

Can Species Distribution Models Aid Bioassessment when Reference Sites are Lacking? Tests Based on Freshwater Fishes

Ben J. Labay¹ · Dean A. Hendrickson¹ · Adam E. Cohen¹ · Timothy H. Bonner² · Ryan S. King³ · Leroy J. Kleinsasser⁴ · Gordon W. Linam⁴ · Kirk O. Winemiller⁵

Received: 13 November 2014 / Accepted: 15 June 2015 / Published online: 20 June 2015
© Springer Science+Business Media New York 2015

Abstract Recent literature reviews of bioassessment methods raise questions about use of least-impacted reference sites to characterize natural conditions that no longer exist within contemporary landscapes. We explore an alternate approach for bioassessment that uses species site occupancy data from museum archives as input for species distribution models (SDMs) stacked to predict species assemblages of freshwater fishes in Texas. When data for estimating reference conditions are lacking, deviation between richness of contemporary versus modeled species assemblages could provide a means to infer relative biological integrity at appropriate spatial scales. We constructed SDMs for 100 freshwater fish species to compare predicted species assemblages to data on contemporary assemblages acquired by four independent surveys that sampled 269 sites. We then compared site-specific observed/predicted ratios of the number of species at sites to scores from a multimetric index of biotic integrity (IBI). Predicted numbers of species were moderately to strongly

correlated with the numbers observed by the four surveys. We found significant, though weak, relationships between observed/predicted ratios and IBI scores. SDM-based assessments identified patterns of local assemblage change that were congruent with IBI inferences; however, modeling artifacts that likely contributed to over-prediction of species presence may restrict the stand-alone use of SDM-derived patterns for bioassessment and therefore warrant examination. Our results suggest that when extensive standardized survey data that include reference sites are lacking, as is commonly the case, SDMs derived from generally much more readily available species site occupancy data could be used to provide a complementary tool for bioassessment.

Keywords Bioassessment · Community modeling · Conservation · Fish biodiversity · Species distribution modeling · Reference sites

Electronic supplementary material The online version of this article (doi:10.1007/s00267-015-0567-0) contains supplementary material, which is available to authorized users.

✉ Ben J. Labay
benlabay@utexas.edu

¹ Department of Integrative Biology, Biodiversity Collections, University of Texas, Austin, TX, USA

² Department of Biology/Aquatic Station, Texas State University, San Marcos, TX, USA

³ Department of Biology, Baylor University, Waco, TX, USA

⁴ River Studies Program, Inland Fisheries, Texas Parks and Wildlife Department, San Marcos, TX, USA

⁵ Department of Wildlife and Fisheries, Texas A&M University, College Station, TX, USA

Introduction

Bioassessment is an important element within efforts to protect, monitor, and restore natural resources. Bioassessment generally relies on benchmarks ideally designed to represent historical, pristine conditions (Karr 1981; Fausch et al. 1984) that, in today's world, are increasingly difficult to find. The degree to which an assessment is considered accurate depends upon how well benchmarks characterize inferred conditions along gradients of environmental degradation. Because it is rare to have accurate data for historical conditions, bioassessment benchmarks often rely on contemporary data reflecting varying degrees of alteration. The practice of applying contemporary, least-disturbed, reference sites as benchmarks is risky and generally

cannot be applied across broad spatial scales (Seegert 2000). This approach also can result in shifting baselines and further environmental degradation (Pauly 1995; Pinnegar and Engelhard 2008; Humphries and Winemiller 2009), especially in developed landscapes where ecological conditions deviate greatly from natural states. The reference site approach to bioassessment can suffer from problems and costs associated with quality and quantity of reference site data (Bowman and Somers 2005; Chessman et al. 2008; Hawkins et al. 2010), inconsistent reference definitions (Stoddard et al. 2006; Whittier et al. 2007; Herlihy et al. 2008), inability to identify causal mechanisms in multimetric (e.g., IBIs—Suter II 1993; Hawkins et al. 2010) and multivariate approaches (e.g., RIVPACS, AUSRIVAS—Dolédéc and Statzner 2010), and non-scalability or non-transferability of indices across regions and studies (Norris and Hawkins 2000; Cao and Hawkins 2011). Transferable metrics scalable to large regions that lack long-term survey data are needed and will require exploration of new methods that reduce reliance on least-disturbed reference sites as benchmarks (Dziocck et al. 2006; Dolédéc and Statzner 2010) and complement current bioassessment techniques.

Criticisms of bioassessment approaches emphasize methods used for estimating reference benchmarks (Suter II 1993; Karr and Chu 1999; Norris and Hawkins 2000; Seegert 2000). The two most common approaches are regionalization and discriminative modeling using environmental covariates. Hawkins et al. (2010) point out that regionalization generally has insufficient precision or numerical criteria to detect ecologically meaningful deviations, and discriminative models, exemplified by The River Invertebrate Prediction and Classification System (RIVPACS—Wright et al. 1984) and its derivatives (e.g., AUSRIVAS—Turak et al. 1999), rely on reference sites and standardization of sampling effort for data for calibration, and therefore do not allow extrapolation beyond the range of calibration data. Both of these shortcomings limit transferability, especially in degraded and data-poor regions. Recent bioassessment literature reviews (Dolédéc and Statzner 2010; Hawkins et al. 2010) identify Chessman and Royal's (2004) approach as an alternative that circumvents the need for reference sites (also see Speight and Castella 2001; Stranko et al. 2005; Chessman 2006; Labay et al. 2011). Chessman and Royal (2004) started with a known pool of potential colonists and used coarse-scale environmental parameters and known tolerances of species as filters to predict assemblage composition for any site. Rather than model historical assemblages, this approach simulates 'null' assemblages based on species' physiological tolerances and biogeographic limitations for comparisons to observed assemblages (Soberón 2007; Soberón and Nakamura 2009). We here use the term 'null' to describe

the predicted assemblages as they are designed for comparison to contemporary surveys, and in this way, it is used analogously to a null hypothesis.

Hawkins et al. (2010) considered Chessman and Royal's (2004) approach conceptually appealing but lacking sufficient validation. Although applications in the stream bioassessment literature are indeed scarce (but see Chessman 2006; Labay et al. 2011; Grown et al. 2013), major advances in the spatial modeling of biodiversity at the community level are enabling bioassessment without use of reference sites (Austin 2002; Scott 2002; Guisan and Thuiller 2005; Ferrier and Guisan 2006; Mateo et al. 2012). Species distribution models (SDMs) are becoming increasingly popular as predictors of biodiversity patterns (Ferrier and Guisan 2006; Costa et al. 2009; Guisan and Rahbek 2011; Peterson et al. 2011; Kuemmerle et al. 2012; Vasconcelos et al. 2012), in large part due to recent improvements in methods (Elith et al. 2006) and advancements in digitization of broad-scale environmental coverages (Turner et al. 2003) and biological occurrence databases (Guralnick and Van Cleve 2005).

SDMs constructed with large databases, such as those associated with major natural history collections, could be used in what Ferrier and Guisan (2006) identify as a "predict first, assemble later" approach to modeling communities. Compared to other community modeling approaches, this method has the advantage of producing individualistic species responses and allowing disparate surveys to be combined (Gioia and Pigott 2000; Guisan and Theurillat 2000; Lehmann et al. 2002), whereas existing bioassessment methods require consistent long-term community sampling across all sites to be assessed. Conceptually, the "predict first, assemble later" approach is not novel (see Guisan and Theurillat 2000; Olden 2003; Peppler-Lisbach and Schröder 2004; Gelfand et al. 2005; Leathwick et al. 2005; Baselga and Araujo 2010; Mateo et al. 2012); however, more research is needed to better determine its bioassessment applications.

Here we test the potential for SDMs to provide a basis for bioassessment of stream fishes in Texas, USA, using SDM-based modeled 'null' fish assemblages that we use as benchmarks for comparisons with contemporary assemblages. Specifically, we (i) produced SDMs for 100 fish species from major river basins of Texas; (ii) created local benchmark assemblages by stacking single-species SDMs for the potential species pool of our study area; (iii) compared the SDM-based assemblage predictions with contemporary fish assemblage data from four independent surveys, two with repetitive sampling protocols that permit relatively robust model evaluation, and two with associated multimetric-based indices of biotic integrity (IBI) scores based on methodologies used by most state and federal agency bioassessment efforts; and (iv) compared ratios of

observed/predicted assemblages and examined their relationship to IBI scores.

Methods

Study Area

The extent used for SDM development was the political boundary of the state of Texas divided into a grid of 931,808 30-arc-second cells (mean cell area 0.73 km²). This extent and resolution encompasses species occurrence records, environmental data used in modeling, and observations used for comparison (Fig. 1). The study region experiences an east–west annual precipitation range of approximately 152–20 cm, resulting in vastly different habitat conditions and a general decline in species richness from east to west (Hubbs et al. 2008).

Fish Occurrence Records for SDMs

Fish occurrences for SDM development were obtained from the recently compiled and publically accessible Fishes of Texas (FoTX) database maintained by Texas Natural History Collection (TNHC), University of Texas at Austin (Hendrickson and Cohen 2012). The FoTX database project compiles and verifies specimen records collected in Texas that are held in 33 museums worldwide (see www.fishesoftexas.org). Specimen records in TNHC are served to the publically accessible national Global Biodiversity Information Facility. We attempted to construct SDMs for all native fishes classified as freshwater by Hubbs et al. (2008) (see Table S1, Online Resource 1) and known to occur in the major Texas river basins for which we had

independent survey data. Non-natives were excluded, as were three species native to the state but known to have high introduction rates from stocking or bait bucket release (*Notemigonus crysoleucas*, *Morone chrysops*, and *Pimephales promelas*), and models were not attempted for species having <10 unique occurrence locations on the environmental layer grids (Phillips and Dudík 2008). Finally, three species groups determined by TNHC staff to have high misidentification rates or recent taxonomic revisions were combined, by aggregating individual species models, and treated as single species for predictions: (i) *Etheostoma lepidum* and *Etheostoma spectabile*, (ii) *Dionda episcopa* and *Dionda argentosa*, and (iii) species of the genus *Gambusia* (*G. affinis*, *G. amistadensis*, *G. georgei*, *G. gaigei*, *G. geiseri*, *G. heterochir*, *G. senilis*, *G. nobilis*, and *G. speciosa*).

Species occurrence records with an estimated georeferencing error (radius) >1 km were excluded from modeling to assure that input species occurrences corresponded in spatial resolution to environmental layers. Occurrence records before 1950 were excluded so that occurrence data were contemporaneous with climatic variables used (1950–2000; Table 1). Although such temporal limitation biases estimation of species' full historical ranges, the quality and size of the FoTX database to some degree alleviate this bias. Further, by modeling with variables not directly influenced by humans, estimated predictions are relatively unbiased with regards to species' historical ranges (Chessman and Royal 2004).

Environmental Variables

The environmental variables used in SDM construction (Table 1) were selected in part on the basis of expert opinion that ranked preliminary models using varied subsets of all available variables that were done for a set of species with well-known distributions (see Labay et al. 2011 for a detailed description). The climatic, hydrologic, and topographic variables were used to attempt to account for broad-scale physiological constraints as determinants of distributions (Graham and Hijmans 2006), and the two hydrology-based geographic variables control for historical zoogeography by categorically constraining predictions of species presence to watersheds from which they are documented. The three hydrological variables sourced from the National Hydrography Dataset Plus (mean annual flow, mean annual velocity, and cumulative drainage; USEPA and USGS 2005) were created by converting NHD catchments (avg. 2 km²) attributed with the select metrics to a 30-arc-second raster grid using ArcGIS polygon to raster tool.

All environmental variables are represented by a grid that covers the entire study area. The raw grid data were

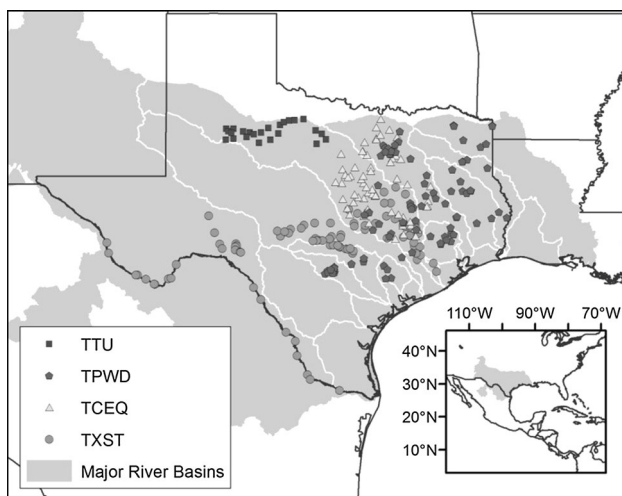


Fig. 1 Distribution of sample localities for the four survey datasets

Table 1 Environmental variables used in models

Layer category	Description	Source
Topological	Aspect	1 km DEM
Topological	Slope	1 km DEM
Topological	Compound topological index [$\ln(\text{acc. flow}/\tan[\text{slope}])$]	1 km DEM
Topological	Altitude	1 km DEM
Climate	Annual mean temperature	Wordclim variable 1
Climate	Mean diurnal range [mean of monthly (max temp – min temp)]	Wordclim variable 2
Climate	Isothermality (P2/P7)(*100)	Wordclim variable 3
Climate	Temperature seasonality (SD × 100)	Wordclim variable 4
Climate	Max temperature of warmest month	Wordclim variable 5
Climate	Min temperature of coldest month	Wordclim variable 6
Climate	Temperature annual range (P5–P6)	Wordclim variable 7
Climate	Annual precipitation	Wordclim variable 12
Climate	Precipitation of wettest month	Wordclim variable 13
Climate	Precipitation of driest month	Wordclim variable 14
Climate	Precipitation seasonality (coefficient of variation)	Wordclim variable 15
Climate	Precipitation of wettest quarter	Wordclim variable 16
Climate	Precipitation of driest quarter	Wordclim variable 17
Climate	Precipitation of warmest quarter	Wordclim variable 18
Climate	Precipitation of coldest quarter	Wordclim variable 19
Geographic	Major river basins	Texas Water Development Board
Geographic	8-Digit hydrologic unit code (HUC)	United States Geologic Survey
Hydrologic	Cumulative drainage	National Hydrology Dataset plus
Hydrologic	Mean annual flow	National Hydrology Dataset plus
Hydrologic	Mean annual velocity	National Hydrology Dataset plus

used for modeling, as opposed to a stream-segment based framework, to retain maximum spatial, and thus data, resolution. In a stream-segment-based framework, attributing environmental data to stream network segments would require aggregation of the raw data at a catchment-scale [<1 km cell data (30 arc-seconds) into a ~ 2.5 km² NHD catchment]. Modeled probability estimates from the SDMs were extracted from the same data cell in which the independent survey data record occurred (see SDM construction and evaluation).

SDM Construction and Evaluation

Models were constructed using the maximum entropy algorithm encoded in the Maxent software package (Version 3.3.4, Phillips et al. 2006), known to be robust for species distribution modeling with presence-only records (Elith et al. 2006) and recently shown to be nearly mathematically equivalent to a Poisson regression model (Renner and Warton 2013). We restricted ourselves to Maxent and refrained from methodological comparisons involving multiple modeling algorithms, parameterizations, and validation procedures in order to focus on our

primary objective of assessing a specific SDM application in bioassessment using common methods. To that aim, we implemented Maxent with default parameterization recommendations (Phillips and Dudík 2008), with models replicated 100 times randomly withholding in each replicate 40 % of localities as ‘test’ records, with the remaining 60 % serving as model ‘training’ records. Individual species’ model performance was evaluated using a receiver operating characteristic (ROC) analysis. The ROC analysis characterizes model performance at all possible thresholds using the area under the curve (AUC). A model with perfect discrimination would have an AUC of 1 while a model that predicted species occurrences at random would have an AUC of 0.5 (Hanley and McNeil 1982). We here recognize that validation using AUC, while intuitively appealing, has been criticized as misleading as it tends to overestimate model quality, may show spurious high performance with small sample sizes, and may reward overparameterization (Lobo et al. 2008). We ultimately utilize the AUC statistic (albeit conservatively, see below) for its commonality and the fact that for most places in the world the kind of data needed over broad scales for better validation are not available. Thus, our models are internally

validated through resubstitution (AUC statistic) and not through independent, reference (historical) data, which are not available for our study area. Exploring how the model predictions compare to contemporary observations across gradients of effort and richness (the four independent survey datasets) is intended to provide approximations of model accuracy and precision.

Species models were considered reliable and retained for modeled community construction if they had an average test AUC over 100 replicates >0.9 and a $<5\%$ difference between average test and training AUC (AUC_{diff}). Despite the maximum entropy algorithm's accounting for correlations among variables, using a large number of variables raises dangers of over-fitting. Utilizing both the AUC_{diff} criterion and the variable set-selection process involving expert opinion conservatively reduces the risk of over-fitting (Warren and Seifert 2010).

For each grid cell, Maxent calculates a continuous probability of occurrence estimate ranging from 0 to 1. To assign species to a predicted community for any particular site (grid cell corresponding to a site sampled in an independent survey), we converted continuous modeled probability estimates to a binary prediction of presence or absence using the minimum training presence threshold (MTP; Pearson et al. 2007; Raxworthy et al. 2007), which is the lowest probability associated with the occurrences used in model construction. This species-specific threshold is used because of its tendency to recognize habitats that are not necessarily the most suitable but nevertheless utilized, providing a broad estimate of potential colonization. Thus, while presence-only modeling, as done here, does not account for estimates of species prevalence in the landscape and thus does not estimate true probabilities (Elith et al. 2011), the use of the MTP threshold is an attempt to account for differential prevalence among species (see Fig. S1, Online Resource 1). This species-specific threshold approach has been shown to produce better predictions across species than does the use of a single threshold (Liu et al. 2005) and has the benefit of allowing the assessment of compositional dissimilarity between observations and predictions that would not be possible if simple richness predictions were used. Additionally, when using Maxent for SDM construction, the use of a threshold has been shown to be more appropriate compared to summing probability values (Cao et al. 2013).

Independent Fish Survey Data

Results from four independent surveys of fishes in wadeable streams were compared to model predictions. Survey data were independent of the museum data used to create SDMs and are generally more recent (post 1998) compared to fish data used to construct models (1950–2000). The four

surveys differed in location (Fig. 1), objectives and methodologies (Table 2), and together sampled 269 unique sites. We refer to the surveys either by the institution of the collectors or by the agency responsible for funding the surveys: Texas Tech University (TTU), Texas State University—San Marcos (TXST), Texas Commission on Environmental Quality (TCEQ), and Texas Parks and Wildlife Department (TPWD).

While methodological differences among surveys caused us to analyze them separately, we provide general conclusions and discussion based on two survey categories. We posit that by having sampled repeatedly at each site over extended periods, TTU and TXST should provide more reliable estimates of species presence/absence at each site and therefore better allow for validation of the modeled assemblages based on stacked SDMs. In contrast, TCEQ and TPWD had fewer replicate samples (none in the case of TPWD), but used standardized sampling methodologies characteristic of state and federal agency bioassessments (Linam et al. 2002) and provided multimetric-based IBI scores (Karr 1981; Fausch et al. 1984). Thus, data from TCEQ and TPWD surveys are appropriate for exploration of the utility of our model-based approach as an alternative or complement to current bioassessment methods.

Comparison of Observed and Predicted

We calculated statistics to help evaluate ability of stacked SDMs to predict local species assemblages. For the two surveys with greater number and frequency of samples (TTU and TXST), we calculated average species model sensitivity (% of occurrences that were successfully predicted) and specificity (% of non-occurrences that were successfully predicted). For each survey, we compared the number of predicted species observed and the number of model-predicted species, hereafter referred to as observed/predicted, for number of species per surveyed site. The observed/predicted metric varies from 0 to 1 and allows for assessing compositional deviance from models among sites. We assessed the linear regression relationship of observed/predicted for each survey. Finally, for numbers of species observed/predicted per site for each survey, we calculated Theil's inequality coefficient (U),

$$U = \frac{\sqrt{\frac{1}{n} \sum (y_i - x_i)^2}}{\sqrt{\frac{1}{n} \sum y_i^2 + \frac{1}{n} \sum x_i^2}},$$

where y_i is the number of predicted species observed and x_i is the model prediction. Theil's inequality coefficient provides a measure of distance between predicted values and observed values (Piñeiro et al. 2008), ranging from 0 to 1 with 0 indicating perfect agreement. This coefficient is

Table 2 Independent fish survey dataset comparison

Dataset	TTU	TXST	TCEQ	TPWD
Contributing institution	Texas Tech University, Lubbock	Texas State University—San Marcos	Texas Commission on Environmental Quality	Texas Parks and Wildlife Department
Background and purpose of survey as it pertained to fishes	To explore reproductive ecology of select cyprinids and population dynamics of fishes within the Upper Brazos River drainage	Multiple independent projects conducted to assess spatial–temporal trends in fish assemblage composition in various river drainages in Texas	Assess trait–environmental relationships in fish assemblages and assess biotic integrity of central and north Texas streams	To determine biotic integrity of east Texas streams
Relevant publications	Wilde (2011)	Bean et al. (2007), Labay (2010), Heard et al. (2012), Kollaus and Bonner (2012)	King et al. (2009), Winemiller et al. (2009), and Pease et al. (2011)	Kleinsasser et al. (2004)
No. sites	24	90	64	91
Year range	2008–2011	2003–2011	2006–2008	1998–2000
Gear types	Seines	Seines and electrofishing	Seines and electrofishing	Seines and electrofishing
Average no. samples per site	7.95	3.68	2.72	1
Average no. samples per year per site	2	3.92	0.91	1
Primary sampling interval	May and August–September	Dependent on study, varies from monthly to 3/year: Feb–Mar, Jun–Sept, Nov–Jan	June through August	May through October
Site selection protocol	Wadeable streams that captured representative habitats throughout target watersheds	Perennial, wadeable streams that captured representative habitats throughout target watersheds	Perennial, wadeable streams that captured representative habitats throughout target ecoregions	Perennial, wadeable streams selected by randomized systematic design: equal No. 2nd, 3rd, and 4th order streams and 30 sites in urban setting

similar to the correlation coefficient but is preferred in this case for its incorporation of slope (Smith and Rose 1995), and thus a goodness-of-fit measure that allows comparison against a theoretically perfect (1:1) relationship.

Relationships between observed/predicted ratios and IBI scores provided in the TPWD and TCEQ survey reports were examined using linear regression within ecoregion groupings described in a 2002 Texas Parks and Wildlife Department study (Linam et al. 2002). The TPWD and TCEQ surveys together sampled in four of Omernik's (1987) ecoregions, which were combined for region-specific metric calibration into two pairs of ecoregion aggregates due to similarities in fish assemblage composition (Linam et al. 2002). The two groups were ecoregions 33 (East Central Texas Plains) and 35 (South Central Plains), hereafter ecoregion 33/35, and ecoregions 29 (Cross Timbers) and 32 (Texas Blackland Prairies), hereafter ecoregion 29/32.

Additionally, to assess the relationship between observed/predicted ratios and IBI with regards to the region's state agency-determined water body beneficial use designations, we transformed IBI scores for all sites in the TCEQ and TPWD datasets into aquatic life use (ALU)

designations (Exceptional, High, Intermediate, and Limited; see Linam et al. 2002 for details) that are widely used by regional state agencies. A one-way ANOVA followed by Tukey's test for unequal sample sizes was used to test for differences ($\alpha = 0.05$) in observed/predicted ratios among the four ALU designations. These analyses were performed with the free software R, version 2.15.1 (R Development Core Team 2012).

Results

From the 131 freshwater fish species or species groups known from the study area, SDMs for 100 met our a priori criteria for model quality (>10 unique geographic records, avg. test AUC >0.9 , and $AUC_{diff.} <5\%$; Table S1). The 31 species for which we were unable to create reliable models are extinct (*Notropis orca* and *Notropis simus simus*), extirpated from our region (*Hybognathus amarus*), had too few precisely georeferenced occurrence points matching unique environmental grid units (e.g., *Macrhybopsis storeriana*, *Notropis chihuahua*, *Pteronotropis hubbsi*, *Cyprinodon eximius*, and *Etheostoma fonticola*), or are

wide-ranging but with low occurrence density in the FoTX database (e.g., *Anguilla rostrata*, *Cycleptus elongatus*, *Atractosteus spatula*, *Lepisosteus osseus*, *Ameiurus melas*, and *Ictalurus furcatus*). Nine of the 31 species not modeled, but observed in the independent surveys (Table S1), were excluded from all analyses.

For TTU and TXST sites, SDM sensitivity averaged 78 and 75 %, respectively, and specificity averaged 79 and 82 %. The linear relationship between number of observed and predicted species per site was significant for all four datasets, with better model fit and steeper slopes for datasets with greater frequency of sampling and fewer potential species (TTU > TXST > TCEQ > TPWD; Fig. 2). Intercepts were closer to zero for TTU and TXST than for TPWD and TCEQ (Fig. 2). Theil’s inequality coefficient was positively correlated with number of sites and sampling frequency, TTU ($U = 0.28$); TXST (0.32); TCEQ (0.33); TPWD (0.47).

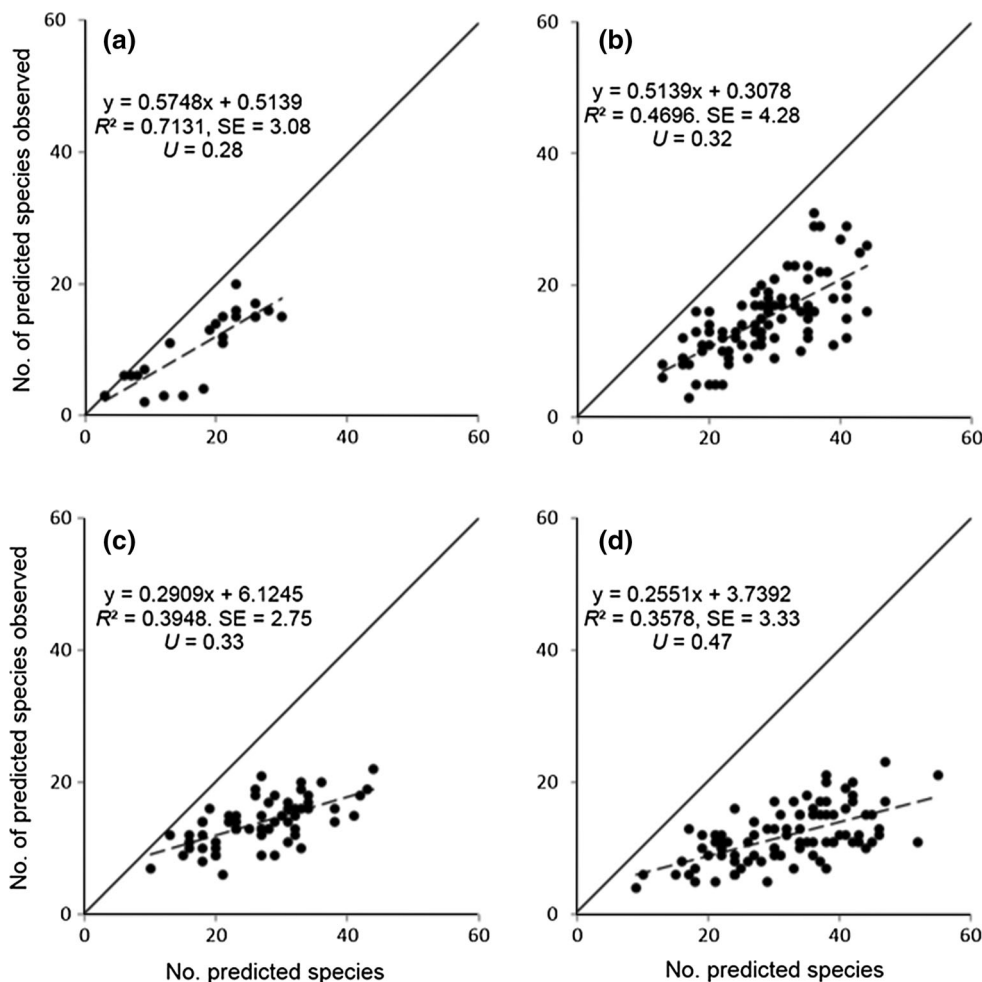
Positive relationships between observed/predicted ratios and IBI scores were detected across all ecoregions sampled by TCEQ and TPWD, with the exception of ecoregion

33/35 for TCEQ ($R = 0.08$, $P = 0.78$; Fig. 3). Observed/predicted ratios differed significantly ($F [3,151] = 6.43$, $P = 0.0004$) among the four ALU designations, showing a decline in mean observed/predicted ratio from “Exceptional” to “Limited,” with these two extremes being different from all other pairwise groupings (Fig. 4) except for a non-significant ($P = 0.16$) difference between “Limited” and “Intermediate.” The two middle groupings, “High” and “Intermediate,” did not differ ($P = 0.77$).

Discussion

This study explored a method for bioassessment that compares modeled estimates of species presence to occurrences from independent field surveys. Model quality varied, but robust SDMs were produced for 92 % of native species (or species groupings) observed by the field surveys. Model-predicted species assemblages at the 269 independently surveyed sites were significantly correlated with the assemblage samples from the surveys. Correlations were

Fig. 2 Relationship between numbers of native fish species predicted and observed for all sites sampled by **a** TTU, **b** TXST, **c** TCEQ, and **d** TPWD surveys. The *solid line* represents equality of predicted and observed species presence. Regression lines fitted to the data are presented with adjusted regression estimates (R^2), standard errors of the regression estimates (SE), and Theil’s inequality coefficient (U)



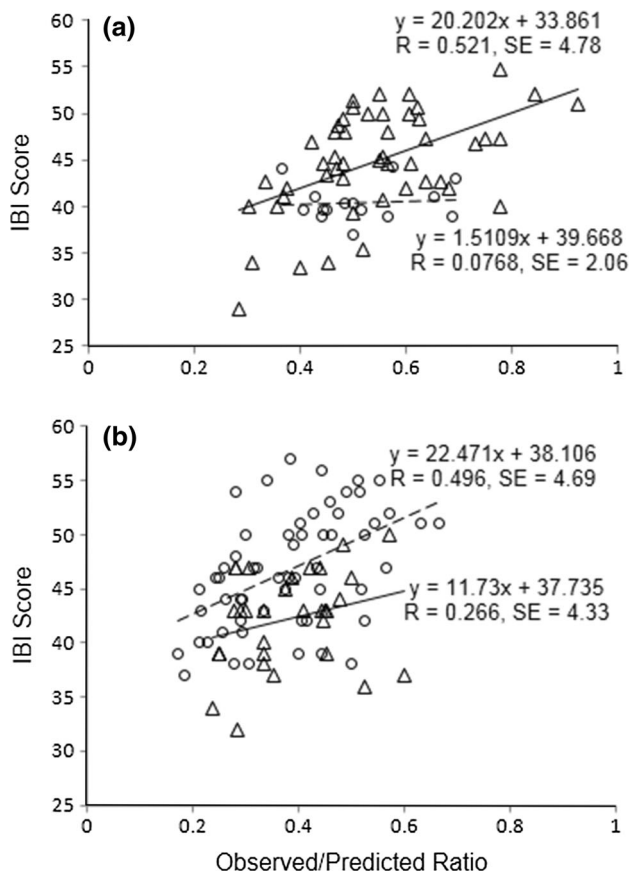


Fig. 3 Relationship between index of biotic integrity scores and observed/predicted ratios for the **a** TCEQ and **b** TPWD surveys. Sites in ecoregion 29/32 are represented by *triangles* with a *solid* linear regression line, and sites in ecoregion 33/35 are represented by *circles* with a *dashed* linear regression line. All within-ecoregion grouping relationships are significant ($\alpha = 0.05$) with the exception of TCEQ ecoregion 33/35. Regression lines fitted to data are presented with regression estimates (R) and standard errors of the regression estimates (SE)

strongest at sites that were repeatedly sampled (3–4 times per year; TTU and TXST datasets), and, as expected, we saw a tendency for stacked models to over-predict richness compared to survey observations, especially in more species-rich eastern parts of the study area and represented by the TCEQ and TPWD datasets. The term ‘over-prediction’ in this case is used to align with terminology within established modeling literature (Piñeiro et al. 2008). Although this tendency to over-predict may be due in part to environmental degradation, over-prediction has been documented in other studies using stacked SDMs to estimate species richness and co-occurrence (Fielding and Bell 1997; Chessman 2006; Mateo et al. 2012; Calabrese et al. 2013).

Apart from potential deviation resulting from environmental degradation, we suspect that over-prediction in this study stems primarily from statistical, model-driven effects. Better understanding of contributing modeling

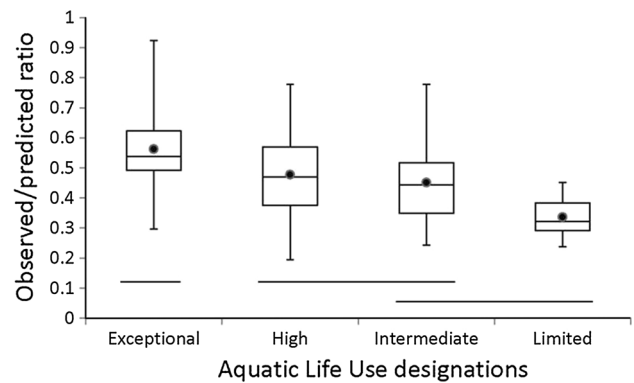


Fig. 4 Boxplot with minimum, 1st quartile, median, 3rd quartile, and maximum of observed/predicted ratios for each aquatic life use (ALU) designation defined by Linam et al. (2002) for TPWD and TCEQ survey sites. Mean is indicated by the *solid black circle*. *Solid horizontal lines* indicate groupings of aquatic life use categories that are significantly different from one another as determined by Tukey’s pairwise tests

artifact factors, especially threshold choice (Calabrese et al. 2013) and model parameter selection, and how these influence patterns of over-fitting should improve accuracy and precision of probability estimates. Additionally, our models do not directly account for all factors that influence assemblage structure at the local level, and therefore the assemblages resulting from model stacking should be interpreted as ‘null’ models of assemblage composition that accounts only for species’ responses to specific abiotic environmental gradients used in model construction. Two factors contributing to over-prediction that the models do not directly account for are biotic interactions (e.g., competitive exclusion, and predation) and differential patterns of species-specific prevalence, and/or detection.

Accounting for effects of biotic interactions in SDMs remains a major challenge (Araújo and Luoto 2007). Ambiguities remain regarding the definition of species niche space for practical applications of SDMs (Soberón 2007), the role of biotic interactions in niche space, and how to parameterize models to account for these ambiguities at various scales (Araújo and Guisan 2006). Similar to this study, Mateo et al. (2012) used stacked SDMs to successfully estimate patterns of species richness across environmental gradients and, as in this study, found greater over-prediction in areas of high richness. They attributed this to the fact that most SDMs produced to date rely predominantly on abiotic variables that exert greater control over biodiversity in harsh climates (e.g., deserts or alpine ecosystems) than do biotic interactions (Pineda and Lobo 2009; Pellissier et al. 2012). As our knowledge of species’ fundamental niches and modeling methods improve, it should be possible to better account for biotic interactions in SDMs (Brooker et al. 2007; Heikkinen et al. 2007; Bateman et al. 2012).

That our models did not directly account for species prevalence likely also contributed to over-prediction, especially in the eastern, more species-rich areas of the study region. Explicitly accounting for species-specific prevalence is difficult without quality, comparable presence/absence data over long-time periods, which, if obtained, would allow more accurate statistical models compared to those obtained from presence-only SDMs. The minimum training presence (MTP) threshold was used in this study to attempt to indirectly account for species-specific prevalence (see Fig. S1, Online Resource 1), and more generally, the influence of imperfect detection of species. In effect, this threshold operationalized the idea that ubiquitous taxa, or habitat generalists, are likely to be observed over a broader range of habitat suitability than are less prevalent specialist taxa that are constrained to relatively narrow ranges of habitat suitability (Peterson et al. 2011). Thus the MTP threshold, although generally broad, partially accounts for documented assemblage response to influences of dispersal and mass effect (Hitt and Angermeier 2008, 2011) for widespread taxa, while tending to restrict presence predictions for less prevalent taxa to their most suitable environmental conditions. This effect also tends to increased over-prediction in more species-rich East Texas. More studies are needed that compare accuracy of stacked SDM richness predictions across varying thresholds and richness gradients.

Regardless of the mechanisms, differential patterns of over-prediction across environmental or species-richness gradients confound attempts to interpret ratios of observed/predicted species. Given varied model prediction success across a range of environmental conditions, predictions should be most accurate when surveys are comprehensive. Of the four surveys examined here, TTU had the lowest species richness together with the highest collection effort. This dataset also produced the highest correspondence between modeled and observed species assemblages, and provides a fairly robust opportunity to explore in more detail how differences between model prediction and observations could be related to sampling error, variation in community composition due to natural or anthropogenic environmental variation, or model-construction artifacts. Several obligate fluvial specialists in the TTU survey dataset were predicted to occur at sites where they were not captured—*Notropis potteri* at 10 sites, *Notropis shumardi* and *Notropis volucellus* at 8 and 5 sites, respectively—yet all 3 have historical records in the area surveyed by TTU (Hendrickson and Cohen 2012). Conversely, some species not predicted and with no historical occurrence records in the area were observed—e.g., *Aplodinotus grunniens* was collected at 10 sites but predicted at 1. We have not determined if it has recently expanded into the upper Brazos River drainage or whether historical sampling efforts simply failed to document it.

In addition to excluding biotic factors, we are aware that the environmental variables used in the models are a subset of all possible abiotic dimensions of fish niches. Saline springs in the Salt Fork watershed of the upper Brazos River basin (Dutton 1989) have been previously noted to limit fish community composition to a handful of salinity-tolerant forms (Hubbs 1957; Ostrand and Wilde 2002), and this might contribute to over-prediction of species presence by our models at four sites in the middle to lower reaches of the Salt Fork surveyed by TTU (Fig. 2a). Unfortunately, we know of no salinity coverages for our study area that would allow us to include salinity in our models.

Given that we are comparing our models, with their previously stated limitations, to the IBI, the basic assumptions of which have been widely criticized and the index noted to filled with ‘noise’ (Suter II 1993; Seegert 2000), it is notable that the overall relationships between our observed/predicted index and IBI scores are significant, albeit admittedly weak. Despite patterns of over-prediction and lack of accounting for all factors determining species distributions, our model predictions provided benchmarks that correlated with prior IBI-derived ALU groupings, effectively distinguishing sites with “excellent,” “average,” and “poor” integrity. Not being able to segregate between the “limited” and “intermediate” designations was likely due to low power related to small sample size ($n = 6$) for the “limited” designation. Failure to segregate between the two middle groupings, “high” and “intermediate,” is consistent with criticisms of the IBI (Suter II 1993; Seegert 2000) suggesting that the multimetric approach is primarily useful in identifying extreme cases where constituent variables are either all low or all high. This lack of detailed discrimination is caused by aggregation of multiple variables which could produce similar IBI scores for highly divergent conditions, effectively obscuring the true condition of the biotic community (Suter II 1993). Lack of stronger concordance could be due, at least in part, to the inadequacy of both IBIs and models to account for all factors that restrict species distributions, so that, consequently, they likely measure different dimensions of habitat and fish assemblage integrity. For example, many of the IBI metrics measure abundances of species or guilds, and even though SDMs have been shown to correlate with species abundance (VanDerWall et al. 2009), they do not explicitly measure abundance.

As a result of the modeling methods and construction decisions (presence-only data, limitation of occurrence records to post-1950, no accounting for biological interactions), the list of species predicted to occur at a given site does not represent an accurate estimate of the historical assemblage, but instead represents a ‘null’ community estimated from best available data. While this over-prediction might limit the stand-alone use of this technique as

a bioassessment tool, it provides managers with a target species list that allows for evaluation of current assemblages.

Conclusions

Interpretation of most bioassessment benchmarks is compromised by the fact that agreed-upon baselines, reference conditions, and/or validation datasets are almost always lacking in highly developed landscapes. Establishment of reference conditions with currently available methodologies is challenging for regions where watersheds in historical condition are rare or absent. We have demonstrated potential for readily available specimen-based datasets compiled by natural history collections to be used with a stacked SDM approach to quickly provide a landscape-scale index useful as a species composition baseline for comparison against contemporary assemblage data, and complementary to standard, regionalized bioassessment techniques such as the IBI or RIVPACS.

However, as with most applications based on modeled systems, conclusions drawn from this technique are subject to data and methodological limitations. If this approach is to be considered as an alternative to traditional bioassessment techniques (e.g., IBI or RIVPACS), additional research is needed to account for methodological uncertainties in probability estimates before this approach can accurately and confidently identify assemblage deviations due solely to anthropogenic influences. However, this approach can immediately offer a complementary perspective allowing resource managers to produce target lists of species based on quantitative and transferable estimation methods, and generate hypotheses regarding specific species exclusions. As an example for implementation, the produced list of estimated species can be compared to what was actually caught, and specific taxa or guild exclusions can be incorporated as individual metrics into a modified IBI composite metric. In other words, IBI-like metrics could benefit from a model-based benchmark of species assemblage composition that could serve as a heuristic tool generating useful species-specific exclusion hypotheses. Currently, for our study region, IBIs are calibrated and restricted to implementation, at an ecoregion scale. Incorporating this model-based methodology could allow calibration at the scale of the SDMs used, and thus automatically incorporate watershed-specific faunal differences related to biogeographic history. Managers could also use this method to improve assessments of whether current reference sites, often used as benchmarks for restoration goal-setting, provide reasonable standards with respect to expected fish community composition.

These methods, and the assemblage predictions they produce, are easily performed at large scales, and the many current initiatives to improve and increase availability of occurrence data and the steadily increasing diversity of environmental coverages, will improve such model-based predictions moving forward. We encourage management agencies to consider adopting this SDM-based bioassessment framework alongside traditional biomonitoring and assessment protocols to start taking advantage of its ability to provide pre-defined target species lists, quantified probabilities of occurrence estimates, and tractable data to help generate and further explore hypotheses about mechanisms controlling species' distributions likely to also help refine the methodology. In practice, integration of this approach could help reduce the widely recognized susceptibility of bioassessment and ecological restoration to shifting baselines.

Acknowledgments This study was supported by The University of Texas at Austin, Texas Parks and Wildlife Department State Wildlife Grant (Data standardization and georeferencing of the Fishes of Texas database, F06AF00007) made available through the United States Fish and Wildlife Service's State Wildlife Grant program (T-106), and Texas Commission on Environmental Quality (Digital Fish Atlas grant contract No. 582-11-99736). Thanks to the thousands of collectors who contributed to the Fishes of Texas database. We also thank Gary Garrett, Tim Birdsong, Kevin Mayes, and Josh Perkin for early discussions and concept reviews, 6 anonymous reviewers, and Meagan Bean, Preston Bean, Steven Curtis, Thom Heard, Douglas Knabe, Kristy Kollaus, Doug Martin, Allison Pease, Zach Shattuck, Jason Taylor, and Gene Wilde for help in compiling portions of the survey databases used in our analyses.

Conflict of interest The authors declare that they have no conflicts of interest.

References

- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *J Biogeogr* 33:1677–1688
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Glob Ecol Biogeogr* 16:743–753
- Austin M (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157:101–118
- Baselga A, Araujo MB (2010) Do community-level models describe community variation effectively? *J Biogeogr* 37:1842–1850
- Bateman BL, VanDerWal J, Williams SE, Johnson CN (2012) Biotic interactions influence the projected distribution of a specialist mammal under climate change. *Divers Distrib* 18:861–872
- Bean PT, Bonner TH, Littrell BM (2007) Spatial and temporal patterns in the fish assemblage of the Blanco River, Texas. *Tex J Sci* 59:179
- Bowman M, Somers K (2005) Considerations when using the reference condition approach for bioassessment of freshwater ecosystems. *Water Qual Res J Can* 40:347–360

- Brooker RW, Travis JMJ, Clark EJ, Dytham C (2007) Modelling species' range shifts in a changing climate: the impacts of biotic interactions, dispersal distance and the rate of climate change. *J Theor Biol* 245:59–65
- Calabrese JM, Certain G, Kraan C, Dormann CF (2013) Stacking species distribution models and adjusting bias by linking them to macroecological models. *Glob Ecol Biogeogr* 23(1):99–112
- Cao Y, Hawkins CP (2011) The comparability of bioassessments: a review of conceptual and methodological issues. *J N Am Benthol Soc* 30:680–701
- Cao Y, DeWalt RE, Robinson JL, Tweddale T, Hinz L, Pessino M (2013) Using Maxent to model the historic distributions of stonefly species in Illinois streams: the effects of regularization and threshold selections. *Ecol Model* 259:30–39
- Chessman BC (2006) Prediction of riverine fish assemblages through the concept of environmental filters. *Mar Freshw Res* 57:601–609
- Chessman BC, Royal MJ (2004) Bioassessment without reference sites: use of environmental filters to predict natural assemblages of river macroinvertebrates. *J N Am Benthol Soc* 23:599–615
- Chessman B, Muschal M, Royal M (2008) Comparing apples with apples: use of limiting environmental differences to match reference and stressor-exposure sites for bioassessment of streams. *River Res Appl* 24:103–117
- Costa GC, Nogueira C, Machado RB, Colli GR (2009) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodivers Conserv* 19:883–899
- Dolédéc S, Stutzner B (2010) Responses of freshwater biota to human disturbances: contribution of J-NABS to developments in ecological integrity assessments. *J N Am Benthol Soc* 29:286–311
- Dutton AR (1989) Hydrogeochemical processes involved in salt-dissolution zones, Texas panhandle, U.S.A. *Hydrol Process* 3:75–89
- Dziok F, Henle K, Foeckler F, Follner K, Scholz M (2006) Biological indicator systems in floodplains: a review. *Int Rev Hydrobiol* 91(4):271–291
- Eliith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A et al (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Eliith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17(1):43–57
- Fausch KD, Karr JR, Yant PR (1984) Regional application of an index of biotic integrity based on stream fish communities. *Trans Am Fish Soc* 113:39–55
- Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *J Appl Ecol* 43:393–404
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Gelfand AE, Schmidt AM, Wu S, Silander JA Jr, Latimer A, Rebelo AG (2005) Modelling species diversity through species level hierarchical modelling. *J R Stat Soc* 54:1–20
- Gioia P, Pigott JP (2000) Biodiversity assessment: a case study in predicting richness from the potential distributions of plant species in the forests of south-western Australia. *J Biogeogr* 27:1065–1078
- Graham CH, Hijmans RJ (2006) A comparison of methods for mapping species ranges and species richness. *Glob Ecol Biogeogr* 15:578–587
- Gronws I, Rourke M, Gilligan D (2013) Toward river health assessment using species distributional modeling. *Ecol Indic* 29:138–144
- Guisan A, Rahbek C (2011) SESAM—a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *J Biogeogr* 38:1433–1444
- Guisan A, Theurillat JP (2000) Equilibrium modeling of alpine plant distribution: how far can we go? *Phytocoenologia* 30:353–384
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8:993–1009
- Guralnick R, Van Cleve J (2005) Strengths and weaknesses of museum and national survey data sets for predicting regional species richness: comparative and combined approaches. *Divers Distrib* 11:349–359
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29
- Hawkins CP, Olson JR, Hill RA (2010) The reference condition: predicting benchmarks for ecological and water-quality assessments. *J N Am Benthol Soc* 29:312–343
- Heard T, Perkin JS, Bonner TH (2012) Intra-annual variation in fish communities and habitat associations in a Chihuahua Desert reach of the Rio Grande/Rio Bravo Del Norte. *West N Am Nat* 72:1–15
- Heikkinen RK, Luoto M, Virkkala R, Pearson RG, Körber JH (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Glob Ecol Biogeogr* 16:754–763
- Hendrickson DA, Cohen AE (2012) Fishes of Texas project and online database. Published by Texas Natural History Collection, a division of the Department of Integrative Biology, University of Texas at Austin
- Herlihy AT, Paulsen SG, Sickle JV, Stoddard JL, Hawkins CP, Yuan LL (2008) Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *J N Am Benthol Soc* 27:860–877
- Hitt NP, Angermeier PL (2008) Evidence for fish dispersal from spatial analysis of stream network topology. *J N Am Benthol Soc* 27:304–320
- Hitt NP, Angermeier PL (2011) Fish community and bioassessment responses to stream network position. *J N Am Benthol Soc* 30:296–309
- Hubbs C (1957) Distributional patterns of Texas fresh-water fishes. *Southwest Nat* 2:89–104
- Hubbs C, Edwards RJ, Garrett GP (2008) An annotated checklist of the freshwater fishes of Texas, with keys to identification of species. *Tex J Sci* 43:1–87
- Humphries P, Winemiller KO (2009) Historical impacts on river Fauna, shifting baselines, and challenges for restoration. *Bio-science* 59:673–684
- Karr JR (1981) Assessment of biotic integrity using fish communities. *Fisheries* 6:21–27
- Karr JR, Chu EW (1999) Restoring life in running waters: better biological monitoring. Island Press, Washington, D.C.
- King RS, KO Winemiller, JM Taylor, JA Back, A Pease (2009) Development of biological indicators of nutrient enrichment for application in Texas streams. Final report to Texas Commission on Environmental Quality, §106 Water Pollution Control Grant #98665304
- Kleinsasser LJ, Jurgensen T, Bowles D, Boles S, Aziz K, Saunders K, Linam G, Tringale J, Mayes K, Rector J et al (2004) Status of biotic integrity, water quality, and physical habitat in Wadeable East Texas streams. Resources Protection Division, Texas Parks and Wildlife Department River Studies Report 19
- Kollaus KA, Bonner TH (2012) Habitat associations of a semi-arid fish community in a karst spring-fed stream. *J Arid Environ* 76:72–79
- Kuemmerle T, Hickler T, Olofsson J, Schurgers G, Radeloff VC (2012) Reconstructing range dynamics and range fragmentation of European bison for the last 8000 years. *Divers Distrib* 18:47–59
- Labay B (2010) The influence of land use, zoogeographic history, and physical habitat on fish community diversity in the lower Brazos

- watershed. Theses and Dissertations-Biology 27, Texas State University, San Marcos
- Labay B, Cohen AE, Sissel B, Hendrickson DA, Martin FD, Sarkar S (2011) Assessing historical fish community composition using surveys, historical collection data, and species distribution models. *PLoS One* 6:e25145
- Leathwick JR, Rowe D, Richardson J, Elith J, Hastie T (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshw Biol* 50:2034–2052
- Lehmann A, Leathwick J, Overton JMC (2002) Assessing New Zealand fern diversity from spatial predictions of species assemblages. *Biodivers Conserv* 11:2217–2238
- Linam GW, Kleinsasser LJ, Mayes KB (2002) Regionalization of the index of biotic integrity for Texas streams. Resources Protection Division, Texas Parks and Wildlife Department River Studies Report 17
- Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385–393
- Lobo JM (2008) More complex distribution models or more representative data? *Biodivers Inform* 5:14–19
- Mateo RG, Felicísimo ÁM, Pottier J, Guisan A, Muñoz J (2012) Do stacked species distribution models reflect altitudinal diversity patterns? *PLoS One* 7:e32586
- Norris RH, Hawkins CP (2000) Monitoring river health. *Hydrobiologia* 435:5–17
- Olden JD (2003) A species-specific approach to modeling biological communities and its potential for conservation. *Conserv Biol* 17:854–863
- Omernik JM (1987) Ecoregions of the conterminous United States. *Ann Assoc Am Geogr* 77:118–125
- Ostrand KG, Wilde GR (2002) Seasonal and spatial variation in a prairie stream-fish assemblage. *Ecol Freshw Fish* 11:137–149
- Pauly D (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends Ecol Evol* 10:430
- Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr* 34:102–117
- Pease AA, Taylor JM, Winemiller KO, King RS (2011) Multiscale environmental influences on fish assemblage structure in central Texas streams. *Trans Am Fish Soc* 140:1409–1427
- Pellissier L, Pradervand JN, Pottier J, Dubuis A, Maiorano L, Guisan A (2012) Climate-based empirical models show biased predictions of butterfly communities along environmental gradients. *Ecography* 35:684–692
- Peppler-Lisbach C, Schröder B (2004) Predicting the species composition of *Nardus stricta* communities by logistic regression modelling. *J Veg Sci* 15:623–634
- Peterson AT, Soberon J, Pearson RG, Anderson RP, Martinez-Meyer E, Nakamura M, Araujo M (2011) Ecological niches and geographic distributions. Princeton University Press, New Jersey
- Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190:231–259
- Pineda E, Lobo JM (2009) Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *J Anim Ecol* 78:182–190
- Piñeiro G, Perelman S, Guerschman JP, Paruelo JM (2008) How to evaluate models: observed vs. predicted or predicted vs. observed? *Ecol Model* 216:316–322
- Pinnegar JK, Engelhard GH (2008) The “shifting baseline” phenomenon: a global perspective. *Rev Fish Biol Fish* 18:1–16
- R Development Core Team (2012) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG (2007) Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Syst Biol* 56:907–923
- Renner IW, Warton DI (2013) Equivalence of MAXENT and poisson point process models for species distribution modeling in ecology. *Biometrics* 69:274–281
- Scott JM (2002) Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C.
- Seeger G (2000) The development, use, and misuse of biocriteria with an emphasis on the index of biotic integrity. *Environ Sci Policy* 3:51–58
- Smith EP, Rose KA (1995) Model goodness-of-fit analysis using regression and related techniques. *Ecol Model* 77:49–64
- Soberón J (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecol Lett* 10(12):1115–1123
- Soberón J, Nakamura M (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proc Natl Acad Sci* 106(Supplement 2):19644–19650
- Speight MCD, Castella E (2001) An approach to interpretation of lists of insects using digitised biological information about the species. *J Insect Conserv* 5:131–139
- Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH (2006) Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecol Appl* 16:1267–1276
- Stranko SA, Hurd MK, Klauda RJ (2005) Applying a large, statewide database to the assessment, stressor diagnosis, and restoration of stream fish communities. *Environ Monit Assess* 108:99–121
- Suter GW II (1993) A critique of ecosystem health concepts and indexes. *Environ Toxicol Chem* 12:1533–1539
- Turak E, Flack LK, Norris RH, Simpson J, Waddell N (1999) Assessment of river condition at a large spatial scale using predictive models. *Freshw Biol* 41:283–298
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends Ecol Evol* 18(6):306–314
- VanDerWal J, Shoo LP, Johnson CN, Williams SE (2009) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *Am Nat* 174:282–291
- Vasconcelos TS, Rodríguez MÁ, Hawkins BA (2012) Species distribution modelling as a macroecological tool: a case study using New World amphibians. *Ecography* 35:539–548
- Warren D, Seifert S (2010) Environmental niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 21:335–342
- Whittier TR, Stoddard JL, Larsen DP, Herlihy AT (2007) Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. *J N Am Benthol Soc* 26:349–360
- Wilde GR (2011) Reproductive ecology and population dynamics of fishes in the upper Brazos river. Texas Parks and Wildlife State Wildlife Grant annual report
- Winemiller KO, King RS, Taylor J, Pease A (2009) Refinement and validation of habitat quality indices (HQI) and aquatic life use (ALU) indices for application to assessment and monitoring of Texas surface waters. final project report, Texas Commission on Environmental Quality Contract 582-6-80304, 81 pp
- Wright JF, Moss D, Armitage PD, Furse MT (1984) A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. *Freshw Biol* 14:221–256