneider 1994), and has been treated in considerable detail elsewhere (Cochran 1977; Thompson 1992; Krebs 1999). The most popular methods for sampling bird communities use a fixed area (hereafter referred to as a quadrat) to standardize the scale of sampling (Emlen 1971; Verner 1985; Recher 1988). A recent extensive review of 224 international journal articles on terrestrial bird distribution published between 1989 and 1998 found that the majority of studies relied upon these methods, with 46 and 29% of studies using point counts and strip transects, respectively (Rosenstock *et al.* 2002). Quadrats may range in size from 5 m^2 (Nettleship 1976) to 50 km^2 (Udvardy 1981; Bibby *et al.* 1992), with 1–2 ha being typical areas that are sampled by strip transects and point counts for patch-scale ecological studies of terrestrial birds (Recher 1988; Bibby *et al.* 1992; Rosenstock *et al.* 2002). As long as quadrats are representative of the overall study site and are sampled with equivalent effort, the results are statistically comparable.

Given differences between species and sites, however, the results may not reflect actual differences in bird distribution between sites, and thus may not be biologically comparable. Using quadrat-based samples necessarily involves ignoring data; birds are invariably seen or heard outside the quadrat but cannot be included in analyses. Adding these incidental records would improve richness estimates, but would bias the survey towards those species that are more detectable (e.g. frequent, far-carrying calls; Mac Nally & Horrocks 2002; Thompson 2002). Ideally, all species occurring within the patch would eventually be recorded from within the quadrat, but this is often not the case (Mac Nally & Watson 1997; Thompson *et al.* 1998). This may be because the quadrat does not represent the variation contained within the overall site (i.e. quadrat is too small or poorly located), but may also be driven by behavioural differences among species. For instance, some species allow much closer approaches than others and are thus more likely to be represented in quadrat counts, whereas others are more wary and are consistently undercounted or overlooked (Conant *et al.* 1981; Verner 1985; Bibby *et al.* 1992; Thompson 2002).

These concerns are exemplified by a previous study I conducted (Mac Nally & Watson 1997; Watson *et al.* 2000), which examined bird distribution in remnants of Buloke woodlands in western Victoria. Twenty-seven remnants were sampled nine times using 20-min surveys within a 1-ha transect. Sixty-five species were recorded during these transect counts, while a further 29 species were only seen incidentally. Despite being observed in the remnants, they never occurred within a transect during a 20-min survey and therefore could not be included in any analyses. More problematic are those species seen during transect counts in some patches, but only seen incidentally in others. Thus, Laughing Kookaburras (*Dacelo novaeguineae*) were recorded within transects in three patches but were seen incidentally in an additional 13. They often perched outside transect boundaries and were thus consistently undercounted. In contrast, Grey Fantails (*Rhipidura fuliginosa*) were more confiding: they were recorded during transect counts in eight patches, with incidental records from an additional three. Transect placement also confounded representation. The main focus of the study was woodland-dependent birds, so transects were placed at least 20 m from remnant edges. Hooded Robins (*Melanodryas cucullata*) were noted during transect counts in 14 patches and were not recorded incidentally at any other sites. Conversely, Australian Ravens (*Corvus coronoides*) were recorded within transects during surveys in three patches but were seen incidentally in an additional 22, often on the edges of woodland remnants. Thus, differing behaviour and habitat preferences of species can have dramatic effects on the likelihood of being sampled adequately by quadrat-based methods.

An inherent consequence of quadrat-based approaches is that the scope of inquiry is shifted from the habitat patch to the quadrat (noted explicitly by Watson *et al.* 2000). This change in scale may not be problematic. Indeed, many worthwhile questions involving density, habitat associations and effects of changing resource availability can best be addressed in fixed-area quadrats (Merikallio 1958; Emlen 1971; Järvinen & Väisänen 1975; Cochran 1977; Udvardy 1981). It becomes a concern, however, when area is a variable of interest (Mac Nally & Horrocks 2002). Thus, the question is not whether more species occur in larger patches; rather, it becomes whether there are more species in quadrats within larger patches, a very different question with a different set of expectations and hypotheses drawing on a different body of theoretical and empirical literature (Holt 1990; Loman & von Shantz 1991; Mac Nally & Watson 1997). Shifting the scale of inquiry from patch to quadrat prevents the researcher from framing results and conclusions in reference to the patches themselves. Although this need not detract from the inherent value of the research, it does diminish the potential application of the findings. Land managers do not restrict grazing or alter fire regimes in a particular quadrat. They operate at the scale of the patch or region. As such, results of quadrat-based research have a reduced ability to inform management reliably, a major drawback for application-orientated research. When quadrat-based studies are the only information source available, results are often inadvertently extrapolated to the patch scale by the reader. Continuing with the above examples, considering Hooded Robins as present in greater than half of the woodland remnants would be accurate, but noting Laughing Kookaburras or Australian Ravens as occurring in fewer than four remnants would misrepresent known distribution patterns.

The alternative to fixed-area approaches is to use variable-area sampling, often using natural features to define the sampling area such as boundaries of territories, habitat patches and catchments. This approach is especially useful for patch-based studies, enabling the researcher to sample the entire patch (see Brooker 2002). Because the patch is typically the unit of biological and conservation relevance, it makes intuitive sense to use it to define the scale of sampling (Schneider 1994). The key drawback of this approach, and the reason why it has not been applied more widely to community-level investigations is standardization of sample completeness. How can sample completeness be standardized if different sized areas are sampled?

SAMPLING COMPLETENESS

Given well-described behavioural responses, effects of habitat complexity, identification issues and various other confounding factors, it is rarely possible to conduct an accurate census (*sensu* Verner 1985) and quantify all bird species occurring in a given area at any one time (Kendeigh 1944; Shields 1979; Ralph *et al.* 1995; Thompson 2002). Most bird surveys are therefore incomplete. Sampling methods aim to maximize completeness (i.e. approach the actual number of species present at that time) and, more importantly, ensure that the level of completeness is equivalent across all sites being sampled. The amount of time required before a sample can be considered complete varies greatly, with published estimates ranging from more than 6 weeks for an intact lowland rainforest (Beehler *et al.* 1995) to 3 min for small roadside verges (Fortin & Arnold 1997). Strictly speaking, an inventory is complete once every species has been recorded in at least two sampling intervals (Colwell & Coddington 1994), but I am unaware of any ornithological application applying this definitive criterion. In reality, sampling is constrained primarily by time and budgetary restrictions, thus deciding when to stop sampling one site and begin sampling another reflects a balance between required resolution and available resources. These decision-making processes are known collectively as stopping rules (Musa *et al.* 1987; Peterson & Slade 1998), and fall into two broad classes: effort-based and results-based.

Effort-based stopping rules

Most bird survey methods use fixed time-spans for all samples (e.g. 30-min transect count, 20-min intensive search; 5-min point count), thereby standardizing sampling effort across all sites. All birds observed within the specified time period are recorded, often after an initial waiting period to minimize disturbance. Given variation among sites, however, this fixed effort will yield samples of differing completeness. Thus, 20-min samples from smaller, low-quality patches will be more complete (i.e. closer to the actual richness of the patch) than 20-min samples from larger, higher quality areas (Watson *et al.* 2000; Mac Nally & Horrocks 2002).

One solution to this problem is to use proportional sampling, where effort is scaled to patch area or some other measure of patch quality. If scaled correctly, this approach has the potential to generate data of equivalent completeness from patches of varying area and quality. In most studies using this approach, sampling effort is scaled to patch area, where proportionally more time is spent in larger patches (Forman *et al.* 1976; Ambuel & Temple 1983; Arnold & Weeldenberg 1990). This only compounds the problem, however, raising the possibility that samples from larger patches are more complete because total sampling effort was greater. Mac Nally and Horrocks (2002) offered an alternative approach, scaling the area sampled to patch area by using multiple transects but holding total effort constant (known as time-balanced area-proportionate sampling). As with other proportional approaches, this presupposes that (i) area is the most important patch-scale variable affecting diversities and (ii) the number of quadrats is scaled appropriately, both of which are highly equivocal for most systems.

Thus, effort-based stopping rules in general are frequently inappropriate for measuring bird richness (and many other kinds of biological sampling). They compare subsets of the patch that are likely to be unrepresentative or they rely on an arbitrary criterion to terminate sampling in all sites, yielding data of differing completeness that cannot be interpreted at the patch-scale. Post-hoc standardization methods such as rarefaction have been advocated by James and Rathbun (1981) and are especially useful for improving the utility of data already collected using fixed-effort procedures (Young *et al.* 1998). The method entails removing data, ignoring the presence of some species in some sites to ensure data from all sites are equivalently complete. Given the time and cost involved in collecting these data initially, discarding information is both wasteful and methodologically problematic. As with quadrat-based sampling, if a researcher knows that a certain species occurs in a particular patch but then discards that information to standardize the data, this leads to conflict when interpreting and applying the results for management and conservation. These concerns notwithstanding, rarefaction is a powerful standardization technique, and can be used to correct existing data sets to improve comparability (Colwell & Coddington 1994; but see Krebs 1999).

Results-based stopping rules

An alternative to using effort-based stopping rules is to use the actual data to determine when sampling is complete. Although results-based stopping rules are routinely used to standardize quality control in the manufacturing industry (e.g. checking software for bugs; Musa *et al.* 1987) and have been discussed widely in the broader ecological literature (Soberón & Llorente 1993; Colwell & Coddington 1994; Peterson & Slade 1998), they have not been embraced by the ornithological community and are not mentioned in any recent methodological reviews (Gilbert *et al.* 1998; Thompson *et al.* 1998; Rosenstock *et al.* 2002; Thompson 2002). The fundamental approach is simple: decide on a robust rule and apply it during sampling. Once the data satisfy the rule, the survey is complete. Thus, 'stop surveying the patch after 15 min in which two previously unencountered species or fewer are observed', or 'continue surveying the habitat until there are three sequential 20-min periods in which no new species are seen' are both tractable stopping rules. The actual time taken to satisfy the rule is immaterial. Provided sampling methods are applied consistently in all sampling periods at every site, the data will be standardized, representative and above all, equivalently complete (Peterson & Slade 1998).

Note, however, that this approach does not specify which methods or combination of methods are used. The only stipulation is that the same methods are used in each sampling interval at all sites. Thus a results-based stopping rule can be used to determine when a point count is complete, when a transect count should finish or when to stop mist-netting. Rather than an alternative method for sampling bird communities, this approach is an alternative means for deciding when to stop sampling. As Peterson and Slade (1998) pointed out, results-based stopping rules need not require greater total sampling effort, and frequently take less time to complete because they minimize under- and over-sampling patches. They do not generate true censuses, but surveys of a predetermined and consistent completeness. However, the data will still be limited by the methods used, and hence may not accurately represent community composition.

THE 'STANDARDIZED SEARCH'

Applying results-based stopping rules to entire-patch surveys, I propose an alternative way to sample bird communities that is ideally suited to comparing richnesses of terresrial birds in patchy landscapes. Called the 'standardized search', this means of surveying combines the comparability and statistical rigour of quadrat-based methods with the flexibility and spatial relevance of variable-time methods, to yield data of standardized completeness at the patch scale. As such, the approach lends itself to studies of bird distribution in patchy landscapes where patch attributes are variables of interest, potentially solving the 'space–time sampling dilemma' identified by Mac Nally and Horrocks (2002).

In order to use this approach, three factors need to be defined: (i) sampling methods; (ii) interval time; and (iii) stopping rule. Just as with Loyn's (1986) approach, I advocate moving freely throughout the area of inquiry, sampling all habitats and features represented within the study site (see also Brooker 2002). For studies in small patches (<30–50 ha) sampling is best conducted by walking throughout the patch in every sampling period, whereas for larger patches, focusing on non-overlapping subsets during each sampling interval is preferable. As long as equivalent sampling effort is conducted during each interval, resultant data will be directly comparable.

Unlike other bird surveying techniques, sampling is divided into discrete units or sampling intervals, and determining the duration of these intervals is the next step. From experience sampling birds in a range of forested habitats, 15 min is the smallest practicable interval, but for larger forest patches (up to 150 000 ha) I have used sampling intervals of 12 h (resulting in up to 7 full days of continuous sampling before attaining the desired completeness; Watson 2003). For surveys in most sites, however, intervals of between 15 and 60 min should suffice, with longer intervals typically needed in denser habitats or areas with higher species richness. Smaller sampling intervals will yield more highly resolved data, and can easily be combined subsequently into larger intervals if required.

The final step is to decide on an appropriate stopping rule, which will depend on the question being addressed and the amount of time available. Thus the questions ‘how many species are found in this habitat?’ and ‘are more species found in this habitat than that one?’ require rules of differing resolution, with the former invariably needing more data to satisfy. Start by collecting data from two or three of the sites you are seeking to compare (patches of different size, stands of varying age, different habitats) and build up a set of species accumulation curves (cumulative number of species recorded *vs* sample interval; Figs 1 and 2), one for each site. As the data approach the level of completeness required (i.e. as the curve begins to plateau) or as you reach the limit of time you can dedicate to sampling one site, the stopping rule can be defined. In addition to estimating species richness, these curves will reveal the magnitude of the difference between the sites (the ‘effect size’) and the effort required before this difference becomes clear. If this difference is large, a relatively lenient stopping rule can be used but if the difference is small, a stricter rule will be required (in order to yield more complete data).

In formulating a stopping rule, the likelihood of prematurely tripping the rule should be minimized. One solution is to use a compound rule. Thus, ‘stop after three sequential periods in which a total of two new species or fewer are encountered’ is more robust than ‘stop once one new species or fewer is encountered in a single sample’, and therefore less liable to be affected by weather, time of day or other confounding factors. Even with these compound rules, however, sample order can affect when sampling stops, particularly when all samples of one site are arranged consecutively. An alternative is to express stopping rules in terms of the ratio of singletons (species seen in only one sampling period) to doubletons (species seen in two sampling periods): the higher this ratio is, the greater the likelihood of the sample being incomplete (applicable only for more than two sampling periods, otherwise all species will be singletons or doubletons). Rules of this type are routinely used to predict total species richness based on inventory data (see the 'Quantifying sample completeness' section) and have been subjected to considerable quantitative evaluation (Colwell & Coddington 1994; Peterson & Slade 1998; Herzog *et al.* 2002; Longino *et al.* 2002), assessing how various equations behave under different conditions. Moreover, these rules are relatively easy to apply: at the end of each sample calculate this ratio and if the value is lower than your threshold, sampling is complete.

In a recent study that explored the long-term consequences of habitat fragmentation, data were collected on the distribution of forest-dependent birds in 17 fragments of humid pine–oak forest in Oaxaca, southern Mexico (Watson 2003). Despite ranging in area from 2 ha to almost 160 000 ha, the forest fragments were all sampled to the same degree of completeness. To the best of my knowledge, this is the only published study to use an explicit results-based stopping rule to standardize samples of avian richness, and the first application of the standardized search. Given the area of some of these forest fragments, sampling intervals were half-days (pre-dawn to midday, midday to after dusk), and sampling was terminated once two consecutive sampling periods yielded a total of zero or one previously unrecorded

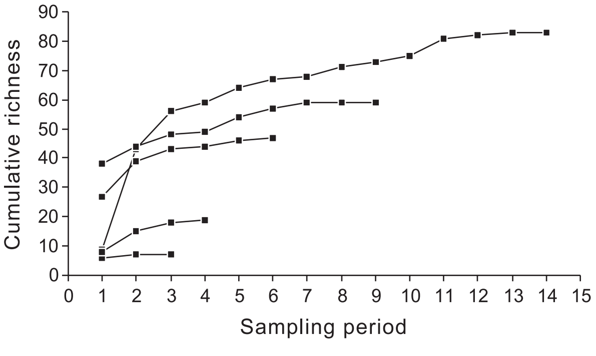

Fig. 1. Species accumulation curves for five fragments of humid pine-oak forest in Oaxaca Mexico, ranging in area from 2 ha to almost 70 000 ha with total observed avian richness ranging from 7 to 83 (Watson 2003). The same stopping rule was used to determine when each site had been sampled adequately, thereby ensuring equivalent completeness for inventories of all sites.

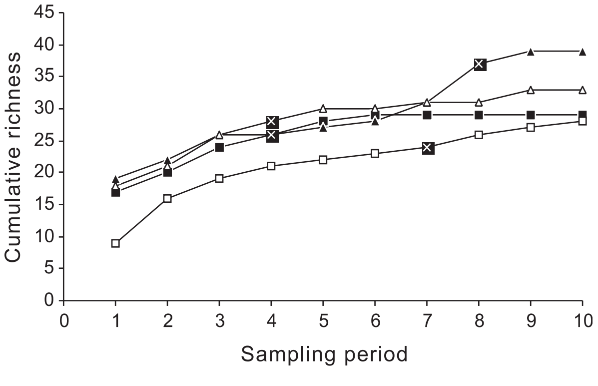

Fig. 2. Species accumulation curves for four inventories of woodland bird species for two adjacent remnants of grassy box woodland in southern New South Wales, Australia, with manipulated mistletoe abundance. (△, ▲), control; (■, □), reduced; (△, □), spring 2000 survey; (▲, ■) summer 2001 survey. Equivalent effort was used for all four inventories, each comprising 10 1-h surveys (Watson 2002a). The cross indicates the sampling period when the inventory satisfied the a posteriori stopping rule, that is, when the ratio of singletons to doubletons ≤1.2.

forest-dependent species. During each sampling interval, I walked throughout the forest both on and off trails, taking a different route each period in all but the smallest patches. Inventories ranged from 14 sampling periods (i.e. 7 days of continuous sampling) to three sampling periods (i.e. 1.5 days), but yielded samples of equivalent completeness (Fig. 1; Watson 2003). Few other methods can be used to sample patches spanning five orders of magnitude in area, highlighting the utility of the standardized search for research into distribution patterns of birds in patchy landscapes.

WORKED EXAMPLE: AVIFAUNA OF TWO WOODLAND REMNANTS IN EASTERN AUSTRALIA

In a recent study of the effect of mistletoe on bird diversity, surveys were conducted in two adjacent woodland remnants in southern New South Wales, Australia. The remnants were similar in all respects except mistletoe density (most mistletoe plants had been removed from one remnant by the landholder; Watson 2002a). Because I was interested in estimating

Table 1. Complete inventory for a 36-ha remnant of grassy box woodland (summer 2001, reduced mistletoe) using the 'standardized search', drawn from Watson (2002a)

		Sampling period									
Common name	Scientific name	1	2	3	4	5	6	7	8	9	10
Australian Magpie	*Gymnorhina tibicen*	×	×	×	×	×	×	×	×	×	×
Eastern Yellow Robin	*Eopsaltria australis*	×	–	–	–	–	–	×	–	–	×
Grey Shrike-thrush	*Colluricincla harmonica*	×	–	×	×	×	×	×	×	–	–
Noisy Friarbird	*Philemon corniculatus*	×	×	×	×	×	×	×	×	×	×
Grey Fantail	*Rhipidura fuliginosa*	×	×	×	×	×	×	×	×	×	×
Buff-rumped Thornbill	*Acanthiza reguloides*	×	×	×	×	×	×	×	×	×	×
White-throated Gerygone	*Gerygone olivacea*	×	×	×	×	×	×	×	×	×	×
Crimson Rosella	*Platycercus elegans*	×	×	×	×	×	×	×	×	×	×
White-throated Treecreeper	*Cormobates leucophaeus*	×	×	×	×	×	×	×	×	×	×
Sacred Kingfisher	*Todiramphus sanctus*	×	–	×	–	–	–	×	–	×	–
Striated Thornbill	*Acanthiza lineata*	×	×	–	–	–	×	–	–	×	–
Rufous Whistler	*Pachycephala rufiventris*	×	×	×	×	×	×	×	×	×	×
Leaden Flycatcher	*Myiagra rubecula*	×	×	×	×	×	×	×	×	×	×
Varied Sitella	*Daphoenositta chrysoptera*	×	–	–	×	×	×	×	×	–	×
Striated Pardalote	*Pardalotus striatus*	×	×	×	–	×	×	×	×	×	×
Noisy Miner	*Manorina melanocephala*	×	×	×	–	×	×	×	×	×	×
Little Pied Cormorant*	*Phalacrocorax melanoleucos*	×	–	–	–	–	–	×	–	–	–
Eastern Rosella	*Platycercus eximius*	×	×	×	×	×	×	–	×	×	×
Laughing Kookaburra	*Dacelo novaeguineae*	×	×	×	×	×	×	×	–	×	×
Wood Duck*	*Chenonetta jubata*	×	–	–	–	–	×	–	–	–	–
Sulphur-crested Cockatoo*	*Cacatua galerita*		×	×	×	×	×	×	×	×	×
Common Bronzewing	*Phaps chalcoptera*		×	–	×	–	–	–	–	–	–
Superb Parrot	*Polytelis swainsonii*		×	×	×	–	×	×	×	×	×
Spotted Pardalote	*Pardalotus punctatus*			×	–	–	×	×	–	–	×
Red-rumped parrot	*Psephotus haematonotus*			×	–	–	–	–	–	×	×
Black-faced Cuckoo-shrike	*Coracina novaehollandiae*			×	×	×	×	×	–	×	×
Brown Thornbill	*Acanthiza pusilla*			×	–	–	–	–	–	–	–
Painted Button-quail	*Turnix varia*				×	–	–	–	–	–	×
Yellow-rumped Thornbill	*Acanthiza chrysorrhoa*				×	–	–	–	–	–	–
Galah*	*Cacatua roseicapilla*				×	×	–	×	–	×	–
Crested Pigeon*	*Ocyphaps lophotes*					×	–	–	–	–	–
Magpie-lark	*Grallina cyanoleuca*					×	×	–	×	×	×
Australian Raven	*Corvus coronoides*					×	–	–	×	–	–
Shining Bronze-cuckoo	*Chrysococcyx lucidus*						×	–	–	–	–
Dollarbird*	*Eurystomus orientalis*						×	–	–	×	×
Tree Martin*	*Hirundo nigricans*							×	–	–	–
Common Skylark*	*Alauda arvensis*							×	–	–	–

The numbers 1–10 correspond to 10 1-h samples, conducted on 5–6 January 2001, with species listed in order of observation. ×, Presence (at least one individual was seen or heard within the remnant during that sampling period); –, absence. *Waterbirds, aerial foragers and open country species that were removed prior to analysis of woodland species richness.

the actual richness of bird assemblages, I sampled both patches with 10 1-h surveys in both spring and summer, yielding a total of 40 h of surveys. Full data for one inventory are presented (summer 2001, reduced mistletoe; Table 1), with all four inventories depicted graphically (Fig. 2). Rather than simply listing all species seen, sampling is divided into discrete intervals, 1 h in this case, with a separate list recorded for each interval (Table 1). Thus for the first period, all species encountered are noted as present in interval one (in this case, 20 species). For interval two, a second list is produced. All species previously seen are noted as either present or absent, and any previously unencountered species added (in this case, three). Sampling thus continues until the stopping rule is satisfied. For this study, effort was standardized to 10 h for each inventory, alternating between the two remnants to minimize any time-of-day effects, but a stopping rule can be applied to the data a posteriori. Thus, 'the survey of woodland birds is complete once the ratio of singletons to doubletons is ≤1.2' represents a realistic stopping rule, easily calculated at the end of each sample. For this survey, the data would satisfy the rule after four samples, whereas the other three surveys would be similarly complete after four, seven and eight samples (see Table 2, Fig. 2). Thus a total of 23 h is needed to obtain four surveys of equivalent completeness, almost half of the actual effort expended. Comparing these standardized richness estimates with values after the full 10 h inventories reveals high consistency, with all four estimates being between 84 and 95% of total measured richness (Table 2), and exhibiting the same relative pattern (Fig. 2).

In addition to differences in methods, it is interesting to compare the sampling effort used in the present study with survey durations typically used in studies of woodland birds in Australia. For the inventory presented here (Table 1), 20 species were recorded during the first hour, subsequent hourly counts ranged from 17 to 24, 27 species were recorded from the first 3 h, and a total of 7 h was required to reach the maximum recorded richness for the site (37 species; Table 1). An additional 15 species were seen during the other 10 h inventory conducted in summer, with new woodland species added in the final sampling interval (Fig. 2). Thus, after 20 h of surveys, 52 species were recorded from the site, including 38 woodland species, with 46 woodland species recorded from the adjoining site (Watson 2002a). These two sites were remnants of grassy box woodland, 36 and 33 ha in area, and are representative of many sites typically surveyed using 3–10 × 20–30 min surveys (Recher 1988; Er *et al.* 1995; Slater 1995; Kutt 1996; Watson *et al.* 2000; Major *et al.* 2001; Mac Nally & Horrocks 2002), however, recorded richness is substantially higher than reported for equivalent sites. This suggests that commonly used sampling efforts may often be inadequate to gain an accurate estimate of bird richness in many studies. Even if relative richness is sought, note that differences between the two sites only became apparent after 8 h for the summer inventories (Fig. 2). Regardless of the sampling methods used, researchers are strongly advised to conduct an exhaustive survey such as this as part of their preliminary fieldwork (Verner 1985), both to determine rates of species accumulation and to ensure that they are aware of how complete and representative their samples are.

DENSITY, RELATIVE ABUNDANCE AND INCIDENCE

So far, I have restricted discussion to species richness, a nominal estimate of how many species occur in a particular area. Other variables that are useful to avian ecologists are density and abundance, ordinal indices that yield a more fine-grained measure of distribution than richness and therefore are more useful for revealing subtle changes in distributional patterns. Density is used more frequently than abundance because it is scaled relative to area and is more useful for between-site comparisons (Verner 1985). All the issues that affect the accurate measurement of species richness are amplified when trying to estimate densities. Having comprehensively reviewed studies using density estimates of birds, Verner (1985; p. 295) commented 'Sample sizes are inadequate; sampling procedures are unbalanced; analyses that assume independent samples

Table 2. Summary of woodland species richness values for the four inventories depicted in Fig. 2

Inventory	Samples to satisfy rule (*n*)	Singletons : doubletons	Richness after *n* samples (*Rn*)	Richness after 10 samples (*R*)	% Completeness (*Rn*/*R*.100)
Reduced 2001	4	5 : 5, 1.0	26	29	89.7
Reduced 2000	7	6 : 5, 1.2	24	27	88.9
Control 2001	8	7 : 7, 1.0	37	39	94.9
Control 2000	4	6 : 6, 1.0	28	33	84.8

These data were subjected to the following stopping rule a posteriori: 'the survey of woodland birds is complete once the ratio of singletons to doubletons is ≤1.2'. Between four and eight 1-h surveys were needed to satisfy this rule, but resultant values of sample completeness are similar for all inventories.

are used when samples are obviously not independent; even recommended standards and guidelines, although cited, are often ignored'. More recently, several reviews have highlighted the difficulties associated with estimating avian density, with factors including apparency and detectability known to differ among species, sites, seasons, and surveyors (Nichols *et al.* 2000; Pagen *et al.* 2002; Rosenstock *et al.* 2002).

Although density cannot be calculated from standardized search data, an alternative measure can be derived that is applicable to many community-level studies. Incidence (or reporting rate) is the number of times a species was encountered (Dawson 1981), that is, the number of intervals in which it was recorded divided by the total number of sampling intervals. This variable is prone to many of the same shortcomings as density, however, and the data used to calculate it are not independent (the presence of a species in one sample at a site is influenced by whether it was there in a previous sample). Nonetheless, given that hearing an individual bird once and seeing large flocks repeatedly during a sampling period are treated as equivalent (both qualify as a presence), this metric is relatively coarse-grained and therefore less affected by behavioural and site-based factors that confound density estimates. Using the data from Table 1, nine species had an incidence of 1 (10 out of 10), five species had an incidence of 0.9, and 11 species had an incidence of 0.1–0.2. Rather than translating to an estimate of true density or abundance (relative or absolute), incidence can be considered an indication of preference or site faithfulness (Dawson 1981; Verner 1985). Hence, of the three species of thornbill recorded during this inventory, it is reasonable to suggest that the site is most important for Buff-rumped Thornbills (incidence of 1), of intermediate importance for Striated Thornbills (0.4) and of least importance for Brown Thornbills (0.1).

If estimates of density are central to the aims of the study, however, the standardized search is not appropriate, nor are most other community-based bird survey methods currently available. Intensive territory mapping, ideally in combination with mark-recapture methods can yield robust density estimates, but they are generally only tractable for single-species studies, and even then are often subject to untestable assumptions (Verner 1985; Bibby *et al.* 1992; Remsen & Good 1996).

QUANTIFYING SAMPLE COMPLETENESS

For some applications, it may not be enough to ensure that the data are equivalently complete, an actual estimate of sample completeness is required. This can be readily derived from predicted total richness of the site, calculated from the inventory data using various estimator procedures (Soberón & Llorente 1993; Colwell & Coddington 1994; Peterson & Slade 1998). Although frequently used by invertebrate biologists to provide an indication of the total richness in a given area (Longino *et al.* 2002), the use of these predictive methods has been more limited by vertebrate biologists, primarily because *which* species are present is often regarded as more important than how many (see MacKinnon & Phillipps 1993). Nonetheless, some recent research has demonstrated the utility of these methods in extrapolating bird richness based on rapid assessment data to generate rapid, robust total richness estimates for tropical areas where complete inventories are prohibitively time-consuming (Herzog *et al.* 2002). In addition to generating predictions of total richness

Table 3. Estimates of sample completeness derived from inventory data presented in Table 1; each sample is 1-h long, thus the number of samples corresponds to the number of hours

Measure	Number of samples	Actual richness (*A*)	Predicted richness* (*P*)	% Completeness, *A*/*P*.100
Total species richness	2	23	27.1	85.0
	3	27	31.3	86.4
	4	30	34.1	83.2
	5	33	36.7	89.9
	6	35	38.6	90.7
	7	37	40.3	91.9
Woodland species richness	2	20	28.1	71.2
	3	24	28.6	83.9
	4	26	29.0	89.6
	5	28	29.4	95.2
	6	29	29.7	97.6
	7	29	29.9	96.8

*Predicted species richness using the Michaelis–Menten richness estimator computed for mean species accumulation curve (Mmeans, Collwell 2001) using 1000 runs, Estimates 6.0b1.

for conservation initiatives, these estimators can be used to quantify sample completeness by dividing observed richness by estimated total richness, yielding percentages analogous to confidence intervals.

Of the many estimators available, the Chao equations are probably the simplest and have gained the widest use (Soberón & Llorente 1993; Peterson & Slade 1998; Colwell 2001). As with most other estimators, they use the ratio of singletons (species recorded in one sample only) to doubletons (species recorded in two samples only) to estimate the number of species that were not recorded. The greater the number of singletons, the higher is the likelihood of species being present but not recorded. Some recent comparative research suggests the Michaelis–Menten equations may be even more robust, despite being slightly more complex to compute (Herzog *et al.* 2002; Longino *et al.* 2002).

Using the data summarized in Table 1, predicted richness was calculated for total species and woodland species after two to seven 1-h samples using the Mmeans estimator (Colwell 2001; Table 3). Dividing observed richness values by predicted richness yields iterative estimates of sample completeness (Table 3). As with most other estimator procedures used by Colwell (2001), the Mmeans procedure uses resampling to prevent sample order from affecting the prediction (I used 1000 iterations here). Note that sample completeness increases rapidly, exceeding 95% after five samples, half of the 10-h inventory. A stopping rule can be designed based on these estimates to ensure that samples of all other sites attain a similar level of sample completeness (e.g. terminate sampling once estimated completeness exceeds 90%). Although this involves some brief calculations, it can be done readily in the field, enabling researchers to collect samples of known completeness quickly and repeatably. At low sample sizes, some of these estimators become unstable, and researchers interested in using them are advised to become familiar with the assumptions of the various methods (Colwell & Coddington 1994; Herzog *et al.* 2002).

CONCLUSIONS AND PROSPECT

Most inventory-based studies of bird communities rely exclusively on point counts or transects: fixed effort methods that generate data of different (and unknown) completeness. Although useful for many aspects of quantitative bird ecology, the data they yield are not representative of the overall site and so they cannot be used to discuss patterns, processes or effects of management at the patch scale. Moreover, typical sampling efforts used (20–30 min or, rarely, 1 h) may be inadequate for all but the smallest or least complex sites. With transect-based studies increasingly used to inform land management, these shortcomings are important and should be given careful consideration when designing a study and interpreting findings from previous studies.

Although results-based stopping rules have been advocated by previous workers as a reliable means of ensuring the comparability of biodiversity inventories (Soberón & Llorente 1993; Peterson & Slade 1998), they have not gained widespread use. One of the main reasons for this, I suspect, is that resultant inventories of variable areas and variable durations seem inconsistent, or somehow not comparable compared with surveys of fixed areas over fixed times. The time taken to conduct a survey is immaterial: as long as surveying methods and sample completeness are equivalent then the inventories are directly comparable and can be analysed using the same analytical tools and statistical tests that are routinely used to infer patterns from quadrat-based data. Having accepted the fact that bird surveys are necessarily incomplete, the next logical step is to ensure that all samples are equivalently complete and the standardized search can help achieve that aim. In addition to generating standardized representative richness data of equivalent (and known) completeness, estimates of incidence can be calculated allowing fine-grained comparison of bird distribution across the study sites. Importantly, most of the factors that confound other sampling approaches, including habitat heterogeneity, secretive species and weather effects (Craig & Roberts 2001), will serve only to extend the amount of effort required before satisfying the stopping rule.

Regardless of whether this approach becomes widely adopted or not, it is essential that careful thought be given to effect size, sample completeness and differences between absolute and relative richness measures; these are critical issues that are often overlooked. Before accepting the standardized search or any other approach as the preferred bird survey method, conduct your own comparisons as part of pilot work in your study area. You can then choose the method or suite of methods that best suits the aims of your research, rather than assuming that a commonly used technique is most appropriate. Finally, I encourage researchers to refine the standardized search, adapt it to suit their needs and contribute to the growing literature on ornithological sampling methods, thereby maximizing the validity and utility of research findings.

ACKNOWLEDGEMENTS

Initial discussions with Andrew (Townsend) Peterson, Jorge Soberón and Norm Slade provided the conceptual basis to this paper, for which I am most grateful. Subsequent input from David Roshier, contributions to my reasoning from Maggie Watson, and comments from three anonymous referees helped clarify the manuscript. Discussions with students in the Graduate Certificate of Ornithology at Charles Sturt University helped to crystallize the thoughts presented here.

REFERENCES

Ambuel B. & Temple S. A. (1983) Area-dependent changes in the bird communities and vegetation of southern Wisconsin forests. *Ecology* **64**, 1057–68.

Anthony R. G., Garrett M. G. & Isaacs F. B. (1999) Double-survey estimates of Bald Eagle populations in Oregon. *J. Wildl. Manage.* **63**, 794–802.

Arnold G. W. & Weeldenberg J. R. (1990) Factors determining the number and species of birds in road verges in the wheatbelt of Western Australia. *Biol. Conserv.* **53**, 295–315.

Beehler B. M., Sengo J. B., Filardi C. & Merg K. (1995) Documenting the lowland rainforest avifauna in Papua New Guinea: Effects of patchy distributions, survey effort and methodology. *Emu* **95**, 149–61.

Bibby C. J., Burgess N. D. & Hill D. A. (1992) *Bird Census Techniques*. Academic Press, London.

Brooker L. (2002) The application of focal species knowledge to landscape design in agricultural lands using the ecological neighbourhood as a template. *Landscape Urban Plan.* **60**, 185–210.

Cochran W. G. (1977) *Sampling Techniques*, 3rd edn. Wiley, New York.

Colwell R. K. (2001) *Estimates: Statistical Estimation of Species Richness and Shared Species from Samples*, Version 6.0b1 [Internet document; updated September 2002; cited 25 September 2002]. Available from: http://viceroy.eeb.uconn.edu/estimates

Colwell R. K. & Coddington J. A. (1994) Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. R. Soc. Lond. B.* **345**, 101–18.

Conant S., Collins M. S. & Ralph C. J. (1981) Effects of observers using different methods upon the total population estimates of two resident birds. *Stud. Avian Biol.* **6**, 12–16.

Craig M. D. & Roberts J. D. (2001) Evaluation of the impact of time of day, weather, vegetation density and bird movements on outcomes of area searches for birds in eucalypt forests of south-western Australia. *Wildl. Res.* **28**, 33–9.

Dawson D. G. (1981) The usefulness of absolute ('census') and relative ('sampling' or 'index') measures of abundance. *Stud. Avian Biol.* **6**, 554–8.

Emlen J. T. (1971) Population densities of birds derived from transect counts. *Auk* **88**, 323–42.

Er K. B. H., Robinson A. P. & Tidemann C. R. (1995) Importance of sampling duration and strip width in use of the fixed-width strip transect method for estimation of bird abundance and species diversity. *Corella* **19**, 109–14.

Forman R. T. T., Galli A. E. & Leck C. F. (1976) Forest size and avian diversity in New Jersey woodlots with some land use implications. *Oecologia* **26**, 1–8.

Fortin D. & Arnold G. W. (1997) The influence of road verges on the use of nearby small shrubland remnants by birds in the central wheatbelt of Western Australia. *Wildl. Res.* **24**, 679–89.

Freudenberger D. (2001) *Bush for the Birds. Biodiversity Enhancement Guidelines for the Saltshaker Project, Booroowa, NSW.* Report for Greening Australia ACT & SE NSW Inc. CSIRO, Canberra.

Gilbert G., Gibbons D. W. & Evans J. (1998) *Bird Monitoring Methods: A Manual of Techniques for Key UK Species*. Royal Society for the Protection of Birds, Bedfordshire.

Herzog S. K., Kessler M. & Cahill T. M. (2002) Estimating species richness of tropical bird communities from rapid assessment data. *Auk* **119**, 749–69.

Holt R. D. (1990) A neglected facet of island biogeography: The role of internal spatial dynamics in area effects. *Theor. Pop. Biol.* **41**, 354–71.

James F. C. & Rathbun S. (1981) Rarefaction, relative abundance, and diversity of avian communities. *Auk* **98**, 785–800.

Järvinen O. & Väisänen R. A. (1975) Estimating relative densities of breeding birds by the line transect method. *Oikos* **26**, 316–22.

Kendeigh S. C. (1944) Measurement of bird populations. *Ecol. Monogr.* **14**, 67–106.

Krebs C. J. (1999) *Ecological Methodology*, 2nd edn. Addison-Wesley Longman, Menlo Park, CA.

Kutt A. S. (1996) Bird populations density in thinned, unthinned and old lowland regrowth forest, East Gippsland, Victoria. *Emu* **96**, 280–4.

Loman J. & von Shantz T. (1991) Birds in a farmland: More species in small than in large habitat islands. *Conserv. Biol.* **5**, 176–86.

Longino J. T., Coddington J. & Colwell R. K. (2002) The ant fauna of a tropical rain forest: Estimating species richness three different ways. *Ecology* **83**, 689–702.

Loyn R. H. (1986) The 20 minute search: A simple method for counting forest birds. *Corella* **1**, 59–60.

MacKinnon S. & Phillipps K. (1993) *A Field Guide to the Birds of Borneo, Sumatra, Java and Bali.* Oxford University Press, Oxford.

Mac Nally R. & Horrocks G. (2002) Proportionate spatial sampling and equal-time sampling of mobile animals: A dilemma for inferring areal dependence. *Austral Ecol.* **27**, 405–15.

Mac Nally R. & Watson D. M. (1997) Distinguishing area and habitat heterogeneity effects on species richness: Birds in Victorian buloke remnants. *Aust. J. Ecol.* **22**, 227–32.

Major R. E., Christie F. J. & Gowing G. (2001) Influence of remnant and landscape attributes on Australian woodland bird communities. *Biol. Conserv.* **102**, 47–66.

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

Musa J. D., Iannino A. & Okumoto K. (1987) *Software Reliability: Measurement, Prediction, Application.* McGraw-Hill, New York.

Nettleship D. N. (1976) Census techniques for seabirds of arctic and eastern Canada. *Can. Wildl. Service Occ. Pap.* **25**, 1–33.

Nichols J. D., Hines J. E., Sauer J. R., Fallon F. W., Fallon J. E. & Heglund P. J. (2000) A double-observer approach for estimating detection probability and abundance from point counts. *Auk* **117**, 393–408.

Pagen R. W., Thompson F. R. III & Burhans D. E. (2002) A comparison of point-count and mist-net detections of songbirds by habitat and time-of-season. *J. Field Orn.* **73**, 53–9.

Peterson A. T. & Slade N. A. (1998) Extrapolating inventory results into biodiversity estimates and the importance of stopping rules. *Divers. Distrib.* **4**, 95–105.

Ralph C. J., Droege S. & Sauer J. R. (1995) Managing and monitoring birds using point counts: Standards and applications. In: *Monitoring Bird Populations by Point Counts* (eds C. J. Ralph, J. R. Sauer & S. Droege) pp. 161–70. United States Forest Service General Technical Report PSW-GTR-149. United States Forest Service.

Recher H. F. (1988) Counting terrestrial birds: Use and application of census procedures in Australia. *Aust. Zool. Rev.* **1**, 25–45.

Remsen J. V. Jr & Good D. A. (1996) Misuse of data from mist-net captures to assess relative abundance in bird populations. *Auk* **113**, 381–98.

Rosenstock S. S., Anderson D. R., Giesen K. M., Leukering T. & Carter M. F. (2002) Landbird counting techniques: Current practices and an alternative. *Auk* **119**, 46–53.

Schneider D. C. (1994) *Quantitative Ecology: Spatial and Temporal Scaling*. Academic Press, San Diego.

Shields W. M. (1979) Avian census techniques: An analytical review. In: *The Role of Insectivorous Birds in Forest Ecosystems* (eds J. G. Dickson, R. N. Connor, R. R. Fleet, J. C. Kroll & J. A. Jackson) pp. 23–51. Academic Press, New York.

Slater P. J. (1995) The interaction of bird communities with vegetation and season in Brisbane Forest Park. *Emu* **95**, 194–207.

Soberón M. J. & Llorente B. J. (1993) The use of species accumulation functions for the prediction of species richness. *Conserv. Biol.* **7**, 480–8.

Tarvin K. A., Garvib M. C., Jawor J. M. & Dayer K. A. (1998) A field evaluation of techniques to estimate density of blue jays. *J. Field Orn.* **69**, 209–22.

Thompson S. K. (1992) *Sampling*. Wiley, New York.

Thompson W. L. (2002) Overview: Towards reliable bird surveys: Accounting for individuals present but not detected. *Auk* **119**, 18–25.

Thompson W. L., White G. C. & Gowan C. (1998) *Monitoring Vertebrate Populations*. Academic Press, San Diego.

Udvardy M. D. F. (1981) An overview of grid-based atlas works in ornithology. *Stud. Avian Biol.* **6**, 103–9.

Verner J. (1985) Assessment of counting techniques. In: *Current Ornithology* **Vol. 2** (ed. R. F. Johnstone) pp. 247–301. Plenum Press, New York.

Watson D. M. (2002a) Effects of mistletoe on diversity: A case-study from southern New South Wales. *Emu* **102**, 276–81.

Watson D. M. (2002b) A conceptual framework for studying species composition in fragments, islands and other patchy ecosystems. *J. Biogeogr.* **29**, 823–34.

Watson D. M. (2003) Long-term consequences of habitat fragmentation: Highland birds in Oaxaca, Mexico. *Biol. Conserv.* **111**, 283–303.

Watson D. M., Mac Nally R. & Bennett A. F. (2000) The avifauna of severely fragmented, Buloke *Allocasuarina luehmannii* woodland in western Victoria, Australia. *Pac. Conserv. Biol.* **6**, 46–60.

Young B. E., DeRosier D. & Powell G. V. N. (1998) Diversity and conservation of understory birds in the Tilarán mountains, Costa Rica. *Auk* **115**, 998–1016.