

Supporting Information

Beale *et al.* 10.1073/pnas.0803506105

SI Text

Full Power Analysis. Statistical power in these null models is expected to be a function of both a species prevalence (determining how much information there is in a pattern) and signal strength (the tightness of the association between species distribution and climate). We assessed the effects of prevalence and signal strength by using the same null model-generating procedure described in the main text: the rank of the climate envelope model AUC score for the focal pattern within the AUC scores of statistical replicates conserving the spatial autocorrelation structure of the focal pattern but independent of climate.

To assess the information effect, using the prevalence for each real species, we generated a pattern completely determined by climate. When we used the null model test described in the main text on these deterministic “species,” all but one were correctly identified as matching climate, indicating that any failure to detect associations between observed species distributions and climate is not simply attributable to failure of the analytical method. The one pattern that was not correctly identified was the pattern with the lowest prevalence: presence in only three squares. This appears to be a result of our climate envelope for this species having an exceptionally low AUC score.

To assess the signal-to-noise effect, for each species we used its prevalence to generate a single artificial distribution that was part deterministic and part noise. These patterns were generated as described in the main text, by first producing a fully deterministic pattern with the same prevalence as each real species, then allowing this species to disperse into all neighboring squares. This generates a pattern with false positives and higher prevalence than the true species distribution, so we sequentially removed presences from these patterns until the pattern had the desired prevalence again, selecting squares to remove that would improve the match between the spatial correlation structure of the simulated distribution and the true species distribution.

This algorithm [implemented in the function `allClim()` below] necessarily results in patterns that have a correlation between prevalence and signal strength: allowing a species that has very low prevalence to disperse into neighboring squares will result in a new pattern with a proportionally higher false positive rate than doing the same to a species with initially much higher prevalence. After this process it is quite possible for patterns with very low initial prevalence to have prevalence several times higher than the initial prevalence: this is clearly impossible for a species with much higher initial prevalence. On “erosion” back to the original prevalence, the pattern with proportionally more false positives will also risk losing more of the original deterministic pattern than a pattern with initially high prevalence. Both these processes and our definition of noise result in higher noise in patterns with low prevalence than those with high prevalence. Although this may be seen as a drawback of using a biologically plausible pattern generation method, it is worth noting that using our definition of signal strength (the proportion of an original deterministic pattern that remains in the final pattern after dispersal and erosion), whatever method is chosen, it is completely impossible to generate patterns simultaneously with high prevalence and low signal strength. Moreover, because the effect of prevalence is symmetrical, ≈ 0.5 our method, when combined with the completely deterministic patterns, samples most of the practical parameter space.

For the patterns generated with both deterministic signal and noise, Kappa for the climate envelope models ranged from 0 to 0.80 (median, 0.54) and AUC values ranged from 0.73 to 0.98

(median, 0.90). We found that AUC scores of ANN models for the deterministic patterns were ranked in the top 5 of the 99 simulations for 75 species. This means that, in our sample, adding a variable amount of noise to a perfect distribution–climate match reduces the number of correctly identified patterns from 99 to 75.

To further explore the relationships between signal strength, prevalence, and power we fitted a binomial logistic regression, modeling the probability of successfully fitting a climate envelope as a quadratic function of both signal strength and prevalence. Because the information contained within a binary pattern is symmetrical and highest at a prevalence of 0.5, we constrained the relationship with prevalence to be symmetrical ≈ 0.5 . As expected from the separate analyses, we found that power was strongly affected by signal strength and weakly affected by prevalence, with particularly good power (>0.8) for patterns that were at least 50% deterministic [supporting information (SI) Fig. S1]. This means that adding 50% noise to a perfect distribution–climate match would still result in our method correctly identifying $>80\%$ of patterns. We consider that as our methods are powerful at detecting all patterns with at least 50% deterministic signal, they should identify meaningful climate envelopes for real species patterns. We note that choosing a less strict α -level for significance testing (e.g., 0.1) will result in increased type I errors, but will further increase the power of these tests. Note also that if the middomain effect mentioned in the main text were influencing our results, we would expect to see low power with intermediate presence, not the high power we observe here.

Congruence of Modeling Methods, Goodness-of-Fit Statistics, and Climate Variables. For each of the 100 real species we generated 99 null distributions with similar spatial structure and prevalence. To each of these 10,000 distribution patterns we fitted climate envelopes by using three methods and two separate datasets, resulting in 90,000 models. We also generated 200 further patterns, each with 99 null distributions during the power analysis, resulting in a total of 150,000 unique climate envelopes, each summarized by an AUC score, a Kappa statistic, and measures of sensitivity and specificity. For the 300 actual climate envelopes for real distributions on real climate variables, Kappa, sensitivity, and specificity values ranged from 0 to 1 (median, 0.53 for Kappa) and AUC values ranged from 0.27 to 0.997 (median, 0.851). Four GAM models (for species with very low prevalence) failed to converge. Within this overall range, models fitted by using GLM and ANN generally resulted in slightly higher AUC than those fitted with GAM ($F_{2,293} = 91.4$, $P < 0.0001$), but the difference was not significant when measured with Kappa statistics ($F_{2,293} = 0.567$, $P = 0.567$). To assess the impact of choosing different modeling methods, different climate variables, and different goodness-of-fit measures, we assessed the congruence between the methods by using Kendall’s W with significance assessed by permutation. All fitting methods, all climate variable sets, and all four goodness-of-fit measures were congruent overall ($W = 0.552$; $P < 0.001$) and each individual ranking was congruent with all others ($P < 0.001$ in all cases). The rankings within the 99 null distributions given to each individual species distributions were therefore similar irrespective of modeling choice, climate variable, or preferred goodness-of-fit statistic. Overall numbers of species identified as significant varied slightly, but our results are qualitatively similar whatever method is chosen (Table S2). Because all results were

broadly similar we concentrate the main text on the results of ANN models and AUC scores, methods that are regularly preferred in comparative studies (1, 2) and were closest to the median in our sample of methods. We anticipate that our results would be broadly similar by using any of the additional modeling methods currently used in climate envelope studies.

The Effects of Choosing Alternative Significance Levels. In the main text we chose a significance level of 0.05. This implies that we expect 5 of the 32 patterns to show significant associations with climate by chance alone. Choosing a more stringent significance level in line with adjustments for multiple testing would further reduce the number of species we identify as significantly associated with climate. In this analysis, however, we chose not to

adjust significance levels for multiple testing, as the more conservative option. Even more conservative would be choosing a significance level of 0.1; this would increase by another five the number of patterns expected to be significantly associated with climate by chance alone, but would increase the power of the test. In our dataset, however, this still identifies only 46 of the patterns as a better match than chance alone, 10 of which may be expected as false positives. Further changes in the significance level result in no more increases in the number of patterns identified as significantly associated with climate than is completely consistent with the increased frequency of false positives, suggesting that the choice of significance level does not substantially affect our findings.

1. Huntley B, *et al.* (2004) The performance of models relating species geographical distributions to climate is independent of trophic level. *Ecol Lett* 7:417–426.

2. McPherson JM, Jetz W, Rogers DJ (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J Appl Ecol* 41:811–823.

Table S1. Species list: Describing the 100 species included in the analysis, the reason for their inclusion (used as a nuisance variable in one analysis), their prevalence, the mean latitude of their distribution in the study area, the AUC score for the climate envelope fitted with neural networks and an indication of whether the AUC score was in the top 5% of AUC scores for the null models

Species	Endemicity (reason for inclusion)	Prevalence	Mean latitude (°N)	AUC	Migrant status	Wetland habitat	Well fitted
<i>Accipiter gentilis</i>	Widespread	0.59	51.9	0.82	PM	FALSE	FALSE
<i>Accipiter nisus</i>	Widespread	0.64	51.52	0.74	PM	FALSE	FALSE
<i>Acrocephalus paludicola</i>	Endemic	0.01	52.49	0.93	LD	TRUE	FALSE
<i>Acrocephalus palustris</i>	Widespread	0.35	51.16	0.91	LD	FALSE	TRUE
<i>Alectoris graeca</i>	Endemic	0.07	42.24	0.82	R	FALSE	FALSE
<i>Alectoris rufa</i>	Endemic	0.18	43.79	0.93	R	FALSE	FALSE
<i>Anthus campestris</i>	Widespread	0.23	45.19	0.82	LD	FALSE	TRUE
<i>Anthus petrosus</i>	Endemic	0.07	58.13	0.89	R	FALSE	FALSE
<i>Anthus pratensis</i>	>60%	0.52	56.82	0.89	PM	FALSE	FALSE
<i>Aquila pomarina</i>	>60%	0.11	50.11	0.92	LD	FALSE	TRUE
<i>Ardea cinerea</i>	Widespread	0.43	51.51	0.81	PM	FALSE	FALSE
<i>Athene noctua</i>	Widespread	0.47	45.41	0.84	R	FALSE	FALSE
<i>Bubulcus ibis</i>	Widespread	0.02	39.81	0.93	R	FALSE	FALSE
<i>Buteo buteo</i>	Widespread	0.72	49.48	0.85	R	FALSE	TRUE
<i>Buteo rufinus</i>	Widespread	0.03	40.37	0.93	R	FALSE	FALSE
<i>Calandrella brachydactyla</i>	Widespread	0.12	40.67	0.89	LD	FALSE	TRUE
<i>Calidris alpina</i>	Widespread	0.11	61.16	0.89	LD	TRUE	TRUE
<i>Calidris maritima</i>	Widespread	0.05	64.5	0.98	SD	TRUE	TRUE
<i>Caprimulgus europaeus</i>	Widespread	0.35	48.46	0.73	LD	FALSE	FALSE
<i>Carduelis carduelis</i>	Widespread	0.69	47.22	0.87	R	FALSE	FALSE
<i>Carduelis chloris</i>	Widespread	0.77	49.69	0.76	R	FALSE	FALSE
<i>Carduelis flavirostris</i>	Widespread	0.05	60.54	0.97	PM	FALSE	TRUE
<i>Certhia brachydactyla</i>	Endemic	0.36	46.08	0.82	R	FALSE	FALSE
<i>Certhia familiaris</i>	Widespread	0.47	53.31	0.88	R	FALSE	TRUE
<i>Cettia cetti</i>	Widespread	0.17	41.62	0.91	R	TRUE	TRUE
<i>Charadrius morinellus</i>	Widespread	0.05	62.12	0.91	LD	FALSE	FALSE
<i>Chlidonias niger</i>	Widespread	0.14	51.29	0.87	LD	TRUE	FALSE
<i>Ciconia ciconia</i>	>60%	0.38	47.93	0.91	LD	TRUE	FALSE
<i>Coccothraustes coccothraustes</i>	Widespread	0.43	49.18	0.89	R	FALSE	TRUE
<i>Corvus corone</i>	Widespread	0.78	50.18	0.77	R	FALSE	FALSE
<i>Corvus frugilegus</i>	Widespread	0.37	50.9	0.9	R	FALSE	TRUE
<i>Cyanopica cyanus</i>	Widespread	0.03	39.27	0.95	R	FALSE	FALSE
<i>Delichon urbica</i>	Widespread	0.88	50.32	0.8	LD	FALSE	FALSE
<i>Dendrocopos leucotos</i>	Widespread	0.09	51.03	0.79	R	FALSE	FALSE
<i>Dendrocopos medius</i>	>60%	0.23	48.5	0.87	R	FALSE	FALSE
<i>Emberiza caesia</i>	Widespread	0.04	38.42	0.95	LD	FALSE	FALSE
<i>Emberiza cia</i>	Widespread	0.19	42.36	0.77	R	FALSE	FALSE
<i>Emberiza cirrus</i>	>60%	0.25	42.37	0.89	R	FALSE	FALSE
<i>Emberiza citrinella</i>	Widespread	0.64	52.29	0.9	PM	FALSE	TRUE
<i>Falco columbaris</i>	Widespread	0.16	63.17	0.97	PM	FALSE	TRUE
<i>Falco eleonorae</i>	>60%	0.02	38.24	0.98	LD	FALSE	TRUE
<i>Falco peregrinus</i>	Widespread	0.23	44.33	0.81	PM	FALSE	FALSE
<i>Falco subbuteo</i>	Widespread	0.43	50.3	0.76	LD	FALSE	FALSE
<i>Ficedula albicollis</i>	>60%	0.1	48.07	0.86	LD	FALSE	FALSE
<i>Ficedula semitorquata</i>	Widespread	0.02	41.08	0.88	LD	FALSE	FALSE
<i>Fringilla coelebs</i>	Widespread	0.84	50.4	0.78	PM	FALSE	FALSE
<i>Fulmarus glacialis</i>	Widespread	0.06	57.92	0.98	PM	FALSE	TRUE
<i>Glareola pratincola</i>	Widespread	0.03	40.52	0.9	LD	FALSE	FALSE
<i>Glaucidium passerinum</i>	Widespread	0.12	57.65	0.91	R	FALSE	TRUE
<i>Gypaetus barbatus</i>	Widespread	0.01	40.14	0.85	R	FALSE	FALSE
<i>Haematopus ostralegus</i>	Widespread	0.22	57.51	0.85	SD	TRUE	FALSE
<i>Himantopus himantopus</i>	Widespread	0.08	41.95	0.88	PM	TRUE	TRUE
<i>Hippolais pallida</i>	Widespread	0.08	40.91	0.91	LD	FALSE	FALSE
<i>Hippolais polyglotta</i>	Endemic	0.17	43.51	0.88	LD	FALSE	FALSE
<i>Lagopus lagopus</i>	Widespread	0.16	63.11	0.99	R	FALSE	TRUE
<i>Lanius minor</i>	Widespread	0.13	44.2	0.87	LD	FALSE	FALSE
<i>Lanius nubicus</i>	Widespread	0.02	40.16	0.97	LD	FALSE	FALSE
<i>Locustella luscinioides</i>	Widespread	0.14	48.38	0.82	LD	TRUE	FALSE
<i>Locustella naevia</i>	Widespread	0.22	52.14	0.84	LD	FALSE	FALSE
<i>Loxia scotica</i>	Endemic	0	57.05	0.99	R	FALSE	FALSE

Species	Endemicity (reason for inclusion)	Prevalence	Mean latitude (°N)	AUC	Migrant status	Wetland habitat	Well fitted
<i>Lullula arborea</i>	>60%	0.42	47.38	0.76	PM	FALSE	FALSE
<i>Luscinia luscinia</i>	Widespread	0.2	54.45	0.92	LD	FALSE	FALSE
<i>Lymnocyptes minimus</i>	Widespread	0.03	64.91	0.92	SD	TRUE	FALSE
<i>Melanocorypha calandra</i>	Widespread	0.1	40.83	0.91	PM	FALSE	TRUE
<i>Merops apiaster</i>	Widespread	0.24	42.73	0.87	LD	FALSE	TRUE
<i>Milvus milvus</i>	>60%	0.15	47.77	0.86	PM	FALSE	FALSE
<i>Motacilla alba</i>	Widespread	0.82	50.78	0.75	SD	FALSE	FALSE
<i>Nyctea scandiaca</i>	Widespread	0.02	67.46	0.98	R	FALSE	TRUE
<i>Otis tarda</i>	Widespread	0.04	42.46	0.91	PM	FALSE	FALSE
<i>Parus cinctus</i>	Widespread	0.04	66.76	0.99	R	FALSE	TRUE
<i>Parus montanus</i>	Widespread	0.48	55.28	0.89	R	FALSE	TRUE
<i>Parus palustris</i>	Widespread	0.46	50.49	0.84	R	FALSE	FALSE
<i>Passer domesticus</i>	Widespread	0.84	50.53	0.79	R	FALSE	FALSE
<i>Passer hispaniolensis</i>	Widespread	0.1	40.31	0.93	R	FALSE	FALSE
<i>Pernis apivorus</i>	>60%	0.38	51.23	0.8	LD	FALSE	TRUE
<i>Phylloscopus bonelli</i>	>60%	0.13	43.47	0.79	LD	FALSE	FALSE
<i>Phylloscopus collybita</i>	Widespread	0.61	51.11	0.8	SD	FALSE	FALSE
<i>Phylloscopus sibilatrix</i>	Widespread	0.45	52.68	0.85	LD	FALSE	FALSE
<i>Pica pica</i>	Widespread	0.83	50.85	0.86	R	FALSE	TRUE
<i>Pinicola enucleator</i>	Widespread	0.04	65.56	0.99	R	FALSE	TRUE
<i>Prunella modularis</i>	>60%	0.57	53.08	0.82	R	FALSE	FALSE
<i>Pyrrhocorax pyrrhocorax</i>	Widespread	0.09	42.05	0.86	R	FALSE	FALSE
<i>Regulus ignicapillus</i>	>60%	0.23	46.34	0.83	PM	FALSE	FALSE
<i>Regulus regulus</i>	Widespread	0.47	53.56	0.82	PM	FALSE	FALSE
<i>Saxicola torquata</i>	Widespread	0.47	45.4	0.85	PM	FALSE	FALSE
<i>Scolopax rusticola</i>	Widespread	0.38	54.97	0.78	SD	FALSE	FALSE
<i>Serinus citrinella</i>	Endemic	0.04	44.27	0.81	R	FALSE	FALSE
<i>Serinus serinus</i>	>60%	0.51	45.73	0.86	PM	FALSE	FALSE
<i>Sitta whiteheadi</i>	Endemic	0	42.08	0.26	R	FALSE	FALSE
<i>Strix aluco</i>	Widespread	0.59	49.55	0.79	R	FALSE	FALSE
<i>Strix nebulosa</i>	Widespread	0.03	63.35	0.96	R	FALSE	TRUE
<i>Strix uralensis</i>	Widespread	0.09	59.19	0.94	R	FALSE	TRUE
<i>Sylvia atricapilla</i>	Widespread	0.72	49.38	0.81	SD	FALSE	FALSE
<i>Sylvia cantillata</i>	Endemic	0.16	40.6	0.93	LD	FALSE	TRUE
<i>Sylvia sarda</i>	Endemic	0.01	40.38	0.98	R	FALSE	FALSE
<i>Sylvia undata</i>	Endemic	0.11	41.57	0.91	R	FALSE	FALSE
<i>Tetrao urogallus</i>	Widespread	0.24	57.94	0.95	R	FALSE	TRUE
<i>Troglodytes troglodytes</i>	Widespread	0.73	50	0.79	R	FALSE	FALSE
<i>Turdus merula</i>	Widespread	0.84	49.63	0.81	PM	FALSE	FALSE
<i>Turdus torquatus</i>	>60%	0.15	53.78	0.85	SD	FALSE	TRUE

Migrant status (PM = partial migrant, LD = long-distance migrant, SD = short-distance migrant, R = resident) and wetland habitat are partially subjective assignments, but neither showed significant relationships in logistic regression models with how well the species distributions were modeled: Migrant status $\chi^2 = 0.722$, $df = 3$, $P = 0.868$; Wetland habitat $\chi^2 = 0.638$, $df = 1$, $P = 0.424$. Because these values are nonsignificant phylogenetic adjustment will not change the results.

