



Testing species assemblage predictions from stacked and joint species distribution models

Damaris Zurell^{1,2}  | Niklaus E. Zimmermann^{2,3}  | Helge Gross² |
Andri Baltensweiler² | Thomas Sattler⁴ | Rafael O. Wüest² 

¹Geography Department, Humboldt-Universität zu Berlin, Berlin, Germany

²Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

³Department of Environmental Systems Science, Swiss Federal Institute of Technology ETH, Zürich, Switzerland

⁴Swiss Ornithological Institute, Sempach, Switzerland

Correspondence

Damaris Zurell, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany.

Email: damaris.zurell@hu-berlin.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: 361/1-1; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 310030L_170059 and PZ00P3_168136/1

Abstract

Aim: Predicting the spatial distribution of species assemblages remains an important challenge in biogeography. Recently, it has been proposed to extend correlative species distribution models (SDMs) by taking into account (a) covariance between species occurrences in so-called joint species distribution models (JSDMs) and (b) ecological assembly rules within the SESAM (spatially explicit species assemblage modelling) framework. Yet, little guidance exists on how these approaches could be combined. We, thus, aim to compare the accuracy of assemblage predictions derived from stacked and from joint SDMs.

Location: Switzerland.

Taxon: Birds, tree species.

Methods: Based on two monitoring schemes (national forest inventory and Swiss breeding bird atlas), we built SDMs and JSDMs for tree species (at 100 m resolution) and forest birds (at 1 km resolution). We tested accuracy of species assemblage and richness predictions on holdout data using different stacking procedures and ecological assembly rules.

Results: Despite minor differences, results were consistent between birds and tree species. Cross-validated species-level model performance was generally higher in SDMs than JSDMs. Differences in species richness and assemblage predictions were larger between stacking procedures and ecological assembly rules than between stacked SDMs and JSDMs. On average, predictions were slightly better for stacked SDMs compared to JSDMs, probabilistic stacks outperformed binary stacks, and ecological assembly rules yielded best predictions.

Main conclusions: When predicting the composition of species assemblages, the choice of stacking procedure and ecological assembly rule seems more decisive than differences in underlying model type (SDM vs. JSDM). JSDMs do not seem to improve community predictions compared to SDMs or improve predictions for rare species. Still, JSDMs may provide additional insights into community assembly and may help deriving hypotheses about prevailing biotic interactions in the system. We provide simple rules of thumb for choosing appropriate modelling pathways. Future

studies should test these preliminary guidelines for other taxa and biogeographical realms as well as for other JSDM algorithms.

KEYWORDS

AUC, community assembly, community composition, ecological niche, habitat suitability models, prediction, sensitivity, species richness, specificity, TSS

1 | INTRODUCTION

Quantifying the reasons why species inhabit some places but not others remains a central question in biogeography. In recent decades, we have seen increasing interest in species distributions modelling and quantitative predictions of where species and communities can be found (D'Amen, Rahbek, Zimmermann, & Guisan, 2017; Zimmermann, Edwards, Graham, Pearman, & Svenning, 2010). Such predictions have aided, for example, reserve selection (Kremen et al., 2008), impact assessments (Thuiller et al., 2014) and invasive species management (Dullinger et al., 2017). This development has been spurred by an increasing availability of environmental and biodiversity data, for example through remote sensing, open-access databases and citizen science. Also, modelling techniques have advanced and new modelling platforms are introduced continually. However, the wealth of possibilities may also complicate modelling choice and different model techniques and frameworks should be benchmarked and tested (continually) to guide practitioners and policy makers.

In this study, we are particularly interested in spatial predictions of community attributes, primarily in the composition but also in the richness of local communities. Such predictions could

be of special interest for species conservation and reserve design. Different spatiotemporal processes influence community assembly (HilleRisLambers, Adler, Harpole, Levine, & Mayfield, 2012). The regional species pool is constrained by (evolutionary and biogeographical) history, the local species pool by dispersal and other stochastic events. Only a subset of the local species pool will then be present in the realized assemblage (Sattler, Obrist, Duelli, & Moretti, 2011). Often, this is metaphorically described as species passing through an abiotic niche and a biotic niche filter (Figure 1; Götzenberger et al., 2012). How we could predict local community richness and composition still remains a challenge even though different modelling frameworks have been proposed (D'Amen et al., 2017; Ferrier & Guisan, 2006). Correlative species distribution models (SDMs) constitute the most widely used tool to predict single species distributions under current and future environments (Elith & Leathwick, 2009; Guisan & Thuiller, 2005; Guisan & Zimmermann, 2000). The apparent and logical next step is then to overlay these single species predictions to obtain community assemblage predictions in so-called stacked species distribution models (S-SDMs) (Guisan & Rahbek, 2011). However, SDMs ignore (or only implicitly consider) biotic interactions (Kisling et al., 2012; Wisz et al., 2013) and, thus, cannot

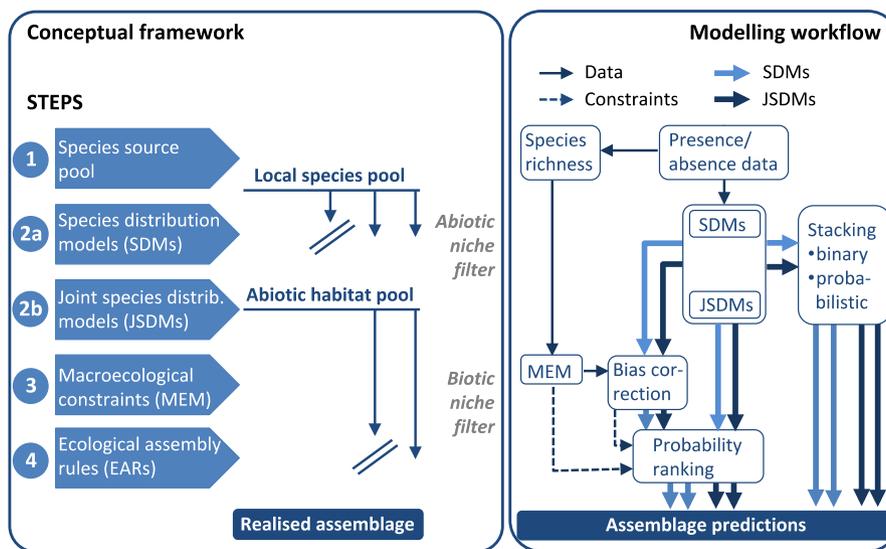


FIGURE 1 Schematic representation of the conceptual framework and our modelling workflow for comparing assemblage prediction success. Left panel: The spatially explicit species assemblage modelling (SESAM) framework applies successive modelling steps to imitate the filtering cascade of community assembly (steps 1, 2a, 3–4; adapted from Guisan & Rahbek, 2011). Joint species distribution models (JSDMs; step 2b) should yield information on both abiotic and biotic constraints. Right panel: The modelling workflow exemplifies how species distribution models (SDMs) and JSDMs could be stacked to directly yield assemblage predictions and how macroecological constraints (with or without bias correction) are used to inform the ecological assembly rules (here, probability ranking rule)



account for the biotic niche filter in community assembly (Guisan & Rahbek, 2011). More recently, joint species distribution models (JSDMs) have been developed that incorporate species co-occurrence data into SDMs and can thus provide a bridge between species distribution modelling and community ecology (Pollock et al., 2014; Warton et al., 2015). Hitherto, not many studies have explicitly compared SDM and JSDM model performance, and the ability of S-SDMs and JSDMs to predict species assemblages (Harris, 2015).

Joint species distribution models simultaneously estimate the species-environment relationship of multiple species and the residual correlation between those species that could be indicative of biotic interactions (Zurell, Pollock, & Thuiller, 2018), but also missing environmental information and other ecological processes. Similar to SDMs, JSDMs generate site-level occurrence probabilities of the different species, which have to be combined ("stacked") in order to yield predictions of local community composition. Previous studies on S-SDMs have shown that the choice of stacking procedure is non-trivial. Binary stacks, for which a threshold is applied to transform continuous species occurrence probabilities to binary maps prior to stacking, often lead to overprediction of species richness (D'Amen, Dubuis, et al., 2015; Dubuis et al., 2011), but not always (D'Amen, Pradervand, & Guisan, 2015; Zurell, Zimmermann, Sattler, Nobis, & Schröder, 2016). This overprediction of species richness has been attributed to biotic interactions that limit realized assemblages (Guisan & Rahbek, 2011) while other authors argue that thresholding site-level occurrence probabilities generally leads to statistical bias, which in turn causes overprediction (Calabrese, Certain, Kraan, & Dormann, 2014). Rather, it has been suggested that probabilistic stacking, for which species presences are obtained from repeated Bernoulli trials (Calabrese et al., 2014; Dubuis et al., 2011; Pottier et al., 2013) should be preferred over binary stacking (Calabrese et al., 2014). However, probabilistic stacking (as well as macroecological models, MEMs, that directly relate species richness to environmental predictors, Guisan & Rahbek, 2011) often lead to overprediction of low species richness and underprediction of high species richness, which can partially be explained by (sub-scale) environmental heterogeneity and species' prevalence (Zurell et al., 2016). Simply summing the predicted occurrence probabilities, which has been proposed as the statistically most correct way of stacking (Calabrese et al., 2014) impedes the analysis of turnover analyses in space or time, as the output does not explicitly reveal species identities. While the accuracy of species richness predictions has been assessed (Dubuis et al., 2011), no study has compared the accuracy of assemblage predictions between binary and probabilistic (Bernoulli trial) stacking methods, even though data on community composition are crucial for adequately assessing potential species turnover in response to global change impacts or for prioritizing conservation actions (Pollock, Thuiller, & Jetz, 2017).

Guisan and Rahbek (2011) proposed the SESAM ("spatially explicit species assemblage modelling") framework to overcome the richness overprediction of binary S-SDMs and to account for biotic filtering. SESAM applies successive filters to select how many and which species from the local species pool could be present in the realized assemblages (Figure 1). Specifically, Guisan and Rahbek

(2011) propose to use SDMs as abiotic niche filter, and macroecological constraints and ecological assembly rules as biotic niche filter. Recent applications indicate that the SESAM framework could improve assemblage predictions compared to binary S-SDMs (D'Amen, Dubuis, et al., 2015; D'Amen, Pradervand, et al., 2015). Nevertheless, Calabrese et al. (2014) caution that SESAM could be of limited value if the species richness predictions (used as macroecological constraint, Figure 1) are strongly biased for low and high species richness sites. To overcome this bias, Calabrese et al. (2014) proposed a maximum-likelihood approach to adjust site-level occurrence probabilities based on observed or estimated species richness. To our knowledge, it has not been tested whether this bias correction improves the accuracy of assemblage predictions. Also, it is unclear where JSDMs can be placed in the SESAM framework and whether assemblage predictions from JSDMs will also, similar to S-SDMs, yield improved accuracies when applying additional biotic filters.

Thus, we propose to compare S-SDMs and JSDMs in their ability to predict species assemblages. Specifically, we test the effect of different stacking procedures (binary vs. probabilistic), the effect of applying macroecological constraints and ecological assembly rules (as proposed in the SESAM framework), and the effect of bias correction in macroecological constraints (as proposed by Calabrese et al., 2014; Figure 1). As case studies, we use data from Swiss breeding bird atlas (Schmid, Luder, Naef-Daenzer, Graf, & Zbinden, 1998) and from Swiss forest inventory (Brassel & Lischke, 2001) that constitute monitoring schemes of high quality and high detection probabilities.

2 | MATERIALS AND METHODS

2.1 | Species and environmental data

Two different taxonomic groups were selected as study systems to test the different modelling choices: (a) forest birds and (b) bush and tree species of Switzerland. We assume that all species in Switzerland can be summarized in the same local species pool and are not further constrained by historical differences or dispersal limitations (Figure 1). We selected climate, topography and vegetation structure as environmental predictor variables.

2.2 | Breeding birds

Bird presence-absence data were obtained from the Swiss breeding bird atlas at a 1×1 km resolution (Schmid et al., 1998). These data were recorded over a 4-year period (1993–1996) in usually three visits per year (two above the tree line) using a simplified territory mapping approach. Previous analyses have shown that this approach ensures high species detectability of approximately 90% (Kéry & Schmid, 2006). We concentrate on forest birds to minimize problems associated with missing environmental covariates, the magnitude of which could differ among species of different broad habitat types (Zurell et al., 2016). To avoid overfitting for rare species (Breiner, Guisan, Bergamini, & Nobis, 2015), we only considered species with at least 50 presences resulting in a total number of 56 forest bird

species in the study region (Table S1). Overall, many forest bird species were rather common (Figure S1).

2.3 | Forest inventory

Tree species presence–absence data were obtained from the Swiss National Forest Inventory (NFI). The NFI samples Switzerland on a regular grid (spacing 1.4 km), and in case the sample falls into forest it records forest characteristics in a maximal area of 50 × 50 m (Brassel & Lischke, 2001). We aggregated the NFI presence–absence data to 100 × 100 m plot size to match the minimum grain of available environmental data (details see below). Analogously to the breeding bird data, we only considered species with at least 50 presences resulting in a total number of 63 tree and shrub species in the study region (hereafter referred to as tree species; Table S2). Overall, many tree species were rare with the majority of species showing a prevalence below 10% (Figure S1).

2.4 | Environmental data

We prepared the environmental predictor variables climate, topography and vegetation structure at the same resolution as the species data, meaning at 1 × 1 km for analyses of forest birds and at 100 × 100 m for analyses of tree species. Topography was represented using Slope, Aspect, the Topographical Position Index (TPI) and Topographical Wetness Index (TWI) (Wilson & Gallant, 2000) based on a 100 m digital elevation model. TPI in a cell corresponds to the difference of the focal cell to the mean of its eight surrounding cells, thus indicating whether the cell is in a depression (negative values) or in a rise (positive values). We further calculated potential monthly solar radiation as outlined in Hofierka, Suri, and Šúri (2002). Monthly average climate data for the period 1981–1990 were generated by interpolating station data from the Federal Office of Meteorology and Climatology MeteoSwiss to a resolution of 100 × 100 m using the DAYMET software (Thornton, Running, & White, 1997). From these monthly averages, we derived a set of 19 bioclimatic predictors (<http://worldclim.org/bio-clim>) as well as so-called degree days. Degree days constitute the sum of all monthly temperature values greater than a given threshold temperature multiplied by the total number of days (where thresholds 0 and 5°C were applied for DDEGO and DDEG5). We further calculated potential evapotranspiration (PET) using radiation as proposed by Makkink (1957), as this method was shown to best approximate PET in Switzerland (Xu & Singh, 2002). For both precipitation and PET we also calculated summer (April–September) and winter (October–March) averages as well as their ratio. We further used PET and precipitation data to calculate the moisture balance (MBAL) as the difference between precipitation and PET, and the moisture index (MIND) as the ratio between PET and precipitation. For the analyses of forest bird species, the topographical and climatic data were aggregated to 1 × 1 km grids using the mean as aggregate-function such that the grid cell-size matched the bird survey data.

Vegetation structure was described using LiDAR-derived data, processed with the LASTOOLS software (Isenburg, 2015). The Swiss-wide LiDAR dataset was acquired during multiple seasons between 2000 and 2007 and consists of discrete first and last pulse returns with a nominal footprint of 0.3 m and a point density of 0.5 points/m² (Artuso, Bovet, & Streilein, 2003). From the terrain corrected and classified LiDAR point cloud (heights of classified vegetation LiDAR returns minus interpolated DTM heights), we calculated the average height, their standard deviation and coefficient of variation as well as the corresponding 10th, 25th and 95th percentiles per 100 m and 1 km pixel, respectively. In addition, we also derived the canopy cover (percentage of first returns above 1 m; COV) and canopy density (ratio of all returns above 1 m divided by all returns; DNS). These variables were used to describe the vertical structure of vegetation. We further calculated both standard deviation and coefficient of variation in order to characterize vertical variation in LiDAR returns. In addition, we estimated the so-called foliage height diversity (FHD), which is the Shannon diversity index based on 5 m vertical bins as $H = -\sum p_i \ln(p_i)$, where p_i is the proportion of LiDAR returns in the 5 m bin i . These LiDAR variables were generated separately as raster datasets for both the 100 × 100 m grid that matches the tree species NFI data, as well as the 1 × 1 km grid that matches the bird survey data.

We calculated additional LiDAR-derived variables for the 1 × 1 km grid as we considered these important predictors for bird distributions in Swiss forests. On the one hand, we calculated understorey height diversity (UHD) analogously to FHD, but using 1 m bins restricted to below 12 m (Zellweger et al., 2016) as well as the ratio between the 95th and 25th percentile. On the other hand, we derived variables accounting for edge effects and fragmentation in order to describe the horizontal structural heterogeneity of the vegetation (Zellweger, Braunisch, Baltensweiler, & Bollmann, 2013). Based on the terrain corrected LiDAR point cloud we generated a gridded Canopy Height Model (CHM) with a grid size of 20 m. The CHM was classified into three classes, which are non-forest (vegetation height <1 m), understorey/midstorey (1–12 m) and canopy (>12 m). The length of edges between two height classes (e.g., non-forest/canopy) was calculated for each grid cell and summed up for the 1 × 1 km grid of the bird survey data. To measure the spatial dispersion or aggregation of the vegetation height classes understorey/midstorey and canopy, a clumpiness index was calculated for the 1 × 1 km grid using the software FRAGSTATS (McGarigal, Cushman, & Ene, 2012).

Because the forest bird data were recorded between 1993 and 1996 and the LiDAR data were recorded after 2000, there is a temporal mismatch between species data and vegetation data. Generally, the forest laws in Switzerland are very strict and we can, thus, rule out any major changes in the vegetation structure between these two periods. The main exception is storm damage due to the cyclone “Lothar” in 1999. In storm-damaged sites we can expect differences in vegetation structure between the recording periods of the bird survey data (before the cyclone) and the LiDAR data (after the cyclone). Hence, we removed all storm-damaged sites ($n = 10$) from the analyses to avoid mismatches in vegetation structure.

After intersecting species and environmental data, we retained 2,535 1 × 1 km cells with breeding bird records and 6,946 100 × 100 m cells with tree species occurrences. From these, we randomly selected 70% (1,774 cells for birds and 4,862 for tree species) for model building and 30% (761 cells for birds and 2,084 cells for tree species) for validation of the community predictions.

2.5 | Model building

We estimated three different kinds of models that served as input for the community predictions (Figure 1). SDMs and JSDMs were estimated to predict species-specific occurrence probabilities per site. MEMs were estimated to predict species richness per site, which then served as macroecological constraint for the application of ecological assembly rules.

Prior to model building, all predictor variables were standardized. In each model, we only included the five most important and weakly correlated variables. To do so, we first assessed cross-validated univariate variable importance for each predictor in a fivefold spatial block cross-validation design. Therefore, for each dataset we split the study region into five rectangular tiles (Figure S2; R package *sperrorest*). The resulting sample sizes per tile ranged 222–568 for forest birds and 570–1,440 for tree species. We then estimated univariate GAMs on four of fivefolds (with logit link for species occurrences and log link for species richness; R package *mgcv*) and cross-predicted to the left-out fold with five repetitions. From the cross-predictions, we calculated the percentage of explained deviance. Then, we reduced the predictor set to variables with bivariate Spearman correlations $|r| < 0.7$, retaining those variables from highly correlated pairs with higher cross-validated univariate importance (Dormann et al., 2013). We thus obtained a predictor ranking for each single species. However, JSDMs require a global set of predictor variables and we thus selected those five variables with highest mean cross-validated univariate importance among all species. SDMs and MEMs were run with the same set of global predictors as in JSDMs. In a sensitivity analysis we additionally estimated SDMs and MEMs using the five most important variables selected individually for each species and species richness, respectively. This did not change the overall results.

We fitted SDMs and JSDMs to single species occurrences and MEMs to species richness counts (Figure 1). Following current standards (Araújo et al., 2019), SDMs and MEMs were fitted using an ensemble approach with four different algorithms: generalized linear models (GLM), generalized additive models (GAM), boosted regression trees (BRT) and random forests (RF) with a binomial error distribution (with logit link) for SDMs (with R packages *mgcv*, *dismo*, *randomForest*) and a Poisson error distribution (with log link) for MEMs. Specifically, GLMs were fitted with linear and quadratic terms and GAMs were fitted with nonparametric cubic smoothing splines with up to four degrees of freedom. BRTs were estimated with a tree complexity of 2, a bag fraction of 0.75 and a variable learning rate such that 1,000–5,000 trees were fitted (Elith, Leathwick, & Hastie, 2008). Random forests were fitted with 1,000 trees, and a minimum node size of 20. JSDMs were fitted with a latent variable model and

a binomial error distribution (with probit link, with R package *boral*; Hui, 2016). In all JSDMs we included linear and quadratic terms and five latent variables. For forest bird species, JSDMs were run with 50,000 iterations, a burnin of 20,000 and a thinning rate of 50. For tree species, JSDMs were run with 100,000 iterations, a burnin of 50,000 and a thinning rate of 50. Convergence was assessed using the Geweke convergence diagnostic.

For all models (SDMs, MEMs, JSDMs), predictive model performance was assessed using a 5-fold spatial block cross-validation (with same spatial tiles as in univariate models, Figure S2) and four different performance measures were estimated: area under the receiver operating characteristic curve (AUC), true skill statistic (TSS), sensitivity and specificity. The latter three measures constitute threshold-dependent performance measures and we calculated them using a TSS-maximization threshold. We compared cross-validated single-species model performance for SDMs and JSDMs and tested whether JSDMs perform better for rare species. Spatial autocorrelation in model residuals was assessed using spline correlograms in the R package *ncf* (Bjornstad, 2016).

2.6 | Community predictions

In order to independently validate the community-level predictions, all models were predicted to the 30% hold-out data that were not used for model building. Consensus predictions from SDMs and MEMs were generated using un-weighted ensemble means (in a sensitivity analysis, we also tested community-level predictions when only GLMs were used to derive SDM predictions). JSDM predictions were derived through marginalizing over the latent variables in the model, which allows making prediction to new sites (Hui, 2016).

Both SDM and JSDM predictions were processed in four different ways to obtain species assemblage predictions: two stacking procedures (binary vs. probabilistic), two applications of SESAM (without and with bias correction; Calabrese et al., 2014) using the probability ranking as ecological assembly rule (D'Amen, Pradervand, et al., 2015) and the species richness predictions from MEMs as site-level constraint (Figure 1). First, binary stacks were derived by converting the species-specific occurrence probabilities from (the consensus predictions of) SDMs and from JSDMs into presence–absence predictions using a TSS-maximization threshold (maximized separately for each species, and for SDMs and JSDMs, respectively) and then stacking the predicted presences. Second, for assessing the accuracy of species assemblage predictions, probabilistic stacks were derived through randomly drawing from binomial trials with the site- and species-level occurrence probabilities as success rates and then stacking the random presences. Random number generation was repeated 100 times, leading to a set of 100 probabilistic stacks. For assessing the accuracy of species richness predictions, we simply summed the site-level occurrence probabilities of all species. Third, the probability ranking rule requires to rank the species in each site according to the occurrence probabilities predicted from SDMs and JSDMs and then only the most probable species are retained in the community prediction with the maximum number of species per

site predicted from MEM. Fourth, we used the maximum-likelihood approach proposed by Calabrese et al. (2014) to adjust the site-level occurrence probabilities such that overprediction and underprediction bias in species richness estimates is reduced and then repeated the probability ranking rule for these bias-corrected occurrence probabilities. Specifically, we estimated the adjustment parameters on the training dataset during model building and then used these parameters together with predicted species richness from MEMs to correct the occurrence probabilities in the validation dataset. After that, the adjusted site-level occurrence probabilities were summed up to obtain adjusted species richness estimates as macroecological constraint and the probability ranking was applied to the adjusted site-level occurrence probabilities. The different model types (SDMs vs. JSDMs), stacking procedures (binary vs. probabilistic) and SESAM applications (probability ranking without and with bias correction of macroecological constraint) resulted in eight different sets of community predictions (Figure 1).

All community predictions were evaluated in terms of accuracy of site-level assemblage prediction and species richness error. Accuracy of assemblage predictions was derived from site-level confusion matrices that classify species into (a) true species presence (observed and predicted), (b) false species presence (not observed but predicted), (c) false species absence (observed but not predicted) and (d) true species absence (not observed and not predicted). From the confusion matrices, we then calculated assemblage sensitivity as $a/(a + c)$, assemblage specificity as $d/(b + d)$, assemblage TSS (true skill statistic) as assemblage sensitivity + assemblage specificity - 1 and assemblage prediction success as $a + d/N$ with N being the total number of species in the local species pool (Pottier et al., 2013). Site-level species richness errors were calculated as the absolute differences between observed and predicted species richness per site divided by N .

We used mixed-effect ANOVAs to test for significant differences in species assemblage and richness predictions between different model choices; specifically, differences between model types (SDMs vs. JSDMs), between stacking procedure (binary vs. probabilistic), between stacking and ecological assembly rule (stacking vs. probability ranking), and between probability ranking without and with bias correction of the macroecological constraint. We corrected for repeated measures by using site as random effect.

All analyses were conducted using R version 3.3.2 (R Core Team, 2016) with packages *sperrorest* (Brenning, 2012), *mgcv* (Wood, 2011), *gbm* (Ridgeway, 2013), *dismo* (Hijmans, Phillips, Leathwick, & Elith, 2017), *randomForest* (Liaw & Wiener, 2002), *boral* (Hui, 2016), *ncf* (Bjornstad, 2016), *PresenceAbsence* (Freeman & Moisen, 2008), *ecospat* (Broennimann, Cola, & Guisan, 2016) and *lme4* (Bates, Maechler, Bolker, & Walker, 2015).

3 | RESULTS

Climate variables showed highest cross-validated univariate importance for most species and, thus, the globally selected predictor

sets contained mainly climate variables (Figure S3). In addition to the climate variables, for both forest birds and tree species, also one LiDAR variable describing the vertical vegetation structure was selected (Figure S3). Spatial block cross-validation showed comparable single species prediction accuracy between JSDMs and GLMs while the ensemble SDMs showed higher accuracy. Mean predictive performance was fair to excellent (Figure S4, Tables S1 and S2). Spatial autocorrelation in model residuals was apparent at relatively short distances (forest birds mean \pm SD: 5,971 m \pm 6,267 m in SDMs, 10,050 m \pm 10,008 m in JSDMs; tree species mean \pm SD: 8,853 m \pm 10,676 m in SDMs, 13,150 m \pm 13,123 m in JSDMs; Figure S5). These distances are short compared to the raw data (forest birds mean \pm SD: 42,687 m \pm 28,451 m; tree species mean \pm SD: 39,410 m \pm 27,018 m; Figure S5). Overall, we judge the two datasets as suitable for the S-SDM and JSDM comparison.

The comparison of single-species performance in SDMs and JSDMs showed that ensemble SDMs generally outperformed JSDMs (JSDMs for birds, mean \pm SD: AUC 0.77 \pm 0.11, TSS 0.44 \pm 0.18; ensemble SDMs for birds: AUC 0.88 \pm 0.04, TSS 0.62 \pm 0.11; JSDMs for trees: AUC 0.78 \pm 0.09, TSS 0.39 \pm 0.20; ensemble SDMs for trees: AUC 0.93 \pm 0.04, TSS 0.75 \pm 0.14). When comparing JSDMs against single-species GLMs, which are comparable in the chosen complexity of the environmental responses, JSDMs performed better in terms of specificity (Tables S1 and S2, Figure S4). Also we found that JSDMs tended to outperform SDMs in terms of specificity mainly for rarer species (Figure S6).

Joint species distribution models indicated that environment is the main driving factor of bird and tree species distribution in Switzerland. Overall, environmental predictors (in particular climate, cf. Figure S2) accounted for 70% and 71% of covariation between bird species and tree species, respectively (calculated by comparing the JSDMs with pure latent variable models, cf. Hui, 2016). We clustered species according to the similarity in their environmental response (using hierarchical clustering with average linkage strategy), which revealed two and three distinct species groups for birds and trees, respectively, that correspond to lowland, montane and alpine species (Figures 2 and 3). Forest birds showed mainly positive residual correlations in montane regions while negative residual correlations were only apparent in lowland species (Figure 2). By contrast, trees also exhibited negative residual correlation in alpine species (Figure 3). Additionally, we found strong negative residual correlations between alpine and lowland tree species while such patterns were less apparent for forest birds.

Deviation between observed species richness and species richness predictions derived from stacking SDMs and JSDMs were largely consistent between birds and trees (Figure 4). Binary stacking led to overestimation of species richness, which was more severe in JSDMs than in SDMs. Probabilistic stacking as well as the probability ranking rule led to overprediction of low species richness and underprediction of high species richness. Underprediction of high species richness sites was stronger for trees compared to birds. Bias correction considerably reduced the overprediction and underprediction in species richness predictions of forest birds but not of trees.

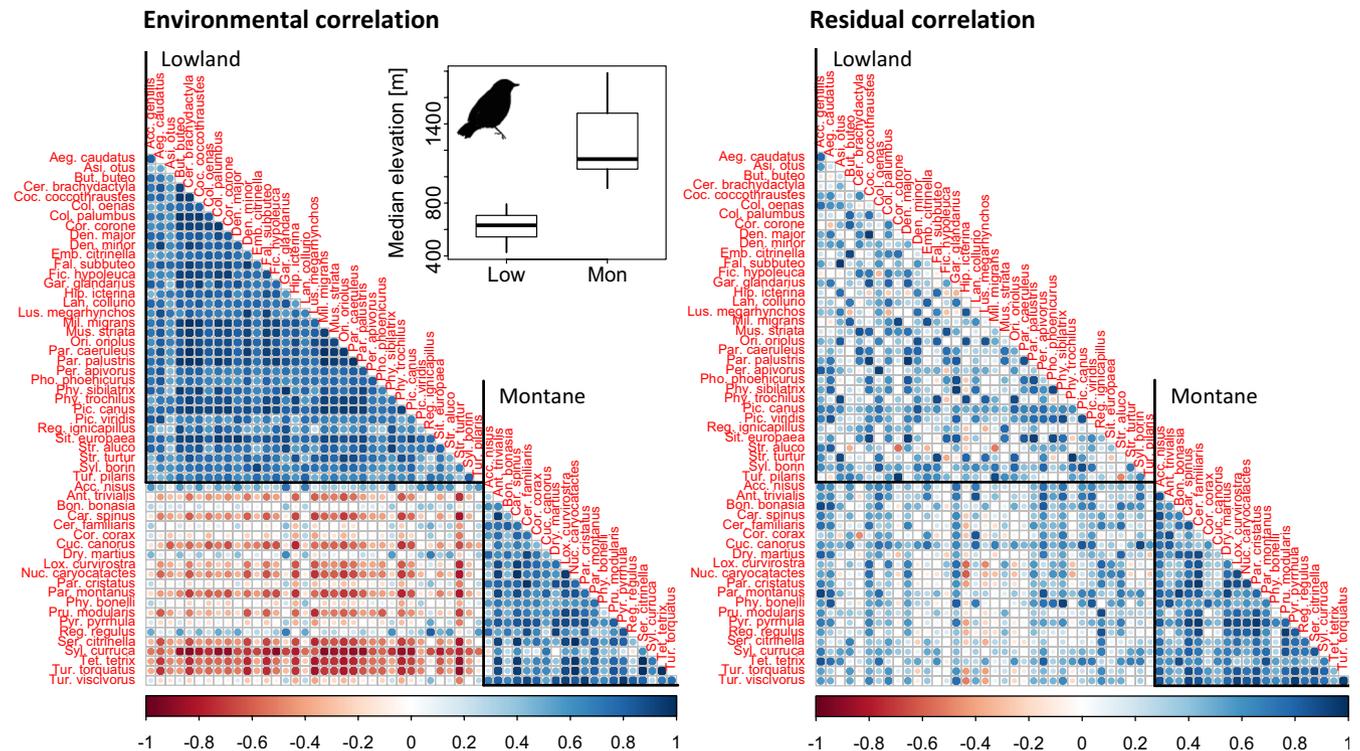


FIGURE 2 Joint species distribution model (JSDM) estimates of environmental correlation (left) and residual correlation among bird species (right). Species were ordered and groups identified by hierarchical clustering of their environmental correlation using an average linking strategy. For comparison, the same order was used for depicting residual correlation. The boxplots summarize the median elevation of each bird species' occurrence records for the different species groups

The accuracy of community assemblage predictions varied considerably between sites ranging from poor to outstanding (Figure 5). Average assemblage TSS values over all sites were 0.50 ± 0.17 for birds and 0.46 ± 0.22 for trees. True absences of species in a site (assemblage specificity) were much better predicted than true presences of species (assemblage sensitivity; Figure S7). On average, birds showed slightly higher assemblage sensitivities than trees while the opposite was true for assemblage specificity. Assemblage prediction success was high with an average of $78 \pm 8\%$ correctly predicted bird species and $87 \pm 10\%$ correctly predicted tree species per site (Figure 5).

Mixed-effect ANOVAs were used to test for significant differences between different modelling choices. Generally, assemblage TSS and assemblage prediction success from stacked SDMs were significantly higher than from stacked JSDMs and deviation in site-level species richness was significantly lower (Figure 5). Probabilistic stacking always outperformed binary stacking in terms of assemblage TSS, assemblage prediction success, assemblage sensitivity and richness deviation. By contrast, it led to lower assemblage specificity (Figure S7). Applying probability ranking as ecological assembly rule yielded significantly better assemblage and richness predictions than either binary or probabilistic stacking ("EAR" in Figure 5). Correcting for species richness bias prior to probability ranking often improved species assemblage predictions and species richness predictions, although the absolute differences cannot be judged as ecologically relevant ("corr" in Figure 5).

4 | DISCUSSION

In this study, we tested the ability of stacked SDMs and JSDMs to predict species assemblages. Our results indicate that although JSDMs show much promise for bridging species distribution modelling and community ecology, they do not necessarily outperform stacked SDMs in predicting site-level species composition and species richness. Rather, the choice of stacking procedure and ecological assembly rules is more important. We thus conclude that JSDMs do not necessarily improve species assemblage and species richness predictions, but yield additional insights about community assembly, which should be further tested for other taxa, ecosystems and JSDM algorithms and possibly also for other questions.

Interestingly, at the species level ensemble SDMs clearly outperformed JSDMs while GLMs, whose response shapes are directly comparable to the parametric JSDMs used here, showed rather similar performance to JSDMs. One important aspect why JSDMs did not improve (or fared worse in terms of) prediction accuracy could be the fact that JSDM predictions to new sites (here, hold-out data) can only be done by marginalizing (i.e., averaging) over the latent variables. For interpolation, JSDMs could potentially achieve higher prediction accuracy compared to SDMs because in such case JSDM predictions of one species can be made conditional on the presence of other species. For extrapolation, however, our results indicate that JSDMs do not improve species-level predictions.

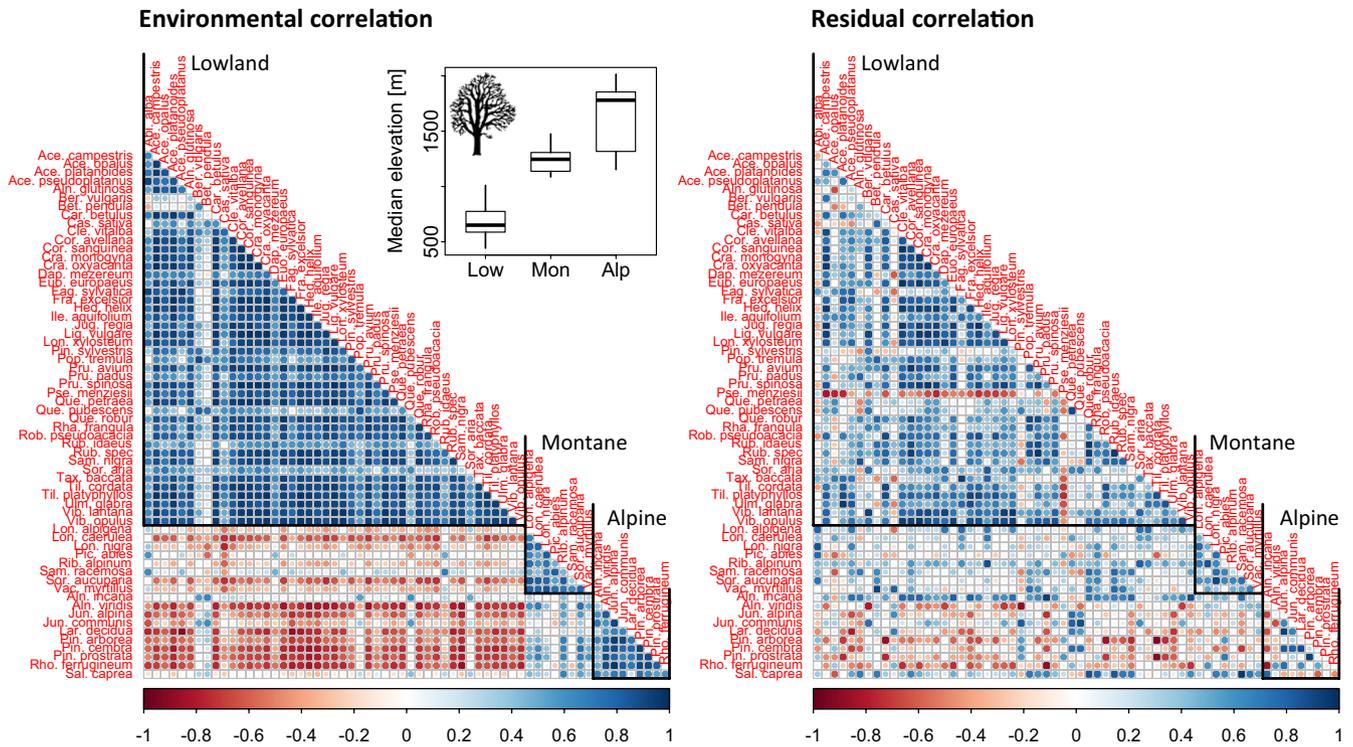


FIGURE 3 Joint species distribution model (JSDM) estimates of environmental correlation (left) and residual correlation (right) among tree species. Species were ordered and groups identified by hierarchical clustering of their environmental correlation using an average linking strategy. For ease of comparison, the same order was used for depicting residual correlation. The boxplots summarize the median elevation of each tree species' occurrence records for the different species groups

It has been suggested that rare species should profit from joint modelling as they can “borrow strength” from estimates of the more common species (Ovaskainen & Soininen, 2011; Pollock et al., 2014). However, our results did not corroborate this proposition and we hardly observed that rare species benefitted from JSDMs. We only found a tendency that JSDMs improved the true absence rate (species-level specificity) in rare species, especially so in tree species. At the same time, SDMs tended to improve the true presence rate (species-level sensitivity) in rare species while results for AUC and TSS were equivocal. Sensitivity analysis showed that these results hold when JSDMs were compared against ensemble SDMs as well as GLMs. Our results do thus not support the idea that JSDMs may generally yield improved predictions for rare species. However, we did only include species with at least 50 presences because SDMs tend to be unreliable for smaller sample sizes (Breiner et al., 2015). In the future, it could be explicitly tested how accurately JSDMs are predicting rare species in comparison to other approaches such as the ensemble of small models (Breiner et al., 2015).

Our sensitivity analysis also indicated a strong effect of model complexity on predictive performance. JSDMs and GLMs only included linear and quadratic terms and thus were bound to yield smooth response surfaces. The ensemble SDMs, on the other hand, also included complex machine-learning approaches like random forests and boosted regression trees that produce much more complex response surfaces (Elith et al., 2006). In our analyses, these ensembles typically yielded improved predictions at the species level. However,

very complex models are not always desirable as this could limit the transferability when extrapolating beyond the sampled environmental conditions (Merow et al., 2014; Wüest, Münkemüller, Lavergne, Pollock, & Thuiller, 2018; Yates et al., 2018). In the future, JSDMs of different complexities (Clark, Nemergut, Seyedsrollah, Turner, & Zhang, 2017; Harris, 2016; Hui, 2016; Ovaskainen et al., 2017) should be compared to investigate how response surface complexity affects the ability of JSDMs to predict community composition.

An advantage of JSDMs compared to SDMs is that they are able to disentangle environmental response from residual correlations among species that could be indicative of biotic interactions between these species or of missing and subscale environmental variability (Dormann et al., 2018; Pollock et al., 2014; Zurell et al., 2018). Our models indicated that environment, in particular climate, is the main driving factor of both bird and tree species distributions in Switzerland. Based on the environmental correlations, we were able to separate species into lowland and montane to alpine species. In forests birds, we found negative residual correlations mainly between lowland species. This could corroborate the stress gradient hypothesis that competition is most severe for low environmental stress (Meier, Edwards Jr, Kienast, Dobbertin, & Zimmermann, 2011). However, for tree species the pattern looked slightly different with strong negative residual correlations also among alpine species. Especially, *Alnus incana*, *Alnus viridis* and *Salix caprea* exhibited strong negative residual correlations with other alpine species. All three of these species are pioneers and thus

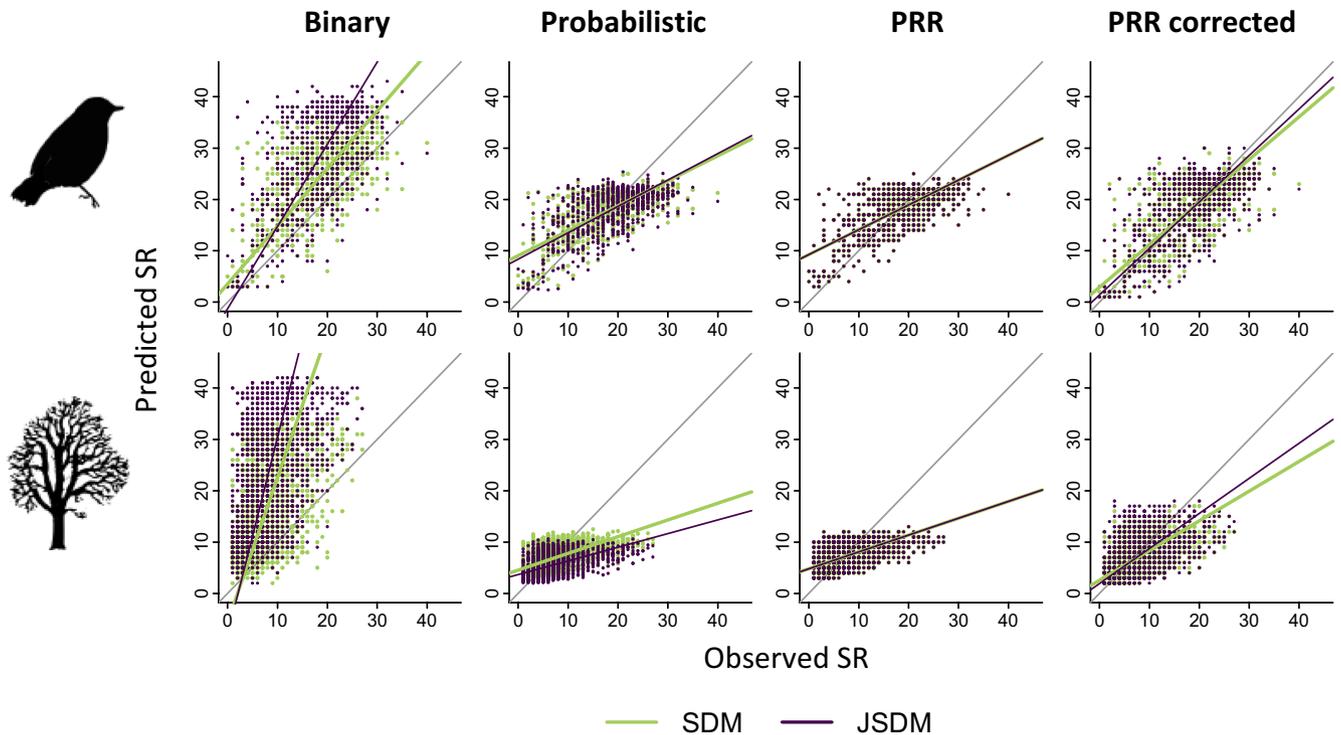


FIGURE 4 Observed versus predicted species richness for different underlying model types (species distribution models [SDMs] vs. joint species distribution models [JSDMs]), different stacking procedures (binary vs. probabilistic), and for the application of ecological assembly rules (PRR, probability ranking rule) without and with bias correction of the macroecological constraints. Coloured lines indicate the major axis regression of observed versus predicted species richness. Grey lines represent the 1:1 line

mainly occur where and when other tree species are not present, which could explain these negative correlations. The positive residual correlations of montane to alpine species, especially in forest birds, could be indicative of facilitative interactions. This is in line with results of Gallien, Zurell, and Zimmermann (2018), who found stronger facilitative interactions in alpine than in lowland species, albeit for vascular plants. We cannot rule out that the positive residual correlations at least partly indicate missing environmental variables or subscale environmental heterogeneity. Interspecific interactions refer to local processes between individuals while here we model species distributions at coarser spatial resolution (100 m for trees and 1 km for forest birds). Previous studies have shown that the signal of competitive interactions is more easily lost at increasingly coarser resolution than the signal of facilitative interactions (Araújo & Rozenfeld, 2014; Zurell et al., 2018). This could pose an additional explanation why we mainly found positive residual correlations, especially in the coarser-grained bird models.

Similar to the species level analyses, JSDMs were not able to outperform S-SDMs in predicting community level properties such as site-level species assemblage and species richness. Specifically, for all performance measures considered, JSDMs produced significantly less accurate community level predictions. However, the absolute differences were so small that they cannot be judged as ecologically relevant. For both SDMs and JSDMs, choices have to be made on how species and site-level occurrence probabilities are translated into predictions of species assemblages and species richness.

Generally, probabilistic stacking yielded more accurate species assemblage and species richness predictions than binary stacking, and ecological assembly rules yielded more accurate predictions than probabilistic stacking. The correction of species richness bias suggested by Calabrese et al. (2014) only marginally improved predictions based on ecological assembly rules. This surprisingly small effect of the bias correction is most likely due to the fact that predictions were made to independent (hold-out) data. In this case, the species richness bias could not be fully corrected because the site-level adjustment of the predictions had to be based on (inherently biased) species richness predictions (from MEMs).

None of the tested approaches for making species assemblage predictions was capable of completely removing the over- and underprediction biases in species richness predictions. Notably, the underprediction of high species richness for probabilistic stacks and probability ranking was much more severe in the tree than in the bird species. These differences are probably related to lower overall prevalence in the tree species (Figure S1), which has been shown to increase underprediction of species richness (Zurell et al., 2016).

Overall, we conclude that for community-level predictions the choice of how species and site-level occurrence probabilities are combined into species assemblage predictions is of higher importance than the choice of the model type used, SDMs or JSDMs. Highest site-level prediction accuracy can be achieved when constructing species assemblages from SDM derived occurrence probabilities using the probability ranking rule and direct

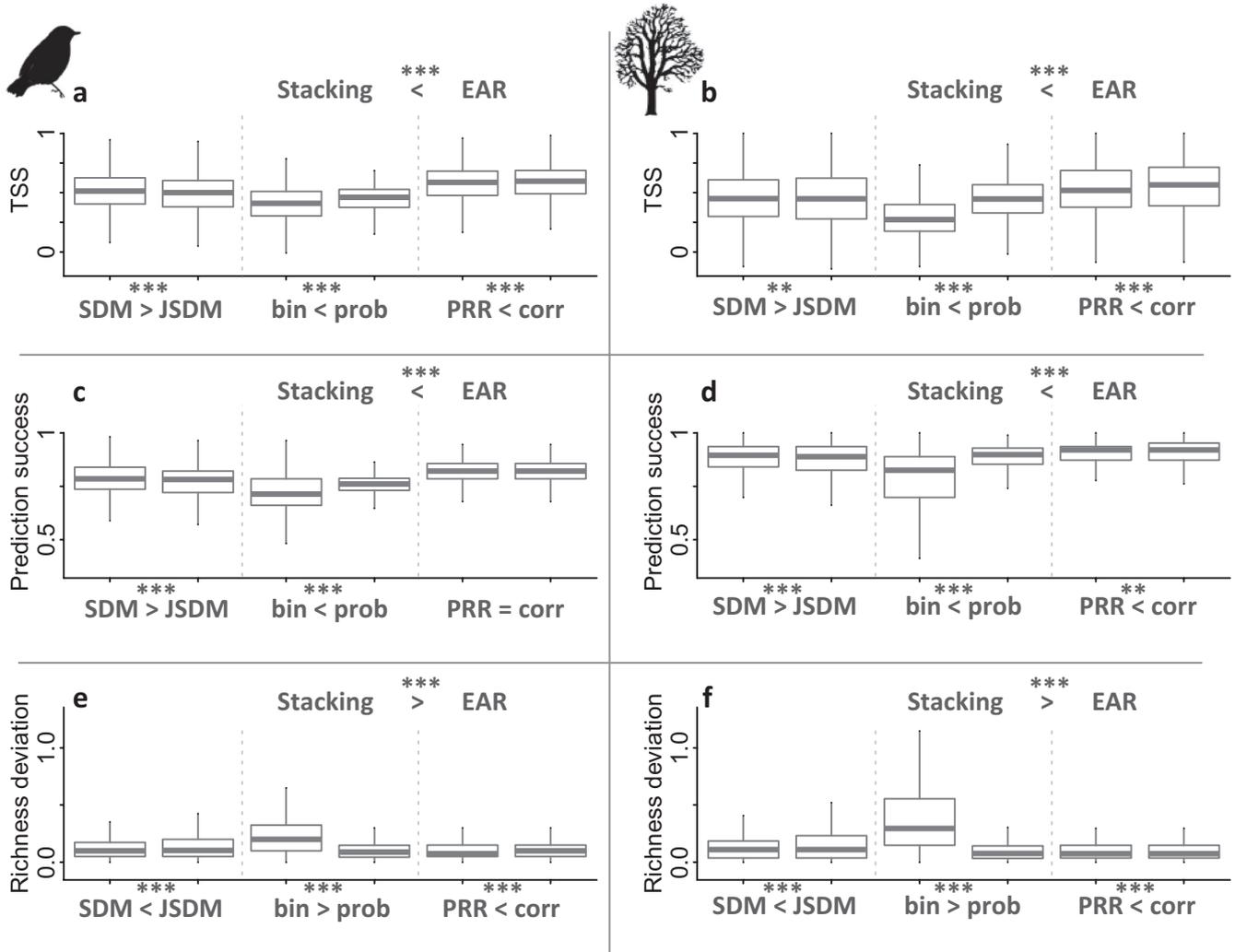


FIGURE 5 Accuracy of community assemblage predictions (a–d) and deviation of predicted from observed community richness predictions (e–f). Boxplots summarize all site-level assemblage (a–d) and richness predictions (e–f). Results are separated for different underlying model types (species distribution models [SDMs] vs. joint species distribution models [JSDMs]), different stacking procedures (binary vs. probabilistic), and for the application of probability ranking as ecological assembly rule (EAR) without and with bias correction of the macroecological constraints. (a–b) TSS (true skill statistic) and (c–d) prediction success represent the accuracy of the assemblage prediction. (e–f) Richness deviation represents the absolute deviation in richness prediction divided by the maximum observed site-level species richness. Significant differences between model choices are evaluated using mixed-effect ANOVAs with site as random effect (significance levels: *** $p < 0.001$, ** $p < 0.01$)

species richness predictions from MEMs as macroecological constraint. Using JSDMs instead of SDMs or correcting for biases in the species richness predictions will only have minor effects on the community-level predictions. Future studies should evaluate these propositions across more taxa and biogeographical realms as well as for other JSDM algorithms (e.g., Harris, 2015; Clark et al., 2017; Ovaskainen et al., 2017). The promise of JSDMs to improve community predictions by incorporating co-occurrence patterns into species distribution models did not hold in our case studies. In fact, JSDMs did not even improve single-species predictions consistently. Specific aspects of JSDMs, such as conditioning predictions on the presence/absence of other species need more attention and may further our ability to successfully predict the composition of species assemblages. Clearly, JSDMs

can help deriving hypotheses about community assembly processes present in the system and could, thus, serve as screening tool for identifying important biotic interactions in local species pools (Ovaskainen et al., 2017; Zurell, 2017). These relative benefits of SDMs and JSDMs in community predictions and hypothesis testing should be further evaluated in the future.

ACKNOWLEDGEMENTS

We thank the dedicated and highly qualified volunteers who conduct the fieldwork of the Swiss breeding bird atlas. DZ received funding from the Swiss National Science Foundation (SNF, grant: PZ00P3_168136/1) and from the German Science Foundation (DFG, grant: ZU 361/1-1). NEZ and ROW acknowledge support from SNF

(grant: 310030L_170059). ROW and AB acknowledge funding from the Swiss National Forest Inventory run by the Swiss Federal Office for the Environment FOEN and the Swiss Federal Research Institute WSL.

DATA ACCESSIBILITY

The species data were provided by courtesy of the Swiss Ornithological Institute and the Swiss Forest Inventory. The presence-absence data (with geographical coordinates removed) and the environmental predictor data are available on Dryad (<https://doi.org/10.5061/dryad.k88v330>).

ORCID

Damaris Zurell  <https://orcid.org/0000-0002-4628-3558>

Niklaus E. Zimmermann  <https://orcid.org/0000-0003-3099-9604>

Rafael O. Wüest  <https://orcid.org/0000-0001-6047-1945>

REFERENCES

- Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., ... Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5, eaat4858. <https://doi.org/10.1126/sciadv.aat4858>
- Araújo, M. B., & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography*, 37, 406–415.
- Artuso, R., Bovet, S., & Streilein, A. (2003). Practical methods for the verification of countrywide terrain and surface models. *International Archives of Photogrammetry and Remote Sensing*, XXXIV, Part 3/ W13.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Bjornstad, O. N. (2016). ncf: Spatial nonparametric covariance functions. Retrieved from <https://CRAN.R-project.org/package=ncf>
- Brassel, P., & Lischke, H. (Eds.) (2001). *Swiss national forest inventory: Methods and models of the second assessment*. Birmensdorf, Switzerland: Swiss Federal Research Institute WSL.
- Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, 6, 1210–1218. <https://doi.org/10.1111/2041-210X.12403>
- Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package 'sperrorest'. IEEE International Symposium on Geoscience and Remote Sensing IGARSS, <https://doi.org/10.1109/igarss.2012.6352393>
- Broennimann, O., Cola, V. D., & Guisan, A. (2016). ecospat: Spatial ecology miscellaneous methods. R Package Version 2.1.1. Retrieved from <https://CRAN.R-project.org/package=ecospat>
- Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, 23, 99–112. <https://doi.org/10.1111/geb.12102>
- Clark, J. S., Nemergut, D., Seyedsnasrollah, B., Turner, P. J., & Zhang, S. (2017). Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecological Monographs*, 87, 34–56. <https://doi.org/10.1002/ecm.1241>
- D'Amen, M., Rahbek, C., Zimmermann, N. E., & Guisan, A. (2017). Spatial predictions at the community level: From current approaches to future frameworks. *Biological Reviews*, 92, 169–187. <https://doi.org/10.1111/brv.12222>
- D'Amen, M., Dubuis, A., Fernandes, R. F., Pottier, J., Pellissier, L., & Guisan, A. (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, 42, 1255–1266. <https://doi.org/10.1111/jbi.12485>
- D'Amen, M., Pradervand, J.-N., & Guisan, A. (2015). Predicting richness and composition in mountain insect communities at high resolution: A new test of the SESAM framework. *Global Ecology and Biogeography*, 24, 1443–1453. <https://doi.org/10.1111/geb.12357>
- Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., ... Kraan, C. (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, 27, 1004–1016. <https://doi.org/10.1111/geb.12759>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.-P., & Guisan, A. (2011). Predicting spatial patterns of plant species richness: A comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, 17, 1122–1131. <https://doi.org/10.1111/j.1472-4642.2011.00792.x>
- Dullinger, I., Wessely, J., Bossdorf, O., Dawson, W., Essl, F., Gattringer, A., ... Dullinger, S. (2017). Climate change will increase the naturalization risk from garden plants in Europe. *Global Ecology and Biogeography*, 26, 43–53. <https://doi.org/10.1111/geb.12512>
- Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distribution from occurrence data. *Ecography*, 29, 129–151.
- 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, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43, 393–404. <https://doi.org/10.1111/j.1365-2664.2006.01149.x>
- Freeman, E. A., & Moisen, G. (2008). PresenceAbsence: An R package for presence absence analysis. *Journal of Statistical Software*, 23, 1–31.
- Gallien, L., Zurell, D., & Zimmermann, N. E. (2018). Frequency and intensity of facilitation reveal opposing patterns along a stress gradient. *Ecology and Evolution*, 8, 2171–2181. <https://doi.org/10.1002/ece3.3855>
- Götzenberger, L., de Bello, F., Bräthen, K. A., Davison, J., Dubuis, A., Guisan, A., ... Zobel, M. (2012). Ecological assembly rules in plant communities - Approaches, patterns and prospects. *Biological Reviews*, 87, 111–127. <https://doi.org/10.1111/j.1469-185X.2011.00187.x>
- Guisan, A., & Rahbek, C. (2011). SESAM - A new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, 38, 1433–1444. <https://doi.org/10.1111/j.1365-2699.2011.02550.x>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8, 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)

- Harris, D. J. (2015). Generating realistic assemblages with a joint species distribution model. *Methods in Ecology and Evolution*, 6, 465–473. <https://doi.org/10.1111/2041-210X.12332>
- Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, 97, 3308–3314. <https://doi.org/10.1002/ecy.1605>
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). *dismo*: Species distribution modeling. R package version 1.1-4. Retrieved from <https://CRAN.R-project.org/package=dismo>
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., & Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution and Systematics*, 43, 227–248. <https://doi.org/10.1146/annurev-ecolsys-110411-160411>
- Hofierka, J., Suri, M., & Šúri, M. (2002). The solar radiation model for Open source GIS: Implementation and applications. Proceedings of the Open source GIS - GRASS users conference, pp. 1–19.
- Hui, F. K. C. (2016). *boral* – Bayesian ordination and regression analysis of multivariate abundance data in R. *Methods in Ecology and Evolution*, 7, 744–750.
- Isenburg, M. (2015). *LAStools* – Efficient tools for LiDAR processing, version 150605.
- Kéry, M., & Schmid, H. (2006). Estimating species richness: Calibrating a large avian monitoring programme. *Journal of Applied Ecology*, 43, 101–110. <https://doi.org/10.1111/j.1365-2664.2005.01111.x>
- Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., McInerney, G., ... O'Hara, R. B. (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, 39, 2163–2178. <https://doi.org/10.1111/j.1365-2699.2011.02663.x>
- Kremen, C., Cameron, A., Moilanen, A., Phillips, S. J., Thomas, C. D., Beentje, H., ... Zihra, M. L. (2008). Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, 320, 222–226. <https://doi.org/10.1126/science.1155193>
- Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, 2, 18–22.
- Makkink, G. F. (1957). Testing the Penman formula by means of lysimeters. *Journal of the Institution of Water Engineers*, 11, 277–288.
- McGarigal, K., Cushman, S., & Ene, E. (2012). *FRAGSTATS v4*: Spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts.
- Meier, E. S., Edwards Jr, T. C., Kienast, F., Dobbertin, M., & Zimmermann, N. E. (2011). Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. *Journal of Biogeography*, 38, 371–382. <https://doi.org/10.1111/j.1365-2699.2010.02405.x>
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., ... Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37, 1267–1281. <https://doi.org/10.1111/ecog.00845>
- Ovaskainen, O., & Soininen, J. (2011). Making more out of sparse data: Hierarchical modeling of species communities. *Ecology*, 92, 289–295. <https://doi.org/10.1890/10-1251.1>
- Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., ... Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, 20, 561–576. <https://doi.org/10.1111/ele.12757>
- Pollock, L. J., Thuiller, W., & Jetz, W. (2017). Large conservation gains possible for global biodiversity facets. *Nature*, 546, 141–144. <https://doi.org/10.1038/nature22368>
- Pollock, L. J., Tingley, R., Norris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., ... McCarthy, M. A. (2014). Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods in Ecology and Evolution*, 5, 397–406. <https://doi.org/10.1111/2041-210X.12180>
- Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C. F., ... Guisan, A. (2013). The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, 22, 52–63. <https://doi.org/10.1111/j.1466-8238.2012.00790.x>
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ridgeway, G. (2013). *gbm*: Generalized boosted regression models. R Package Version 2.1.3. Retrieved from <https://CRAN.R-project.org/package=gbm>
- Sattler, T., Obrist, M. K., Duelli, P., & Moretti, M. (2011). Urban arthropod communities: Added value or just a blend of surrounding biodiversity? *Landscape and Urban Planning*, 103, 347–361. <https://doi.org/10.1016/j.landurbplan.2011.08.008>
- Schmid, H., Luder, R., Naef-Daenzer, B., Graf, R., & Zbinden, N. (1998). *Schweizer Brutvogelatlas. Verbreitung der Brutvögel in der Schweiz und im Fürstentum Liechtenstein 1993–1996*. Sempach, Switzerland: Swiss Ornithological Institute.
- Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, 190, 214–251. [https://doi.org/10.1016/S0022-1694\(96\)03128-9](https://doi.org/10.1016/S0022-1694(96)03128-9)
- Thuiller, W., Pironon, S., Psomas, A., Barbet-Massin, M., Jiguet, F., Lavergne, S., ... Zimmermann, N. E. (2014). The European functional tree of bird life in the face of global change. *Nature Communications*, 5, 3118–3118. <https://doi.org/10.1038/ncomms4118>
- Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. C. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology & Evolution*, 30, 766–779. <https://doi.org/10.1016/j.tree.2015.09.007>
- Wilson, J. P., & Gallant, J. C. (Eds.) (2000). *Terrain analysis: Principles and applications*. New York, NY: John Wiley & Sons.
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., ... Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88, 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>
- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73, 3–36. <https://doi.org/10.1111/j.1467-9868.2010.00749.x>
- Wüest, R. O., Münkemüller, T., Lavergne, S., Pollock, L. J., & Thuiller, W. (2018). Integrating correlation between traits improves spatial predictions of plant functional composition. *Oikos*, 127, 472–481. <https://doi.org/10.1111/oik.04420>
- Xu, C. Y., & Singh, V. P. (2002). Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resources Management*, 16, 197–219.
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, 33, 790–802. <https://doi.org/10.1016/j.tree.2018.08.001>
- Zellweger, F., Baltensweiler, A., Ginzler, C., Roth, T., Braunisch, V., Bugmann, H., & Bollmann, K. (2016). Environmental predictors of species richness in forest landscapes: Abiotic factors versus vegetation structure. *Journal of Biogeography*, 43, 1080–1090. <https://doi.org/10.1111/jbi.12696>
- Zellweger, F., Braunisch, V., Baltensweiler, A., & Bollmann, K. (2013). Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. *Forest Ecology and Management*, 307, 303–312. <https://doi.org/10.1016/j.foreco.2013.07.023>



- Zimmermann, N. E., Edwards, T. C., Graham, C. H., Pearman, P. B., & Svenning, J.-C. (2010). New trends in species distribution modelling. *Ecography*, 33, 985–989. <https://doi.org/10.1111/j.1600-0587.2010.06953.x>
- Zurell, D. (2017). Integrating demography, dispersal and interspecific interactions into bird distribution models. *Journal of Avian Biology*, 48, 1505–1516. <https://doi.org/10.1111/jav.01225>
- Zurell, D., Pollock, L. J., & Thuiller, W. (2018). Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography*, 41, 1812–1819. <https://doi.org/10.1111/ecog.03315>
- Zurell, D., Zimmermann, N. E., Sattler, T., Nobis, M. P., & Schröder, B. (2016). Effects of functional traits on the prediction accuracy of species richness models. *Diversity and Distributions*, 22, 905–917. <https://doi.org/10.1111/ddi.12450>

BIOSKETCH

Damaris Zurell is broadly interested in macroecology and biodiversity modelling. Her research aims at improving biodiversity predictions under global change by taking into account the complex effects of demography, dispersal and interspecific interactions on niche and range dynamics of species.

Rafael O. Wüest explores the generation, preservation and future fate of biodiversity. He assesses community structure and assembly, models diversity along environmental gradients, and analyses how evolutionary, biogeographical and ecological processes contribute to shape biodiversity patterns across scales.

Author contributions: DZ and ROW conceived the general idea and designed the study with help of NEZ. ROW, AB and HG pre-processed the environmental data and helped with preliminary analyses. DZ performed the main analyses and designed the figures. DZ led the writing, with extensive contributions from all authors.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zurell D, Zimmermann NE, Gross H, Baltensweiler A, Sattler T, Wüest RO. Testing species assemblage predictions from stacked and joint species distribution models. *J Biogeogr.* 2020;47:101–113. <https://doi.org/10.1111/jbi.13608>