

# Using temporal occupancy to predict avian species distributions

Sara Snell Taylor<sup>1</sup>  | Grace Di Cecco<sup>1</sup>  | Allen H. Hurlbert<sup>1,2</sup> 

<sup>1</sup>Department of Biology, University of North Carolina, Chapel Hill, North Carolina, USA

<sup>2</sup>Environment, Ecology, and Energy Program, University of North Carolina, Chapel Hill, North Carolina, USA

## Correspondence

Sara Snell Taylor, Department of Biology, CB 3280, University of North Carolina, Chapel Hill, NC 27599, USA.  
Email: ssnell6@gmail.com

## Funding information

National Science Foundation, Grant/Award Number: DEB-1354563

Editor: Xuan Liu

## Abstract

**Aim:** Species distribution models (SDMs) are ubiquitous in ecology to predict species occurrence throughout their range. Typically, SDMs are created using presence-only or presence-absence data. We hypothesize that the continuous metric of temporal occupancy, the proportion of time a species is observed at a given site, provides more detail about species occurrence than binary presence-based SDMs.

**Location:** North America.

**Methods:** We compared SDMs for 189 focal species using four modelling methods to determine whether North American avian species distributions are better predicted using temporal occupancy over presence-absence. We used the North American Breeding Bird Survey and built SDMs based on all sites sampled consecutively between 2001 and 2015, as well as on a subset of only five time points within the 15-year sampling window. Each model used the same environmental inputs to predict species range. Each SDM was cross-validated temporally and spatially.

**Results:** Species distributions were generally better predicted using temporal occupancy rather than presence-absence when using either a five-year or fifteen-year sampling window. Species that occurred in a smaller proportion of their predicted range were particularly better predicted with SDMs using temporal occupancy. Temporal occupancy SDMs had lower false discovery and false-positive rates but higher false-negative rates than presence-absence models.

**Main conclusions:** Temporal occupancy is a valuable metric that can improve predictions of species occurrence for birds and may improve conservation planning and design efforts.

## KEYWORDS

species distribution models, species range, temporal occupancy

## 1 | INTRODUCTION

One of the primary methods for classifying a species range is the species distribution model (SDM), which predicts where species are most likely to occur based on a set of environmental factors (Freeman &

Mason, 2015; Guisan & Thuiller, 2005). These model outputs have important implications for ecological research and conservation management (Elith & Leathwick, 2009; Engler et al., 2017; Franklin, 2010; Guisan et al., 2013), but depend critically on methods that produce accurate model predictions by minimizing false positives and false

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

negatives. Most SDM modelling methods use species presence and absence as input to generate a probability surface of predicted occurrences (Elith & Leathwick, 2009; Gomes et al., 2018; Guillera-Arroita et al., 2015; Liu et al., 2011). However, the binary nature of presence-absence data provides less information than continuous measures that can reflect gradations of habitat suitability. Some SDM approaches use continuous information on species abundance, which may account for how many individuals of the species occur at a given site and allow the model to account for more specific site preferences (Bahn & McGill, 2007; Gomes et al., 2018). Another continuous measure with the potential to better characterize site suitability is temporal occupancy, or the proportion of years a species is observed at a given site. While temporal occupancy and abundance are usually positively correlated, species commonly exist at sites with high temporal occupancy but at low abundance (Coyle et al., 2013), meaning these two measures may contain unique information. Specifically, temporal occupancy can help distinguish sites that are regularly versus intermittently occupied over time, and thus may reflect longer-term population viability at a site (Magurran & Henderson, 2003; Ulrich & Ollik, 2004; Coyle et al., 2013; Supp et al., 2015; Umaña et al., 2017; Snell Taylor et al., 2018; Jenkins et al., 2018).

This potential difference between geographic patterns in temporal occupancy and abundance has consequences for questions considering temporal turnover and richness-environment correlations (Belmaker, 2009; Coyle et al., 2013; Snell Taylor, et al., 2018; Umaña et al., 2017). For example, species with low temporal occupancy exhibit higher temporal turnover, shallower species-area relationships and richness patterns that are less correlated with the local environment compared to species with high temporal occupancy (Snell Taylor, et al., 2018). Additionally, environmental predictors explained more variation in temporal occupancy than variation in abundance on average across nearly 200 bird species (Snell Taylor, Umbanhowar, et al., 2020). This suggests that temporal occupancy may indeed capture important information about habitat suitability, and yet, modelling spatial variation in temporal occupancy remains understudied in ecology (but see Green et al., 2019; Rushing et al., 2019; Zuckerberg et al., 2011). We seek to determine whether using temporal occupancy data can provide more reliable SDMs than using presence-absence data for a large suite of avian species in North America.

We also want to ensure that any observed differences in performance between SDMs created with temporal occupancy versus presence-absence data are independent of the modelling approach used. SDMs have been implemented using a wide range of approaches, and the optimal method may depend on the available data (Araújo et al., 2014; Elith et al., 2006; Guillera-Arroita et al., 2015; Merow et al., 2014; Phillips et al., 2006). The simplest SDMs (e.g. MaxEnt; Phillips et al., 2006) rely on species presence-only and environmental data as the inputs. In contrast, some SDM approaches also take into account absences where species are sampled but not observed (Elith & Leathwick, 2009; Guillera-Arroita et al., 2015). Traditional occupancy models use more detailed information about species occurrence, generally spatial or temporal occurrence, which can be used to infer a more detailed gradient of suitability (Comte &

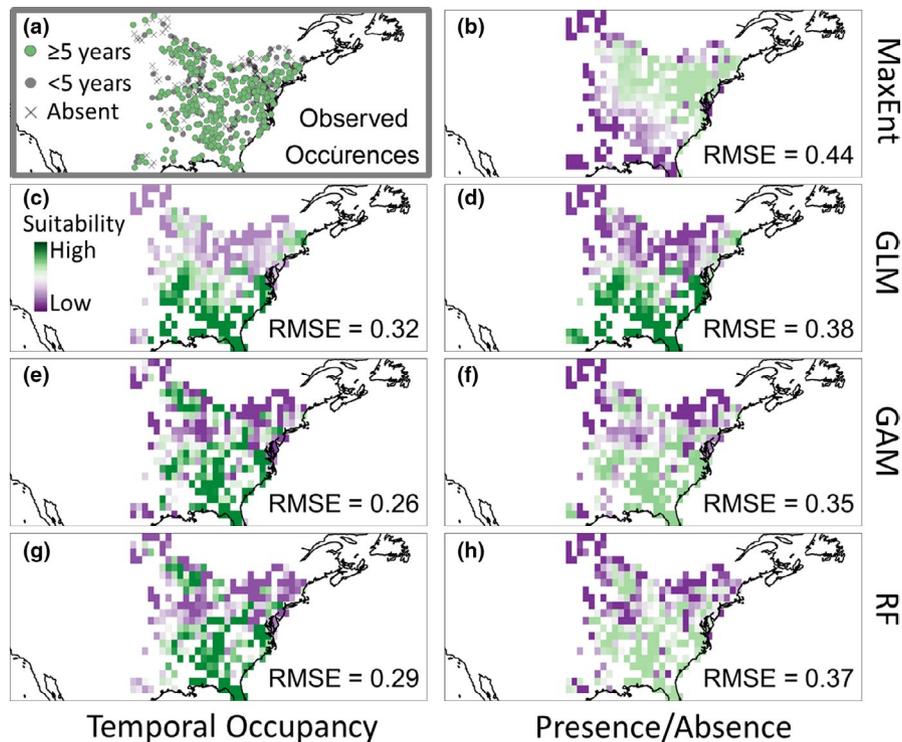
Grenouillet, 2013; Gomes et al., 2018; Guillera-Arroita et al., 2015). In this study, we evaluate the utility of using temporal occupancy data for building SDMs relative to traditionally used presence-absence or presence-background data, comparing across four common modelling approaches (generalized linear models, generalized additive models, random forest and MaxEnt). We assess the relative performance of presence-absence and temporal occupancy data using both spatial cross-validation and temporal cross-validation.

## 2 | METHODS

We used the North American Breeding Bird Survey (BBS) to characterize geographic variation in species presence-absence and temporal occupancy. BBS surveys are distributed across the continent and monitor breeding birds via a series of fifty 3-min point counts spaced at 0.8-km intervals along a roadside route (USGS Patuxent Wildlife Research Center 2012). Each survey is conducted during the breeding season (typically June) by a single observer who records all avian species seen or heard along the route. We used the 1,004 BBS survey routes that were continuously surveyed from 2001 to 2015 and excluded species that were poorly sampled by the survey design such as waterbirds, raptors and nocturnal species (Figure S1; Butcher, 1987). In total, we identified 189 focal land bird species that were present on at least 60 BBS routes over 15 sampling years. These thresholds were chosen to ensure that a model could be fit for each focal species.

We constructed SDMs based on seven environmental variables summarized at the route level: temperature seasonality (intra-annual standard deviation \*100), mean maximum temperature of the warmest month, mean minimum temperature of the coolest month, mean precipitation during the wettest month, mean precipitation during the driest month, mean elevation from the 30-m resolution National Elevation Dataset ([ned.usgs.gov](http://ned.usgs.gov)) and normalized difference vegetation index (NDVI) from the Global Inventory Mapping and Monitoring Studies database (<https://glam1.gsfc.nasa.gov/>). Because each BBS route is 40 km long and because only the starting coordinates are known for many of the routes, we calculated the means of each environmental variable within a 40-km buffer around each BBS route's starting coordinates. The five climate variables included in the models were acquired from Bioclim at 2.5-min resolution and have been commonly used in previous distribution modelling studies (Hijmans et al., 2005; Thuiller et al., 2014), while elevation and NDVI are also known to play a role in determining bird distributions (Bahn & McGill, 2007; Coyle et al., 2013; Hurlbert & Haskell, 2003; Rich & Currie, 2018; Snell Taylor, Umbanhowar, et al., 2020). The NDVI data were 250 m, monthly resolution, averaged over the summer months of June, July and August (<https://gimms.gsfc.nasa.gov/MODIS/>).

We selected four commonly used methods for building SDMs: generalized linear models (GLMs), generalized additive models (GAMs), random forest (RF) and MaxEnt (Elith & Leathwick, 2009; Merow et al., 2014). We used R packages `glm2` (Marschner, 2019; GLM), `MGCV` (Wood, 2017; GAM), `RANDOMFOREST` (Liaw and Wiener,



**FIGURE 1** Map depicting the observed distribution and seven species distribution models (SDMs) of the Yellow-throated Vireo (*Vireo flavifrons*). (a) Presence–absence points for the Yellow-throated Vireo over the 15-year sampling period. Green dots represent sites where the species was present five or more sampling years (temporal occupancy of  $\geq 33\%$ ), grey dots indicate sites where the species was present less than five sampling years, and Xs indicate no occurrences observed during the sampling period. SDMs in the left column (c, e, g) are based on temporal occupancy, while SDMs in the right column are based on (b) presence-only or (d, f, h) presence–absence data. Modelling approaches include (b) MaxEnt, (c, d) generalized linear models (GLMs), (e, f) generalized additive models (GAMs) and (g, h) random forest (RF). Purple indicates high suitability or relative probability of presence and is scaled independently for each plot based on quantile bins

2002; RF) and DISMO (Hijmans et al., 2011; MAXENT) for each of the modelling approaches, respectively. For GAM, GLM and RF, we created SDMs using each method for each focal species based on presence–absence. We included all presences for each species during the 15-year period based on BBS data, meaning that any detection of a species during this period was recorded as a presence. We included as absences all survey routes within each species' breeding range map polygon from BirdLife International ([www.birdlife.org](http://www.birdlife.org)) on which the species was not detected during the 15-year period.

GLM, GAM and RF models were additive and used a binomial logit link function. We generated GAMs with cubic-smooth splines, and the smoothing parameter was chosen by generalized cross-validation (Godsoe et al., 2017; Oliver et al., 2012). We used regression random forests with the following default settings: 500 trees, the number of variables tried at each split was the number of predictor variables divided by 3, sampling was conducted with replacement, node size was 5, and sampling class was not stratified (Steen et al., 2019). All model fits were assessed using root mean square error (RMSE). We also calculated the range occupancy of each species by dividing the number of routes on which a species was present by the total number of routes within their breeding range polygon, which has been used to characterize the

extent to which a range is fully versus patchily occupied (Hurlbert & White, 2007).

For MAXENT, we created SDMs for each focal species using presence points only. Although MAXENT cannot take the continuous temporal occupancy data as input, we include it in this study as a benchmark for comparison to the other methods because it is so widely used (Melo-Merino et al., 2020; Merow et al., 2013). We ran MAXENT using the default settings, which was shown to outperform a range of alternative settings for avian species using the BBS dataset specifically (Ralston et al., 2016), as well as in other applications (Elith et al., 2010; Qiao et al., 2015). Just as in the GAM, GLM and RF model, environmental data were characterized over a 40-km buffer.

Next, we created SDMs using temporal occupancy for the three modelling approaches that accept continuous variables as input: GLM, GAM and RF. We calculated temporal occupancy using a 15-year window from 2001 to 2015, which provides greater resolution to temporal occupancy estimates and inclusion of more focal species, and a five-year window, which provides a more realistic representation of typical ecological sampling efforts. We used sampling years 2001, 2004, 2007, 2010 and 2013 for the five-year temporal occupancy calculation in order to sample evenly from our 15-year time period. Our intention for the 5-year analysis was to maintain the temporal extent of the analysis (15 years) and evaluate the efficacy

of using ecological datasets with more limited temporal sampling, and results were qualitatively similar using five consecutive years (not shown). The 15-year analysis included 189 focal species, but for the five-year analysis 16 species no longer met presence and route occurrence thresholds with the restricted time window and were excluded. Thus, for each species we used presence or temporal occurrence data (Figure 1a) to generate seven SDMs based on the different methodological approaches (Figure 1b–h). We calculated root mean square error (RMSE) for each model using the hydroGOF package (Zambrano-Bigiarini, 2017), with values closer to zero denoting a better fit to the observed values. This method is commonly used to describe the predictive performance of SDMs (Bell & Schlapfer, 2016; Liu et al., 2011; Norberg et al., 2019; Tobler et al., 2019).

Finally, we conducted both temporal cross-validation and spatial cross-validation of the SDMs. First, we examined how well our models were able to predict species presence in 2016 using a model built from occurrence data from 2001 to 2015. The output of each model was a set of continuous probabilities of occurrence. In order to convert the continuous output of the predicted values to binary presence–absence, we first rescaled the predicted values for a given SDM by the maximum. All sites with a rescaled probability greater than 0.5 were predicted to be presences for 2016. We chose the threshold of 0.5 to increase the true-positive rate, with the cost of a decreased true-negative rate, and we verified this trade-off by running the cross-validations using alternative thresholds of 0.25 and 0.75 (Figures S2 and S3). For each species and modelling approach, we created a confusion matrix describing the number of correct and incorrect classifications of presence–absence by model. Then, we calculated the summary metrics of false-positive and false-negative rates and false discovery and false omission rates to assess how robust each model was at correctly identifying presences and absences (Table 1). The validations included 173 species since 16 species had only predicted presences, meaning we could not generate false omission rates for those species because the denominator, predicted absences, would be zero.

For spatial cross-validation, we split the data into ten randomly generated bins, and for each bin, we evaluated how well the SDM based on the other nine bins could predict occurrences in the bin

that was withheld. In order to convert the continuous output of the predicted values to binary presence–absence, we again rescaled predicted values by the maximum for each species and then used a threshold of 0.5 to generate the predicted presences for each bin and created confusion matrices and summary metrics for each species.

### 3 | RESULTS

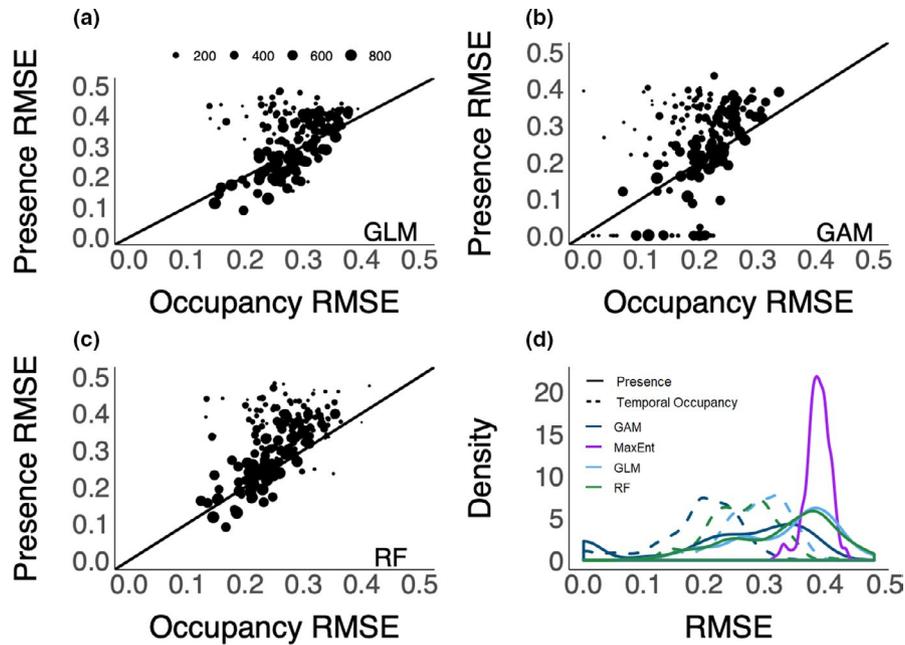
For most species, root mean square error (RMSE) for SDMs based on temporal occupancy was lower than RMSEs for SDMs based on presence–absence regardless of modelling approach (Figure 2a–c). This difference was greatest for RF SDMs, for which 85% (161) of species performed better with temporal occupancy than presence–absence ( $t = -9.33$ ,  $df = 324$ ,  $p$ -value  $< 2.2e-16$ ). Temporal occupancy-based models performed better than presence–absence for 72%–73% (137–138) of species for both GLM and GAM (GLM:  $t = -6.45$ ,  $df = 315$ ,  $p$ -value =  $4.17e-10$ , GAM:  $t = -4.81$ ,  $df = 291$ ,  $p$ -value =  $2.49e-06$ ). For fewer than 20 species, the GAM presence–absence models had a near perfect fit with the predictor variables, suggesting model overfitting. Because these species could be properly fit with the other modelling methods or using temporal occupancy instead of presence, we retained them in the analysis. RF SDMs generally had the lowest RMSE values and MaxEnt had the highest RMSE values across species (Figure 2d). All model and cross-validation outputs are reported in Table S1.

RMSE values were slightly lower for models using temporal occupancy from five rather than 15 years of data. For 5-year models, 91% (157) of species performed better with temporal occupancy than presence–absence with the RF SDMs (Figure 3a), and 81%–82% (140–142) of species performed better using temporal occupancy than presence–absence for GLM and GAM (Figure 3b,c). Again, MaxEnt models had the highest RMSE values of all modelling approaches examined, especially compared with models using temporal occupancy (Figure 3d). Overall, the extent to which temporal occupancy-based models performed better than presence-based models was similar whether using 5-year or 15-year windows.

Metric	Definition	Best performing variable
False discovery rate	Proportion of predicted presences that were incorrectly identified	Temporal occupancy
False omission rate	Proportion of predicted absences that were incorrectly identified	Presence
False-negative rate (1 – sensitivity)	Proportion of actual absences that were incorrectly predicted	Presence
False-positive rate (1 – specificity)	Proportion of actual presences that were incorrectly predicted	Temporal occupancy

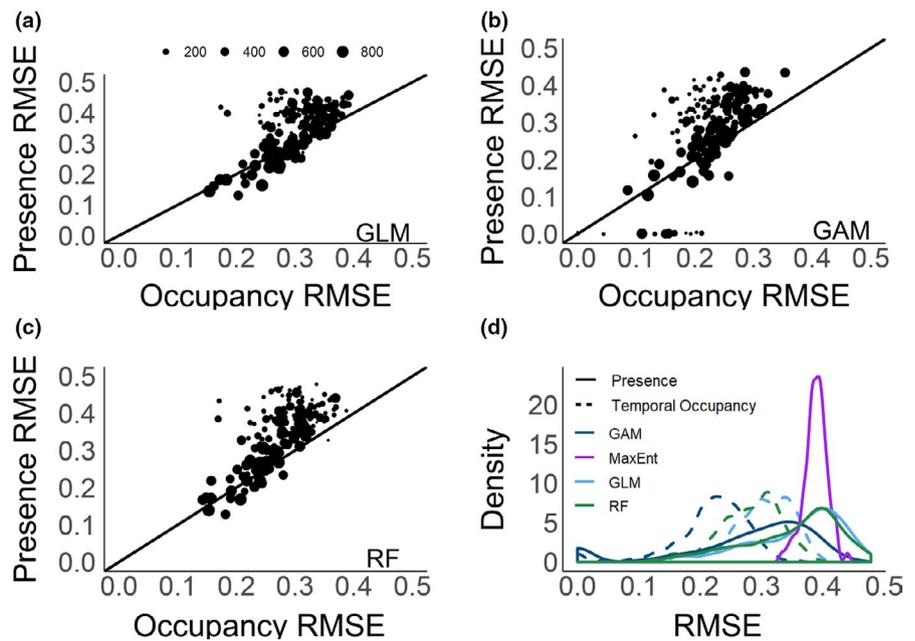
**TABLE 1** Definitions of error rates and the best performing input variable for each metric

Note: The term presence indicates presence–absence for GLM/GAM/RF, and presence-only for MaxEnt.



**FIGURE 2** Comparisons of root mean square error (RMSE) for species distribution models using temporal occupancy based on 15 years of data versus presence over the 15-year period. Points represent each species and are sized by the number of sites where the species occurs. The solid black line is the 1:1 line. (a) Comparison of RMSE for temporal occupancy versus presence–absence models using generalized linear models (GLMs). (b) Comparison of RMSE for temporal occupancy versus presence–absence models using generalized additive models (GAM). (c) Comparison of RMSE for temporal occupancy versus presence–absence models using random forest models (RFs). (d) Density plots of RMSE for each method using temporal occupancy (solid line) or presence (dashed line)

**FIGURE 3** Comparisons of root mean square error (RMSE) for species distribution models using temporal occupancy based on 5 years of data versus presence. Panels as in Figure 2

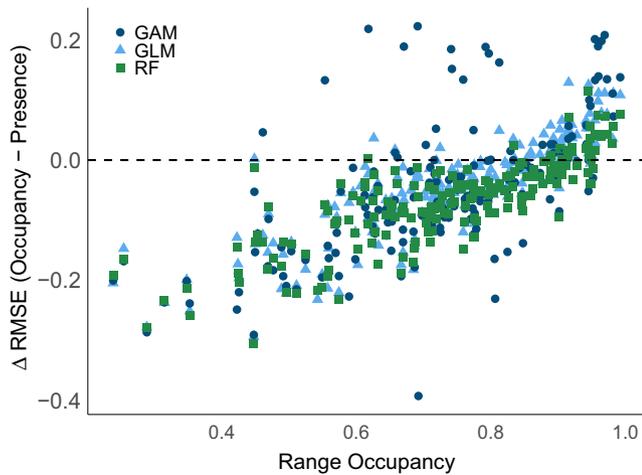


Next, in an effort to understand why some species are better predicted (have lower RMSE) using temporal occupancy and other species are better predicted by presence–absence, we examined the relationship between range occupancy (the proportion of routes within the breeding range of a species on which it was actually observed) and the difference in RMSE between 15-year temporal occupancy-based and presence-based models.

We found that the species for which temporal occupancy-based models had much lower RMSE (i.e.  $\Delta$ RMSE was most negative) generally had lower range occupancy independent of modelling approach (Figure 4).

For the 15-year temporal cross-validation, SDMs based on temporal occupancy had lower false discovery rates ( $t = -6.7, p = .02$ ) and false-positive rates ( $t = -41.6, p = .0006$ ), and higher false omission

rates ( $t = 14.7, p = .005$ ) and false-negative rates ( $t = 22.0, p = .002$ ) compared with models based on presence-absence (Figure 5a). While presence-based models had very low (2%–8%) false-negative and false omission rates, false-positive rates exceeded 60% on average. In contrast, temporal occupancy-based models had comparable false-negative (23%–27%) and false-positive (27%–37%) rates across modelling approaches. False discovery rates were high for both temporal occupancy-based (34%–43%) and presence-based SDMs (51%–54%).

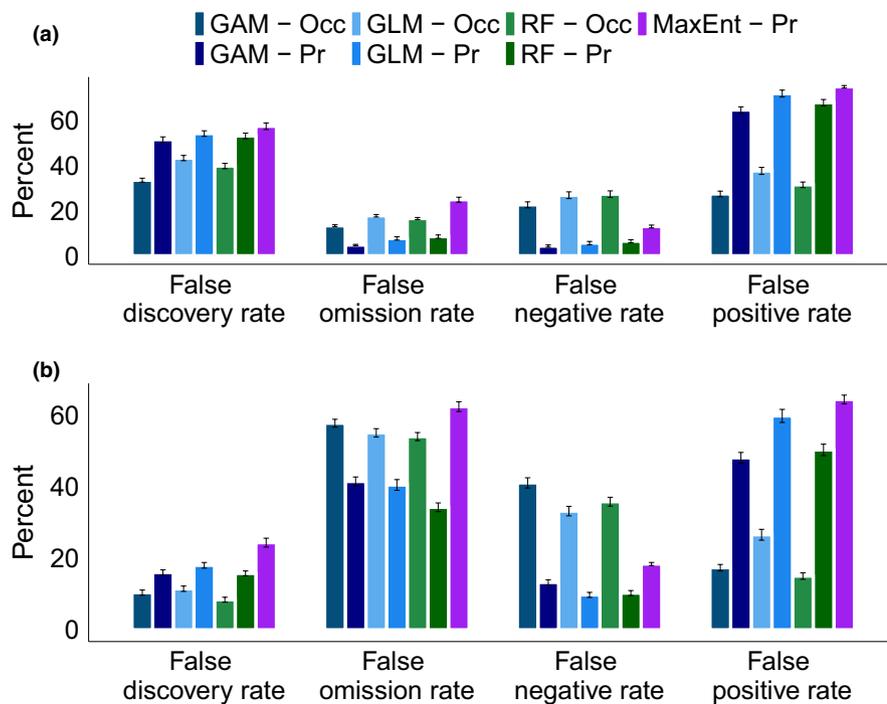


**FIGURE 4** The difference between RMSE for 15-year temporal occupancy models and RMSE for 15-year presence-absence models as a function of range occupancy (number presences/number presences and absences across the range) for each species. Points are coloured by model method, and values below zero represent species that had lower RMSE for temporal occupancy-based models than presence-absence models

Results of 15-year spatial validation were similar to those of temporal validation. SDMs based on temporal occupancy had lower false discovery rates ( $t = -13.1, p = .006$ ) and false-positive rates ( $t = -25.1, p = .002$ ), and higher false omission rates ( $t = 11.0, p = .008$ ) and false-negative rates ( $t = 20.1, p = .002$ ) compared with models based on presence-absence (Figure 5b). Presence-based models had lower false-negative rates (10%–13%) and false omission rates (34%–41%). Temporal occupancy-based models had comparable false-negative rates (33%–41%) across modelling approaches. False discovery rates (8%–11%) and false-positive rates (15%–26%) were lower for temporal occupancy-based SDMs than presence-based false discovery rates (16%–18%) and false-positive rates (48%–60%).

## 4 | DISCUSSION

SDMs are important tools for conservation and resource management, and many studies have considered how to optimize them (Elith et al., 2006; Guillera-Aroita et al., 2015; Liu et al., 2011; Merow et al., 2014). Using a dataset of 189 avian species in North America and a variety of commonly employed modelling approaches, we found that for most species, SDMs based on temporal occupancy generally outperformed SDMs based on presence-absence. SDMs using temporal occupancy were more conservative in predicting presences relative to presence-based SDMs because locations where a species was observed infrequently were down-weighted by this measure. Additionally, the improvement in performance was greatest for species with lower range occupancy, which had more patchy distributions (Hurlbert & White, 2007) within their breeding range.



**FIGURE 5** Average error rates ( $\pm$  standard error) across all species for (a) 15-year temporal cross-validation and (b) 15-year spatial cross-validation of presence-based ("Pr") or temporal occupancy-based ("Occ") species distribution models using four modelling approaches (GAM, GLM, RF and MaxEnt). Error rates are defined in Table 1

Initially, we used a 15-year time series to calculate temporal occupancy because a long sampling window provides a more precise estimate of occupancy. However, it is uncommon to have datasets with such a long continuous time series, so we compared the 15-year temporal occupancy results to those from a shorter sampling window of five non-contiguous years. We saw similar results between the five-year and 15-year time series, with the five-year time series performing slightly better. In the 5-year time series, temporal occupancy models outperformed presence-absence models for even more species than in the 15-year time series. Finally, species with large differences between temporal occupancy and presence-absence RMSEs performed the same regardless of the duration of sampling window, whereas species with smaller ranges or fewer occurrences at a site were more likely to be excluded from the analysis with the shorter sampling window. Overall, there appears to be no reduction in effectiveness of temporal occupancy-based SDMs using only five sampling years, indicating that SDMs using temporal occupancy for sites or taxonomic groups with only modest monitoring efforts over time likely still provide useful habitat suitability data.

There was interspecific variation in whether GLM, GAM or RF models had the lowest RMSE, but MaxEnt was consistently the poorest performer for most species. This is concerning because MaxEnt is a widely accepted method of creating SDMs (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). Our results indicate that MaxEnt may be adequate when a species is expected to be distributed uniformly in space with few "range holes," but may be problematic for species with few presences or low range occupancy. More broadly, several species with many presences and few absences were overfit using the presence-absence GAMs, while the same species were not overfit using the GAM temporal occupancy SDMs. These species had either complete or partial separation, meaning the outcome values were almost perfectly fit by the predictors. All overfit species had range occupancy values of 70% or more. This overfitting demonstrates a potential drawback of using presence to create SDMs, and a potential advantage of using a continuous response variable like temporal occupancy.

One of the challenges of creating SDMs is balancing types of errors in order to maximize correct prediction of presences while not under-predicting correct absences. Any statistical approach for developing SDMs will be subject to a trade-off between errors of omission (false negatives) and errors of commission (false positives; Barry & Elith, 2006; Elith et al., 2006; Wintle et al., 2005). Our results demonstrate that these trade-offs need to be considered carefully when creating SDMs, as the various modelling methods we examined performed differently with respect to these two types of errors and depending on whether performance was evaluated based on spatial or temporal cross-validation. One consistent result was a much lower false-positive rate for models using temporal occupancy than for those using presence-absence. As such, temporal occupancy might be the superior metric when striving to minimize errors of commission, such as when creating SDMs for rare species

or species with low range occupancy that are known to patchily occupy the landscape.

Differences in error rates between temporal and spatial cross-validation were consistent across all models. When predicting presence-absence at the same set of sites but outside of the temporal window in which models were trained, false discovery rates were undesirably high (40%–60% on average), while false omission rates were low. On the one hand, this implies that presences were being predicted too liberally. Sites at which a species had been observed previously might be predicted to be presences again, and yet, the species might fail to be detected in a particular year, resulting in apparent false positives. On the other hand, absences were being predicted conservatively, as sites at which a species had never been observed previously could reliably be predicted as absences by the models. In contrast, when making predictions in new spatial locations during the same time window, false discovery rates were low, while false omission rates were 40%–60%. This means that, spatially, SDMs are good at predicting a core set of environmental conditions strongly associated with species presences, but fail to adequately recognize the full set of conditions over which presences are possible. We encourage others working with SDMs to compare results from both spatial cross-validation and temporal cross-validation of their models when possible to shed more light on how, where and why the models perform as they do.

Temporal occupancy is easy to calculate for annual monitoring datasets in which individual sites have been surveyed over multiple years, and our work suggests that the information encoded in this simple metric can improve species distribution models relative to presence-absence-based models. However, it is important to realize that temporal occupancy will necessarily underestimate true occupancy at a site given that species are not always detected when present. Dynamic occupancy models (MacKenzie et al., 2005) provide a framework for explicitly accounting for imperfect detection and have been increasingly used for range-wide species distribution models (Clement et al., 2016; Green et al., 2019; Kéry et al., 2013; Rota et al., 2011; Rushing et al., 2019). These dynamic occupancy models provide the additional advantage of explicitly modelling the colonization and extinction processes (Kéry et al., 2013; MacKenzie et al., 2003), and thus, the further development of such models is clearly a productive area of continued research. In particular, it would be useful to understand how the error rates from such models were compared to those from the temporal occupancy-based models we examined here.

Nevertheless, the failure to account for imperfect detection does not impede our ability to compare SDM performance between presence-absence and temporal occupancy models, the primary goal of this manuscript. Furthermore, the use of temporal occupancy as an indicator of whether a species truly maintains a viable population at a given site despite imperfect detection has been validated with simulation models (Snell Taylor, Coyle, et al., 2020). High temporal occupancy accurately reflected sites

with high suitability except in cases where the focal site was an unsuitable population sink surrounded by a landscape of more than 60% suitable habitat. Low temporal occupancy accurately reflected sites with low suitability as long as the probability of detection was greater than ~30%, which is usually the case in birds based on several multispecies compilations (Boulinier et al., 1998; Johnston et al., 2014). That said, developing reliable SDMs for species known to be difficult to detect may benefit from (a) the use of data obtained from targeted monitoring efforts rather than generic surveys and (b) analysis by more sophisticated dynamic occupancy modelling approaches.

Overall, the use of temporal occupancy can improve SDMs by prioritizing areas where the species persists more regularly and using only a few years of data over a 15-year sampling period proved to be equally as effective for bird species. Temporal occupancy does not necessarily improve the performance of SDMs for abundant and wide-ranging species, but is beneficial for modelling species that occur infrequently or occupy a smaller proportion of their range. We hope this study encourages scientists to consider the role time series data can play in improving SDMs using this simple framework, and to assess whether it is appropriate to use temporal occupancy in their analyses when multiple years of surveys are available.

#### ACKNOWLEDGEMENTS

We are grateful to the volunteers who collect data for the North American Breeding Bird Survey. We thank J. Umbanhowar for statistical consultation and E. White for suggestions regarding methodology. This work was made possible by funding from the National Science Foundation through DEB-1354563. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

#### PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ddi.13296>.

#### DATA AVAILABILITY STATEMENT

Data and code associated with this study are available from the Zenodo Digital Repository (<https://doi.org/10.5281/zenodo.4672696>).

#### ORCID

Sara Snell Taylor  <https://orcid.org/0000-0002-0353-9613>

Grace Di Cecco  <https://orcid.org/0000-0003-1872-5117>

Allen H. Hurlbert  <https://orcid.org/0000-0002-5678-9907>

#### REFERENCES

- Bahn, V., & McGill, B. J. (2007). Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography*, 16(6), 733–742. <https://doi.org/10.1111/j.1466-8238.2007.00331.x>
- Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43(3), 413–423. <https://doi.org/10.1111/j.1365-2664.2006.01136.x>
- Bell, D. M., & Schlaepfer, D. R. (2016). On the dangers of model complexity without ecological justification in species distribution modeling. *Ecological Modelling*, 330, 50–59. <https://doi.org/10.1016/j.ecolmodel.2016.03.012>
- Belmaker, J. (2009). Species richness of resident and transient coral-dwelling fish responds differentially to regional diversity. *Global Ecology and Biogeography*, 18(4), 426–436. <https://doi.org/10.1111/j.1466-8238.2009.00456.x>
- Boulinier, T., Nichols, J. D., Sauer, J. R., Hines, J. E., & Pollock, K. H. (1998). Estimating Species Richness: The importance of heterogeneity in species detectability. *Ecology*, 79(3), 1018–1028.
- Butcher, G. S., Robbins, C. S., Bystrak, D., & Geissler, P. H. (1987). The breeding bird survey: Its first fifteen years, 1965–1979. *The Condor*, 89(3), 681–682. <https://doi.org/10.2307/1368666>
- Clement, M. J., Hines, J. E., Nichols, J. D., Pardieck, K. L., & Ziolkowski, D. J. (2016). Estimating indices of range shifts in birds using dynamic models when detection is imperfect. *Global Change Biology*, 22(10), 3273–3285. <https://doi.org/10.1111/gcb.13283>
- Comte, L., & Grenouillet, G. (2013). Species distribution modelling and imperfect detection: Comparing occupancy versus consensus methods. *Diversity and Distributions*, 19(8), 996–1007. <https://doi.org/10.1111/ddi.12078>
- Coyle, J. R., Hurlbert, A. H., & White, E. P. (2013). Opposing mechanisms drive richness patterns of core and transient bird species. *The American Naturalist*, 181(4), E83–E90. <https://doi.org/10.1086/669903>
- de Araújo, C. B., Marcondes-Machado, L. O., & Costa, G. C. (2014). The importance of biotic interactions in species distribution models: A test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography*, 41(3), 513–523. <https://doi.org/10.1111/jbi.12234>
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. M., Peterson, A. T., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330–342. <https://doi.org/10.1111/j.2041-210X.2010.00036.x>
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Engler, J. O., Stiels, D., Schidelko, K., Strubbe, D., Quillfeldt, P., & Brambilla, M. (2017). Avian SDMs: Current state, challenges, and opportunities. *Journal of Avian Biology*, 48(12), 1483–1504. <https://doi.org/10.1111/jav.01248>
- Franklin, J. (2010). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, 16(3), 321–330. <https://doi.org/10.1111/j.1472-4642.2010.00641.x>
- Freeman, B. G., & Mason, N. A. (2015). The Geographic Distribution of a Tropical Montane Bird Is Limited by a Tree: Acorn Woodpeckers (*Melanerpes formicivorus*) and Colombian Oaks (*Quercus humboldtii*) in the Northern Andes. *PLoS One*, 10(6), e0128675. <https://doi.org/10.1371/journal.pone.0128675>
- Godsoe, W., Franklin, J., & Blanchet, F. G. (2017). Effects of biotic interactions on modeled species' distribution can be masked by environmental gradients. *Ecology and Evolution*, 7(2), 654–664. <https://doi.org/10.1002/ece3.2657>
- Gomes, V. H. F., IJff, S. D., Raes, N., Amaral, I. L., Salomão, R. P., de Souza Coelho, L., de Almeida Matos, F. D., Castilho, C. V., de Andrade Lima Filho, D., López, D. C., Guevara, J. E., Magnusson, W. E., Phillips, O. L., Wittmann, F., de Jesus Veiga Carim, M., Martins, M. P., Irumé, M. V., Sabatier, D., Molino, J. F., ... Steege, H. (2018). Species Distribution

- Modelling: Contrasting presence-only models with plot abundance data. *Scientific Reports*, 8, 1003.
- Green, A. W., Pavlacky, D. C., & George, T. L. (2019). A dynamic multi-scale occupancy model to estimate temporal dynamics and hierarchical habitat use for nomadic species. *Ecology and Evolution*, 9(2), 793–803. <https://doi.org/10.1002/ece3.4822>
- Guillera-Aroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3), 276–292. <https://doi.org/10.1111/geb.12268>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2011). Package 'dismo'. Available at <https://cran.r-project.org/web/packages/dismo/index.html>.
- Hurlbert, A. H., & Haskell, J. P. (2003). The effect of energy and seasonality on avian species richness and community composition. *The American Naturalist*, 161(1), 83–97. <https://doi.org/10.1086/345459>
- Hurlbert, A. H., & White, E. P. (2007). Ecological correlates of geographical range occupancy in North American birds. *Global Ecology and Biogeography*, 16(6), 764–773. <https://doi.org/10.1111/j.1466-8238.2007.00335.x>
- Jenkins, M. F., White, E. P., & Hurlbert, A. H. (2018). The proportion of core species in a community varies with spatial scale and environmental heterogeneity. *PeerJ*, 6, e6019.
- Johnston, A., Newson, S. E., Risely, K., Musgrove, A. J., Massimino, D., Baillie, S. R., & Pearce-Higgins, J. W. (2014). Species traits explain variation in detectability of UK birds. *Bird Study*, 61(3), 340–350. <https://doi.org/10.1080/00063657.2014.941787>
- Kéry, M., Guillera-Aroita, G., & Lahoz-Monfort, J. J. (2013). Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography*, 40(8), 1463–1474. <https://doi.org/10.1111/jbi.12087>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by random Forest. *R News*, 2(3), 18–22. Available at <https://CRAN.R-project.org/package=randomForest>
- Liu, C., White, M., & Newell, G. (2011). Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography*, 34(2), 232–243. <https://doi.org/10.1111/j.1600-0587.2010.06354.x>
- MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, 84(8), 2200–2207. <https://doi.org/10.1890/02-3090>
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2005). *Occupancy estimation and modeling* (1st ed.). Academic Press.
- Magurran, A. E., & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422(6933), 714–716.
- Marschner, I. C. (2019). glm2: Fitting generalized linear models with convergence problems. *The R Journal*, 3(2), 12–15. Available at <https://CRAN.R-project.org/package=glm2>
- Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, 415, 108837. <https://doi.org/10.1016/j.ecolmodel.2019.108837>
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37(12), 1267–1281. <https://doi.org/10.1111/ecog.00845>
- Merow, C., Smith, M. J., & Silander, J. A. Jr (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10), 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo, M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., ... Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, 89(3), e01370. <https://doi.org/10.1002/ecm.1370>
- Oliver, T. H., Gillings, S., Girardello, M., Rapacciuolo, G., Brereton, T. M., Siriwardena, G. M., Roy, D. B., Pywell, R., & Fuller, R. J. (2012). Population density but not stability can be predicted from species distribution models. *Journal of Applied Ecology*, 49(3), 581–590. <https://doi.org/10.1111/j.1365-2664.2012.02138.x>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Qiao, H., Soberón, J., & Peterson, A. T. (2015). No silver bullets in correlative ecological niche modelling: Insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and Evolution*, 6(10), 1126–1136. <https://doi.org/10.1111/2041-210X.12397>
- Ralston, J., DeLuca, W. V., Feldman, R. E., & King, D. I. (2016). Realized climate niche breadth varies with population trend and distribution in North American birds. *Global Ecology and Biogeography*, 25(10), 1173–1180. <https://doi.org/10.1111/geb.12490>
- Rich, J. L., & Currie, D. J. (2018). Are North American bird species' geographic ranges mainly determined by climate? *Global Ecology and Biogeography*, 27(4), 461–473. <https://doi.org/10.1111/geb.12708>
- Rota, C. T., Fletcher, R. J., Evans, J. M., & Hutto, R. L. (2011). Does accounting for imperfect detection improve species distribution models? *Ecography*, 34(4), 659–670. <https://doi.org/10.1111/j.1600-0587.2010.06433.x>
- Rushing, C. S., Royle, J. A., Ziolkowski, D. J., & Pardieck, K. L. (2019). Modeling spatially and temporally complex range dynamics when detection is imperfect. *Scientific Reports*, 9(1), 12805. <https://doi.org/10.1038/s41598-019-48851-5>
- Snell Taylor, S. J., Coyle, J. R., White, E. P., & Hurlbert, A. H. (2020). A simulation study of the use of temporal occupancy for identifying core and transient species. *PLoS One*, 15(10), e0241198. <https://doi.org/10.1371/journal.pone.0241198>
- Snell Taylor, S. J., Evans, B. S., White, E. P., & Hurlbert, A. H. (2018). The prevalence and impact of transient species in ecological communities. *Ecology*, 99(8), 1825–1835. <https://doi.org/10.1002/ecy.2398>
- Snell Taylor, S. J., Umbanhowar, J., & Hurlbert, A. H. (2020). The relative importance of biotic and abiotic determinants of temporal occupancy for avian species in North America. *Global Ecology and Biogeography*, 29(4), 736–747. <https://doi.org/10.1111/geb.13064>
- Steen, V. A., Elphick, C. S., & Tingley, M. W. (2019). An evaluation of stringent filtering to improve species distribution models from citizen

- science data. *Diversity and Distributions*, 25(12), 1857–1869. <https://doi.org/10.1111/ddi.12985>
- Supp, S. R., Koons, D. N., & Ernest, S. K. M. (2015). Using life history trade-offs to understand core-transient structuring of a small mammal community. *Ecosphere*, 6(10), 1–15. <https://doi.org/10.1890/ES15-00239.1>
- Thuiller, W., Pironon, S., Psomas, A., Barbet-Massin, M., Jiguet, F., Lavergne, S., Pearman, P. B., Renaud, J., Zupan, L., & Zimmermann, N. E. (2014). The European functional tree of bird life in the face of global change. *Nature Communications*, 5, 3118. <https://doi.org/10.1038/ncomms4118>
- Tobler, M. W., Kéry, M., Hui, F. K. C., Guillera-Arroita, G., Knaus, P., & Sattler, T. (2019). Joint species distribution models with species correlations and imperfect detection. *Ecology*, 100, e02754. <https://doi.org/10.1002/ecy.2754>
- Ulrich, W., & Ollik, M. (2004). Frequent and occasional species and the shape of relative-abundance distributions. *Diversity and Distributions*, 10(4), 263–269. <https://doi.org/10.1111/j.1366-9516.2004.00082.x>
- Umaña, M. N., Zhang, C., Cao, M., Lin, L., & Swenson, N. G. (2017). A core-transient framework for trait-based community ecology: An example from a tropical tree seedling community. *Ecology Letters*, 20(5), 619–628. <https://doi.org/10.1111/ele.12760>
- Wintle, B. A., Kavanagh, R. P., McCarthy, M. A., & Burgman, M. A. (2005). Estimating and dealing with detectability in occupancy surveys for forest owls and aboreal marsupials. *Journal of Wildlife Management*, 69(3), 905–917.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman and Hall/CRC. <https://CRAN.R-project.org/package=mgcv>
- Zambrano-Bigiarini, M. (2017). *hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series*. R package version 0.3-10. <http://hzambran.github.io/hydroGOF/>
- Zuckerberg, B., Bonter, D., Hochachka, W., Koenig, W., DeGaetano, A., & Dickinson, J. (2011). Climatic constraints on wintering bird distributions are modified by urbanization and weather. *Journal of Animal Ecology*, 80(2), 403–413. <https://doi.org/10.1111/j.1365-2656.2010.01780.x>

## BIOSKETCHES

**Sara Snell Taylor** recently graduated from the University of North Carolina with her PhD. Her research considers how species interactions and the environment work to shape species distributions at large scales. **Grace Di Cecco** is a PhD student at the University of North Carolina. Her research focuses on understanding how populations, communities and species distributions are shaped by environmental change. **Allen Hurlbert** is interested in the processes governing the distribution of organisms in space and time.

Author contributions: SST, GDC, AHH conceived and designed the research study. SST executed the study and wrote the manuscript. All authors contributed to the drafts and gave final approval for publication.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Snell Taylor S, Di Cecco G, Hurlbert AH. Using temporal occupancy to predict avian species distributions. *Divers Distrib*. 2021;00:1–12. <https://doi.org/10.1111/ddi.13296>

## APPENDIX 1

Column headings and descriptions for Table S1

Column name	Column description
AOU	Unique species ID from the Breeding Bird Survey
CommonName	Species common name
ScientificName	Species scientific name
Migclass	Species migratory class
n_routes	Number of routes a species occurred on over the 15-year study period
rmse_occ_15yr	RMSE value from GLM 15-year temporal occupancy SDM
rmse_pres_15yr	RMSE value from GLM 15-year presence-absence SDM
rmse_gam_15yr	RMSE value from GAM 15-year temporal occupancy SDM
rmse_gam_pres_15yr	RMSE value from GAM 15-year presence-absence SDM
rmse_rf_15yr	RMSE value from RF 15-year temporal occupancy SDM
rmse_rf_pres_15yr	RMSE value from RF 15-year presence-absence SDM
rmse_me_pres_15yr	RMSE value from MaxEnt 15-year presence-only SDM
rmse_occ_5yr	RMSE value from GLM 5-year temporal occupancy SDM
rmse_pres_5yr	RMSE value from GLM 5-year presence-absence SDM
rmse_gam_5yr	RMSE value from GAM 5-year temporal occupancy SDM
rmse_gam_pres_5yr	RMSE value from GAM 5-year presence-absence SDM
rmse_rf_5yr	RMSE value from RF 5-year temporal occupancy SDM
rmse_rf_pres_5yr	RMSE value from RF 5-year presence-absence SDM
rmse_me_pres_5yr	RMSE value from MaxEnt 5-year presence-only SDM
sensitivity_glmocc_temporal	Temporal cross-validation false-negative rate from GLM 15-year temporal occupancy SDM
specificity_glmocc_temporal	Temporal cross-validation false-positive rate from GLM 15-year temporal occupancy SDM
pp_glmocc_temporal	Temporal cross-validation false discovery rate from GLM 15-year temporal occupancy SDM
np_glmocc_temporal	Temporal cross-validation false omission rate from GLM 15-year temporal occupancy SDM
sensitivity_glmpr_temporal	Temporal cross-validation false-negative rate from GLM 15-year presence-absence SDM

Column name	Column description
specificity_glmpr_temporal	Temporal cross-validation false-positive rate from GLM 15-year presence-absence SDM
pp_glmpr_temporal	Temporal cross-validation false discovery rate from GLM 15-year presence-absence SDM
np_glmpr_temporal	Temporal cross-validation false omission rate from GLM 15-year presence-absence SDM
sensitivity_gamocc_temporal	Temporal cross-validation false-negative rate from GAM 15-year temporal occupancy SDM
specificity_gamocc_temporal	Temporal cross-validation false-positive rate from GAM 15-year temporal occupancy SDM
pp_gamocc_temporal	Temporal cross-validation false discovery rate from GAM 15-year temporal occupancy SDM
np_gamocc_temporal	Temporal cross-validation false omission rate from GAM 15-year temporal occupancy SDM
sensitivity_gampr_temporal	Temporal cross-validation false-negative rate from GAM 15-year presence-absence SDM
specificity_gampr_temporal	Temporal cross-validation false-positive rate from GAM 15-year presence-absence SDM
pp_gampr_temporal	Temporal cross-validation false discovery rate from GAM 15-year presence-absence SDM
np_gampr_temporal	Temporal cross-validation false omission rate from GAM 15-year presence-absence SDM
sensitivity_rfocc_temporal	Temporal cross-validation false-negative rate from RF 15-year temporal occupancy SDM
specificity_rfocc_temporal	Temporal cross-validation false-positive rate from RF 15-year temporal occupancy SDM
pp_rfocc_temporal	Temporal cross-validation false discovery rate from RF 15-year temporal occupancy SDM
np_rfocc_temporal	Temporal cross-validation false omission rate from RF 15-year temporal occupancy SDM
sensitivity_rfpr_temporal	Temporal cross-validation false-negative rate from RF 15-year presence-absence SDM
specificity_rfpr_temporal	Temporal cross-validation false-positive rate from RF 15-year presence-absence SDM
pp_rfpr_temporal	Temporal cross-validation false discovery rate from RF 15-year presence-absence SDM

Column name	Column description
np_rfpr_temporal	Temporal cross-validation false omission rate from RF 15-year presence-absence SDM
sensitivity_max_temporal	Temporal cross-validation false-negative rate from MaxEnt 15-year presence-only SDM
specificity_max_temporal	Temporal cross-validation false-positive rate from MaxEnt 15-year presence-only SDM
pp_max_temporal	Temporal cross-validation false discovery rate from MaxEnt 15-year presence-only SDM
np_max_temporal	Temporal cross-validation false omission rate from MaxEnt 15-year presence-only SDM
sensitivity_glmocc_spatial	Spatial cross-validation false-negative rate from GLM 15-year spatial occupancy SDM
specificity_glmocc_spatial	Spatial cross-validation false-positive rate from GLM 15-year spatial occupancy SDM
pp_glmocc_spatial	Spatial cross-validation false discovery rate from GLM 15-year spatial occupancy SDM
np_glmocc_spatial	Spatial cross-validation false omission rate from GLM 15-year spatial occupancy SDM
sensitivity_glmpr_spatial	Spatial cross-validation false-negative rate from GLM 15-year presence-absence SDM
specificity_glmpr_spatial	Spatial cross-validation false-positive rate from GLM 15-year presence-absence SDM
pp_glmpr_spatial	Spatial cross-validation false discovery rate from GLM 15-year presence-absence SDM
np_glmpr_spatial	Spatial cross-validation false omission rate from GLM 15-year presence-absence SDM
sensitivity_gamocc_spatial	Spatial cross-validation false-negative rate from GAM 15-year spatial occupancy SDM
specificity_gamocc_spatial	Spatial cross-validation false-positive rate from GAM 15-year spatial occupancy SDM
pp_gamocc_spatial	Spatial cross-validation false discovery rate from GAM 15-year spatial occupancy SDM
np_gamocc_spatial	Spatial cross-validation false omission rate from GAM 15-year spatial occupancy SDM

Column name	Column description
sensitivity_gampr_spatial	Spatial cross-validation false-negative rate from GAM 15-year presence-absence SDM
specificity_gampr_spatial	Spatial cross-validation false-positive rate from GAM 15-year presence-absence SDM
pp_gampr_spatial	Spatial cross-validation false discovery rate from GAM 15-year presence-absence SDM
np_gampr_spatial	Spatial cross-validation false omission rate from GAM 15-year presence-absence SDM
sensitivity_rfocc_spatial	Spatial cross-validation false-negative rate from RF 15-year spatial occupancy SDM
specificity_rfocc_spatial	Spatial cross-validation false-positive rate from RF 15-year spatial occupancy SDM
pp_rfocc_spatial	Spatial cross-validation false discovery rate from RF 15-year spatial occupancy SDM
np_rfocc_spatial	Spatial cross-validation false omission rate from RF 15-year spatial occupancy SDM
sensitivity_rfpr_spatial	Spatial cross-validation false-negative rate from RF 15-year presence-absence SDM
specificity_rfpr_spatial	Spatial cross-validation false-positive rate from RF 15-year presence-absence SDM
pp_rfpr_spatial	Spatial cross-validation false discovery rate from RF 15-year presence-absence SDM
np_rfpr_spatial	Spatial cross-validation false omission rate from RF 15-year presence-absence SDM
sensitivity_max_spatial	Spatial cross-validation false-negative rate from MaxEnt 15-year presence-only SDM
specificity_max_spatial	Spatial cross-validation false-positive rate from MaxEnt 15-year presence-only SDM
pp_max_spatial	Spatial cross-validation false discovery rate from MaxEnt 15-year presence-only SDM
np_max_spatial	Spatial cross-validation false omission rate from MaxEnt 15-year presence-only SDM