THE ROYAL SOCIETY

Dynamics of extinction and the selection of nature reserves

Miguel B. Araújo[1,2]*, Paul H. Williams[1] and Robert J. Fuller[3]

[1]*Biogeography and Conservation Laboratory, The Natural History Museum, London SW5 5BD, UK*
[2]*Macroecology and Conservation Unit, University of Évora, Largo dos Colegiais, 7000 Évora, Portugal*
[3]*British Trust for Ornithology, The Nunnery, Norfolk IP24 2PU, UK*

Familiar quantitative reserve-selection techniques are tailored to simple decision problems, where the representation of species is sought at minimum cost. However, conservationists have begun to ask whether representing species in reserve networks is sufficient to avoid local extinctions within selected areas. An attractive, but previously untested idea is to model current species' probabilities of occurrence as an estimate of local persistence in the near future. Using distribution data for passerine birds in Great Britain, we show that (i) species' probabilities of occurrence are negatively related to local probabilities of extinction, at least when a particular 20-year period is considered, and (ii) local extinctions can be reduced if areas are selected to maximize current species' probabilities of occurrence. We suggest that more extinctions could be avoided if even a simple treatment of persistence were to be incorporated within reserve selection methods.

Keywords: area selection; persistence; probabilities of extinction; probabilities of occurrence; range contraction; range expansion

1. INTRODUCTION

Reserves are needed to ensure persistence of species within regions (Pimm & Lawton 1998; Margules & Pressey 2000). However, because knowledge of the local population dynamics for many species is poor, explicit criteria for persistence are usually not incorporated within quantitative reserve-selection methods (Nicholls 1998; Williams 1998; Cabeza & Moilanen 2001). Consequently, reserves may be selected where species have low local probabilities of persistence (Margules *et al.* 1994; Virolainen *et al.* 1999; Rodrigues *et al.* 2000; Araújo & Williams 2001). An attractive, but previously untested idea is to model current species' probabilities of occurrence as an estimate of local persistence in the near future (Araújo & Williams 2000). The rationale is that intrinsic factors affecting occurrence of species at one time t_1 (i.e. habitat suitability plus dispersal from neighbour populations) are likely to correlate to some extent with the factors affecting its occurrence at a time t_2, further in the future (i.e. local persistence). Departures from this model (i.e. local extinctions in areas with high probabilities of occurrence) are expected to indicate the degree to which biotic interactions, natural stochasticity or extrinsic human-induced threatening factors affect persistence of species in a given area. We are able the test the usefulness of this idea by comparing distribution atlas records for breeding passerine birds in Great Britain in a particular 20-year period.

2. METHODS

(a) *Data*

As an example, we considered occurrence records of 78 species of breeding passerine birds in Britain, mapped onto a 10 km × 10 km grid for two time-periods (1968–1972 and 1988–1991) (Sharrock 1976; Gibbons *et al.* 1993). Occurrence data include a total of 123 682 records in the first period (minimum number of records per species, 5; median, 1941; maximum, 2774) and a total of 118 358 records in the second period (minimum, 5; median, 1559; maximum, 2747). Overall, we recorded a total of 14 468 local extinctions, i.e. losses of records in grid cells, between the two time-periods (minimum, 2; median, 143; maximum, 862). Two species of passerine birds were excluded from analysis because they had completely non-overlapping distributions in the two time-periods considered. These were the serin *Serinus serinus* and the brambling *Fringila montifringilla*.

(b) *Modelling occurrence and extinction probabilities*

Species' probabilities of occurrence are often modelled in relation to environmental predictors. However, when the aim is to rank observed occurrences according to their probability value rather than to predict occurrences outside a species' known range, or discriminate among possible environmental determinants, then species' local probabilities of occurrence can be modelled using patterns of 'contagion' among records (Araújo & Williams 2000). This uses the idea of positive spatial autocorrelation, in which individuals of a species are more likely to occur in one area if the species also occurs in many surrounding areas. Positive spatial autocorrelation among occurrence records may arise due both to dispersal and environmental similarities among neighbouring areas.

Aggregated or spatially autocorrelated distributions are a problem for statistical testing because they violate the assumption of independence required by most standard statistical procedures (Legendre 1993). This is a severe problem as individuals of most species are typically aggregated, except at very low abundances (e.g. Greig-Smith 1983; McArdle *et al.* 1990; He & Gaston 2000). Modelling probabilities of occur-

Table 1. Number of non-extinctions (persist) and extinctions (extinct) among the British passerine bird species analysed.

(We tested to see if species went extinct more often in cells with lower probabilities of occurrence than in cells in which species persisted. A K–S test for two independent samples was used to test for differences in frequency distribution of probability values in samples (cells) where species persisted and went extinct. Species with significant differences have extinctions occurring significantly more often in cells with lower probability of occurrence than cells in which species persisted. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; n.s., not significant.)

species	K–S test	persist/extinct	species	K–S test	persist/extinct
Lanius collurio	n.s.	5/106	*Sylvia communis*	***	2119/222
Garrulus glandarius	***	1587/156	*Sylvia borin*	***	1618/216
Pica pica	***	1833/103	*Sylvia atricapilla*	***	1812/109
Pyrrhocorax pyrrhocorax	n.s.	61/19	*Panurus biarmicus*	n.s.	31/14
Corvus monedula	***	2286/128	*Cettia cetti*	n.s.	3/2
Corvus frugilegus	***	2136/111	*Locustella naevia*	**	1012/862
Corvus corax	***	995/252	*Locustella luscinioides*	n.s.	5/7
Sturnus vulgaris	***	2599/118	*Regulus regulus*	***	2171/171
Turdus torquatus	***	483/262	*Regulus ignicapillus*	n.s.	10/10
Turdus merula	***	2638/77	*Phylloscopus trochilus*	***	2542/44
Turdus pilaris	n.s.	4/32	*Phylloscopus collybita*	***	1838/163
Turdus philomelos	***	2595/80	*Phylloscopus sibilatrix*	***	894/335
Turdus iliacus	**	49/63	*Acrocephalus schoenobaenus*	***	1688/378
Turdus viscivorus	***	2329/116	*Acrocephalus scirpaceus*	***	622/155
Muscicapa striata	***	2291/145	*Lullula arborea*	***	57/137
Ficedula hypoleuca	***	406/141	*Alauda arvensis*	n.s.	2711/63
Saxicola torquata	***	797/413	*Prunella modularis*	***	2465/135
Saxicola rubetra	***	1183/493	*Passer domesticus*	***	2512/154
Oenanthe oenanthe	***	1536/329	*Passer montanus*	***	1249/424
Phoenicurus phoenicurus	***	1154/510	*Motacilla flava*	***	909/247
Phoenicurus ochruros	*	36/32	*Motacilla cinerea*	***	1674/175
Erithacus rubecula	***	2568/36	*Motacilla alba*	***	2594/68
Luscinia megarhynchos	***	404/233	*Anthus trivialis*	***	1321/475
Sitta europaea	***	1076/99	*Anthus pratensis*	***	2430/194
Certhia familiaris	***	2018/269	*Anthus petrosus*	***	598/137
Cinclus cinclus	***	1191/243	*Plectrophenax nivalis*	n.s.	8/6
Troglodytes troglodytes	n.s.	2713/42	*Emberiza citrinella*	***	2196/235
Parus palustris	***	1056/306	*Emberiza cirlus*	***	22/150
Parus montanus	***	910/310	*Emberiza schoeniclus*	***	2095/382
Parus cristatus	n.s.	28/18	*Miliaria calandra*	***	836/521
Parus ater	***	2235/159	*Fringilla coelebs*	**	2553/29
Parus caeruleus	***	2446/67	*Carduelis chloris*	***	2239/147
Parus major	***	2385/70	*Carduelis carduelis*	***	2013/83
Aegithalos caudatus	***	1988/203	*Carduelis spinus*	***	515/111
Riparia riparia	***	1434/609	*Pyrrhula pyrrhula*	***	2093/232
Hirundo rustica	***	2546/52	*Coccothraustes coccothraustes*	***	183/277
Delichon urbica	***	2310/116	*Carduelis cannabina*	***	2201/176
Sylvia undata	**	20/8	*Carduelis flavirostris*	***	501/157
Sylvia curruca	***	969/126	*Carduelis flammea*	***	1574/402

rence using a predictor term for contagion is one possible way to account for this problem (Smith 1994; Augustin *et al.* 1996).

High contagion among records might also be linked to local persistence if two important ideas are accepted. First, aggregation is a response of populations to local conditions and this reflects the extent to which local environments meet species' ecological requirements (e.g. Hutchinson 1957; Brown 1984; Lawton 1993). In other words, individuals may aggregate because they find suitable environmental conditions that are themselves spatially aggregated. It follows that the greater an area's environmental suitability for a species, the greater the predicted abundance, i.e. the ratio between a population's net increase and net loss (e.g. Lawton 1993). It is indeed encouraging to find that patterns of occupancy can provide useful predictions of abundance (He & Gaston 2000; Pearce & Ferrier

2001), where these estimates are difficult or expensive to obtain more directly. Second, spatial aggregation of species' records is likely to correlate positively with the ability of individuals to disperse from one area to another (i.e. the ability to disperse is partially associated with the distance travelled), hence with the ability to establish stable metapopulation dynamics (e.g. Hanski *et al.* 1996; Wahlberg *et al.* 1996; Gonzalez *et al.* 1998). This may not be the case for most birds that disperse widely during the non-breeding season, but it is likely to be an important factor for other taxa and some of the most sedentary birds. For one reason or another, areas with high contagion are expected to have lower probabilities of local extinction, at least where this results from demographic or environmental stochasticity (Pimm 1991, pp. 135–172; Curnutt *et al.* 1996).

We model probabilities of occurrence for 78 species of breed-

Table 2. *Nagelkerke's r^2 measuring the strength of association between the response variables (i.e. occurrence and extinction) and the predictor variable (i.e. contagion); direction of the association (i.e. positive or negative) between the predictor and response variable, which is given through the β coefficient of equations (2.1) and (2.3); and Wald statistic testing the significance of individual logistic regression coefficients for each predictor variable.*

(Change in the total number of records in the two time-periods, measured as no. records t_2 − no. records t_1; negative values for increasing number of local extinctions over local expansions, positive values for the opposite pattern. ****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; n.s., not significant.)

	occurrence models		extinction models		
	r^2	coefficient/Wald	r^2	coefficient/Wald	change in number of records
Lanius collurio	0,386	(+) ***	0,077	(−) n.s.	−96
Garrulus glandarius	0,787	(+) ***	0,206	(−) ***	−30
Pica pica	0,801	(+) ***	0,52	(−) ***	22
Pyrrhocorax pyrrhocorax	0,617	(+) ***	0,036	(−) n.s.	8
Corvus monedula	0,674	(+) ***	0,258	(−) ***	−70
Corvus frugilegus	0,605	(+) ***	0,244	(−) ***	−10
Corvus corax	0,739	(+) ***	0,171	(−) ***	−116
Sturnus vulgaris	0,264	(+) ***	0,173	(−) ***	−97
Turdus torquatus	0,675	(+) ***	0,162	(−) ***	−201
Turdus merula	0,313	(+) ***	0,081	(−) ***	−51
Turdus pilaris	0,061	(+) ***	0,009	(−) n.s.	67
Turdus philomelos	0,507	(+) ***	0,062	(−) ***	−55
Turdus iliacus	0,47	(+) ****	0,151	(−) **	24
Turdus viscivorus	0,666	(+) ****	0,06	(−) ***	−48
Muscicapa striata	0,545	(+) ***	0,065	(−) ***	−58
Ficedula hypoleuca	0,487	(+) ***	0,345	(−) ***	185
Saxicola torquata	0,521	(+) ***	0,197	(−) ***	−176
Saxicola rubetra	0,488	(+) ***	0,205	(−) ***	−272
Oenanthe oenanthe	0,552	(+) ***	0,328	(−) ***	−127
Phoenicurus phoenicurus	0,511	(+) ***	0,218	(−) ***	−337
Phoenicurus ochruros	0,332	(+) ***	0,12	(−) *	35
Erithacus rubecula	0,611	(+) ***	0,193	(−) ***	25
Luscinia megarhynchos	0,79	(+) ***	0,045	(−) ***	−180
Sitta europaea	0,799	(+) ***	0,22	(−) ***	95
Certhia familiaris	0,561	(+) ***	0,226	(−) ***	−167
Cinclus cinclus	0,722	(+) ***	0,152	(−) ***	−124
Troglodytes troglodytes	0,265	(+) ***	0,045	(−) ***	−8
Parus palustris	0,807	(+) ***	0,144	(−) ***	−229
Parus montanus	0,689	(+) ***	0,153	(−) ***	−120
Parus cristatus	0,559	(+) ***	0,182	(−) *	5
Parus ater	0,606	(+) ***	0,068	(−) ***	−79
Parus caeruleus	0,678	(+) ***	0,081	(−) ***	−33
Parus major	0,649	(+) ***	0,087	(−) ***	−12
Aegithalos caudatus	0,543	(+) ***	0,175	(−) ***	−85
Riparia riparia	0,376	(+) ***	0,074	(−) ***	−484
Hirundo rustica	0,472	(+) ***	0,212	(−) ***	28
Delichon urbica	0,592	(+) ***	0,276	(−) ***	−33
Sylvia undata	0,491	(+) ***	0,389	(−) *	17
Sylvia curruca	0,735	(+) ***	0,212	(−) ***	176
Sylvia communis	0,629	(+) ***	0,232	(−) ***	−155
Sylvia borin	0,673	(+) ***	0,208	(−) ***	33
Sylvia atricapilla	0,694	(+) ***	0,256	(−) ***	127
Panurus biarmicus	0,293	(+) ***	0,101	(−) n.s.	15
Cettia cetti	0,2	(+) ***	0,496	(−) n.s.	81
Locustella naevia	0,478	(+) ***	0,011	(−) ***	−685
Locustella luscinioides	0,144	(+) ***	0	(−) n.s.	15
Regulus regulus	0,493	(+) ***	0,029	(−) ***	−15
Regulus ignicapillus	0,183	(+) ***	0,046	(−) n.s.	79
Phylloscopus trochilus	0,575	(+) ***	0,103	(−) ***	16
Phylloscopus collybita	0,633	(+) ***	0,261	(−) ***	99
Phylloscopus sibilatrix	0,469	(+) ***	0,15	(−) ***	41
Acrocephalus schoenobaenus	0,386	(+) ***	0,107	(−) ***	−179
Acrocephalus scirpaceus	0,664	(+) ***	0,079	(−) ***	13

(*Continued.*)

Table 2. (*Continued.*)

	occurrence models		extinction models		
	r^2	coefficient/Wald	r^2	coefficient/Wald	change in number of records
Lullula arborea	0,419	(+) ***	0,18	(−) ***	−121
Alauda arvensis	0,182	(+) ***	0,054	(−) ***	−45
Prunella modularis	0,535	(+) ***	0,138	(−) ***	−89
Passer domesticus	0,35	(+) ***	0,161	(−) ***	−141
Passer montanus	0,702	(+) ***	0,296	(−) ***	−327
Motacilla flava	0,747	(+) ***	0,232	(−) ***	−109
Motacilla cinerea	0,611	(+) ***	0,106	(−) ***	130
Motacilla alba	0,46	(+) ***	0,042	(−) ***	7
Anthus trivialis	0,51	(+) ***	0,105	(−) ***	−272
Anthus pratensis	0,198	(+) ***	0,138	(−) ***	−85
Anthus petrosus	0,749	(+) ***	0,095	(−) ***	−81
Plectrophenax nivalis	0,276	(+) ***	0,291	(−) n.s.	28
Emberiza citrinella	0,611	(+) ***	0,175	(−) ***	−207
Emberiza cirlus	0,492	(+) ***	0,293	(−) ***	−143
Emberiza schoeniclus	0,332	(+) ***	0,079	(−) ***	−289
Miliaria calandra	0,633	(+) ***	0,299	(−) ***	−436
Fringilla coelebs	0,662	(+) ***	0,097	(−) ***	20
Carduelis chloris	0,622	(+) ***	0,186	(−) ***	−63
Carduelis carduelis	0,777	(+) ***	0,225	(−) ***	113
Carduelis spinus	0,559	(+) ***	0,151	(−) ***	532
Pyrrhula pyrrhula	0,631	(+) ***	0,102	(−) ***	−152
Coccothraustes coccothraustes	0,363	(+) ***	0,214	(−) ***	−145
Carduelis cannabina	0,682	(+) ***	0,387	(−) ***	−109
Carduelis flavirostris	0,703	(+) ***	0,205	(−) ***	−7
Carduelis flammea	0,617	(+) ***	0,056	(−) ***	−222

ing passerine birds in Great Britain by fitting univariate logistic regressions with the software Spss for Windows v. 10, where occurrence records in 1968–1972 are the response variable and contagion among these records is the predictor. This is expressed using the equation:

$$\log\left(\frac{p(\text{Occur}_{i,a,t1})}{1 - p(\text{Occur}_{i,a,t1})}\right) = \alpha + \beta(\text{Cont}_{i,a,t1}), \qquad (2.1)$$

where $p(\text{Occur}_{i,a,t1})$ is the probability of occurrence of a species i in area a at time t_1, in relation to a coefficient β of contagion ($\text{Cont}_{i,a,t1}$) among records of species i in area a and at time t_1; α is a constant.

Models of occurrence using contagion as a predictor variable have been shown to produce very similar results to those using environmental variables (Araújo & Williams 2000; Segurado & Araújo 2002). This is because contagion already incorporates environmental information (species only occur in greater aggregations where the environment is suitable), although it adds additional factors that may not relate to the equilibrium assumptions of environmental envelope approaches (e.g. dispersal and local history). Contagion is measured as a weighted average of the number of occupied grid cells among a set of k_a neighbours of a central grid cell y_a, so that:

$$\text{Contagion} = \left(\frac{\sum_{b=1}^{k_a} w_{ab} y_b}{\sum_{b=1}^{k_a} w_{ab}}\right), \qquad (2.2)$$

where the weight given to the grid cell y_b is $w_{ab} = 1/d_{ab}$ and d_{ab}

is the distance between grid cells y_a and y_b. We used two orders of neighbours, assigning a weight of $d = 1$ to the first-order and a weight of $d = 2$ to the second-order neighbours. Neighbours in the first order were the eight adjacent cells touching the central cell along the edges and at the corners within a rectangular grid. The second-order neighbours were the next group of cells concentric to first order with 16 grid cells.

The outcomes of these models are interpreted as estimates of species' local probabilities of persistence. Because we are not interested in predicting occurrences outside a species' known range, all predicted probability values that did not overlap with an observed record are excluded. This should also reduce uncertainty from selecting areas with predicted but no observed records of occurrence (type I errors) (Araújo & Williams 2000).

Observed occurrences for a grid cell in 1968–1972 that are not recorded in 1988–1991 are treated as extinction events. The probabilities of these events occurring are modelled using the same approach as for occurrences; that is, extinction events for a given species in 1988–1991 are treated as the response variable, while patterns of contagion among records in 1968–1972 are treated as the predictor variable. This is expressed as

$$\log\left(\frac{p(\text{Extinct}_{i,a,t2})}{1 - p(\text{Extinct}_{i,a,t2})}\right) = \alpha + \beta(\text{Cont}_{i,a,t1}), \qquad (2.3)$$

where $p(\text{Extinct}_{i,a,t2})$ is the probability of extinction of a species i in area a at time t_2, in relation to a coefficient β of contagion ($\text{Cont}_{i,a,t1}$) among records of species i in area a and at time t_1; α is a constant.

We used the fitted models to predict species' local probabilities of extinction in relation to initial patterns of aggregation among records. Because we were modelling local extinctions,
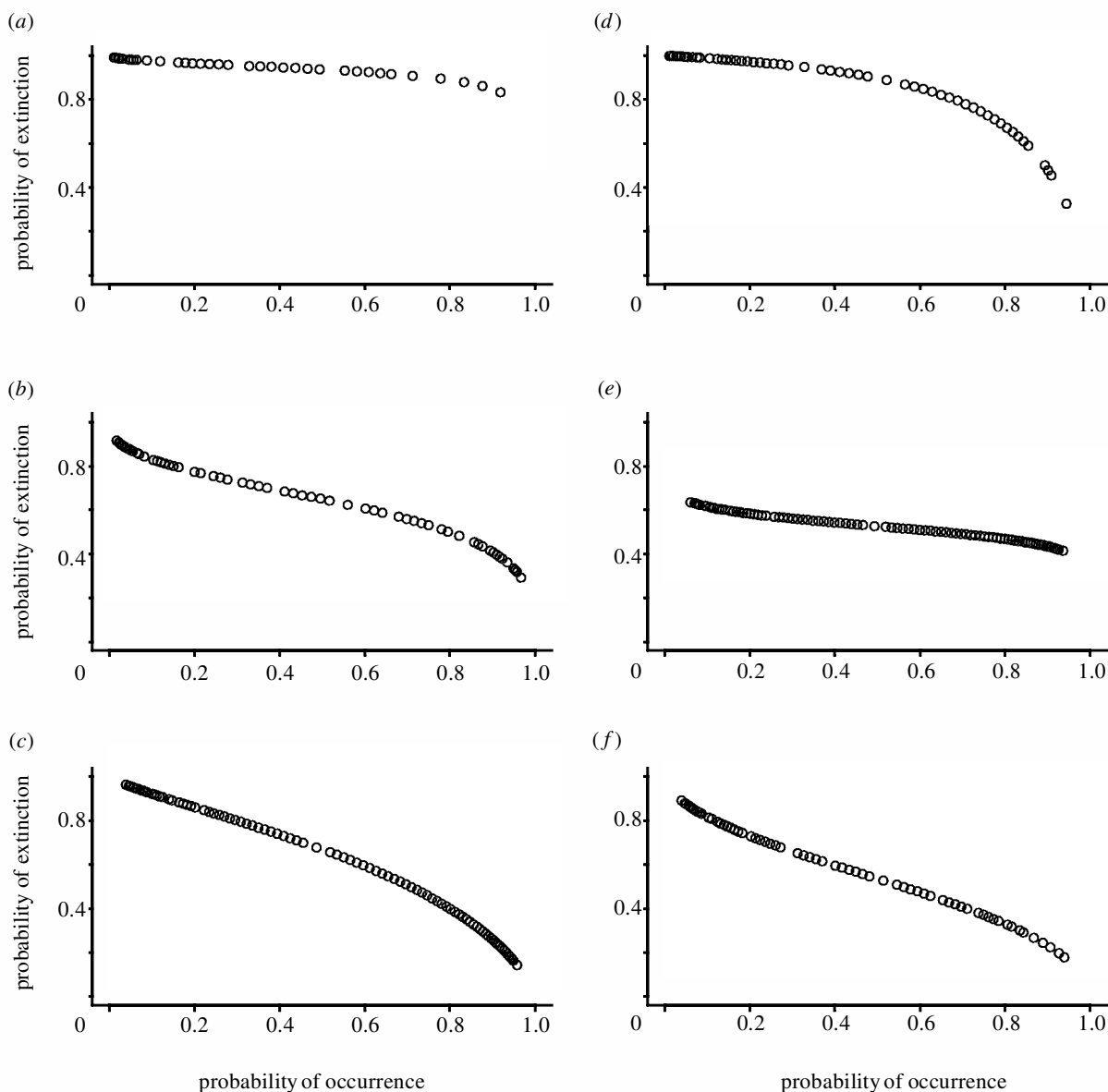
Figure 1. Relationship between the probabilities of occurrence and probabilities of extinction for six species of passerines in Great Britain: (*a*) *Lanius colurio*; (*b*) *Lullula arborea*; (*c*) *Miliaria calandra*; (*d*) *Emberiza cirlus*; (*e*) *Locustella naevia*; and (*f*) *Coccothraustes coccothraustes*. Probabilities of occurrence were modelled using distribution data for 1968–1972. Probabilities of extinction were modelled using information on extinction events occurring in the time-periods 1968–1972 and 1988–1991.

not expansions, models of probabilities of extinction only included areas with occurrence records at time-period 1968–1972.

(c) *Area selection*

Two complementarity-based reserve-selection strategies are used for comparison: a probability-based and an occurrence-based area selection method. The first approach uses a heuristic method for seeking reserve networks that maximize probabilities of occurrence across passerine bird species in 1968–1972, while minimizing the total area reserved. We used the 'goal-gap' algorithm of Williams & Araújo (2000) that seeks to obtain a given combined probability value for species, while minimizing the total area, but with an additional step to ensure that the areas selected were restricted to those with top probability values for each species (see also Williams & Araújo 2002). To achieve this, records with probabilities of occurrence below the top 95% frequency distribution of the maximum probability observed for

every species are excluded (this is an arbitrary value but conservative enough to favour only occurrence records with high probability values). In the second step, the goal-gap algorithm is set with the goal of representing all species with a total combined probability ($p(\text{Occur}_i)$) of at least 0.95. Local probabilities of occurrence ($p(\text{Occur}_{i,a})$) for species i in area a are combined among n areas using the product of probabilities of non-occurrence,

$$p(\text{Occur}_i) = 1 - \Pi_{a=1...n}(1 - p(\text{Occur}_{i,a})), \quad (2.4)$$

where p_{ij} is the probability that species i occurs in area a.

The algorithm begins by selecting all of the areas that are essential to achieving the representation goal. The method then chooses one area at each iteration by examining how much choosing each area would contribute incrementally to reaching the representation goal for each species and choosing the area that contributes the most across all species. If ties occur for any choice, then the area with the highest sum of probabilities across
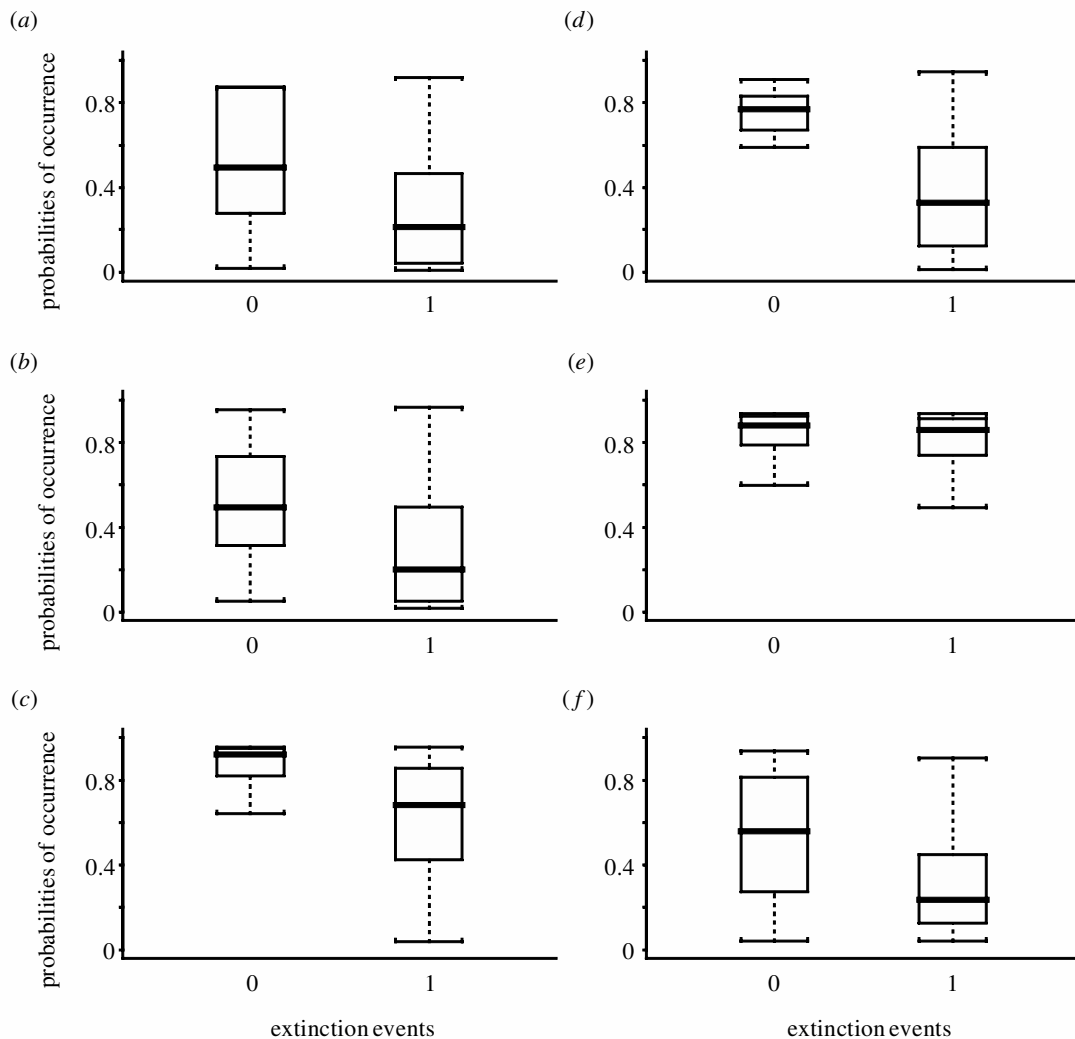
Figure 2. Frequency distribution of probabilities of occurrence in relation to (0) non-extinction; and (1) extinction events: (*a*) *Lanius colurio*; (*b*) *Lullula arborea*; (*c*) *Miliaria calandra*; (*d*) *Emberiza cirlus*; (*e*) *Locustella naevia*; and (*f*) *Coccothraustes coccothraustes*. The shaded bar represents the interquartile range and the median is marked within this; the line extensions from each box are the largest and smallest values, excluding outliers (points more than 1.5 times the interquartile range distant from the box edges). Probability value distributions for *Lanius colurio* are not significantly different between extinction and non-extinction events (K–S test = 0.841, n.s.), but the sample size for non-extinctions ($n = 5$) is too small for a valid test. For the remaining species the frequency distribution of probabilities between extinction and non-extinction is different at $p < 0.0001$, except for *Locustella naevia* (K–S test = 0.1681, $p < 0.01$). (see table 2 for all results).

all species without complementarity is selected. One of the consequences of restricting selection to top probability values for each species is that fewer records are included for selection. This has the consequence of potentially increasing the overall number of areas required to attain a given representation goal. Nevertheless, the algorithm proposed by Williams & Araújo (2000) proved to be superior to the related method of Margules & Nicholls (1987).

The second method that uses a heuristic algorithm that treats all non-zero probabilities overlapping with observed records as presence data to solve two problems: (i) find the set of areas that represents all species at least once in the minimum area possible (i.e. minimum set); and (ii) find the set of areas that represents as many species as possible given the same number of areas as the probability-based method (i.e. maximum-coverage solution). The algorithm used is based on a popular heuristic technique (Margules *et al.* 1988) of selecting those areas richest in the rarest species at each step. Essentially, the algorithm starts

by selecting all areas with taxa that are equally or more restricted than the representation goal. For example, for a goal of representing each species at least once (i.e. minimum set for one representation of all species), it begins by selecting all areas that have species recorded in only one grid cell. Then the algorithm follows a set of simple rules, applied iteratively to select areas that are richest in the rarest taxa. First it selects grid cells with the greatest complementary richness in just the rarest taxa (ignoring other taxa). If there are ties, it proceeds by selecting areas among ties that are richest in the next-rarest taxa. If there are still ties, it then selects those areas among ties with the lowest grid-cell number. This is an arbitrary rule used in place of random choice among ties in order to ensure repeatability in tests. Finally, these are repeated as necessary until the representation goal is achieved. Checks to exclude redundant areas have been added to improve efficiency, together with a re-ordering procedure that provides approximate solutions to maximum-coverage problems (Williams 1999).
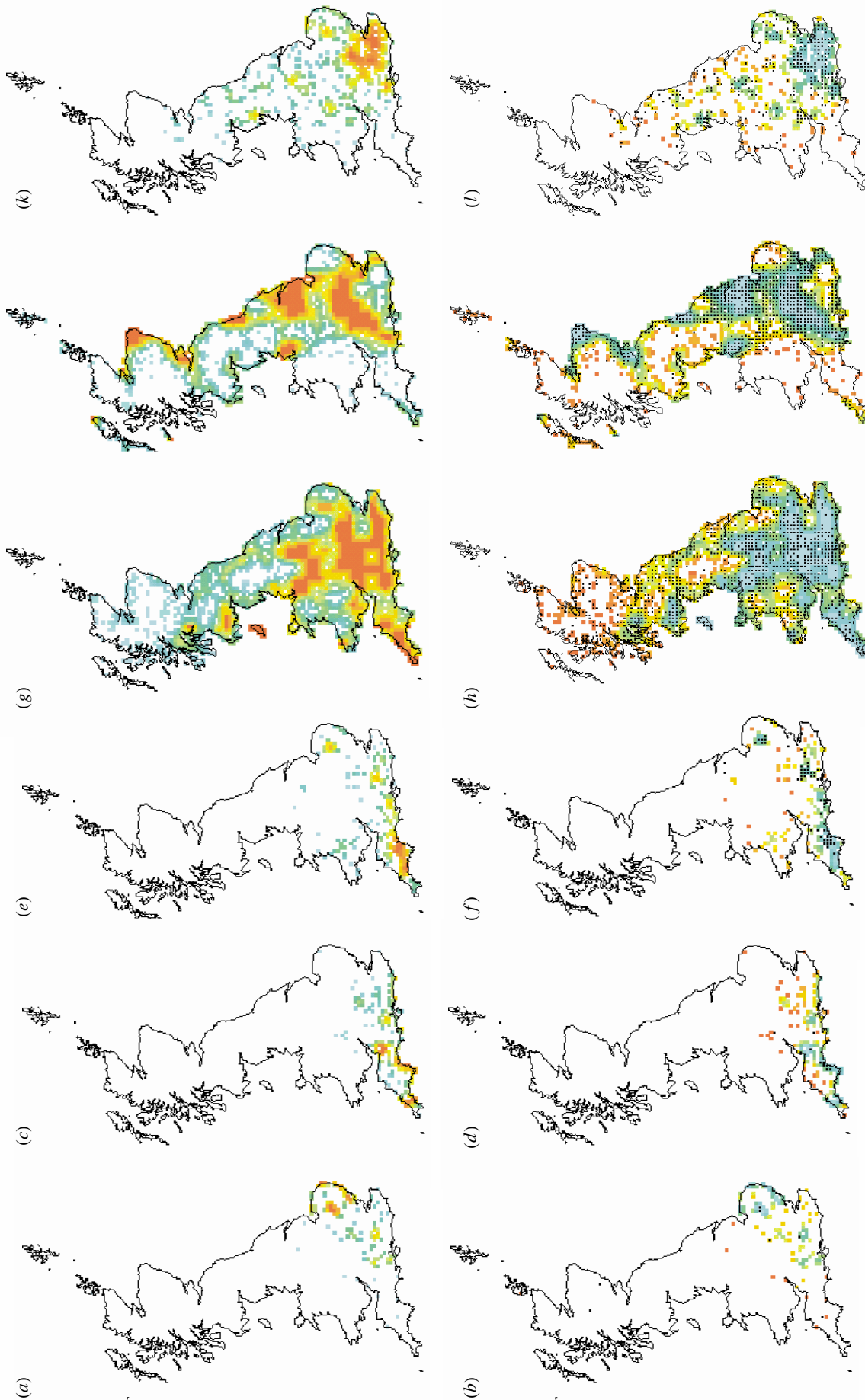
Figure 3. Probabilities of occurrence (first row) and extinction (second row) for six passerines with the highest rate of decline in Great Britain, during the two times considered (i.e. 1968–1972 and 1988–1991). (*a,b*) *Lanius colurio*; (*c,d*) *Emberiza cirlus*; (*e,f*) *Lullula arborea*; (*g,h*) *Locustella naevia*; (*i,j*) *Miliaria calandra*; and (*k,l*) *Coccothraustes coccothraustes*. Quantitative scores were divided into 33 equal-frequency colour classes with maximum-probability scores shown in orange and minimum (non-zero) probability scores shown in blue. Black dots in the second row are observed occurrences in 1988–1991. Extinction probabilities without occurrence records are associated with extinction events. Occurrence records without probability values are range expansions.

3. RESULTS

We found a negative relationship between probabilities of occurrence and extinction for all species, although the slope and precise shape of the curves differ among species (see examples in figure 1). For 81% of the species, extinction events occurred significantly more often in areas where species had lower probability of occurrence than in areas where species persisted (Kolmogorov–Smirnov (K–S) tests, $p < 0.0001$; see figure 2 and table 2). For most of the remaining 19% of species, the sample size was too small for the test of difference between distributions of values to be applicable. One exception is the grasshopper warbler *Locustella naevia*, which had the greatest absolute number of local extinctions amongst the passerine birds (862 events) and a relatively uniform pattern of extinction throughout its range (figure 3). This particular pattern of range contraction had already been described by others (Gates & Donald 2000; Donald & Greenwood 2001), who suggested that factors other than population dynamics might be responsible for this pattern.

For all passerine bird species, probabilities of occurrence were positively correlated with contagion (Wald tests, $p < 0.001$) and for 85% of the species there is a significant negative relationship between contagion and probability of extinction (Wald tests, $p < 0.001$) (see table 3). The strengths of association between contagion and occurrence or extinction are measured with Nagelkerke's r^2, which is a modification of the Cox and Snell coefficient to assure that values range between zero and unity (Nagelkerke 1991). Generally, the association between occurrence events and contagion is greater than the association between extinction events and contagion (see table 3). However, differences between r^2-values in logistic regressions should be interpreted cautiously, because the variance of a dichotomous variable (here, presence and absence of occurrences or extinctions) depends on the frequency distribution of that variable and this is greatly affected by sample size; and there were generally fewer extinction events than occurrences. Nevertheless, results show that, within Britain, most species' ranges tended to collapse towards areas with high contagion (which can be interpreted as regional population cores), rather than towards areas with low contagion (which can be interpreted as regionally isolated or marginal populations) (for visual inspection of results, see examples in figure 3). Almost all species with no significant association (i.e. $p > 0.001$) between contagion and local extinction (15% of passerine species) had expanding ranges within Britain and very low numbers of local extinctions (table 3). The exception was the red-backed shrike *Lanius colurio*, which is the passerine with the greatest relative decline in Britain (a contraction of -86.5% of its range) during the period studied. For this bird, persistence was too low for any test of the difference between means in areas of high contagion and low contagion to be statistically valid.

In agreement with results obtained in other studies (Margules *et al.* 1994; Virolainen *et al.* 1999; Rodrigues *et al.* 2000), a minimum-set strategy to represent all species at least once was insufficient to retain all species in the near future (six species lost from reserves, which represented 7.7% of the total number of species). We also found that if areas were selected on the basis of higher estimated probabilities of occurrence for species within the first time-period (1968–1972), then species would have a greater probability of persisting within reserves in the second time-period (1988–1991) (table 1). Indeed, no species would be lost from areas selected using the probability method (table 1), although one species (1.29%) was lost when the same number of areas was selected using occurrence data alone. The mean number of representations of species within reserves was also higher with the probability-based method than with the maximum-coverage solution identified with presence data. This is because, in order to achieve a given combined probability of occurrence for every species in the set of selected areas (using equation (2.4)), the probability-based algorithm addresses the multiple-representation problem by representing species with low local probabilities of occurrence many times (if the total combined probability among a species is lower than the representation goal, then all records of the species are represented). Given that areas with the highest probabilities of occurrences for species were generally less extinction prone locally, it is unsurprising to find that the percentage change in the mean or lower quartile number of representations per species is also lower with the probability method.

4. DISCUSSION

Results are consistent with the prediction that local extinctions are more common among species records associated with low probabilities of occurrence (low contagion) than among records with high probabilities of occurrence (high contagion) (Araújo & Williams 2000, 2001). Similar results were obtained for a smaller sample of bird species in the UK by Gates & Donald (2000) and by Donald & Greenwood (2001), although they modelled probabilities of occurrence with environmental rather than spatial predictors. Given that contagion already accounts for some of the important environmental variation for each species (through spatial autocorrelation among environmental predictors), it is unsurprising that the broad pattern from these models coincides.

It is particularly encouraging that selecting reserves from populations with high probabilities of occurrence improves the probability that areas would retain species in the near future (e.g. 20 years), although population stability can never be assumed as shown previously for limestone pavement floras (Margules *et al.* 1994). Of course, local management within reserves is likely to offset some of the threats that species might face in the wider countryside. Lower extinction rates would be expected if intensive local management were undertaken to manage local populations and their habitats. However, given the relatively low cost of presence–absence inventories as compared with more detailed demographic, genetic or behavioural studies, this kind of modelling approach is likely to provide useful information for real-world conservation planning. However, the question remains as to whether selecting populations with high probabilities of occurrence at a given time would ensure high probabilities of persistence in the longer term (e.g. 100 years). Certainly, this would depend on the species' spatio-temporal dynamics of range contraction and expansion, which are dependent on many other factors, including local management, environmental change and contagious threatening processes. Indeed,

Table 3. Species-representation results for three reserve-selection strategies at time t_1 (1968–1972) and consequences for persistence of species at time t_2 (1988–1991).

	reserve-selection strategy		
	minimum-set, probability method	maximum-coverage presence data	minimum-set, presence data
number of areas selected	64	64	6
percentage of species represented in 1968–1972	100	100	100
percentage of species represented in 1988–1991	100	98.71	92.30
number of representations per species in 1968–1972 (mean)	43.07	38.79	3.45
number of representations per species in 1988–1991 (mean)	40.94	37.00	3.32
percentage change in number of representations per species (mean)	5.76	−5.94	−5.55
percentage change in number of representations per species (lower quartile)	−8.06	−15.09	−16.66

although range contractions often start in areas with low contagion (e.g. at range peripheries) (Channell & Lomolino 2000*a*), species respond to environmental changes by extending their ranges, also from the peripheries (e.g. Parmesan *et al.* 1999; Thomas & Lennon 1999; Thomas *et al.* 2001). Consequently, what constitutes a local population core (with high contagion) at present could turn into a peripheral population in the future, and vice versa. However, contagious threatening processes may cause some populations to persist in the locally isolated and marginal populations rather than in local cores of their distributions (Channell & Lomolino 2000*a,b*). Even when contagious threats begin at one periphery of a species' range, the last place to be affected is likely to be the region most isolated from the initial point of contagion, i.e. along an opposite edge range (Channell & Lomolino 2000*a*).

Estimates of persistence from simple probability-of-occurrence models could also be improved if knowledge of threatening processes and species' corresponding vulnerabilities were available to be incorporated in the models. However, threats are not static and this approach would only be useful if future threats were predictable. More dynamic models of persistence could be fitted by using environmental-change scenarios to predict the future distribution of core populations (e.g. Huntley *et al.* 1995; Peterson *et al.* 2002), although the usefulness of such approaches remains, to our knowledge, untested (Davis *et al.* 1998; but see Hodkinson 1999). Our ability to predict patterns of species' persistence is likely to improve, but it is unlikely that models will ever be able to remove all kinds of uncertainty arising both from our knowledge and from the way that criteria are combined (Araújo *et al.* 2002). Indeed, natural stochasticity and other kinds of unpredictable threats may cause models to fail to a variable, but not always negligible extent (Ludwig 1999). Nevertheless, for 81% of the species in this study, local extinctions occurred significantly more often in areas with low probabilities of occurrence than in areas where species persisted. This is a reasonable achievement and supports the idea that simple probability-based approaches may be considered an alternative to more complex population viability analysis (e.g. Hanski *et al.* 1996; Wahlberg *et al.* 1996; Brook *et al.* 1997; Lindenmayer *et al.* 2001), when conservation decisions involve large numbers of species and there is little time and few resources available. Nevertheless, models such as those outlined here are likely to be seen mainly as heuristic tools for priority setting in conservation. Uncertainty due to population variability is likely to increase with time (Pimm & Redfearn 1989; Bengtsson *et al.* 1997), so the larger the time-scale considered the greater the expected error in the predictions. If the precautionary principle is to be applied to prevent the loss of biodiversity from reserves, then we need both large proportions of our territories to be set aside to conserve species (Soulé & Sanjayan 1998) and flexible management tools to allow conservationists to re-assess priorities for area conservation and management as knowledge on the species' requirements for persistence improve.

REFERENCES

Araújo, M. B. & Williams, P. H. 2000 Selecting areas for species persistence using occurrence data. *Biol. Conserv.* **96**, 331–345.

Araújo, M. B. & Williams, P. H. 2001 The bias of complementarity hotspots towards marginal populations. *Conserv. Biol.* **15**, 1710–1720.

Araújo, M. B., Williams, P. H. & Turner, A. 2002 A sequential approach to minimise threats within selected conservation areas. *Biodiv. Conserv.* **11**, 1011–1024.

Augustin, N. H., Mugglestone, M. A. & Buckland, S. T. 1996 An autologistic model for spatial distribution of wildlife. *J. Appl. Ecol.* **33**, 339–347.

Bengtsson, J., Baillie, S. R. & Lawton, J. H. 1997 Community variability increases with time. *Oikos* **78**, 249–256.

Brook, B. W., Lim, L. H., Harden, R. & Frankham, R. 1997 Does population viability software predict the behaviour of real populations? A retrospective analysis of the Lord Howe

Island woodhen, *Tricholimnas sylvestris* (Sclater). *Biol. Conserv.* **82**, 119–128.

Brown, J. H. 1984 On the relationship between abundance and distribution of species. *Am. Nat.* **124**, 255–279.

Cabeza, M. & Moilanen, A. 2001 Design of reserve networks and the persistence of biodiversity. *Trends Ecol. Evol.* **16**, 242–248.

Channell, R. & Lomolino, M. V. 2000*a* Dynamic biogeography and conservation of endangered species. *Nature* **403**, 84–86.

Channell, R. & Lomolino, M. V. 2000*b* Trajectories of extinction: spatial dynamics of the contraction of geographical ranges. *J. Biogeogr.* **27**, 169–179.

Curnutt, J. L., Pimm, S. L. & Maurer, B. A. 1996 Population variability of sparrows in space and time. *Oikos* **76**, 13–44.

Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. 1998 Making mistakes when predicting shifts in response to global warming. *Nature* **391**, 783–786.

Donald, P. F. & Greenwood, J. J. D. 2001 Spatial patterns of range contraction in British breeding birds. *Ibis* **143**, 593–601.

Gates, S. & Donald, P. F. 2000 Local extinction of British farmland birds and the prediction of further loss. *J. Appl. Ecol.* **37**, 806–820.

Gibbons, D. W., Reid, J. B. & Chapman, R. A. 1993 *The new atlas of breeding birds in Britain and Ireland: 1988–1991*. London: Poyser.

Gonzalez, A., Lawton, J. H., Gilbert, F. S. & Blackburn, T. M. 1998 Metapopulation dynamics, abundance, and distribution in a microsystem. *Science* **281**, 2045–2047.

Greig-Smith, P. 1983 *Quantitative plant ecology*, 3rd edn. Oxford: Blackwell.

Hanski, I., Moilanen, A., Pakkala, T. & Kuussaari, M. 1996 The quantitative incidence function model and persistence of an endangered butterfly metapopulation. *Conserv. Biol.* **10**, 578–590.

He, F. & Gaston, K. J. 2000 Estimating species abundance from occurrence. *Am. Nat.* **156**, 553–559.

Hodkinson, I. D. 1999 Species response to global environmental change or why ecophysiological models are important: a reply to Davis *et al*. *J. Anim. Ecol.* **68**, 1259–1262.

Huntley, B., Berry, P. M., Cramer, W. & McDonald, A. P. 1995 Modelling present and potential future ranges of some European higher plants using climate response surfaces. *J. Biogeogr.* **21**, 967–1001.

Hutchinson, G. E. 1957 Concluding remarks. In *Symp. Quantitative Biol., Cold Spring Harbor*, vol. 22, pp. 415–427.

Lawton, J. H. 1993 Range, population abundance and conservation. *Trends Ecol. Evol.* **8**, 409–413.

Legendre, P. 1993 Spatial autocorrelation: trouble or new paradigm. *Ecology* **74**, 1659–1673.

Lindenmayer, D. B., Ball, I., Possingham, H. P., Mccarthy, M. A. & Pope, M. L. 2001 A landscape-scale test of the predictive ability of a spatially explicit model for the population viability analysis. *J. Appl. Ecol.* **38**, 36–48.

Ludwig, D. 1999 Is it meaningful to estimate a probability of extinction? *Ecology* **80**, 298–310.

McArdle, B. H., Gaston, K. J. & Lawton, J. H. 1990 Variation in the size of animal populations: patterns, problems and artefacts. *J. Anim. Ecol.* **59**, 439–454.

Margules, C. R. & Nicholls, A. O. 1987 Assessing the conservation value of remnant habitat 'islands': mallee patches on the western Eyre Peninsula, South Australia. In *Nature conservation: the role of remnants of native vegetation* (ed. D. A. Saunders, G. W. Arnold, A. A. Burbidge & A. J. M. Hopkins), pp. 89–102. Canberra, Australia: CSIRO.

Margules, C. R. & Pressey, R. L. 2000 Systematic conservation planning. *Nature* **405**, 243–253.

Margules, C. R., Nicholls, A. O. & Pressey, R. L. 1988 Selecting networks of reserves to maximise biological diversity. *Biol. Conserv.* **43**, 63–76.

Margules, C. R., Nicholls, A. O. & Usher, M. B. 1994 Appar-

ent species turnover, probability of extinction and the selection of nature reserves: a case study of the Ingleborough Limestone pavements. *Conserv. Biol.* **8**, 398–409.

Nagelkerke, N. J. D. 1991 A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691–692.

Nicholls, A. O. 1998 Integrating population abundance, dynamics and distribution into broad scale priority setting. In *Conservation in a changing world: integrating process into priorities for action* (ed. G. Mace, A. Balmford & J. R. Ginsberg), pp. 251–272. Cambridge, UK: Cambridge University Press.

Parmesan, C. (and 12 others) 1999 Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* **399**, 579–583.

Pearce, J. & Ferrier, S. 2001 The practical value of modelling abundance of species for regional conservation planning: a case study. *Biol. Conserv.* **98**, 33–43.

Peterson, A. T., Ortega-Huerta, M. A., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, R. H. & Stockwell, D. R. B. 2002 Future projections for Mexican faunas under global climate change scenarios. *Nature* **416**, 626–629.

Pimm, S. L. 1991 *The balance of nature? Ecological issues in conservation of species and communities*. Chicago Press.

Pimm, S. L. & Lawton, J. H. 1998 Planning for biodiversity. *Science* **279**, 2068–2069.

Pimm, S. L. & Redfearn, A. 1989 The variability of population densities. *Nature* **334**, 613–614.

Rodrigues, A. S., Gregory, R. D. & Gaston, K. 2000 Robustness of reserve selection procedures under temporal species turnover. *Proc. R. Soc. Lond.* B **267**, 49–55. (DOI 10.1098/rspb.2000.0965.)

Segurado, P. & Araújo, M. B. 2002 Modelling species probabilities of occurrence: what method performs best and when? *J. Appl. Ecol.* (Submitted.)

Sharrock, J. T. R. 1976 *The atlas of breeding birds of Britain and Ireland*. Berkhamsted, UK: Poyser.

Smith, P. A. 1994 Autocorrelation in logistic regression modelling of species distributions. *Global Ecol. Biogeogr. Lett.* **4**, 47–61.

Soulé, M. E. & Sanjayan, M. A. 1998 Conservation targets: do they help? *Science* **279**, 2060–2061.

Thomas, C. D. & Lennon, J. J. 1999 Birds extend their ranges northwards. *Nature* **399**, 213.

Thomas, C. D., Bodsworth, E. J., Wilson, R. J., Simmons, A. D., Davies, Z. G., Musche, M. & Conradt, L. 2001 Ecological and evolutionary processes at expanding range margins. *Nature* **411**, 577–581.

Virolainen, K. M., Virola, T., Suhonen, J., Kuitunen, M., Lammin, A. & Siikamaki, P. 1999 Selecting networks of nature reserves: methods do affect the long term outcome. *Proc. R. Soc. Lond.* B **266**, 1141–1146. (DOI 10.1098/rspb.1999.0755.)

Wahlberg, N., Moilanen, A. & Hanski, I. 1996 Predicting the occurrence of endangered species in fragmented landscapes. *Science* **273**, 1536–1538.

Williams, P. H. 1998 Key sites for conservation: area-selection methods for biodiversity. In *Conservation in a changing world: integrating process into priorities for action* (ed. G. Mace, A. Balmford & J. R. Ginsberg), pp. 211–249. Cambridge, UK: Cambridge University Press.

Williams, P.H. 1999 WORLDMAP *v. 4.2 WINDOWS: Software and user document 4.1*. See http://www.nhm.ac.uk/science/projects/worldmap. London: Privately distributed.

Williams, P. H. & Araújo, M. B. 2000 Using probabilities of persistence to identify important areas for biodiversity. *Proc. R. Soc. Lond.* B **267**, 1959–1966. (DOI 10.1098/rspb.2000.1236.)

Williams, P. H. & Araújo, M. B. 2002 Apples, oranges and probabilities: integrating multiple factors into biodiversity conservation with consistency. *Environ. Model. Assess.* **7**, 139–151.