A Landscape Approach to Fisheries Database Compilation and Predictive Modeling

Submitted to: Bureau of Reclamation Agreement Number: R12AP80917
Prepared by: William Stewart and Nicole Eiden, Arizona Game and Fish Department and Julian Olden, University of Washington

June 2015

Acknowledgements

We would like to thank Jessica Latzko (Arizona Game and Fish) and Keith Fritschie (University of Washington) for their tedious effort at compiling and cleaning up datasets. We would also like to thank Taylor Cotton for coordinating field surveys along with his technicians Jonathan Miller and Chad Rubke. Additionally we would like to thank Dave Weedman, Mike Lopez, Antonio Lopez, Scott Rogers, Matt Rinker, Chuck Benedict, Matt Chmiel, Greg Cummins and Curtis Gill for their valuable expertise and input into this project.

Suggested citation:

Stewart, W. T., N.L. Eiden, and J.D Olden. 2015. A Landscape Approach to Fisheries Database Compilation and Predictive Modeling. Final Report submitted to the Bureau of Reclamation, Denver, Colorado, by Arizona Game and Fish Department, Phoenix, Arizona 68p.

Executive Summary

Fisheries data compilation efforts for this project fell within two large watersheds in Arizona; the Verde River watershed (Desert LCC) and the Little Colorado River watershed (Southern Rockies LCC). We divided the project into two phases; 1) data compilation for the Arizona Game and Fish Fisheries Information Systems (FINS) and 2) a demonstration of FINS through model development and species distribution data. During phase 1, we compiled, cleaned, assigned National Hydrography Dataset (NHD) reachcodes to historical data for 113,230 fish records in the Verde River watershed and 43,828 fish records from the Little Colorado River watershed. These records were standardized to meet the Arizona Game and Fish Department FINS standards. From these records we developed three types of geospatial data layers: 1) individual species, 2) reach information, and 3) fish collection records.

In phase 2, we demonstrated the utility of FINS by employing a multivariate adaptive regression splines method to model species presence/absence throughout the entire Verde River watershed. Catchments boundaries were used to represent the area of land contributing to individual stream segments as defined by the NHD+ database, and represent an ecologically relevant scale to quantify environmental factors likely to influence local species occurrence. Environmental and physical attributes associated with each catchment was used to predict presence/absences of species across the watershed. Distribution models were created for 12 native fish species and 18 non-native fish. To test the models accuracy, we electro-fished, seined and/or set hoop nets at 58 sites that had not been previously surveyed. The models ability to predict presence/absence was good. With the exception of rainbow trout, the model predicted species presence/absence in greater than 86% of the previously unsurveyed location.

Contents

Introduction 1
Verde River Watershed 1
Little Colorado River Watershed 2
Fisheries Information System (FINS) 5
Previous Data Compilation Efforts 5
Previous model developments 6
Methods 7
Use of Established Data Standards (Phase 1) 7
Data compilation methods 7
Geodata layer development 8
Model development (Phase 2) 8
Site selection for ground truthing 15
Results 15
Data compilation 15
Verde Watershed 15
Little Colorado River Watershed 18
Geodata layer development 19
Model results 21
Ground truthing results 23
Conclusion/Discussion 25
Utility of data compilation efforts 26
Where to Get the Data 27
References 28
APPENDICES 31

Introduction

This report is in fulfillment agreement number R12AP80917 titled A Landscape Approach for Fisheries Database Compilation and Predictive Modeling. We divided the project into two phases; 1) data compilation for the Arizona Game and Fish Fisheries Information Systems (FINS) and 2) a demonstration of FINS through model development and species distribution data.

Addressing the management and conservation challenges for native and sport fishes will require the ability to data mine the extensive existing information on distribution and abundance of species available from Arizona Game and Fish Department's (AZGFD) ten sport and native fish programs and many other government and non-governmental sources. Such a synthesis can be used to assess the current conservation status of native fishes, quantify the extent of species invasions, and establish baseline distributions with which to evaluate the effects changing environments.

Arizona's highly endemic ichthyofauna has declined over the 20th century. Historically, 36 species of native fish occurred in varied habitats ranging from small springs to raging flood waters of the lower Colorado River. One native fish species has become extinct, 20 are federally listed under the Endangered Species Act, and 34 are recognized as Species of Greatest Conservation Need under Arizona's State Wildlife Action Plan (AGFD 2012). Aquatic habitat loss and alteration through irrigation diversions, dams, and groundwater pumping creates challenges for fisheries management for both sportfish and native fish not only in Arizona, but the entire arid southwestern region of the United States.

The information here describes the compilation effort for data that spans 60 years. Data that can be used to inform managers about current and future threats identify areas that lack connectivity, and help identify priority areas for conservation of species and water resources. For the purposes of this project, we focus data compilation efforts on two large watersheds in Arizona; the Verde River watershed and the Little Colorado River (LCR) watershed. Each of these watersheds fall within the boundaries of the Desert (Verde) and Southern Rockies (LCR) LCCs.

Verde River Watershed

The Verde River watershed is located in north-central Arizona and drains an area of approximately 6,623 square miles (Figure 1). It is the ninth largest of the 18 watersheds (6-digit Hydrologic Unit Categories defined by USGS) in Arizona. The area includes the 2,153 square miles of the Big Chino Wash watershed that drains into the Verde River mainstem upstream of Sullivan Dam. Perennial sections of the mainstem Verde River are commonly divided by fisheries managers into three reaches: the upper Verde is from Sullivan Dam to the TAPCO Diversion Dam; the middle Verde is from TAPCO Diversion Dam to Beasley Flats; and the lower Verde is from Beasley Flat to the confluence with the Salt River at the most downstream end of the watershed.

The upper Verde River flows eastward from Sullivan Dam to Perkinsville, and then flows southeastward to the TAPCO Diversion Dam. The Verde River is intermittent immediately below Sullivan Dam and water

begins pooling at the upstream end of Stillman Lake, about 1 mile downstream of Sullivan Dam. Perennial flow extends approximately 140 miles from this area to its confluence with the Salt River just east of Phoenix. In the upper reach, numerous agricultural diversions exist. The TAPCO Diversion Dam, also called the Pecks Lake Diversion near the community of Clarkdale, is a significant diversion owned by the Freeport-McMoRan Corporation. Within the middle reach other diversion dams provide water for the following ditch systems: Cottonwood, Woods, Ok and Eureka.

The lower Verde River, beginning at Beasley Flats, becomes more isolated as it winds through the Prescott, Coconino, and Tonto National Forests. Fossil Creek enters the Verde River approximately 24 miles from Beasley Flats. The river continues in a southerly direction before it reaches Horseshoe and Bartlett reservoirs, two large impoundments operated by the Salt River Project to provide water to central Arizona. From Bartlett Reservoir the Verde continues through the Fort McDowell Indian Reservation until it meets the Salt River.

Little Colorado River Watershed

The Little Colorado River (LCR) watershed is located in north-eastern Arizona and drains an area of approximately 21,707 square miles (Figure 2). It is the largest of the 18 watersheds (6-digit Hydrologic Unit Categories defined by USGS) in Arizona. Major, north-flowing basins produce most of the streamflow in the LCR. They are the Upper LCR above Woodruff, Silver Creek, Chevelon Creek, and Clear Creek. The major water-bearing tributaries in the eastern part of the LCR watershed arise in the volcanic rocks of the White Mountains. These include Nutrioso Creek, the South, East, and West forks of the LCR, and Silver Creek. To the west, two basins--Chevelon Creek and Clear Creek—arise in the marine sedimentary rocks which emerge at the Mogollon Rim. The main LCR drainage continues north and becomes intermittent from Lyman Lake until it reaches blues springs approximately 10 miles upstream from the confluence of the Colorado River. A large portion of the lower section of the Little Colorado River watershed is on tribal land.

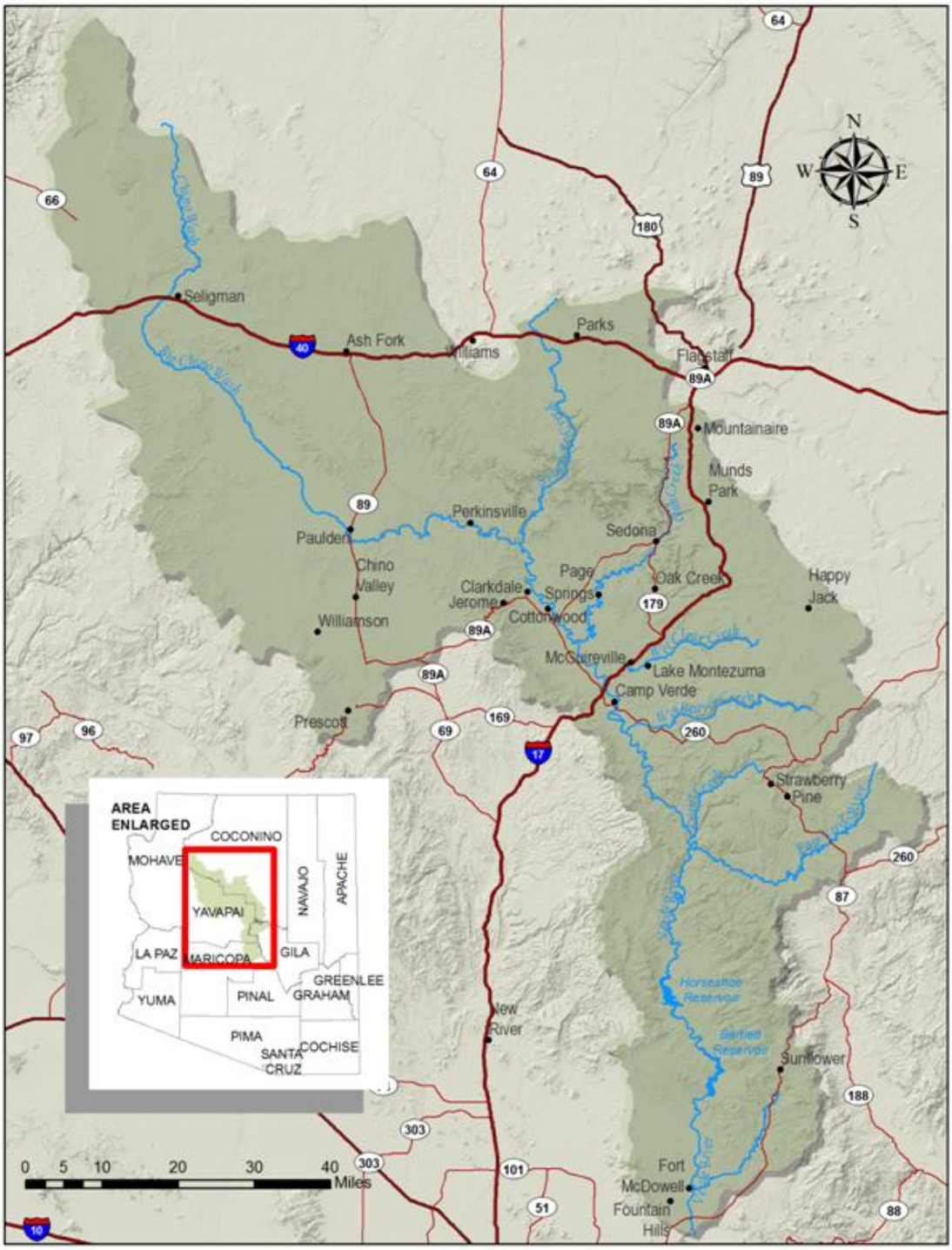


Figure 1 Verde River Watershed in north central Arizona.

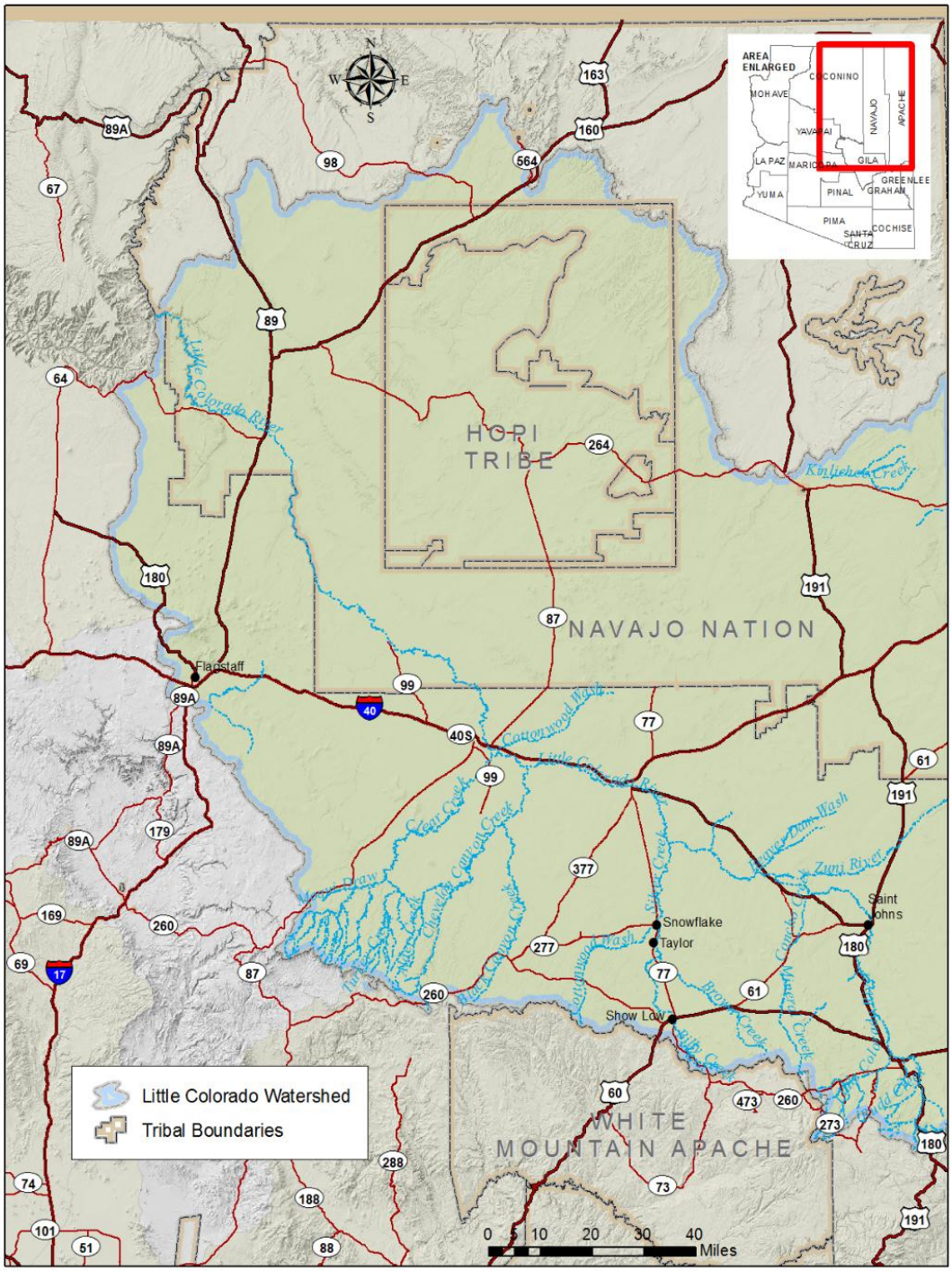


Figure 2. Little Colorado River Watershed in north-eastern Arizona.

Fisheries Information System (FINS)

AZGFD has a wealth of fish records with decades of fish surveys that are currently contained in reports and unpublished datasets that are available to be compiled. In 2010, AZGFD assessed the feasibility of developing a comprehensive fisheries data management system to support watershed-based fisheries management planning. This planning initiative balances AZGFD's dual mandate of managing for sport fish recreation while recovering native aquatic populations.

The AZGFD recognizes the need for a strong data foundation to inform science-based decisions for fisheries management at a watershed level. In preparation for a shift towards comprehensive watershed-scale planning, AZGFD committed staff and financial resources towards developing Fisheries Information System (FINS). FINS is an enterprise-level information system, allowing data to be entered in the field and easily retrieved by managers across the agency.

The first module of FINS went live in January 2015, and five state hatcheries, as well as the urban stocking program, are recording sport fish stockings. Native fish stockings went online in May 2015 with the release of FINS version 1.5.0. Fish stocking data is no longer siloed in disparate individual program datasets (spreadsheets, MS Access databases) but rather is accessed via a single web-based application that supports not only data entry, but also data filtering and export, and report generation. Import of historic fish stocking records will be complete by June 2015.

With the FINS infrastructure in place, the challenge was finding the resources to standardize and populate historical survey data gathered from hundreds of sources throughout Arizona. While the FINS-Fish Stockings was being developed, phase 1 of this project agreement focused on compiling and formatting several hundred thousand fish observation records dating back to the 1950s. FINS integrates fish data into a Geographic Information System (GIS) by georeferencing observations to an existing national spatial framework (see National Hydrography Dataset). This allows broader transferability to watersheds shared with neighboring states, creating a seamless layer not limited by state boundaries.

Previous Data Compilation Efforts

One of the challenges to effective science-based decision-making in Arizona is the lack of a high-quality, accessible data that is continually updated over the long-term. Several recent data compilation efforts have resulted in large compilations of fish records for the desert southwest.

The Sonoran Fishes (SONFISHES) database was initially developed by the late ichthyologist W.L. Minckley. SONFISHES provides GIS-compatible native fish occurrences for the desert southwest and northwestern Mexico over a 160 year timespan (1843-2005). Data sources include museum records, the AZGFD Nongame Wildlife Branch Native Fish Database, and peer-reviewed or grey literature sources. In 2005, this was the most comprehensive dataset for native fishes in the desert southwest.

Another effort was initiated in 2004 by the USGS Gap Analysis Program, a national program that develops key datasets to conduct coarse–filter assessments for the

conservation of biological diversity within terrestrial and aquatic ecosystems (Whittier et al. 2011). The primary goal of the Lower Colorado River Basin Aquatic Gap Analysis Project (LCRGAP) was to identify riverine fish species that were inadequately represented within the existing network of protected lands in the Lower Colorado River Basin (LCRB). Fish location data were compiled from several state and federal agencies, natural heritage programs, universities, and museums. Records included at the minimum: species name, number captured, location, date, and year. Excluding hybrids, the LCRB database contains 771,600 records for non-native species and 1,005,905 records for native species. The total number of sites sampled is 104,486. About 60 of the non-native species have established in the LCRB. Several native species have been extirpated or are now extinct. The earliest records are from the 1700's with about 80% occurring after 1979. This project was completed in 2011.

While the SONFISHES and LCRGAP efforts produced large datasets, these efforts were project-based and did not have a long-term objective to keep the data updated with new records. The fish sampling data used for this project is referenced spatially to the existing National Hydrography Dataset, resulting in a relational database of fish observations for individual reachcodes. In addition to waterbodies and flowlines, the NHD data also includes attributes to characterize waters and catchments at the local and watershed levels. These characteristics can be used to model aquatic species affinities based on preferred habitat values.

An example of a national effort to serve as a repository for fish observations is the Multistate Aquatic Resources Information System (MARIS). MARIS is a cooperative effort between state and federal agencies to share fisheries information collected as part of ongoing sampling programs. MARIS data is owned and provided by participating state natural resource management agencies, while technical support and hosting are currently provided by the Core Science Analytics and Synthesis Program of the U.S. Geological Survey. Currently 25 states participate in MARIS by submitting fish sampling and/or water quality data.

Previous model developments

The benefits of a system such as FINS is that is allows user to develop models to predict species distribution and forecast future trends. A particular promising approach involves ecological niche modeling in conjunction with the power of GIS to predict the geography of species' invasions (Elith and Leathwick 2009). With the recent development of conservation tools specifically oriented to aquatic systems (Hermoso et al. 2011), conservation planning and the development of freshwater protected areas have become a burgeoning field of research, incorporating aspects of complementarity, connectivity, and network prioritization specific to riverine systems (Abell et al. 2007, Lawrence et al. 2011). Systematic conservation planning identifies areas with high native species diversity and complementarity that are under minimal threat from habitat loss and fragmentation, flow regime modification, and non-native species introductions. Such areas can help guide regional and local management

efforts to conserve or restore critical habitat for native fishes (Strecker et al. 2011). However, despite recent advances, systematic conservation planning of freshwater ecosystems remain limited because of the paucity of large and spatially-extensive databases. Here, we will demonstrate the utility of the FINS to inform conservation planning outcomes (Phase 2).

Methods

Use of Established Data Standards (Phase 1)

We used established national data standards wherever possible during data compilation to facilitate data sharing. Data standards use rules to ensure consistency across multiple sources, resulting in greater return on investment since more resource managers can use the data, rather than just the project that collected the data.

FINS uses the following standards to help ensure interoperability within AZGFD and external agencies:

- **Integrated Taxonomic Information System (ITIS):** ITIS provides authoritative taxonomic information on plants, animals, fungi, and microbes of North America and the world. Fish and other aquatic species observations are referenced within FINS by its authoritative ITIS name and classification to minimize ambiguity in species names between data sources. http://www.itis.gov
- **National Hydrography Dataset (NHD):** A geospatial framework containing features organized into polygons, lines and points to represent lakes and ponds, streams and rivers and springs (http://nhd.usgs.gov/). Surface water features are labeled with nationally unique and permanent identifiers known as reach codes. Reach codes are used in FINS to link fish observations to a stream or a lake. GIS features are labeled with a place name approved by the U.S. Board of Geographic Names (BGN) and implemented through the Geographic Names Information System.
- **Geographic Names Information System (GNIS):** The Geographic Names Information System (GNIS) is the federal and national standard for geographic nomenclature. The U.S. Geological Survey developed the GNIS in support of the U.S. Board on Geographic Names as the official repository of domestic geographic names data. GIS features in FINS are named using the authoritative GNIS name.

In addition to these national standards, the AZGFD has a standard sampling protocol (AGFD 2004) for fish surveys that provide guidelines to AZGFD biologists for collecting, analyzing, reporting fish data.

Data compilation methods

We compiled fisheries data through 2013 by contacting each of the individual fisheries programs that conduct fish surveys in the Verde (Desert LCC) and LCR (Southern Rockies) watersheds. If the data was not in digital format, the data was entered from field notebooks and data sheets. Once the data were compiled into a common format, we standardized each individual fish observation record to match the established FINS database standards. For

example, species names were coded to established FINS species codes based on ITIS standards. Typically the species code is the first two letters of the genus and species name i.e. Razorback Sucker *(Xyrauchen texanus)* is coded to XYTE (Appendix A).

Once tabular data is compiled and standardized, reachcodes are added. Each observation record is georeferenced to the National Hydrography Dataset using an NHD reachcode. We used a subset of the NHD GIS dataset by selecting features where a fish record is found. NHD GIS features were stored and edited in a multiuser Esri ArcSDE SQL Server environment. The NHDPlus feature dataset for Hydrography was subset for the purposes of addressing fish events. The features were checked for accuracy, and stream names added based on the Geographic Names Information System and verified against USGS 7.5" topological maps.

Geodata layer development

For the purposes of geodata layer development, we selected FINS attributes that provide information about species type, time, location, how the species was collected and frequency of surveys. Because some surveys happen over multiple days or weeks, we defined a survey in any given reachcode to be unique if the time between surveys was greater than 30 days. Additionally, we used each unique survey to summarize the frequency in which reachcodes were being sampled and summed up the number of times each species was positively identified in each reachcode. To ensure we had a robust number of observations we chose to develop data layers for species (including hybrids or fish identified to family) with greater than 1,000 records. Finally for each species we calculated the proportion of positive identifications in a reachcode over the number of times that reach code was sampled. This was done for each watershed with historical data through 2013.

Model development (Phase 2)

Data Description

In order to develop robust species distribution models, we selected fish collection events from the FINS Database based on the following criteria. First, we chose collection events that occurred post-1980 to represent a time period reflecting present-day fish assemblages subsequent to major physical alterations to waterways in the Lower Colorado River Basin (Fagan et al. 2002, Olden and Poff 2005). Second, we selected collection events that occurred in lotic (flowing-water) channel reaches. Third, we included collection events where at least one species had been observed assuming that these indicated survey efforts aimed to quantify community composition and not just population trends. We acknowledge that not all of the events included in our analysis may have been sampled for the entire fish community; however, it is not unexpected that many sites may have very few species given the low diversity that generally characterize streams of the Lower Colorado River Basin (Pool et al. 2010) and specifically the Verde River Basin (Rinne 2012, Gibson et al. 2014). Analyses conducted using collection events reporting two or more fish species produced very similar results and for the sake of brevity are not shown here. Fish taxonomy followed Minckley and Marsh (2009) such that records of hybrids were discarded and unrecognized subspecies

were grouped with the higher-level species designation. We also excluded fish records classified as reintroductions.

Collection events were assigned to reachcode identifiers according to the National Hydrography Dataset Plus stream network (1:100,000 scale). Only the most recent collection event to occur in each reachcode was retained. Next, each collection event was assigned to a GAP catchment identifier representing hydro-geomorphologic discrete watershed units. Catchments boundaries represent the area of land contributing to individual stream segments as defined by the NHD+ database, and represent an ecologically relevant scale to quantify environmental factors likely to influence local species occurrence. We refer the reader to Whittier et al. (2011) for additional information regarding the delineation of catchments. In the rare cases where a catchment containing multiple reach codes with fish collection data, a species was considered present in that catchment if it occurred in any of the reachcode collection events. After enforcing the data selection protocols above, our final dataset contained data for 12 native species (Table 1) and 18 nonnative species (Table 2) collectively occupying 323 catchments in the Verde River Basin.

Table 1. Native fish species (n=12) recorded in the Verde River Basin from 1980 to present according to the AZGFD FINS Database. Species included in the MARS model are indicated; all other species are included as point occurrences in the mapping of species distributions.

Family	Scientific name	Common name	Code	Modeled
Catostomidae	*Catostomus clarkii*	desert sucker	CACL	Y
	Catostomus insignis	Sonora sucker	CAIN	Y
	Xyrauchen texanus	razorback sucker	XYTE	Y
Cyprinidae	*Agosia chrysogaster*	longfin dace	AGCH	Y
	Gila intermedia	Gila chub	GIIN	Y
	Gila nigra	headwater chub	GINI	Y
	Gila robusta	roundtail chub	GIRO	Y
	Meda fulgida	spikedace	MEFU	
	Ptychocheilus lucius	Colorado pikeminnow	PTLU	
	Rhinichthys osculus	speckled dace	RHOS	Y
Poeciliidae	*Poeciliopsis occidentalis*	Gila topminnow	POOC	Y
Salmonidae	*Oncorhynchus gilae*	Gila trout	ONGI	

Table 2. Non-native fish species (n=18) recorded in the Verde River Basin from 1980 to present according to the Arizona Game and Fish Department Database Species. Species included in the MARS model are indicated; all other species are included as point occurrences in the mapping of species distributions.

Family	Scientific name	Common name	Code	Modeled
Centrarchidae	*Lepomis cyanellus*	green sunfish	LECY	Y
	Lepomis macrochirus	bluegill	LEMA	Y
	Micropterus dolomieu	smallmouth bass	MIDO	Y
	Micropterus punctulatus	spotted bass	MIPU	
	Micropterus salmoides	largemouth bass	MISA	Y
	Pomoxis annularis	white crappie	POAN	
Cyprinidae	*Cyprinella lutrensis*	red shiner	CYLU	Y
	Cyprinus carpio	common carp	CYCA	Y
	Pimephales promelas	fathead minnow	PIPR	Y
Ictaluridae	*Ameiurus melas*	black bullhead	AMME	Y
	Ameiurus natalis	yellow bullhead	AMNA	Y
	Ameiurus nebulosus	brown bullhead	AMNE	
	Ictalurus punctatus	channel catfish	ICPU	Y
	Pylodictis olivaris	flathead catfish	PYOL	Y
Poeciliidae	*Gambusia affinis*	Western mosquitofish	GAAF	Y
Salmonidae	*Oncorhynchus mykiss*	rainbow trout	ONMY	Y
	Salmo trutta	brown trout	SATR	Y
	Salvelinus fontinalis	brook trout	SAFO	

Environment Data

We modeled species occurrence in each catchment as a function of eight predictor variables describing channel topography, watershed hydrology, land use, and regional climate (Table 3). These variables were selected based on their demonstrated importance in shaping patterns of fish species occurrence in the region (Pool et al. 2010, Strecker et al. 2011, Pool et al. 2013).

Channel topography variables included Shreve link value as a measure of stream size (NHD+, http://www.horizon-systems.com/nhdplus/); specifically, a measure of stream ordering by magnitude (Shreve 1966) where all links with no tributaries are assigned a magnitude (order) of one and magnitudes are additive downslope (i.e., when two links intersect, their magnitudes are added and assigned to the downslope link). We also included average elevation of stream segment (m) and the Topographic Wetness Index representing the topographic control on hydrological processes and ranging from 0 (xeric) to 1 (mesic) (Theobald 2007). Land use and watershed physiographic variables included agricultural land use (%) in the upstream contributing watershed and urban land use (%) in the upstream contributing watershed based on the National Land Cover Database (Multi-Resolution Land Characteristics Consortium 2001). Variables describing regional climatic conditions included mean US annual precipitation (mm); variation in spring precipitation calculated as the coefficient of variance (CV) in monthly upstream precipitation for March to April), and mean annual temperature (degrees Celsius) calculated as the midpoint of maximum and minimum air temperatures recorded across each watershed. All precipitation and temperature variables were based on records from 1971 to 2000. Data source was Climate Source (USDA – National Resources Conservation Service 2007).

Table 3. Predictor environmental variables used in the species distribution models (MARS). Transformations were considered but deemed not necessary to meet assumptions of variable normality.

Category	Metric (Abbreviation)	Definition	Reference
Topography	Shreve stream order (StOrd)	measure of stream size	National Hydrography Dataset Plus (2014)[1]
	Gradient (Grad)	gradient or slope (m/m)	US Geological Survey (2004)[5]
Hydrology	Topographic Wetness Index (TWetI)	length weighted mean of the topographic wetness index	Theobald (2007)[3]
Land use	Agriculture (Agric)	proportion of land that is in agricultural use	Multi-Resolution Land Characteristics (2001)[2]
	Urban (Urban)	proportion of land that is in urban use	Multi-Resolution Land Characteristics (2001)[2]
Climate	Mean temperature (TempMn)	average annual temperature (°C)	US Department of Agriculture (2007)[4]
	Mean precipitation (PrecMn)	average annual precipitation (mm)	US Department of Agriculture (2007)[4]
	Spring precipitation CV (PrecSCV)	coefficient of variation (CV) for spring precipitation (March – April)	US Department of Agriculture (2007)[4]

[1]National Hydrography Dataset Plus. 2014. http://www.horizon-systems.com/nhdplus/
[2]Multi-Resolution Land Characteristics. 2001. National land cover database. Multi-Resolution Land Characteristics Consortium. Washington, USA. http://www.mrlc.gov/
[3]Theobald, D.M. 2007. LCaP v1.0: landscape connectivity and pattern tools for ArcGIS. Colorado State University, Fort Collins, USA.
[4]United States Department of Agriculture. 2007. PRISM climate mapping project. United States Department of Agriculture - Natural Resources Conservation Service. Corvallis, USA. http://www.prism.oregonstate.edu/
[5]United States Geological Survey. 2004. The national elevation dataset. United States Geological Survey. Washington, USA. http://ned.usgs.gov/

Other potential environmental variables were excluded as a result of (1) demonstrating high collinearity with those selected above or (2) found to be poor predictors based on a series of preliminary analyses. We assessed variable collinearity by computing the variance inflation factor (VIF), where VIF values greater than five are considered to be indicative of high collinearity (Menard 1995, O'Brien 2007). None of our environmental variables had VIF values greater than this conservative threshold, indicating an acceptable level of collinearity in subsequent analyses (Figure 3). We tested for spatial dependency among our environmental variables in geographic space by regressing the pairwise similarity in environmental characteristics between catchments (based on Euclidean distance) against the geographical distance separating them (according to Euclidean distance between centroids of each catchment) using a Mantel test. We found very little spatial structure in our environmental variables among catchments ($r = 0.081$, $P = 0.143$), suggesting that catchments could be treated as independent observations in subsequent statistical analyses. Summary statistics are reported in Table 4.

Table 4. Summary statistics for environmental variables used in the MARS modeling. Variable descriptions provided in Table 3.

Variable	Min	1st q.	Med	Mean	3rd q.	Max
Shreve (stream order)	1.0	15.0	82.0	504.6	878.5	1971.0
Elevation (m)	452.0	827.0	1040.0	1070.0	1256.0	2043.0
Topographic Wetness I.	7.8	13.3	15.4	16.2	18.5	32.2
Agriculture (%)	0.0	0.0	0.0	0.0	0.1	0.3
Urban (%)	0.0	0.5	1.0	1.1	1.0	22.0
Mean Temperature (deg C)	7.5	11.4	12.2	12.5	12.9	20.1
Mean Precipitation (mm)	34893.0	47076.0	53345.0	56410.0	66487.0	92770.0
Spring Precipitation CV	0.0	0.1	0.2	0.2	0.4	0.8

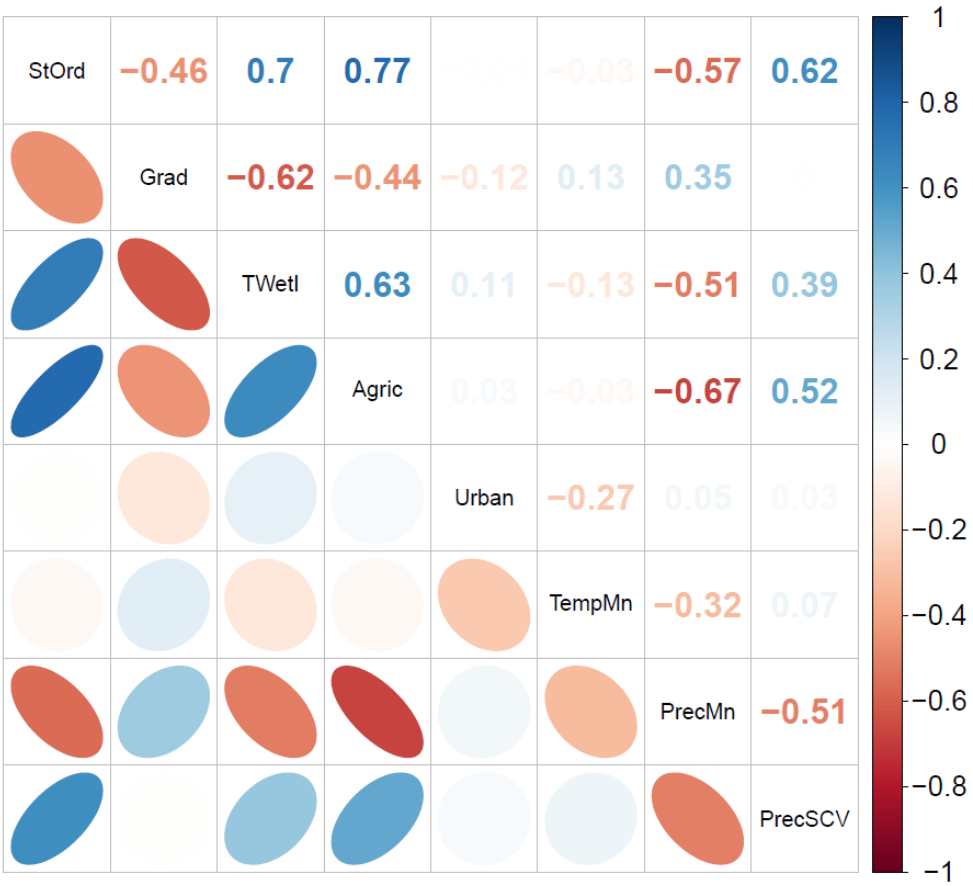


Figure 3. Correlations among the environmental predictor variables considered in the MARS modeling. Upper diagonal depicts Pearson correlation coefficients and lower diagonal depicts ellipses scaled to correlation values. Variable abbreviations are presented in Table 3.

Species Distribution Modeling

A review and comparison of statistical approaches is provided in Appendix B (also see Olden et al. 2008). For our purposes we deployed multivariate adaptive regression splines (MARS: Friedman 1991) to model species presence-absence in each catchment as a function of the environmental variables described above. Most survey events do not give data on sample area or effort; therefore we modeled species presence/absence and not species abundance. MARS is a procedure for fitting adaptive non-linear regression that uses piece-wise linear basis functions to define relationships between a response variable and some set of predictors (Friedman 1991). The resulting regression surface is piecewise linear and continuous. Breaks between segments are defined by a knot in a model that initially over-fits the data, and then simplified using a backwards/forwards stepwise cross-validation procedure to identify terms to be retained in the final model.

MARS allows simultaneous analysis of multiple species by identifying a common set of basis functions, with knots (and therefore variables) selected according to

their improvement in predictive power of the model, averaged across all species. Therefore, this method was selected based on its ability to incorporate data from multiple species to inform model development for data-poor species (i.e., a multi-response model: Elith et al. 2006, Leathwick et al. 2006, Elith and Leathwick 2007). One requirement of the MARS multi-response model is knowledge of true absences of species in sampled locations. To accommodate this requirement, we only included catchments where at least one species was recorded (n = 323 catchments). All MARS models were fitted in R (R Development Core Team 2010) and utilized functions in the earth library (Milborrow and Maintainer 2007). The predictive performance of models was evaluated using area under the Receiver Operating Characteristic curve (AUC: Fielding and Bell 1997), which was calculated using 3-fold cross-validation. We considered an AUC ≥ 0.75 as indicative of a useful model (Pearce and Ferrier 2000). Probabilities of species occurrence were converted to presence or absence (decision threshold = 0.5).

MARS models include a backwards elimination feature selection routine that looks at reductions in the generalized cross-validation (GCV) estimate of error. Overall variable importance was based on changes in GCV when each predictor's feature was added to the model. If a predictor was never used in any MARS basis function, it has an importance value of zero. Next, we conducted a model sensitivity analysis where the predicted probability of occurrence for each species was plotted as a function of varying each predictor variable individually while holding all other predictors at their observed median value.

Correct and incorrect classification of species presence/absence from the MARS models were tabulated according to a confusion matrix (model prediction, actual occurrence): predicted, observed (true presence); not predicted, not observed (true absence); predicted, not observed (false presence); not predicted, observed (false absence). Each element of confusion matrix has a different ecological interpretation (Figure 4). For example, catchments in which species were predicted but not observed (false absent) are considered areas that are environmentally suitable (according to the environmental variables and MARS modeling approach used here) yet may be absent due to other limited factors related to natural and human-related factors. Lastly, the environmental predictors of all 4,244 catchments for the entire Verde River Basin where inputted into the MARS models to predict fish species occurrence. Catchments in which species were predicted based on this analysis are considered areas that are environmentally suitable and may represent current populations (in the case where the catchment has not be surveyed) or may represent catchments that are truly absent of that species. All five catchment classification described above and illustrated in Figure 4 were displayed spatially using a Geographic Information System.

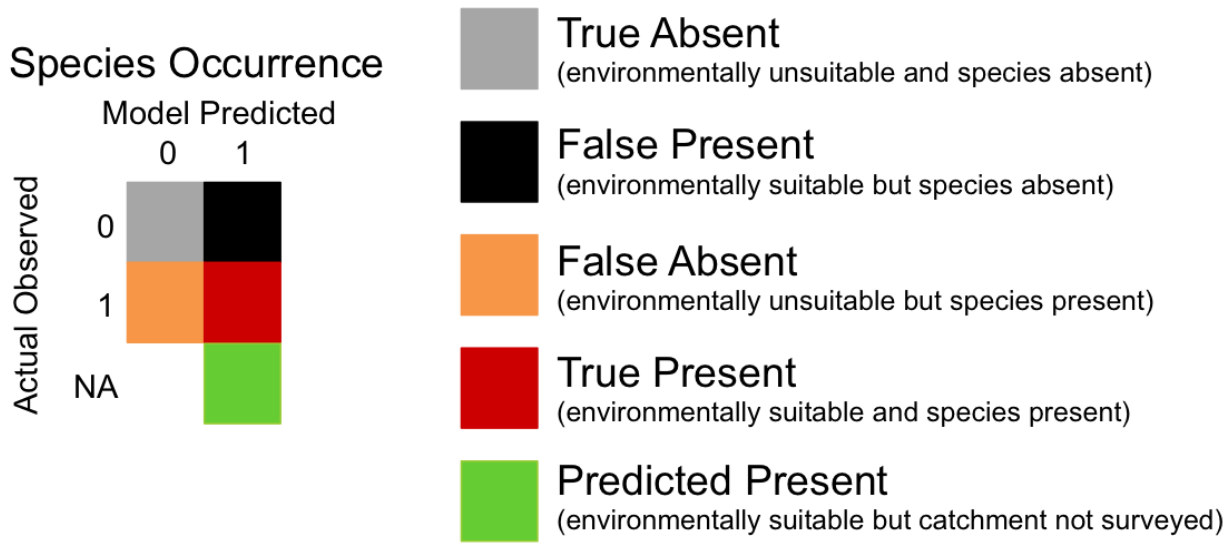


Figure 4. Confusion matrix detailing the interpretation of model predictions and actual species occurrence.

Site selection for ground truthing

To assess the model's accuracy we added a ground truthing component to this project. Ground truthing sites were identified as unsurveyed reachcodes in the Verde River watershed. The unsurveyed reachcodes were prioritized based on the information needs of our regional fish program managers. Initial filtering of site selection was done via satellite imagery to determine if priority reaches had water. After the initial filter, sites were then visited and surveyed, if water was present. Even after this initial filter, a number of the reaches that lacked fish data and made it through the initial filter were dry or intermittent. Surveys were conducted between April and July 2014 using standard fish sampling protocols (AGFD 2004). Multiple gear types were used including backpack electro-fishing, dip netting and hoop netting. Fish were identified to species, measured and returned to the stream.

Results

Data compilation (Phase 1)

Data compilation derived from the following sources: Nongame Wildlife Branch Native Fish database, Aquatic Research Program, datasets from Pinetop, Flagstaff, Kingman and Mesa regional fish programs, Gila River Basin Conservation (CAP) datasets. Additionally, AZGFD entered data from Scientific Collecting Permits spanning 1991-2013.

Verde Watershed

The Verde River watershed encompasses 4.23 million acres for which we identified 183 water bodies that include rivers, streams, creeks, springs, ponds, tanks, and lakes. For sake of analysis water bodies were divided into impoundments and streams. The streams ranged in length from 0.2 km to 15.01 km and the impoundments ranged from 0.2 acres to 2,033 acres at full pool. Within the entire watershed 1,319 reachcodes were identified and 406 (31%) were identified as

having been surveyed. From 1954 to 2013, 2,466 surveys have been conducted in this watershed (Table 5). The majority of the surveys (74%) were conducted in reaches within two streams; the Verde River and Fossil Creek (Figure 5). The most heavily surveyed reach has been surveyed 70 times dating back to 1954 in the Verde River near Perkinsville, AZ.

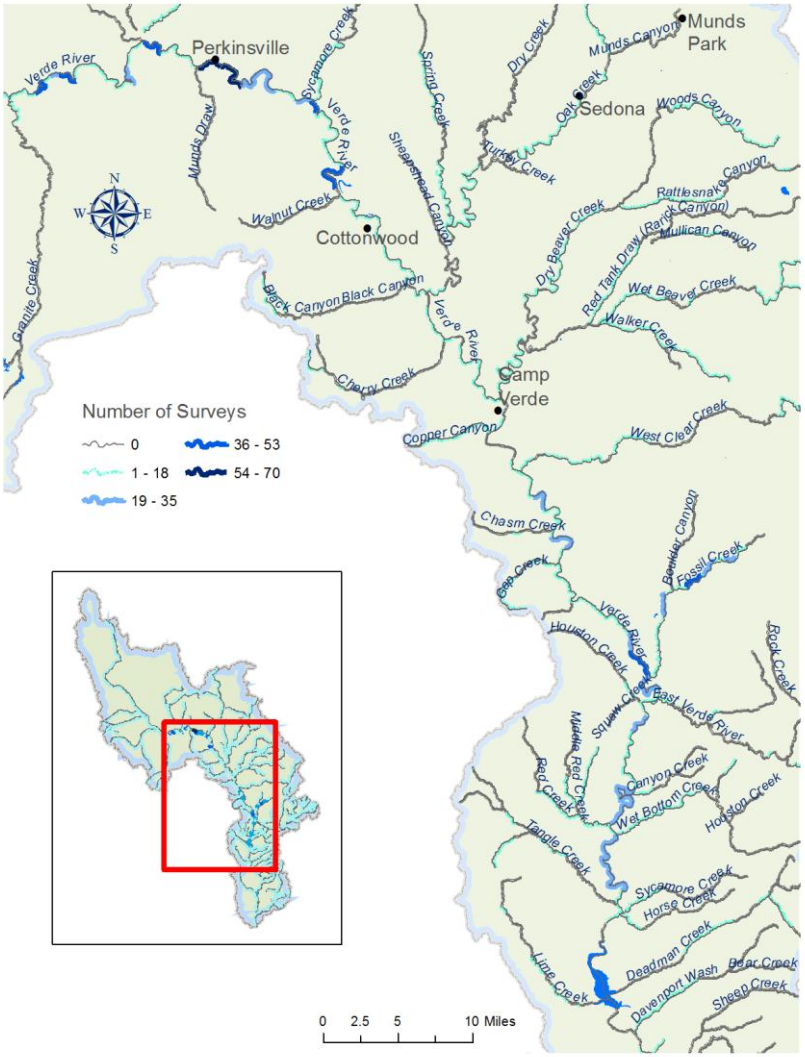


Figure 5. A map of the Verde River watershed that displays the number of fish surveys conductied starting in 1954.

A total of 113,230 individual records that included 57 fish (either identified to species or families) were identified for this dataset. Many records had counts greater than one for a total of 494,050 fish that have been sampled from 1954-2013 (Table 5). Red shiner, Desert sucker and Sonora sucker were the species with the most records (Table 6).

Table 5. Summary of watershed information related to bodies of water and fish survey information.

	Verde	LCR
Watershed area (Acres)	4,229,553	13,892,293
Number of water bodies	183	186
Number of reaches	1,319	1,042
Number of reaches with surveys	406	293
Range of stream reach length	0.02-15.01	0.02-24.41
Range of impoundment size	0.02-2,033	0.13-3,099
Number of surveys in watershed	2,466	993
Total fish records	113,230	43,828
Total fish count	494,050	246,114
Total species	57	48
Range of survey years	1954-2013	1958-2013

Table 6. Total number of fish counted by species in the Verde watershed from 1954-2013.

Common Name	Total Fish count	Common Name	Total Fish count
Red shiner	78,918	Green/Bluegill Sunfish hybrid	1,229
Desert sucker	38,999	Headwater chub	1,099
Sonora sucker	37,432	Brown trout	890
Western mosquitofish	36,808	Unknown	712
Roundtail chub	29,027	Goldfish	578
Longfin dace	28,131	Sonoran sucker/Desert sucker hybrid	412
Speckled dace	20,598	Brown bullhead	353
Green sunfish	18,665	Little Colorado sucker	249
Smallmouth bass	18,403	Rock bass	232
Desert pupfish	14,803	Sunfish hybrid	189
Threadfin shad	14,470	Black bullhead	127
Common carp	14,281	Redear sunfish	96
Largemouth bass	12,896	Centrarchid family	61
Spikedace	11,065	Bluehead sucker	44
Catostomid family	8,381	Yellow perch	34
Razorback sucker	7,331	Apache trout	33
Bluegill	6,687	Sonoyta pupfish	30
Northern crayfish	6,326	Walleye	28
Loach minnow	4,127	Warmouth	25
Gila topminnow	3,399	Brook trout	20
Colorado pikeminnow	3,353	Grass carp	16
Yellow bullhead	3,168	Sonora Mud Turtle	14

Black crappie	3,082	Cutthroat/rainbow hybrid	11
Crayfish spp.	2,927	Gila trout	9
Fathead minnow	2,780	Smallmouth buffalo	8
Flathead catfish	2,460	Spotted bass	7
Channel catfish	2,332	White crappie	4
Rainbow trout	2,020	Israeli carp	2
Gila chub	1,645	Cyprinid family	1
Golden shiner	1,324	Flannelmouth sucker	1

Little Colorado River Watershed

The LCR watershed encompasses 13.89 million acres for which we identified 186 water bodies that include rivers, streams, creeks, springs, ponds, tanks, and lakes. The streams ranged in length from 0.2 km to 24.41 km and the impoundments ranged from 0.13 acres to 3,099 acres at full pool. Within the entire watershed 1,042 reaches were identified and 293 (28%) were identified as having been surveyed. From 1958 to 2013, 993 surveys have been conducted in this watershed (Table 5). The three most surveyed streams were Little Colorado River, East Clear Creek, and Nutrioso Creek. They made up 33% of all reached surveyed in the LCR watershed. The most heavily surveyed water is Becker Lake which has been surveyed 18 times since 2005.

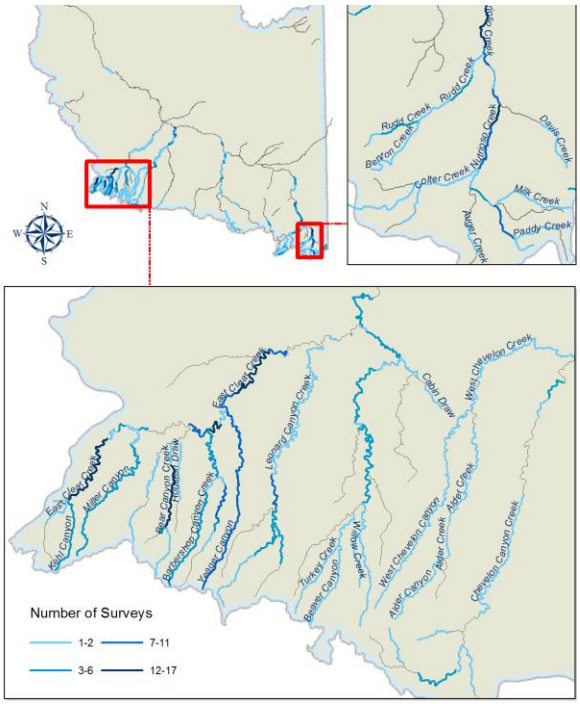


Figure 6. A map of the Little Colorado River watershed that displays the umber of fish surveys conductied starting in 1958;

A total of 48 fish (either identified to species or families) were identified making up 43,828 individual records for this dataset. Many records had counts greater than one making a total of 246,114 fish that have been sampled from 1958-2013 (Table 5). Speckle dace, fathead minnow, rainbow trout, and bluehead sucker were the most frequently survey fish (Table 7)

Table 7. Total number of fish counted by species in the LCR watershed from 1958-2013.

Common Name	Total Fish count	Common Name	Total Fish count
Speckled dace	108,128	Black crappie	291
Fathead minnow	64,359	Western mosquitofish	270
Humpback chub	16,611	Mosquitofish spp.	217
Bluehead sucker	14,541	Longfin dace	215
Rainbow trout	7,321	Roundtail chub	207
Little Colorado spinedace	6,259	Sucker spp.	197
Green sunfish	3,738	Desert sucker	166
Red shiner	3,355	Catostomid family	158
Little Colorado sucker	2,338	Yellow bullhead	130
Brown trout	1,850	Sonoran sucker/Desert sucker hybrid	101
Flannelmouth sucker	1,713	Arctic grayling	95
Golden shiner	1,636	Yellow perch	70
Northern pike	1,568	Goldfish	36
Unknown	1,524	Sonora sucker	19
Brook trout	1,449	Sunfish	16
Smallmouth bass	1,142	Shiner spp	8
Black bullhead	1,134	Bullhead spp.	4
Largemouth bass	967	Cutthroat trout	4
Common carp	936	Gila chub	1
Channel catfish	758	Gila trout	1
Walleye	739	Oncorhynchus spp.	1
Plains killifish	603	Pumpkinseed	1
Bluegill	490	Rock bass	1
Apache trout	400	Sonoyta pupfish	1
Crayfish spp.	344	Sunfish hybrid	1

Geodata layer development

We developed three types of geodata layers; fish collection data layer, reach data layer, and individual species data layers.

The fish data layer is the most comprehensive of the three. This dataset has 10 attributes that contain information about individual fish records including species captured, gear type used, reach code, and date collected (Table 8).

The reach data layer contains information about each reachcode in the two watersheds (1,319 in the Verde watershed

and 1,042 in the LCR watershed). This dataset contains seven attributes including the number of surveys conducted in each reach, series of years each reach was surveyed, and species observed in each reach (Table 8).

The individual species data layers are a series of data layers for each prevalent species in the watersheds. There are 29 species layers in the Verde watershed and 16 in the LCR watershed. Each layer contains eight attributes for each reachcode including the number of surveys conducted in each reach, the number of surveys that species was positively detected, and the proportion of positive detections (Table 8).

Table 8. List of attributes and description for each geodata layer.

Individual Species Data Layer	
Name	**Description**
ReachCode	Unique Identifier
Watername	Name of waterbody
Number of Surveys with Positive Detection	Number of surveys where species was positively identified
Number of Surveys in Reach	Number of times reach has been uniquely surveyed. Surveys done less than one month apart are not considered unique.
Proportion of positive detection	Number of surveys with positive detection divided by number of times reach has a unique survey.
Species code	Species code typically is the first two letters of genus and first two letters of species (see species code table).
Common Name	Common name associated with each species
Watershed	HUC6 watershed
Reach Data Layer	
Watershed	HUC6 watershed
Waterbody	Name of waterbody
ReachCode	National hydrography dataset reach code ID
ReachCode Length (Km) or Acres	streams are measured by stream length, impoundments are perimeter length)

Years Surveyed	Series of years surveyed
TotalSurveysInReach	Number of times reach has been uniquely surveyed. Surveys done less than one month apart are not considered unique.
Species Observed	Series of common names of fish species found in that reach
Fish Collection Data Layer	
DataSource	Location or dataset from which data came from:
ActivityDate	Date that survey took place
ActivityYear	Year that survey took place
Watershed	HUC6 watershed
WaterBody	Name of water
SppCode	Species code typically is the first two letters of genus and first two letters of species (see species code table)
CommonName	Common name associated with each species
FishCount	Count of species
GearCode	Indicates the type of gear used to collect each species
ReachCode	National hydrography dataset reachcode ID

Model results

We found that fish species occurrence in the Verde River Basin was highly predictive according to channel- and catchment-scale descriptors of topography, watershed hydrology, land use, and regional climate. The MARS models demonstrated robust predictive performance across species as indicated by AUC values ranging from 0.54 to 0.95 (mean = 0.79) and internal R^2 values ranging from 0.02 to 0.48 (mean = 0.18) (Table 9). Native species that were highly predictable across the basin included longfin dace, speckled dace and Gila chub; similarly nonnative common carp, smallmouth bass and rainbow trout demonstrated strong predictions based on the MARS models. As expected, we found a

strong relationship (R=0.68, P<0.05) between model performance (R^2) and species prevalence (proportion of catchments present); suggesting that modeling rare species is challenged by limited data needed to develop robust occurrence-environment associations. See Olden et al. (2002) for discussion of this common modeling artifact.

The spatial distributions of species model predictions are presented in Appendices C and D. These maps display those catchments where the model correctly predicted presence and absence (red and grey) from those catchments where occurrence was incorrectly predicted (orange and black). Also indicated are catchments across the Verde River basin where environmental conditions are predicted to the suitable for species occurrence but have not been surveyed (according to FINS) or where streams may be intermittent or ephemeral in their hydrology (green).

Table 9. MARS model results, including species prevalence (Prev) and model performance according to R-sq (internal validation based on training data), R-sq (external validation based on cross-validation) and area under the Receiver Operating Characteristic curve (AUC) based on cross-validation). NA values indicate species that were not able to be modeled due to low prevalence.

Common name	Code	Prev	R-sq (int)	R-sq (CV)	AUC (CV)
NATIVE SPECIES					
longfin dace	AGCH	0.23	0.47	0.41	0.89
desert sucker	CACL	0.37	0.14	0.09	0.69
Sonora sucker	CAIN	0.36	0.22	0.18	0.73
Gila chub	GIIN	0.03	0.03	0.01	0.77
headwater chub	GINI	0.02	0.07	0.00	0.54
roundtail chub	GIRO	0.17	0.23	0.12	0.72
Gila topminnow	POOC	0.02	0.04	0.02	0.76
speckled dace	RHOS	0.17	0.31	0.22	0.87
razorback sucker	XYTE	0.02	0.05	0.01	0.73
NONNATIVE SPECIES					
black bullhead	AMME	0.02	0.02	NA	NA
yellow bullhead	AMNA	0.16	0.25	0.19	0.82
rock bass	AMRU	0.04	0.04	0.02	0.79
white sucker	CACO	0.02	0.05	NA	NA
common carp	CYCA	0.23	0.39	0.31	0.92
red shiner	CYLU	0.32	0.26	0.23	0.82
western mosquitofish	GAAF	0.08	0.15	0.04	0.76
channel catfish	ICPU	0.11	0.22	0.18	0.88
green sunfish	LECY	0.37	0.16	0.08	0.69
bluegill	LEMA	0.06	0.07	0.04	0.76

smallmouth bass	MIDO	0.42	0.48	0.45	0.89
largemouth bass	MISA	0.17	0.30	0.14	0.85
rainbow trout	ONMY	0.10	0.23	0.20	0.86
fathead minnow	PIPR	0.08	0.08	0.04	0.72
flathead catfish	PYOL	0.16	0.26	0.25	0.87
brown trout	SATR	0.02	0.08	0.06	0.95

Across all species, Shreve link value was the most important predictor of species occurrence, followed by mean annual temperature, total precipitation and coefficient of variance in monthly spring precipitation for the contributing upstream watershed (Table 10). All remaining variables played no significant influence in the MARS model. Variable importance varied across species indicating that environmental variables are differentially important for influencing species occurrence. Probability of occurrence for speckled dace, desert sucker, Gila chub, roundtail chub, decreased with warming temperatures, whereas the opposite relationship was true for longfin dace, Sonora sucker and Gila topminnow (Appendix C). All native species with the exception of Sonora sucker showed either lower or no change in probability of occurrences with increasing coefficient of variance in monthly spring precipitation. Nonnative species showed many different associations with the environmental factors. For example, most nonnative species exhibited higher probabilities of occurrence in those catchments characterized by higher coefficient of variance in monthly spring precipitation (Appendix D); the only exception to this pattern was for rainbow trout that demonstrated a negative association between occurrence and coefficient of variance in monthly spring precipitation. Most nonnative species exhibited a negative relationship between probability of occurrence and mean annual temperature, suggesting that high temperature streams of Verde river may not favor nonnative species persistence. Unsurprisingly, exceptions to this relationship were the red shiner, western mosquitofish and fathead minnow (Appendix D).

Table 10. MARS variable importance. Variable descriptions provided in Table 3. Variable importance using the "number of subsets" criterion. Is the number of subsets that include the variable. Generalized Cross Validation (GCV) of the model (summed over all responses) The Residual sum-of-squares (RSS) of the model (summed over all responses if y has multiple columns).

Variable	# of subsets	GCV	RSS
Shreve (stream order)	7	100.0	100.0
Mean Temperature	6	56.2	67.3
Mean Precipitation	3	16.2	35.9
Spring Precipitation CV	2	3.8	27.2
Gradient	0	0.0	0.0
Topographic Wetness I.	0	0.0	0.0
Agriculture	0	0.0	0.0
Urban	0	0.0	0.0

Ground truthing results

We surveyed or visited 58 reaches in 17 streams between April and July 2014 that had no prior records and had been identified as priority by fisheries managers (Table 11). Many of the unsurveyed sites in the Verde River watershed were either dry or intermittent. Using electro-fishing,

seining, dip nets, and/or hoop nets, we captured a total of 1,768 individuals that represented 14 different species. Of the 17 streams we visited, we only captured fish in six of them. All fish surveys were entered into the FINS database, but are not included in the dataset used for model development. Additionally, draft trip reports were completed for surveys conducted in the Verde River drainage, Oak Creek, and Rock Creek (Appendix E)

Table 11. Stream reaches that have not been surveyed and were visited and either surveyed or determined to be dry. Number of species is the number of unique species found in that stream and the water field for each stream was one of three categories; intermittent, dry, or water present.

Stream Name	Number of reaches	Number of species	Water
Alder Creek	2	1	Intermittent
Black Canyon	1	No fish	Dry
Chase Creek	3	No fish	Dry
Deadman Creek	1	No fish	Water present
Dry Beaver creek	2	No fish	Dry
E.F. Sycamore	1	No fish	Intermittent
East Verde	8	No fish	Water present
Ellison Creek	10	No fish	Dry
Lime Creek	1	No fish	Dry
North Fork Deadman Creek	1	No fish	Water present
Oak Creek	6	13	Water present
Pine Creek	1	No fish	Intermittent
Rock Creek	1	7	Water present
Sycamore Creek	7	3	Water present
Walnut Creek	1	1	Water present
Webber Creek	11	No Fish	Intermittent
West Clear Creek	1	5	Water present

To validate the MARS model, we compared species collected during field surveys (in areas previously unsurveyed) to predicted presence/absence. For most species, predictions were correct in greater than 90% of the 58 sites that were surveyed as part of this exercise (Table 12). The clear exception was rainbow trout (ONMY) where it was correctly predicted in 24% of the sites. At many of the sites the model predicted rainbow trout presence when the species was not observed during surveys. These streams where this occurred were East Verde River, Webber Creek, Ellison Creek, and Chase Creek. Many of these sites were dry or appeared to be intermittent.

Table 12. Percentage of ground truthing sites that correctly predicted species presence. Based on 58 sites.

Species	Percent Predicted	Species	Percent Predicted
AGCH	95%	ICPU	100%
AMME	100%	LECY	86%
AMNA	92%	LEMA	100%
AMRU	100%	MIDO	97%
CACL	95%	MISA	97%
CACO	100%	ONMY	24%
CAIN	95%	PIPR	100%
CYCA	100%	POOC	100%
CYLU	97%	PYOL	100%
GAAF	100%	RHOS	89%
GIIN	100%	SATR	92%
GINI	95%	XYTE	97%
GIRO	97%		

Conclusion/Discussion

Biologists today are collecting GPS points in the field to identify survey locations. Much of the historical location data was identified through topographic maps or general descriptions were used to identify a sample location. What sets this effort apart from previous data compilation efforts is both the magnitude of data compiled from each watershed and assignment of a reachcode to each fish record. The benefit of assigning reachcodes allows researchers to integrate the physical and environmental geospatial data provided by the National Hydrography Dataset.

One of the greatest challenges with data management is fitting it all together. Nearly 1,000 investigators collected fisheries data over the last 60 years in the Verde River and Little Colorado River watersheds. Investigators are employed with various state and federal agencies, universities, and non-governmental consulting firms and are generally standardized at local levels (Bonar and Hubert 2002). The types of data collected and reported vary widely. Comparing data across years, surveys, and investigators is difficult to do unless data standards exist and are accurately reported. Gear type and effort, for example, can bias the fish species which are caught leaving an under representative picture of the fish community (Hubert 1996). While the methods of fisheries standardization have been developed (Bonar et al 2009), fisheries data standards are still lacking. The likely more challenging issue is assuring that fisheries standardized sampling and data standards are implemented. The Arizona Game and Fish has developed fisheries standardized sampling methods that fall in line with the national standard developed by Bonar et al (2009) and have recently began implementing standardized data collection and reporting in the form of FINS. The funding associated with this agreement

has allowed AZGFD to complete the arduous task of standardizing historical fisheries records and compile them into a centralized location.

The intent of the project was to make data available for the science community to better help conserve our fisheries resources. Applications are numerous. The data we provide serves as a basis for understanding species interactions and environmental suitability by species. We provided one example as demonstrated by the Verde River watershed model. The use of NHDplus attributes allowed us to determine which physical and environmental factors drive species presence. There are a number of threats that exist in Arizona, but water availability is the most critical factor impacting the ability of wildlife and land management agencies to provide anglers sport fishing opportunities while conserving native fishes. Few, if any, ponds, lakes, streams and springs are managed strictly for fish or wildlife benefits. Most are managed to supply water for various human and livestock needs, with fish and wildlife needs a lesser priority. It was evident from our ground truthing attempts that water is scarce and the options for suitable fish habitat are at risk. All economic activity, including mining, agriculture, and urban growth, relies on a dependable water supply. Fish and wildlife must compete with these resources for water supply. Having a robust geospatial fisheries database will help managers make data driven decisions and indentify important areas for water conservation at a watershed and statewide scale. Efforts have been underway to identify and prioritize fisheries conservation areas in the southwest (Paukert et al. 2011; Stecker et al. 2011) and recently AZGFD has begun produced watershed base management plans (AGFD 2009).

Utility of data compilation efforts

The utility of this database is especially evident in a recent Verde River watershed based fisheries management planning effort. The planning effort used data driven management decisions to balance AZGFD's dual mandates to provide sport fish opportunities and to conserve native fish (AGFD 2015 in prep). A Verde River Watershed Fisheries Management Plan web map was developed as the primary communication tool for Department fisheries and aquatic management decisions in the Verde River watershed. The web map displays each individual management unit on an interactive map, allowing users to easily identify high priority management units and the primary management emphasis for that unit. The web tool also provides photos of waters within the management units and information about primary and secondary managed species. Even more information about the management units is provided in Adobe PDF documents, that can be downloaded and printed. The web map is available online at: http://phx-gis-dev-ags/verdewatershed/. With clean digitally referenced historic fish records now entered into FINS, future watershed based fish management efforts will have easy access to historical records.

Where to Get the Data

Datasets that resulted from this project will be posted at the following locations:

- Desert LCC Conservation Planning Atlas http://dlcc.databasin.org/
- Southern Rockies Conservation Planning Atlas http://srlcc.databasin.org/
- MARIS (Multistate Aquatic Resources Information System) http://www.marisdata.org/
- AZGEO (Arizona's GIS Clearinghouse) https://azgeo.az.gov/AZGEO/

References

Abell, R., J. D. Allan, and B. Lehner. 2007. Unlocking the potential of protected areas for fresh waters. Biological Conservation 134:48–63.

AGFD (Arizona Game and Fish Department) 2004. Standard fish sampling protocol for state of Arizona Waters. Phoenix, AZ.

AGFD (Arizona Game and Fish Department) 2009. Statewide Fisheries Management Team. Phoenix, AZ.

AGFD (Arizona Game and Fish Department) 2012. Arizona's State Wildlife Action Plan: 2012-2022. Arizona Game and Fish Department, Phoenix, Arizona.

AGFD (Arizona Game and Fish Department) 2015. Verde River Watershed Fish Management Plan. Phoenix, AZ.

Bonar, Scott A., and Wayne A. Hubert. "Standard sampling of inland fish: benefits, challenges, and a call for action." Fisheries 27.3 (2002): 10-16.

Bonar, Scott A., Wayne A. Hubert, and David W. Willis, eds. Standard methods for sampling North American freshwater fishes. Bethesda, Maryland: American Fisheries Society, 2009.

Friedman, J. H. 1991. Multivariate adaptive regression splines. Annals of Statistics 19:1–67.

Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith, J. and Leathwick, J. 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. Diversity and Distributions 13: 265–275

Elith, J. and Leathwick, J. 2009. Species distribution models across space and time. Annu. Rev. Ecol. Evol. Syst. 40:677-97.

Fagan, W.F., Unmack, P.J., Burgess, C., and Minckley, W.L. 2002. Rarity, fragmentation, andextinction risk in desert fishes. Ecology, 83(12): 3250–3256.

Fielding, Alan H., and John F. Bell. "A review of methods for the assessment of prediction errors in conservation presence/absence models." Environmental conservation 24.01 (1997): 38-49.

Friedman, Jerome H. "Multivariate adaptive regression splines." The annals of statistics (1991): 1-67.

Gibson, Polly P., Julian D. Olden, and Matthew W. O'Neill. "Beaver dams shift desert fish assemblages toward dominance by non-native species (Verde River, Arizona, USA)." Ecology of Freshwater Fish (2014).

Hermoso, V., S. Linke, J. Prenda, and H. P. Possingham. 2011. Addressing

longitudinal connectivity in the systematic conservation planning of fresh waters. Freshwater Biology 56:57–70.

Hubert, W. A. 1996. Passive capture techniques. Pages 157-192 in B. R. Murphy and D. W. Willis, eds. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

Lawrence, D. J., Larson, E. R., Liermann, C. A. R., Mims, M. C., Pool, T. K. and Olden, J. D. 2011.National parks as protected areas for U.S. freshwater fish diversity. Conservation Letters,4: 364–371.

Leathwick, J.R., 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. Freshwater Biology 50, 2034–2052.

Leathwick, J.R., Elith, J. & Hastie, T. 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecological Modelling 199, 188–196.

Menard, S., 1995. Applied logistic regression analysis. Sage University Paper Series 07–106. Sage Publications, Thousand Oaks, CA

Milborrow, Stephen, and Maintainer Stephen Milborrow. "The earth Package." The R Project for Statistical Computing (2007).

Minckley, W. L., and Paul C. Marsh. Inland fishes of the greater Southwest: chronicle of a vanishing biota. University of Arizona Press, 2009.

Multi-Resolution Land Characteristics Consortium. "National Land Cover Database." Sioux Falls, SD: US Geological Survey (2001).

O'brien, Robert M. "A caution regarding rules of thumb for variance inflation factors." Quality & Quantity 41.5 (2007): 673-690.

Olden, Julian D., Donald A. Jackson, and Pedro R. Peres-Neto. "Predictive models of fish species distributions: a note on proper validation and chance predictions." Transactions of the American Fisheries Society 131.2 (2002): 329-336.

Olden, J.D. & Poff, N.L. 2005. Long-term trends of native and non-native fish faunas in the American Southwest. Animal Biodiversity and Conservation 28: 75-89.

Olden, Julian D., Joshua J. Lawler, and N. LeRoy Poff. "Machine learning methods without tears: a primer for ecologists." The Quarterly review of biology 83.2 (2008): 171-193.

Paukert, C. P., K. L. Pitts, J. B. Whittier, and J. D. Olden. 2011.Development and assessment of a landscape-scale ecological threat index for the Lower Colorado River Basin. Ecological Indicators 11:304–310

Pearce, Jennie, and Simon Ferrier. "Evaluating the predictive performance of habitat models developed using logistic regression." Ecological modelling 133.3 (2000): 225-245.

Pool, T. K., Olden J. D., Whittier J. B., and Paukert C. P. 2010. Environmental drivers of fishfunctional diversity and composition in the Lower Colorado River Basin. Can. J. Fish. Aquat. Sci. 67: 1791–1807.

Pool, T.K., Strecker, A.L. & Olden, J.D. 2013. Identifying preservation and restoration priority areas for desert fishes in an increasingly invaded world. Environmental Management 3:631-641.

Rinne, J.N. 2012. Fish and aquatic organisms. In: Neary, D.G., Medina, A.L. & Rinne, J.N., eds. Synthesis of Upper Verde River Research and Monitoring 1993-2008. Fort Collins: USDA Forest Service, Gen. Tech. Rep. RMRS-GTR-291, pp 189-232.

Shreve, Ronald L. "Statistical law of stream numbers." The Journal of Geology (1966): 17-37.

Strecker, A. L., J. D. Olden, J.B. Whittier, C.P. Paukert. 2011. Defining conservation priorities forfreshwater fishes according to taxonomic, functional, and phylogenetic diversity.Ecological Applications, 21(8), pp. 3002–3013.

Theobald, D.M. 2007. LCaP v1.0: Landscape Connectivity and Pattern Tools for ArcGIS. Colorado State University, Fort Collins.

Whittier, J.B., C.P. Paukert, J.D. Olden, K.L. Pitts, and A.L. Strecker. 2011 Lower Colorado River Basin aquatic gap analysis project final report. U.S. Geological Survey, Gap Analysis Program, Reston Virginia, US

APPENDICES

Appendix A. List of species code, common name and scientific name of aquatic species found in Arizona.

Species Code	Common Name	Scientific Name
AGCH	Longfin dace	Agosia chrysogaster
AMME	Black bullhead	Ameiurus melas
AMNA	Yellow bullhead	Ameiurus natalis
AMNE	Brown bullhead	Ameiurus nebulosus
AMRU	Rock bass	Ambloplites ruprestris
AMSPP	Bullhead spp.	Ameiurus spp.
ANWO	Woodhouse's Toad	Anaxyrus woodhousii
BUPU	Red-spotted Toad	Bufo punctatus
CAAU	Goldfish	Carassius auratus
CACL	Desert sucker	Catostomus clarki
CADI	Bluehead sucker	Catostomus discobolus
CADIYA	Zuni Bluehead Sucker	Catostomus discobolus yarrow
CAFM	Catostomid family	Catostomidae
CAIN	Sonora sucker	Catostomus insignis
CAINxCACL	Sonoran sucker/Desert sucker hybrid	Catostomus insignis/Catostomus clarki hybrid
CALA	Flannelmouth sucker	Catostomus latipinnis
CASP	Little Colorado sucker	Catostomus spp.
CASPP	Unidentified sucker	Catostomidae spp.
CEFM	Centrarchid family	Centrarchidae
CRAY	Crayfish spp.	Unknown crayfish species

CTID	Grass carp	Ctenopharyngodon idella
CYCA	Common Carp	Cyprinus carpio
CYCACA	Israeli carp	Cyprinus carpio carpio
CYER	Sonoyta/Quitobaquito pupfish	Cyprinidon eremus
CYFM	Cyprinid family	Cyprinidae
CYLU	Red shiner	Cyprinella lutrensis
CYMA	Desert pupfish	Cyprinidon macularius
CYSPP	Unknown minnow	Cyprinidae spp.
DOPE	Threadfin shad	Dorosoma petenense
ESLU	Northern pike	Esox Lucius
FUZE	Plains killifish	Fundulus zebrinus
GAAF	Western mosquitofish	Gambusia affinis
GASPP	Mosquitofish spp.	Gambusia spp.
GICY	Humpback chub	Gila cypha
GIIN	Gila chub	Gila intermedia
GINI	Headwater chub	Gila nigra
GIRO	Roundtail chub	Gila robusta
GISPP	Chub spp.	Gila spp.
ICBU	Smallmouth buffalo	Ictiobus bubalus
ICPU	Channel catfish	Ictalurus punctatus
KISO	Sonora Mud Turtle	Kinosternon sonoriense
LECY	Green sunfish	Lepomis cyanellus
LECYxLEMA	Green/Bluegill Sunfish hybrid	Lepomis cyanellus/Lepomis microlophus hybrid

LEGI	Pumpkinseed	Lepomis gibbosus
LEGU	Warmouth	Chaenobryttus gulosus
LEHY	Sunfish hybrid	Lepomis sp. x Lepomis sp.
LEMA	Bluegill	Lepomis macrochirus
LEMI	Redear sunfish	Lepomis microlophus
LESPP	Sunfish	Lepomis spp.
LEVI	Little Colorado spinedace	Lepidomeda vittata
LICAAD	Bullfrog-Adult	Lithobates catesbeiana
LICATA	Bullfrog-Tadpole	Lithobates catesbeiana
LIFMTA	Ranid frog tadpole	Ranidae tadpole
LISPP	Leopard frog	Lithobates spp.
LIYA	Lowland leopard frog	Lithobates yavapaiensis
MEFU	Spikedace	Meda fulgida
MIDO	Smallmouth bass	Micropterus dolomieu
MIPU	Spotted bass	Micropterus punctulatus
MISA	Largemouth bass	Micropterus salmoides
MISPP	Bass spp.	Micropterus spp
NF	No fish	No Fish
NOCR	Golden shiner	Notemigonus crysoleucas
NOSPP	Shiner spp.	Notemigonus spp.
ONAP	Apache trout	Oncorhynchus apache
ONCL	Cutthroat trout	Oncorhynchus clarki
ONCLxONMY	Cutthroat/rainbow hybrid	Oncorhynchus clarki/Oncorhynchus mykiss hybrid
ONGI	Gila trout	Oncorhynchus gilae

ONMY	Rainbow trout	Oncorhynchus mykiss
ONSPP	Oncorhynchus spp.	Oncorhynchus spp.
ORVI	Northern crayfish	Orconectes virilis
PEFL	Yellow perch	Perca flavescens
PIPR	Fathead minnow	Pimephales promelas
POAN	White crappie	Pomoxis annularis
PONI	Black crappie	Pomoxis nigromaculatus
POOC	Sonora topminnow	Poeciliopsis occidentalis
POOCOC	Gila topminnow	Poeciliopsis occidentalis occidentalis
PTLU	Colorado pikeminnow	Ptychocheilus Lucius
PYOL	Flathead catfish	Pylodictis olivaris
RHOS	Speckled dace	Rhinichthys osculus
SAFO	Brook trout	Salvelinus fontinalis
SATR	Brown trout	Salmo trutta
SAVI	Walleye	Sander vitreum
THAR	Arctic grayling	Thymallus arcticus
TICO	Loach minnow	Tiaroga cobitis
TISPP	Tilapia spp.	Tilapia spp.
UF	Unidentified fish	Actinopterygii (bony fishes)
UNK	Unknown	Unknown
XYTE	Razorback sucker	Xyrauchen texanus

Appendix B. Summary of major approaches to species distribution models

Method	Class of Model	Data Requirements
ANN	artificial neural network	pa, community
BIOCLIM	envelope model	P
DOMAIN	multivariate distance	P
GARP	rule sets from genetic algorithms	pa
GAM	generalized additive model	pa
GDM	generalized dissimilarity model	pa, community
GLM	generalized linear model	pa
MARS	multivariate adaptive regression splines	pa, community
MAXENT	maximum entropy	pa
MRT	multivariate regression trees (random forest, boosted)	pa

p = only presence data used; pa =presence and some form of absence required (e.g. a background sample); community = community data contribute to model fitting.

Optimal properties of a modeling approach

1. the ability to model non-linear associations between predictor and response variables
2. no requirement of specific assumptions concerning the distributional characteristics of the predictor variables (i.e., nonparametric)
3. the ability to include spatial structure in model
4. the accommodation of predictor variable interactions without a priori specification
5. the ability to model multiple responses variables in a single integrative model

Candidate approaches with key references

Generalized Additive Models

A general class of regression models that allow for independent and interactive effects of independent variables.

- Guisan, A., T. C. Edwards, Jr., and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157:89-100.

- Leathwick, J.R., Elith, J., and T. Hastie. 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecological Modelling 199:188-196.

Artificial Neural Networks
An artificial neural network (ANN), or, more generally, a multilayer perception, is a modeling approach inspired by the way biological nervous systems process complex information. The key element of the ANN is the novel structure of the information processing system, which is composed of a large number of highly interconnected elements called neurons, working in unity to solve specific problems. The feed-forward neural network trained by the back-propagation algorithm is the common approach to developing ANNs. During network training, observations are sequentially presented to the network and the back-propagation algorithm adjusts the weights in a backwards fashion, layer by layer, in the direction of steepest descent in minimizing the error function. ANNs can accommodate multiple response variables.

- Olden, J.D., Joy, M.K., and R.G. Death. 2006. Rediscovering the species in community-wide modeling. Ecological Applications 16:1449-1460.
- Olden, J.D. 2003. A species-specific approach to modeling biological communities and its potential for conservation. Conservation Biology 17:854-863.
- Olden, J.D., Lawler, J.J., and N.L. Poff. 2008. Machine learning methods without tears: a primer for ecologists. The Quarterly Review in Biology 83: 171-193.

Multivariate Adaptive Regression Splines
Multivariate adaptive regression splines (MARS) is a procedure for fitting adaptive non-linear regression that uses piece-wise linear basis functions to define relationships between a response variable and some set of predictors (Friedman, 1991). The resulting regression surface is piecewise linear and continuous. Breaks between segments are defined by a knot in a model that initially over-fits the data, which is then simplified using a backwards/forwards stepwise cross-validation procedure to identify terms to be retained in the final model. MARS allows simultaneous analysis of multiple species by identifying a common set of basis functions, with knots (and therefore variables) selected according to their improvement in predictive power of the model, averaged across all species.

- Friedman, J. H. 1991. Multivariate adaptive regression splines. Annals of Statistics 19:1–67.
- Elith, J. and Leathwick, J. 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. Diversity and Distributions 13: 265–275.
- Leathwick, J.R., 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. Freshwater Biology 50, 2034–2052.
- Leathwick, J.R., Elith, J. & Hastie, T. 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecological Modelling 199, 188–196.

Multivariate Regression Trees

Multivariate regression trees (MRT) are a natural extension to univariate classification and regression trees. Univariate regression trees explain the variation of a single numeric response variable using explanatory variables that may be numeric and/or categorical. They do this by growing a tree structure that partitions the data set into mutually exclusive groups, each of which has similar values of the response variable. Starting with all the data represented by a single node at the top of the tree, the tree is grown by repeated binary splitting of the data. Each split is defined by a simple rule, usually based on a single explanatory variable, and forms two nodes. Splits are (generally) chosen to maximize the homogeneity (minimize the impurity) of the resulting two nodes. MRT replace the univariate response with a multivariate response, and to redefine the impurity of a node by summing the univariate impurity measure over the multivariate response. Thus, to extend the sum of squares URT, impurity is defined as the sum of squares about the multivariate mean. Geometrically, this is simply the sum of squared Euclidean distances of sites about the node centroid. Random forests and boosted regression trees are two viable approaches that involve MRTs.

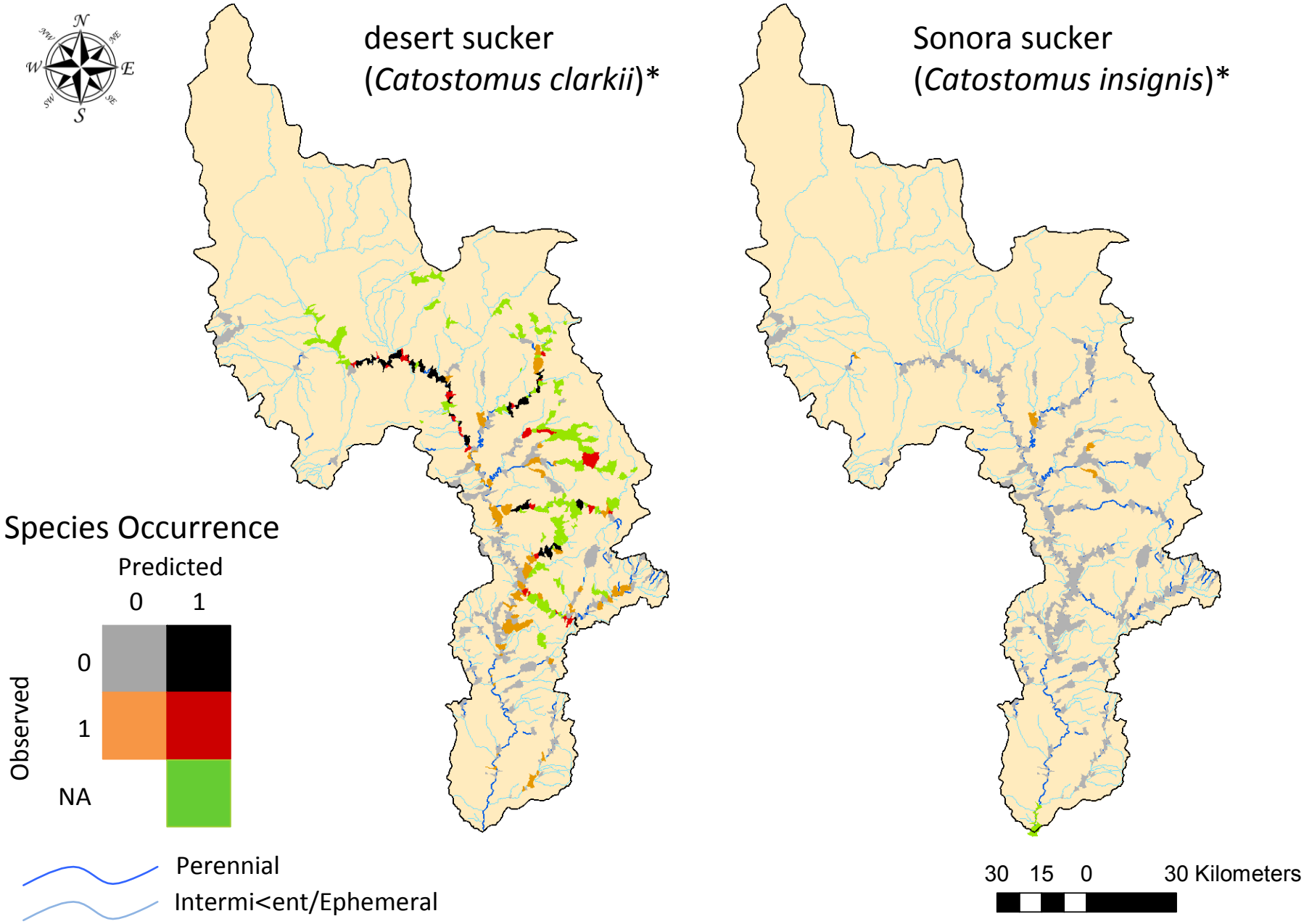
- De'ath, G. 2002. Multivariate regression trees: A new technique for modeling species-environment relationships. Ecology 83:1105-1117.
- De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for the analysis of complex ecological data. Ecology 81:3178-3192.
- Larsen, D. and P. Speckman. 2004. Multivariate regression trees for analysis of abundance data. Biometrics 60: 543-549.
- Olden, J.D., Lawler, J.J., and N.L. Poff. 2008. Machine learning methods without tears: a primer for ecologists. The Quarterly Review in Biology 83: 171-193.

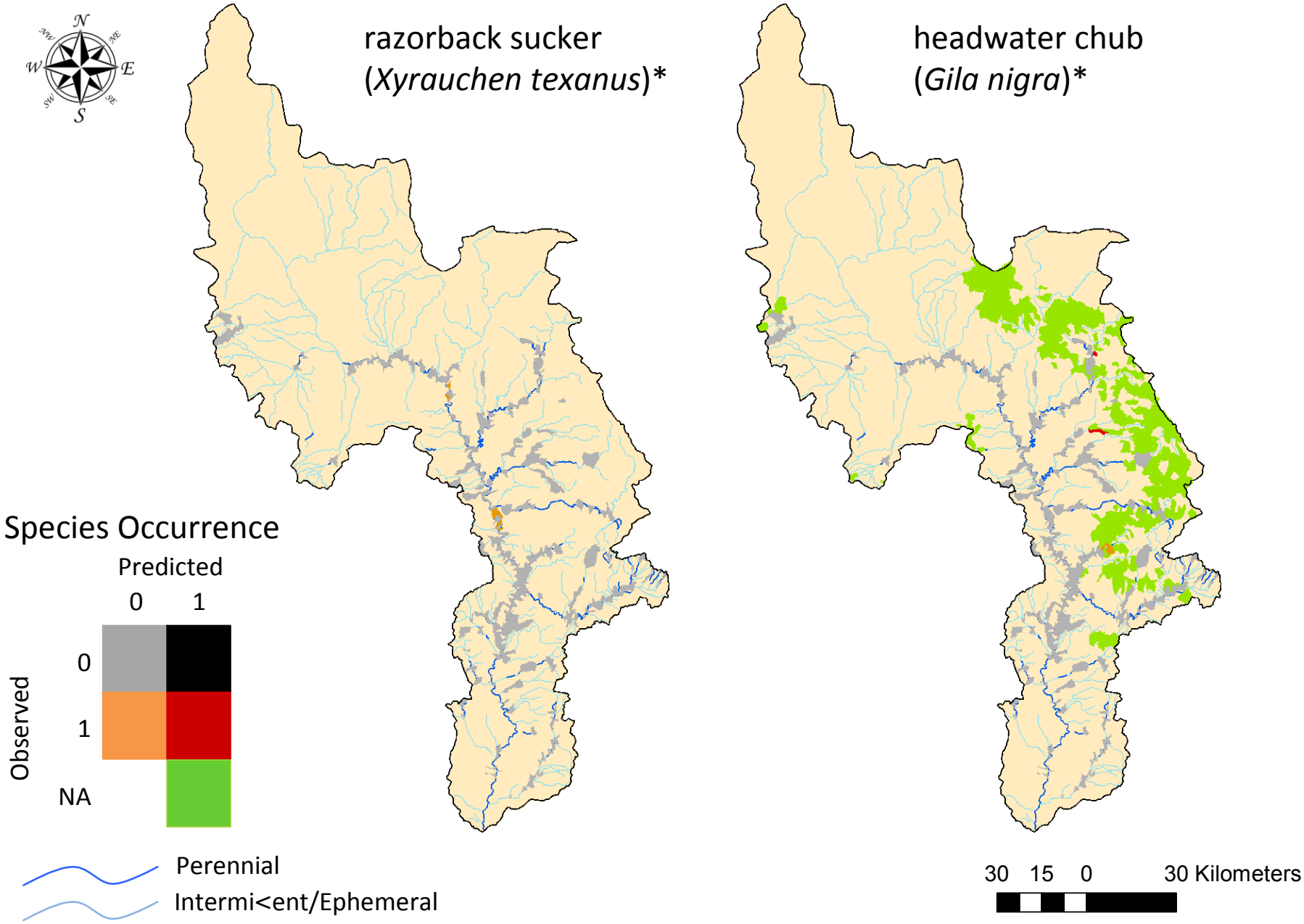
Maximum Entropy Model

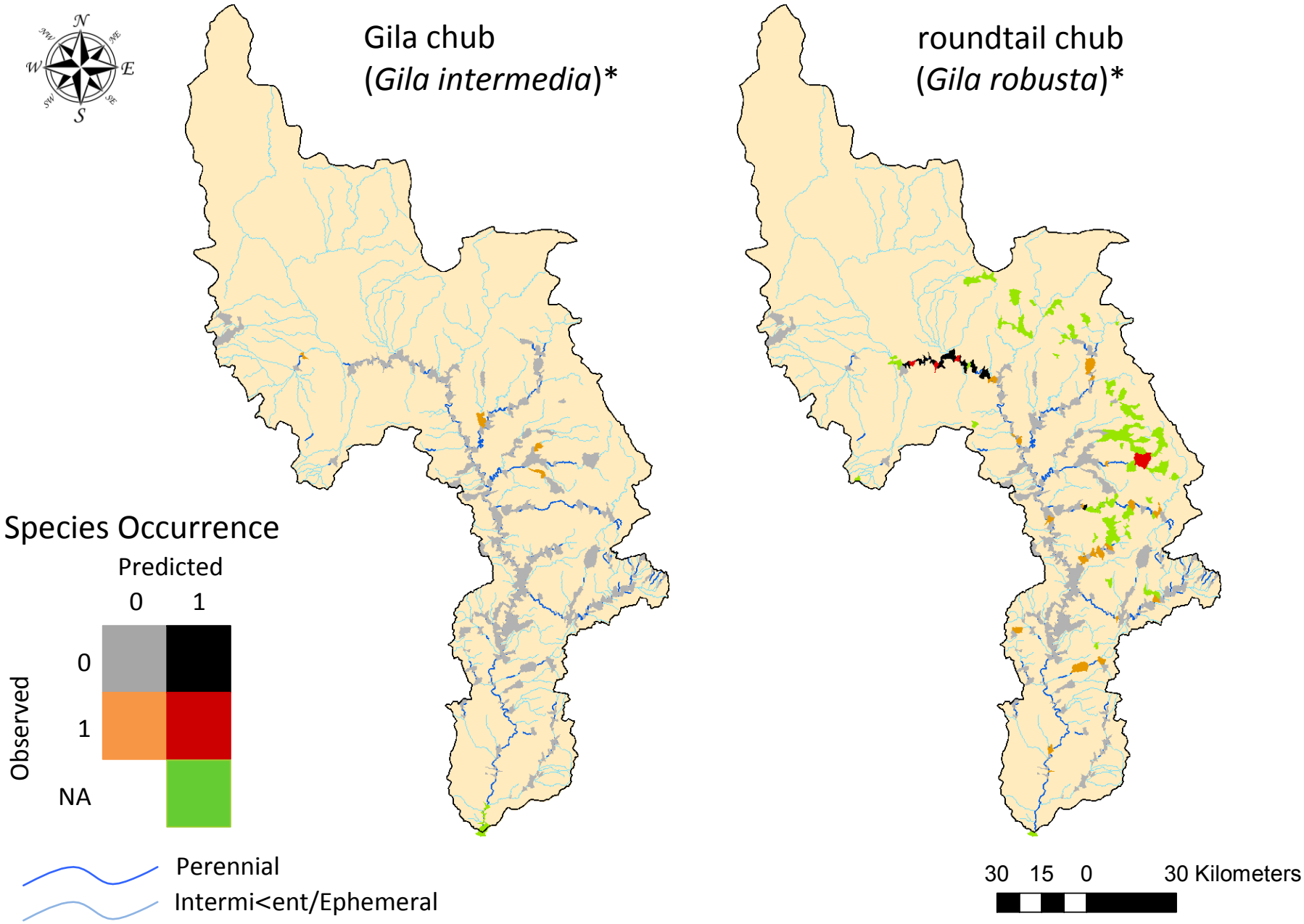
Maximum Entropy (MaxEnt), is a general-purpose method for making predictions or inferences from incomplete information. The idea of Maxent is to estimate a target probability distribution by finding the probability distribution of maximum entropy (i.e., that is most spread out, or closest to uniform), subject to a set of constraints that represent our incomplete information about the target distribution. The information available about the target distribution often presents itself as a set of real-valued variables, called “features”, and the constraints are that the expected value of each feature should match its empirical average (average value for a set of sample points taken from the target distribution).

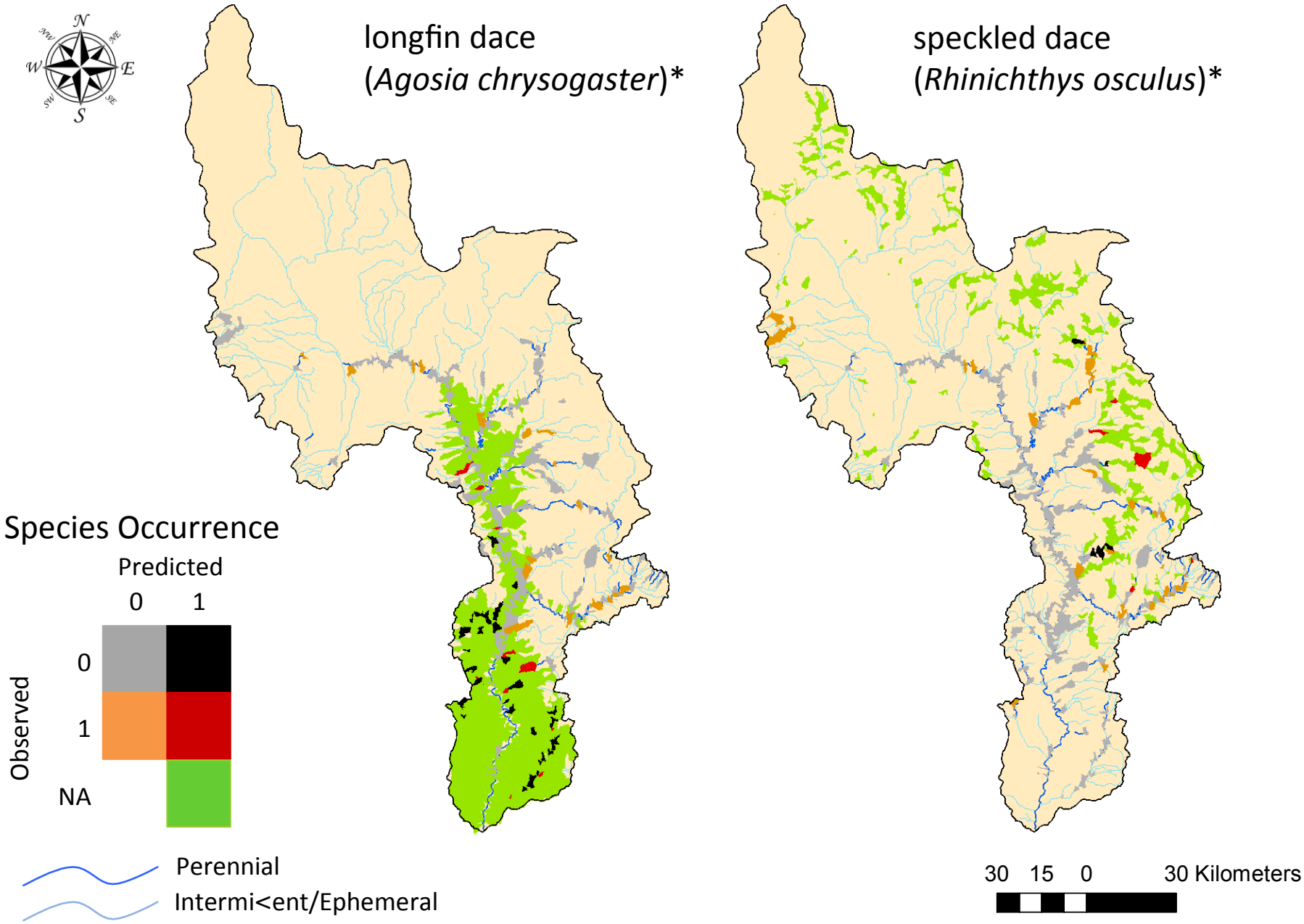
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. and Yates, C. J. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, 17: 43–57.
- Elith, J. & Leathwick, J.R. (2009a) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics, 40: 677–697.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190: 231–259.

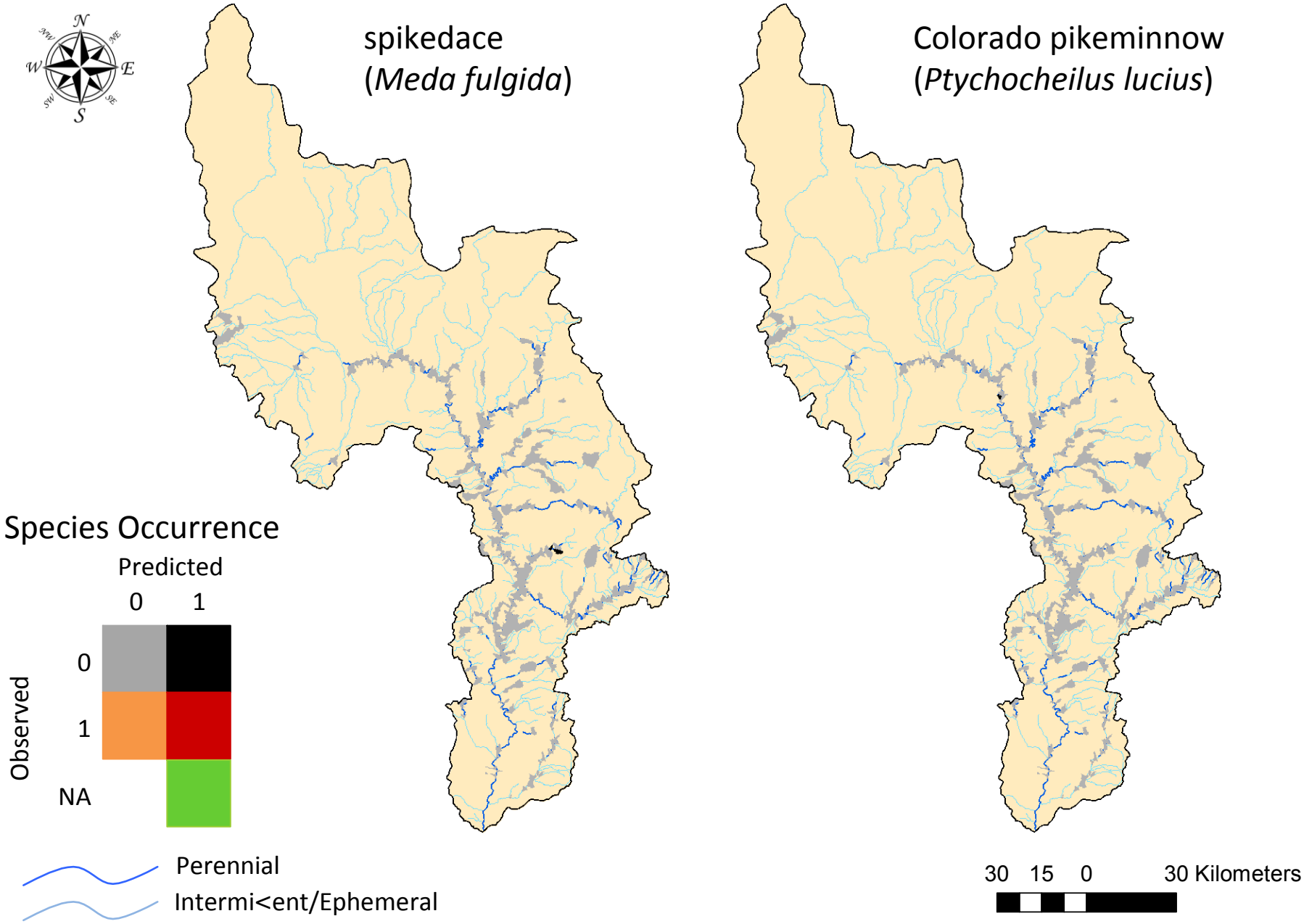
Appendix C. MARS predictions for native species of the Verde River Basin.

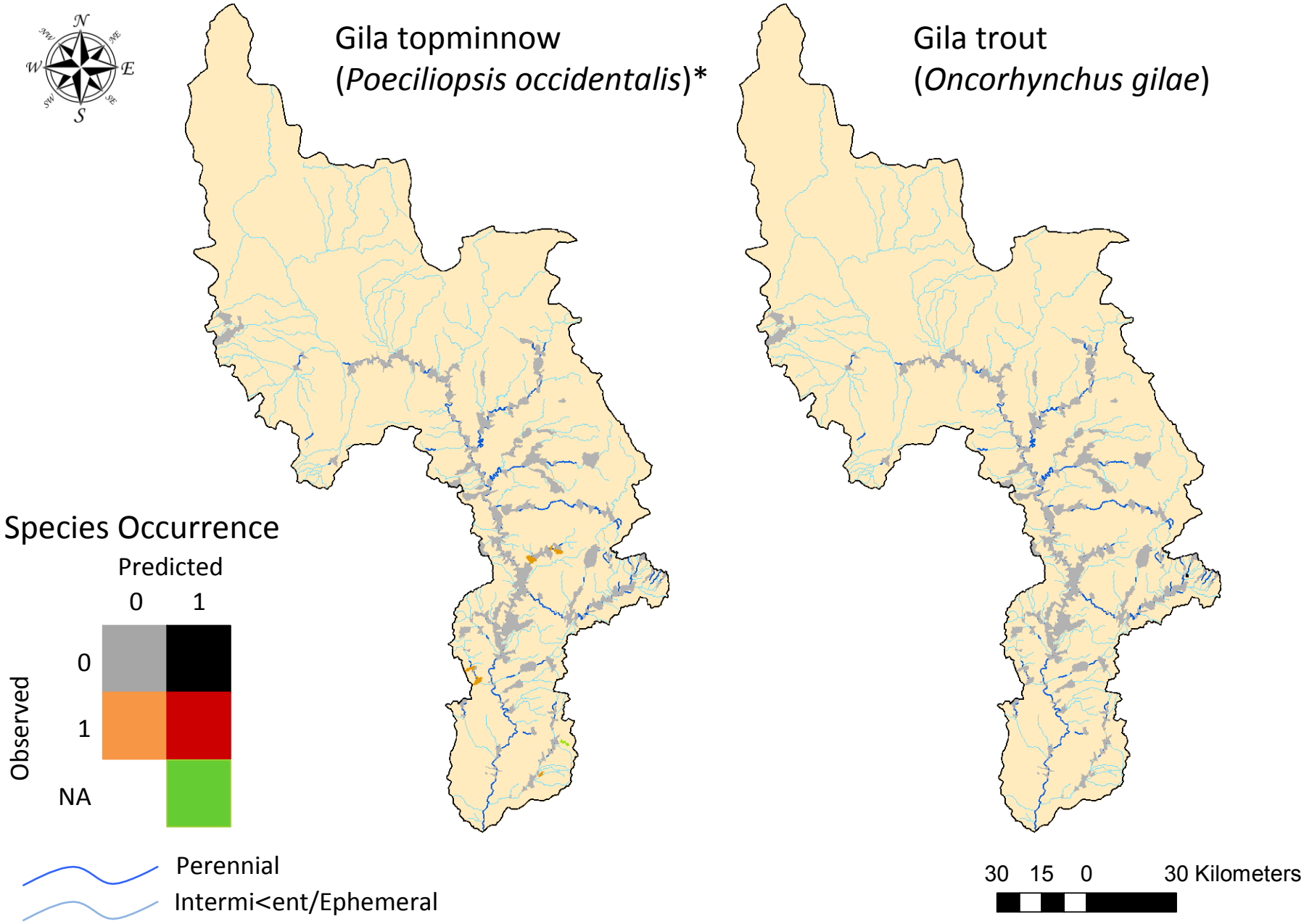


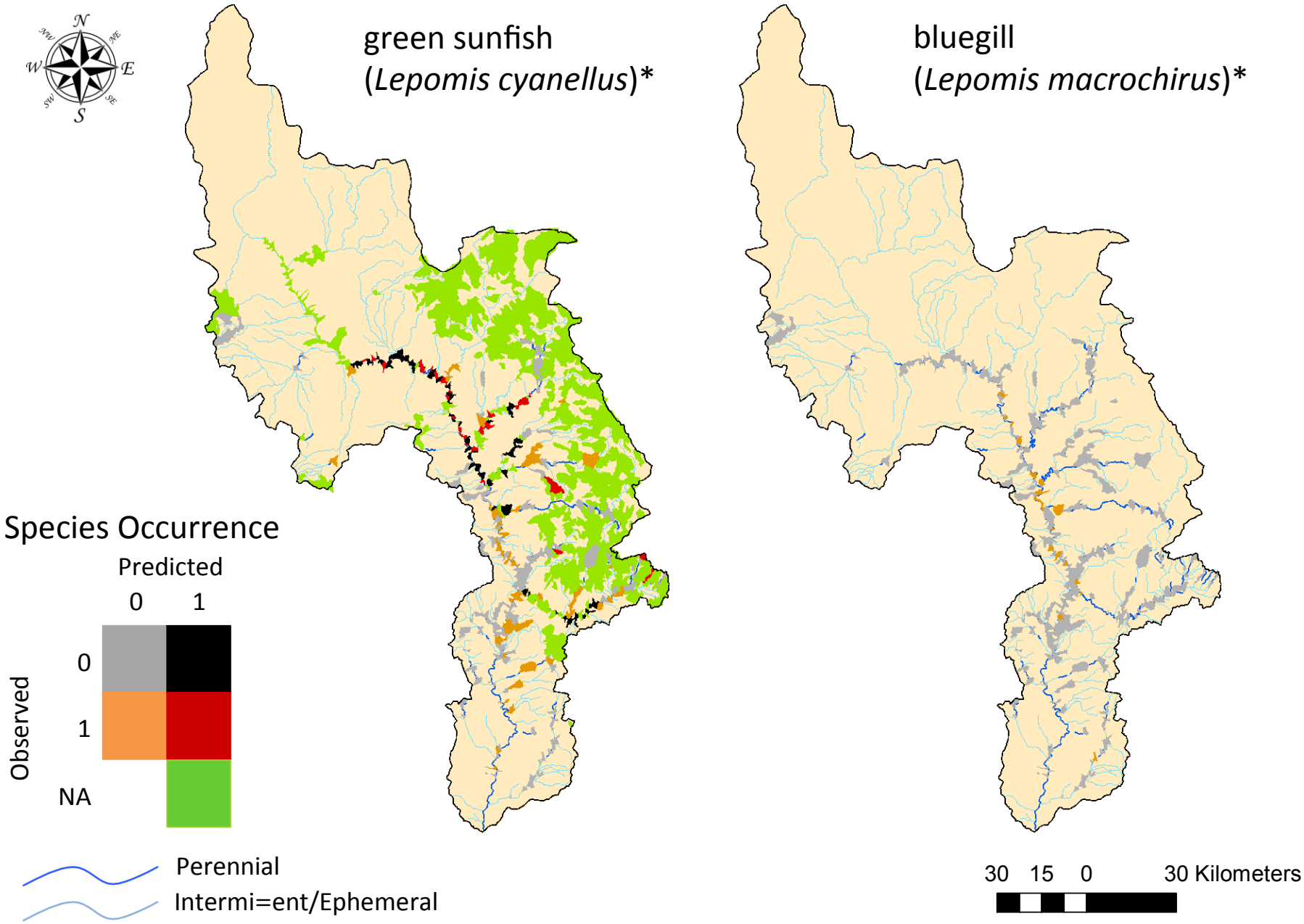

razorback sucker
(*Xyrauchen texanus*)*
headwater chub
(*Gila nigra*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi<ent/Ephemeral
30 15 0 30 Kilometers

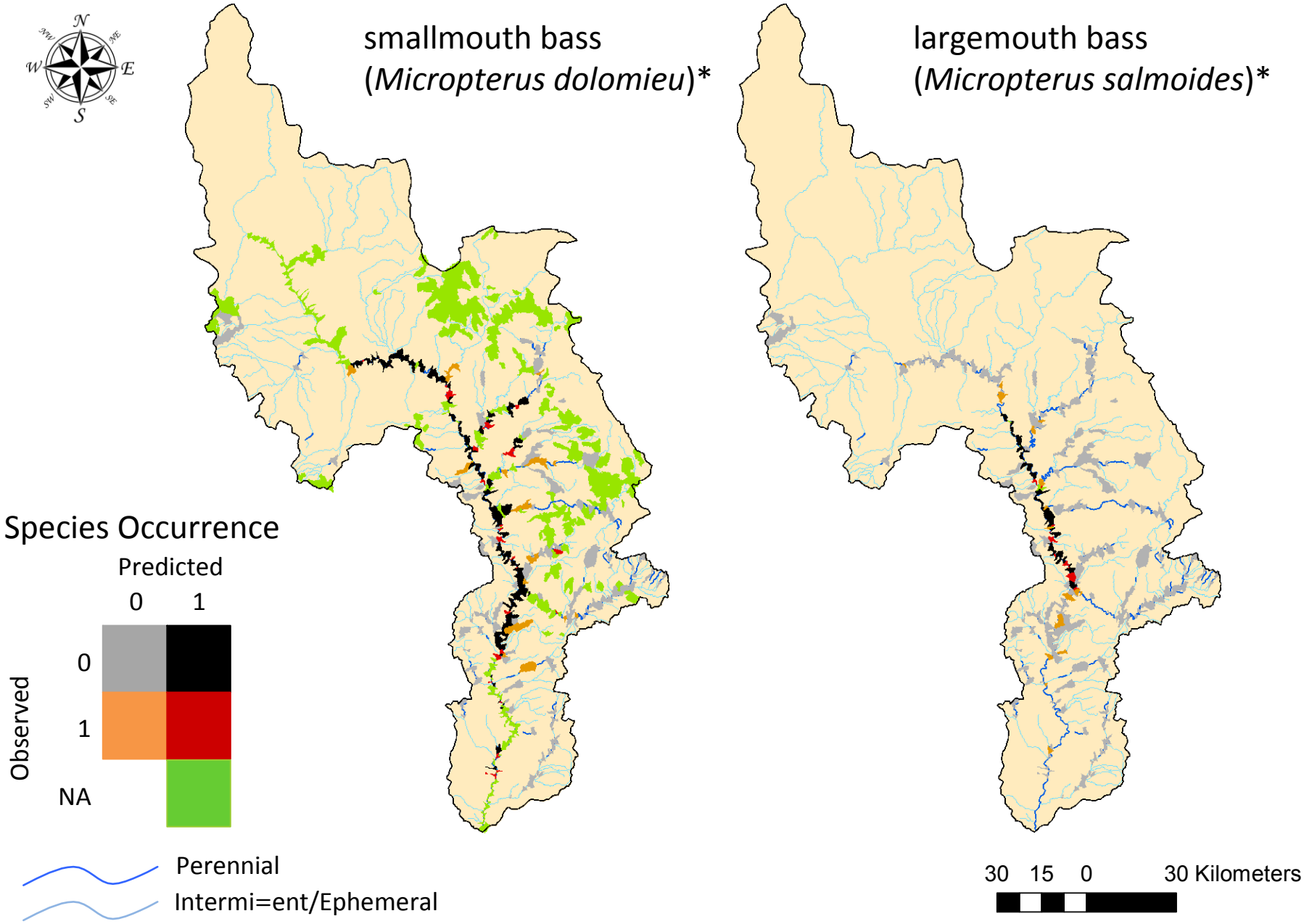

Gila chub
(*Gila intermedia*)*
roundtail chub
(*Gila robusta*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi<ent/Ephemeral
30 15 0 30 Kilometers

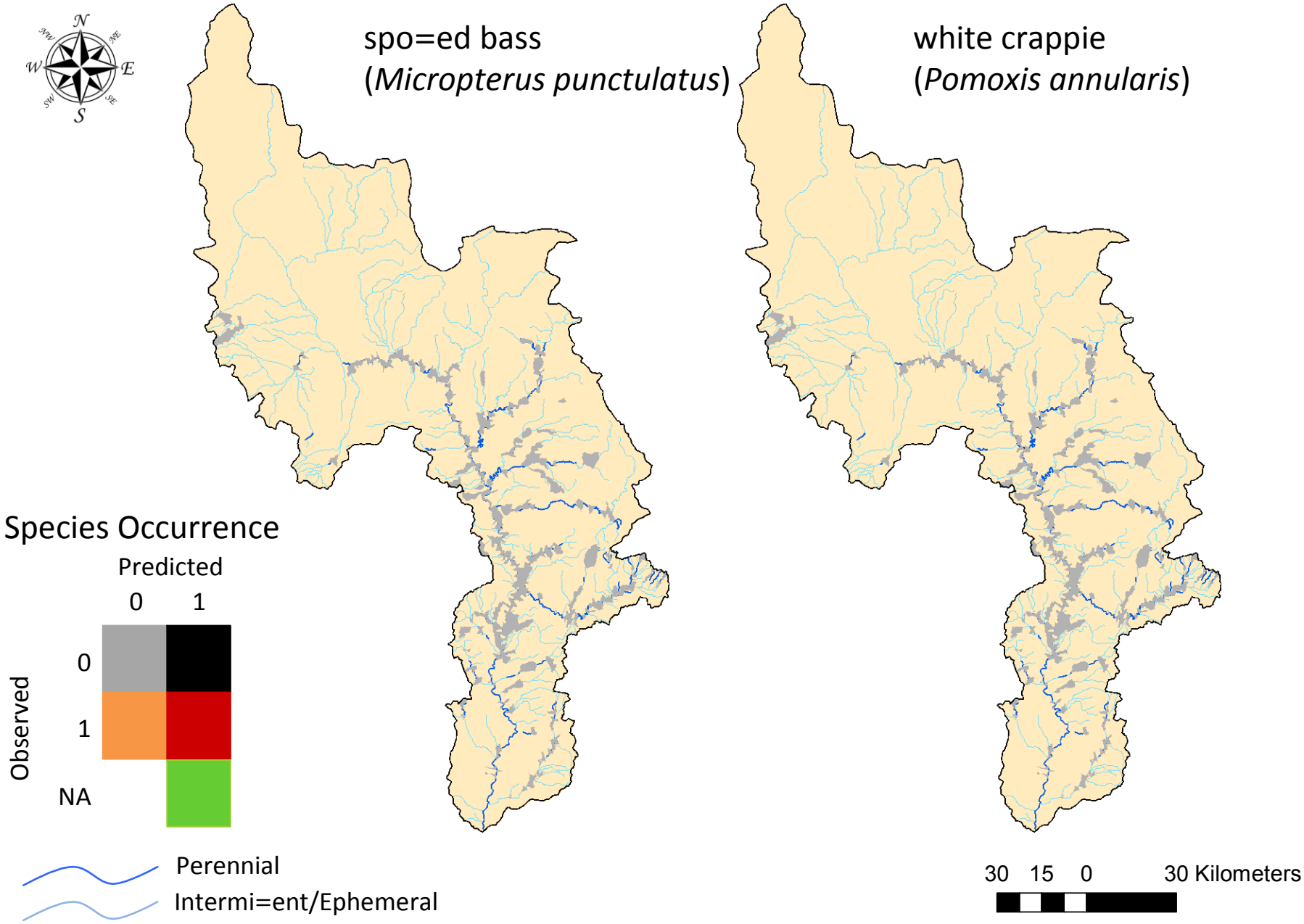

longfin dace
(*Agosia chrysogaster*)*
speckled dace
(*Rhinichthys osculus*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi<ent/Ephemeral
30 15 0 30 Kilometers

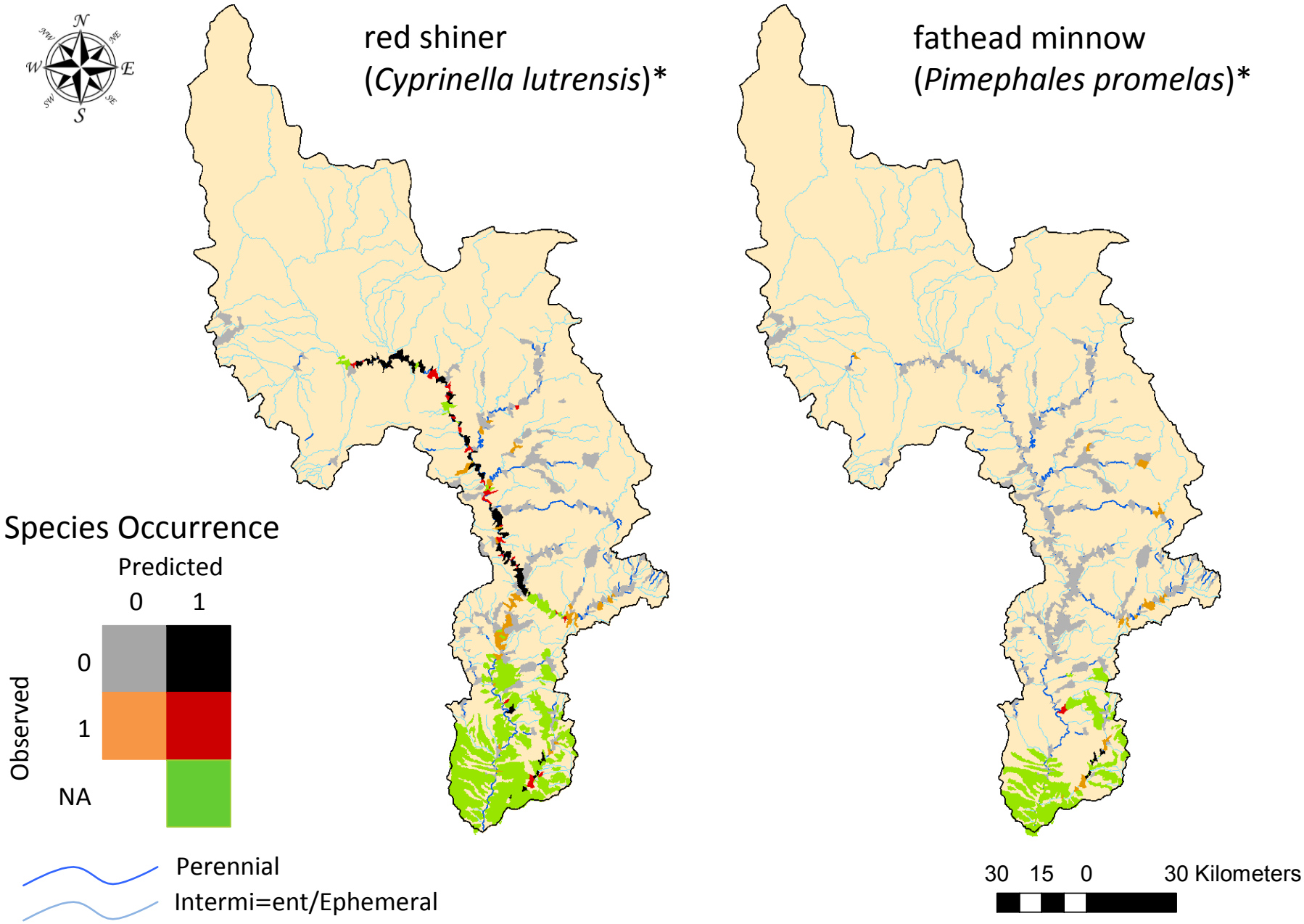

spikedace
(*Meda fulgida*)
Colorado pikeminnow
(*Ptychocheilus lucius*)
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi<ent/Ephemeral
30 15 0 30 Kilometers

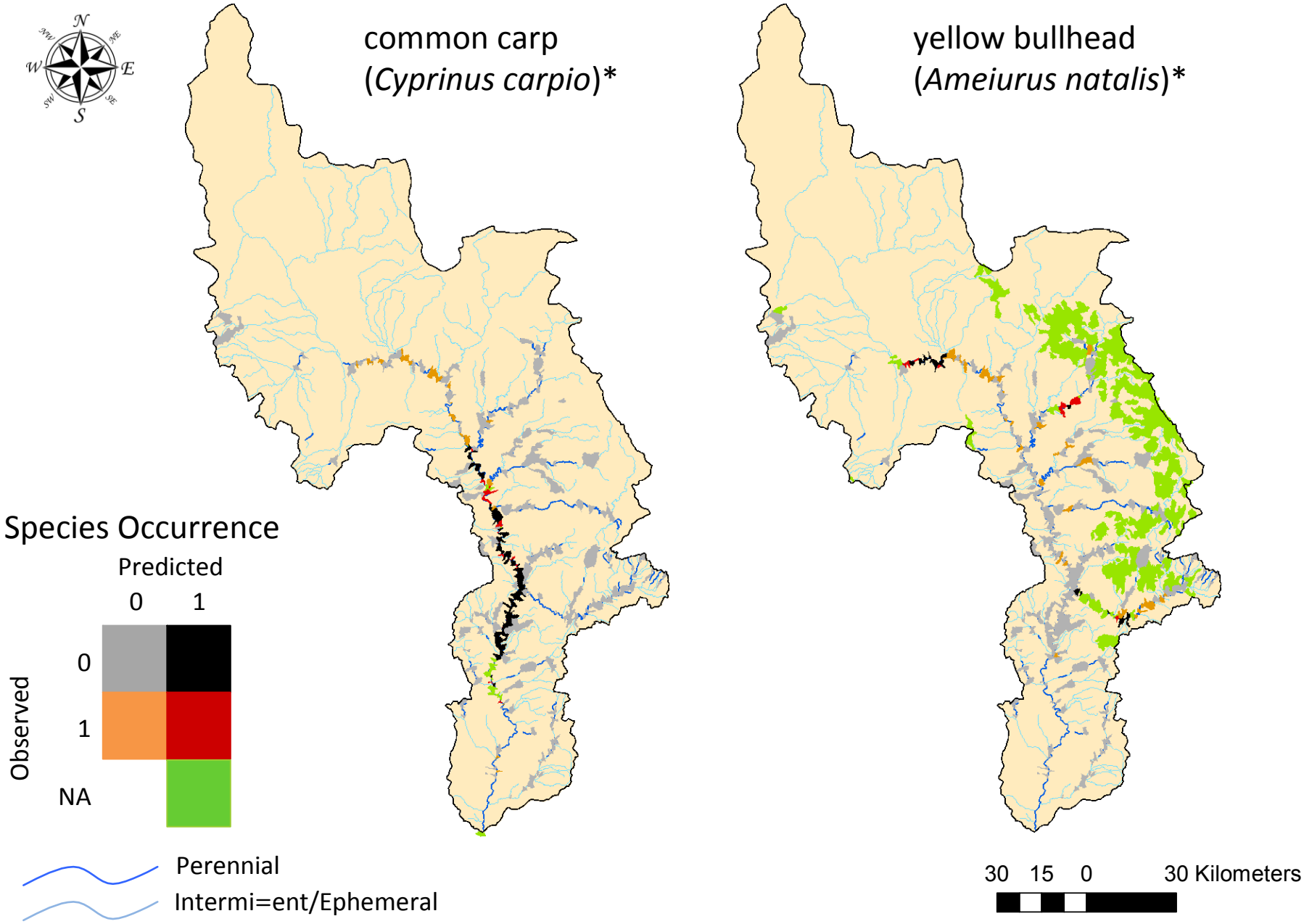

Gila topminnow
(*Poeciliopsis occidentalis*)*
Gila trout
(*Oncorhynchus gilae*)
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi<ent/Ephemeral
30 15 0 30 Kilometers

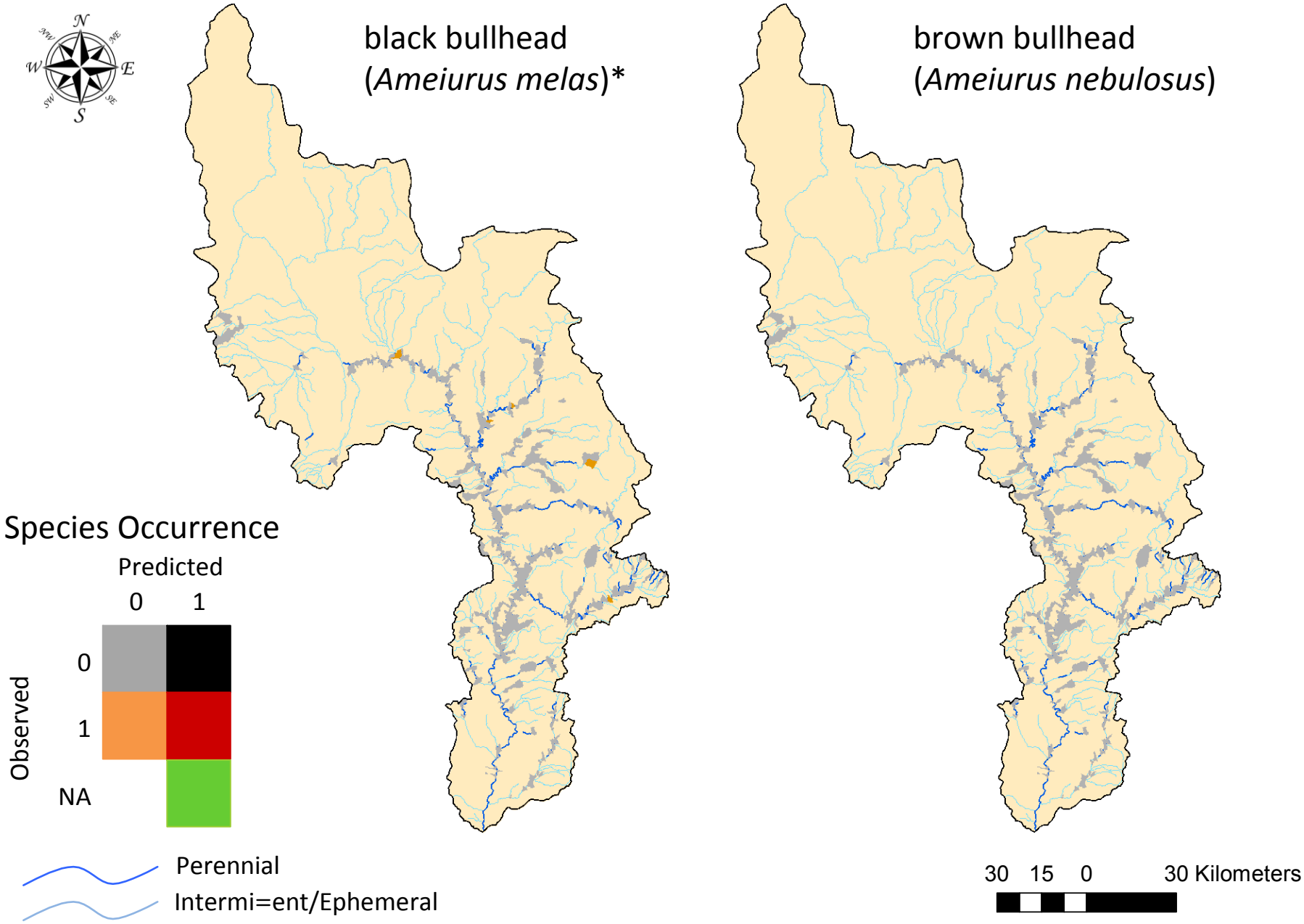

N
NW
NE
W
E
SW
SE
S
green sunfish
(*Lepomis cyanellus*)*
bluegill
(*Lepomis macrochirus*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers

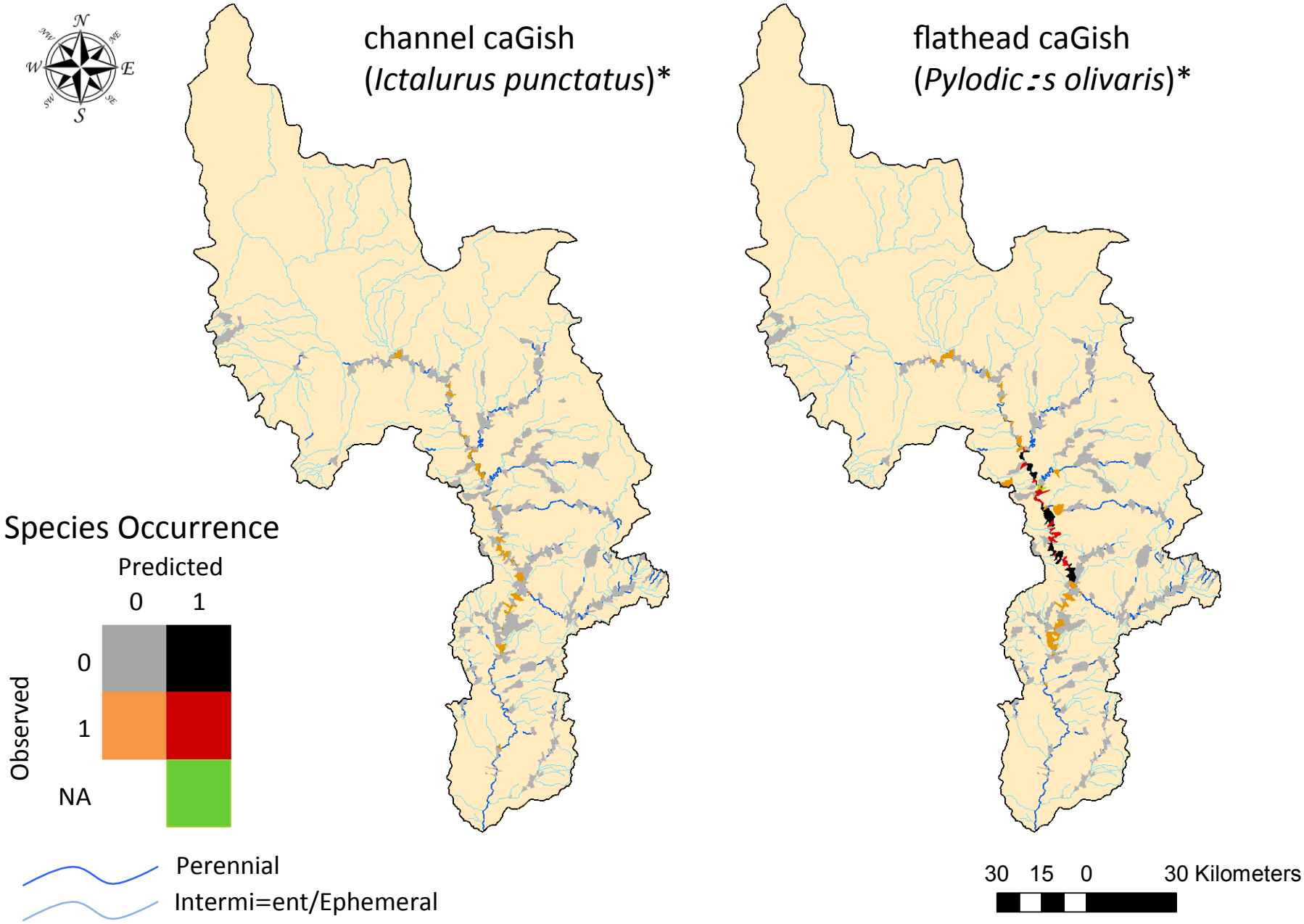

smallmouth bass
(*Micropterus dolomieu*)*
largemouth bass
(*Micropterus salmoides*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers

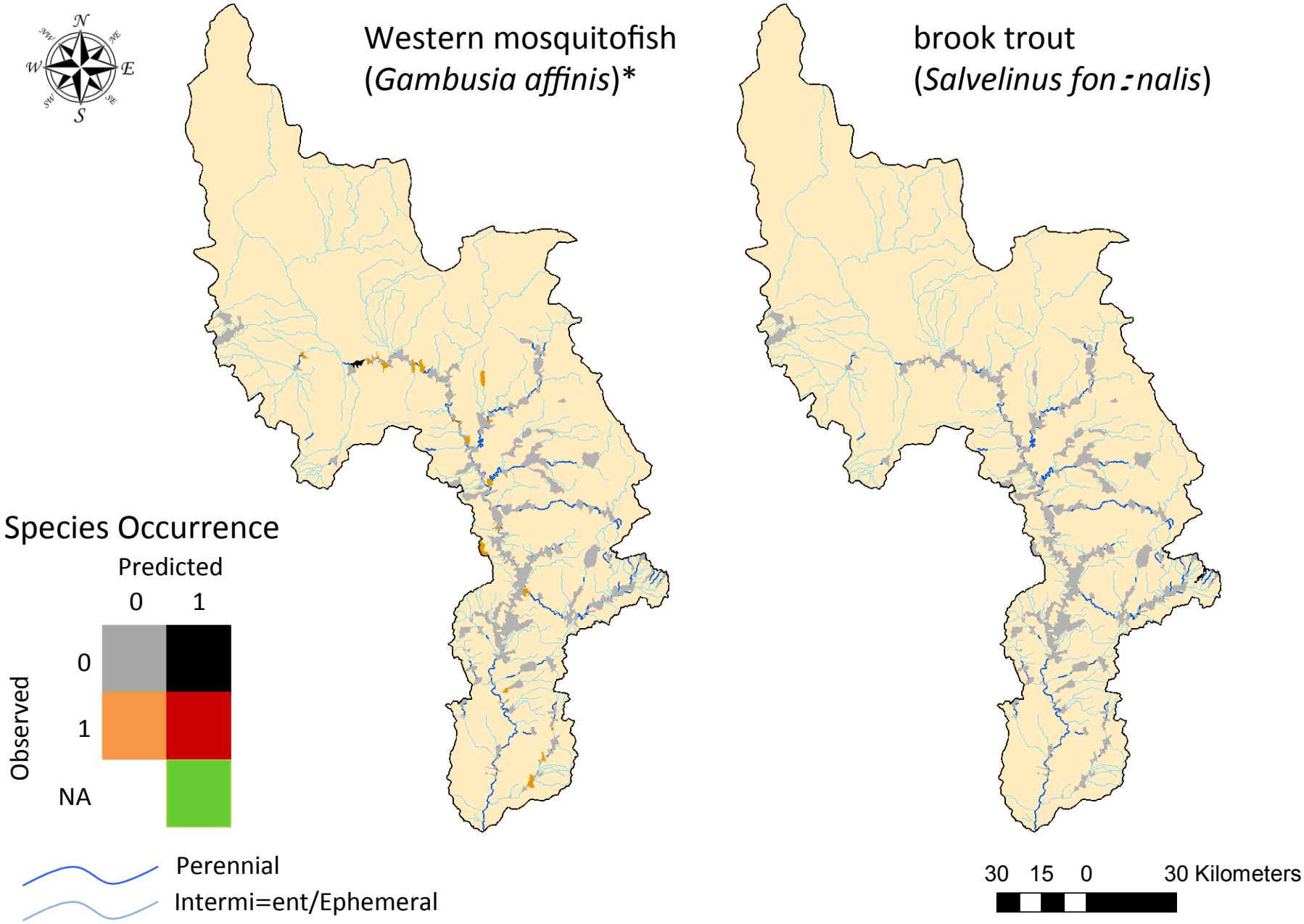

spo=ed bass
(*Micropterus punctulatus*)
white crappie
(*Pomoxis annularis*)
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers

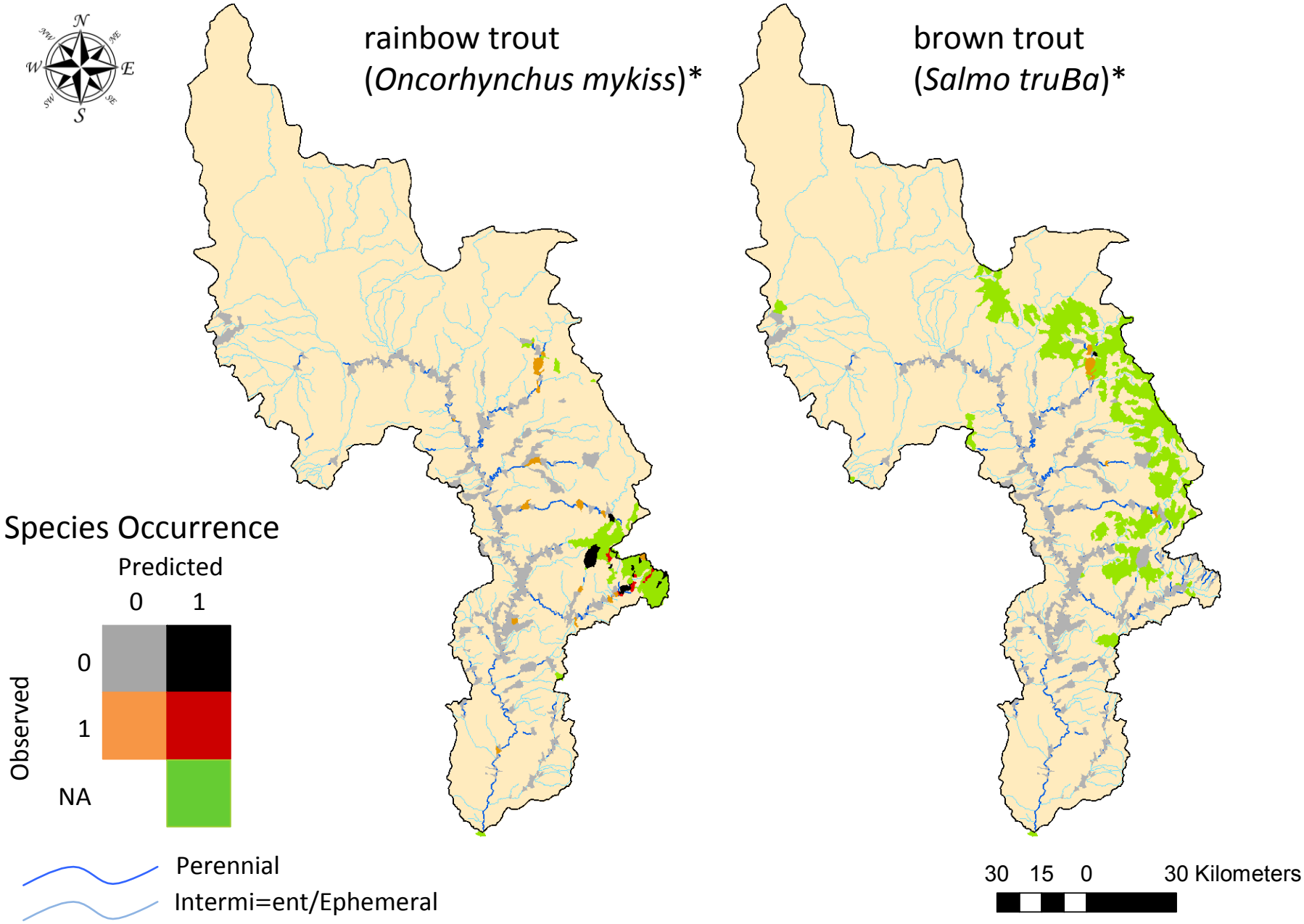

red shiner
(*Cyprinella lutrensis*)*
fathead minnow
(*Pimephales promelas*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers


common carp
(*Cyprinus carpio*)*
yellow bullhead
(*Ameiurus natalis*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers


black bullhead
(*Ameiurus melas*)*
brown bullhead
(*Ameiurus nebulosus*)
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers


channel caGish
(*Ictalurus punctatus*)*
flathead caGish
(*Pylodic:s olivaris*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers


N
NW
NE
W
E
SW
SE
S
Western mosquitofish
(*Gambusia affinis*)*
brook trout
(*Salvelinus fon:nalis*)
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers


rainbow trout
(*Oncorhynchus mykiss*)*
brown trout
(*Salmo truBa*)*
Species Occurrence
Predicted
0
1
Observed
0
1
NA
Perennial
Intermi=ent/Ephemeral
30 15 0 30 Kilometers

Appendix E. Trip reports from ground truth field surveys.

Oak Creek Fish Survey

Fish surveys were conducted on Oak Creek between May 14, 2014 and May 15, 2014 by Jonathan Miller, Taylor Cotten and Dan Leavitt. Reaches 15060202000118, 15060202000119 and 15060202000120 were surveyed approximately 9km southwest of Sedona, AZ (Figure 1). Reaches 15060202000128, 15060202000130, 15060202000133 and 15060202000136 were surveyed located 1km to 13km northeast of Sedona, AZ (Figure 1). Surveys were conducted using a Halltech HT-2000 backpack electroshocker. Fish caught in the southwestern reaches consisted of flathead catfish (PYOL), rock bass (AMRU), smallmouth bass (MIDO), green sunfish (LECY), yellow bullhead (AMNA), black bullhead (AMME), red shiner (CYLU), fathead minnow (PIPR), Sonora sucker (CAIN) and brown trout (SATR) (Figure 2). Non-native fish were the most numerous fish present with LECY, MIDO and AMRU abundant. CAIN was the only native fish found and it was only found in reach 15060202000120. All other species were intermittently found within the reaches. Crayfish were abundant in all reaches. Bullfrogs were detected in reach 15060202000118. The creek consisted of slower moving water than reaches north of Sedona. Width was generally between 10-15m. Deep and long pools were fed by slow to moderate flow with very few riffles. The substrate consists of silt, mud, cobble and small boulders. The water temperature ranged from 17.6°C to 20.2°C. Canopy cover was moderate with much of the stream seeing ample sunlight throughout the day. The reaches Northeast of Sedona contained SATR, rainbow trout (ONMY), speckled dace (RHOS), CAIN, CACL and AMRU (Figure 3). Within the northeastern reaches SATR and RHOS were abundant. CAIN and CACL were both common. SATR was infrequent and AMRU was only found in reach 15060202000128 just north of Sedona. Crayfish were seen throughout all of the reaches surveyed. Water flow in these reaches was swift with many riffles. Large boulders were found throughout these stretches which provided ample habitat. Small pools were encountered as well; however they were less frequent and smaller than those found in the southwestern reaches. Water temperatures for the northeastern reaches ranged from 11.5°C to 18.2°C.

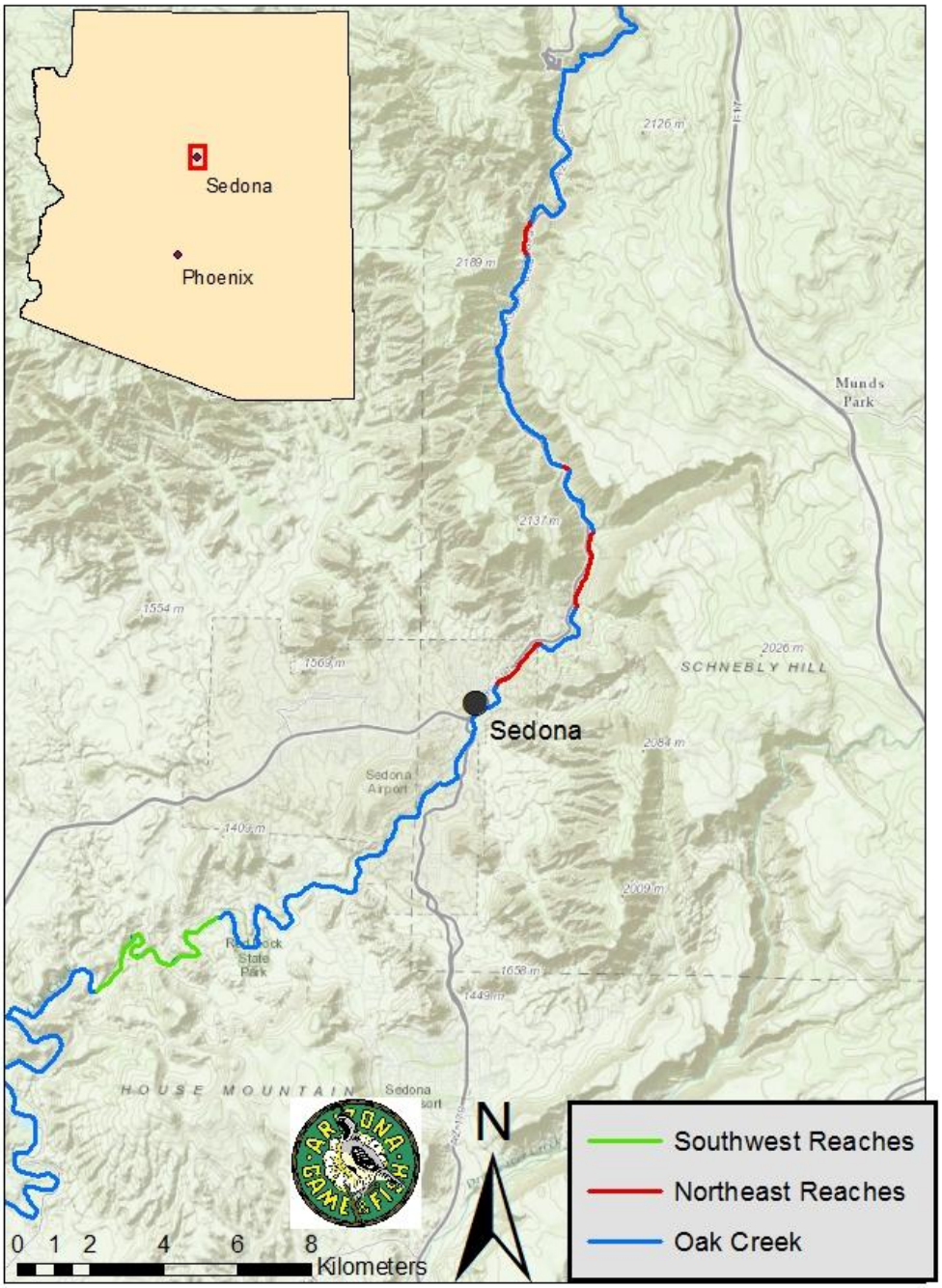


Figure 1 Survey Locations along Oak Creek

Figure 2 Brown Trout (SATR)

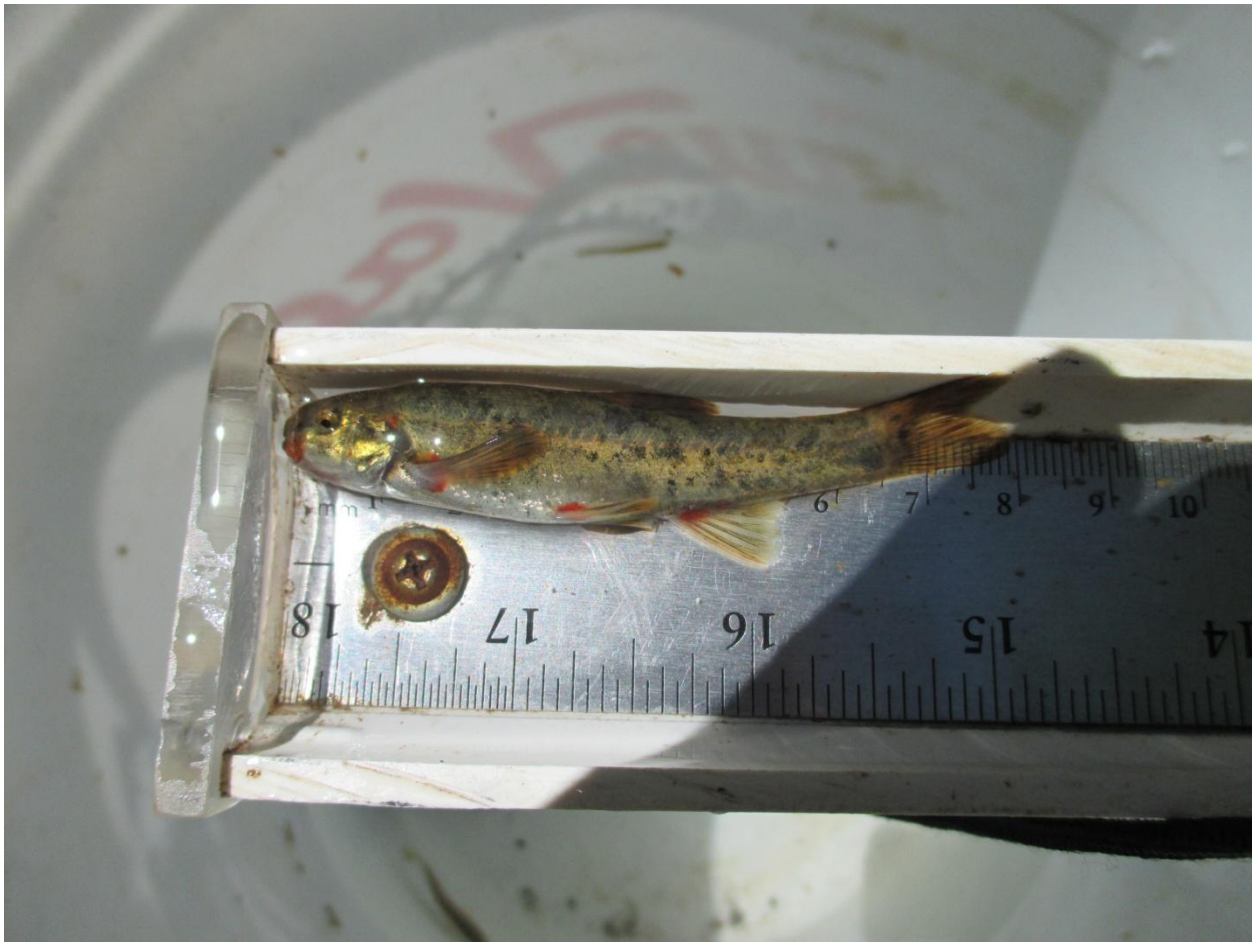


Figure 3 Speckled Dace (RHOS)

Rock Creek Fish Survey

On May 19, 2014 an electroshocking survey was conducted on the southern most reach (15060203000462) of Rock Creek (E. Verde tributary approximately 15 kilometers west of Payson) by Taylor Cotten and Jonathon Miller. We covered approximately 1.8km of water over six hours using a Halltech HT-2000 backpack electrofisher (Figure 1). Fish captured in the stream included longfin dace (AGCH), desert sucker (CACL), red shiner (CYLU), yellow bullhead (AMNA), green sunfish (LECY), speckled dace (RHOS), and chub (GINI). LECY was the most abundant species, especially in the lower portion closer to the confluence with the E. Verde. LECY numbers and crawfish numbers appeared to decrease the further upstream we went. AGCH was relatively common throughout with RHOS becoming more prominent the further upstream we went as the water between pools became shallower. Rock Creek was generally shallow throughout and usually less <3m wide with staggered pools. Average depth for the entire survey stretch was around 0.25m or less. Dominant riparian plant species included sycamore (*Platanus* sp.), alder (*Alnus* sp.), cottonwood (*Populus* sp.), willow (*Salix* sp.), ash (*Fraxinus* sp.) and columbine (*Aqualegia* sp.). The habitat continued intermittently upstream from the end point of this survey with staggered small pools, but the volume of water was markedly decreasing.

Chub Observations

Pool at 12S 453209 3788479 (NAD 83), lies directly adjacent to a spring coming out of the west bank (Figure 2). The pool was approximately 4x25m and approximately 1.5m in depth at the deepest points. The water was very murky and we managed to shock one 120mm chub at 12:57 (Figure 3). Other species in the pool included AGCH, RHOS, and LECY. Water temperature was 19.8°C and air temperature was 27.3°C. The spring on the west bank was very active and likely a major contributor to downstream flow on Rock Creek, the creek became noticeably narrower and intermittent in places beginning upstream from this pool.

Pool at 12S 453196 3788591 (NAD 83), we captured one chub (132mm) in this pool at 13:31 and observed at least 2 others (Figure 4). The second pool was approximately 3x12m and >2m in depth at the deepest point and wraps around a large boulder on the west bank (Figure 5). LECY were also observed and captured in the second pool. Water temperature was 19.4°C and air temperature was 28.4°C.

Pool at 453213 3788699 (NAD 83), extremely clear plunge pool with little to no vegetation directly on the banks and approximately 2m at the deepest points. We were not able to shock any fish from this pool but did observe at least two GINI as well as numerous LECY and at least one CACL. The clarity and depth of the water and adjacent large undercut boulder on the north side of the pool made sampling with the backpack shocker difficult.

Pool at 12S 453224 3788873 (NAD 83), this pool was extremely murky with a large over-hanging boulder on the west bank and we were not able to confirm the presence of chub. We did record LECY from this pool and the habitat was analogous to the other pools supporting chub within the reach.

Pool at 12S 453330 3788917 (NAD 83), the pool was approximately 3mx12m and approximately 1m in depth at the deepest points with a substrate consisting primarily of exposed bedrock and sand (Figure 6). We captured two GINI (105mm and 122mm) in this pool T 15:09 along with LECY and AGCH (Figures 7-8).

Pool at 12S 453333 3789080 (NAD 83), this site was approximately 3x16m and around 1m deep at the deepest points (Figure 9). We captured one GINI (68mm) AT 15:26 and observed at least two more (Figure 10). We also captured and observed LECY and AGCH. The stream bed and banks around the pool consisted of sand and cobble with exposed root networks. This was also the upstream terminus of our survey. Water temperature was 21.8°C and air temperature was 28.0°C.

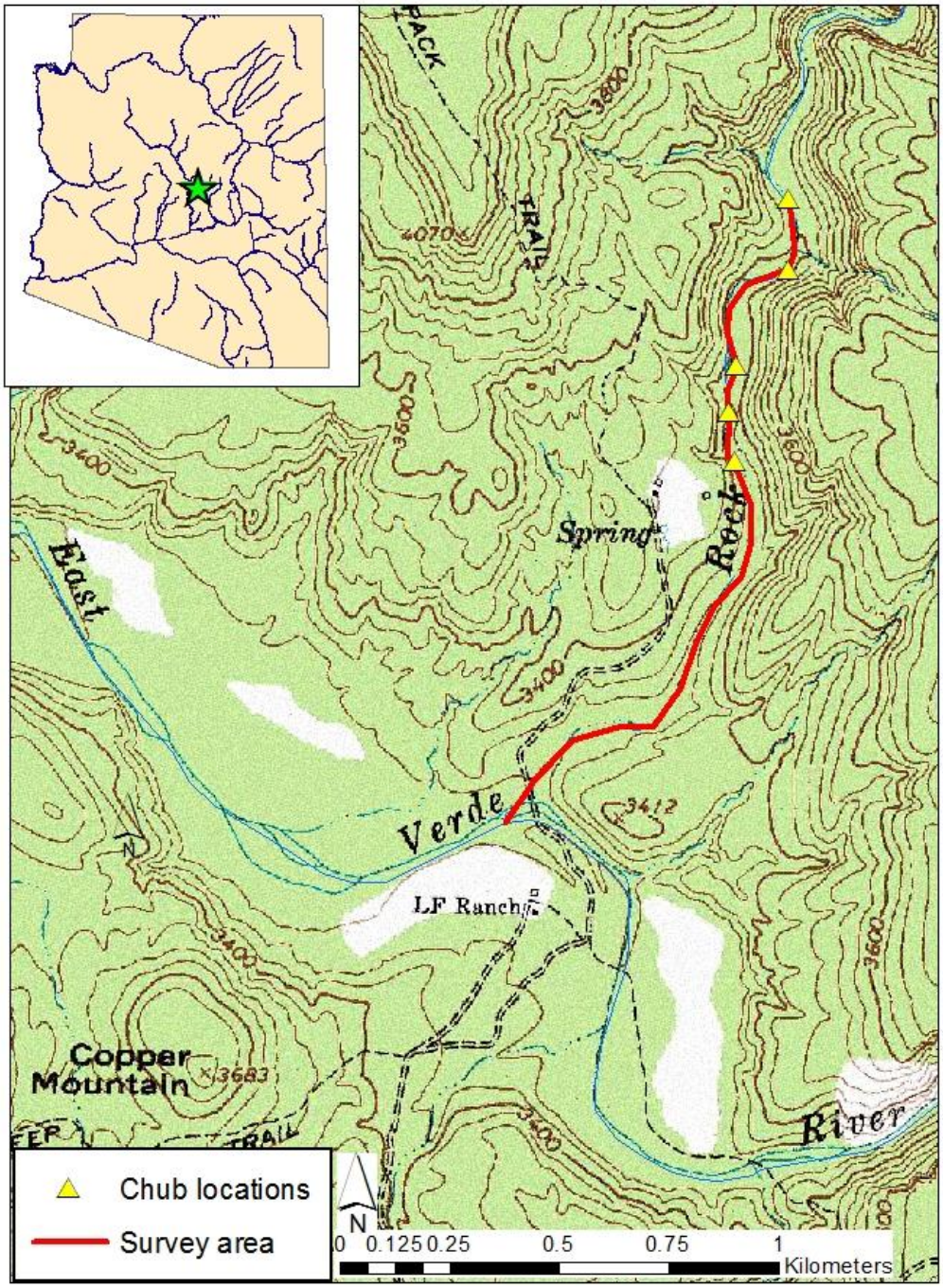


Figure 1. Survey location and location of GINI captures/observations.

Figure 2. First pool encountered on Rock Creek where GINI were caught.

Figure 3. GINI caught in the first pool.

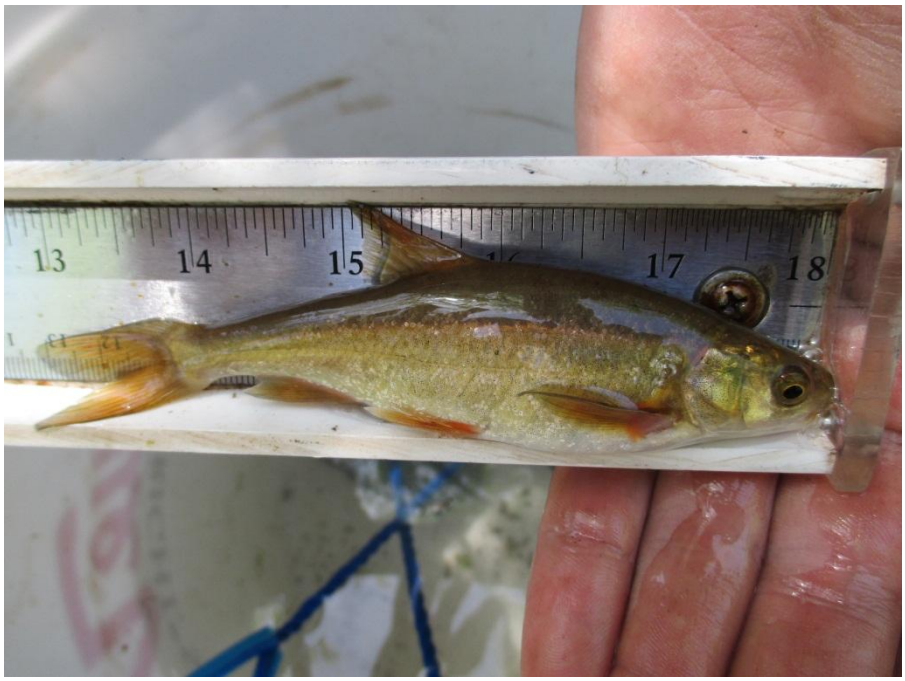


Figure 4. GINI caught in the second pool.

Figure 5. Second pool encountered on Rock Creek where GINI were caught.

Figure 6. Third pool encountered on Rock Creek where GINI were caught.

Figure 7. One of two GINI caught in the third pool.

Figure 8. Second of two GINI caught in the third pool.

Figure 9. Final pool encountered on Rock Creek survey where GINI were caught.

Figure 10. GINI caught in final pool of Rock Creek survey.

Verde River Drainage Fish Surveys

Fish surveys were conducted in the Verde River drainage for the purpose of bolstering the Fisheries Information Systems (FINS) for the Arizona Game and Fish in order to provide more data to make scientific fisheries decisions in the future. Many portions of streams in this system do not contain fish records and collecting records from previously un-surveyed sites in streams is the goal of this project. Reconnaissance and fish surveys were conducted on tributaries of the Verde River from April 14 to April 19, 2014. Reconnaissance consisted of locating streams with a moderate to high probability of containing water throughout the year through Geographic Information Systems (GIS), historic fish surveys and regional biologist's knowledge of potential streams. Streams were divided into reaches which consisted of a section between two tributaries. Those were then narrowed down into reaches that had no previous fish records. Reaches that had moderate to high probability of water and had no fish surveys on record were identified as primary target areas for surveys. Based on the factors listed above, streams that were identified as potential survey sites included Sycamore creek, East and West Fork Sycamore Creek, Alder creek and Lime creek (Figure 1). Reaches without fish records are highlighted in blue in Figure 1.

Sycamore Creek

Sycamore Creek flows along much of state route 87 south of Payson, AZ and enters the Verde north of the state route 87/Verde River crossing, east of Fountain Hills, Arizona. On April 14 Jonathan Miller and Chad Rubke conducted initial site visits to determine access to reaches in the stream and if water and/or fish were present for future electroshocking surveys. On April 15, 2014 through April 17, 2014 electroshocking surveys were conducted on four reaches and one dip net survey was conducted on a fifth reach by Taylor Cotten and Tyler Nord. Fish caught in the stream included Longfin dace (AGCH), Fathead minnow (PIPR) and Desert sucker (CACL). AGCH was abundant in all reaches surveyed while PIPR was fairly common and CACL was uncommon and only found in one reach. The reach that was dip netted was previously dry upon our initial site visit but had seasonal flow enough for a dip net survey where unidentified fry and Canyon tree frogs (HYAR) tadpoles were caught. Some fry were collected for further identification. Black-necked Garter snakes (THCY) and Lowland Leopard frogs (LIYA) were also seen in several places. Evidence of Bullfrog and crayfish were also documented.

East and West Fork Sycamore Creek

The east and west fork of sycamore creek are located west of state route 87 north of sunflower, AZ. They consist of the headwaters for Sycamore creek. They were surveyed for access and presence of water and fish by Jonathan Miller and Chad Rubke. A follow up dip net survey was conducted on the East fork by Taylor Cotten and Tyler Nord on April 17, 2014. The stream had very little flow or depth and no fish were observed.

Alder Creek

Alder Creek flows west from Sunflower to Bartlett Lake. It was initially surveyed by Jonathan Miller and Chad Rubke for access and presence of water and fish. Water was seen at 2 separate locations, one of

which longfin dace were caught and presumably HYAR tadpoles were seen and collected. Taylor Cotten and Jonathan Miller conducted a dip net survey of two reaches on April 18, 2014. One reach started with a pool of water at 452597E 3751296N that had flow downstream from it for approximately 291m until water disappeared at 452741E 3751416N. This reach had no fish present. The second reach had low flow for approximately 785m before it disappeared at 448838E 3747199N. Longfin dace were sampled in this reach almost all of which were collected from a pool where the surface water began at 448131E 3747204. HYAR tadpoles were also seen in this same pool.

Lime Creek

Lime Creek flows from a drainage east of New River and flows east into Horseshoe Lake. This stream was surveyed on April 19, 2014. The upstream most reaches were looked at but no flow existed as was the case with the downstream most reach. To conduct surveys into the middle reaches of this stream will require more time as there is limited access to those reaches. Potential overnight backpack trip would be needed for a future survey in this area.

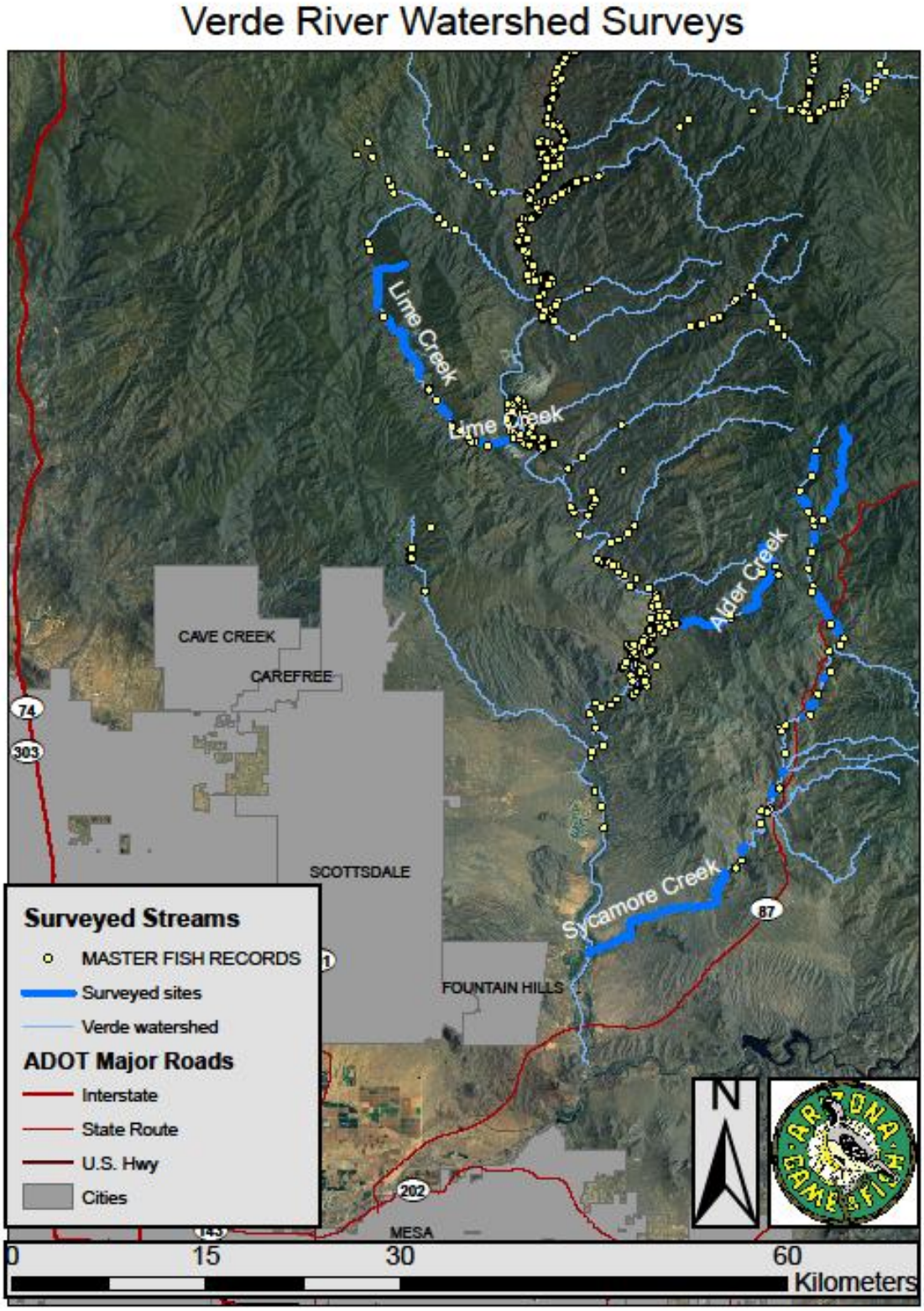


Figure 1. Surveyed sites within the Verde River Watershed