CONSERVATION BIOGEOGRAPHY OF HEADWATER CATFISH (*Ictalurus lupus*) IN THE

UNITED STATES

A Thesis

by

STEPHANIE D. GEORGE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Joshuah S. Perkin
Committee Members,	Kevin W. Conway
	X. Ben Wu
	Preston T. Bean
Head of Department,	Kirk O. Winemiller

December 2019

Major Subject: Wildlife and Fisheries Sciences

ABSTRACT

The aim of this study was to review the conservation status of Headwater catfish (*Ictalurus lupus*) in the United States with emphasis on Texas populations. This status assessment included evaluating the change in geographic distribution and measuring introgression and hybridization with Channel catfish (*Ictalurus punctatus*) to inform conservation prioritization in the United States with emphasis in Texas. I used machine learning methods (i.e., random forest and boosted regression tree) to construct species distribution models (SDMs) based on historical and contemporary presence-absence data using 23 environmental predictors based on remotely sensed stream network data. I measured introgression and hybridization with the widely introduced Channel catfish using external morphology and molecular markers. The sub-basin (8-digit hydrologic unit code) from which collections were made was the most important predictor of Headwater catfish occurrence across all models. Species distribution models illustrated temporal changes in Headwater catfish occurrence. Historically, Headwater catfish occurrence was higher among streams with higher slopes, closer distances from spring outflows, broader ranges of annual precipitation, and with higher portions of the network catchment classified as water. These shifts are likely related to both range contraction of the species and temporal variation in sampling locations. Morphological and molecular data revealed four non-introgressed and isolated locations where conservation of Headwater catfish are likely to be most successful. Species distribution models provide critical assessments of where a species might persist, but they require careful validation and cannot account for introgression. Pairing targeted sampling efforts with locations highlighted by SDMs can be used to promote systematic conservation planning for rare and threatened species.

DEDICATION

I dedicate this thesis to my wonderful husband, Andy, who has supported me from day one. Without your guidance, encouragement, and sense of humor, this would have been a much more difficult journey.

ACKNOWLEDGEMENTS

<div align="center">CONTRIBUTORS AND FUNDING SOURCES</div>

Contributors

This work was supervised by a thesis committee consisting of Dr. Joshuah Perkin and Dr. Kevin Conway of the Department of Wildlife and Fisheries Sciences, Dr. X. Ben Wu of the Department of Ecosystem Science and Management, and Dr. Preston Bean of the Texas Parks and Wildlife Department.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Rapid human population growth and increased density are key factors contributing to the loss of biodiversity on a global scale, and conservation efforts now require systematic planning to be effective (McKee, Sciulli, Fooce, & Waite, 2004; Wilson, 1989). Habitat loss and degradation due to anthropogenic activities present the greatest threats to biodiversity, including introduction of non-native species, municipal and industrial development, increased agricultural land use, and human-induced increases in concentrations of greenhouse gases (Sala et al., 2000). Though biodiversity loss in terrestrial ecosystems is increasingly well-documented (Aerts & Honnay, 2011; Sala et al., 2000), diversity loss in aquatic ecosystems (i.e., marine and freshwater) historically received comparatively less attention despite widely documented anthropogenic degradation to physical, biological, and chemical components of aquatic systems (Moyle & Leidy, 1992). In marine ecosystems, human activities are frequently linked to biodiversity loss in coastal zones (Gray, 1997; Sala et al., 2000), while freshwater biodiversity in lakes, rivers, and springs is disproportionally threatened by human activities (Dudgeon et al., 2006). Freshwater animals represent 9.5% of known species, though freshwater covers only 0.8% of the Earth's surface (Dudgeon et al., 2006; Strayer & Dudgeon, 2010). Consequently, there is a growing urgency to address threats to biodiversity in freshwater ecosystems (Reid et al., 2019). Across ecosystem types, systematic conservation planning tools that allow for prioritizing allocation of limited resources towards habitat and species preservation are necessary for the long-term protection of intact ecosystems and their biota (Poiani, Richter, Anderson, & Richter, 2000; Possingham, Bode, & Klein, 2015; Williams et al., 2011).

Systematic conservation assessments for freshwater systems and their biota are relatively new because of the spatial and temporal challenges that are posed by streams and their drainage networks (Barmuta, Linke, & Turak, 2011; Erős, O'Hanley, & Czeglédi, 2018). Natural stream systems maintain connectivity across four dimensions, including longitudinal (upstream to downstream), lateral (main channels to floodplains), vertical (groundwater to surface water) and temporal (flow variability through time; Ward, 1998). Human alterations to riverscapes affect each of these dimensions of connectivity and result in alteration to natural structuring mechanisms for biotic assemblages (Cooper et al., 2017; Dudgeon et al., 2006; Perkin, Gido, Costigan, Daniels, & Johnson, 2015). In addition to habitat fragmentation, stream biodiversity is threatened by anthropogenic overexploitation, pollution, species invasions, land use change, and alterations to hydrologic regimes through surface water diversion, groundwater extraction, and climate change (Dudgeon et al., 2006; Gido, Whitney, Perkin, & Turner, 2016; Saunders, Meeuwig, & Vincent, 2002). Among these alterations, freshwater organisms are most strongly affected in places where habitat destruction (such as the loss of water availability) and the release of non-native species occur together (Arlinghaus, Lorenzen, Johnson, Cooke, & Cowx, 2016; Dudgeon et al., 2006; Vörösmarty et al., 2010). Consequently, the ever-changing field of freshwater conservation planning with focus on fishes seeks to identify areas where anthropogenic alterations can be mitigated to benefit the greatest number of species (Hermoso et al., 2015). For example, freshwater stream fishes may be useful biological indicators of water quantity and quality and can be used to develop proposed reserves (Araújo & Williams, 2000; Karr, 1981; Wellemeyer, Perkin, Fore, & Boyd, 2018). Comparison of fish occurrence data over time with corresponding environmental parameters may provide a better understanding of watershed characteristics and fish community structure while providing direction for site-specific

2

conservation areas. New research approaches to conservation planning using fish species distribution modeling are necessary to determine priority areas for biodiversity preservation and restoration actions (D'amen, Rahbek, Zimmermann, & Guisan, 2017; Hermoso et al., 2016).

Fish populations and assemblages in arid riverscapes around the world are in need of conservation. Drylands, which consist of arid, semi-arid, and desert regions, make up ~40% of Earth's surface and contain nearly one-third of the global human population (James et al., 2013). Streams in arid riverscapes rely on groundwater to maintain base flow, which usually has limited spatial extent and can vary seasonally (Murray, Zeppel, Hose, & Eamus, 2003). Consequently, preservation of perennial water sources in dryland regions is a major challenge facing aquatic biodiversity conservation (Davis, Kerezsy, & Nicol, 2017). This is because increases in human water demand and frequency and intensity of extreme hydrologic events such as drought affect availability of water for fishes in dryland freshwater ecosystems (Heino, Virkkala, & Toivonen, 2009; Perkin et al., 2019). Furthermore, the deliberate or accidental introduction of non-native fish species can have major cumulative effects on native fish species through competition or introgression in highly isolated dryland waterbodies (Cambray, 2003; Orians, 1995). A recurring theme among freshwater fishes inhabiting dryland riverscapes in North America is that threats emerge from combined habitat degradation and non-native species invasions (Jelks et al., 2008; Ketmaier & Bianco, 2015). This means conservation planning for dryland fishes must address habitat integrity while including consideration of natural patterns of genetic diversity for populations that use quality habitats (Echelle, 1991; Hermoso et al., 2015; Meffe & Vrijenhoek, 1988). Such approaches require multidisciplinary frameworks focused on applying modeling tools, molecular techniques, and spatial conservation planning to benefit species persistence in occupied habitats or reintroduction into restored habitats (Malone et al., 2018).

3

The Headwater catfish (*Ictalurus lupus*) is a member of the Channel catfish (*Ictalurus punctatus*) species complex and occurs in riffles, runs, and pools of spring-fed streams and moderate-sized rivers in the American Southwest (Figure 1). Gilbert and Burgess (1980) stated that Headwater catfish was among the least known and studied of North American freshwater fishes. Since that time, further study revealed the geographic distribution of Headwater catfish in the United States was declining as a result of habitat degradation and competition and introgression with the widely introduced Channel catfish (Bean, Jackson, McHenry, Bonner, & Forstner, 2011; Kelsch & Hendricks, 1990; McClure-Baker et al., 2010). As a result, Headwater catfish was listed as a species of Special Concern by Williams et al. (1989) and Hubbs et al. (2008), and listed as Threatened by Jelks et al. (2008). The historical range of Headwater catfish in the United States included the Pecos and the Rio Grande basins of Texas and New Mexico and the upper Nueces, Guadalupe, and Colorado basins in Texas (Kelsh & Hendricks, 1990). However, Headwater catfish are extirpated from a large but currently unquantified portion of their historical range in the United States, with only limited portions of the Pecos River, Rio Grande, and Frio River basins currently inhabited as of 10 years ago (Bean et al., 2011; Kelsch & Hendricks, 1990; McClure-Baker et al., 2010). The river systems where Headwater catfish persists in the United States are threatened by the growth of irrigated agriculture, declining groundwater tables and thus spring discharges, and continued introduction of non-native Channel catfish via reservoir stockings (Contreras-Balderas & Escalante, 1984; Souza et al., 2006). Systematic conservation planning for Headwater catfish is needed to determine the watershed conditions associated with historical and contemporary occurrences, where these conditions currently exist, the occurrence of the species at these locations, the genetic integrity of

4

populations where they exist, and how protective measures might be spatially allocated to ensure persistence of the species.

The goal of this study is to provide a comprehensive review of the conservation status of Headwater catfish in the United States with emphasis on Texas populations, which constitute the core range of the species in the United States. My first objective was to evaluate change in the geographic distribution of Headwater catfish using machine learning methods to construct species distribution models (SDMs) based on historical (1980–1999) and contemporary (2000–2018) presence-absence data and remotely sensed stream network data. Identifying areas suitable for Headwater catfish will support decision-making challenges for conservation managers, including habitats correlated with occurrence, where these habitats occur across the landscape, and the occurrence of Headwater catfish at these habitats both historically and recently. My second objective was to measure introgression and hybridization with the widely introduced Channel catfish at locations where Headwater catfish persist (as highlighted in the SDMs). Defining the genetic integrity of extant populations will inform conservation managers of where non-hybridized population remain and how these strongholds might be included in restoration planning.
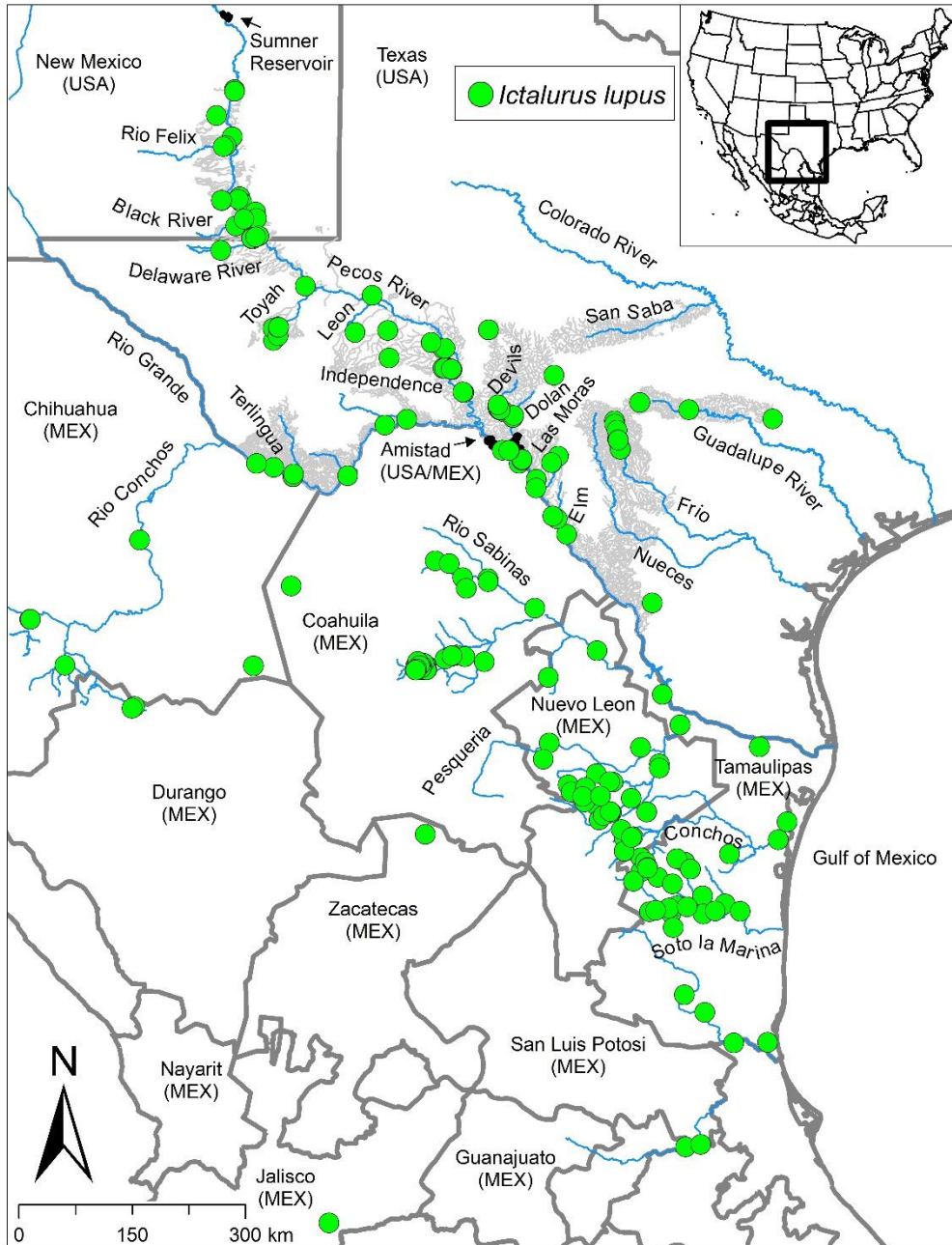
Figure 1. Geographic range of Headwater catfish (*Ictalurus lupus*). Presences compiled from Global Biodiversity Information Facility database. Grey shaded area represents the area emphasized by this study in Texas and New Mexico, USA.

6

METHODS

Study Area

I defined the spatial extent of this study based on level IV ecoregions of the conterminous United States (Omernik & Griffith, 2014) and the locations of Headwater catfish occurrences (Figure 1). The Rio Grande, or Rio Bravo del Norte, is located in the southwestern United States and northern Mexico. It is a water supply source for agriculture, industry, municipalities, and wildlife (Ward, Booker, & Michelsen, 2006). The Rio Grande flows through multiple biomes, including deserts, wetlands, mountains, and subtropical coastal regions. The river forms a 1,248-mile international border between Mexico and the United States from El Paso, TX to the Gulf of Mexico (Benke & Cushing, 2011). Other tributaries such as the Pecos River and the smaller Devils River join the Rio Grande upstream to or directly into the Amistad Reservoir in Val Verde County, Texas. Headwater catfish range spreads east to the Edwards Plateau region in Texas, where it was once found in the upper Nueces, Frio, Guadalupe, and San Saba rivers (Edwards, Garrett, & Allan, 2004; Kelsch & Hendricks, 1990). The Pecos River headwaters are in New Mexico and the river flows south for 500 miles before it joins with the Rio Grande upstream to Amistad Reservoir. Headwater catfish range in New Mexico includes the Pecos River downstream from Sumner Reservoir and the Black River as well as other southeastern-flowing tributaries to the Pecos River. Elsewhere in Mexico, Headwater catfish range includes various localities in Coahuila, Durango, Nuevo Leon, and Tamaulipas, southward to Rio Soto la Marina (Kelsch & Hendricks, 1986; Miller, Minckley, Norris, & Gach, 2005; Sublette, Hatch, & Sublette, 1990).

Spatial Data Collection

I downloaded occurrence data from the Global Biodiversity Information Facility (GBIF) to model the distribution of Headwater catfish. The GBIF is a portal that organizes digitized collection and survey data and is the largest online distributional database (Beck, Böller, Erhardt, & Schwanghart, 2014). Within the United States, records for Headwater catfish became increasingly prevalent beginning in the 1980s through the 1990s because of works by authors such as Kelsch & Hendricks (1986) and Kelsch & Hendricks (1990), and more recent collections were made by authors such as McClure-Baker et al. (2010). Given the temporal nature of occurrences, I used GBIF records from across the study area split into historical (1980-1999) and contemporary (2000-2018) time periods. Because these data represent occurrence-only data, I used the target-group absence (TGA) approach described by Mateo et al. (2010) to assign Headwater catfish absences at locations where Channel catfish, but not Headwater catfish, were collected. The TGA framework is a method for assigning pseudo-absences in a manner that is more effective than generation of random absences, and Mateo et al. (2010) showed that models built with TGA are more accurate than models based on traditional pseudo-absences. Previous works predicting fish distributions have also utilized the TGA approach (Huang & Frimpong, 2015; Malone et al., 2018). For each time period, georeferenced data downloaded from GBIF were linked to National Hydrography Dataset (NHD) Version 2 Plus inter-confluence stream segments (EPA, 2012) using a geographic information system (GIS). Specifically, I assigned GBIF occurrence data to the nearest segment within 50 m (Frimpong et al., 2005) using ArcGIS 10.5 (ESRI, The Redlands, CA), such that any segment that was linked to a Headwater catfish record was denoted as a presence location (class = 1) and any segment that was linked to

only a Channel catfish record was denoted as an absence location (class = 0). Segments with both Headwater catfish and Channel catfish were denoted as presence locations (class = 1).

I used 23 environmental predictor variables to test as covariates for Headwater catfish occurrence (Table 1). These covariates were based on previous studies linking catchment attributes to the occurrence of stream fishes (Malone et al., 2018). Predictor variables described hydrologic, climatic, physiographic, and anthropogenic land uses for each stream segment across the study area. Hydrologic variables included upstream watershed area (km^2), stream order (Strahler, 1957), stream channel maximum and minimum elevation (m), stream channel slope (m/m), discharge (cubic feet per second), water velocity (feet per second), and distance to the nearest known spring outflow (km). Climatic variables included mean annual air temperature (°C) and mean annual precipitation (mm). Physiographic variables included the major underlying geologic features and the identity of the 8-digit hydrologic unit code (HUC) in which the stream segment occurred. Land use data were from Falcone et al. (2015) and described the area of upstream watershed covered by 11 land use classes (Table 1) prior to the historical (survey year 1974) period and during the contemporary (survey year 2002) period. Although the survey years for land cover does not exactly match the historical and contemporary periods defined for fish collections, they do represent antecedent conditions for the period from which fish records were compiled (Perkin, Murphy, Murray, Gibbs, & Gebhard, 2019).

Table 1. Twenty-three environmental predictor variables used for species distribution models, data description, and source. Land use subclasses of the NAWQA Wall-to-Wall Anthropogenic Land Use Trends dataset (NAWQA, U.S. Geological Survey's National Water-Quality Assessment Program, Falcone, 2015).

	Habitat Parameter	Description	Source
1	Watershed	Upstream area of drainage (km^2) based on DivDASqKM attribute from PlusFlowlineVAA attribute table	EPA (2012)
2	Stream order	Strahler (1957) stream order classification based on StreamOrde from PlusFlowlineVAA attribute table (>2)	EPA (2012)
3	Min. Elevation	Min elevation (m) of stream segment based on MINELEVSMO attributes from elevslope attribute table.	EPA (2012)
4	Max Elevation	Max elevation (m) of stream segment based on MAXELEVSMO attributes from elevslope attribute table.	EPA (2012)
5	Slope	Slope (m/m) of stream segment based on SLOPE attribute from elevslope attribute table.	EPA (2012)
6	Discharge	Flow from gage adjustment (cubic feet per second, cfs) based on Q0001E attribute from EROMExtension attribute table.	EPA (2012)
7	Velocity	Velocity from gage adjustment (feet per second, fps) based on V0001E attribute from EROMExtension attribute table.	EPA (2012)
8	Air temperature	Mean annual air temperature (°C*10) based on Temp attribute from CatchmentAttributesTempPrecip attribute table	EPA (2010)
9	Precipitation	Mean annual precipitation (mm) based on Precip from the CatchmentAttributesTempPrecip attribute table	EPA (2010)
10	Major geology	Major geology of the Conterminous United States	Schruben et al. (1994)
11	HUC8	8-digit Hydrologic Unit Code (Subbasin level)	USDA-NRCS (2015)
12	Distance to springs	Distance to springs (Euclidian distance in km)	Heitmuller and Reece (2003)

Table 1. Continued

	Habitat Parameter	Description	Source
13–23	Land Cover/Land Use	Land cover and land use (1974, 2002) Lu_11—Water; Lu_21—Major Transportation; Lu_22—Commercial/Services; Lu_23—Industrial/Military; Lu_26—Residential, Low-Medium Density; Lu_27—Developed, Other; Lu_32—Urban Interface Low Medium; Lu_43—Crops; Lu_44—Pasture/Hay; Lu_45—Grazing Potential; Lu_50—Low Use	Falcone (2015)

Species Distribution Models

   I summarized changes in the relationships between watershed covariates and Headwater

catfish occurrences during historical, contemporary, and combined time periods using SDMs. I

grouped the data to create a combined time period to show presence-absence of Headwater

catfish during the period from 1980–2018. Elith et al. (2006) suggested presence-absence models

outperform presence-only models, thus I used two forms of presence-absence modelling: boosted

regression tree (BRT; Elith, Leathwick, & Hastie, 2008) and random forest (RF; Breiman, 2001)

models. I fit models to historical, contemporary, and combined periods separately for each model

type, resulting in a total of six models. Datasets assigned to each time period suffered from class

imbalance such that presence records (class = 1) were fewer compared with absence records

(class = 0) in the historical (Figure 2a; 1 = 31, 0 = 154), contemporary (Figure 2b; 1 = 19, 0

=114), and combined (1= 45, 0 = 234) periods.
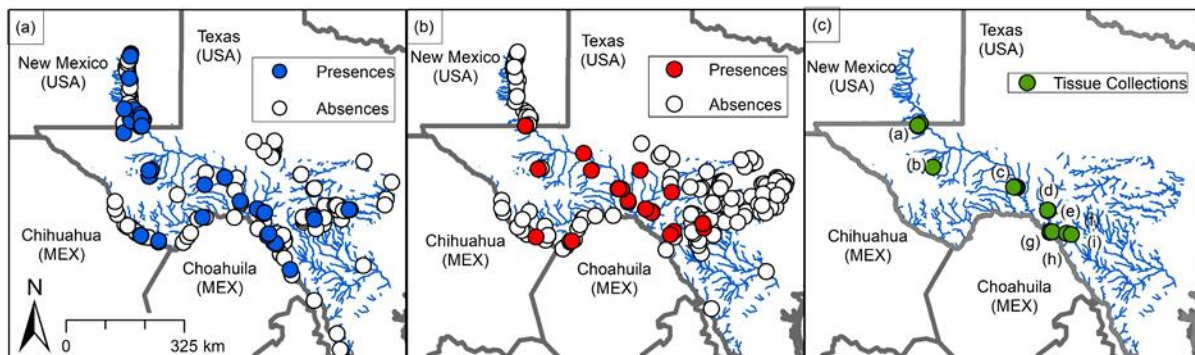
Figure 2. Historical sites of Headwater catfish presence (blue solid circles) and absence (open circles) in the United States (a). Contemporary sites of Headwater catfish presence (red solid circles) and absence (open circles) in the United States (b). Sites where tissue of Headwater catfish (solid green circles) were collected in 2018 (c). See Table 4 for identification of locations with adjacent letters.

I addressed class imbalance using the synthetic minority over-sampling technique (SMOTE) described by Chawla et al. (2002) and applied the 'SMOTE' function from the 'DMwR' package in R (Torgo, 2010). This process resulted in balanced numbers of occurrences for historical (1 = 93, 0 = 93), contemporary (1 = 76, 0 = 76), and combined (1 = 175, 0 = 180) periods. I then fit BRT and RF models to the SMOTE-adjusted datasets using the 'gbm.step' function from the 'dismo' package (Greenwell, Boehmke, Cunningham, & GBM Developers, 2019) and the 'randomForest' function from the 'randomForest' package (Liaw & Wiener, 2002) in R. For the historical BRT model, I used 10,650 trees, tree complexity = 5, bagging fraction = 0.30, and learning rate = 0.0005, for the contemporary BRT model I used 10,100 trees, tree complexity = 5, bagging fraction = 0.25, and learning rate = 0.001, and for the combined BRT model I used 9,600 trees, tree complexity = 5, bagging fraction = 0.25, and learning rate = 0.001. I verified these parameters by using the process of hypertuning (Kuhn, 2008) in order to detect any improvements to the models performance. I tuned random forest models using the 'tuneRF' function from the 'rfUtilities' package (Evans & Murphy, 2015) to determine the number of variables to try at each split. For the historical RF model, I used 1,200 trees with 4 variables tried at each split, for the contemporary RF model I used 250 trees with 4 variables tried at each split, and for the combined RF model I used 250 trees with 4 variables tried at each split. I assessed model performance using k-fold cross validation (k = 5) and used the 'gbm.step' and 'rf.crossvalidation' functions for cross-validation analyses and the 'confusion.matrix' function from the 'dismo' package to assess model performance. I report model accuracy and omission and commission rates for the cross-validations (Table 2). Partial dependence plots were constructed for the historical and contemporary BRT models using the 'gbm.plot' function and I viewed variable importance using the 'summary' function in the 'dismo' package. Partial

dependence plots were constructed for the historical and contemporary RF models using the 'rf.partial.prob' function in the 'rfUtilities' package and I viewed variable importance using the 'varImpPlot' function in the 'randomForest' package. Finally, I developed an ensemble model based on averaged predictions from BRT and RF models for historical, contemporary, and combined periods to estimate occurrence predictions informed by both models (Araújo & New, 2007). All analyses were performed in R version 3.6.0 (R Core Team, 2019).

Table 2. Headwater catfish species distribution model performance statistics based on 5-fold cross validation for boosted regression tree and random forest historical, contemporary, and combined time periods.

Performance metric	Historical BRT	Contemporary BRT	Both BRT	Historical RF	Contemporary RF	Both RF
AUC	0.90	0.97	0.92	0.93	0.99	0.97
Kappa	0.66	0.83	0.76	0.72	0.89	0.83
Accuracy	0.83	0.91	0.88	0.86	0.95	0.91
Sensitivity	0.85	0.91	0.86	0.86	0.95	0.90
Specificity	0.81	0.92	0.90	0.86	0.95	0.92

Tissue Collections

In 2018, my field assistants and I sampled 36 sites within the historical range of Headwater catfish (Figure 2c). These sites included the Rio Grande and tributaries that directly empty into the Rio Grande, including Cibolo Creek, Alamito Creek, Terlingua Creek, Tornillo Creek, Devils River, Dolan Creek, San Felipe Creek, Pinto Creek, Las Moras Creek, and Elm Creek. Our team also sampled tributaries to the Pecos River, including the Delaware River, Salt Creek, Independence Creek, and San Soloman Springs in Balmorhea State Park. Specimens were collected using seines, mini gillnets, and backpack electrofishing equipment. Catfishes were euthanized in a lethal solution of Tricaine methanesulfonate (MS-222) and tissue was extracted from the adipose fin and the right maxillary barbel and preserved in 95% non-denatured ethanol. Specimens were tagged using Floy T-bar tags with unique ID numbers, fixed in 10% formaldehyde solution, and later transferred to 70% ethanol alcohol for final storage. All specimens were deposited in the Biodiversity Research and Teaching Collections (TCWC) at Texas A&M University, College Station, TX, USA.

For each of the catfish specimens collected during 2018, I recorded anal fin ray count, standard length, pectoral spine length, caudal peduncle depth, and mouth width. These were measured with digital calipers and recorded to the nearest 0.1 mm, except that standard length in large specimens were measured to the nearest mm with a meter ruler. The data for each fish was inserted into the following formulae to obtain a linear canonical discriminant function score based on Kelsch (1995) to separate Headwater catfish and Channel catfish by external characters. Individual morphology scores (S) were calculated using the equations:

$$S = S_1 + S_2 + S_3 + S_4;$$
$$S_1 = \frac{1.244(anl - 25.2)}{1.66};$$
$$S_2 = \frac{0.2\{log_e(psl) - [-1.254 + 0.864 + log_e(stdl)] + 0.077\}}{0.144};$$
$$S_3 = \frac{-0.276\{log_e(cpd) - [-2.127 + 0.952 * log_e(stdl)] - 0.057\}}{0.087};$$
$$S_4 = \frac{-0.673\{log_e(mw) - [-2.996 + 1.129 * log_e(stdl)] + 0.096\}}{0.163}$$

where *anl* is anal fin ray count, *psl* is pectoral spine length, *stdl* is standard length, *cpd* is caudal

peduncle depth, and *mw* is mouth width (Kelsch, 1995). I assigned specimens to species

classifications based on these scores and paired this information with molecular data. For each

tissue sample paired with a specimen, I sequenced the mitochondrial cytochrome b gene (mt-*cyt*

b) and resolved taxa diagnostic single nucleotide polymorphisms (SNPs) from the nuclear

recombination activating 2 gene (nuc-RAG2). Cytochrome b is commonly used as a region of

mitochondrial DNA for determining phylogenetic relationships between organisms and is most

useful for the comparison of species in the same genus or the same family because of its

sequence variability (Castresana, 2001; Ketmaier & Bianco, 2015). To gain insight of species

boundaries, the use of nuclear DNA can detect areas of hybridizations and introgression events

likely to go unnoticed based only on mitochondrial DNA (Ketmaier & Bianco, 2015). In this

study, Headwater catfish revealed three mt-*cyt b* haplotypes (Ha1, Ha2, and Hb) and Channel

catfish revealed four haplotypes (Ca1, Ca2, Cb, and Cc1). Mitonuclear genotypes further defined

parental and admixed hybrid individuals. Genotypes of each specimen were classified as pure

Headwater catfish (H) and pure Channel catfish (C). Genotypes were further classified as first

generation hybrid with Headwater catfish (HC_F1), later generation hybrid with Headwater

catfish (HC_Fx), first generation hybrid with Channel catfish (CH_F1), and later generation

17

hybrid with Channel catfish (CH_Fx). Phenotypes and genotypes were compared by plotting

frequency histograms of linear discriminant function scores colored by genotype.

RESULTS

Species Distribution Models

Headwater catfish SDMs differed among time periods and model types. Model performances reported for each model and time period included AUC values 0.90–0.99, Kappa 0.66–0.89, accuracy 0.83–0.95, sensitivity 0.85–0.95, and specificity 0.81-0.95 (Table 2). Contemporary models tended to perform better than their historical counterparts while the combined time period model performance statistics were intermediate. The sub-basin (HUC 8) from which collections were made was the most important predictor variable across all models (Table 3).

Table 3. Twenty-three environmental predictor variables used for species distribution models, code, BRT relative influence values and RF variable importance (mean decrease in Gini and mean decrease in accuracy) values for historical, contemporary, and combined time periods.

Variable	Code	Historical			Contemporary			Combined		
		BRT	RF		BRT	RF		BRT	RF	
		Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy
Watershed	WSHED	0.587	4.664	18.138	0.529	3.681	8.366	0.576	8.055	12.276
Stream order	SO	0.726	2.999	16.518	0.052	2.339	7.290	0.823	7.073	11.201
Minimum elevation	MINELEV	0.786	3.778	19.323	0.130	1.691	7.709	0.449	10.358	12.328
Maximum elevation	MAXELEV	2.376	3.469	20.189	0.706	2.305	7.016	1.275	9.207	13.193
Slope	SLOPE	6.676	9.327	29.377	0.552	2.246	7.518	3.922	12.365	16.310
Discharge	DIS	2.264	6.537	24.000	0.358	2.160	7.189	1.047	8.297	12.574
Velocity	VEL	1.478	4.360	19.467	0.667	2.718	7.227	2.496	7.156	14.117
Air temperature	Temp	1.084	3.105	17.405	1.014	1.622	6.597	1.510	8.330	11.865
Precipitation	Precip	2.210	6.799	25.362	0.846	8.312	11.79	2.599	16.443	16.953
Major geology	GEOL	7.513	4.030	14.116	0.457	1.610	5.653	7.603	9.977	13.404
HUC8	HUC8	62.010	21.861	39.534	83.210	20.960	16.520	61.70	38.382	21.763

Table 3. Continued

Variable	Code	Historical			Contemporary			Combined		
		BRT	RF		BRT	RF		BRT	RF	
		Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy
Distance to springs	NEAR_ DIST	2.164	2.890	14.323	1.301	3.053	7.891	3.719	7.539	12.609
Land cover/Land use										
	Lu_11	3.767	2.143	12.365	8.507	8.048	12.00	5.825	9.674	14.579
	Lu_21	0.768	1.559	11.806	0.189	0.638	4.245	1.051	2.386	6.985
	Lu_22	0.235	1.805	13.264	0.022	0.791	4.337	0.463	1.929	6.710
	Lu_23	0.277	0.765	8.747	0.000	0.229	2.626	0.024	0.690	2.932
	Lu_26	0.003	0.753	6.257	0.000	0.327	2.394	0.013	0.773	3.187
	Lu_27	2.373	3.239	17.861	0.966	4.920	7.583	1.673	4.435	10.507
	Lu_32	0.189	1.568	12.826	0.000	0.287	3.192	0.016	0.608	3.368
	Lu_43	0.329	1.144	10.519	0.008	1.210	5.272	0.269	1.632	6.270

Table 3. Continued

Variable	Code	Historical			Contemporary			Combined		
		BRT	RF		BRT	RF		BRT	RF	
		Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy	Relative influence	Mean decrease Gini	Mean decrease accuracy
Land cover/Land use										
	Lu_44	0.145	1.038	10.402	0.000	0.165	2.064	0.014	1.266	5.743
	Lu_45	0.417	0.924	8.828	0.157	3.770	8.044	0.331	3.300	9.278
	Lu_50	1.610	3.571	13.663	0.317	1.175	3.110	2.587	6.606	10.642

The BRT models showed greater importance of geology and stream channel slope during the historical period compared with the contemporary period, but greater importance in percentage of land cover occupied by water during the contemporary period compared with the historical period. In the combined period, geology showed importance similar to the historical period and intermediate importance of stream channel slope and percentage of land cover occupied by water (Figure 3a). The RF models provided two different importance measures to rank the predictor variables, mean decrease in Gini and mean decrease in accuracy. Mean decrease in Gini showed greater importance of stream channel slope and discharge during the historical period compared with the contemporary period, but greater importance of precipitation and percentage of land cover occupied by water during the contemporary period compared with the historical period. Overall, the same predictor variables ranked as having greater importance in the combined period compared with the historical and contemporary periods separately (Figure 3b). Stream channel slope and discharge ranked highest for the mean decrease in accuracy importance, although these variables showed much greater importance in the historical period compared with the contemporary and combined period (Figure 3c).
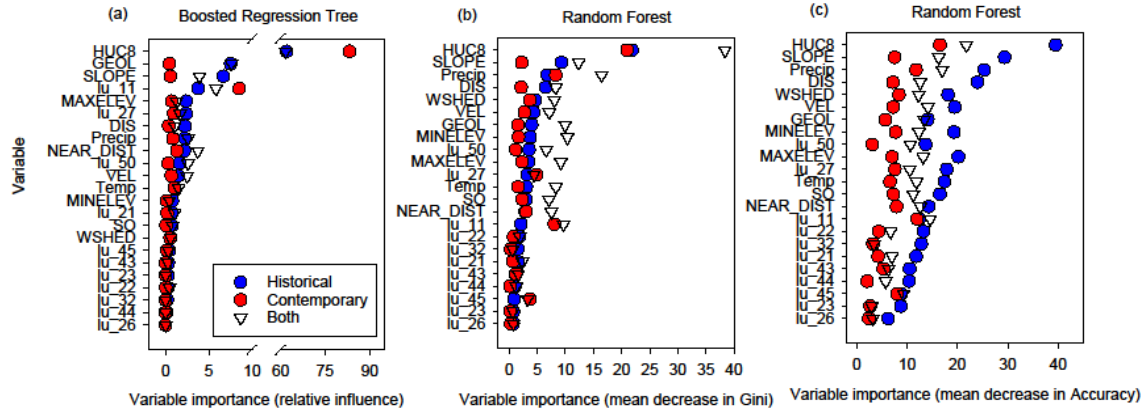
Figure 3. Plot of 24 environmental variable importance values (a) relative influence from boosted regression tree model output and (b) mean decrease in Gini and (c) mean decrease in accuracy from random forest model output. Solid blue circles represent historical periods (1980-1999), solid red circles represent contemporary periods (2000-2018), and open triangles represent combined periods.

Partial dependence plots provided insight into relationships between environmental variables and Headwater catfish occurrence. For both BRT and RF models, Headwater catfish occurrence was greater in streams with steeper slopes (>0.002 m/m), and both model types showed reduced occurrence within these higher gradient streams during the contemporary period compared with the historical period (Figure 4a and b). Combined periods show the highest probability of occurrence in streams with slopes approximately 0.002-0.003 m/m. Headwater catfish occurrence was greatest in close proximity to spring outflows, and declined with distance from springs (Figure 4c and d). The influence of distance to the nearest spring was much more pronounced in the contemporary period compared with the historical period, such that occurrences were always greatest within 0.1 km of known springs. Headwater catfish occurrence was greater in stream segments that received 375–550 mm/year across all time periods, but occurrence was most probable during the contemporary period among stream segments that received approximately 500 mm of precipitation (Figure 4e and f). Models illustrated a general reduction in the percent of watersheds dominated by water land cover between historical and contemporary periods (Figure 4g and h). Headwater catfish probability of occurrence was lowest among contemporary stream segments with <1% of land covered by water.
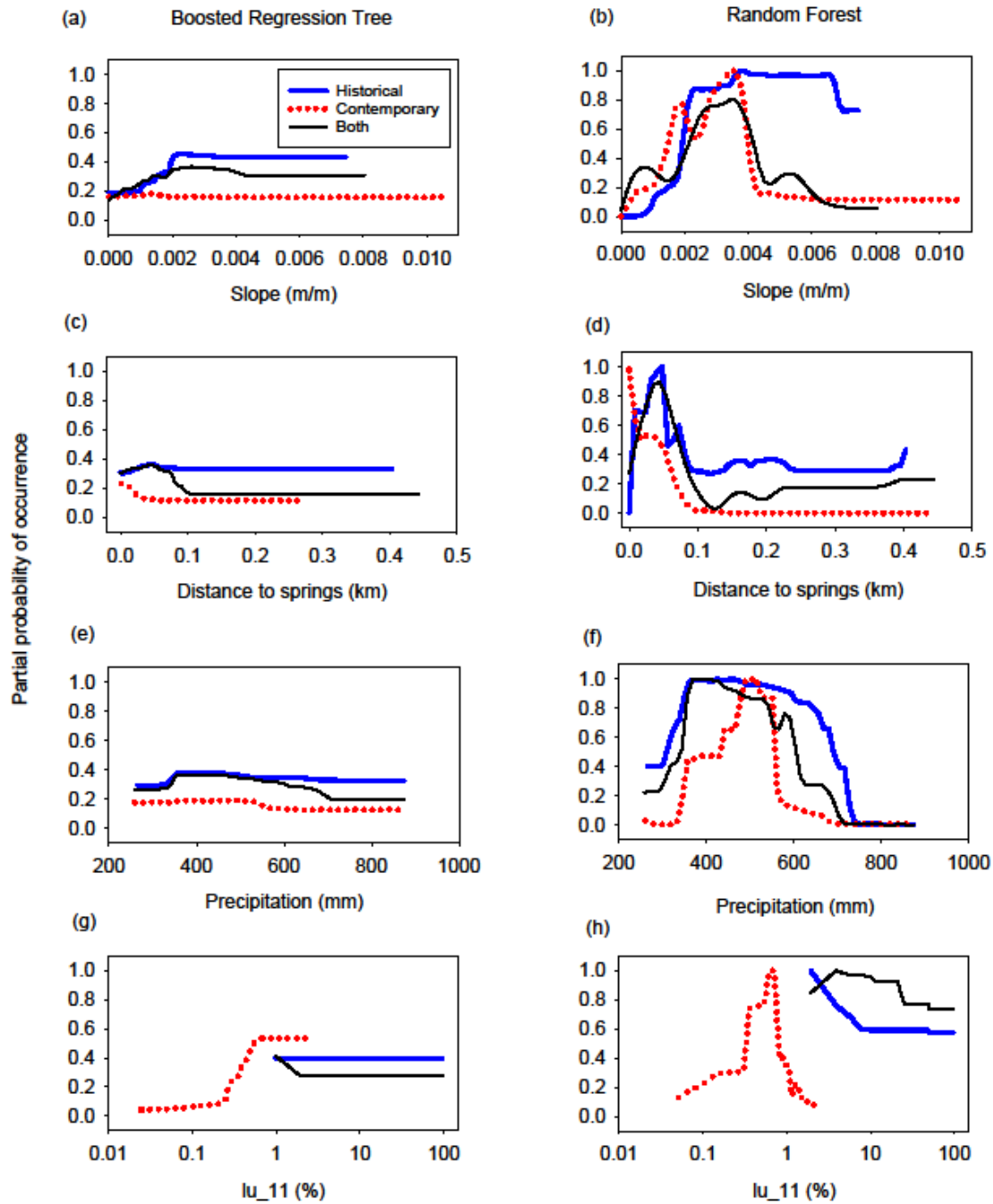
Figure 4. Partial dependence plots for boosted regression tree and random forest analyses relating species occurrence to the top influential environmental predictors for historical, contemporary, and combined periods.

Species distribution maps fit to all study area stream segments with sufficient data varied by model type and time period. The BRT, RF, and ensemble models fit to historical data captured regional variation in occurrence hotspots, including streams in the upper Pecos River, Delaware River and Rio Felix, Toyah Creek, Independence Creek, Dolan Creek, and Terlingua Creek in the Big Bend area (Figure 5a-c). The models highlighted additional streams where Headwater catfish historically occurred, including Sycamore Creek, Pinto Creek, Frio River, Sabinal River, and sections of the Guadalupe River. Analyses from all contemporary models showed a general pattern in hotspots around the lower Pecos River and surrounding tributaries such as Live Oak Creek and Independence Creek as well as the Devils River and Dolan Creek (Figure 5d-f). The combined time period models emphasized occurrence hotspots including Rio Felix, the Delaware River, Toyah Creek, the Devils River, Dolan Creek, and sections of the lower Pecos River near Independence Creek. A few hotspots existed along tributaries of the Rio Grande, including San Felipe Creek, Las Moras Creek, Cienegas Creek, and Pinto Creek (Figure 5g-i). Overall, the ensemble model using both time periods resulted in the most comprehensive map of Headwater catfish range in the United States over the last 40 years (Figure 5i).
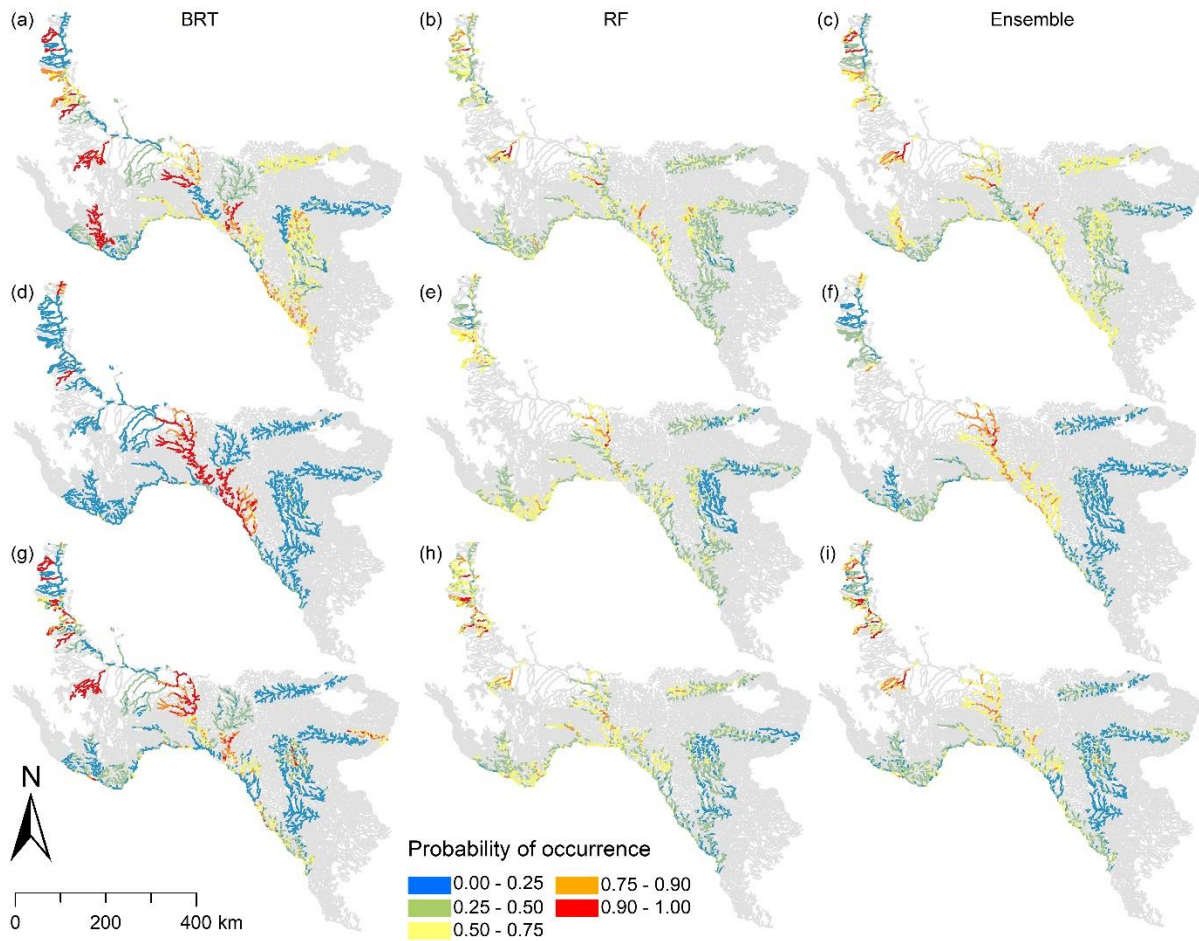
Figure 5. Species distribution models showing predicted probability of occurrence of Headwater catfish for historical (1980-1999) (a-c), contemporary (2000-2018) (d-f), and combined periods (1980-2018) (g-i) using a boosted regression tree model (a, d, g), a random forest model (b, e, h), and an ensemble of both models (c, f, i).

External Morphology vs. Molecular Status

Sampling during 2018 yielded 145 catfishes. The canonical discriminant function analysis identified 131 as Headwater catfish and 11 as Channel catfish. There were seven sites where Headwater catfish were present based on morphology and genetic results: Dolan Creek, San Felipe Creek, Las Moras Creek, Pinto Creek, Delaware River, Independence Creek, and San

Soloman Springs at Balmorhea State Park. Comparison of canonical scores from phenotypic measurements with results from mitonuclear analyses revealed 43 catfishes classified as Headwater catfish based on phenotypes also had Headwater catfish genotypes. Seventy-six individuals with Headwater catfish phenotypes were classified as genotypic hybrids and four individuals with Headwater catfish phenotypes had Channel catfish genotypes (Figure 6a). All specimens from Balmorhea State Park had Headwater catfish phenotypes and genotypes (Figure 6b), two channel catfish and one hybrid were collected from Cienegas Creek (Figure 6c), and a mixture of phenotypes and genotypes were collected from the Delaware River (Figure 6d). Six Headwater catfish with intermediate phenotypes and one hybrid were collected from the Devils River (Figure 6e), six specimens from Dolan Creek had Headwater catfish phenotypes and genotypes (Figure 6f), and only Headwater catfish and hybrids were collected at Independence Creek (Figure 6g). Two Headwater catfish were collected from Las Moras Creek (Figure 6h), three Headwater catfish and one hybrid were collected from Pinto Creek (Figure 6i) and four Headwater catfish were collected from San Felipe Creek (Figure 6j).
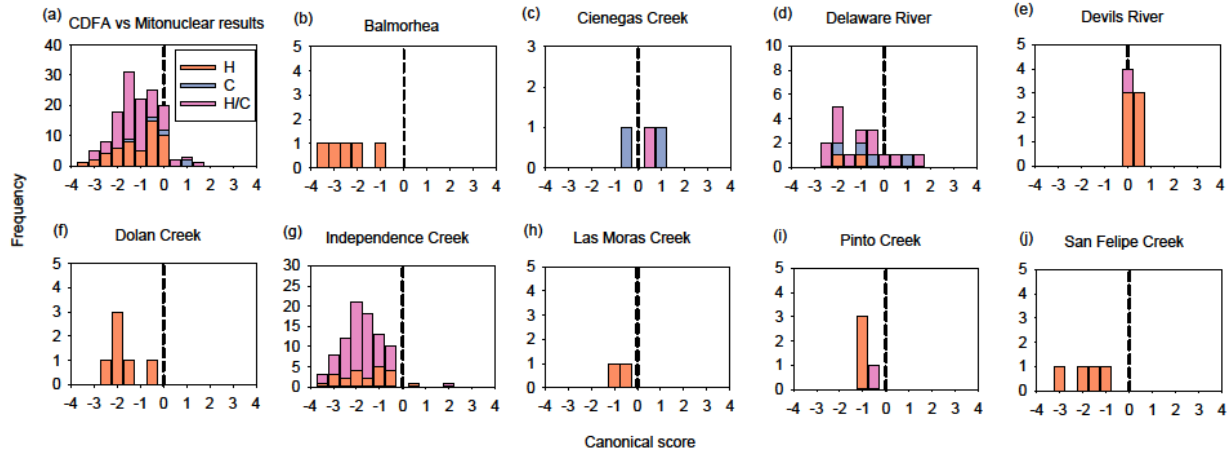
Figure 6. Color coded mitonuclear genotypes and canonical discriminant function analysis values for all specimens (a) and per individual sites surveyed (b-j). Canonical scores less than 0 represent Headwater catfish and scores greater than 0 represent Channel catfish. Headwater catfish (H) represented in orange, Channel catfish (C) represented in purple, and Headwater/Channel catfish hybrid (H/C) represented in pink. Specimens too small for external morphology analysis were included and given a score of zero.

Landscape Genetics

The use of mitochondrial and nuclear DNA sequencing techniques proved instrumental in detecting introgression and hybridization events of Headwater catfish and Channel catfish in this region (Table 4).

Table 4. Individual molecular results from 2018 survey showing the number of haplotypes for mt-*cyt b* sequences, genotypes from nuc_RAG2 gene, and mitonuclear genotypes.

		mt-*cyt b* haplotypes							nuc-RAG2-genotypes					mitonuclear genotypes			
		Ha1	Ha2	Hb	Ca1	Ca2	Cb	Cc1	H	C	H/C	H	C	CHF1	CHFx	HCF1	HCFx
a)	Delaware River	7	-	-	11	-	-	-	2	4	12	2	4	5	2	4	1
b)	Balmorea	-	5	-	-	-	-	-	5	-	-	5	-	-	-	-	-
c)	Independence Creek	23	-	-	59	5	-	-	22	-	65	22	-	3	61	1	-
d)	Dolan Creek	-	-	6	-	-	-	-	6	-	-	6	-	-	-	-	-
e)	Devils River	-	-	6	1	-	-	-	7	-	-	6	-	-	1	-	-
f)	San Felipe Creek	-	-	4	-	-	-	-	4	-	-	4	-	-	-	-	-
g)	Cienegas Creek	-	-	-	-	-	2	1	-	2	1	-	2	1	-	-	-
h)	Pinto Creek	-	-	4	-	-	-	-	3	-	1	3	-	-	-	1	-
i)	Las Moras Creek	-	-	2	-	-	-	-	2	-	-	2	-	-	-	-	-

Multiple mitochondrial-CytB haplotypes were found for Headwater catfish and Channel catfish. Three Headwater catfish mt-*cyt b* haplotypes occurred among three regions, including the Pecos River corridor (Delaware River and Independence Creek; Ha1), Balmorhea State Park (Ha2), and the Devils River and Rio Grande tributaries downstream of Amistad Reservoir (Hb; Figure 7a). Headwater catfish mt-*cyt b* haplotypes were found in isolation (i.e., no Channel catfish present) at Balmorhea State Park, Dolan Creek, San Felipe Creek, Pinto Creek, and Las Moras Creek. Four Channel catfish mt-*cyt b* haplotypes were detected, and these co-occurred with Headwater catfish mt-*cyt b* haplotypes at four locations, including the Delaware River, Independence Creek, Devils River, and Cienegas Creek. Nuclear RAG2 SNPs indicated exclusively Headwater catfish alleles at Balmorhea State Park, the Devils River, Dolan Creek, San Felipe Creek, and Las Moras Creek (Figure 7b). Combined mitonuclear genotypes indicated only Headwater catfish at Balmorhea State Park, San Felipe Creek, Las Moras Creek, and Dolan Creek (Figure 7c). The Delaware River, Independence Creek, Devils River, Cienegas Creek and Pinto Creek exhibited varying levels of hybridization. Non-introgressed Channel catfish samples were only found at Cienegas Creek and the Delaware River, and Channel catfish combined mitonuclear genotypes out-numbered Headwater catfish genotypes in the Delaware River, Independence Creek, and Cienegas Creek (where no Headwater catfish mt-*cyt b* haplotypes were detected).
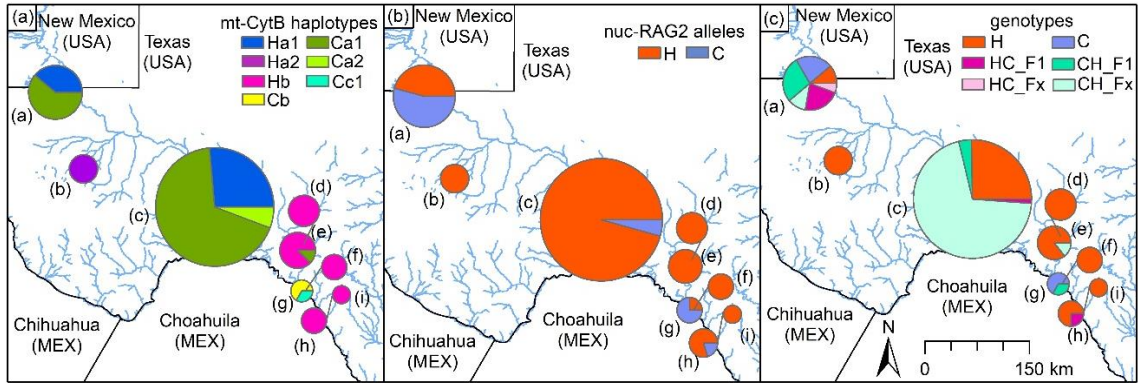
Figure 7. Genetic results from 2018 Headwater catfish surveys. Haplotypes for mt-*cyt b* sequences. First alpha character: H = Headwater catfish, C = Channel catfish. Subsequent alphanumeric characters differentiate each haplotype (a); alleles from the nuc-RAG2 gene (b); and mitonuclear genotypes (c). Graphs scaled by sample size at each site. H= Headwater catfish (genotype), C = Channel catfish, HC_F1 = first generation hybrid with Headwater catfish mitochondria, HC_Fx = later generation hybrid with Headwater catfish mitochondria, CH_F1 = first generation hybrid with Channel catfish mitochondria, CH_Fx = later generation hybrid with Channel catfish mitochondria. See Table 4 for identification of locations with adjacent letters.

CONCLUSIONS

Historical decline in the geographic range of Headwater catfish was the result of combined effects of habitat degradation and introductions of non-native Channel catfish (Bean et al., 2011; Kelsch & Hendricks, 1990; McClure-Baker et al., 2010). However, identifying the nature of habitat degradation as well as the habitat features that should be conserved is a challenge. Currently, modeling provides the most flexible and comprehensive method to predict changes in species populations under predicted or observed environmental change (D'amen et al., 2017). Results identified streams with steeper slopes in close proximity to springs, with intermediate precipitation levels, and greater amounts of water within the catchment as habitats that are most likely to be presently inhabited by Headwater catfish. This temporal assessment showed that contemporary distributions tended to have lower probability of occurrence among stream segments with steep slopes and further away from spring outflows. This pattern is consistent with loss of Headwater catfish from their namesake headwater stream habitats. Where headwater stream habitats persist, the resident populations of Headwater catfish are also threatened by the stocking of non-native Channel catfish. For example, Channel catfish were historically stocked into ponds on the Independence Creek preserve prior to establishment of the preserve, and the high-quality habitats that persist there are inhabited by a mix of pure-strain Headwater catfish and hybrids with Channel catfish. Similar signals of hybridization within relatively pristine habitats were apparent for Cienegas Creek, the Delaware River, the Devils River, and Pinto Creek. Fortunately, some high-quality habitats maintain populations of Headwater catfish that have been isolated from introgression with Channel catfish, including Balmorhea State Park, Dolan Creek, Las Moras Creek, and San Felipe Creek. Each of these

habitats are threatened by anthropogenic water extractions that could destroy habitats, and in the cases of Balmorhea State Park, Las Moras Creek, and San Felipe Creek, human alterations to habitats included construction of swimming pools, irrigation canals, and golf courses. Successful conservation of Headwater catfish will require maintenance of both high quality habitat and protection from future introductions of Channel catfish at these locations.

Species distribution models are beneficial tools for conservation biogeography because of their ability to project future time periods and across landscapes to identify suitable habitat important for species conservation (Araújo & New, 2007; Elith & Leathwick, 2009). Systematic conservation planning relies on accurate species occurrence records as well as the knowledge of population size and connectivity to provide a complete picture of both current status and future management actions (Hermoso, Kennard, & Linke, 2012; Moritz, 1995). The use of occurrence records obtained from GBIF paired with remotely sensed environmental predictor variables provided a basis for modeling suitable habitat for Headwater catfish. Although the resulting models provided excellent predictive power, I recognize these projections are subject to some limitations. First, bias in the collection of species occurrence records could result in model predictions that are biased towards environments that have received more intense sampling (Araújo & Guisan, 2006). This issue confounds temporal shifts in species distribution because differences between historical and contemporary time periods could be due to either shifts in species distributions or the areas sampled. I addressed this issue by constructing a combined time period model that made use of all available data to highlight habitats most important to Headwater catfish. Second, GBIF data represent occurrence only records and therefore limit the use of presence-absence modeling, a method demonstrated to be more accurate compared with presence-only modeling (Elith et al., 2006). I addressed the issue of 'no absences' by assigning

'target group absences' identified as sites where Channel catfish but not Headwater catfish were collected (Mateo et al., 2010). I also point out that the degree to which competition between Channel catfish and Headwater catfish influences the range of Headwater catfish may be worth exploring for future SDMs (Araújo & Guisan, 2006). Third, SDMs should be interpreted as predictions of potential habitat that are useful as guiding information for conservation planning, but occurrences outside of the predicted areas are still possible. Recent documented occurrence of Headwater catfish in the Frio River where the species was previously believed to be extirpated (Bean et al., 2011) illustrates that the species may persist (though at low densities) in other areas of the riverscape not identified in the SDM. By the same token, low detection among commonly used gears (e.g., seines) could result in absence of Headwater catfish among sites with high probability of occurrence (Budy, Conner, Salant, & Macfarlane, 2015). Fourth, the SDMs showed discrepancies in predictions between the BRT and RF models. I used the process of hypertuning (Kuhn, 2008) to adjust the models parameters in order to improve the model performance. I found that hypertuning the models did not result in prediction improvements and therefor prediction patterns in the BRT are a result of a strong effect of important variables (e.g., HUC8) and stochastic boosting (Elith et al., 2008). Finally, the predictions rely on the assumption that current modeled environmental conditions are the primary drivers of Headwater catfish distributions and that these relationships will persist in the future (Araújo, Pearson, Thuiller, & Erhard, 2005; Guisan & Thuiller, 2005). Other environmental variables might be identified in the future as data becomes available and could be included in future research of Headwater catfish distribution. Significant habitat alterations in portions of the study area make Headwater catfish persistence unlikely in some locations (e.g., the now desiccated Toyah Creek), and further surveys are necessary before local extinctions can be confirmed in these locations.

36

There is a high degree of morphological similarity between Headwater catfish and Channel catfish and this creates a disadvantage to identifying specimens in field collections. Although Kelsch (1995) correctly identified 91% of individuals on the basis of canonical discriminant analyses, this study only identified 36% of individuals correctly because most hybrids shared Headwater catfish phenotypes. This could mean that external morphological characters alone are insufficient for species identification in areas were hybrids occur. In these areas, proper identification of Headwater catfish relies on the collection of genetic data to address introgressed and hybrid individuals (Ketmaier & Bianco, 2015). Populations where introgression has taken place still remain areas of conservation priority and it is possible that hybridization has played an important role in maintaining this diversity (Demarais, Dowling, Douglas, Minckley, & Marsh, 1992; McClure-Baker et al., 2010). On the other hand, populations where the genetic integrity remains intact can be useful to facilitate gene flow and potentially reverse the process of range reduction and introgression (Crispo, Moore, Lee-Yaw, Gray, & Haller, 2011). Future work in the field of landscape genetics is promising towards this goal and information acquired from genetic diversity is essential to accurately assess future conservation planning and possible reintroduction initiatives (Malone et al., 2018).

The range contraction of Headwater Catfish documented here illustrates the need to conserve limited perennial water resources in arid riverscapes (Davis et al., 2017). Many tributaries in west Texas have become isolated from larger rivers, and headwater springs that once connected pools and perennial streams have shrunk to become isolated pools and ciénegas (Hoyt, 2002). Habitat loss in the form of declining spring outflows has a great effect on Headwater catfish occurrence and is one of the biggest threats to the species persistence (Bean et al., 2011; Kelsch & Hendricks, 1990; Figure 4c and d). With groundwater declines, flows from

headwater springs decrease and, in turn, cause reductions in flow to tributaries that serve as essential habitat for Headwater catfish. Another form of habitat loss comes from the decrease of surface water as shown in the models by the percent of watersheds dominated by water land cover between historical and contemporary periods (Figure 4g and h). Surface water loss is indicative of the changes aquatic ecosystems are facing in many parts of the American southwest. For example, Toyah Creek and the surrounding spring-fed system in the vicinity of Balmoreha once was an area of high Headwater catfish occurrence, but currently the creek is void of surface water due to the overdraft of groundwater and lowering of the water table (Sharp, 2001). In some circumstances, alteration of habitats for recreation or domestic water supply can result in indirect protection of springs and their biotas (Unmack & Minckley, 2008). The area once surrounded by multiple artesian springs now has diminished flows, but San Soloman Springs in Balmoreha State Park serves as a formal refuge where a Headwater catfish population, among other threatened desert fishes (Comanche Springs pupfish, *Cyprinodon elegans* and Pecos gambusia, *Gambusia nobilis*), persists. Las Moras springs create the headwaters of Las Moras Creek and was converted into a recreational swimming pool in 1971. However, just below the dammed pool a population of Headwater catfish was discovered where no recent records existed. Another refuge worth noting is the San Felipe Country Club where Headwater catfish were collected immediately downstream from the spring outflow. These human-altered springs not only serve as recreational areas but also as important biological refuges, and my analyses suggest these locations will play critical roles in preserving Headwater catfish in the United States.

Systematic conservation planning informs decision-makers of the most effective and efficient ways to achieve conservation goals (Hermoso et al., 2015). This study approach further demonstrates the potential for using systematic methods for conservation planning in freshwater

ecosystems to identify the locations of appropriate habitats, (2) the most likely current

distribution of Headwater catfish, (3) the locations of non-introgressed populations for

conservation, and (4) the distribution of haplotypes that can be used in genetic restoration and

management. The species distribution models generated here as well as the information acquired

from molecular data aide in identifying rehabilitation efforts and conservation in priority areas.

This study represents the most current evaluation of Headwater catfish in the United States with

emphasis on Texas, but I point out conservation status in Mexico requires additional research.

My findings emphasize the magnitude of change for natural aquatic systems in water-limited

environments and the need for on-the-ground management actions to achieve the most effective

conservation of intact populations of Headwater catfish. Declining spring flows and reduced

surface waters, competition with non-native species, and loss of genetic integrity due to

introgression and hybridization with introduced species all plague Headwater catfish persistence

in Texas. The status of this little-known species shares a similar story with that of the Yaqui

catfish (*Ictalurus pricei*) in that, without proper protection, it may face irreversible declines

(Stewart, Butler, Harris, & Radke, 2017). Maintaining balance between watershed management

for natural resources and anthropogenic usages remains a significant challenge. Even among

larger municipalities where there is a reasonably strong message of water conservation, the

surrounding rural areas show signs of increasing agricultural production and water consumption

(Edwards, Garrett, & Marsh-Matthews, 2002; Edwards et al., 2004). However, conservation

measures to protect aquatic habitats and native fishes in the region are in development with the

formation of Native Fish Conservations Areas (James, 2011; Williams et al., 2011), guidance

from organizations such as the Desert Fish Habitat Partnership, as well as increased public

awareness of the value of cooperative conservation.

REFERENCES

Aerts, R., & Honnay, O. (2011). Forest restoration, biodiversity and ecosystem functioning. *BMC Ecology*, 11(1), 29.

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677–1688.

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*(1), 42-–47.

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, *11*(9), 1504–1513.

Araújo, M. B., & Williams, P. H. (2000). Selecting areas for species persistence using occurrence data. *Biological Conservation*, *96*(3), 331–345.

Arlinghaus, R., Lorenzen, K., Johnson, B. M., Cooke, S. J., & Cowx, I. G. (2016). Management of freshwater fisheries: addressing habitat, people and fishes. *Freshwater Fisheries Ecology*, 1, 557-–579.

Barmuta, L. A., Linke, S., & Turak, E. (2011). Bridging the gap between 'planning' and 'doing' for biodiversity conservation in freshwaters. *Freshwater Biology*, 56(1), 180–195.

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10–15.

Bean, P. T., Jackson, J. T., McHenry, D. J., Bonner, T. H., & Forstner, M. R. (2011). Rediscovery of the headwater catfish *Ictalurus lupus* (Ictaluridae) in a western Gulf-Slope drainage. *The Southwestern Naturalist*, 56(2), 285–289.

Benke, A. C., & Cushing, C. E. (Eds.). (2011). *Rivers of North America*. Academic Press, San Diego, CA.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Budy, P., Conner, M. M., Salant, N. L., & Macfarlane, W. W. (2015). An occupancy-based quantification of the highly imperiled status of desert fishes of the southwestern United States. *Conservation Biology*, *29*(4), 1142–1152.

Cambray, J. A. (2003). Impact on indigenous species biodiversity caused by the globalisation of alien recreational freshwater fisheries. *Hydrobiologia*, 500(1–3), 217–230.

Castresana, J. (2001). Cytochrome b phylogeny and the taxonomy of great apes and mammals. *Molecular Biology and Evolution*, 18(4), 465–471.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.

Contreras-Balderas, S. & Escalante, M. A. (1984). Distribution and known impacts of exotic fishes in Mexico. In Courtenay, W.R. & Stauffer, J.R. (Eds), *Distribution, Biology, and Management of Exotic Fishes (*pp.102–130). Johns Hopkins University Press: Baltimore, MD.

Cooper, A. R., Infante, D. M., Daniel, W. M., Wehrly, K. E., Wang, L., & Brenden, T. O. (2017). Assessment of dam effects on streams and fish assemblages of the conterminous USA. *Science of the Total Environment*, *586*, 879–889.

Crispo, E., Moore, J. S., Lee-Yaw, J. A., Gray, S. M., & Haller, B. C. (2011). Broken barriers: human-induced changes to gene flow and introgression in animals: an examination of the ways in which humans increase genetic exchange among populations and species and the consequences for biodiversity. *BioEssays*, 33, 508–518.

D'amen, M., Rahbek, C., Zimmermann, N. E., & Guisan, A. (2017). Spatial predictions at the community level: from current approaches to future frameworks. *Biological Reviews*, 92(1), 169–187.

Davis, J. A., Kerezsy, A., & Nicol, S. (2017). Springs: conserving perennial water is critical in arid landscapes. *Biological Conservation*, 211, 30–35.

Demarais, B. D., Dowling, T. E., Douglas, M. E., Minckley, W. L., & Marsh, P. C. (1992). Origin of Gila seminuda (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proceedings of the National Academy of Sciences*, 89(7), 2747–2751.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., … Sullivan, C.A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*(2), pp.163–182.

Echelle, A. A. (1991). Conservation genetics and genic diversity in freshwater fishes of western North America. In W. L., Minckley & Deacon, J. E. (Eds.), *Battle Against Extinction* (pp. 141–153). University of Arizona Press, Tucson.

Edwards, R. J., Garrett, G. P., & Marsh-Matthews, E. (2002). Conservation and status of the fish communities inhabiting the Rio Conchos basin and middle Rio Grande, Mexico and USA. *Reviews in Fish Biology and Fisheries*, *12*(2-3), 119–132.

Edwards, R. J., Garrett, G. P., & Allan, N. L. (2004). Aquifer-dependent fishes of the Edwards Plateau region. In R.E., Mace, Angle, E.S., & Mullican, W.F. (Eds.), *Aquifers of the Edwards Plateau. Texas Water Development Board*, Austin, TX, 253–268.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., ... Li, J. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151.

Elith, J., & Leathwick, J. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813.

EPA (Environmental Protection Agency), USGS (United States Geological Survey), Horizon Systems Corporations (2010) Database: *NHDplusV1 Data*. Available at: http://www.horizon-systems.com/nhdplus/data.php.

EPA (Environmental Protection Agency), USGS (United States Geological Survey), Horizon Systems Corporations (2012) Database: *NHDplusV2 Data*. Available at: http://www.horizon-systems.com/nhdplus/NHDplusV2_data.php.

Erős, T., O'Hanley, J. R., & Czeglédi, I. (2018). A unified model for optimizing riverscape conservation. *Journal of Applied Ecology*, 55, 1871–1883

Evans, J.S., and Murphy, M.A. 2015. rfUtilities: random forests model selection and performance evaluation. *R package version* 1.0–2, http://cran.r–project.org/pack–age=rfUtilities.

Falcone, J.A. (2015). *U.S. conterminous wall-to-wall anthropogenic land use trends* (NWALT), 531 1974–2012: U.S. Geological Survey Data Series 948, 33 p. plus appendixes 3–6 as532 separate files. Available at: http://dx.doi.org/10.3133/ds948.

Frimpong, E. A., Sutton, T. M., Lim, K. J., Hrodey, P. J., Engel, B. A., Simon, T. P., ... Le Master, D. C. (2005). Determination of optimal riparian forest buffer dimensions for stream biota landscape association models using multimetric and multivariate responses. *Canadian Journal of Fisheries and Aquatic Sciences*, *62*(1), 1–6.

Gido, K. B., Whitney, J. E., Perkin, J. S., & Turner, T. F. (2016). Fragmentation, connectivity and fish species persistence in freshwater ecosystems. *Conservation of Freshwater Fishes*, 292–323.

Gilbert, C. R., & Burgess, G. H. (1980). *Ictalurus lupus* (Girard), headwater catfish. In Lee, D. S., Gilbert, C. R., Hocutt, C. H., Jenkins, R. E., McAllister, D. E., & Stauffer Jr, J. R. (1980). *Atlas of North American freshwater fishes* (pp. 146). North Carolina State Museum of Natural History. Raleigh, North Carolina.

Gray, J. S. (1997). Marine biodiversity: patterns, threats and conservation needs. *Biodiversity & Conservation*, *6*(1), 153–175.

Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers (2019). gbm: Generalized Boosted Regression Models. *R package version* 2.1.5. https://CRAN.R project.org/package=gbm

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009.

Heino, J., Virkkala, R., & Toivonen, H. (2009). Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biological Reviews*, *84*(1), 39–54.

Hermoso, V., Kennard, M. J., & Linke, S. (2012). Integrating multidirectional connectivity requirements in systematic conservation planning for freshwater systems. *Diversity and Distributions*, *18*(5), 448–458.

Hermoso, V., Linke, S., Januchowki-Hartley, S.R., & Kennard, M.J. (2015). Freshwater conservation planning. In Closs, G. P., Krkosek, M., & Olden, J. D. (Eds.), *Conservation of Freshwater Fishes* (pp. 437– 466). Cambridge University Press, Cambridge.

Hermoso, V., Abell, R., Linke, S., & Boon, P. (2016). The role of protected areas for freshwater biodiversity conservation: challenges and opportunities in a rapidly changing world. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *26*(S1), 3–11.

Heitmuller, F.T., & Reece, B.D. (2003). Database of historically documented springs and spring flow measurements in Texas: U.S. Geological Survey Open-File Report 03-315. Available for download through Databasin.org [accessed May 8 2018]

Hoyt, C. A. (2002). The Chihuahuan Desert. *Endangered Species Bulletin*, *27*, 03–06.

Huang, J., & Frimpong, E. A. (2015). Using historical atlas data to develop high-resolution distribution models of freshwater fishes. *PloS One*, *10*(6), e0129995.

Hubbs, C., Edwards, R. J., & Garrett, G. P. (1991). An annotated checklist of the freshwater fishes of Texas, with keys to identification of species. *Texas Journal of Science*, Supplement 43: 1–56.

James R. (2011). Native fish conservation areas: a vision for large-scale conservation of native fish communities. *Fisheries*. 36(6): 267–277.

James, J. J., Sheley, R. L., Erickson, T., Rollins, K. S., Taylor, M. H., & Dixon, K. W. (2013). A systems approach to restoring degraded drylands. *Journal of Applied Ecology*, 50(3), 730–739.

Jelks, H. L., Walsh, S. J., Burkhead, N. M., Contreras-Balderas, S., Diaz-Pardo, E., Hendrickson, D. A., ... & Platania, S. P. (2008). Conservation status of imperiled North American freshwater and diadromous fishes. *Fisheries*, 33(8), 372–407.

Karr, J. R. (1981). Assessment of biotic integrity using fish communities. *Fisheries*, 6(6), 21–27.

Kelsch, S. W. (1995). Patterns of morphometric variation in the channel and headwater catfishes. *Transactions of the American Fisheries Society*, *124*(2), 272–279.

Kelsch, S. W., & Hendricks, F. S. (1986). An electrophoretic and multivariate morphometric comparison of the American catfishes *Ictalurus lupus* and *I. punctatus*. *Copeia*, 646–652.

Kelsch, S., & Hendricks, F. (1990). Distribution of the Headwater Catfish *Ictalurus lupus* (Osteichthyes: Ictaluridae). *The Southwestern Naturalist, 35*(3), 292–297. doi:10.2307/3671942

Ketmaier, V., & Bianco, P.G. (2015). Understanding and conserving genetic diversity in a world dominated by alien introductions and native transfers: the case study of primary and peripheral freshwater fishes in southern Europe. In Closs, G. P., Krkosek, M., & Olden, J. D. (Eds.), *Conservation of Freshwater Fishes* (pp. 506–534). Cambridge University Press, Cambridge.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26.

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News* 2(3), 18–22.

Malone, E. W., Perkin, J. S., Leckie, B. M., Kulp, M. A., Hurt, C. R., & Walker, D. M. (2018). Which species, how many, and from where: Integrating habitat suitability, population genomics, and abundance estimates into species reintroduction planning. *Global Change Biology*, *24*(8), 3729–3748.

Mateo, R. G., Croat, T. B., Felicísimo, Á. M., & Muñoz, J. (2010). Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions, 16*(1), 84–94.

Meffe, G. K., & Vrijenhoek, R. C. (1988). Conservation genetics in the management of desert fishes. *Conservation Biology*, *2*(2), 157–169.

McClure-Baker, S. A., Echelle, A. A., Van den Bussche, R. A., Echelle, A. F., Hendrickson, D. A., & Garrett, G. P. (2010). Genetic status of headwater catfish in Texas and New Mexico: a perspective from mtDNA and morphology. *Transactions of the American Fisheries Society*, 139(6), 1780–1791.

McKee, J. K., Sciulli, P. W., Fooce, C. D., & Waite, T. A. (2004). Forecasting global biodiversity threats associated with human population growth. *Biological Conservation*, 115(1), 161–164.

Miller, R. R., Minckley, W. L., Norris, S. M., & Gach, M. H. (2005). *Freshwater Fishes of Mexico* (No. QL 629. M54 2005). Chicago: University of Chicago Press.

Moritz, C. (1995). Uses of molecular phylogenies for conservation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *349*(1327), 113–118.

Moyle, P. B., & Leidy, R. A. (1992). Loss of biodiversity in aquatic ecosystems: evidence from fish faunas. In *Conservation Biology* (pp. 127–169). Springer, Boston, MA.

Murray, B.B.R., Zeppel, M.J., Hose, G.C., & Eamus, D. (2003). Groundwater-dependent ecosystems in Australia: It's more than just water for rivers. *Ecological Management & Restoration*, 4(2), pp.110–113.

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environmental Management*, 54(6), 1249–1266.

Orians, G. H. (1995). Thought for the morrow: cumulative threats to the environment. *Environment: Science and Policy for Sustainable Development*, 37(7), 6–36.

Perkin, J. S., Gido, K. B., Costigan, K. H., Daniels, M. D., & Johnson, E. R. (2015). Fragmentation and drying ratchet down Great Plains stream fish diversity. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *25*(5), 639–655.

Perkin, J. S., Starks, T. A., Pennock, C. A., Gido, K. B., Hopper, G. W., & Hedden, S. C. (2019) Extreme drought causes fish recruitment failure in a fragmented Great Plains riverscape. *Ecohydrology*, e2120.

Perkin, J. S., Murphy, S. P., Murray, C. M., Gibbs, W. K., & Gebhard, A. E. (2019). If you build it, they will go: A case study of stream fish diversity loss in an urbanizing riverscape. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(4), 623–638.

Poiani, K. A., Richter, B. D., Anderson, M. G., & Richter, H. E. (2000). Biodiversity conservation at multiple scales: functional sites, landscapes, and networks. *BioScience*, 50(2), 133–146.

Possingham, H.P., Bode, M., & Klein, C.J. (2015). Optimal conservation outcomes require both restoration and protection. *PLoS Biology* 13(1): e1002052.

R Core Team (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL http://www.R-project.org/.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T., ... & Smol, J. P. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873.

Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., … Leemans, R., 2000. Global biodiversity scenarios for the year 2100. *Science*, 287(5459), pp.1770–1774.

Saunders, D. L., Meeuwig, J. J., & Vincent, A. C. J. (2002). Freshwater protected areas: strategies for conservation. *Conservation Biology*, *16*(1), 30–41.

Sharp Jr, J. M. (2001). Regional groundwater flow systems in Trans-Pecos Texas. *Texas Water Development Board Report*, *356*(356), 41–55.

Schruben, P. G., Arndt, R. E., Bawiec, W. J., King, Philip B., & Beikman, H. M. (1994). Geology of the Conterminous United States at 1:2,500,000 Scale -- A Digital

Representation of the 1974 P.B. King and H.M. Beikman Map: *U.S. Geological Survey Digital Data Series DDS-11*, U.S. Geological Survey, Reston, VA. Available at: https://mrdata.usgs.gov/geology/us/

Smith, K. L., Corujo Flores, I., & Pringle, C. M. (2008). A comparison of current and historical fish assemblages in a Caribbean island estuary: conservation value of historical data. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18(6), 993–1004.

Souza, V., Espinosa-Asuar, L., Escalante, A. E., Eguiarte, L. E., Farmer, J., Forney, L., ... Elser, J. J. (2006). An endangered oasis of aquatic microbial biodiversity in the Chihuahuan desert. *Proceedings of the National Academy of Sciences*, *103*(17), 6565–6570.

Stewart, D. R., Butler, M. J., Harris, G., & Radke, W. R. (2017). Mark-recapture models identify imminent extinction of Yaqui catfish *Ictalurus pricei* in the United States. *Biological Conservation*, 209, 45–53.

Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *Eos, Transactions American Geophysical Union*, 38(6), 913–920.

Strayer, D. L., & Dudgeon, D. (2010). Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, 29(1), 344–358.

Sublette, J.E., Hatch, M.D., & Sublette, M. (1990). *The Fishes of New Mexico*. University of New Mexico Press, Albuquerque, New Mexico.

Torgo, L. (2011). *Data mining with R: learning with case studies*. Chapman and Hall/CRC Press.

United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), the United States Geological Survey (USGS), and the Environmental Protection Agency (EPA). (2015). Watershed Boundary Dataset for Texas and New Mexico. Available at: https://databasin.org/datasets/f57141fecb9849bfa7ccd6ad6634c118

Unmack, P. J., & Minckley, W. L. (2008). The demise of desert springs. In Stevens, L.E. & Meretsky, V.J. (Eds), *Aridland springs in North America: ecology and conservation*. University of Arizona Press, Tucson, pp.11–34.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., ... Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, 467(7315), 555.

Ward, J. (1998). Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. *Biological Conservation*, 83(3), 269–278.

Ward, F. A., Booker, J. F., & Michelsen, A. M. (2006). Integrated economic, hydrologic, and institutional analysis of policy responses to mitigate drought impacts in Rio Grande Basin. *Journal of Water Resources Planning and Management*, 132(6), 488–502.

Wellemeyer, J. C., Perkin, J. S., Fore, J. D., & Boyd, C. (2018). Comparing assembly processes for multimetric indices of biotic integrity. *Ecological Indicators*, 89, 590–609.

Williams, J. E., Johnson, J. E., Hendrickson, D. A., Contreras-Balderas, S., Williams, J. D., Navarro-Mendoza, M., ... & Deacon, J. E. (1989). Fishes of North America endangered, threatened, or of special concern: 1989. *Fisheries*, 14(6), 2–20.

Williams, J.E., Williams, R.N., Thurow, R.F., Elwell, L., Philipp, D.P., Harris, F.A., … Frissell, C.A. (2011). Native Fish Conservation Areas: A Vision for Large-Scale Conservation of Native Fish Communities. *Fisheries*, *36*(6), pp.267–277.

Wilson, E. O. (1989). Threats to biodiversity. *Scientific American*, 261(3), 108–11