ECOGRAPHY

Research

Factors influencing transferability in species distribution models

Josée S. Rousseau and Matthew G. Betts

J. S. Rousseau (https://orcid.org/0000-0001-5797-3565) ✉ (josee.s.rousseau@gmail.com) and M. G. Betts (https://orcid.org/0000-0002-9131-8844), Forest Biodiversity Research Network, Forest Ecosystems and Society, Oregon State Univ., Corvallis, OR, USA. JSR also at: Cornell Lab of Ornithology, Ithaca, NY, USA.

Ecography
2022: e06060
doi: 10.1111/ecog.06060
Subject Editor: Carsten Dormann
Editor-in-Chief: Miguel Araújo
Accepted 2 February 2022

Species distribution models (SDMs) provide insights into species' ecology and distributions and are frequently used to guide conservation priorities. However, many uses of SDMs require model transferability, which refers to the degree to which a model built in one place or time can successfully predict distributions in a different place or time. If a species' model has high spatial transferability, the relationship between abundance and predictor variables should be consistent across a geographical distribution. We used Breeding Bird Surveys, climate and remote sensing data, and a novel method for quantifying model transferability to test whether SDMs can be transferred across the geographic ranges of 129 species of North American birds. We also assessed whether species' traits are correlated with model transferability. We expected that prediction accuracy between modeled regions should decrease with 1) geographical distance, 2) degree of extrapolation and 3) the distance from the core of a species' range. Our results suggest that very few species have a high model transferability index (MTI). Species with large distributions, with distributions located in areas with low topographic relief, and with short lifespans are more likely to exhibit low transferability. Transferability between modeled regions also decreased with geographical distance and degree of extrapolation. We expect that low transferability in SDMs potentially resulted from both ecological non-stationarity (i.e. biological differences within a species across its range) and over-extrapolation. Accounting for non-stationarity and extrapolation should substantially increase the prediction success of species distribution models, therefore enhancing the success of conservation efforts.

Keywords: extrapolation, model transfer, species distribution models, stationarity, traits

Introduction

The accuracy of species distribution models (SDMs) is essential for sound conservation decision making. By predicting the presence of a species through time and space, biologists and managers can use SDMs to forecast biological invasions, identify critical habitats, prioritize the locations of reserves, appropriately translocate endangered species (Guisan et al. 2013), and inform large-scale land management practices (Thomas et al.

NORDIC SOCIETY OIKOS
www.ecography.org

2006, Millar et al. 2007). While SDMs are a ubiquitous tool, they require several assumptions that may not be realistic, decreasing prediction accuracy and therefore model transferability (Guisan and Thuiller 2005, Pearman et al. 2008, Elith and Leathwick 2009).

First, the underlying habitat associations need to exhibit 'stationarity' in space and time. We define stationarity here as a process for which the parameters are temporally and spatially consistent (Miller 2012, Dale and Fortin 2014). In other words, to have a stationary SDM, a species needs to be associated with the same abiotic and biotic factors, i.e. have a consistent ecological niche (Hutchinson 1957), over time and at all locations. However, theory and empirical evidence suggest it is often not the case. Spatial variation in species' niches may occur if there are genetic differences among populations (Pearman et al. 2008) or if a species is facing different degrees and types of competition, predation and diseases across their range (Araújo and Luoto 2007, Daskin and Alford 2012, Chamberlain et al. 2014, Vergnon et al. 2017). There are some indications that the assumption of stationarity may not always hold true (Whittingham et al. 2007, Fink et al. 2010, Schmidt et al. 2014, Shirk et al. 2014, Howard et al. 2015, Laube et al. 2015, Gómez et al. 2016, Zuckerberg et al. 2016, Wan et al. 2017), and while the inclusion of biotic factors in SDMs would likely reduce the problem of non-stationarity, it is often impractical to include them.

Second, models must be well specified, such that the entire range of potential environmental variables are present in data used for SDMs. For instance, if a model is trained in a location with low to moderate values of a strongly limiting variable (e.g. precipitation) the variable will emerge as an important SDM predictor. However, if the same model is extrapolated to a region where precipitation is no longer limiting for the species, the model will likely fail (Owens et al. 2013, Betts et al. 2021). In this case, using a mechanistic model, rather than a correlative model, while also including the full range of potential environmental values is likely to improve model transferability (Higgins et al. 2020). In practice, it is challenging to separate difficulties in model transferability that are due to either low stationarity (i.e. 'true' differences in species niches across the geographic range) versus issues relating to extrapolation (Dale and Fortin 2014).

Regardless of the mechanism, research is scant on which species, and more specifically which species traits, may be associated with lower model transferability. For example, life history traits, such as clutch size, number of broods and longevity may correlate with stationarity. Other traits such as the spatial extent and topography of a distribution, as well as primary habitat used, may result in models for some species' SDMs being better specified than others. Overall, understanding such correlates of poor model transferability will provide advance knowledge about the likely reliability of species distribution model predictions.

Within a species' geographic range, it is also likely that geographical, ecological and environmental features impact prediction accuracies. In theory, a regional model should predict equally well within and between regions of a distribution. However, at least three mechanisms may decrease prediction accuracy of models transferred in space.

First, it is widely known that ecological similarity decreases with increasing distance – a phenomenon known as Tobler's law in geography (Tobler 1970). Therefore, under this *geographic distance hypothesis* (Fig. 1a), we would expect species distribution model predictions to become less accurate with increasing physical distances between where a model has been built (i.e. trained) and where it is applied (i.e. tested; Yates et al. 2018). The degree to which extrapolation occurs is also likely to increase with geographic distance (for instance, topography, climate and landcover features are less likely to be similar as distance increases). Second, prediction accuracies may decrease between models that are located at different distances from a distribution's core (i.e. the geographic center of the range). Under this *core-boundary hypothesis* (Fig. 1b), prediction accuracies may increase if a model is transferred between core or boundary regions. This hypothesis is derived from the premise that species may have different levels of fitness, competition, gene flow and abundance across their ranges, following a gradient along the range's core to boundary (Sexton et al. 2009, Orme et al. 2019). This hypothesis is different from the *geographical distance hypothesis* because two regions may both be located at the boundary of a distribution yet be very geographically distant from each other. Lastly, models transferred between regions with overlapping ranges of environmental values (i.e. similar environmental space, termed 'interpolation') should have higher predictive accuracy than models transferred between regions where the environments are not analogous (Elith and Leathwick 2009, Qiao et al. 2019). We refer to this hypothesis as the *analogue environment hypothesis* (Fig. 1c). Whereas the *core-boundary hypothesis* tends to be associated with the concept of stationarity, non-analogue environments across space are likely to be more associated with poor model specification (Franklin 2010, Dale and Fortin 2014).

Here, we used 129 species of North American breeding birds to address three main questions: 1) Can we transfer SDMs across species' ranges? Because we expected variation across species, and to clarify under what conditions SDMs failed to predict species abundance, we addressed two additional questions. 2) What geographic and life history predictors can explain variation in species-level model transferability? Here, our results provide guidance on which species, based on their traits, are more likely to have high prediction accuracies across their range. 3) For species with lower prediction accuracies, what causes the decreases in model predictions across space? To answer this question, we tested three hypotheses associated with features of a distribution – the *geographical distance hypothesis*, the *core-boundary hypothesis* and the *analogue environment hypothesis* (Fig. 1).

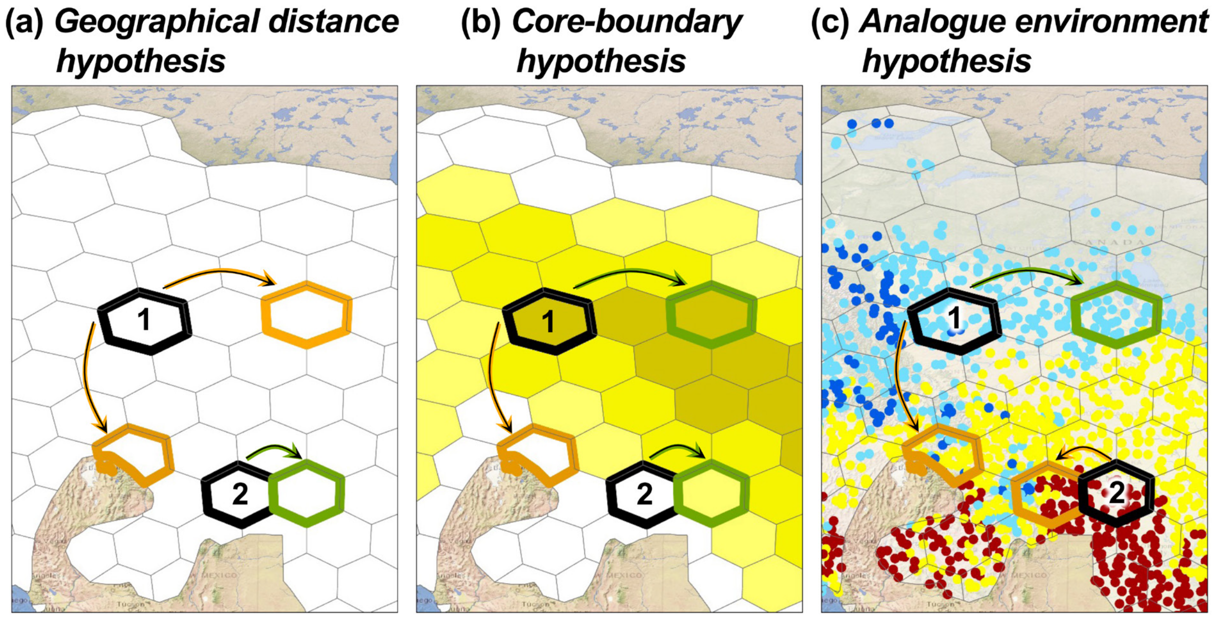

Figure 1. Graphical representation of the three hypotheses tested in our model transferability analysis. Black hexagons represent regions where a model was trained. Green hexagons represent regions where we expect good model transferability, and orange hexagons where we expect poor model transferability. (a) Under the *geographical distance hypothesis*, model transferability will be greatest over shorter geographical distances. A model trained in 1 would do poorly in the orange regions and a model trained at 2 would predict well the abundance in the neighboring green region. (b) The *core-boundary hypothesis* is symbolized by a yellow gradient across regions, where core regions are dark yellow and boundary regions are white. Under this hypothesis, a model at the core, such as model 1 would predict well other core regions such as the nearby green region. The same applies for two boundary regions such as 2 and its neighboring region, but a core region such as 1 would poorly predict the region in orange because it is at the range boundary. (c) Under the *analogue environment hypothesis*, regions with a similar range of environmental conditions should show high transferability. Here the color dots represent temperature, from warm in red, to cold in dark blue. Region 1 would predict well the green region because they have the same range of temperature (light blue dots) but poorly the regions in orange. Region 2 would also poorly predict the regions in orange, even if within close geographic proximity.

Material and methods

Data – bird species and abundance

We used data from the USGS Breeding Bird Survey (BBS) to extract abundances of bird species in Canada and USA (Pardieck et al. 2019). BBS consists of routes surveyed once a year during the breeding season (typically June). A BBS route includes 50 three-minute point counts, separated by a distance of at least 0.5 miles (Sauer et al. 2003).

We selected bird species (n = 138) using three criteria. First, a species had to be detected in at least 30 different routes per year. This provided for a minimum sample size in the number of routes where a species is present and helped ensure our models would predict reasonably well (Hernandez et al. 2006, Wisz et al. 2008). Second, we selected species with a prevalence (i.e. the percentage of routes where the species was present) of at least 20%. We also excluded highly common species (prevalence > 75%). These prevalence values are recommended to improve the fit of SDMs (McPherson et al. 2004). Third, a minimum of 80% of the breeding distribution of each species had to be within the area covered by the BBS routes. We used breeding bird geographic distributions from BirdLife International (2018). The area covered by the BBS routes was determined using a minimum convex polygon surrounding all BBS routes.

For each BBS route and species, we used the mean abundance for the years of 2013–2017 (inclusively; Howard et al. 2014). This range of years represents the latest five years available at the time of download. Using the mean abundance across a short time frame enabled us to reduce the noise caused by yearly changes in detections, while limiting the impact caused by long-term changes in habitat and climate on bird abundance (Gutiérrez-Illán et al. 2014, Betts et al. 2019).

Data – environmental covariates

We used climatic and land cover covariates known to be correlated with bird abundance (Austin 2002, Shirley et al. 2013, Howard et al. 2015). Data were obtained from Google Earth Engine (Gorelick et al. 2017) and were summarized for each BBS route and year (2013–2017), using a 400 m buffer (Bahn and McGill 2013). Datasets were selected based on their availability across North America. Climatic covariates were obtained from Daymet V3 (Thornton et al. 2017) and included summer precipitation (prcpSummer), winter precipitation (prcpWinter), maximum summer temperature (tMax) and minimum winter temperature (tMin). We used the equivalent of the band 3 (B3) and 4 (B4) of Landsat 7, from Landsat 5, 7 and 8 as land cover variables. These land cover data were summarized using the LandTrendr tools (Kennedy et al. 2018). LandTrendr includes pre-processing of the images including geometric rectification and cloud and shadow screening. It creates a yearly surface reflectance composite which we used to summarize data for each BBS route. We used B3 to discriminate between built-up environments and vegetation, and B4 to compare rates of chlorophyll absorption which is useful to distinguish between conifer and broadleaf as well as young versus senescent vegetation (Cohen and Goward 2004). The climatic and land cover covariates used in the analysis are summarized in Table 1. To be consistent with bird data and to increase model transferability (Tuanmu et al. 2011), covariates for each BBS route were then averaged over the period from 2013 to 2017. The number of BBS routes used for each species was dependent on the size of its breeding distribution (mean number of BBS routes per species: 1568.5 ± 1061.4 SD).

Model transferability analysis

We divided the distribution of each species into equal-sized hexagons of 106 088 km^2 (apothem of 175 km; we refer to these as 'regions'; Fig. 2). We established a hexagon size that jointly maximized both the within-hexagon number of BBS routes, and number of hexagons within each distribution. We determined this optimum via sensitivity analysis, in which we used different sizes of hexagons and species to assess the greatest number of hexagons containing at least 30 BBS routes (Luan et al. 2020). Out of 138 species, 129 had at least one pair of hexagons with a minimum of 30 BBS routes per hexagon.

We trained a Random Forest regression model (Breiman 2001, Liaw and Wiener 2002) on each hexagon, using mean abundance of each bird species per BBS route as response variables and the BBS route environmental variables (Table 1) as explanatory variables. For the Random Forest models, we used an ensemble of 1000 trees. To avoid overfitting, each tree was split using two randomly selected variables. We zero-centered and scaled by the standard deviations all predictors to avoid bias in the variable selection process (Boulesteix et al. 2012). To further lower risks of overfitting, we divided each hexagon into blocks of 130 km and used a 5-fold block validation approach (Roberts et al. 2017) to predict bird abundance within the same hexagon (referred to 'within-region predictions'). We assessed for the presence of bias in our prediction by plotting observed ~ predicted and found our models to be well-calibrated (Supporting information). The within-region prediction accuracy, the relationship between observed and predicted across blocked areas within the same hexagon, was calculated using a Pearson coefficient of correlation (r_i) (Yates et al. 2018). We used Random Forest because it is commonly used in SDMs (Syphard and Franklin 2009, Goetz et al. 2010, Matthews et al. 2011, Mainali et al. 2015, Rich and Currie 2018). It also has the advantage of modeling multiple interactions and relationships between variables (Evans et al. 2011, Boulesteix et al. 2012).

We paired all hexagons, keeping track of the directionality of each pair (i.e. keeping the pairs for hexagon 1 to 2 and hexagon 2 to 1). For each pair of hexagons, we used the training model from one hexagon to predict abundance in the second hexagon (referred to 'between-regions prediction'). We also assessed between-regions prediction accuracy using a Pearson coefficient of correlation (Qiao et al. 2019). Each species had a different number of Pearson coefficients based on the number of pairs of predictions (i.e. number of pairs of hexagons) within their distribution.

We used the within-region predictions per species to calculate a mean within-region Pearson ($\overline{r_i}$). Species with a high $\overline{r_i}$ were species for which models generally predicted well within the same hexagon. We selected species with a $\overline{r_i}$ > 0.2 for the remainder analyses. This value is associated with models that have a small positive predictive strength of association between predicted and observed values (Cohen 1992, Møller and Jennions 2002, Betts et al. 2019). We also completed a sensitivity analysis to assess if different cutoffs for $\overline{r_i}$ would yield different results. The direction of the results did not qualitatively change with different cutoff values.

We created a model transferability index (*MTI*) per species. This index represents the mean change in predictive value between the within-region and between-regions, such that:

$$\text{Model transferability index}(\text{MTI}) = 1 - \left(\frac{1}{i \times j} \sum_{i \times j}^{1} \left| r_i - r_{ij} \right| \right)$$

Table 1. Summary of the environmental covariates used in the modeling analysis.

Variable	Source	Pixel size	Cadence; range used	Summary per pixel, per year	Summary per route, per year
Band 3*	Landsat 5, 7 and 8	30 m averaged to 90 m	16 days; 15 May–15 July	Medoid**	Mean and SD
Band 4*	Landsat 5, 7 and 8	30 m averaged to 90 m	16 days; 15 May–15 July	Medoid**	Mean and SD
Precipitation summer	Daymet V3 – prcp	1000 m	Daily; 15 May–15 July	Sum	Mean
Precipitation winter	Daymet V3 – prcp	1000 m	Daily; 1 Jan.–15 March	Sum	Mean
Maximum summer temperature	Daymet V3 – tmax	1000 m	Daily; 15 May–15 July	Medoid**	Mean
Minimum winter temperature	Daymet V3 – tmin	1000 m	Daily; 1 Jan.–15 March	Medoid**	Mean

* Used LandTrendr to summarize the values.

** The mediod is obtained by selecting the image pixel with the smallest difference between the median from all layers and the observation, per band.

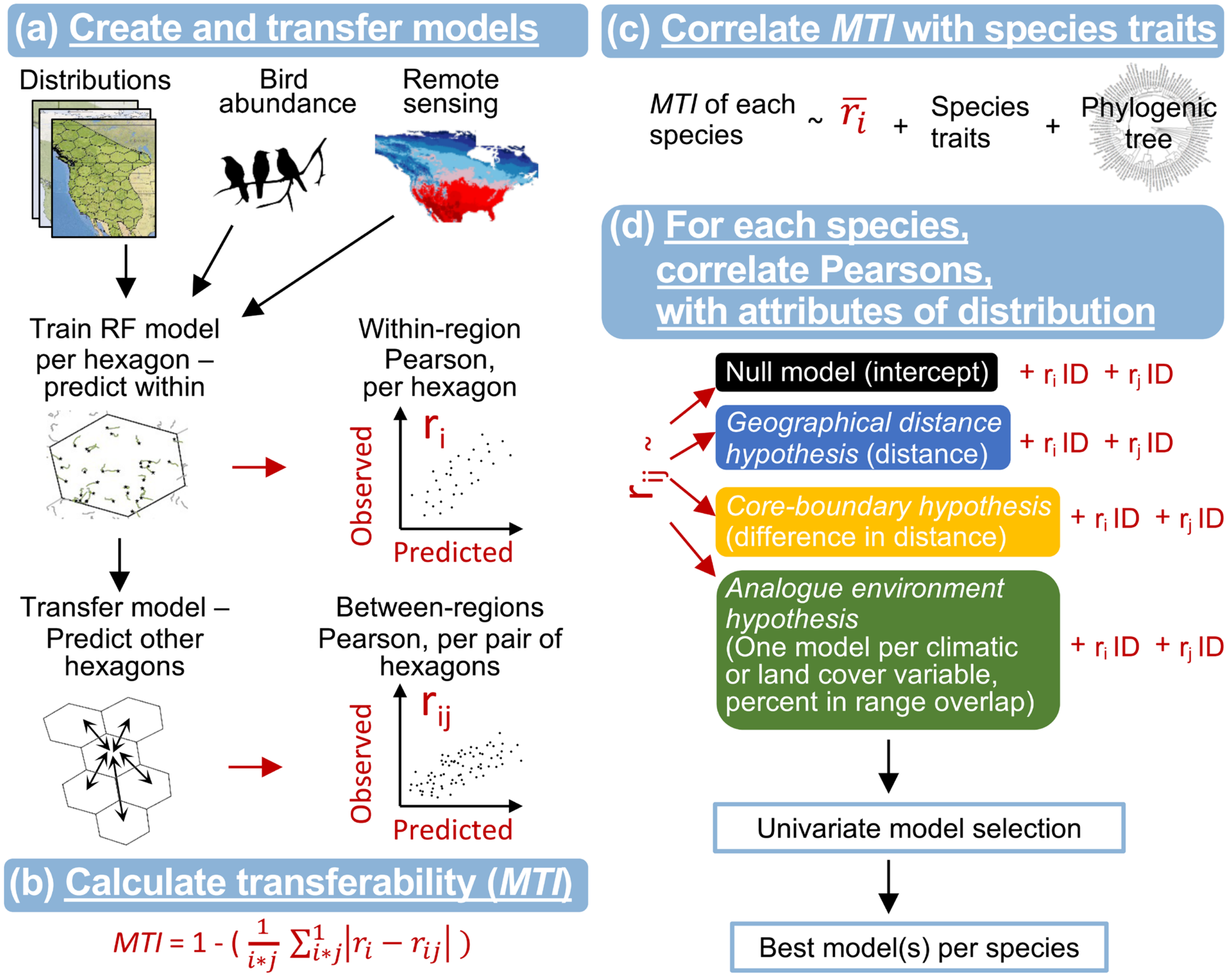

Figure 2. Summary of the methodology used in this paper. (a) We used bird distributions, abundance and remote sensing data to train models within small regions (hexagons) of each distribution. For each model, we calculated a Pearson coefficient of correlation (r_i) representing the within-region prediction accuracy. Each trained model was transferred to all other regions of the distribution. We calculated a between-regions prediction accuracy using Pearson (r_{ij}). (b) For each species, we calculated a model transferability index (*MTI*) using the average difference between r_i and r_{ij}. (c) We assessed whether *MTI* was correlated with species traits using a phylogenetic linear model. (d) For each species, we completed a univariate model selection analysis, using the between-region predictions (r_{ij}) as response variables and variables representing our three hypotheses as explanatory variables. We then compiled the best model(s) from all species to assess which hypothesis best fitted our species.

where r_i is the within-region Pearson coefficient of correlation for region i of each species, r_{ij} is the between-regions Pearson for a model trained in region i and tested in region j, and $i \times j$ represent the total number of pairs of regions. Species for which we can predict abundance between regions with high accuracy will have a small difference in Pearson coefficients between the within-region and between-regions. The index ranges from -1 to 1, but is most commonly positive, with mean differences closer to 1 representing overall high model transferability across a species range. Values close to 1 can either represent a high model transferability across a species range (even more so when the r_i is large), or a species for which the training model did not account well for the processes regulating abundance (i.e. species with a low r_i). In other words, the index is more meaningful for those species with a higher within-region mean Pearson, which is one reason we removed species with a very low mean Pearson ($\overline{r_i} \leq 0.2$). This index represents a gradient of model transferability and one would need to consider sample size (number of independent pairs) to assess the level of significance in *MTI* for each species (Cohen 1992).

Geographical and life history traits associated with high model transferability

We expected that the level of model transferability (*MTI*) would be predictable based on the geographical and life-history traits of species. For example, shorter lived species, and species with a greater number of broods and higher clutch size may have more rapid evolution and greater potential for adaptation to their environment (Vedder et al. 2013) as a result of quick turnover between generations; this could result in reduced model transferability across their ranges. Alternatively, long-lived species, which tend to have bigger brains (Minias et al. 2017), have a higher capacity for behavioral adaptation and would thus have a lower model transferability in their distributions. The geographic range size of a species and its topographic heterogeneity may also impact model transferability. Larger geographic ranges should have more varied populations across space (Phillimore et al. 2007) and would likely cross a more diverse set of habitats and climatic conditions. Further, in North America, western regions tend to have more topographic relief, which also offers greater diversity of habitats and climatic conditions than in the midwest region of North America. Spatial autocorrelation in biophysical features should occur at finer scales in the west where habitat heterogeneity is greater; we expect each of these factors could impact model transferability because they decrease the likelihood that models are extrapolated outside the bounds of environmental variables within the focal training hexagons. We also expected that the maximum summer temperature tolerated by a species within their ranges, i.e. the 10% warmest locations where the species is present (referred to as maximum temperature tolerance), could be correlated with model transferability. Jiguet et al. (2010) showed that bird species with higher thermal maxima are more likely to have positive population trends. These species may also be better adapted to a range of environmental conditions and have lower model prediction accuracies. Lastly other species traits such as migration status and primary habitat have been correlated with species' abilities to adapt to new environments (Pulido and Widmer 2005, Colles et al. 2009) and may affect model transferability.

We used a phylogenetic linear model and the R package 'phylolm' (Ho and Ané 2014) to assess if the response variable, *MTI*, was correlated with species traits (explanatory variables). We accounted for differences in within-region model fit of each species by including $\overline{r_i}$ as an explanatory variable (Fig. 2). A total of 107 species had phylogenetic data available and a $\overline{r_i} > 0.2$. The phylogenetic data were extracted from Jetz et al. 2012 (<www.birdtree.org>). We used 5000 trees from Ericson et al. (Ericson et al. 2006, Hackett et al. 2008) and 5000 trees from Hackett et al. (2008). We used the R package 'ape' (Paradis and Schliep 2019) to calculate one consensus tree from the combined 10 000 trees. This consensus tree was used in the linear model to account for the lack of independence between closely related species.

We included six traits in the phylogenetic linear model. Each corresponds to the hypotheses above: distribution size, topographic variation of the geographic range, maximum temperature tolerance, breeding habitat type, migration status and a single combined variable representing a slow–fast continuum in life history traits (Bennett and Owens 2002). We calculated topographic variation by calculating the standard deviation of the elevation pixels overlapping each distribution. We calculated maximum temperature tolerance for each species using the mean temperature of the warmest 10% BBS routes where the species was detected. We extracted breeding habitat information from the Partners In Flight database (Partners in Flight 2020). All species with the word 'forest' within the name of their primary breeding habitat were associated with the category 'forest', all others were associated with the category 'non-forest'. Based on Partners in Flight (2020), birds without the word forest as a primary habitat did not rely on forests during breeding. We visually assessed migration status using the distribution maps in Birds of the World (Billerman et al. 2020). All species with a summer distribution that is estimated to be more than 10% different from their distribution on other seasons were categorized as migrants, the others as residents. Lastly, we combined three life history traits into a principal component analysis (PCA): maximum longevity, number of broods and maximum clutch size. The life history traits for most species ($n = 102$) were extracted from Amniote (Myhrvold et al. 2015). Whenever data were absent from Amniote, we used the information from Birds of the World (Billerman et al. 2020). The PCA axis 1 explained 85.9% of the variance, with the 'longevity' responsible for 99.9% of its weight, therefore representing a 'slow-to-fast' continuum in avian life-histories (Bennett and Owens 2002). We implemented PCA using the R function prcomp (<www.r-project.org>). The axis 1 values were then used as input in the phylogenetic linear model.

Attributes of a distribution

In total, 96 species satisfied our criteria for inclusion in the analysis (between-regions $\overline{r_i} > 0.2$, 30 hexagon pairs). For each species, we tested whether the r_{ij} (response) were correlated with three aspects of a species distribution, representing our three hypotheses above (*geographic distance hypothesis*, *core-boundary hypothesis* and *analogue environment hypothesis*; Fig. 1). We tested whether the r_{ij}, for each species, were correlated with geographical distance between hexagons. We used the center point of each hexagon to calculate a great circle distance (as the crow flies) between paired regions.

We tested if the location of the hexagons, whether at the core of a distribution or towards its boundary (i.e. *core-boundary hypothesis*) affected prediction accuracy between regions. We first calculated the distance between a hexagon's center to the closest range boundary line. We then used the absolute difference between the two hexagons' distances to the boundary. Under this hypothesis, hexagons with the greatest distances between core and range edge should have the lowest prediction accuracy.

We tested the *analogue environment hypothesis* for each of the covariates used in the training and predicting models: prcpSummer, prcpWinter, tMax, tMin, B3 mean and standard deviation, and B4 mean and standard deviation. We first calculated the minimum and maximum value of each covariate within each hexagon, resulting in a covariate range of values. We then divided the amount of overlap between the two ranges (one per hexagon), by the range of values from the two hexagons combined. A range overlap of 100% could either represent a perfectly identical environmental range between paired hexagons, or more likely, that the range of the hexagon used for training included 100% of the values represented in the predicted hexagon. A percentage of overlap < 100% is associated with predictions that involved at least some degree of extrapolation. Negative percentages represent ranges that did not overlap. We calculated the percentage of extrapolation across regions, considering all 8 variables, for each species and pairs of hexagons (Supporting information).

We tested which of these three hypotheses were associated with each species using a univariate model selection approach (Fig. 2). This allowed us to include a greater number of species, at the cost of excluding possible interactions between variables. Each univariate model consisted of a linear mixed model with r_{ij} as a response, one feature as an explanatory variable, and two random effects, the training hexagon ID and the predicted hexagon ID. The two random effects were needed since each hexagon was used multiple times as both the training and as the predicted hexagon, resulting in a lack of independence among predictions. We used a Bayesian information criterion (BIC) approach to select the top model(s) per species because it controls for differences in sample size in model selection (Burnham and Anderson 2003, Vrieze 2012); this enabled comparisons among species with different sample sizes. BIC is also more consistent when simple models are considered (Vrieze 2012). We compared 11 univariate models: a null (intercept) model, a model representing geographical distance, one for the core versus boundary difference, and eight models representing the *analogue environment hypothesis*, one per climatic or land cover variable's percent range overlap. For each species, we retained the model(s) that had a delta BIC < 2. If the null model was among the best models, we only retained the null model.

Results

Model transferability

Across 129 species, we found a consistent decrease in model prediction accuracy when testing models within-region ($\overline{r}_i$) than when testing between-regions ($\overline{r}_{ij}$; Fig. 3a). A total of 109 species, out of the 129, had a $\overline{r}_i > 0.2$. These species had a mean decrease in prediction accuracy of 43.4% ± 30.0 between $\overline{r}_i$ and $\overline{r}_{ij}$, and a mean model transferability index of 0.72 ± 0.09 (Fig. 4b; Supporting information).

Species traits

As expected, species with larger geographic ranges tended to have lower prediction accuracies and associated *MTIs* (Table 2; Supporting information). Topographic variation within distributions was also important; species inhabiting regions with greater topographic variation are more likely to have higher prediction accuracies across their SDMs (Table 2; Supporting information). Finally, we found that SDMs for short-lived species tended to have lower *MTIs* (Table 2; Supporting information).

Attributes of distributions

Several attributes of species' geographic ranges were useful predictors of model transferability between-regions. Of

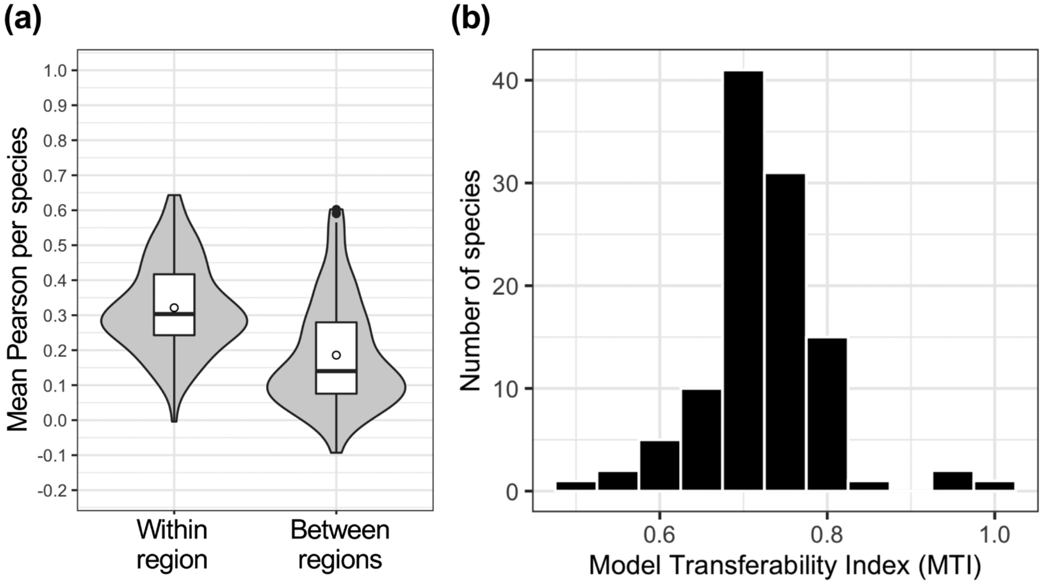

Figure 3. (a) Comparisons of the mean within-region Pearson ($\overline{r}_i$) per species and mean between-regions Pearson ($\overline{r}_{ij}$). The models predicted within the same region had substantially higher prediction accuracy than between regions. The white circles represent the mean per category. The violin plots represent the distribution of means among species and the boxplot shows the quartile spread of the data. (b) Histogram of the model transferability index (*MTI*) per species, where large values of *MTI* indicate high model transferability.

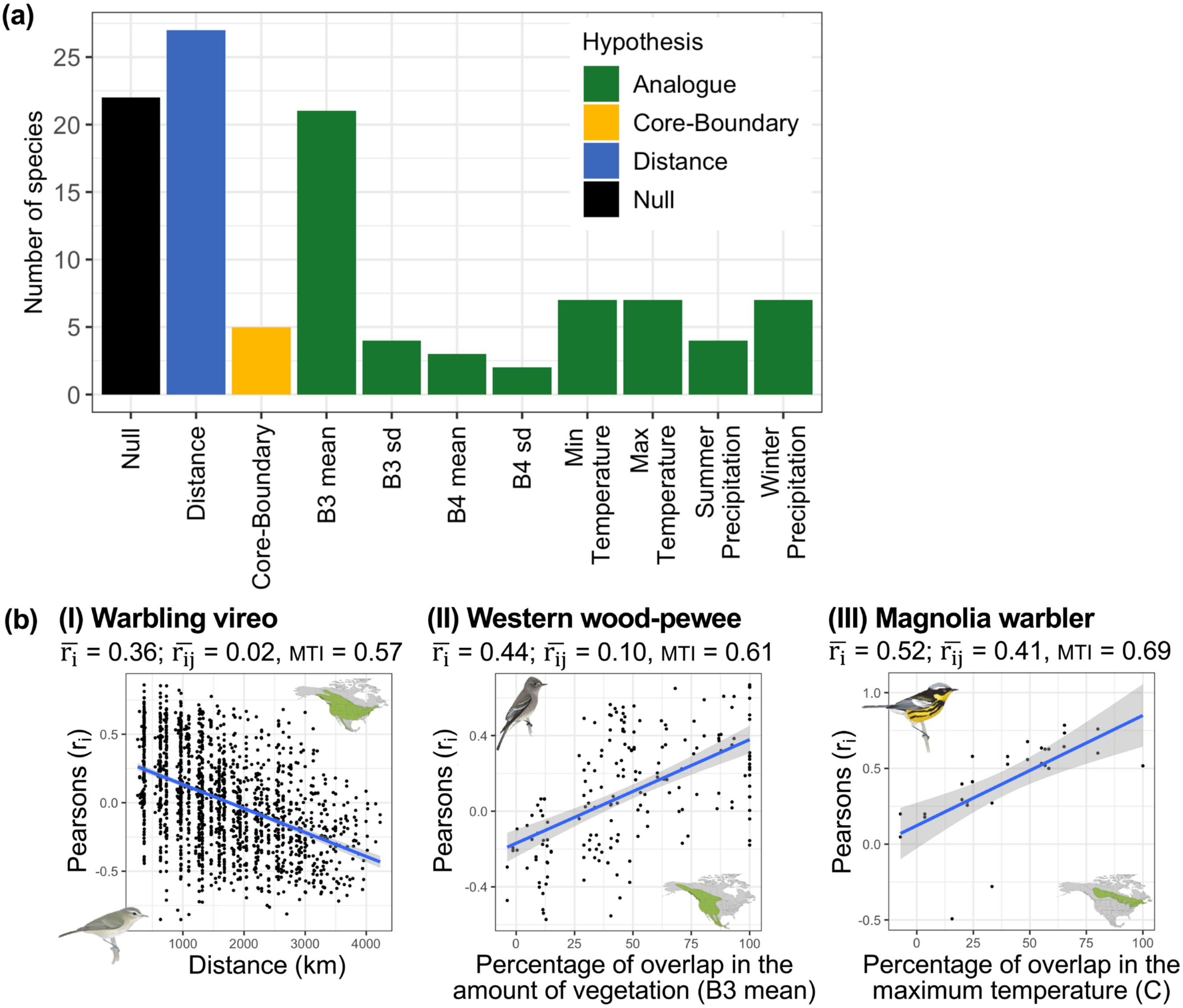

Figure 4. For each species, we compared 11 models in a model selection analysis. The top model(s) of each species is represented by a bar in this graph. The colors represent the hypothesis tested by each model. The blue bar represents the *geographic distance hypothesis*, the yellow bar represents the *core-boundary hypothesis* and the green bars represent each variable tested as part of the *analogue environment hypothesis*. We added a null model (in black) to each species. Out of 96 species, most species (77.1%) had model transferability results that fitted at least one of the three hypotheses. Most species had prediction accuracies correlated with geographical distance between regions and with extrapolation in the amount of vegetation (B3 mean) between regions. (b) Example results showing how the attributes of a distribution influence model transferability among study regions (hexagons). (I) Increases in geographic distances between training and testing regions decreased model transferability for warbling vireo (*Vireo gilvus*) (II). Model transferability was improved with a decrease in extrapolation in the amount of vegetation (B3 mean) for the western wood-pewee (*Contopus sordidulus*) and (III) model transferability was improved with a decrease in extrapolation in maximum summer temperature for the magnolia warbler (*Setophaga magnolia*).

the 96 species used in this analysis, 77.1% (74 species) had prediction data fitting at least one of the three hypotheses tested (Fig. 4a; Supporting information). We found strong support for the *geographic distance hypothesis* for 28.1% of species (Fig. 4a; example in Fig. 4b). As expected, larger distances between regions decreased model transferability. Further, the *analogue environment hypothesis* was supported for 54.2% of species; the greater the amount of extrapolation from the environmental conditions observed in the model training region, the greater the fall-down in prediction success. However, one covariate stood out as being the most impactful; 21.9% of species had model transferability decreased by extrapolation in the amount of vegetation (B3 mean) between regions (Fig. 4a; example in Fig. 4b). Overall, 31.3% of species were affected by extrapolation in at least one land-cover covariate and 26.0% of species by extrapolation in at least one climate covariate. Unexpectedly, the *core-boundary hypothesis* was supported for only 5.2% of species.

Table 2. Results of phylogenetic linear models assessing relationship between model transferability index *MTI* and species traits. Traits in bold had 95% confidence intervals that did not include zero.

Species trait	Estimate	95% CI
(Intercept)	0.675	0.356 to 0.994
Within-region mean Pearson	0.140	0.005 to 0.275
Distribution size (1 000 000 km^2)	**−0.009**	**−0.014 to −0.003**
Topographic variation (km)	**0.065**	**0.002 to 0.127**
Maximum temperature (°C)	−0.0004	−0.006 to 0.0005
Habitat – forest versus non-forest	−0.009	−0.064 to 0.045
Migration status – migrants versus residents	−0.009	−0.109 to 0.092
PCA axis 1 (longevity)	**0.007**	**0.001 to 0.013**

Overall, 8.3% of species were affected by both distance and extrapolation, i.e. these species had two hypotheses reflected among the top models (delta BIC < 2), one representing the *geographic distance hypothesis*, one representing the *analogue environment hypothesis*.

Discussion

Overall, we found that most species exhibited low prediction accuracies when transferring models across their geographic ranges. Between-regions predictions were 43% poorer than models making predictions within-regions – suggesting poor model transferability and possibly a violation of the stationarity assumption. While selection of modeling tool can impact between-region model performance (Heikkinen et al. 2012, Qiao et al. 2019), our results support previous studies indicating that an increase in geographical distances (Qiao et al. 2019) and degree of extrapolation (Torres et al. 2015, Bell and Schlaepfer 2016, Qiao et al. 2019, Charney et al. 2021) between model training and testing areas reduce prediction accuracies. Moreover, our results suggest that extrapolating between regions with different amounts of vegetation is especially problematic for model transferability.

Species' model transferability indices (*MTI*) were correlated with three ecological and life history traits. We found that short-lived species and species with large geographic distributions were more likely to have lower model transferability prediction accuracies. As predicted, we also found that species with distributions in areas with low topographic relief are more likely to have lower prediction accuracies across their SDMs. We recommend caution when transferring models across space for such species. If the objective is prediction to new areas, we recommend that model training areas include as broad a range of environmental and land-cover variables as possible. For instance, in mountain landscapes, environmental conditions in a region can often reflect those encountered by species across their entire geographic ranges. This limits the risks posed by extrapolation. The inclusion of mechanistic variables (e.g. dispersal, factors influencing survival and reproduction) may also decrease the risks posed by extrapolation regardless of the species (Evans et al. 2016, Higgins et al. 2020, Betts et al. 2021).

Similar mechanisms likely drove our findings that 1) species with larger geographical ranges are more likely to have low model transferability, and 2) increases in geographic distance reduce prediction accuracies. However, the negative effect of geographic distance was not consistent across all species. In other words, species with large distributions were not always more affected by long distances. This could be due to the presence subspecies or subpopulations across space, with each population responding slightly differently to environmental conditions. While an increase in distance elevates the likelihood of encountering different populations, in some circumstances (e.g. in the presence of a geographical barrier), populations could be relatively close to each other (Pomara et al. 2014, Kopuchian et al. 2020).

Two mechanisms could influence the degree to which models can be transferred across space. First, the actual biological processes governing a species' distribution could vary across geographic space; such factors influence what we term *ecological stationarity*. For instance, different biotic interactions in the form of, for example, new competitors, could arise in particular parts of the species range, which could alter relationships between observed predictors and species abundances (Jankowski et al. 2010). Similarly, geographic isolation and different climatic conditions could combine to enable subpopulations within species – each with somewhat different environmental niches (Bush 1994, Peterson and Holt 2003, Rollinson et al. 2021). In short, a lack of ecological stationarity could be caused by the presence of unaccounted differences in biotic interactions or processes such as emergence of sub-population structure across a species' range.

However, even if a species exhibits a high degree of ecological stationarity, extrapolating models to new environments can also impinge severely on model prediction success (Fig. 5). Indeed, there are substantial risks of model failure due to over-extrapolation to new environments; this situation is particularly likely to occur under global change. Indeed, several of our results (including support for the *analogue environment hypothesis*) indicate that reduced model transferability is due to over-extrapolation.

Model misspecification is yet another mechanism that likely affects model stationarity and transferability (Miller 2012, Dale and Fortin 2014, Rollinson et al. 2021). For instance, SDMs often use predictor variables collected at broad spatial scales, but organisms often respond to the environment at finer scales (Potter et al. 2013, Frey et al. 2016). Species also likely have changing biotic interactions across their range (Dormann et al. 2018), but such data, while important, are often unavailable across large spatial extents. Moreover, adding additional covariates to a model could help, but it involves costs, both statistically in terms of model complexity, and the economic costs of gathering additional data. Considering limited conservation and management resources, it becomes important to balance the level of accuracy in model prediction with the cost involved in improving the model accuracy (Montanari and Koutsoyiannis 2014).

Our results emphasize the importance of assessing the degree of extrapolation used in model transferability analysis

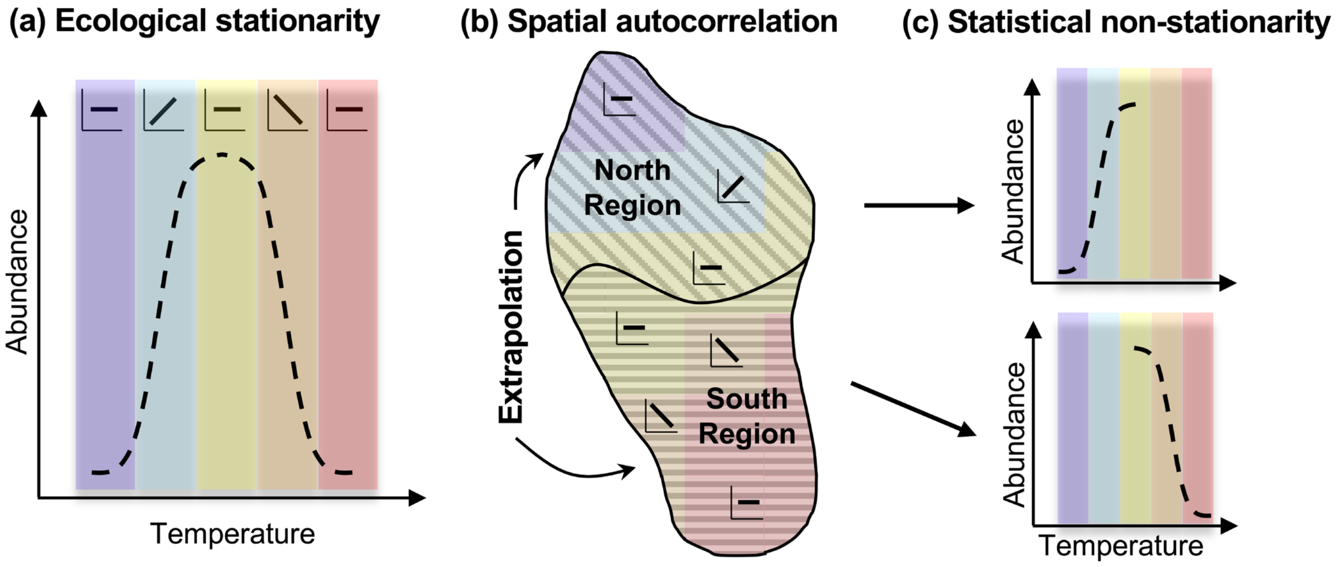

Figure 5. Example of a situation where extrapolation and statistical non-stationarity are related through the concept of spatial autocorrelation. (a) Assume a species has a true unimodal relationship between abundance and temperature, referred as ecological stationarity. (b) If a distribution is characterized by spatial autocorrelation in temperature, with colder temperatures in the northern region and warmer temperatures in the southern region, then predicting the relationship from one region to the other will require extrapolation. Note that spatial autocorrelation is not necessary to have extrapolation. (c) The starting model of the relationship between abundance and temperature will be reversed, yielding statistical nonstationarity.

and documenting the results in publications (Qiao et al. 2019, Guillaumot et al. 2020). We also recommend testing the stationarity assumption in SDMs prior to transferring models to new geographic locations or into the future (Legendre and Fortin 1989, Betts et al. 2006, Dormann et al. 2007, Hawkins 2012, Ghorbani 2013, Sequeira et al. 2018). While it may be difficult to disentangle the effects of extrapolation and non-stationarity, it may be possible to account for each separately. A first option is to limit extrapolation in model specifications by selecting regions with analogue environments (Owens et al. 2013).

Management and conservation applications

Our results highlight the challenges associated with predicting abundance across space for most species of birds breeding in North America. If unaccounted for, poor model transferability could mislead conservation efforts by inaccurately predicting species occurrences and abundances across space and time. The problems associated with extrapolation are well known, and yet many applications of SDMs ignore the consequences. Moreover, only rarely are attempts made to directly test for stationarity in space or time. Once extrapolation is accounted for, ecological non-stationarity (Legendre and Fortin 1989, Ghorbani 2013, Dale and Fortin 2014) may not be a problem in itself but a source of information to be quantified. It can elucidate potential patterns in habitat use and selection across geographic ranges that could imply population structure (Hawkins 2012).

As mentioned by Milly et al. (2008, 2015), 'In a non-stationary world, continuity of observations is critical'. To improve the accuracy of SDMs and to better identify stationary versus non-stationary ecological processes, will require dependence upon an ever-expanding pool of biodiversity data – both in terms of responses (e.g. species' abundance) and covariates (Yates et al. 2018) – which are increasingly quantified at fine resolutions using remote sensing (Goetz et al. 2010, He et al. 2015). The future of conservation and success of management actions must reflect chronic non-stationarity in species responses to their environment (Wiens 2012, Wolkovich et al. 2014). Continual and consistent monitoring of our environment and constituent biodiversity is therefore essential (Yates et al. 2018).

Acknowledgements – We would like to thank Rebecca Hutchinson, Ariel Muldoon, Christopher Wolf and Taal Levi for their guidance and helping us refine our ideas. Thanks also to Rebecca Hutchinson, Laurel Hopkins, Justin Braaten and Robert Kennedy for their help accessing computer nodes and/or their help with extracting data from the cloud.

Funding – The author JSR was supported by a Natural Sciences and Engineering Research Council (NSERC) Postgraduate Scholarship, Oregon State Univ. Richardson Family Graduate Fellowship and a Diversity Advancement Pipeline Fellowship. This research was supported in part by a US National Science Foundation (grant NSF-CDI-0913560, and NSF-LTER8-DEB-2025755) to MGB.

Author contributions

Josée S. Rousseau: Conceptualization (equal); Formal analysis (lead); Funding acquisition (equal); Methodology (equal); Visualization (lead); Writing – original draft (lead);

Writing – review and editing (equal). **Matthew G. Betts**: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Writing – original draft (supporting); Writing – review and editing (equal).

Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.06060>.

Data availability statement

All data and scripts used in the analyses are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.08kprr54c> (Rousseau and Betts 2022). Raw Breeding Bird Survey data are available from: <https://www.pwrc.usgs.gov/BBS/RawData/>. Original, unprocessed Landsat Images and climate data are available from Google Earth Engine (<https://developers.google.com/earth-engine/datasets/catalog/landsat> and <https://developers.google.com/earth-engine/datasets/tags/daymet, respectively).

Supporting information

The supporting information associated with this article is available from the online version.

References

Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – Global Ecol. Biogeogr. 16: 743–753.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. – Ecol. Model. 157: 101–118.

Bahn, V. and McGill, B. J. 2013. Testing the predictive performance of distribution models. – Oikos 122: 321–331.

Bell, D. M. and Schlaepfer, D. R. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. – Ecol. Model. 330: 50–59.

Bennett, P. M. and Owens, I. P. F. 2002. Evolutionary ecology of birds: life histories, mating systems and extinction. – Oxford Univ. Press.

Betts, M. G. et al. 2006. The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. – Ecol. Model. 191: 197–224.

Betts, M. G. et al. 2019. Synergistic effects of climate and land-cover change on long-term bird population trends of the western USA: a test of modeled predictions. – Front. Ecol. Evol. 7: 186.

Betts, M. G. et al. 2021. When are hypotheses useful in ecology and evolution? – Ecol. Evol. 11: 5762–5776.

Billerman, S. M. et al. 2020. Birds of the World. – Cornell Laboratory of Ornithology, Ithaca, NY, USA.

BirdLife International and Handbook of the Birds of the World 2018. Bird species distribution maps of the world, Ver. 2018.1. – <http://datazone.birdlife.org/species/requestdis>.

Boulesteix, A.-L. et al. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics: random forests in bioinformatics. – WIREs Data Mining Knowl. Discov. 2: 493–507.

Breiman, L. 2001. Random forests. – Mach. Learn. 45: 5–32.

Burnham, K. P. and Anderson, D. R. 2003. Model selection and multimodel inference: a practical information–theoretic approach. – Springer.

Bush, G. L. 1994. Sympatric speciation in animals: new wine in old bottles. – Trends Ecol. Evol. 9: 285–288.

Chamberlain, S. A. et al. 2014. How context dependent are species interactions? – Ecol. Lett. 17: 881–890.

Charney, N. D. et al. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. – Front. Ecol. Evol. 9: 689295.

Cohen, J. 1992. A power primer. – Psychol. Bull. 112: 155.

Cohen, W. B. and Goward, S. N. 2004. Landsat's role in ecological applications of remote sensing. – BioScience 54: 535–545.

Colles, A. et al. 2009. Are specialists at risk under environmental change? Neoecological, paleoecological and phylogenetic approaches. – Ecol. Lett. 12: 849–863.

Dale, M. R. T. and Fortin, M.-J. 2014. Spatial analysis: a guide for ecologists. – Cambridge Univ. Press.

Daskin, J. H. and Alford, R. A. 2012. Context-dependent symbioses and their potential roles in wildlife diseases. – Proc. Biol. Sci. 279: 1457–1465.

Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.

Dormann, C. F. et al. 2018. Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. – Global Ecol. Biogeogr. 27: 1004–1016.

Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.

Ericson, P. G. P. et al. 2006. Higher-level phylogeny and morphological evolution of tyrant flycatchers, cotingas, manakins and their allies (Aves: Tyrannida). – Mol. Phylogenet. Evol. 40: 471–483.

Evans, J. S. et al. 2011. Modeling species distribution and change using random forest. – In: Drew, C. A. et al. (eds), Predictive species and habitat modeling in landscape ecology. Springer, pp. 139–159.

Evans, M. E. K. et al. 2016. Towards process-based range modeling of many species. – Trends Ecol. Evol. 31: 860–871.

Fink, D. et al. 2010. Spatiotemporal exploratory models for broad-scale survey data. – Ecol. Appl. 20: 2131–2147.

Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. – Cambridge Univ. Press.

Frey, S. J. K. et al. 2016. Microclimate predicts within-season distribution dynamics of montane forest birds. – Divers. Distrib. 22: 944–959.

Ghorbani, M. 2013. Testing the weak stationarity of a spatio-temporal point process. – Stoch. Environ. Res. Risk Assess. 27: 517–524.

Goetz, S. J. et al. 2010. Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. – Ecology 91: 1569–1576.

Gómez, C. et al. 2016. Niche-tracking migrants and niche-switching residents: evolution of climatic niches in New World warblers (Parulidae). – Proc. R. Soc. B 283: 20152458.

Gorelick, N. et al. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. – Remote Sens. Environ. 202: 18–27.

Guillaumot, C. et al. 2020. Extrapolation in species distribution modelling. Application to Southern Ocean marine species. – Prog. Oceanogr. 188: 102438.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.

Guisan, A. et al. 2013. Predicting species distributions for conservation decisions. – Ecol. Lett. 16: 1424–1435.

Gutiérrez-Illán, J. et al. 2014. Precipitation and winter temperature predict long-term range-scale abundance changes in western north American birds. – Global Change Biol. 20: 3351–3364.

Hackett, S. J. et al. 2008. A phylogenomic study of birds reveals their evolutionary history. – Science 320: 1763–1768.

Hawkins, B. A. 2012. Eight (and a half) deadly sins of spatial analysis. – J. Biogeogr. 39: 1–9.

He, K. S. et al. 2015. Will remote sensing shape the next generation of species distribution models? – Remote Sens. Ecol. Conserv. 1: 4–18.

Heikkinen, R. K. et al. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? – Ecography 35: 276–288.

Hernandez, P. A. et al. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. – Ecography 29: 773–785.

Higgins, S. I. et al. 2020. Predictive ability of a process-based versus a correlative species distribution model. – Ecol. Evol. 10: 11043–11054.

Ho, L. si T. and Ané, C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. – Syst. Biol. 63: 397–408.

Howard, C. et al. 2014. Improving species distribution models: the value of data on abundance. – Methods Ecol. Evol. 5: 506–513.

Howard, C. et al. 2015. The drivers of avian abundance: patterns in the relative importance of climate and land use. – Global Ecol. Biogeogr. 24: 1249–1260.

Hutchinson, G. E. 1957. Concluding remarks. – Cold Spring Harbor Symp. Quant. Biol. XXII: 415–427.

Jankowski, J. E. et al. 2010. Squeezed at the top: interspecific aggression may constrain elevational ranges in tropical birds. – Ecology 91: 1877–1884.

Jetz, W. et al. 2012. The global diversity of birds in space and time. – Nature 491: 444–448.

Jiguet, F. et al. 2010. Population trends of European common birds are predicted by characteristics of their climatic niche. – Global Change Biol. 16: 497–505.

Kennedy, R. E. et al. 2018. Implementation of the LandTrendr algorithm on Google Earth Engine. – Remote Sens. 10: 691.

Kopuchian, C. et al. 2020. A test of the riverine barrier hypothesis in the largest subtropical river basin in the Neotropics. – Mol. Ecol. 29: 2137–2149.

Laube, I. et al. 2015. Niche availability in space and time: migration in Sylvia warblers. – J. Biogeogr. 42: 1896–1906.

Legendre, P. and Fortin, M. J. 1989. Spatial pattern and ecological analysis. – Vegetatio 80: 107–138.

Liaw, A. and Wiener, M. 2002. Classification and regression by randomForest. – R News 2: 6.

Luan, J. et al. 2020. The predictive performances of random forest models with limited sample size and different species traits. – Fish. Res. 227: 105534.

Mainali, K. P. et al. 2015. Projecting future expansion of invasive species: comparing and improving methodologies for species distribution modeling. – Global Change Biol. 21: 4464–4480.

Matthews, S. N. et al. 2011. Changes in potential habitat of 147 North American breeding bird species in response to redistribution of trees and climate following predicted climate change. – Ecography 34: 933–945.

McPherson, J. M. et al. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? – J. Appl. Ecol. 41: 811–823.

Millar, C. I. et al. 2007. Climate change and forests of the future: managing in the face of uncertainty. – Ecol. Appl. 17: 2145–2151.

Miller, J. A. 2012. Species distribution models: spatial autocorrelation and non-stationarity. – Prog. Phys. Geogr. 36: 681–692.

Milly, P. C. D. et al. 2008. Stationarity is dead: whither water management? – Science 319: 573–574.

Milly, P. C. D. et al. 2015. On critiques of 'stationarity is dead: whither water management?' – Water Resour. Res. 51: 7785–7789.

Minias, P. et al. 2017. Longevity is associated with relative brain size in birds. – Ecol. Evol. 7: 3558–3566.

Møller, A. P. and Jennions, M. D. 2002. How much variance can be explained by ecologists and evolutionary biologists? – Oecologia 132: 492–500.

Montanari, A. and Koutsoyiannis, D. 2014. Modeling and mitigating natural hazards: stationarity is immortal! – Water Resour. Res. 50: 9748–9756.

Myhrvold, N. P. et al. 2015. An amniote life-history database to perform comparative analyses with birds, mammals and reptiles. – Ecology 96: 3109–3109.

Orme, C. D. L. et al. 2019. Distance to range edge determines sensitivity to deforestation. – Nat. Ecol. Evol. 3: 886–891.

Owens, H. L. et al. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. – Ecol. Model. 263: 10–18.

Paradis, E. and Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. – Bioinformatics 35: 526–528.

Pardieck, K. L. et al. 2019. North American Breeding Bird Survey dataset 1966–2018, ver. 2018.0. – U.S. Geological Survey, Patuxent Wildlife Research Center.

Partners in Flight 2020. Partners in flight avian conservation assessment database, ver. 2020. – <https://pif.birdconservancy.org/>.

Pearman, P. B. et al. 2008. Niche dynamics in space and time. – Trends Ecol. Evol. 23: 149–158.

Peterson, A. T. and Holt, R. D. 2003. Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. – Ecol. Lett. 6: 774–782.

Phillimore, A. B. et al. 2007. Biogeographical basis of recent phenotypic divergence among birds: a global study of subspecies richness. – Evolution 61: 942–957.

Pomara, L. Y. et al. 2014. Avian species composition across the Amazon River: the roles of dispersal limitation and environmental heterogeneity. – J. Biogeogr. 41: 784–796.

Potter, K. A. et al. 2013. Microclimatic challenges in global change biology. – Global Change Biol. 19: 2932–2939.

Pulido, F. and Widmer, M. 2005. Are long-distance migrants constrained in their evolutionary response to environmental change? Causes of variation in the timing of autumn migration in a blackcap (*S. atricapilla*) and two garden warbler (*Sylvia borin*) populations. – Ann. N. Y. Acad. Sci. 1046: 228–241.

Qiao, H. et al. 2019. An evaluation of transferability of ecological niche models. – Ecography 42: 521–534.

Rich, J. L. and Currie, D. J. 2018. Are North American bird species' geographic ranges mainly determined by climate? – Global Ecol. Biogeogr. 27: 1–13.

Roberts, D. R. et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical or phylogenetic structure. – Ecography 40: 913–929.

Rollinson, C. R. et al. 2021. Working across space and time: nonstationarity in ecological research and application. – Front. Ecol. Environ. 19: 66–72.

Rousseau, J. S. and Betts, M. G. 2022. Data from: Factors influencing transferability in species distribution models. – Dryad Digital Repository, <https://doi.org/10.5061/dryad.08kprr54c>.

Sauer, J. R. et al. 2003. Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. – J. Wildl. Manage. 67: 372–389.

Schmidt, A. E. et al. 2014. Non-stationary seabird responses reveal shifting ENSO dynamics in the northeast Pacific. – Mar. Ecol. Prog. Ser. 499: 249–258.

Sequeira, A. M. M. et al. 2018. Transferring biodiversity models for conservation: opportunities and challenges. – Methods Ecol. Evol. 9: 1250–1264.

Sexton, J. P. et al. 2009. Evolution and ecology of species range limits. – Annu. Rev. Ecol. Evol. Syst. 40: 415–436.

Shirk, A. J. et al. 2014. Spatiotemporal variation in resource selection: insights from the American marten (*Martes americana*). – Ecol. Appl. 24: 1434–1444.

Shirley, S. M. et al. 2013. Species distribution modelling for the people: unclassified landsat TM imagery predicts bird occurrence at fine resolutions. – Divers. Distrib. 19: 855–866.

Syphard, A. D. and Franklin, J. 2009. Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. – Ecography 32: 907–918.

Thomas, J. W. et al. 2006. The Northwest Forest Plan: origins, components, implementation experience and suggestions for change. – Conserv. Biol. 20: 277–287.

Thornton, P. E. et al. 2017. Daymet: daily surface weather data on a 1-km grid for North America, ver. 3. – ORNL DAAC, Oak Ridge, Tennessee, USA.

Tobler, W. R. 1970. A computer movie simulating urban growth in the detroit region. – Econ. Geogr. 46: 234–240.

Torres, L. G. et al. 2015. Poor transferability of species distribution models for a pelagic predator, the grey petrel, indicates contrasting habitat preferences across ocean basins. – PLoS One 10: e0120014.

Tuanmu, M.-N. et al. 2011. Temporal transferability of wildlife habitat models: implications for habitat monitoring. – J. Biogeogr. 38: 1510–1523.

Vedder, O. et al. 2013. Quantitative assessment of the importance of phenotypic plasticity in adaptation to climate change in wild bird populations. – PLoS Biol. 11: e1001605.

Vergnon, R. et al. 2017. Complex relationships between competing guilds along large-scale environmental gradients. – Am. Nat. 189: 407–421.

Vrieze, S. I. 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). – Psychol. Methods 17: 228.

Wan, H. Y. et al. 2017. Meta-replication reveals nonstationarity in multi-scale habitat selection of Mexican spotted owl. – Condor 119: 641–658.

Whittingham, M. J. et al. 2007. Should conservation strategies consider spatial generality? Farmland birds show regional not national patterns of habitat association. – Ecol. Lett. 10: 25–35.

Wiens, J. A. et al. (ed.) 2012. Historical environmental variation in conservation and natural resource management. – Wiley.

Wisz, M. S. et al. 2008. Effects of sample size on the performance of species distribution models. – Divers. Distrib. 14: 763–773.

Wolkovich, E. M. et al. 2014. Temporal ecology in the Anthropocene. – Ecol. Lett. 17: 1365–1379.

Yates, K. L. et al. 2018. Outstanding challenges in the transferability of ecological models. – Trends Ecol. Evol. 33: 790–802.

Zuckerberg, B. et al. 2016. Novel seasonal land cover associations for eastern North American forest birds identified through dynamic species distribution modelling. – Divers. Distrib. 22: 717–730.