Insect Conservation and Diversity

Insect Conservation and Diversity (2021) **14,** 52–66 doi: 10.1111/icad.12444

Predicting the distributions of regional endemic dragonflies using a combined model approach

WADE A. BOYS,[1] ADAM M. SIEPIELSKI,[1] BRENDA D. SMITH,[2] MICHAEL A. PATTEN[2] and JASON T. BRIED[3] [1]Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA, [2]Oklahoma Biological Survey, University of Oklahoma, Norman, OK, USA and [3]Illinois Natural History Survey, University of Illinois at Urbana-Champaign, Champaign, IL, USA

Abstract. 1. Climate warming is predicted to have large effects on insects, yet several data shortfalls, including distributional information, impede effective conservation strategies.

2. Knowledge of species distributions is a critical component for assessing conservation need but is often lacking for endemic or rare taxa, especially invertebrates.

3. One approach to better inform this gap is by using species distribution modelling (SDM) to predict suitable habitat and guide field surveys.

4. Here, we combine the predictions of two machine learning algorithms, maximum entropy and Random Forest, to estimate the current and future distributions of two endemic dragonflies of the Ozark-Ouachita Interior Highlands region in the southcentral United States.

5. Current suitable areas predicted by both algorithms largely overlapped for each species, but different environmental variables were most important for predicting their distributions. Field validation of these models resulted in new detections for both species showing their utility in guiding subsequent field surveys.

6. Future projections under two climate change scenarios support maintaining current suitable areas as these are predicted to be strongholds for these species. Our results suggest that combining outputs of multiple species distribution models is a useful tool for better informing the distributions of geographically limited or rare species.

Key words. Anisoptera, aquatic insects, endemicity, interior highlands, species distribution modelling.

Introduction

Despite their abundance and diverse ecosystem roles, invertebrates are disproportionately understudied, even though they comprise the vast majority of animal diversity (Stork, 1988). In the freshwater realm, insects comprise around 60% of known animal species (Dijkstra *et al*., 2014). Yet, climate change is predicted to have substantial effects on insects. Indeed, insect biomass has already declined in various regions throughout the world, the impacts of which have cascaded throughout biological communities (Thomas, 2004; Hallmann *et al*., 2017; Lister & Garcia, 2018). Species richness and biomass of aquatic insects, especially within orders Ephemeroptera, Plecoptera, Trichoptera, and Odonata, have considerably declined over the past 40 years (Sánchez-Bayo & Wyckhuys, 2019). Nevertheless, a recent meta-analysis revealed an overall increase in aquatic insect abundance and biomass, with this trend becoming more positive in the past 30 years in Europe, Asia and North America (Klink *et al*., 2020). Such conflicting results may arise, in part, from several shortfalls of insect biodiversity information (taxonomic, distribution, abundance, etc.) that impede our ability to accurately predict how these taxa will respond to climate change and thus apply effective conservation strategies (Cardoso *et al*., 2011).

Organisms, especially insects, are predicted to exhibit temporal shifts in phenology and spatial shifts in geographic distributions as the climate continues to warm (Parmesan, 2006; Hassall, 2015). Many insect species have already rapidly shifted their ranges poleward as suitable habitat becomes available at higher latitudes (Hickling *et al*., 2005, 2006; Lancaster, 2016). Odonates, for example, appear to respond strongly to warming temperatures by shifting distributions poleward and advancing

Correspondence: Wade A. Boys, Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701. E-mail: wadeboys@email.uark.edu

their phenology (Hickling *et al.*, 2005; Hassall, 2015). Still, even more species are predicted to undergo range shifts and contractions under future climate scenarios (Domisch *et al.*, 2011, 2013a).

A species' range size is considered a strong predictor of extinction risk, thus it is important to understand the factors that shape a species' current distribution, as well as how its distribution may shift in response to climate change (Gaston & Fuller, 2009). For insect species, we often have limited occurrence data and poor knowledge of distributions at all geographic scales (Cardoso *et al.*, 2011). This has been coined the 'Wallacean shortfall' and poses challenges in developing and implementing effective conservation strategies which rely on robust distributional data to inform where efforts should be focused (Margules & Pressey, 2000; Cardoso *et al.*, 2011). Habitat specialists and regional endemics are of particular concern and may be the most vulnerable to warming as predictions frequently show large reductions in suitable habitat under various climate change scenarios (Domisch *et al.*, 2013a; Li *et al.*, 2014; Markovic *et al.*, 2014).

Using species distribution modelling (SDM) to predict distributions of rare or endemic species can be of great benefit but remains challenging due to the data demands of different algorithms (Papeş & Gaubert, 2007). Regional endemicity imposes a geographic constraint on extent of occurrence, and rarity lowers the occupancy rate or area occupied within that limited extent (Gaston, 1991; Rocha-Ortega *et al.*, 2020). Although geographically restricted species often show better model performance compared to widespread species that are found across many different habitat types (Tessarolo *et al.*, 2014; Proosdij *et al.*, 2016), models for species that are rare and regionally endemic may suffer (i.e. increased uncertainty around predictions) from small sample sizes (Gaston, 1991; Wisz *et al.*, 2008). Despite this limitation, two models in particular, maximum entropy (MaxEnt) (Phillips *et al.*, 2006) and Random Forest (Breiman, 2001), have consistently provided robust estimates for species with limited presence data (Hernandez *et al.*, 2006; Williams *et al.*, 2009; Mi *et al.*, 2017). Combining the predictions of these models may further help target surveys to locations with a high probability of presence as predicted by both algorithms (Tronstad *et al.*, 2018).

The use of SDM in conservation assessments has become more prevalent for estimating potential habitat, helping guide field surveys, and predicting changes in distributions through climate change scenarios (Guisan *et al.*, 2006, 2013; Papeş & Gaubert, 2007; Young *et al.*, 2019). An analysis of odonate diversity and overlap with current and planned conservation areas showed little protection for most species, especially those with small range sizes (Nóbrega & De Marco, 2011). This approach was used to model the distribution of a rare odonate in relation to deforestation of the Amazon forest, informing which areas were at the most risk of habitat destruction and extirpation (De Marco *et al.*, 2015). Conservation status, as assessed by the International Union for Conservation of Nature criteria, for various endemic damselflies has been updated following extent of occurrence estimations from SDM (Almeida *et al.*, 2010; Torres-cambas *et al.*, 2016; Rangel-Sánchez *et al.*, 2018). Furthermore, field surveys guided by predictions from SDM for rare taxa have shown promise in detecting new populations, including those of a rare stonefly (Young *et al.*, 2019) and cryptic crayfish (Rhoden *et al.*, 2017). Other studies have modelled distributions of odonates for conservation purposes, including many that have appeared since the review by Collins and McIntyre (2015), but there has been little attention to regionally endemic species of concern.

Here, we apply SDM for two rare and regional endemic species of dragonflies (Odonata: suborder Anisoptera) that are listed as Species of Greatest Conservation Need (SGCN; Bried & Mazzacano, 2010) *Gomphurus ozarkensis* (Westfall, 1975) (Ozark Clubtail) and *Somatochlora ozarkensis* (Bird, 1933) (Ozark Emerald). These species are highland endemics, mainly restricted to streams in the Ozark-Ouachita mountains spanning parts of Arkansas, Oklahoma, and Missouri, with sparse outlying occurrences in the Wichita mountains and Osage Hills of Oklahoma and the Flint Hills of Kansas (Figure 1). Both species are understudied and of conservation concern due to limited knowledge of their distributions, habitat requirements, and life histories (Patten & Smith-Patten, 2013; Smith & Patten, 2020). To better understand the environmental features associated with the distributions of these two species, we first employed the use of two machine-learning SDM algorithms shown to perform well with limited presence data and combined their predictions to help guide targeted field surveys. We then projected these distributions into the future under various climate change scenarios to assess where conservation efforts should be focused.

Materials and methods

We conducted field surveys and distribution modelling for the two focal species. Field surveys were aimed at checking previous record localities, finding new presence records, and verifying model predictions. The distribution modelling had two parts, one based on current distributions and one on future distributions, each requiring different environmental datasets and completed independently from each other. The current distribution modelling focused on local, stream-level environmental variables and the future distribution modelling on different climate change scenarios.

Field surveys

Surveys occurred in 2017 and 2018 during the known flight seasons: May to July for *G. ozarkensis* and June–September for *S. ozarkensis*. The 2017 surveys provided a baseline check of existing localities while trying to find new localities, and the 2018 surveys were guided by the current distribution modelling (described below). Sampling effort at each location involved two people (W.A. Boys and technician) searching for adults, exuviae, and larvae. Observations were not independent as experience in detecting and identifying species varied per individual.

Adult surveys were conducted when air temperatures reached 15.5 °C and above and not in heavy or steady rainfall. Starting from the site access point, both observers walked a 50 m stream section (one on each bank when accessible) back and forth for an hour. Observers searched for the focal species in flight and perching on nearby vegetation or other substrates. Adults were identified to species following capture with aerial nets. A

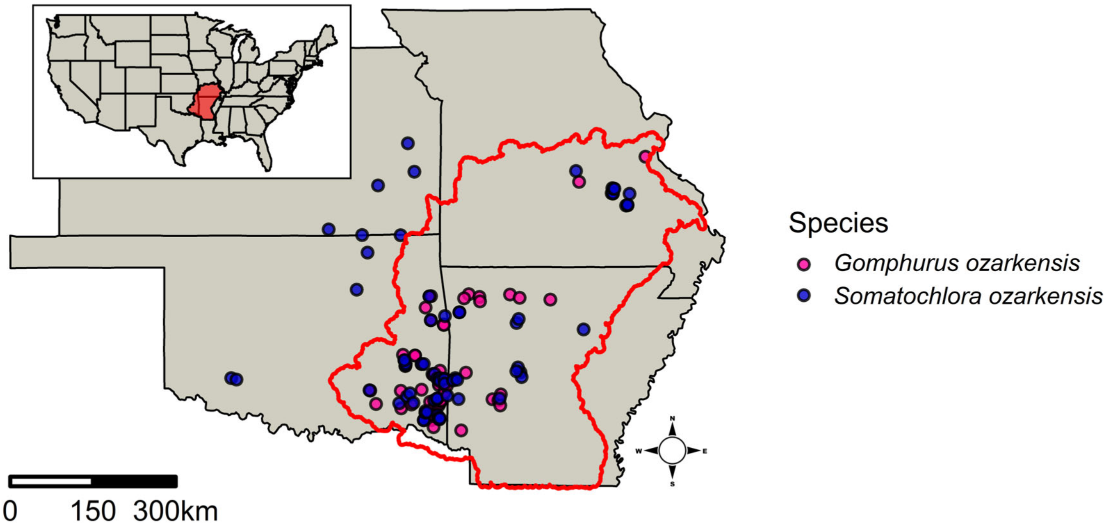

Figure 1. Presence localities used in current distribution models. Ozark-Ouachita Interior Highlands region outlined in red. [Color figure can be viewed at wileyonlinelibrary.com]

voucher specimen was collected and deposited at the University of Arkansas, although, because the focal species are considered SGCN, photographs of these individuals were taken in the field and served as vouchers in some instances. Observations of sex, mating pairs, tenerals, and oviposition behaviour were recorded, as they can help identify breeding sites (Patten *et al.*, 2019). Additionally, observers stayed alert for the focal species during hikes to and from the sites and while driving among sites; *S. ozarkensis* is known to congregate (swarm feeding) along roadways and other open areas at dawn and dusk.

Exuviae, or the exoskeleton of the final instar left behind after adults emerge, were searched for during the one-hour adult survey. Observers also opportunistically collected exuviae found later during larval surveys. All exposed substrate, including rocks, emergent vegetation, and sticks were searched for the presence of exuviae. If found, exuviae were collected for identification in the laboratory and the substrate they were attached to was noted.

Streams were searched for *S. ozarkensis* larvae under large instream rocks and cobbles, as well as along the banks in areas with exposed tree roots. *G. ozarkensis* larvae were searched for in mud, gravel, and detritus. All available habitat within the 50 m transect was sampled with an aquatic D-frame dip net with mesh size of 0.5 μm. All odonates were sorted out in the field and preserved in 70% ethanol for further identification in the lab.

Current distribution modelling

Our first modelling objective was to estimate the contemporary distributions of the focal species as a tool to guide field surveys and to get a better understanding of habitat requirements. For this objective, we used a stream-level approach aligned with the scale of environmental predictors.

Presence data were obtained from the online database OdonataCentral (Abbott, 2006) and complemented with unpublished data from a database compiled and maintained by the Oklahoma Biological Survey (Patten & Smith-Patten, 2014). OdonataCentral is a citizen science database and allows users to upload occurrences of adult odonates throughout the Western hemisphere, which are then vetted by regional experts. In addition, records from museum specimens and previous literature are included; however, we removed records with only county centroid coordinates unless specific location notes were included. We used these notes to georeference records in Google Earth (Google Inc., 2015) to obtain more precise geographic coordinates.

In total, we identified 55 presences for *G. ozarkensis* and 50 presences for *S. ozarkensis* (Figure 1). Spatial autocorrelation of presences used in distribution modelling can result in inflated measures of prediction accuracy (Veloz, 2009; Kramer-Schadt *et al.*, 2013). Yet, previous studies have shown that when modelling geographically limited or rare species, each presence location matters and can largely influence prediction outcomes (Almeida *et al.*, 2010; Silva *et al.*, 2013, 2016). Therefore, we did not explicitly control for spatial autocorrelation among presence localities, however, we did ensure that only a single presence locality was associated to each unit of prediction, which were individual stream segments for this modelling effort. Still, we acknowledge likely sampling bias in these data, particularly in the southwestern region of our study area, thus caution should be given when interpreting model predictions. *S. ozarkensis* localities that fell outside of the Ozark-Ouachita region (Fig. 1) were removed as these records are considered dubious or are of disjunct populations (Smith & Patten, 2020). Removal also facilitated ease of generating pseudo-absences from the same area for both species (Barbet-Massin *et al.*, 2012). Omission of these locations may result in biased predictions but given our immediate goal of increasing detections only within the study region, we included presences falling within this same region. Still, caution should be given to the model predictions for *S. ozarkensis* due to this omission.

For environmental predictors we used the U.S. Environmental Protection Agency's national StreamCat database (Hill *et al*., 2016). This database includes over 517 natural and anthropogenic environmental features collated from other sources for 2.6 million stream segments within the conterminous U.S. at a resolution of 30 m (Hill *et al*., 2016). These data are based on the National Hydrography Dataset Plus V2 geospatial framework and allow for modelling of individual stream segments, which were clipped to the target region based on the study by Leasure *et al*. (2016). StreamCat environmental variables were pre-selected based on knowledge of odonate biology and previous literature on odonate distribution modelling (Domisch *et al*., 2011, 2013b; Hassall, 2012; Kuemmerlen *et al*., 2014; Collins & McIntyre, 2015, 2017). To reduce overfitting, environmental predictors were removed if highly correlated ($r > 0.7$) (Dormann *et al*., 2013) resulting in a total of 12 predictors used in model training (Table 1). Dragonfly presence locations were uploaded into QGIS (QGIS Development Team 2018) and snapped to the nearest stream segment by using the Snap Geometries to Layer tool. Each point location was inspected to ensure proper stream associations and reassigned manually when necessary.

Two machine-learning methods, MaxEnt and Random Forest, were used to model the distributions of *G. ozarkensis* and *S. ozarkensis*. MaxEnt models use species presence data and background environmental variables to associate known presences to a unique set of environmental conditions, and then predict the probability of presence onto other locations (Phillips *et al*., 2006; Elith *et al*., 2011). Random Forest models are built on the classification and regression Ttree (CART) framework, but fit many classification trees to a dataset and combine the predictions from all trees (Breiman, 2001; Cutler *et al*., 2007). Both of these models have been used to estimate odonate distributions, though MaxEnt has been implemented more frequently (Collins & McIntyre, 2015).

MaxEnt models were executed in the open source software interface (Phillips *et al*., 2017). Model parameters were kept at default settings, except models were run with a 10-fold cross validation because we lacked enough presence data to create a testing set. Predictions were made onto the background StreamCat data across the entire region in the format of logistic probabilities. Random Forest models were executed using the *Caret* (Kuhn 2018) package in the statistical program R (R Core Team 2018). Random Forest models require absence data, which were not available. Therefore random, pseudo-absences were generated using the *spsample* function in the R package *sp* (Pebesma & Bivand 2005), constrained by the study region boundary (Fig. 1). An equal number of pseudo-absences to presences were generated for both modelled species (Barbet-Massin *et al*., 2012). A total of 500 trees were created and models were run with a 10-fold cross-validation. Binary predictions (present/absent) were projected onto background StreamCat data.

Table 1. StreamCat environmental predictors used in training stream-level current distribution models.

Hydrologic predictors	Landscape predictors	Climatic predictors
Stream base flow (USGS)	% High urbanisation (NLCD)	Mean annual temperature (PRISM)
Hydrologic conductivity (USGS)	% Cropland (NLCD)	Mean annual precipitation (PRISM)
	% Deciduous forest (NLCD)	
	% Coniferous forest (NLCD)	
	% Grassland (NLCD)	
	% Herbaceous wetland (NLCD)	
	Population density 2010 (US Census)	
	Road density (Tiger Lines)	

Original data sources are provided in parentheses.

To combine model outputs, MaxEnt predictions were transformed into the binary format of Random Forest predictions by applying the minimum training presence threshold (Pearson *et al*., 2007). This allowed for streams predicted as suitable under a binary format to be extracted for both models. Since MaxEnt predicted a larger area for both species, we ultimately clipped the MaxEnt predictions to the Random Forest predictions, which allowed us to convert back into the logistic format (see Supporting Information SI for binary predictions maps generated by the MaxEnt models). The resulting maps were used to guide field surveys for both species in the second (2018) field season; because stream accessibility varied a range of sites were chosen across the probability spectrum (0–1).

To assess SDM accuracy, we used the area under the curve (AUC) which is a threshold-independent discrimination metric that represents the probability that a random presence or absence is correctly assigned by the model (Phillips *et al*., 2006). An AUC score of 0.5 means the model is no better than random chance in correctly assigning presence or absence of species. This metric may be misleading, however, and is highly dependent on species prevalence (Lobo *et al*., 2008). High AUC values do not necessarily reflect good model fit as it is possible for a poorly fitted model to have high discrimination power and vice versa (Lobo *et al*., 2008). Yet, even with recent suggestions to use other metrics such as the True Skill Statistic (TSS) (Allouche *et al*., 2006), these metrics are still dependent on species prevalence and require quality presence–absence data to obtain an accurate estimate of model performance (Leroy *et al*., 2018). Therefore, we suggest interpreting these values with caution and combining them with the results of our field surveys to assess model performance. We also used the out-of-sample (OOS) error rate to assess Random Forest models, which is a measure of prediction error generated by the cross-validation technique.

Variable importance and response curves for the top three important environmental variables were compared across models for both species. MaxEnt models estimate variable importance using two different metrics and we compared the percent contribution of environmental predictors, which estimates the contribution of each predictor by measuring how much the predictor influences the model across each iteration. Marginal response curves, which show how a single predictor variable affects the prediction while all others are held at their mean, were compared for the top three contributing variables. Variable importance for the Random Forest models was calculated using the function *varImp* in the Caret package (Kuhn, 2018) in R (R Core Team, 2018). Partial dependence plots for these top three predictors from the Random Forest models were created with R package *pdp* (Greenwell, 2017). These plots show how the values of a single predictor affect the prediction of the correct class from the model. At a specific value of a single predictor variable, a negative value on the y-axis means the model is less likely to predict the correct class whereas a positive value on the y-axis means the model is positively influenced by that predictor variable.

Future distribution modelling

StreamCat environmental predictors are not available for future time periods, thus models to create predictions for the year 2070 were trained on a widely used bioclimatic dataset including 19 temperature and precipitation variables available through the WorldClim database (Hijmans *et al*., 2005). Future climatic environments were generated from global climate models that are based on different representative concentration pathways (RCPs). These pathways represent various levels of future greenhouse gas emissions, and thus differing levels of severity of climate warming. RCP 8.5 accounts for continuous rising in carbon dioxide emissions into the 21st century, while RCP 6.0 accounts for a peak in emissions around 2080 followed by slight decline. We characterised future environments within the study region by including bioclim variables representing these two emission scenarios generated by the Community Climate System Model 4. Current and future bioclim variables were downloaded at a spatial resolution of 2.5 min, and variables were removed if highly correlated ($r > 0.7$) (Dormann *et al*., 2013), resulting in a total of six predictors (Table 2).

MaxEnt and Random Forest algorithms were used to model the future distributions of *G. ozarkensis* and *S. ozarkensis* for both RCP scenarios. New occurrence data collected following field surveys (2017 and 2018) were incorporated, resulting in 63 presences for *G. ozarkensis* and 58 for *S. ozarkensis* for model training. MaxEnt models were executed as described above; however, when using raster predictors, MaxEnt removes duplicate presence records as to retain only one location per pixel. The bioclim predictors used had a spatial resolution of $\sim$5 km, and the numbers of presences ultimately used for MaxEnt training were 53 for *G. ozarkensis* and 47 for *S. ozarkensis* (Figure 1). To obtain the average prediction from the cross-validation models, logistic probabilities were averaged across each fold for every projected scenario. Random Forest models were executed as above, but to remain consistent with MaxEnt models, presence data were thinned based on a distance of 5 km (R package *spThin*; Aiello-Lammens *et al*., 2015). An equal number of random pseudo-absence points were generated (R package *sp*; Pebesma & Bivand 2005) within the Ozark-Ouachita Interior Highlands region.

Table 2. BioClim predictors used to train climate-based models for future projections.

Code	Climatic variable
Bio 8	Mean Temperature of Wettest Quarter
Bio 9	Mean Temperature of Driest Quarter
Bio 10	Mean Temperature of Warmest Quarter
Bio 16	Precipitation of Wettest Quarter
Bio 18	Precipitation of Warmest Quarter
Bio 19	Precipitation of Coldest Quarter

To combine model outputs, we used the Raster Calculator tool in QGIS (QGIS development team 2018) to multiply binary prediction rasters produced by Random Forest models with the binary raster predictions generated by MaxEnt after applying the minimum training presence threshold (Pearson *et al*., 2007) to the logistic output. The resulting maps only show areas both models predicted as suitable rather than just one or the other. In addition, we estimated the area lost or gained under both climate change scenarios for each species as predicted by both MaxEnt and Random Forest models (Supporting Information SI and see Results). To do so, we used the QGIS Raster Calculator to subtract the current distribution maps from future predictions. Then, we reprojected raster predictions to an equal area projection EPSG:102003 and used the QGIS Raster Layers Unique Values Report tool to estimate suitable area lost or gained in km^2.

Models were evaluated as above, using AUC and OOS error rates to assess model discrimination and accuracy. In addition, measures of variable importance were also compared across models using the same methods stated above. The top three predictors for each model were estimated and variable response curves were constructed to determine the influence of those predictors on the models.

To evaluate potential areas of extrapolation in future predictions, a multivariate environmental similarity surfaces (MESS) analysis was conducted (Elith *et al*., 2010). This method compares the similarity between the environment used in model training to future environments the model is being projected to which allows for identification of areas where careful interpretation may be needed (Elith *et al*., 2010).

Results

Field surveys

A total of 36 sites were surveyed throughout the Ozark and Ouachita mountains during the first season (summer 2017), including a mix of known localities and opportunistic sites. We detected *G. ozarkensis* twice, one adult male at a known locality and one adult female at an opportunistic site. A gravid female of *S. ozarkensis* was detected at a known locality, adding evidence of a new breeding location. There were no other records of either species reported in 2017.

In 2018, a total of 77 sites were surveyed throughout Arkansas, Oklahoma, and Missouri, guided by the combined predictions of the current distribution models (see below). We detected eight new presence locations for *S. ozarkensis* (Figure 2), including one new breeding site. Seven of these detections were predicted by both models and occurred across a wide range of probabilities (0.2–1.0). One site was only

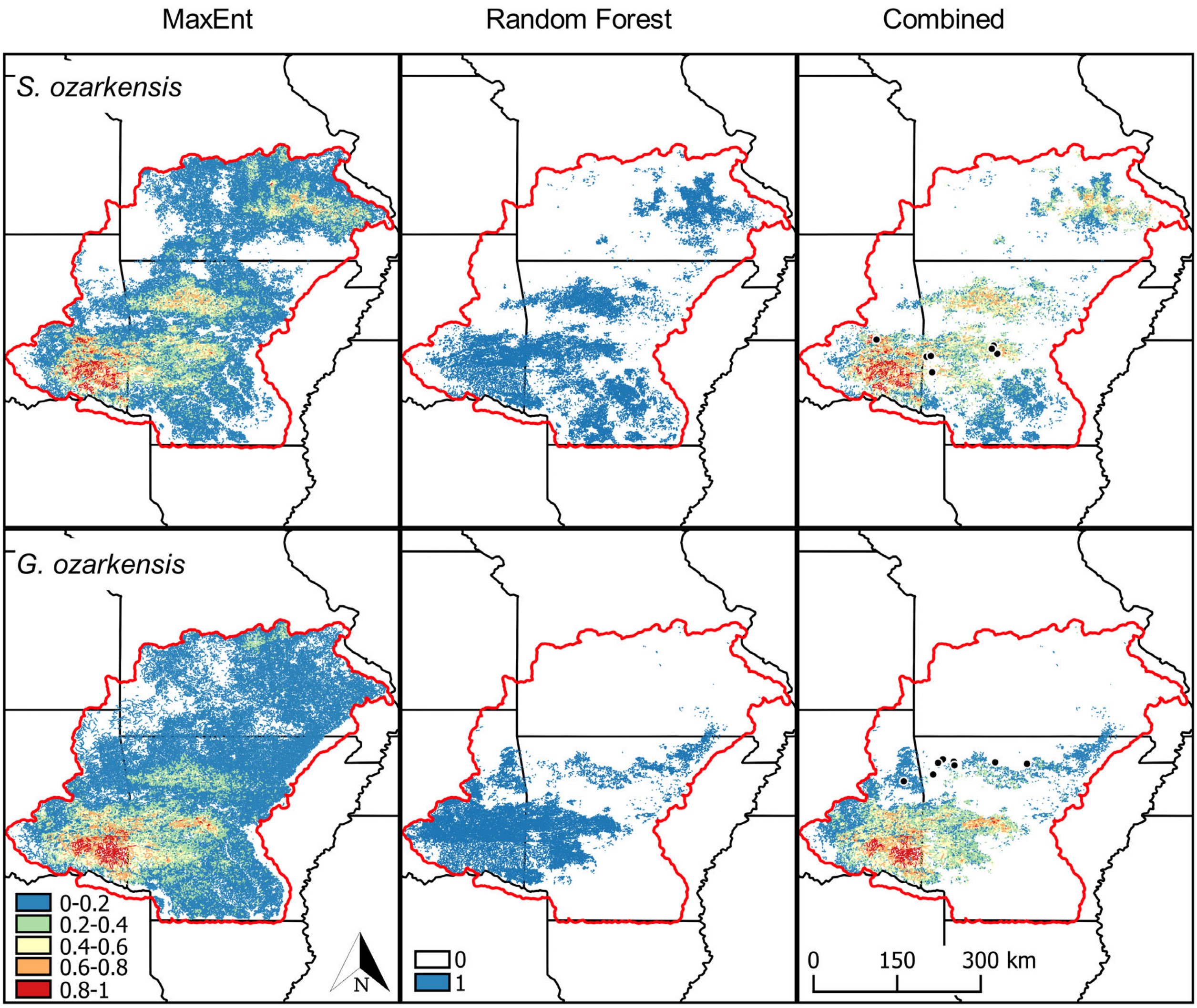

Figure 2. Stream level predictions of both focal species. MaxEnt-only, Random Forest-only, and combined model predictions are presented as probability of presence (0–1) represented by colour classifications from low probability, Blue (0–0.2) to high probability, Red (0.8–1). Ozark-Ouachita Interior Highlands region outlined in red. New detections discovered through field surveys of each species are shown with black points. [Color figure can be viewed at wileyonlinelibrary.com]

predicted by MaxEnt. We also detected eight new presence locations for *G. ozarkensis* (Figure 2). Two of these were predicted by both models, while the remaining were located at low probability (0–0.2) streams predicted by MaxEnt. Overall, these predictions increased detections for both species compared to the previous year.

Current distribution models

MaxEnt and Random Forest models showed high discrimination power for *G. ozarkensis* with AUC scores of 0.88 and 0.84, respectively, and for *S. ozarkensis* with scores of 0.87 and 0.74, respectively (Table 3). Overall, MaxEnt predicted more streams as suitable compared to Random Forest (Fig. 2). Random Forest models largely overlapped with the highest probability streams from the Maxent models (Fig. 2) yet had lower AUC scores. After combining model predictions, *S. ozarkensis* is predicted to occur mostly in the Ozark mountains in Missouri and Arkansas, as well as the Ouachita mountains in southern Arkansas and Oklahoma. *G. ozarkensis* is predicted to occur mainly in the Ouachita mountains of Arkansas and Oklahoma (Fig. 2).

An analysis of variable contribution showed similarity between species (Table 4). Variable importance for the Random Forest models showed congruence with MaxEnt variable contributions for *S. ozarkensis*, but slightly different variables for *G. ozarkensis* (Table 4). Percent coniferous forest, human population density and stream base flow were the top three important variables for *S. ozarkensis* for both models. This largely coincides with what we know from its observed distribution, which are typically streams found in steep, mixed-forest habitats. These three variables were also the top three important predictors for

Table 3. Accuracy metrics for current (stream-level) and future (climate-based) distribution models.

	Species	Presences	MaxEnt AUC	Random Forest AUC	OOS error rate
Current models	*Somatochlora ozarkensis*	50	0.868	0.74	29
	Gomphurus ozarkensis	55	0.876	0.85	15.6
Future models					
	Somatochlora ozarkensis	58	0.978	0.84	22.5
	Gomphurus ozarkensis	63	0.977	0.79	27.17

the MaxEnt model of *G. ozarkensis*; however, stream base flow, mean annual precipitation, and mean annual temperature were the top three variables for the Random Forest model. From what is known or observed about the habitat of *G. ozarkensis*, it prefers medium sized streams with open canopy, cobble riffles, and exposed rocks. Marginal response curves of the top three predictors from the MaxEnt models showed similar relationships across species (Supplemental Information Fig. S1). Partial dependence plots of the top three predictors from the Random Forest models showed that Stream Base Flow had the most positive impact on correct class identification by the models for both species (Supplemental Information Fig. S2).

Future distribution models

MaxEnt and Random Forest models showed high discrimination power for *G. ozarkensis* with AUC scores of 0.98 and 0.79, respectively, and for *S. ozarkensis* with scores of 0.98 and 0.84, respectively (Table 3). An analysis of variable contribution showed similarity in important environmental characteristics between species regardless of algorithm (Table 5). The top three important variables for both species, albeit in different orders of importance, were precipitation of the coldest quarter, precipitation of the wettest quarter, and mean temperature of the driest quarter. The only exception to this was for the Random Forest model of *S. ozarkensis* where precipitation of the warmest quarter was more important than precipitation of the coldest quarter (Table 5). Marginal response curves of the top three predictors from the MaxEnt models showed similar relationships across species (Supplemental Information Fig. S1). Partial dependence plots of the top three predictors from the Random Forest models showed that all top three predictors had a positive impact on correct class identification by the models for both species (Supplemental Information Fig. S2).

Overall, MaxEnt predicted more area as suitable under all climate change scenarios compared to Random Forest (Figs. 3 and 4). MaxEnt predictions remained similar across all scenarios for both species, however, Random Forest models predicted smaller suitable areas under the RCP 6.0 scenario (Figs. 3 and 4). Predicted suitability then increased slightly under the RCP 8.5 scenario (Figs. 3 and 4). Both MaxEnt and Random Forest models show distributions are predicted to shift slightly west further into

Table 4. Variable importance or percent contribution for stream-level distribution models.

Species	Variable	%Contribution or importance
Somatochlora ozarkensis		
MaxEnt	% Coniferous forest	27.7
	Stream base flow	23.6
	Population density	13
Random Forest	Population density	100
	Stream base flow	99.9
	% Coniferous forest	94.2
Gomphurus ozarkensis		
MaxEnt	Population density	32.9
	Stream base flow	21.9
	% Coniferous forest	14.3
Random Forest	Stream base flow	100
	Mean annual precipitation	43.8
	Mean annual temperature	34.5

Table 5. Variable importance or percent contribution for future distribution models.

Species	Variable	% Contribution or Importance
Somatochlora ozarkensis		
MaxEnt	Precip. of coldest quarter	52.5
	Precip. of wettest quarter	21.2
	Mean temp. of driest quarter	14.9
Random Forest	Precip. of wettest quarter	100
	Mean temp. of driest quarter	36
	Precip. of warmest quarter	17.4
Gomphurus ozarkensis		
MaxEnt	Precip. of coldest quarter	42.5
	Mean temp. of driest quarter	31.7
	Precip. of wettest quarter	17.6
Random Forest	Precip. of wettest quarter	100
	Mean temp. of driest quarter	77.4
	Precip. of coldest quarter	43.1

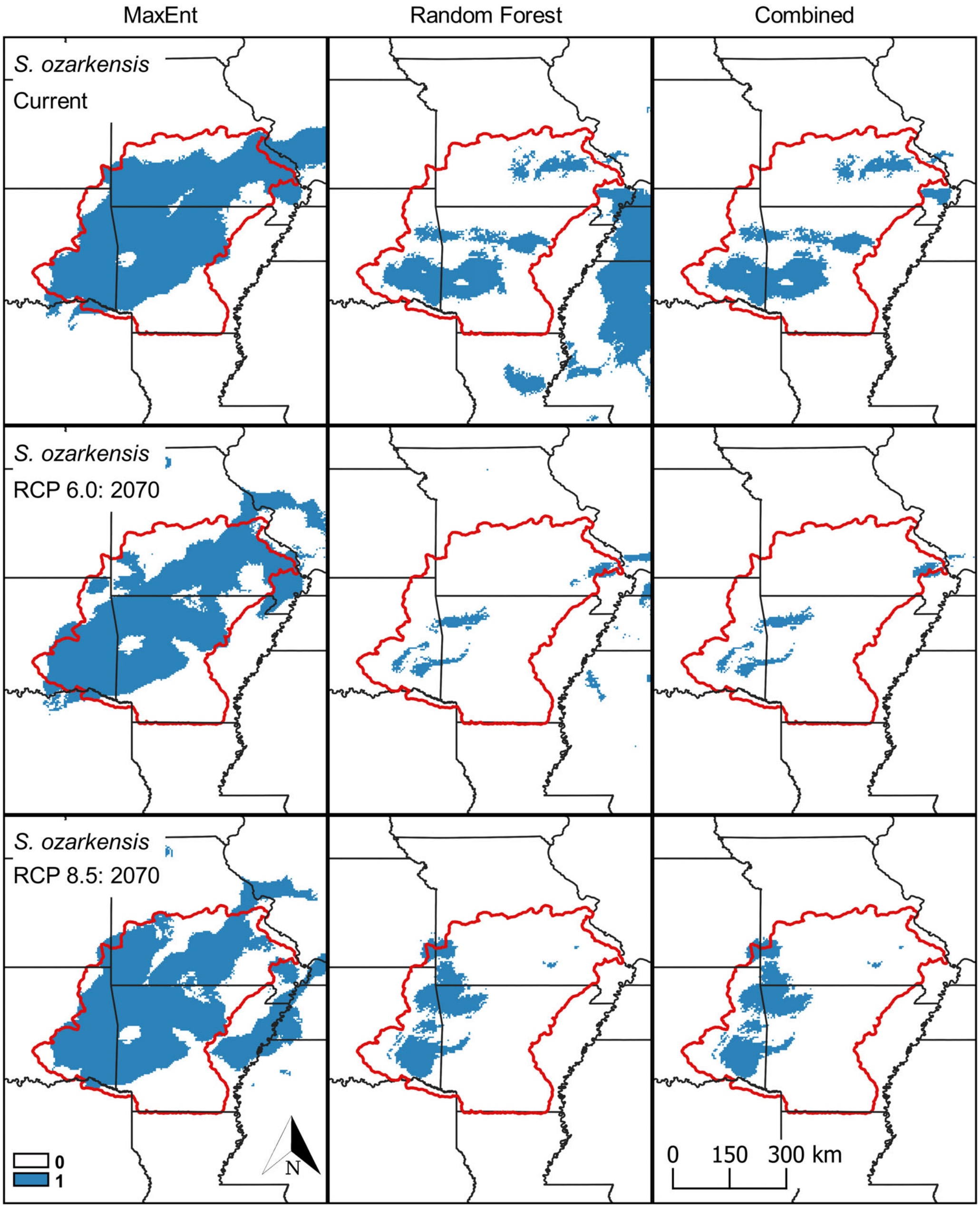

Figure 3. Current and future predictions for the mid-level (RCP 6.0) and most severe (RCP 8.5) climate change scenarios of *S. ozarkensis* using WorldClim bioclimatic variables. Ozark-Ouachita Interior Highlands region outlined in red. Blue represents predicted presence and white represents predicted absence. [Color figure can be viewed at wileyonlinelibrary.com]

Oklahoma and north further into the Ozark mountains for both species. Random Forest predictions, although smaller than MaxEnt predictions, showed large overlap with the highest probability areas predicted by MaxEnt models as shown by combining model predictions (Figs. 3 and 4). Under the RCP 8.5 scenario, MaxEnt models predicted *G. ozarkensis* to have a net loss of 63,781 km^2 of suitable area and *S. ozarkensis* to have a net gain of 2522 km^2 of suitable area (Supplemental Information Fig. S4). Random Forest models for the same scenario predicted *G. ozarkensis* to have a net loss of 138 372 km^2, and *S. ozarkensis* to have a net loss of 175 618 km^2 of suitable area (Supplemental Information Fig. S5).

A MESS analysis showed that, in some areas, future environments within the study region are novel compared to the model training environment (Fig. 5). Thus, future predictions should be interpreted with caution, particularly for the RCP 8.5 scenario, as extrapolating beyond the calibration data can result in overpredictions in potentially unsuitable habitat.

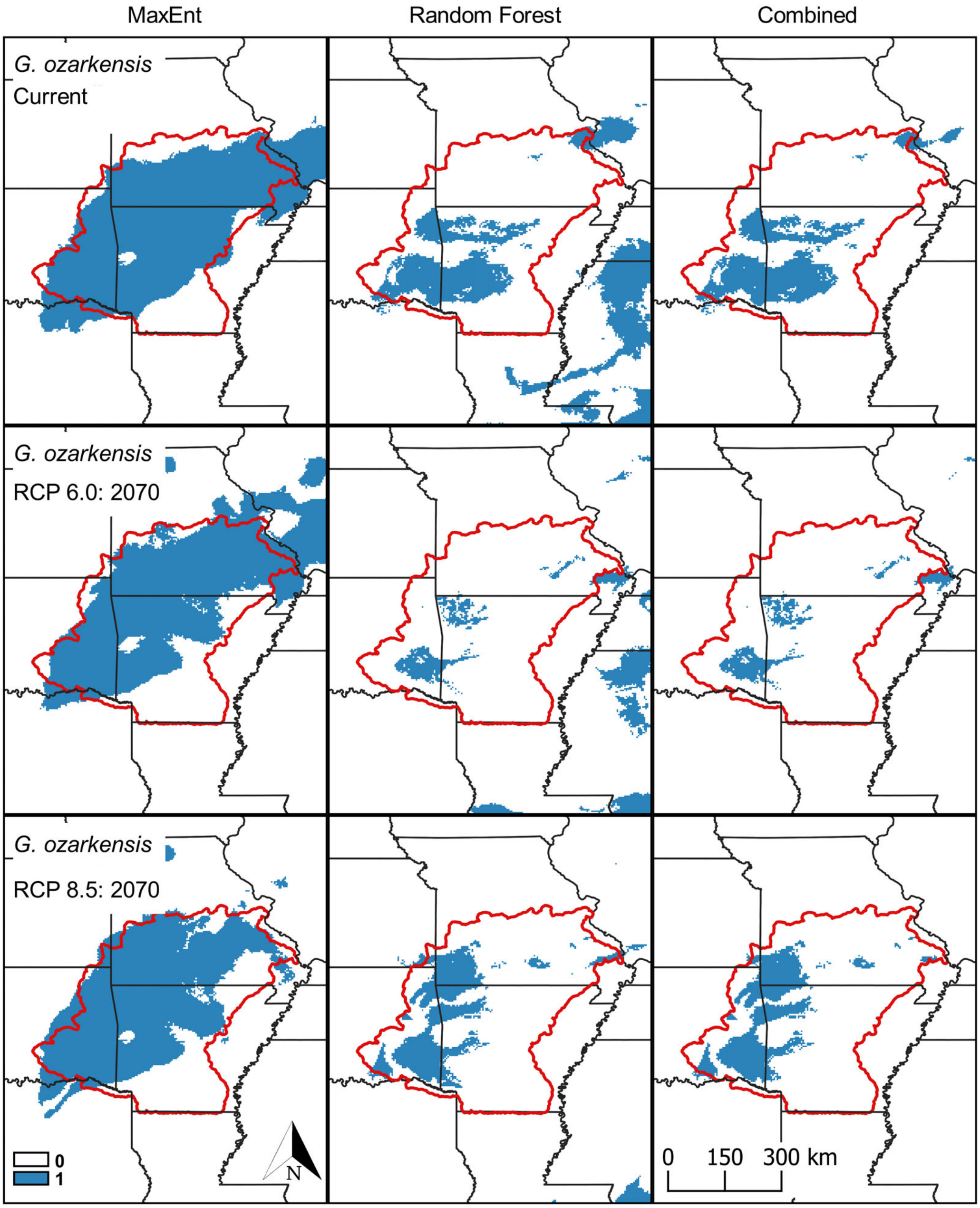

Figure 4. Current and future predictions for the mid-level (RCP 6.0) and most severe (RCP 8.5) climate change scenarios of *G. ozarkensis* using WorldClim bioclimatic variables. Ozark-Ouachita Interior Highlands region outlined in red. Blue represents predicted presence and white represents predicted absence. [Color figure can be viewed at wileyonlinelibrary.com]

Discussion

Conservation efforts are increasingly being applied to insects, but assigning status is challenging due to limited distribution data. The challenges can be amplified when species are regionally endemic or have very limited geographic ranges. Using SDM to predict insect distributions is becoming more common; however, fewer than 25% of odonate species have been modelled in this way (Collins & McIntyre, 2015). The aquatic to terrestrial life cycle of these insects allows for modelling approaches that focus on water bodies, terrestrial landscapes, or both. Small-scale stream level modelling can fit within endemic range limits and provide more accurate predictions for stream species than coarse scale landscape models and may also better inform management practises as local catchments are often the unit targeted for conservation (Kuemmerlen *et al*., 2014). We demonstrated the value of stream-level predictions for increasing detections of rare and regionally endemic dragonflies that can aid in guiding conservation efforts.

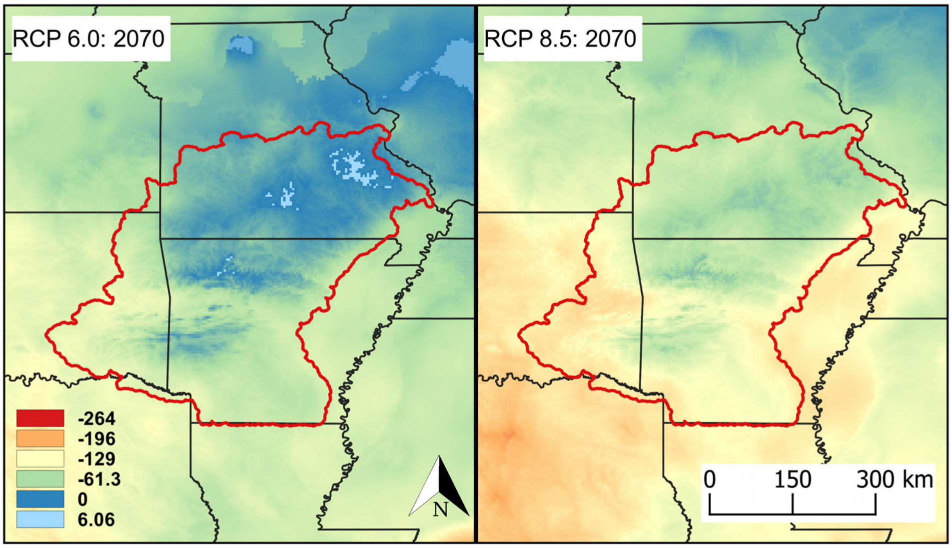

Figure 5. MESS analysis showing areas of environmental dissimilarity between training and future environments. Negative values represent environments that are novel compared to the training environment. Positive or 0 represents similar environments to the training environment and are shown in the blue colours. [Color figure can be viewed at wileyonlinelibrary.com]

There is no shortage of modelling techniques available, including ensemble approaches that combine predictions from multiple models (Thuiller *et al*., 2009). Although the suite of algorithms continues to grow, no single algorithm has outperformed others and choice of algorithm is context dependent (Qiao *et al*., 2015). We therefore chose to combine predictions from two machine-learning algorithms that have consistently provided robust predictions for species with limited presence only data as in this study. Another method is to use an iterative model approach. For example, use of this later approach led to the discoveries of new populations of a rare stonefly (Young *et al*., 2019), but such an approach requires resources for multi-year sampling efforts to successively incorporate new detections into these models, which is often not feasible. Our results suggest there is utility in combining predictions from two robust algorithms as we increased detections for both modelled species compared to baseline field surveys the previous year. Any additional detections for geographically limited, rare, or cryptic taxa, such as our focal species, are especially valuable to the understanding of their distributions and potential habitat requirements.

As the climate continues to warm, many species will be affected, with ectotherms potentially being most sensitive as their basic physiological functions are strongly influenced by external temperature (Deutsch *et al*., 2008). These species may respond to warming temperatures in many ways, such as advancing their phenologies or shifting their geographic distributions to track optimal thermal ranges (Parmesan, 2006). For example, numerous insects and other taxa in Britain have largely shifted northward during a period of climate warming (Hickling *et al*., 2005, 2006). The ability of a species to track thermal habitat depends on various factors including dispersal capabilities, degree of habitat specialisation, and thermal tolerance breadth (Rocha-Ortega *et al*., 2020). Although it has been hypothesised that ectotherms at high latitudes exhibit a greater range of thermal tolerance, a global analysis on insect data showed this only held for species currently undergoing range expansions (Lancaster, 2016). Insular or endemic species showed no latitudinal variation in thermal tolerance breadth suggesting that these species are unable to tolerate warming at higher latitudes (Lancaster, 2016).

Instead, endemic species may experience further contraction of an already narrow range or shifts to novel conditions to which they are maladapted. For example, a damselfly endemic to the Pampa region in South America was projected to undergo range contractions through multiple climate change scenarios (Pires *et al*., 2018). Contractions to lower elevations are also predicted for odonates; out of 14 dragonfly species modelled, 12 were predicted to experience reduced extents at higher elevations (Simaika & Samways, 2015). Future projections of our focal taxa also indicate range contractions and biogeographic shifts to the west and north in response to climate change, and potentially elevational shifts that were ignored by our analysis. The contractions and shifts may seem slight, but regional endemic species, by definition, already have limited distributions and often limited ecological amplitude or niche breadth. Our models for the worst-case scenario RCP 8.5 showed net losses of suitable area for both species except the MaxEnt predictions for *S. ozarkensis* that predicted a slight gain in suitable area. Previous studies suggest regional endemic species might persist for a time (Pires *et al*., 2018), but the very nature of endemicity makes them more vulnerable to environmental change, whether catastrophic or gradual. Taken together, these studies imply that vulnerability to climate change is largely context dependent, and that resiliency may trade off with capacity to shift or adapt.

Conservation implications

For strategic conservation planning to occur, knowledge of current and future species distributions is critical but remains one of many data shortfalls impacting invertebrate conservation

(Cardoso *et al*., 2011; Samways, 2015). Extinction risk is largely based on extent of occurrence or range size (Rocha-Ortega *et al*., 2020), thus the lack of distributional data hinders our ability to assess threat status (Gaston & Fuller, 2009). Recent advances in ecological modelling have enhanced our ability to estimate and predict where species occur in relation to environmental characteristics and can help fill in data gaps when assessing conservation need. In this study, future projections suggest conservation efforts should focus within the study region given the limited predicted shifts in these species' distributions. In contrast, other studies have shown potential range shifts of odonates in response to climate change, particularly lentic species (Christian *et al*., 2012). Although our models predict a limited range shift, they do estimate a considerable reduction of suitable area for both species under the RCP 8.5 scenario. Thus, further efforts to understand current population trends are imperative for the conservation of these species.

Regardless of whether projections show shifts in a species distribution or not, we caution against these efforts being used to preclude conservation plans. That is, if a species is predicted to show a shift in its distribution outside of its current management jurisdiction, that agency should not abandon current efforts as predictions are never perfect and detections of new populations may continue to be discovered. Instead, we suggest these techniques can be used to implement cross-boundary or regional collaborations that can pool resources, diversify strategies, and hopefully create more accurate conservation assessments (Rodrigues & Gaston, 2002). These collaborations will likely become increasingly common as species distributions shift in response to global change. Development of regional conservation plans for *G. ozarkensis* and *S. ozarkensis* is feasible thanks to state wildlife grants dedicated to species of greatest conservation need (Bried & Mazzacano, 2010). While there may be challenges for allocating funds and responsibility, open communication and structured decision making between jurisdictions and stakeholders can lead to actionable conservation (Guisan *et al*., 2013).

Limitations

Two terminologies are often used in the literature when modelling species–environment relationships: SDM and Environmental Niche modelling. These two terms are fundamentally different in what they aim to model. Species distribution models are a correlative approach that associate species presences to environmental data and one limitation to this method is that biotic interactions are not considered, thus many aspects that constitute a species niche are not addressed (Elith & Leathwick, 2009). As such, we refer to SDM in the present study since we acknowledge the underlying processes resulting in observed occurrence patterns are not modelled (McInerny & Etienne, 2013). That is, the results are entirely phenomenological and not mechanistic.

The predicted distributions of these species were modelled using climatic variables, and though these are appropriate given the life history traits of odonates (Corbet, 1999), they do not account for dispersal ability or species interactions which are also important in determining species distributions. For example, biotic interactions may restrict the area predicted as suitable in which a species can occupy and even under suitable abiotic and biotic conditions, a species can only occur in areas that have been accessible to them for some time (Soberon & Peterson, 2005; Soberón, 2007; Barve *et al*., 2011; Saupe *et al*., 2012). Another limitation of using SDM to predict future distributions is the assumption that the current distribution of a species is in equilibrium with the current environmental conditions, which is often not the case and cannot be determined with only climatic variables (Lobo, 2016). Models trained solely on presence-only data and pseudo-absences do not reflect causal mechanisms of a species distribution, but rather the density of observations used to train the model (Lobo, 2016).

Taken together, simply using presence-only data and environmental predictors suggests we are predicting areas of similar climatic conditions in which a species was observed and not true future distributions (Lobo, 2016), thus cautious interpretation is recommended. Furthermore, the presence of a species in an area does not necessarily mean that area is suitable breeding habitat (Siepielski & Mcpeek, 2013; Patten *et al*., 2019). Species distributions likely include sink populations where the presence of a species may be a result of ongoing dispersal and not of long-term positive population growth rate resulting from a suitable environment (Pulliam, 1988).

Further limitations of this approach include using adult records as presences, since odonates have a complex life cycle and thus different habitat requirements as nymphs and adults. Using only adult records as input for distribution models can produce misleading predictions of habitat suitability (Patten *et al*., 2015). Nevertheless, our focal taxa are data deficient and we generally lack information about nymph, exuviae, tenerals or breeding behaviour at occupied localities. Thus, we acknowledge that using adult records as input for these models may not accurately represent the distribution or habitat suitability for nymphs, which is often the critical life stage influencing population regulation (McPeek & Peckarsky, 1998). Nonetheless, these models resulted in higher detections of adults, including one new breeding location for *S. ozarkensis*, and are therefore useful as a means to guide future field surveys that may target the aquatic stage or capture evidence of breeding.

Conclusions

This study provided extensive field surveys and distribution modelling for two understudied dragonflies of conservation concern throughout the Ozark-Ouachita Interior Highlands region. We demonstrated the utility of SDM in guiding field surveys and increasing detections of endemic or rare species. These surveys can lead to better informed conservation assessments for species of concern. Further, these same models can be used with current and future climatic scenarios to project species distributions into the future and serve as a resource to develop conservation plans in light of different climate change scenarios. Projections of our focal taxa suggest that not all regional endemics are particularly vulnerable to changes in future climate as there still remain areas of high predicted suitability. Given the current method of assessing conservation need at the state level, we recommend using SDM techniques to facilitate cross-governance-boundary collaborations since species ranges often do not coincide with geopolitical boundaries. Finally, we caution against using these methods to forego conservation planning as there are limitations to these models and other factors such as dispersal and biotic interactions that will certainly affect species distributions in the future.

Acknowledgements

The authors would like to thank the Arkansas Game and Fish Commission (State Wildlife Grant Agreement # T-70) and The Alongside Wildlife Foundation for providing funding for this work. Thanks to Bruce Henry and Brett Landwer of the Missouri Department of Wildlife Conservation for their help with field surveys. The authors appreciate the Missouri Natural Heritage Program and their contributions to this work. This study would not have been possible without foundational research on these species provided for by the Oklahoma Department of Wildlife Conservation. Also, many thanks to Lauren Wishard and Zachary Bragg for their immense help with field surveys. The authors also thank Dr. Mona Papeş for input in regards to modelling techniques. Finally, the authors thank Daniel de Paiva Silva and one anonymous reviewer of this manuscript for providing many excellent suggestions that improved this work.

Data availability statement

The data that support the findings of this study are available at OdonataCentral.org. R code available upon request.

Supporting information

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

Figure S1. Variable response curves for current (A-C, and H-J) and future (D-F, and K-M) MaxEnt models. Variables shown in order of importance. Red represents how the logistic prediction changes as each predictor is varied while keeping all other predictors at their average value. Blue represents the variable response +/− 1 standard deviation. PctConif2011Cat represents percent coniferous forest, BFICat represents stream base flow, and PopDen2010Cat represents human population density. Bio19 represents precipitation of coldest quarter, bio9 represents mean temperature of driest quarter, bio16 represents precipitation of wettest quarter.

Figure S2. Partial dependence plots of the top three predictor variables from the current Random Forest model (A-C) and future model (D-F) for *G. ozarkensis* and current (G-I) and future model (J-L) for *S. ozarkensis*. Black lines represent the actual response values of a single predictor and blue lines represent the LOESS smoothed function of those values. Values above 0 on the Y-axis mean that predictor had a positive effect on correct class identification by the model. The model is less likely to assign the correct class when these values fall below 0.

Figure S3. Binary maps of MaxEnt predictions for the stream-level model. Blue represents streams predicted as presence locations after applying the minimum training presence threshold.

Figure S4. Areas gained and lost under both climate change scenarios as predicted by MaxEnt models. Orange represents areas lost, white represents areas unchanged, and blue represents area gained.

Figure S5. Areas gained and lost under both climate change scenarios as predicted by Random Forest models. Orange represents areas lost, white represents areas unchanged, and blue represents area gained.

References

Abbott, J.C. (2006) OdonataCentral: an online resource for the distribution and identification of Odonata.

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. & Anderson, R.P. (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, **38**, 541–545.

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Almeida, M.C.D., Côrtes, L.G. & Júnior, P.D.M. (2010) New records and a niche model for the distribution of two Neotropical damselflies: Schistolobos boliviensis and Tuberculobasis inversa (Odonata: Coenagrionidae). *Insect Conservation and Diversity*, **3**, 252–256.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810–1819.

Bird, R.D. (1933) *Somatochlora ozarkensis*, A new species from Oklahoma (Odonata-Cordulinae). Occassional papers of the museum of zoology, University of Michigan, 261.

Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5–32.

Bried, J.T. & Mazzacano, C.A. (2010) National review of state wildlife action plans for Odonata species of greatest conservation need. *Insect Conservation and Diversity*, **3**, 61–71.

Cardoso, P., Erwin, T.L., Borges, P.A.V. & New, T.R. (2011) The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, **144**, 2647–2655.

Collins, S.D. & McIntyre, N.E. (2015) Modeling the distribution of odonates: a review. *Freshwater Science*, **34**, 1144–1158.

Collins, S.D. & McIntyre, N.E. (2017) Extreme loss of diversity of riverine dragonflies in the northeastern U.S. is predicted in the face of climate change, 14.

Corbet, P. (1999) *Dragonflies: Behavior and Ecology of Odonata*. Cornell University Press.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007) Random Forests for classification in ecology. *Ecology*, **88**, 2783–2792.

De Marco, P.D.M., Nóbrega, C.C., de Souza, R.A. & Neiss, U.G. (2015) Modeling the distribution of a rare Amazonian odonate in relation to future deforestation. *Freshwater Science*, **34**, 1123–1132.

Deutsch, C.A., Tewksbury, J.J., Huey, R.B., Sheldon, K.S., Ghalambor, C.K., Haak, D.C. & Martin, P.R. (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences*, **105**, 6668–6672.

Dijkstra, K.-D.B., Monaghan, M.T. & Pauls, S.U. (2014) Freshwater biodiversity and aquatic insect diversification. *Annual Review of Entomology*, **59**, 143–163.

Domisch, S., Araújo, M.B., Bonada, N., Pauls, S.U., Jähnig, S.C. & Haase, P. (2013a) Modelling distribution in European stream macroinvertebrates under future climates. *Global Change Biology*, **19**, 752–762.

Domisch, S., Jähnig, S.C. & Haase, P. (2011) Climate-change winners and losers: stream macroinvertebrates of a submontane region in Central Europe. *Freshwater Biology*, **56**, 2009–2020.

Domisch, S., Kuemmerlen, M., Jähnig, S.C. & Haase, P. (2013b) Choice of study area and predictors affect habitat suitability projections, but not the performance of species distribution models of stream biota. *Ecological Modelling*, **257**, 1–10.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27–46.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

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.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Gaston, K.J. (1991) How large is a species' geographic range? *Oikos*, **61**, 434–438.

Gaston, K.J. & Fuller, R.A. (2009) The sizes of species' geographic ranges. *Journal of Applied Ecology*, **46**, 1–9.

Google Inc. (2015) Google Earth Pro, version 7.3.

Greenwell, B.,.M. (2017) pdp: an R package for constructing partial dependence plots. *The R Journal*, **9**, 421.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. & Zimmerman, N.E. (2006) Using niche-based models to improve the sampling of rare species. *Conservation Biology*, **20**, 501–511.

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., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424–1435.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D. & de Kroon, H. (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One*, **12**, e0185809.

Hassall, C. (2012) Predicting the distributions of under-recorded Odonata using species distribution models. *Insect Conservation and Diversity*, **5**, 192–201.

Hassall, C. (2015) Odonata as candidate macroecological barometers for global climate change. *Freshwater Science*, **34**, 1040–1049.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773–785.

Hickling, R., Roy, D.B., Hill, J.K., Fox, R. & Thomas, C.D. (2006) The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology*, **12**, 450–455.

Hickling, R., Roy, D.B., Hill, J.K. & Thomas, C.D. (2005) A northward shift of range margins in British Odonata. *Global Change Biology*, **11**, 502–506.

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**, 1965–1978.

Hill, R.A., Weber, M.H., Leibowitz, S.G., Olsen, A.R. & Thornbrugh, D.J. (2016) The stream-catchment (StreamCat) dataset: a database of watershed metrics for the conterminous United States. *JAWRA Journal of the American Water Resources Association*, **52**, 120–128.

Hof Christian, Brändle Martin, Dehling D. Matthias, Munguía Mariana, Brandl Roland, Araújo Miguel B., Rahbek Carsten (2012) Habitat stability affects dispersal and the ability to track climate change. *Biology Letters*, **8** (4), 639–643.

Klink, R., van Bowler, D.E., Gongalsky, K.B., Swengel, A.B., Gentile, A. & Chase, J.M. (2020) Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science*, **368**, 417–420.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., Macdonald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J. W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H. & Wilting, A. (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, **19**, 1366–1379.

Kuemmerlen, M., Schmalz, B., Guse, B., Cai, Q., Fohrer, N. & Jähnig, S. C. (2014) Integrating catchment properties in small scale species distribution models of stream macroinvertebrates. *Ecological Modelling*, **277**, 77–86.

Kuhn (2018) caret: Classification and Regression Training. R package version 6.0-81.

Lancaster, L.T. (2016) Widespread range expansions shape latitudinal variation in insect thermal limits. *Nature Climate Change; London*, **6**, 618–621.

Leasure, D.R., Magoulick, D.D. & Longing, S.D. (2016) Natural flow regimes of the Ozark–Ouachita interior highlands region. *River Research and Applications*, **32**, 18–35.

Leroy, B., Delsol, R., Hugueny, B., Meynard, C.N., Barhoumi, C., Barbet-Massin, M. & Bellard, C. (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *Journal of Biogeography*, **45**, 1994–2002.

Li, F., Kwon, Y.-S., Bae, M.-J., Chung, N., Kwon, T.-S. & Park, Y.-S. (2014) Potential impacts of global warming on the diversity and distribution of stream insects in South Korea. *Conservation Biology*, **28**, 498–508.

Lister, B.C. & Garcia, A. (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences*, **115**, E10397–E10406.

Lobo, J.M. (2016) The use of occurrence data to predict the effects of climate change on insects. *Current Opinion in Insect Science*, **17**, 62–68.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

Margules, C.R. & Pressey, R.L. (2000) *Systematic Conservation Planning*, Vol. **405**, pp. p. 11.

Markovic, D., Carrizo, S., Freyhof, J., Cid, N., Lengyel, S., Scholz, M., Kasperdius, H. & Darwall, W. (2014) Europe's freshwater biodiversity under climate change: distribution shifts and conservation needs. *Diversity and Distributions*, **20**, 1097–1107.

McInerny, G.J. & Etienne, R.S. (2013) 'Niche' or 'distribution' modelling? A response to Warren. *Trends in Ecology & Evolution*, **28**, 191–192.

McPeek, M.A. & Peckarsky, B.L. (1998) Life histories and the strengths of species interactions: combining mortality, growth, and fecundity effects. *Ecology*, **79**, 867–879.

Mi, C., Huettmann, F., Guo, Y., Han, X. & Wen, L. (2017) Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*, **5**, e2849.

Nóbrega, C.C. & De Marco, P.D. (2011) Unprotecting the rare species: a niche-based gap analysis for odonates in a core Cerrado area. *Diversity and Distributions*, **17**, 491–505.

Papeş, M. & Gaubert, P. (2007) Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity and Distributions*, **13**, 890–902.

Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.

Patten, M.A., Bried, J.T. & Smith-Patten, B.D. (2015) Survey data matter: predicted niche of adult vs breeding Odonata. *Freshwater Science*, **34**, 1114–1122.

Patten, M.A., Hjalmarson, E.A., Smith-Patten, B.D. & Bried, J.T. (2019) Breeding thresholds in opportunistic Odonata records. *Ecological Indicators*, **106**, 105460.

Patten, M.A. & Smith-Patten, B.D. (2013) Odonata species of special concern for Oklahoma, USA. *International Journal of Odonatology*, **16**, 327–350.

Patten, M.A. & Smith-Patten, B.D. (2014) The Oklahoma Odonata project: progress and trends. *Argia*, **26**, 19–25.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) ORIGINAL ARTICLE: predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.

Pebesma, E.J. & Bivand, R.S. (2005) Classes and methods for spatial data in R. *R News*, **5**(2).

Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E. & Blair, M.E. (2017) Opening the black box: an open-source release of Maxent. *Ecography*, **40**, 887–893.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pires, M.M., Périco, E., Renner, S. & Sahlén, G. (2018) Predicting the effects of future climate change on the distribution of an endemic damselfly (Odonata, Coenagrionidae) in subtropical South American grasslands. *Journal of Insect Conservation*, **22**, 303–319.

van Proosdij, A.S.J., Sosef, M.S.M., Wieringa, J.J. & Raes, N. (2016) Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, **39**, 542–552.

Pulliam, H.R. (1988) Sources, Sinks, and Population Regulation. *The American Naturalist*, **132**, 652–661.

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**, 1126–1136.

QGIS Development Team (2018) QGIS Geographic Information System. Open Source Geospatial Foundation Project. December 11, 2018. http://qgis.osgeo.org.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Accessed on March 8, 2018. https://www.R-project.org/.

Rangel-Sánchez, L., Nava-Bolaños, A., Palacino-Rodríguez, F. & Córdoba-Aguilar, A. (2018) Estimating distribution area in six Argia damselflies (Insecta: Odonata: Coenagrionidae) including *A. garrisoni*, a threatened species. *Revista Mexicana de Biodiversidad*, **89**.

Rhoden, C.M., Peterman, W.E. & Taylor, C.A. (2017) Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. *PeerJ*, **5**, e3632.

Rocha-Ortega, M., Rodríguez, P., Bried, J., Abbott, J. & Córdoba-Aguilar, A. (2020) Why do bugs perish? Range size and local vulnerability traits as surrogates of Odonata extinction risk. *Proceedings of the Royal Society B: Biological Sciences*, **287**, 20192645.

Rodrigues, A.S.L. & Gaston, K.J. (2002) Rarity and conservation planning across geopolitical units. *Conservation Biology*, **16**, 674–682.

Samways, M.J. (2015) Future-proofing insect diversity. *Current Opinion in Insect Science*, **12**, 71–78.

Sánchez-Bayo, F. & Wyckhuys, K.A.G. (2019) Worldwide decline of the entomofauna: a review of its drivers. *Biological Conservation*, **232**, 8–27.

Saupe, E.E., Barve, V., Myers, C.E., Soberón, J., Barve, N., Hensz, C. M., Peterson, A.T., Owens, H.L. & Lira-Noriega, A. (2012) Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. *Ecological Modelling*, **237–238**, 11–22.

Siepielski, A.M. & Mcpeek, M.A. (2013) Niche versus neutrality in structuring the beta diversity of damselfly assemblages. *Freshwater Biology*, **58**, 758–768.

Silva, D., Aguiar, A., Melo, G., Anjos-Silva, E. & Marco, P. (2013) Amazonian species within the Cerrado savanna: new records and potential distribution for *Aglae caerulea* (Apidae: Euglossini). *Apidologie*, **44**, 673–683.

Silva, D.P., Aguiar, A.G. & Simião-Ferreira, J. (2016) Assessing the distribution and conservation status of a long-horned beetle with species distribution models. *Journal of Insect Conservation*, **20**, 611–620.

Simaika, J.P. & Samways, M.J. (2015) Predicted range shifts of dragonflies over a wide elevation gradient in the southern hemisphere. *Freshwater Science*, **34**, 1133–1143.

Smith, B.D. & Patten, M.A. (2020) *Dragonflies at a Biogeographical Crossroads: The Odonata of Oklahoma and Complexities Beyond its Borders*. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.

Soberon, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**.

Stork, N.E. (1988) Insect diversity: facts, fiction and speculation. *Biological Journal of the Linnean Society*, **35**, 321–337.

Tessarolo, G., Rangel, T.F., Araújo, M.B. & Hortal, J. (2014) Uncertainty associated with survey design in species distribution models. *Diversity and Distributions*, **20**, 1258–1269.

Thomas, J.A. (2004) Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science*, **303**, 1879–1881.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.

Torres-cambas, Y., Cabana-otero, M., Lorenzo-carballa, M.O. & Cordero-rivera, A. (2016) Conservation status and protection of three Antillean endemic damselflies. *Journal of Insect Conservation*, **20**, 277–284.

Tronstad, L.M., Brown, K.M. & Andersen, M.D. (2018) Using species distribution models to guide field surveys for an apparently rare aquatic beetle. *Journal of Fish and Wildlife Management*.

Veloz, S.D. (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, **36**, 2290–2299.

Williams, J.N., Seo, C., Thorne, J., Nelson, J.K., Erwin, S., O'Brien, J. M. & Schwartz, M.W. (2009) Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, **15**, 565–576.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H. & Guisan, A. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Young, N.E., Fairchild, M., Belcher, T., Evangelista, P., Verdone, C.J. & Stohlgren, T.J. (2019) Finding the needle in the haystack: iterative sampling and modeling for rare taxa. *Journal of Insect Conservation*, **23**, 589–595.

Accepted 17 August 2020

Editor: Christopher Hassall; Associate Editor: Göran Sahlen