Increasing the value of bird-habitat studies in tropical forests: choice of approach and habitat measures

DAVID C. LEE and STUART J. MARSDEN

Summary

An important component of many conservation studies is the assessment of bird-habitat relationships, but limited resources often lead to constraints on study design, quality and quantity of bird data, and restrict the number and types of habitat variables gathered. The aim of this study was to identify habitat features that were both relatively easy and quick to collect and powerful in identifying bird-habitat relationships. We also discuss some issues with our study and alternative approaches that may help in future bird-habitat studies in tropical forests. Twenty-four habitat measures representing geographical (e.g. altitude, topography, X and Y coordinates), vegetation structure (e.g. tree sizes), and tree floristics (abundance of 28 indicator tree species) features were collected in association with bird presence/absence data from point transects within a 1,500 ha Philippine lowland forest. We used hierarchical partitioning of regression analyses to assess which of these geographical and structural variables along with four floristics axes derived from DECORANA were the most important variables for explaining the occurrence of individual bird species and guilds. The ten most powerful variables for a range of bird species included seven geographical and three floristic variables, while the ten least important were all structural variables. There were differences in importance of individual variables across guilds, with, for example, floristics very important in canopy frugivores, and geographical variables more important for upperstorey gleaning insectivores. We stress the importance of geographical variables in linking birds to habitat at this local scale, but also suggest that efforts are made to collect some floristics data, perhaps a subset of species that represent resources for birds (e.g. *Ficus* spp.), people's use of the forest (e.g. dipterocarps), and indicators of forest type. While the habitat variables and approach in this study adequately identified bird-habitat relationships for most species, we suggest improvements and alternative methods that may improve results in other studies.

Introduction

Understanding bird-habitat relationships is often a requirement of conservation studies (Bibby *et al.* 1998), allowing us to gauge species' responses to habitat alteration (MacNally and Bennett 1997), and to predict the occurrence of species at unsurveyed sites (Miller and Cale 2000, Suarez-Seoane *et al.* 2002). Bird-habitat studies at the local scale generally require detailed measurements of the physical (Pearson 1975), structural (MacArthur *et al.* 1962), and vegetation composition or floristics (Rotenberry 1985) of sites. These are then used to assess local patterns in species richness (e.g. Rotenberry 1985, Berry and Bock 1998), or the occurrence (e.g. James 1971, Marsden and Fielding 1999), relative abundance (e.g. Farina 1997, Kirk and Hobson 2001), or density (e.g. Mills *et al.* 1991, Doherty and Grubb 2000) of individual species.

Bird-habitat relationships are thought to be closely correlated with structural characteristics (MacArthur *et al.* 1962, James 1971, Cody 1981) or plant species composition of the habitat (Wiens and Rotenberry 1981, Rotenberry 1985), and may work alongside physical factors acting on physiological limitations (Karr and Freemark 1983). The relative importance of each habitat factor in structuring bird communities does not seem to be fixed (Rotenberry 1985, MacNally 1990), with individual bird species exhibiting unique responses to habitat attributes (MacNally 1990) and to habitat alteration (e.g. Dellasala *et al.* 1996, Twedt *et al.* 1999). In the context of short, multi-species conservation surveys, comprehensive characterisation of bird-habitat associations may be difficult since obtaining fine-grained assessments of habitat structure or composition can be expensive and labour-intensive (e.g. Fleishman *et al.* 2002). We may not be able to gather bird data of sufficient quantity and quality to allow accurate modelling, and may be constrained by shortfalls in study design, or lack of knowledge about the ecology of the study species. Importantly, given that resources for field research are usually limited, it is crucial to record habitat data relatively quickly, while at the same time choosing the right habitat measures to be able to detect important bird-habitat relationships (Bibby *et al.* 1992).

In this paper, we assess the relative capacity of a range of geographical, vegetation structure, and floristic variables to explain the local occurrence of individual species and groups of species. We also considered the 'costs' of collecting each habitat variable, and autocorrelations among variables, to allow us to recommend guild-specific suites of tropical forest habitat variables that are both independently powerful in explaining bird-habitat relationships and are easy to collect. We also discuss briefly the performance of our bird-habitat methods and suggest alternative or complementary approaches.

Methods

Study area

Data were collected at the 1,500 ha Mt. Siburan Important Bird Area (12°48′ N 120°55′ E; 50–420 m; Mallari *et al.* 2001), the largest remaining tract of lowland forest on the island of Mindoro, Philippines. The site consists of one lowland dipterocarp forest block, that comprises closed-canopy primary forest (mature trees covering > 50%), secondary forest, and second growth scrub. These latter habitats are a result of previous commercial logging that ceased in c.1960 (N.A.D. Mallari, pers. comm. 2003), and current small-scale activities in some areas of the forest (Condeno and Gaerlan 2002, D. Lee pers. obs. 2002, 2003).

The site is surrounded by cultivated land mixed with scrub and grassland, both natural and resulting from shifting agriculture (Condeno and Gaerlan 2002). Several patches of open-canopy forest (mature trees covering < 50%), and a few fragments of closed-canopy forest (SSC 1988, Evans *et al.* 1993, CGCEO 2001, D. Lee pers. obs. 2002) are found within this landscape but are not connected to the study site.

Bird survey data collection and analysis

Straight line transects were positioned at the site using a random systematic approach to proportionally represent (and stratify by) broad habitat types at the site (Buckland *et al.* 1993) identified from satellite imagery (SSC 1988, CGCEO 2001). Three broad habitat types were identified: primary forest, secondary forest and second growth. Transects were sited to largely avoid any edge effects associated with disproportionately sampling along roads, paths, ridge tops or watercourses (Jones 1998).

We collected bird data from point transects (Buckland *et al.* 1993) positioned every 200 m along transects of 1.6–2.0 km length. In total, 128 points, which can be regarded as a random sample, were surveyed along 13 transects and revisited four times ($n = 512$) during the main breeding season (April–June) on Mindoro (Dickinson *et al.* 1991). All distance data for bird

encounters were right-truncated at 50 m. This removed outlying bird records and maintained an association between bird encounters and the habitat features surrounding each point. A bird species was recorded as 'present' at a point transect if it was detected, either aurally or visually, at least once from the four ten-minute survey repeats. Only species encountered from > 5% of point transects were considered for analysis ($n = 39$ species).

Bird species were grouped into guilds broadly following the scheme of Karr *et al.* (1990), which includes broad dietary type, main foraging method, and principal vegetation stratum used, but also including detection-based characteristics such as movement rates, foraging behaviour, calling rates, and body size (for details see Lee unpublished PhD Thesis 2005). The final arrangement comprised seven guilds: 'Canopy frugivores' ('CF'), 'Coucals' ('CL'), 'Ground-dwellers' ('GD'), 'Omnivores' ('O'), 'Nectarivores' ('N'), 'Understorey insectivores' ('UI'), and 'Upperstorey gleaning insectivores' ('UGI'). A guild was recorded as present at a point if any member species was recorded at that point.

Habitat data collection and analysis

Three groups of habitat variables were collected within a 20 m radius (unless otherwise stated in Table 1) of each survey point: geographical variables, vegetation structural characteristics, and tree species composition (floristics) (Table 1). Therefore, habitat variables were measured across the same plot size as each other and across a scale commensurate with the bird data collected. These variables are hereafter referred to as 'Geographical', 'Structure', and 'Floristics' variables, respectively.

'Geographical' and 'Structure' variables are described fully in Table 1. Models of bird-habitat relationships will, of course, depend on the habitat features recorded in the field. In our case, we did not target our habitat measures towards any particular group of birds (very little is known about the ecologies of almost all the species studied anyway). Instead, we used a range of geographical and vegetation structural measures that have been used in past studies of multi-species bird-habitat relationships (e.g. Marsden and Fielding 1999, Jones *et al.* 2003), along with inventories of tree species.

Attempts were made to identify to species level all tree stems > 40 cm girth (circumference) at breast height (gbh), within a 20 m radius of each point transect. In total, 39 tree species were identified at the survey points consistently and independently by two field assistants. Tree species recorded from < 5 survey points were removed from the data set ($n = 11$ species), since ordination methods are sensitive to such outlying cases (Seaby and Henderson 2004). Detrended Correspondence Analysis (DCA: Hill 1979, Hill and Gauch 1980) using Community Analysis Package v.3.0 (Seaby and Henderson 2004) was used to describe the floristics at each point and the underlying tree composition gradients within the study area. DCA ordinates samples (point transects) and species (trees) simultaneously, producing a series of orthogonal axes that maximise site dispersion along each axis (ter Braak 1995). Greater distances between points along DCA axes indicate greater differences in tree species composition. Initial ordination was based on four axes, and the major axis was divided into segments ('detrending by segments') to remove the problems of the quadratic relationship between axes 1 and 2 (the 'arch' effect) and end point compression along the first axis (Hill and Gauch 1980). Rare species were not down-weighted, having been removed from the data set.

DCA extracted four gradients of tree floristics, which are described below. Although it was relatively straightforward to describe the major axis, we acknowledge that not having a full understanding of the complex relationships that underpin tree species composition (e.g. Gentry 1988) meant our interpretations of axes 2–4 might have overlooked some more subtle interactions (e.g. Brokaw 1985).

DCA Axis 1 ('F1', eigen value = 0.708: the strength of/amount of variation along an axis]: termed 'Forest succession') described a gradient of natural habitat succession. Species with low

Table 1. Description of the habitat variables measured or estimated, the method used and sampling radius they were recorded within at each point transect. Time required to collect each habitat variable is also listed. n/a = not applicable.

Habitat variable[1]	Code	Sampling radius	Mean time taken (mins)	Description of variable	Variable type
Altitude	AL	n/a	5	Taken from GPS fix, or triangulated from topographic maps in areas of poor satellite coverage	Geographical
Distance to forest edge	FE	n/a	3	Measured directly, or indirectly when mapping point transect locations; 'forest edge' is defined here as the boundary between the forest block of the site and the surrounding agricultural and scrub habitat	Geographical
Distance to water[a]	WT	n/a	3		Geographical
Topography[b]	TP	n/a	3	Point transect position classified as 'valley bottom' (0–25% of the distance from the bottom of the valley the point was located in), 'mid-slope' (26–75%), or 'ridge top' (> 75%)	Geographical
Gradient[c]	GR	20 m	1	Mean gradient from two measurements, one up and one down any general slope pattern, using a clinometer	Geographical
Y coordinate	Y	n/a	5	Point transects were mapped along	Geographical
X coordinate	X	n/a	5	west-east and south-north axes and assigned x and y coordinates for the study site	Geographical
Rock cover[b]	RK	20 m	1	Estimate of percentage rock cover	Geographical
Floristics DCA 1	F1	20 m	30	Derived from field identification of 28 indicator tree species: See Habitat data collection and analysis	Floristic
Floristics DCA 2	F2	20 m	30		Floristic
Floristics DCA 3	F3	20 m	30		Floristic
Floristics DCA 4	F4	20 m	30		Floristic
Very large trees[a]	VT	20 m	2	Number of woody stems with > 320 cm girth at breast height (gbh)	Structural
Large trees[a]	LT	20 m	3	Number of woody stems with > 80–320 cm gbh	Structural
Medium trees	MT	10 m	4	Number of woody stems with 20–80 cm gbh	Structural
Small trees	ST	10 m	5	Number of woody stems with < 20 cm gbh	Structural
Canopy height	CH	20 m	2	Mean canopy height at each plot calculated from four clinometer measurements, one from each quarter.	Structural
Palm stems	PL	10 m	2	Number of palm stems ≥2 m tall	Structural
Canopy cover[b]	CC	20 m	2	Estimate of percentage foliage cover at a height of ≥ 20 m	Structural
Midstorey cover	MC	20 m	2	Estimate of percentage foliage cover at a height of 5–20 m	Structural
Understorey cover[b]	UC	20 m	2	Estimate of percentage foliage cover at a height of 1–5 m	Structural

Table 1. Continued.

Habitat variable[1]	Code	Sampling radius	Mean time taken (mins)	Description of variable	Variable type
Ground cover[c]	GC	20 m	2	Estimate of percentage foliage cover at a height of <1 m	Structural
Bamboo cover	BM	20 m	1	Estimate of percentage of bamboo cover over all vegetation strata	Structural
Mean gbh	GB	20 m	5	Measured gbh of the five largest trees. For buttress-rooted trees, gbh was taken at a point immediately above the buttressing	Structural

[1]Habitat variable transformations: a = log (x + 1); b = arcsine.$\sqrt{[(x+1)/100]}$; c = $\sqrt{(x + 0.5)}$

scores were indicative of primary forest (e.g. *Koordersiodendron pinnatum*), and those with high scores were associated with second growth (e.g. *Macaranga tanarius*).

DCA Axis 2 ('F2', eigen value = 0.372: 'Logging intensity') appeared to explain an axis of increasing intensity in historical logging disturbance: tree species with high scores (e.g. *Antidesma ghaesembilla*) were those associated with areas of forest that had been cleared heavily in the past, and low-scoring species (e.g. *Bischofia javanica*) were associated with areas that had only been selectively logged in the past.

DCA Axis 3 ('F3', eigen value = 0.275: 'Topography') seemed to relate to a topographic gradient in tree species composition. Low scoring species (e.g. *Lagerstroemia speciosa*) tended to be associated with valley forest and/or more level areas, while high scoring species (e.g. *Dillenia philippinensis*) were generally associated with slopes and ridge tops.

DCA Axis 4 ('F4', eigen value = 0.213: 'Canopy gaps and emergent trees'). Species with low scores (e.g. *Pterocymbium tinctorium*) were associated with open forest, regardless of anthropogenic or topographic influences, and high scoring species (e.g. *Intsia bijuga*) tended to be emergent trees within closed canopy forest.

Bird-habitat associations

The capabilities of the habitat variables (predictors) to explain variation in the occurrence of species and guilds (response variables) were determined using hierarchical partitioning (HP) of logistic regression analysis (Chevan and Sutherland 1991, MacNally 1996). The response variable was the presence (1) or absence (0) of a given species/guild at a sampling point. Hierarchical partitioning was implemented using the 'hier.part package' v.0.5–1 (Walsh and MacNally 2003) in the R Project statistical software v.2.0.1 (R Development Core Team 2004). Habitat variables were transformed where necessary prior to regression analyses (see Table 1).

Hierarchical partitioning (Chevan and Sutherland 1991) removes the problem of multi-colinearity between predictor variables (MacNally 1996) by considering all possible combinations of predictor variables to identify those variables whose independent correlation with a response variable (I_A) may be important and distinct from variables that have little independent effect, but may have a high correlation with the response variable because of joint contributions (J_A) shared with other predictor variables. Thus, the total contribution of each predictor variable to the variation in a response variable is separated into independent (I_A) and joint components (J_A) (MacNally 2002). Those habitat variables that have high independent correlations with bird occurrences are most likely to be effective for assessing bird-habitat relationships during conservation surveys, and extremely valuable to conservation practitioners aiming to understand the impact of habitat management strategies on birds (MacNally 2002).

The analysis is limited to 12 variables at a time (Walsh and MacNally 2003). Therefore, each group of habitat variables (geographical, structure, floristics) was entered separately into regression analysis for each species or guild. The twelve variables from all three analyses that explained the most *independent* variance (I_A) in the occurrence of each species or guild were identified and entered into a final 'mixed variable' HP model. Z scores were calculated for each of these 12 independent contributions, with significance based on the upper 95% confidence limit, i.e. $z \geqslant 1.65$ (MacNally 2002), to identify those variables that explained independently a significant amount of the variance in occurrence of that species/guild.

The total explanatory power of structural variables was compared to the explanatory power of structural variables in combination with other types of habitat variables. Explanatory powers were log-transformed and compared using two-way parametric ANOVA. Although one data set includes 12 structure variables and the other between one and eight structural variables, variables had, from the outset, an equal likelihood of being selected for these later analyses.

Results

Field data collection issues

During three months of fieldwork, DL was able to undertake 512 point transect surveys at Mt. Siburan IBA. Altogether, 2,486 bird records were accumulated within 50 m of point transects, at a rate of 8.4 $\pm$ 3.3 SD individuals and 4.5 $\pm$ 1.7 SD species per count period. A total of 79 bird species were recorded during this survey, with 68 of these being described as forest species. There were few records of *Lonchura* spp. ($n = 2$), *Megalurus* spp. ($n = 2$) and *Accipiter* spp. ($n = 3$). Five species were recorded at least once at more than 50 point transects (Table 2), 15 species at 20–49 points, and 33 species at 5–19 points. Eight of 11 lowland (excluding nocturnal) species of high conservation interest (Mindoro endemics and Threatened/Near-threatened species) known historically from the site were recorded (three at $\geqslant$ 20 points, five at $<$ 20 points), and one threatened island endemic montane species (Mindoro Imperial-pigeon *Ducula mindorensis*; $n = 1$).

On average, it took 30–35 minutes at each point transect to collect the 12 structure variables (Table 1). Time spent collecting geographical characteristics at each point was more variable as six of the eight variables were derived at least in part from the GPS fix at the point, and the ease of obtaining this depended on local canopy cover. Once a GPS fix was obtained, mapping the position of a point and calculating the various distance variables was not time consuming. Time spent collecting floristic data of this type ranged from 5–50 minutes per guide per survey plot, depending on broad habitat type.

Autocorrelations among habitat variables

There were some strong autocorrelations among the habitat variables collected (Figure 1). There were nine pairwise correlations with very strong autocorrelation (r_s values $>$ 0.5). These were Altitude vs Distance to forest edge, Topography, Canopy height and Ground cover; Distance to forest edge vs Topography and Ground cover; Gbh vs Very large trees; and Canopy cover vs Understorey cover and Ground cover. Most highly correlated overall with other habitat variables were three geographical variables (Altitude, Distance to Forest Edge, and Topography), six structural variables (especially vegetation covers and numbers of small/medium trees), and only one floristic variable (F1, the principal DCA axis). The least strongly autocorrelated variables included the other three floristic axes, along with some variables describing quite specific vegetative (bamboos and palms) or geographical (distance to water, rock cover) features (Figure 1).

Performance of habitat variables in explaining the distribution of species

The amount of variance explained by the geographical, structural and floristic variables differed greatly between species (Table 2). There was no correlation between the number of points at

Table 2. Maximum variance explained by a single group of habitat variables ('F' = Floristics, 'G' = Geographical, 'S' = Structure), and important habitat variables for each species (codes for habitat variables are shown in Table 1). '*n*' is the number of 'positive' points. Maximum variance = percentage of variance explained by the most powerful group of habitat variables.

Species	*n*	Guild	Maximum variance (%)	Important variables
Oriolus chinensis	14	O	79.8 (G)	AL, FE, MT
Ptilinopus occipitalis	32	CF	54.0 (S)	CH, Y, MT
Centropus viridis	67	CL	44.0 (G)	FE, UC, AL
Hypothymis azurea	30	UI	39.0 (G)	AL, CH, UC
Phapitreron leucotis	42	CF	35.1 (G)	TP, FE, UC
Hemiprocne comata	9	SI	31.2 (G)	AL, FE, X
Corvus enca	10	O	27.2 (G)	MT, AL, Y
Prioniturus discurus	42	CF	27.0 (G)	CC, TP, MT
Cuculus fugax	8	UGI	19.5 (F)	WT, F3, GR
Treron curvirostra	11	CF	19.5 (F)	F3, MT
Lalage melanoleuca	15	UGI	18.9 (G)	–
Aethopyga shelleyi	13	N	17.9 (G)	TP, FE, AL
Penelopides mindorensis	36	CF	17.6 (S)	MT, F2, GC
Ducula poliocephala	18	CF	17.1 (F)	F4
Eudynamys scolopacea	34	CL	16.8 (F)	CH, F3, F4
Centropus steerii	25	CL	15.6 (G)	AL
Surniculus lugubris	17	UGI	15.6 (G)	TP
Zosterops nigrorum	7	UGI	15.6 (S)	CC, WT, VT
Loriculus philippensis	12	N	14.8 (S)	ST, TP
Tanygnathus lucionensis	7	CF	14.7 (F)	ST, F3
Hypsipetes philippinus	105	O	14.6 (S)	MT, UC, LT
Dicrurus balicassius	107	O	14.1 (F)	CH
Ducula aenea	23	CF	12.1 (G)	X, F4, FE
Pachycephala albiventris	8	UGI	11.6 (G)	VT, TP, GB
Terpsiphone cinnamomea	9	UI	11.4 (F)	GR
Ptilinopus leclancheri	14	CF	11.3 (S)	MT
Dryocopus javensis	55	UGI	11.2 (S)	–
Gallus gallus	16	GD	10.2 (F)	F4, UC
Dicaeum retrocinctum	41	N	9.9 (F)	F4
Megalaima haemacephala	7	CF	9.9 (F)	RK, GC, GR
Coracina striata	23	UGI	9.6 (G)	WT
Chalcophaps indica	22	GD	9.0 (F)	–
Cyornis rufigastra	28	UI	7.4 (S)	VT, PL
Dendrocopos maculatus	13	UGI	7.2 (F)	–
Dicaeum bicolor	14	N	7.1 (G)	FE
Parus elegans	17	UGI	7.0 (S)	CC
Macropygia phasianella	19	CF	6.6 (F)	BM
Dicaeum pygmaeum	8	N	5.7 (F)	CC
Sarcops calvus	54	O	3.1 (S)	–

which species were recorded and the mean percentage of variance explained for those species ($r = +0.06$, $P = 0.70$). There was a strong correlation across species between the percentage variance explained by structural variables and that explained by geographical variables ($r = +0.83$, $P < 0.0001$). Variance for species explained by floristic variables was less strongly correlated with either geographical ($r = +0.50$, $P = 0.001$) or structural variables ($r = +0.55$, $P < 0.001$).

The ten most important variables for explaining independently the occurrence of species included seven geographical and three floristics variables (Figure 2).
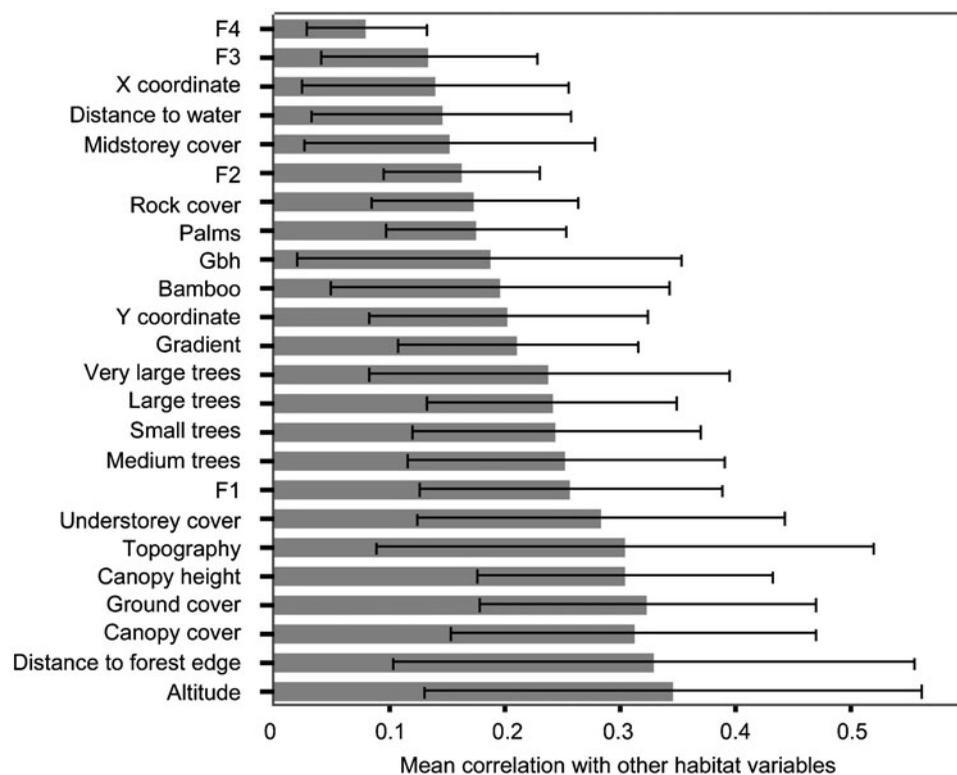
Figure 1. Autocorrelations (Mean values of absolute Spearman's rank coefficients $\pm$ SD) between habitat variables.

Performance of habitat variables in explaining the distribution of guilds

Eleven habitat variables appeared in list of the five most powerful predictors for at least one guild. These included all four floristic variables and all but one geographical variable – the list included no structural variables (Table 3). Some habitat variables were important for several guilds with 'Altitude' appearing in the top five predictors for all but one guild, 'Distance to forest edge' in five guilds, and 'Gradient' and 'Floristics F3' in four guilds. We looked for correlations across the ranked order of habitat variables (Table 3) to examine the similarities across guilds in the types of variables that best predicted species presence. There were some strong correlations between the ranks, for example between Ground dwellers and Nectarivores ($r_s = + 0.74$, $P < 0.001$), and Canopy frugivores and Omnivores ($r_s = + 0.66$, $P < 0.001$). Some guilds, however, had sets of habitat predictors that were different from that of most other guilds. The ranked order of habitat variable importance for Understorey gleaning insectivores was significantly correlated with that for just one other guild (Nectarivores) and Understorey insectivores with just two other guilds (Omnivores and Canopy frugivores).

The 12 structural variables alone had less explanatory power (mean = 11.6% $\pm$ 13.2 SD) than the 12 most powerful habitat variables (which included 1–8 structure variables for each guild; mean = 20.7% $\pm$ 18.6 SD) across all guilds. A two-way parametric ANOVA (on log-transformed explanatory power) indicated a significant effect of variable type ($F_{1,62} = 14.6$, $P < 0.001$), near-significant differences in explanatory power across guilds ($F_{6,62} = 2.1$, $P = 0.06$), but no interaction between explanatory power and guild ($F_{6,62} = 0.10$, $P = 0.99$).
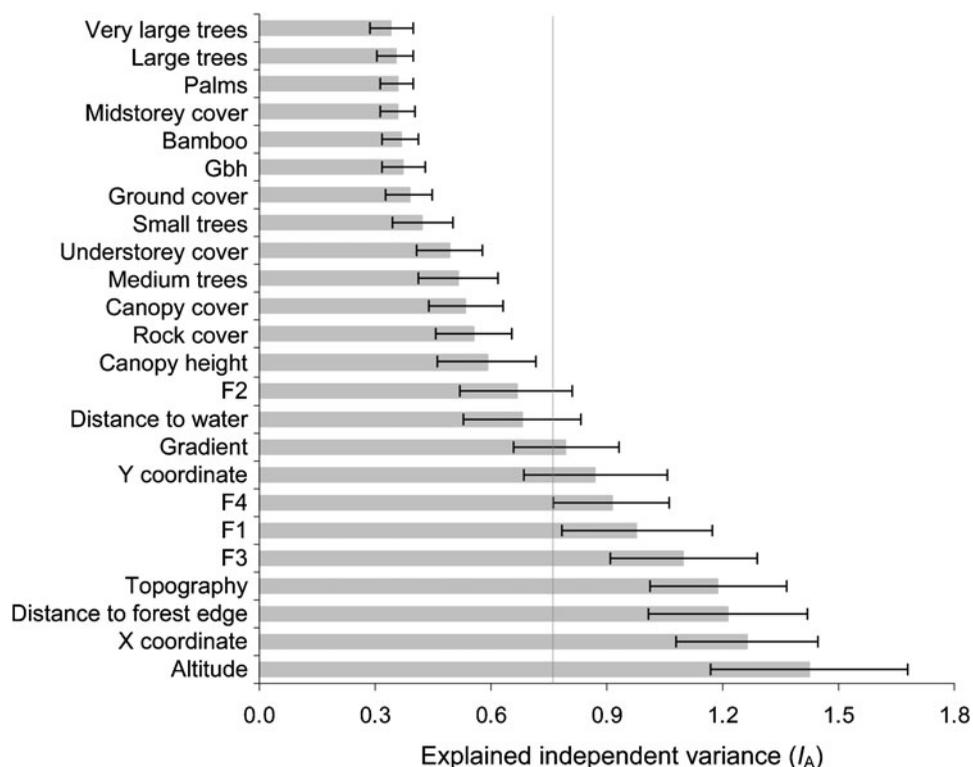
Figure 2. Mean independent explanatory power of each habitat variable ($\pm$ SE). The grey line represents the mean independent variance explained by predictor variables.

Discussion

Issues with the present study: improvements and alternative approaches

Obtaining information on fruiting and flowering phenology, prey abundance, and potential nesting sites is beyond the capacity of most conservation-orientated surveys, since time in the field, training capacity, and experience required are often limited (Fleishman *et al.* 2003). Therefore, it is important to identify surrogates for these data that can go some way towards explaining bird-habitat relationships which are both biologically interpretable and useful for conservation management. Certainly the habitat variables achieved this - combinations of the variables collected accounted for significant variation in the occurrence of 34 of the 39 bird species, and all bar one (midstorey cover) of the habitat variables explained significant variation in the occurrence of at least one bird species. Having said this, explanatory power was generally low with the variables used, and we were unable to link the occurrence of some species to habitat features.

The first issue is that several species (including some species of conservation concern) were recorded very few times during the fieldwork. There are severe limits to the number of candidate explanatory variables that can be included when developing bird-habitat models for species recorded at just a handful of plots and also for ubiquitous species that were present at all but a few plots. It is the smaller of the two numbers, the number of presences and the number of absences, which is important here. This is a difficult problem and especially frustrating with rare

Table 3. Mean ranked independent (I_A) percentage variance in the occurrence of similar species explained by individual habitat variables. Independent variance of habitat variables is ranked within guilds, and total rank of each habitat variable made across guilds (1 = most powerful variable). The number of species representing each guild is in parentheses.

Variable Name	Type	'O' (5)	'N' (5)	'UGI' (9)	'CF' (11)	'CL' (3)	'GD' (2)	'UI' (3)	Total Mean ± SD	Rank
Altitude	G	1	4	9	5	2	2	5	4.0 ± 2.7	1
Distance to forest edge	G	2	3	6	8	1	4	7	4.4 ± 2.6	2
Topography	G	3	1	1	7	7	13	6	5.4 ± 4.2	3
Floristics F3	F	9	10	4	3	4	12	4	6.6 ± 3.6	4
Gradient	G	11	5	7	14	5	5	2	7.0 ± 4.1	5
Floristics F1	F	8	7	14	2	15	7	1	7.7 ± 5.3	6
Floristics F4	F	24	2	8	4	3	1	15	8.1 ± 8.5	7
y coordinate	G	6	6	19	6	9	6	8	8.6 ± 4.8	8
x coordinate	G	4	9	3	1	24	21	3	9.3 ± 9.4	9
Floristics F2	F	5	11	24	12	6	3	24	12.1 ± 8.7	10
Canopy height	S	10	21	11	10	10	19	10	13.0 ± 4.8	11
Rock cover	G	14	8	5	21	13	9	23	13.3 ± 6.7	12.5
Understorey cover	S	15	17	23	13	8	8	9	13.3 ± 5.6	12.5
Distance to water	G	21	12	2	15	12	10	22	13.4 ± 6.8	14
Canopy cover	S	12	15	12	9	11	18	20	13.9 ± 4.0	15
Ground cover	S	17	16	17	18	20	14	16	16.9 ± 1.9	16
Medium trees	S	7	24	21	11	22	23	13	17.3 ± 6.8	17.5
Midstorey cover	S	18	19	15	23	17	11	18	17.3 ± 3.7	17.5
Bamboo cover	S	20	20	18	17	16	17	14	17.4 ± 2.1	19
Palms	S	22	18	16	22	19	16	11	17.7 ± 3.9	20
Small trees	S	13	13	22	16	21	20	21	18.0 ± 3.9	21
Large trees	S	16	14	20	20	23	15	19	18.1 ± 3.0	22.5
Mean gbh of largest trees	S	19	23	13	19	14	22	17	18.1 ± 3.5	22.5
Very large trees	S	23	22	10	24	18	24	12	19.0 ± 5.4	24

and threatened species (e.g. the 'Critically Endangered' Mindoro Bleeding-heart *Gallicolumba platenae* in this study). Partial solutions include firstly, collection of more bird data, be it through a larger survey effort within a random sampling design, or perhaps more appropriately, targeting surveys to likely areas or habitats for the species. In particularly rare species, habitat recordings could be made around spots where individuals are recorded during targeted searches for the species, and these related to a series of randomly placed 'negative' sites. A second approach might be to reduce the number of candidate habitat variables using prior knowledge of likely important habitat features from previous studies and from basic understanding of the ecology of the species. We do not, at present, have detailed information of this kind for most tropical forest bird species, but we can use commonsense or knowledge of related species from better known regions (e.g. woodpeckers and snags; Walter and Maguire 2005). Some useful examples of how to whittle down the number of independent variables using reasoning from prior knowledge and commonsense are given by Burnham and Anderson (2002). A third option is to change the analytical approach: classification trees (e.g. Gray and Fan 2008) or presence only techniques such as GARP (e.g. Pearson *et al.* 2007) might be useful in offering clues as to which environmental features a species might be associated with, or at least guide future detailed surveys of a species into the right areas or habitats.

Our response variable was rather crude in that it relates to presence or absence of a species during any of four ten-minute long visits to each point, so many false negative plots are to be expected (plots where we happened not to record a species that uses it). That presence-absence

data for birds should be more powerful than presence-only data is not doubted (e.g. Brotons *et al.* 2004), but improving our measures of bird presence/usage, through use of point density estimates (Buckland *et al.* 2004) or occupancy (MacKenzie 2006) should be a general aim of bird-habitat studies. A final issue was that we related bird species recorded within 50 m of a plot's central point to habitat recordings taken within 20 m – a practicality to make the bird data workable and the habitat data collectable.

The above factors considered, it may also have been that rather simple models based upon the habitat variables collected did not reflect the true habitat requirements of a species, making it difficult to explain their occurrence either at this time of year or at this spatial scale (Wiens 1989). Second, some species may have very general habitat requirements, which spanned the range of habitats represented at the study site. Coleto (*Sarcops calvus*) may be one such example, and it may be better to describe its habitat characteristics at a coarser scale such as broad land use or vegetation type (e.g. forest vs scrub vs grassland). Third, the analysis technique used here requires us to assume that a species' reaction to a habitat measure is monotonic, so species with multimodal relationships with habitat variables or preferences for mid-range values of the variables can be easily missed (e.g. Bell 1999). Techniques to deal with multimodal responses, curvilinear relationships and interactions include general additive models (e.g. Hastie and Tibshirani 1990) and polynomial regression models. We also assumed that there were no statistical interactions among variables. However, in species for which we know details of ecological requirements we often find that interactions are important; as for example when a species only occurs if there are both sufficient feeding sites and roost sites within easy reach of one another. Where knowledge permits, we recommend the assembly of a set of candidate models that include biologically plausible interactions.

Which measures were most useful in explaining bird-habitat associations?

Seven of the ten most independently powerful habitat variables were geographical variables, with Altitude, X coordinate, Distance to forest edge and Topography being most powerful. Altitude has, for a long time, been known as one of the fundamental drivers/indicators of bird distribution (e.g. Wiens 1989), but the importance for forest birds of topographical variation over such fine scales is less well demonstrated. Distance to forest edge, notwithstanding difficulties in definition or measurement, should be important in many bird species, whether as an indicator of general forest 'quality' (e.g. Gehlhausen *et al.* 2000), nesting success (e.g. Flaspohler *et al.* 2001), or a correlate of disturbance or hunting. Geographic location on a west-east axis (especially important in Canopy frugivores, Omnivores, Understorey insectivores, and Ground-dwellers) was not significantly correlated with any of the other geographical variables (r_s max $= -0.15$, $P_{min} = 0.06$; altitude) but may reflect an important geographical gradient in vegetation or resource availability (e.g. Jones *et al.* 2003), although it may also indicate differences in human disturbance patterns (the west of the site is more regularly disturbed by hunters, or collectors of forest products (D. Lee pers. obs. 2002, 2003).

Aside from geographical variables, the most important variables were floristic, of which tree species composition linked to topography (F3) was the most powerful. The greater importance of tree species composition than structural variables is likely because it not only provides information on fruiting, flowering, and nesting resources, it reflects differences in habitat structure to some extent too. It also supports the idea that floristics are more influential at the local scale than habitat structure (Power 1975, Wiens and Rotenberry 1981, Rotenberry 1985, Wiens *et al.* 1987), a pattern that has been found especially with frugivorous birds (Hasui *et al.* 2007). Interestingly, the major floristics axis (F1: 'forest quality') linked fewer bird species to habitat than the secondary floristics axes. This suggests that, although this axis (F1) defined the most obvious gradient in tree species composition, any distinct species associations with that axis were better explained by geographical or structural variables, with which F1 was correlated

strongly. The secondary axes (F2, F3 and F4) were among the variables least strongly correlated with the other habitat variables.

Eleven of the 12 least important variables were structural, and this is important since many studies focus on relating birds to structural characteristics of habitat (e.g. Berry and Bock 1998, Doherty and Grubb 2000). In addition, structure variables as a group of habitat variables had less explanatory power than the most powerful combinations of different types of habitat variables. Therefore, structure variables alone did not provide as much information about local scale bird-habitat associations then when using a combination of habitat variable types. In the absence of more powerful habitat variables, structural measures are no doubt useful, especially given the general ease with which they can be collected. Many of the structural variables were highly autocorrelated with other habitat variables, perhaps indicating that several of the measures contain similar pieces of information (e.g. vegetation cover at canopy and ground levels, number of small versus large trees). The importance of structural variables may also depend on the habitat studied (e.g. monocultural plantations; Bassett-Touchell and Stouffer 2006) or types of birds under study (e.g. parrots/hornbills with strong nest-site preferences; Marsden and Fielding 1999). However, we suggest that in local-scale studies within tropical forests, if geographical and floristic data can be gathered, then structural variables can become somewhat redundant.

Recommendations

Wiens *et al.* (2002) suggested that identifying groups of species that shared similar responses to landscape structure would benefit conservation management. A similar approach is adopted here (Table 4) to recommend priority habitat variables that should be collected during studies similar to this one. If a study is restricted by time and personnel then every effort should be made to collect at least the following 'priority' habitat variables: altitude, distances to forest edge and water, geographic position, topography, gradient, and floristics variables. Consideration should

Table 4. The most important habitat variables linking different groups of species to habitat. Type of habitat variable is in parentheses (G - Geographical, S - Structural, F - Floristics). Irreplaceability indicates how strongly the variable was correlated with other habitat variables (Low indicates strong correlations with other variables).

Habitat variable	Time/difficulty to collect variable	Irreplaceability	Important for
Altitude (G)	Moderate, Easy	Low	Coucals, Ground-dwellers, Nectarivores, Omnivores
Distance to forest edge (G)	Moderate, Easy	Low	Coucals, Ground-dwellers, Nectarivores, Omnivores
Floristics F3 (F)	Lengthy, Difficult	High	Canopy frugivores, Coucals, Understorey insectivores, Upperstorey gleaning insectivores
Floristics F4 (F)	Lengthy, Difficult	High	Canopy frugivores, Coucals, Ground-dwellers, Nectarivores
X or Y coordinate (G)	Moderate, Medium	Medium/High	Canopy frugivores, Omnivores, Understorey insectivores, Upperstorey gleaning insectivores
Topography (G)	Moderate, Easy	Low	Omnivores, Nectarivores, Upperstorey gleaning insectivores
Floristics F1 (F)	Lengthy, Difficult	Low	Canopy frugivores, Understorey insectivores
Gradient (G)	Fast, Easy	Medium	Understorey insectivores
Distance to water (G)	Moderate, Medium	High	Upperstorey gleaning insectivores
Floristics F2 (F)	Lengthy, Difficult	High	Ground-dwellers

at the same time be given to the degree of autocorrelation between habitat variables. In Table 4, we include a column 'Irreplaceability' based on the magnitude of correlations with a range of other habitat variables (see Figure 1). Some pairwise correlations were very strong indicating first that multicollinearity will be a problem in deciphering drivers of bird distribution, but also that effort might be poorly directed if two autocorrelated variables are recorded in the field (as information will be largely duplicated). Vegetation covers at various strata, counts of very large trees and tree Gbh, and altitude and several other measures are examples of variables with a high degree of duplicated information.

There may be some reluctance, however, to spend time collecting floristics data at the expense of structural variables. Structural variables are quick and easy to collect, and require little expertise and, therefore, should certainly be collected, particularly over larger study areas where their importance to bird-habitat relationships should be more pronounced (e.g. Wiens *et al.* 1987, Jones *et al.* 2003). Conversely, the time and expertise required to identify tree species that occur at low densities in tropical forests (Bibby *et al.* 1998) mean floristics are harder to collect. However, some of the floristic axes were little-correlated with the other habitat variables considered, making them, to a degree, 'irreplaceable', and adding to their importance (see Table 4). Based on the evidence of this study, collecting data on a small indicator group of only 28 tree species is extremely useful for identifying bird-habitat relationships. The tree species this study focused on were well known by local people and easy to identify, were important resources for birds (e.g. *Palaquium luzoniense, Ficus* spp.) and reflected aspects of forest structure, including local disturbance (e.g. *Koordersiodendron pinnatum*) and stages of forest regeneration (e.g. *Bischofia javanica, Macaranga tanarius*). Therefore, we recommend that data on similar groups of indicator tree species are collected during short conservation surveys.

Acknowledgements

References

Bassett-Touchell, C. A. and Stouffer, P. C. (2006) Habitat selection by Swainson's warblers breeding in loblolly pine plantations in southeastern Louisiana. *J. Wildl. Manage.* 70: 1013–1019.

Bell, J. F. (1999) Tree-based methods. Pp. 89–105 in A. Fielding, ed. *Machine learning methods for ecological applications*. Dordrecht: Kluwer.

Berry, M. E. and Bock, C. E. (1998) Effects of habitat and landscape characteristics on avian breeding distributions in Colorado foothills shrub. *Southwest. Nat.* 43: 453–461.

Bibby, C. J., Burgess, N. D. and Hill, D. A. (1992) *Bird census techniques*. London: Academic Press.

Bibby, C. J., Marsden, S. J. and Fielding, A. J. (1998) Bird-habitat studies. Pp. 99–114 in C. J. Bibby, M. J. Jones and S. J. Marsden, eds. *Expedition field techniques: Bird surveys*. London: Royal Geographic Society.

Brokaw, N. V. L. (1985) Treefalls, regrowth, and community structure in tropical forests. Pp. 53–69 in S. T. A. Pickett and P. S. White, eds. *Natural disturbance and patch dynamics*. Toronto: Academic Press.

Brotons, L., Thuiller, W., Araujo, M. B. and Hirzel, A. H. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27: 437–448.

Buckland, S. T., Anderson, D. R., Burnham, K. P. and Laake, J. L. (1993) *Distance sampling: estimating abundance of biological populations.* London: Chapman & Hall.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. and Thomas, L., eds. (2004) *Advanced distance sampling.* Oxford: Oxford University Press.

Burnham, K. P. and Anderson, D. R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach.* Second Edition. Heidelberg: Springer-Verlag.

CGCEO (2001) Landsat.org. Center for Global Change and Earth Observations, Michigan State University. <http://landsat.org> (updated 2001, accessed 22 August 2002, and 20 January 2005).

Chevan, A. and Sutherland, M. (1991) Hierarchical partitioning. *Amer. Statist.* 45: 90–96.

Cody, M. L. (1981) Habitat selection in birds: the roles of vegetation structure, competitors, and productivity. *Biosci.* 31: 107–113.

Condeno, R. and Gaerlan, F. (2002) *Full biosurvey report (Flora). Mt. Siburan, Sablayan, Occidental Mindoro: Full survey, 2002.* Quezon City, Philippines: The Haribon Foundation. Unpub. Report.

Dellasala, D. A., Hagar, J. C., Engel, K. A., McComb, W. C., Fairbanks, R. L. and Campbell, E. G. (1996) Effects of silvicultural modifications of temperate rainforest on breeding and wintering bird communities, Prince of Wales Island, southeast Alaska. *Condor* 98: 706–721.

Dickinson, E. C., Kennedy, R. S. and Parkes, K. C. (1991) *The birds of the Philippines: an annotated check-list.* Tring, UK: British Ornithologists' Union. British Ornithologists' Union Check-list No. 12.

Doherty, P. F. Jr. and Grubb, T. C. Jr. (2000) Habitat and landscape correlates of presence, density, and species richness of birds wintering in forest fragments in Ohio. *Wilson Bull.* 112: 388–394.

Evans, T. D., Dutson, G. C. L. and Brooks, T. M. (1993) *Cambridge Philippines rainforest project 1991: Final report.* Cambridge, UK: BirdLife International. Study Report No. 54.

Farina, A. (1997) Landscape structure and breeding bird distribution in a sub-Mediterranean agro-ecosystem. *Landscape Ecol.* 12: 365–378.

Flaspohler, D. J., Temple, S. A. and Rosenfield, R. N. (2001) Species-specific edge effects on nest success and breeding bird density in a forested landscape. *Ecol. Applic.* 11: 32–46.

Fleishman, E., MacNally, R. and Fay, J. P. (2002) Validation tests of predictive models of butterfly occurrence based on environmental variables. *Conserv. Biol.* 17: 806–817.

Fleishman, E., McDonal, N., MacNally, R., Murphy, D. D., Walters, J. and Floyd, T. (2003) Effects of floristics, physiognomy and non-native vegetation on riparian bird communities in a Mojave Desert watershed. *J. Anim. Ecol.* 72: 484–490.

Gehlhausen, S. M., Schwartz, M. W. and Augspurger, C. K. (2000) Vegetation and microclimatic edge effects in two mixed-mesophytic forest fragments. *Plant Ecol.* 147: 21–35.

Gentry, A. H. (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann. Mo. Bot. Gard.* 75: 1–34.

Gray, J. B. and Fan, G. Z. (2008) Classification tree analysis using TARGET. *Computational Statistics & Data Analysis* 52: 1362–1372.

Hastie, T. J. and Tibshirani, R. J. (1990) *Generalized additive models.* London: Chapman & Hall/CRC.

Hasui, E., Gomes, V. S. D. and Silva, W. R. (2007) Effects of vegetation traits on habitat preferences of frugivorous birds in Atlantic rain forest. *Biotropica* 39: 502–509.

Hill, M. O. (1979) DECORANA – a FORTRAN program for detrended correspondence analysis and reciprocal averaging Ithaca, New York: Cornell University.

Hill, M. O. and Gauch, H. G. (1980) Detrended correspondence analysis, an improved ordination technique. *Vegetatio* 42: 47–58.

James, F. C. (1971) Ordinations of habitat relationships among breeding birds. *Wilson Bull.* 83: 215–236.

Jones, M. J. (1998) Study design. Pp. 15–34 in C. J. Bibby, M. J. Jones and S. J. Marsden, eds. *Expedition field techniques: Bird surveys*. London: Expedition Advisory Centre.

Jones, M. J., Marsden, S. J. and Linsley, M. D. (2003) Effects of habitat change and geographical variation on the bird communities of two Indonesian islands. *Biodivers. Conserv.* 12: 1013–1032.

Karr, J. R. and Freemark, K. E. (1983) Habitat selection and environmental gradients: Dynamics in the ''stable'' tropics. *Ecology* 64: 1481–1494.

Karr, J. R., Robinson, S. K., Blake, J. G. and Bierregaard, R. O. (1990) Birds of four Neotropical forests. Pp. 237–272 in A. H. Gentry, ed. *Four Neotropical rainforests*. New Haven, CT: Yale University Press.

Kirk, D. A. and Hobson, K. A. (2001) Bird-habitat relationships in jack pine boreal forests. *Forest Ecol. Manage.* 147: 217–243.

Lee, D. (2005) Improving methods for conservation-based assessments of abundance and habitat use in tropical forest birds. Unpublished PhD thesis. Manchester Metropolitan University, UK.

MacArthur, R. H., MacArthur, J. W. and Preer, J. (1962) On bird species diversity. II. Prediction of bird census from habitat measurements. *Amer. Nat.* 96: 167–174.

MacKenzie, D. I. (2006) Modeling the probability of resource use: The effect of, and dealing with, detecting a species imperfectly. *J. Wildl. Manage.* 70: 367–374.

MacNally, R. (1990) The roles of floristics and physiognomy in avian community composition. *Aust. J. Ecol.* 15: 321–327.

MacNally, R. (1996) Hierarchical partitioning as an interpretative tool in multivariate inference. *Aust. J. Ecol.* 21: 224–228.

MacNally, R. (2002) Multiple regression and inference in ecology and conservation biology: Further comments on retention of independent variables. *Biodivers. Conserv.* 11: 1397–1401.

MacNally, R. and Bennett, A. F. (1997) Species-specific predictions of the impact of habitat fragmentation: local extinction of birds in the box-ironbark forests of central Victoria, Australia. *Biol. Conserv.* 82: 147–155.

Mallari, N. A. D., Tabaranza Jr., B. R. and Crosby, M. J. (2001) *Key conservation sites in the Philippines*. Makati City, Philippines: Bookmark Inc.

Marsden, S. J. and Fielding, A. H. (1999) Habitat associations of parrots on the islands of Buru, Seram and Sumba. *J. Biogeogr.* 26: 439–446.

Miller, J. R. and Cale, P. (2000) Behavioural mechanisms and habitat use by birds in a fragmented agricultural landscape. *Ecol. Appl.* 10: 1732–1748.

Mills, G. S., Dunning, J. B. Jr. and Bates, J. M. (1991) The relation between breeding bird density and vegetation volume. *Wilson Bull.* 103: 468–479.

Pearson, D. L. (1975) The relation of foliage complexity to ecological diversity of three Amazonian bird communities. *Condor* 77: 453–466.

Pearson, R. G., Raxworthy, C. J., Nakamura, M. and Peterson, A. T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.* 34: 102–117.

Power, D. M. (1975) Similarity among avifaunas of the Galapagos Islands. *Ecology* 56: 616–626.

R Development Core Team (2004) *R: A language and environment for statistical computing.* Version 2.3.1. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, http://www.R-project.org (accessed 25 April 2005).

Rotenberry, J. T. (1985) The role of habitat in avian community composition: Physiognomy or floristics? *Oecol.* 67: 213–217.

Rotenberry, J. T. and Wiens, J. A. (1980) Habitat structure, patchiness, and avian communities in North American steppe vegetation: a multivariate approach. *Ecology* 61: 1228–1250.

Seaby, R. M. H. and Henderson, P. A. (2004) *Community analysis package 3.0.* Lymington, UK: Pisces Conservation Ltd.

Suarez-Seoane, S., Osborne, P. E. and Alonso, J. C. (2002) Large-scale habitat selection by agricultural steppe birds in Spain:

identifying species-habitat responses using generalized additive models. *J. App. Ecol.* 39: 755–771.

SSC (1988) *Mapping the Natural Conditions of the Philippines: Final Report*. Solna, Sweden: Swedish Space Corporation.

ter Braak, C. J. F. (1995) Ordination. Pp. 91–164 in R. H. G. Jongman, C. J. F. ter Braak and O. F. R. van Tongeren, eds. *Data analysis in community and landscape ecology*. Biddles, UK: Cambridge University Press.

Twedt, D. J., Wilson, R. R., Henne-Kerr, J. L. and Hamilton, R. B. (1999) Impact of forest type and management strategy on avian densities in the Mississippi Alluvial Valley, USA. *Forest Ecol. Manage.* 123: 261–274.

Walsh, C. and MacNally, R. (2003) *The hier. part package Version 0.5–1 hierarchical partitioning*. Vienna, Austria: R Foundation for Statistical Computing. http://cran.r-project.org (accessed 25 April 2005).

Walter, S. T. and Maguire, C. C. (2005) Snags, cavity-nesting birds, and silvicultural treatments in Western Oregon. *J. Wildl. Manage.* 69: 1578–1591.

Wiens, J. A. (1989) *The ecology of bird communities*. Vol. 1 & 2. Cambridge: Cambridge University Press.

Wiens, J. A. and Rotenberry, J. T. (1981) Habitat associations and community structure of birds in shrub-steppe environments. *Ecol. Monograph.* 51: 2l–41.

Wiens, J. A., Rotenberry, J. T. and van Home, B. (1987) Habitat occupancy patterns of North American shrubsteppe birds: the effects of spatial scale. *Oikos* 48: 132–147.

Wiens, J. A., van Horne, B. and Noon, B. R. (2002) Integrating landscape structure and scale into natural resource management. Pp. 23–67 in J. Liu and W. W. Taylor, eds. *Integrating landscape ecology into natural resource management*. Cambridge, UK: Cambridge University Press.

DAVID C. LEE*, STUART J. MARSDEN
Department of Environmental and Geographical Sciences, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, U.K.

*Author for correspondence; E-mail: David.Lee@rspb.org.uk