8-2023

Comparing Commonly Used Aquatic Habitat Modeling Methods for Native Fishes

Eryn K. Turney
Utah State University

COMPARING COMMONLY USED AQUATIC HABITAT MODELING METHODS FOR NATIVE FISHES

by

Eryn K. Turney

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Ecology

Approved:

Sarah E. Null, Ph.D. Major Professor	Timothy E. Walsworth, Ph.D. Committee Member
W. Carl Saunders, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah
2023

ABSTRACT

Comparing Commonly Used Aquatic Habitat Modeling Methods for Native Fishes

by

Eryn K. Turney, Master of Science

Utah State University, 2023

Major Professor: Dr. Sarah E. Null
Department: Watershed Sciences

Accurate estimates of species distribution and habitat are critical to effectively incorporate ecological objectives, which protect native fish, into water management. However, no standard methods have been developed to compare predictive accuracy of models developed for different objectives, extents, and with different outputs. In this study, I compared three commonly used aquatic habitat modeling methods which predicted Bonneville Cutthroat Trout and Bluehead Sucker distribution in the Bear River Watershed (UT, ID, WY) at a monthly timestep. Models included an existing hydraulic-habitat model, an existing habitat threshold model, and a geospatial model developed as part of this study. Environmental conditions used as predictors in each model were validated with field observations where applicable. Model predicted suitable habitat for both species was validated using three metrics of predictive performance conducive to presence-only datasets; including proportion of reaches with fish presence observations correctly classified given a binary threshold, weighted proportion of reaches with fish presence observations correctly classified given continuous suitability predictions, and weighted proportion of reaches with fish presence observations correctly classified with an adjustment factor based on model parsimony. For the geospatial model, total upstream catchment area was the most important predictor of both Bonneville Cutthroat Trout and Bluehead Sucker habitat suitability, and land use was a secondary important predictor for Bonneville Cutthroat Trout. Validation of environmental predictors reflected satisfactory to poor fit in all models—no observed conditions were well represented by model estimates—a function of either outdated, incorrect, or over-generalized input data. Validating models with fish presence data showed habitat threshold and geospatial models were similarly accurate for Bonneville Cutthroat Trout, and the habitat threshold model performed best for Bluehead Sucker. However, model performance was sensitive to threshold and performance criteria selection. Habitat predictions from simple, generalizable methods which incorporate biological characteristics of the species of interest are most useful to incorporate native fish conservation into water management models as ecological objectives.

(121 pages)

PUBLIC ABSTRACT

Comparing Commonly Used Aquatic Habitat Modeling Methods for Native Fishes

Eryn Turney

Water resources are managed for a variety of human needs, including agriculture, industrial and municipal consumption, hydropower generation, and recreation. There has been a recent push to incorporate habitat needs of aquatic wildlife into water management models alongside these other uses, particularly as competition for limited water resources in a changing climate has reduced instream flow and contributed to declining native fish populations. Habitat models are used to estimate species distributions and differentiate between suitable and unsuitable habitat based on variables important to a given species, but are not usually incorporated into water management models. Because there are many ways of modeling habitat and no standard way to compare model accuracy, for this research I used three methods of comparing the accuracy of three commonly used habitat modeling approaches to identify best methods for estimating Bonneville Cutthroat Trout and Bluehead Sucker habitat in the Bear River Watershed (UT, ID, WY). I also explored how well variables used in making each model's predictions compared with real-world conditions based on field observations. I determined total upstream catchment area was the most important large-scale variable for predicting both Bonneville Cutthroat Trout and Bluehead Sucker habitat suitability, and nearby land use was also important for Bonneville Cutthroat Trout. I showed none of the models' variables reflected real-world conditions observed in summer 2022, which suggests data commonly used to build habitat models like these can be outdated, incorrect, or over-simplified. Finally, I determined simple habitat models which incorporated aspects of water quality or species biology, rather than simply available water quantity, best predicted both Bonneville Cutthroat Trout and Bluehead Sucker presence, though performance metrics chosen to evaluate model accuracy influenced results. Simpler methods that incorporate species-specific biological criteria are best to include in water management models so fish conservation can be easily and accurately included as a demand for water resources alongside other uses.

ACKNOWLEDGMENTS

I would first like to express much gratitude to the National Science Foundation (Grant #1653452) for funding this work, and my advisor, Sarah Null, for her many hours of patient guidance throughout my experience here at Utah State. Her work is immediately relevant and inspiring in the face of an unknown future, and it has been an undoubtedly life-changing experience to observe her approaches to exploring nuanced and interdisciplinary questions facing watershed scientists in the western U.S. I had a loose tie to watershed sciences before coming to pursue this degree, and I am ever-grateful for the opportunity she provided me to grow as both a scientist and person. Beyond advising, I am also grateful for her as a living example of being a female pushing boundaries in STEM, and her friendship. I hope we get to have a ski-day reunion in the coming years.

Additionally, I am indebted to my committee members, Tim Walsworth and Carl Saunders, for their time and thoughtful insights towards components of this work outside of mine and Sarah's expertise. I thoroughly enjoyed the time the four of us spent talking through implications of this work, and greatly appreciate both of their kind and supportive natures.

I count myself as quite lucky to have been a part of the Aquatic Habitat, Climate, and Water Analysis (ACWA) Lab. Christina, Greg, Ali, Gabi, Saddy, and Ratha: I admire you all as both people and scientists. I am indebted to the support, friendship, and laughter you all have given me, and the example you all provided in how to make it out of this experience with a smile.

I am grateful for many other friends both within and outside of USU who helped me complete field work, as well as friendly residents all around the Bear River Watershed who provided insights, property access, and familiar waves from their yards during the summer of 2022.

Many thanks as well to members of UDWR, IDGF, WYFG, TU, and Brandy Smith for sharing fish presence data with me. This work would have been completely impossible without these datasets!

Finally, I give much love and forever thank yous to my family, and partner. I am beyond fortunate to have family who is interested and supportive of my life and work from 2000 miles away, and has encouraged me in every wild thing I've ever attempted. I will never be able to thank my partner, Dylan, enough for his unrelenting support during my time at USU. I am doubtful I would have made it through this experience healthy or happy without him, which is something I do not take for granted for a second.

Eryn Turney

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Supplementary Figures

INTRODUCTION

Water management and reservoir operations models quantify water that is stored and released to satisfy dynamic human demands. These models have traditionally represented ecosystem needs as minimum instream flows, which are regulatory targets often with little ecological meaning, but increasingly incorporate quantifiable aquatic habitat extent as ecological objectives (Alafifi and Rosenberg, 2020; Jager and Smith, 2008; Null et al., 2021, 2020). For this reason, developing simple and accurate habitat suitability models is an ongoing challenge for water resources management.

Habitat models use physical and biotic predictors which most influence species occupation to estimate habitat suitability in a given area, and these predictions can be used to estimate potential species distribution (Kearney, 2006; Mouton et al., 2007). Hereafter, “habitat models” generally refers to both models that estimate habitat suitability, as well as species distribution models (SDM).

Accurate habitat model estimates of native fish habitat extent throughout watersheds are crucial for conservation managers to identify reaches to preserve, and then prioritize restoration in degraded reaches (Barry et al., 2006; Suding and Hobbs, 2009; Yi et al., 2017). However, there are a variety of methods and predictors used to create aquatic habitat models, which can yield different suitability or distribution estimates over the same area. This begs the question: do differences in accuracy or precision justify the use of particular models or approaches? Specifically, do generalizable methods that are simpler to develop have both sufficient accuracy and precision to be used in place of data-hungry methods? If habitat models created with alternative approaches and/or predictors yield similar predictions, scientists and managers should use those which are least data-intensive and simplest to meet management objectives. However, if more data intensive and complex approaches yield significantly higher predictive accuracy or a heightened level of precision needed for a specific management goal, then those approaches may be justified.

Previous comparisons regarding predictive performance of habitat models focus within certain ‘families’. For instance, there have been many studies comparing predictive capabilities statistical habitat modeling approaches (Barry et al., 2006; Guisan and Zimmermann, 2000; Knudby et al., 2010; Palialexis et al., 2011; Steen et al., 2006; Valavi et al., 2021), but entirely separate comparisons for hydraulic-habitat models developed with preferences functions to quantify usable habitat area (Yi et al., 2014). Though there are review papers which discuss pros and cons of differing model families (Ahmadi-Nedushan et al., 2006; Jowett and Davey, 2007; Yi et al., 2017), comparing predictive performance across families has not been explored. This is partially because there is no standard method of comparing approaches given different assumptions and outputs across model types.

For this study, I compared the predictive accuracies of a hydraulic-habitat model (HYD), a habitat threshold model (THRESH), and a geospatial model (GEO). All are commonly used methods for estimating habitat suitability or extent and have not previously been compared to one another. All models estimated suitable habitat extent for Bonneville Cutthroat Trout (*Oncorhynchus clarkii utah*) and Bluehead Sucker (*Catostomus discobolus*) in the Bear River Watershed (UT, WY, ID, USA). The HYD model was developed by Alafifi and Rosenberg (2020), the THRESH model was developed by Goodrum and Null (2022), and I developed the GEO model. I validated each model's habitat predictions using three metrics of predictive accuracy, including: 1) proportion of occupied reaches correctly classified given a binary threshold, 2) weighted proportion of correctly classified occupied reaches given continuous suitability predictions, and 3) a variant of method 2 with an adjustment factor based on degree of model parsimony. Performance metrics used in model comparison were calculated given each model's respective extent. My primary research questions included:

1. What geospatial variables best predict Bonneville Cutthroat Trout and Bluehead Sucker habitat distribution in the Bear River Watershed?
2. How accurately do the HYD, THRESH, and GEO models predict Bonneville Cutthroat Trout and Bluehead Sucker presence given each model's total respective extent?

1.1 Habitat Modeling Approaches and Literature Review

Most aquatic habitat modeling for fishes has been conducted at the reach scale using species preference functions to estimate suitability (a gradient from unsuitable (0), to suitable(1)) for a given continuous variable; most commonly site-specific streamflow, velocity, depth, or substrate size (Jowett and Davey, 2007; Lamouroux and Capra, 2002; Parasiewicz and Dunbar, 2001; Vadas and Orth, 2001; Vezza et al., 2015; Yi et al., 2017). These methods are generally referred to as hydraulic-habitat models. This modeling method has been increasingly criticized in recent years for inaccurately representing potential suitable habitat due to inaccuracies in either hydraulic data used as predictors or in preference functions (Booker, 2016; Fausch et al., 2002; Steen et al., 2008a). Further, though high-resolution data can be very precise, it is difficult to generalize across large spatial extents, to other regions, or to other species (Boavida et al., 2014; Freeman et al., 1997; Moir et al., 2005).

Models created for larger spatial extents may incorporate biological or water quality predictors into preference functions to delineate suitable versus unsuitable habitats. Commonly used non-hydraulic predictors include stream temperature, dissolved oxygen levels, concentration of Total Suspended Solids (TSS) or nutrients, landform characteristics, stream connectivity, predation risk, and food availability (Boavida et al., 2014, 2013; Brenden et al., 2007; Elith and Leathwick, 2009; Lamouroux et al., 1998; Mouton et al., 2007; Vezza et al., 2015; Wurtsbaugh et al., 2014). These models could also be site-specific, or basin- or state-scale from publicly available sources, leading to more generalized approaches. Again, however, these methods are dependent on knowledge of species biology.

Geospatial models typically use predictive algorithms to estimate species distribution and/or habitat suitability over large spatial extents. These models are often coarser in resolution than site-specific hydraulic-habitat models. Data used as predictors are widely available and often open-source (Steen et al., 2008a) and can be used as proxies when a more specific variable of interest is unavailable (Argent et al., 2003; Boavida et al., 2014; Brenden et al., 2006; Dauwalter et al., 2011; Kristensen et al., 2012; Meixler, 2021). These models have become increasingly popular in recent years to estimate fish and macroinvertebrate habitat suitability (Brotons et al., 2004; Elith et al., 2006; Falke, 2006; Guisan and Zimmermann, 2000; Hirzel et al., 2002; Mugodo et al., 2006; Peterson and Vieglais, 2001; Phillips and Dudık, 2008; Smith and Kraft, 2005; Valavi et al., 2021; Zorn et al., 2002), and are commonly referred to as species distribution models (SDM). Unlike approaches based on species preference functions, SDM methods utilize species presence data and algorithms to determine patterns based on predictor variables to calculate probability of occurrence. SDMs can be constructed with either presence only or presence-absence data (Brotons et al., 2004; Valavanis et al., 2008a; Ward et al., 2009). For example, MaxEnt, Support Vector Machines (SVM), and the Genetic Algorithm for Rule-Set Prediction (GARP) work with presence-only observational data, whereas Random Forest, Artificial Neural Networks (ANN), Boosted Trees, and Classification Trees require use of absence data as well. While these methods are designed to cooperate with collinear variables and work with combinations of continuous and categorical variables (Armstrong et al., 2003; Jorde et al., 2001), they are collectively criticized for their black box methodology (Phillips et al., 2017) and both algorithm selection and species prevalence are known to significantly impact model performance (Benkendorf and Hawkins, In Review; Li and Wang, 2013; Manel et al., 2001; Valavi et al., 2021).

METHODS

2.1 Species of Interest in the Bear River Watershed

Study extent was the Bear River Watershed, which is located in the northeastern corner of the Great Basin and spans 19,425 km^2 through Utah, Idaho, and Wyoming (Fig. 1). The Bear River Watershed is the largest watershed in the U.S. that does not terminate to an ocean ("The Bear River Watershed Information System," 2017; U.S. Fish and Wildlife Service, 2013). Instead, the mainstem Bear River flows to Bear River Migratory Bird Refuge then to Great Salt Lake. The Bear River contributes the most streamflow to Great Salt Lake, with a mean annual flow (MAF) of 1.2 million acre-feet at Corinne, just upstream of Great Salt Lake.

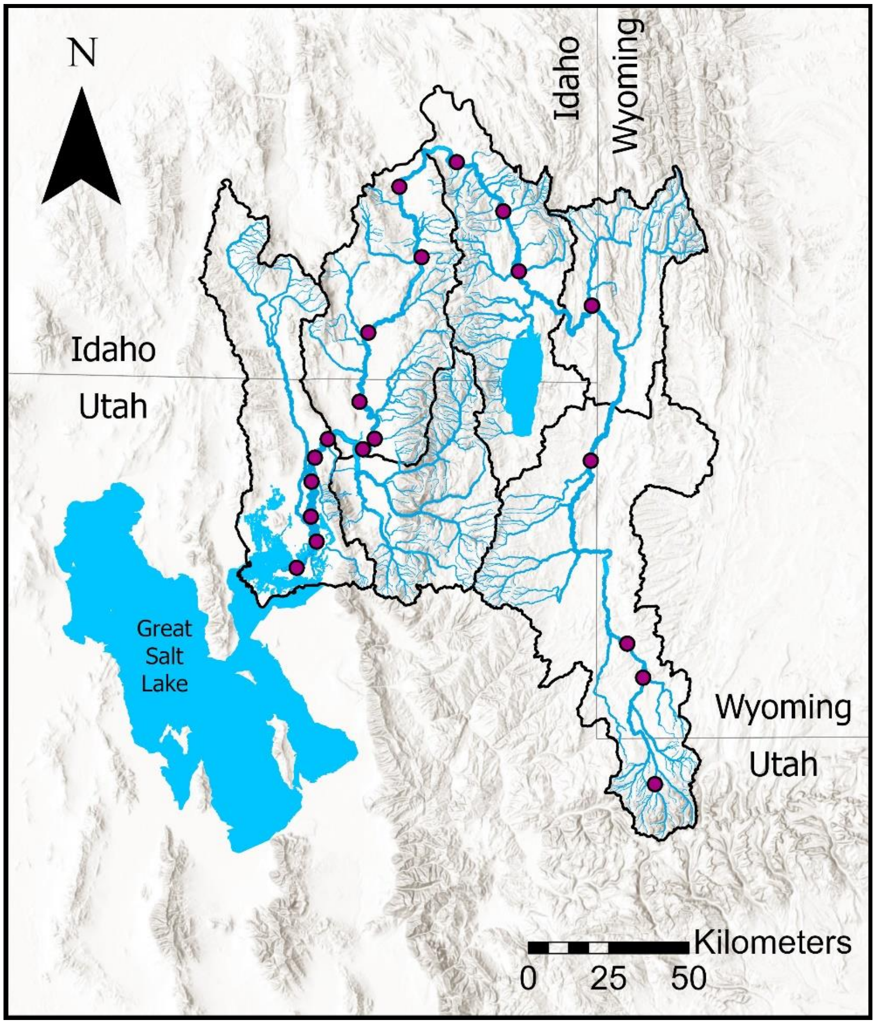

Fig. 1. The Bear River Watershed with summer 2022 field sampling sites.

In many ways, the Bear River Watershed is representative of western U.S. watersheds. Land within the Bear River Watershed has been altered through its multi-use history, with land uses including extensive agriculture, grazing, logging, urban development, and oil and gas exploration ("The Bear River Watershed Information System," 2017; Toth et al., 2009, 2005; U.S. Fish and Wildlife Service, 2013). The flow regime has been manipulated to satisfy needs for both irrigation and hydropower, with a rapidly expanding human population. This ever-expanding development of the basin's land and water resources have resulted in adverse effects on native fish abundance and overall suitable habitat availability within the basin (Toth et al., 2005; U.S. Fish and Wildlife Service, 2013). Competition for a limited water supply for both human and ecological needs is a typical problem in the American West; including the Great Basin, Colorado Basin, and Columbia Basins (Buraeu of Reclamation, 2021; DeRose et al., 2015; Hall, S.A., et al., 2022; Utah Water Science Center, 2018).

The Bear River Watershed is characterized by hot, dry summers and cold winters, with an overall climate classified as semi-arid continental (Hillyard and Keeley, 2012; “The Bear River Watershed Information System,” 2017). Snowmelt is the primary source of water to the Bear River (Goodrum, 2020; Toth et al., 2005). With competing demands for river water, climate change further exacerbates the struggle to maintain adequate instream conditions for aquatic species due to reduced streamflow and warming temperatures (Ficklin et al., 2018; U.S. Fish and Wildlife Service, 2013).

Bonneville Cutthroat Trout (BCT) and Bluehead Sucker (BHS) are two fish species native to the Bear River Watershed and larger Great Basin. Throughout the Great Basin, BHS are estimated to occupy approximately 47% of their historic range, and BCT are estimated to occupy as little as 33% of their historic range (Budy et al., 2015, 2007; Maloney, 2017). While range reductions have not been documented for either species within the Bear River Watershed (Fig. 2), both species have experienced substantial decline in population abundance, attributed at least in part to high sensitivity to anthropogenic disturbance (Budy et al., 2012).

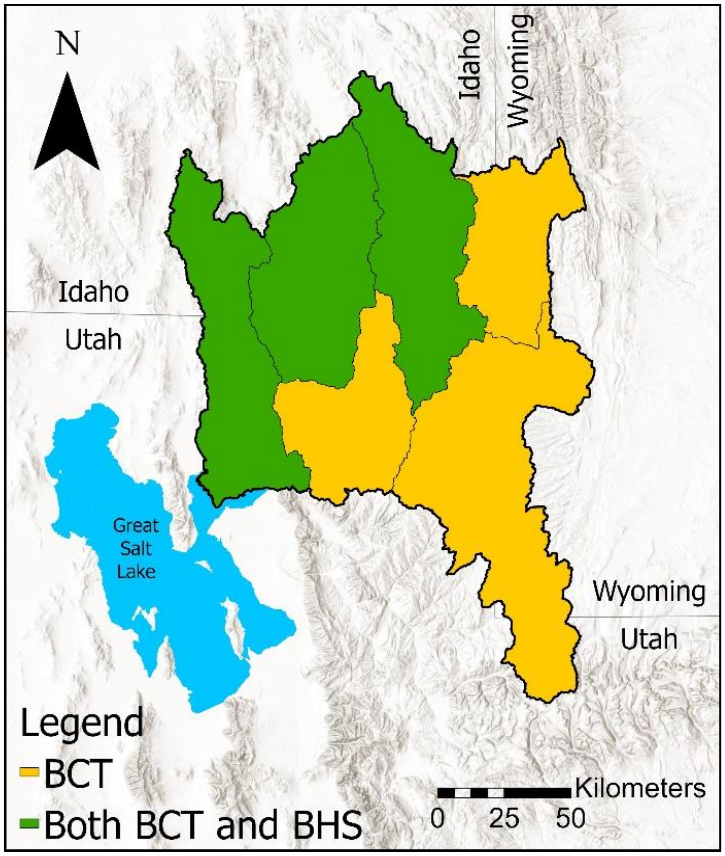

Fig. 2. Historic and current distribution of Bonneville Cutthroat Trout (BCT) and Bluehead Sucker (BHS) in the Bear River Watershed (Western Division of the American Fisheries Society, 2011a, 2011b).

BCT are a charismatic game species and Utah's state fish, leading to considerable advocacy for their conservation within the watershed (Utah Wildlife Action Plan Joint Team, 2015). BHS are a non-game species rarely focused on in media or conservation campaigns, but are considered an indicator species of overall aquatic ecosystem health (Thompson, 2016). Both species are listed as Species of Greatest Conservation Need (SGCN), to preclude listing under the federal Endangered Species Act (ESA), in Utah and Wyoming, though not in Idaho (Idaho Department of Fish and Game, 2016; Utah Wildlife Action Plan Joint Team, 2015; Wyoming Game and Fish Department, 2017). Listing as a SGCN spurs management actions designed to restore critical habitat, which are detailed in State Wildlife Action Plans (SWAP). Inconsistent listing and subsequent management planning across state boundary lines within the watershed makes cohesive habitat conservation initiatives and land/water management difficult.

BCT and BHS prefer pool-riffle habitats with vegetative cover and gravel to cobble size substrate for spawning. Adults of both species use swift-moving riffles for foraging, where larger substrates are more common and support algae growth and invertebrate populations (Budy et al., 2012, p. 2; Ptacek et al., 2005; Walsworth and Budy, 2015; Webber et al., 2012). Adult BCT feed in pool-adjacent riffles where they can conserve energy, and juveniles and fry of both species prefer backwaters refuges as warmer water temperatures aid growth (Lokteff, 2014; Webber et al., 2012; White and Rahel, 2008). BCT, in particular, prefer cold and clean headwater streams, though both species tolerate and have been observed in the turbid, warmer mainstem Bear River Watershed—as long as there is a functional riparian zone to supply adequate structure, cover, shade, and abate warm water temperatures (Budy et al., 2007, 2006; Ptacek et al., 2005).

Stream temperature, specifically, is important for habitat suitability for BCT (Goodrum and Null, 2022). The lethal temperature threshold for BCT is 24.2 degrees Celsius and chronic exposures to greater than 22 degrees Celsius is stressful (Budy et al., 2007; Hillyard and Keeley, 2012; Johnstone, 2000; Johnstone and Rahel, 2003). Though BHS are thought to be more temperature tolerant, the temperature tolerance of BHS native to the Bear River Watershed is unknown because most research has been conducted for Colorado River BHS, which are actively being split into a distinct subspecies (Hopken et al., 2013; Maloney, 2017; Unmack et al., 2014; Webber et al., 2012). There are likely local adaptations to habitat conditions between BHS native to the Colorado River Basin and those native to the Bear River Watershed (Bangs et al., 2020; Webber et al., 2012).

Dams and diversions reduce streamflow which can warm stream temperatures. It is common for managed reaches of the Bear River to approach or exceed the lethal temperature threshold for BCT in hottest months of the year (Hillyard and Keeley, 2012). In-stream barriers also fragment habitats and reduce overall habitat complexity (Kraft et al., 2019). Generally, alteration of the natural flow regime through human development and/or changing environmental conditions creates conditions favorable to non-native species (Budy et al., 2007; Dobos et al., 2016; Dzara et al., 2019; Jager and Smith, 2008; Kraft et al., 2019; U.S. Fish and Wildlife Service, 2013; Walsworth and Budy, 2015; Worthington et al., 2016). While streamflow in the Bear River Watershed was manipulated to satisfy agricultural demands as early as the 1850s (DeRose et al., 2015; Null and Wurtsbaugh, 2020), increasing consumptive water uses and a warming climate have increased the abundance of non-native species such as Brook, Rainbow, and Brown Trout, as well as Largemouth and Smallmouth Bass (Wyoming Game and Fish Department, 2017). While BCT and BHS sometimes co-exist with non-native species, increased abundance of non-natives, particularly Brown Trout, reduces the abundance of native BCT and BHS in the Bear River Watershed (Budy et al., 2008, 2007; Budy and Gaeta, 2017; Dauwalter et al., 2011; Thompson, 2016; Walsworth and Budy, 2015; Webber et al., 2012; Wyoming Game and Fish Department, 2017).

2.2 Modeling

Workflow for this research is shown in Fig. 3, and includes developing the GEO model, validating environmental conditions for all three models, and quantifying model accuracy. Fig. 4 details predictors for each model.

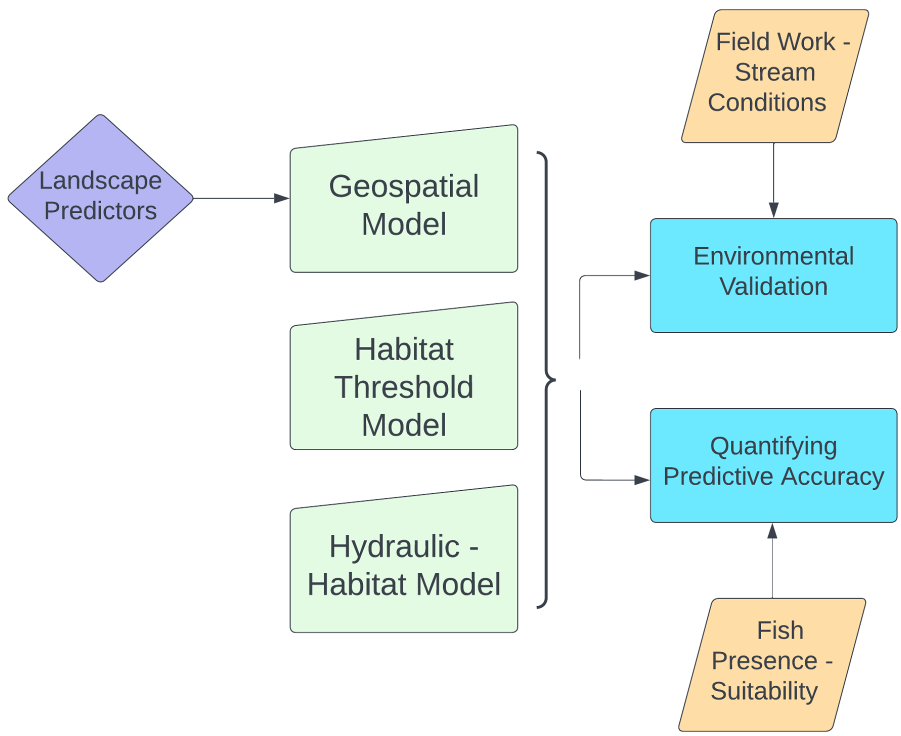

Fig. 3. Research workflow with data used for each step.

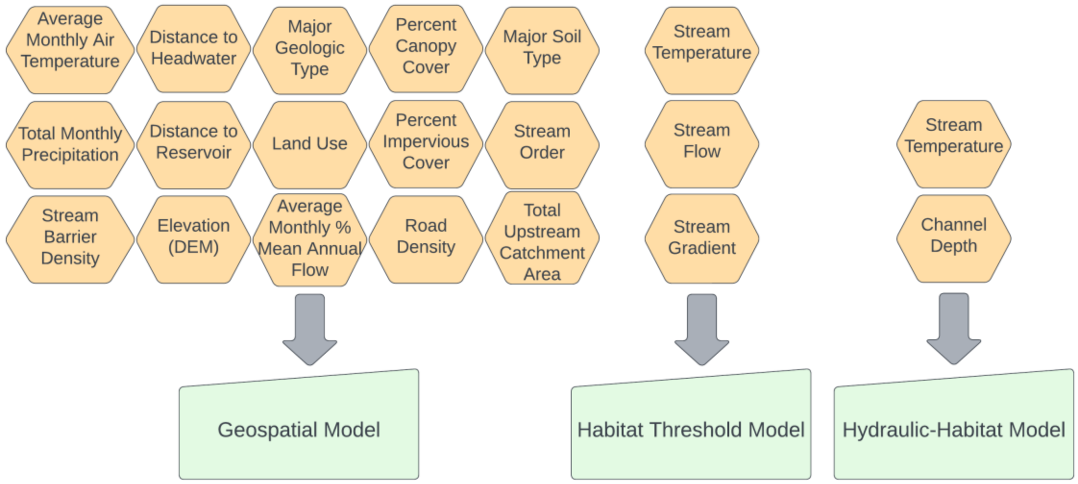

Fig. 4. Predictors for each model.

2.2.1 Existing Models

2.2.1.1 Hydraulic-Habitat (HYD) Model

The one-dimensional HYD model predicted habitat suitability at a monthly timestep for BCT and Brown Trout (*Salmo trutta*), using stream temperature and channel depth based on 2003 streamflow as predictors. The extent of the model was the Lower Bear River Watershed. For this model, habitat suitability was expressed as a continuous value from 0-1, where 0 represented unsuitable conditions, and 1 was perfectly suitable. This value was derived from probabilistic species- and life stage-specific preferences for channel depth constrained by a maximum temperature threshold, and represents the probability of species occurrence based on environmental conditions in a given reach (Alafifi and Rosenberg, 2020). Should the maximum temperature threshold be exceeded in a reach, habitat suitability was assumed to be 0.

A monthly timestep was used because both streamflow and stream temperature fluctuate depending on the time of year, which results in different habitat suitability estimates for the same reaches over the course of a year. This coincides with aquatic species habitat use and preferences changing throughout the year.

Habitat suitability was ultimately used in this model to calculate Weighted Usable Area (WUA) and quantify benefits of environmental water allocations. This model was developed using General Algebraic Modeling System (GAMS) software. In an experimental scenario, BHS habitat preference functions were applied to a reduced extent to compare total WUA to that calculated for Brown Trout (Alafifi, 2018; Alafifi and Rosenberg, 2020), but were not included in the published model. For model comparison, calculations were extended to the model's full extent.

The model's spatial extent was based on observed fish presence locations. It included the mainstem Bear River between Idaho-Utah state line and Great Salt Lake, and portions of major tributaries. The model assumed uniform channel width and depth within reaches. This model represents a much larger spatial extent than most hydraulic models are designed for, and with that has an unusually coarse resolution. Median reach length was 24.3 km (15.1 miles), with a minimum of 4.5 km (2.8 miles), and a maximum of 127.2 km (79 miles), though this was an outlier *(*Fig. 5*)*. The model's habitat quality estimates were validated with expert opinion and fish presence data from Trout Unlimited (TU) and were not made publicly available (Alafifi and Rosenberg, 2020).

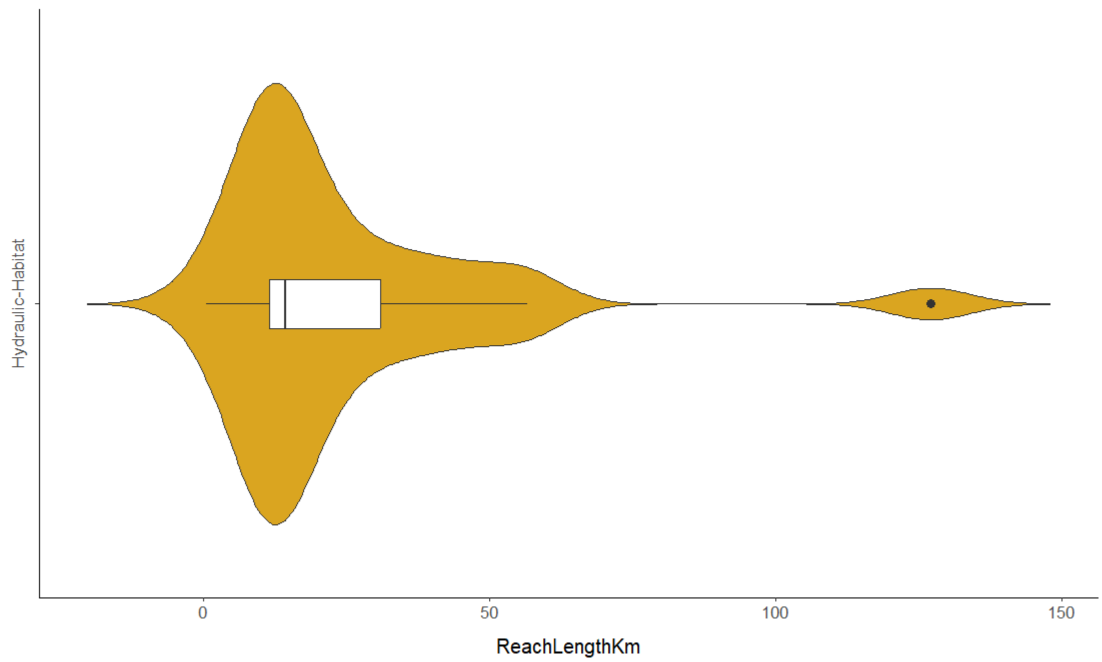

Fig. 5. Modeled reach lengths for the Hydraulic-Habitat (HYD) model, with the mean and outliers shown with the boxplot, and the overall distribution shown with the violin plot.

2.2.1.2 Habitat Threshold (THRESH) Model

The THRESH model predicted habitat suitability for BCT and BHS at a monthly timestep in streams throughout Utah (Goodrum, 2020; Goodrum and Null, 2022). Species-specific thresholds for stream gradient and streamflow were used as predictors in classifying suitable BCT and BHS habitat, and stream temperature was also included as a predictor for BCT models. Again, a monthly timestep was used because both streamflow and stream temperature change over the course of a year, resulting in temporally dynamic habitat suitability estimates. Habitat suitability for this model was expressed as a binary classification for each month, with 0 and 1 reflecting unsuitable and suitable habitat, respectively. In this case, habitat suitability was not a function of species preference, but related to biological limits for modeled species making habitat either lethal or non-lethal. Stream temperature data was from 2000-2018, and streamflow data was based on the National Hydrography Dataset (NHD), from 1971-2000. The model was developed in ArcGIS Pro and R Programming Language to test accurate, simple, and generalizable models which could be paired with water management models.

The model's spatial extent includes all perennial Utah streams, using data from the NHD stream network. Reach divisions were altered to begin and end at points of known and expected instream barriers, resulting in far more, and shorter, reaches than the NHD. Median reach length was 2.2 km (1.4 miles), with a minimum of 0.02 m (0.07 ft), and a maximum of 200.4 km (124.5 miles) (Fig. 6). The majority of modeled reaches were comparable in extent to the median reach length, with reaches less than a meter representing less than 0.2% of modeled reaches, and reaches exceeding approximately 6 km considered outliers. Environmental input data was validated with measured conditions and model performance was validated with fish presence data using Chi-Square Effect Size tests (Goodrum and Null, 2022).

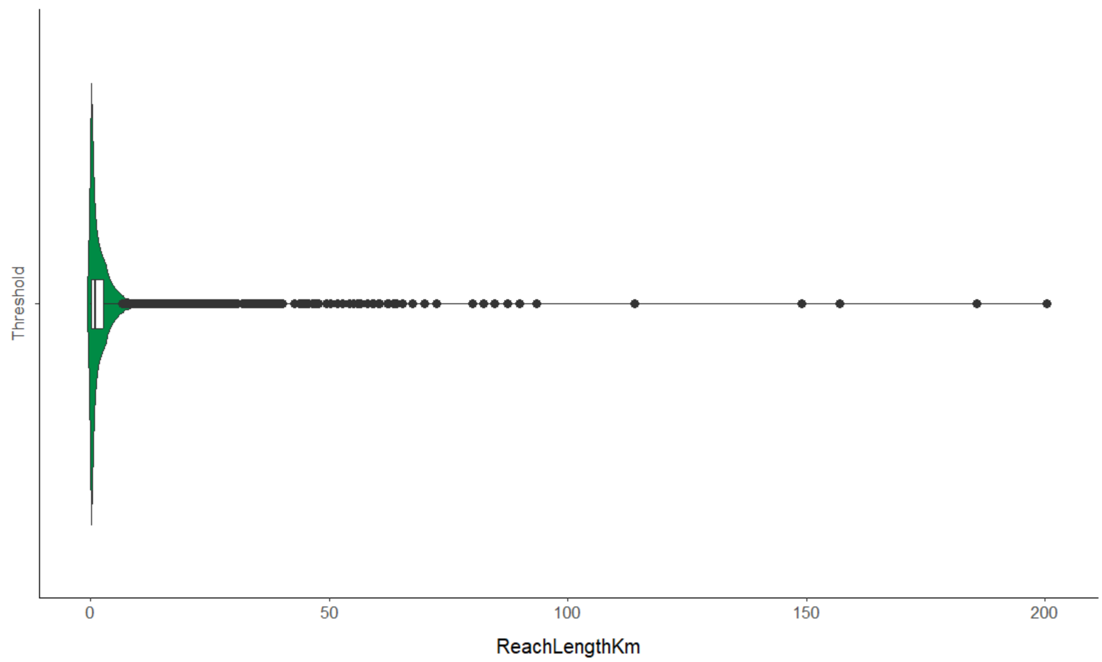

Fig. 6. Modeled reach lengths for the Habitat Threshold (THRESH) model, with the mean and outliers shown with the boxplot, and the overall distribution shown with the violin plot (note that x-axis scale differs from Fig. 5).

2.2.2 Geospatial (GEO) Model Development

I developed the GEO Model to estimate habitat suitability for both BCT and BHS throughout the Bear River Watershed using open-source geospatial data, fish presence data acquired from state agencies, and the maximum entropy (MaxEnt) algorithm. Habitat suitability was expressed as a continuous value from 0-1, where 0 represented unsuitable conditions, and 1 was perfectly suitable; the same as the HYD model. Again, this value referred to the probability of species occupation given the effects of predictors. Month-specific models were developed to match the format of the HYD and THRESH models used in model comparison efforts, as well as to explore how species habitat preferences may change as a result of predictors with monthly or seasonal variability. The GEO model was constructed in R version 4.1.2 using the *dismo* and *SDMtune* packages

(Hijmans et al., 2022; R Core Team, 2021; Vignali et al., 2020a) using a vector-based format.

Model extent was identical to the Bear River Watershed portion of the THRESH model, though reach resolution differed. The GEO model used default reaches determined by NHD, an effect of junction points between stream network lines; which includes intermittent and ephemeral streams. Median reach length for the GEO model was 1.9 km (1.2 miles), with a minimum of 1.3 m (4.3 feet), and a maximum of 26.9 km (16.7 miles) (Fig. 7*)*. Reaches greater than approximately 5 km were considered outliers. Workflow detailing creation of this model is shown in Fig. 8.

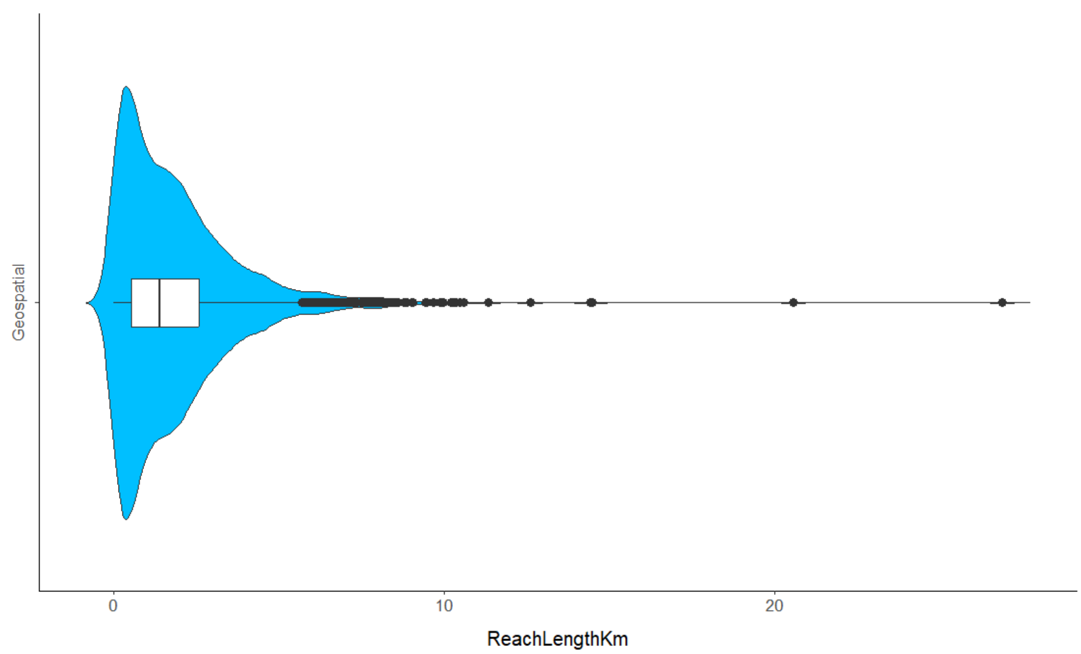

Fig. 7. Modeled reach lengths for the Geospatial (GEO) model, with the mean and outliers shown with the boxplot, and the overall distribution shown with the violin plot (note that x-axis scale differs from Figs. 5 and 6).

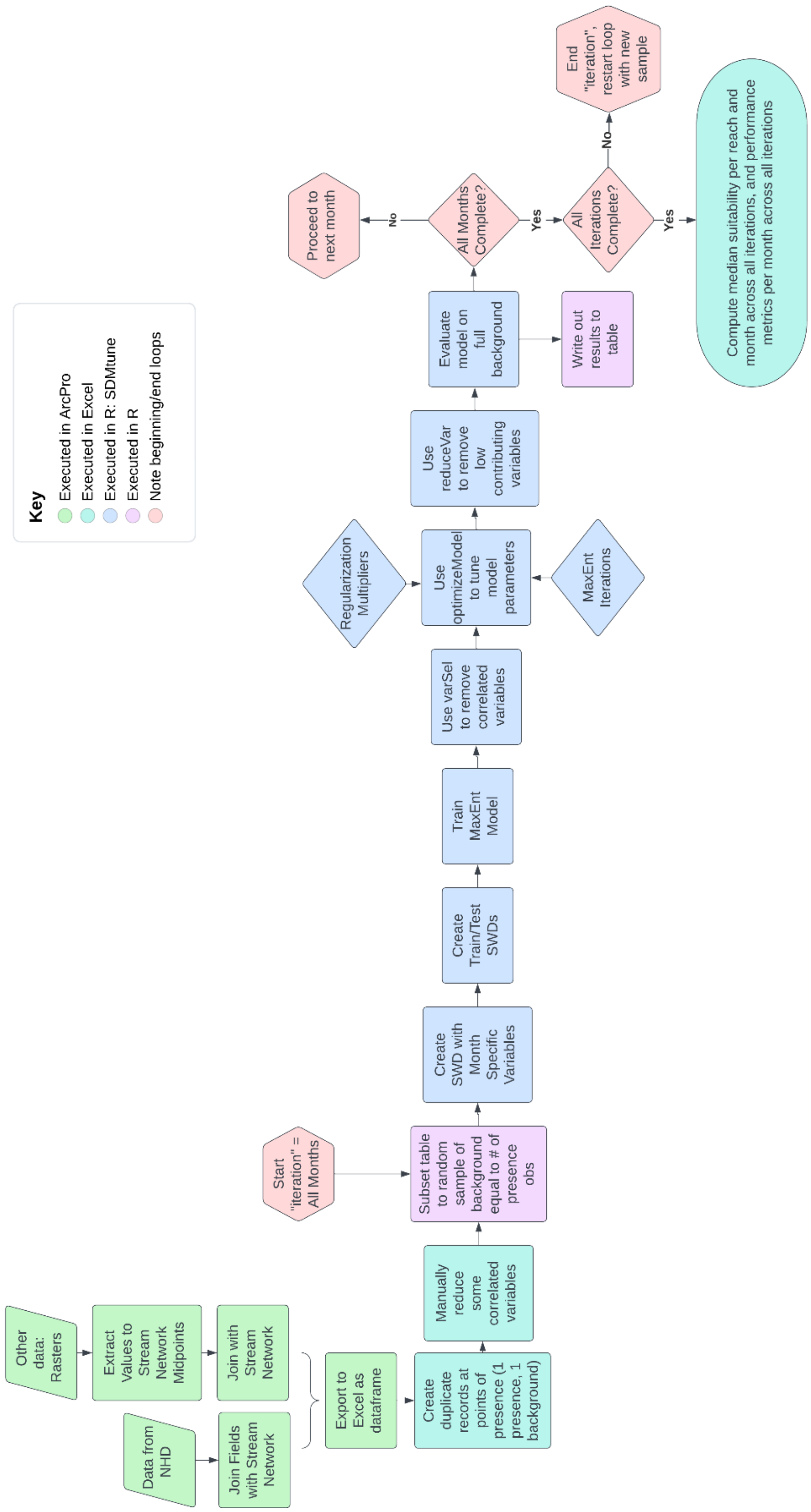

Fig. 8. Geospatial (GEO) model development workflow.

2.2.2.1 Habitat Prediction with MaxEnt

MaxEnt is an open-source, machine learning algorithm and utilizes presence-only observations to predict species distribution by determining the maximum possible extent of suitable habitat given the calculated influence of environmental predictors (Phillips et al., 2006). MaxEnt is best known for predicting terrestrial species habitat suitability and distribution, where raster data are used with species presence points to predict suitable area, but has also been used in fish distribution modeling, where model extent is restricted to a river network using a vector-based approach. Each vector (or row of a data frame), represents a reach that is either occupied or unoccupied by a species of interest (Schmidt et al., 2020; Taylor et al., 2020; Worthington et al., 2016). I used a vector-based approach.

In MaxEnt models, habitat suitability is determined by comparing the range of environmental conditions in species occupied areas to the total range of environmental conditions within a model's extent; hereafter referred to as the background (Elith et al., 2011; Merow et al., 2013; Phillips et al., 2017; Phillips and Dudık, 2008; Phillips and Elith, 2013). A ratio between the range of conditions at points or reaches with species presence observations to the range of those same conditions over entire background is calculated, where environmental conditions in occupied areas become limiting factors for estimating habitat suitability and potential distribution (Elith et al., 2011; Merow et al., 2013). Predictor importance is determined by permuting the tolerable limits of each predictor through iterative model runs, where the larger the range of conditions, the less important the predictor (Phillips, 2017).

MaxEnt runs quickly with a built-in regularization strategy to avoid overfitting, which is needed for correlated ecological variables and small sample sizes (Merow et al., 2013; Valavi et al., 2021; Worthington et al., 2016). Additionally, R packages such as *SDMtune* have been specifically developed to reduce variable correlation and improve parameter tuning when training algorithm-based models like MaxEnt. These packages use iterative model fitting processes to compare the performance of tuned models to an original model. I selected the simplest model without reducing performance (Dorji et al., 2020; Valavi et al., 2021; Vignali et al., 2020b). However, MaxEnt relies heavily on species prevalence—or number of presence observations across a study area the algorithm uses to learn habitat preferences from—and prevalence has a greater effect on SDM performance than algorithm selection or parameter tuning (Benkendorf and Hawkins, In Review). When compared to newer SDM techniques such as Support Vector Machines (SVM) or the Genetic Algorithm for Rule-Set Prediction (GARP), MaxEnt remains among the best performing methods when used in combination with methods to adjust for sampling bias and low prevalence (Barber et al., 2021; Benkendorf and Hawkins, In Review; Fourcade et al., 2014; Kramer-Schadt et al., 2013).

2.2.2.2 Model Workflow

2.2.2.2.1 Data Preparation

Variables considered for reach scale model fitting were either directly available from NHD data tables, created using the *riverDist* package in R, or derived from continuous rasters (Table 1). Variables with cumulative impacts on downstream reaches were summarized for all reaches within HUC-12 basins using the Zonal Statistics tool in

ArcPro to calculate either the mean or mode for continuous or categorical variables, respectfully. I summarized predictors this way as opposed to delineating individual contributing watershed polygons for each reach to reduce computational effort and to focus on local effects. Climatic, hydrologic, and land cover data were the only predictors which had a temporal component. Most of the variables included were common proxies for in-stream conditions; such as air temperature as a stand-in for water temperature, which are highly correlated (Goodrum, 2020; Goodrum and Null, 2022).

Table 1
Geospatial (GEO) model predictor variables, data sources, resolution, intended proxy, and calculation methods.

Variable	Data Source	Raster Data Resolution	Proxy	Method for Calculated Variables
Average monthly air temperature	PRISM 2020	30m	Water temperature	-
Total monthly precipitation	PRISM 2020	30m	Water availability, streamflow	-
Barrier density by HUC-12	Calculated, Goodrum 2020	-	Connectivity	Kernel density of "Instream barriers" layer, search radius based on Silberman's Rule of Thumb 1986; mean over HUC-12
Distance to headwater	Calculated, USGS NHD	-	-	*riverDist* package: Line distance via stream network
Distance to nearest reservoir	Calculated, USGS NHD	-	-	*riverDist* package: Line distance via stream network

Elevation	USGS National Map	10m	-	-
Major geologic type by HUC-12	USGS SGMC	-	Substrate size	Majority class by HUC-12
Land cover by HUC-12	Reclassified, National Land Cover Dataset 2016	30m	Water quality	Reclassified into 6 categories: Open Water/Ice, Developed, Forested, Shrub/Scrub, Agricultural/Rangeland, Wetlands; majority class by HUC-12
Percent Mean Annual Flow (MAF)	Calculated, USGS NHD Stream Network	-	Streamflow, velocity, continuity of groundwater	Monthly percent of annual flow by reach
Percent canopy cover by HUC-12	National Land Cover Dataset 2016	30m	Temperature regulation, LWD input	Mean over HUC-12
Percent impervious cover by HUC-12	National Land Cover Dataset 2016	30m	Human development	Mean over HUC-12
Road density by HUC-12	Calculated, Bureau of Transportation Statistics	10m	Human development	Kernel density of "Detailed Roads" layer, search radius based on Silberman's Rule of Thumb 1986; mean over HUC-12
Soils	Calculated, USEPA Basins	-	Substrate size	Mean over HUC-12
Stream order	USGS National Hydrography Dataset	-	Habitat diversity and stream temperature	-
Total contributing catchment area	USGS National Hydrography Dataset	-	Velocity, width, depth, gradient	-

In order to develop separate monthly models, I created input data for each month by intersecting the 15 predictor variables relevant to each month, if applicable, with my stream network. I also intersected fish presence data for each species with the stream network, to determine if a given reach could be classified as species present. All reaches were viable background in all months, but reaches could be classified as species present if at least one fish presence observation was located within that reach.

Fish presence data included observations acquired over the last 20 years from state wildlife agencies including Utah Division of Wildlife Resources (UDWR), Idaho Department of Fish and Game (IDFG), Wyoming Department of Game and Fish (WDGF), Trout Unlimited, as well as private researchers (Smith, 2022) and are considered representative of both species' current distribution. Fish presence data was not time stamped, so models fitting for each month reflected identical presence reaches, even though it is likely these data were collected between May-October of each year. Here I assumed suitable habitat extent would be far more restricted in sampled months, making these occupied reaches at least viable, if not preferred, over the full year.

In sum, there were 549 of 3055 reaches with at least one BCT observation, and 19 of 3055 reaches with at least one BHS observation in all months of the year. With such a substantial difference in the number of presence reaches between species, it should be noted while there are certainly more sampling efforts targeted at BCT, it is also likely there is extremely low prevalence of BHS in the Bear River Watershed (Bangs and Douglas, 2017; Thompson, Paul, 2015).

In an initial effort to create more parsimonious models, I fit several MaxEnt models for both species using all 15 variables to determine which were least influential on results, as well as used the *plotCor* function available in the *SDMtune* package in R to identify highly correlated variables. With this information, I manually eliminated eight variables from my input dataset (barrier density per HUC-12, distance to headwater, distance to nearest reservoir, major geologic type per HUC-12, percent canopy cover per HUC-12, percent impervious cover per HUC-12, soils, and stream order) which were highly correlated or consistently unimportant in these preliminary model runs. The following seven variables were included in model runs: average monthly air temperature, total monthly precipitation, elevation, land use per HUC-12, percent MAF, road density per HUC-12, and total upstream catchment area.

2.2.2.2.2 Model Fitting

I used multiple iterations of MaxEnt models in an ensemble to best understand the influence of different landscape-scale variables for BCT and BHS distribution in each month. For a single iteration, MaxEnt models were created for all months, which compared environmental conditions at a random sample of the background (static for all months) to conditions in occupied reaches. Conditions at a new random sample of the background were compared to conditions in occupied reaches for each subsequent iteration. Number of iterations was based on the number of background random samples needed to adequately represent the study area, or where I began to observe stability in predictions. I used 10 iterations to estimate suitable habitat for BCT, and 1000 iterations for BHS.

Random samples of the background were created by downsampling based on a BCT density layer, similar to using a bias file in raster-based applications of MaxEnt

(Barber et al., 2021; Fourcade et al., 2014; Kramer-Schadt et al., 2013), to adjust for low prevalence and sampling bias. Downsampling reduced extent of the background by selecting a random subset of background reaches equal the number of reaches with species presence (Benkendorf and Hawkins, In Review). Doing so based on a species density layer forces MaxEnt to draw a greater proportion of these random background samples from areas with more species observations, reducing the inclusion of potentially unsampled reaches in the background sample; where lack of presence may not reflect unsuitable habitat and could skew model results (Elith and Leathwick, 2009; Hirzel et al., 2002; Li et al., 2011; Phillips, 2012; Phillips and Elith, 2013). I used the BCT density layer for BHS as well because there were not enough presence observations to create a density layer for BHS. This assumes BCT and BHS utilize similar habitat in the basin, and/or that sampling efforts for both species would have likely been conducted in similar areas. This coincides with understanding of common limiting factors for both species' distributions in the Bear River Watershed; such as instream barriers fragmenting habitats and excluding potential suitable habitat area from species occupation (Budy et al., 2007; Kraft et al., 2019; Walsworth and Budy, 2015).

With a random sample selected, I fit an initial model for each month using a training partition representing 75% of available presence and background reaches, where the remaining 25% of reaches were withheld as a test partition to be used in model tuning steps. It is important to note creating training/test partitions and cross-validating for model fitting and evaluation are the most common approaches for creating and validating geospatial models (Elith et al., 2011; Merow et al., 2013; Vignali et al., 2020b). Cross-validation methods are heavily utilized when species occurrence data is limited, but has received criticism for overfitting and/or inflating indicators of model performance (Olden et al., 2002). Ultimately I elected to use training and test partitions because with limited species presence data and given subsequent steps of this research, I wanted to limit bias from validating the model with the very data used to create it.

Tuning steps removed correlated variables (using *varSel* within the *SDMtune* package), optimized two MaxEnt parameters (the regularization multiplier, or “smoothing” effect, and the number of within model iterations used in fitting; using *optimizeModel* within the *SDMtune* package), and reduced variables with low contribution effects (using *reduceVar* within the *SDMtune* package) by using optimization algorithms to iteratively compare performance of a tuned model to performance of the initial model (Vignali et al., 2020b). The initial model was replaced if performance improved.

When model tuning was completed, I used the function *evaluate* within the *SDMtune* package to apply the final tuned model to the full dataset, and determine suitability predictions expressed as a continuous value between 0 (unsuitable) and 1 (suitable), performance metrics, MaxEnt suggested thresholds for calculating binary suitability predictions, and model parameter values. Performance metrics included Area Under the Receiver Operating Characteristic Curve (AUC), True Skill Statistic (TSS), and Symmetric Extremal Dependence Index (SEDI). AUC and TSS can be poor performance indicators for presence-only models, especially those with extremely low species prevalence (Ferro and Stephenson, 2011; Wunderlich et al., 2019). I included them because they have traditionally been standards of SDM or GEO model performance, and are generally still reported (Allouche et al., 2006; Leroy et al., 2018;

Lobo et al., 2008). SEDI is useful for presence-background models with low prevalence (Ferro and Stephenson, 2011; Wunderlich et al., 2019).

2.2.2.2.3 Ensemble Summary

Median monthly suitability, performance metrics, thresholds, and parameter values were calculated across all MaxEnt iterations. I used median suitability estimates across all iterations for comparisons with the HYD and THRESH models in validation steps.

As predictor variables and their importance could vary among iterations, variable importance is reported as a Weighted Permutation Importance (WPI) by season. To calculate this, I first counted the frequency of each variable's inclusion in each month's final model across all iterations, then summarized this count by season: spring, summer, fall, and winter. Next, I averaged the permutation importance of each variable included in a month's final model across all iterations by season. I multiplied this average seasonal permutation importance by the seasonal inclusion count to determine seasonal WPI. Using an ensemble approach, I was able to account for more of the background than if I only used one downsampled partition, and effects of low prevalence and sampling bias were reduced in individual model runs. By summarizing across many iterations and unique presence/background partitions, I reduced the likelihood of overfitting models based on overfit outliers (and conversely, particularly poorly fit outliers).

2.3 Model Validation

For this research, I validated both environmental predictors used in models, as well as predictive accuracy of the models. Both types of validation are uncommon (Elith and Leathwick, 2009; Vezza et al., 2015), but important as validation ensures reliability of predictions and provides a basis for recommending model applicability to other locations and/or species (Olden et al., 2002; Valavanis et al., 2008b).

2.3.1 Environmental Data Collection

Again, though fish presence data were not time-stamped, they are considered representative of both species' current distribution by state agencies and Trout Unlimited. Environmental predictors of all models were validated with observed conditions from 2022 in an effort to understand how accurately models represented current habitat conditions, and could from there reliably predict habitat quality and species distribution potential.

I sampled stream depth, velocity, and temperature in the mainstem Bear River between June-October 2022 to validate environmental variables in the HYD model (stream temperature, stream depth, and streamflow), two of three variables in the THRESH model (stream temperature and streamflow), and percent mean annual flow as a function of streamflow in the GEO model. I used a HACH FH950 flowmeter, a Teledyne RiverPro Acoustic Doppler Current Profiler (ADCP), and Onset HOBO water temperature Pro v2 and water level data loggers to take measurements at 20 sites where there was public access or an adjacent road (Fig. 1).

Sampling efforts were focused on the mainstem Bear River where models overlapped, and which makes up most of the modeled extent of the HYD model.

Hydraulic models are infrequently field-validated because data collection efforts are time-consuming and expensive (Barker et al., 2018; Brenden et al., 2007; Rosenfeld et al., 2011; Williams et al., 2013). Though the HYD model was previously calibrated with field-collected data from three sites (Alafifi and Rosenberg, 2020), I hoped to further validation efforts by determining how comparable stream conditions in 2022 were to estimates based on streamflows in 2003. Environmental conditions in the THRESH model were previously extensively validated with environmental data (Goodrum and Null, 2022). Because both the THRESH and GEO model use the NHD to estimate percent MAF, I could have used results of Goodrum & Null's (2022) validation efforts to represent quality of fit for this variable in GEO model. However, because the THRESH model included streams throughout all of Utah, I wanted to recalculate fit metrics specifically within the Bear River Watershed, the extent of the GEO model, where I expected results to be different than the trend based on all perennial streams across Utah.

2.3.2 Environmental Data Validation

I used the coefficient of determination (R2), Nash–Sutcliffe efficiency (NSE) index, and percent bias (PBIAS) to compare observed conditions at point locations from 2022 to model estimates for each variable (D. N. Moriasi et al., 2015, 2007; Goodrum and Null, 2022). Reach lengths varied between model types, so point estimates applied to the reach the point fell within. The HYD model had long reaches, so this validation assumed site-scale conditions were uniform throughout reaches.

Though the THRESH an GEO model estimates were based on monthly averages from the NHD (1971-2000), it is useful to show how well 1971-2000 NHD estimates fit current streamflow conditions. HYD model estimates were based on streamflow observed in 2003, so I ranked mean annual flows (MAF) for the full period of record at USGS Station 10126000 (Bear River near Corinne, UT) to determine if streamflows were similar enough in 2003 and 2022 to be compared (Oregon State University, 2002). Corinne is the farthest downstream stream gage in the Bear River Watershed, and is proximal to Cutler Dam and Reservoir, a small, but major reservoir in the basin. There is limited water storage in Cutler Reservoir, equivalent to 8,563 acre-feet (Olson, 2022). This makes up less than 1% of the total MAF for the Bear River Watershed. Streamflow measurements at Corinne do neglect the effects of some additional diversions and tributary inputs downstream. However, these effects scale accordingly across wet and dry years, allowing streamflow here to stand as a proxy for flows for the full basin.

Across the full streamflow record at Corinne (1950-2022 with 5 years missing between 1958-1963; a total of 67 years), 2003 was a dry water year, equivalent to ~33.0 % MAF. 2022 was also dry, but slightly less so, and MAF was ~40.7% of average. 2003 was the second driest year on record and 2022 was seventh driest (U.S. Geological Survey, 2016). I also analyzed years from 2000-present, a period where the western USA has experienced a prolonged drought (Williams et al., 2020). Within this period, 2003 streamflow was ~47.5% of average and was the driest year on record. 2022 was the fifth driest. While 2003 and 2022 were close enough to be comparable, validation results were expected to reflect 2022 streamflow and depth as slightly greater than modeled estimates from 2003.

2.3.3 Quantifying Model Accuracy

There were no commonly used metrics of predictive accuracy applicable for all models compared here. The three models used in this comparison were created for different purposes with distinct extents and reach lengths, and outputs in different forms. Fish absence data were unavailable, making quantifying model predictive accuracy in reaches without presence observations impossible (for models both incorrectly predicting unoccupied reaches as suitable habitat, as well as correctly predicting unoccupied reaches as unsuitable habitat). Sparse presence observations (particularly for the HYD model extent), eliminated traditional statistical methods such as Generalized Linear Models, Linear Mixed Effects Models, and Generalized Linear Mixed Effects Models to determine differences in presence-only model predictive accuracy.

My validation efforts were restricted to those reaches with presence observations, in which I could correctly classify occupied reaches that predicted suitable habitat, and incorrectly classify occupied reaches that predicted unsuitable habitat. This correct/incorrect classification method worked regardless of whether model output was a probability of occurrence or an expected lethal/non-lethal classification. I developed three metrics to compare model predictive accuracy given differences in model extent, threshold selection, and type of model output in occupied reaches. Extents of models used in this comparison with fish presence data are shown in Fig. 9. Table 2 reports counts of reaches with fish presence observations for model type.

Table 2
Reaches with fish presence observations per model extent.

Model	Total Reaches	Reaches with BCT Observations	Reaches with BHS Observations
HYD	24	9	4

GEO	3055	549	19
THRESH	31147	1321	73

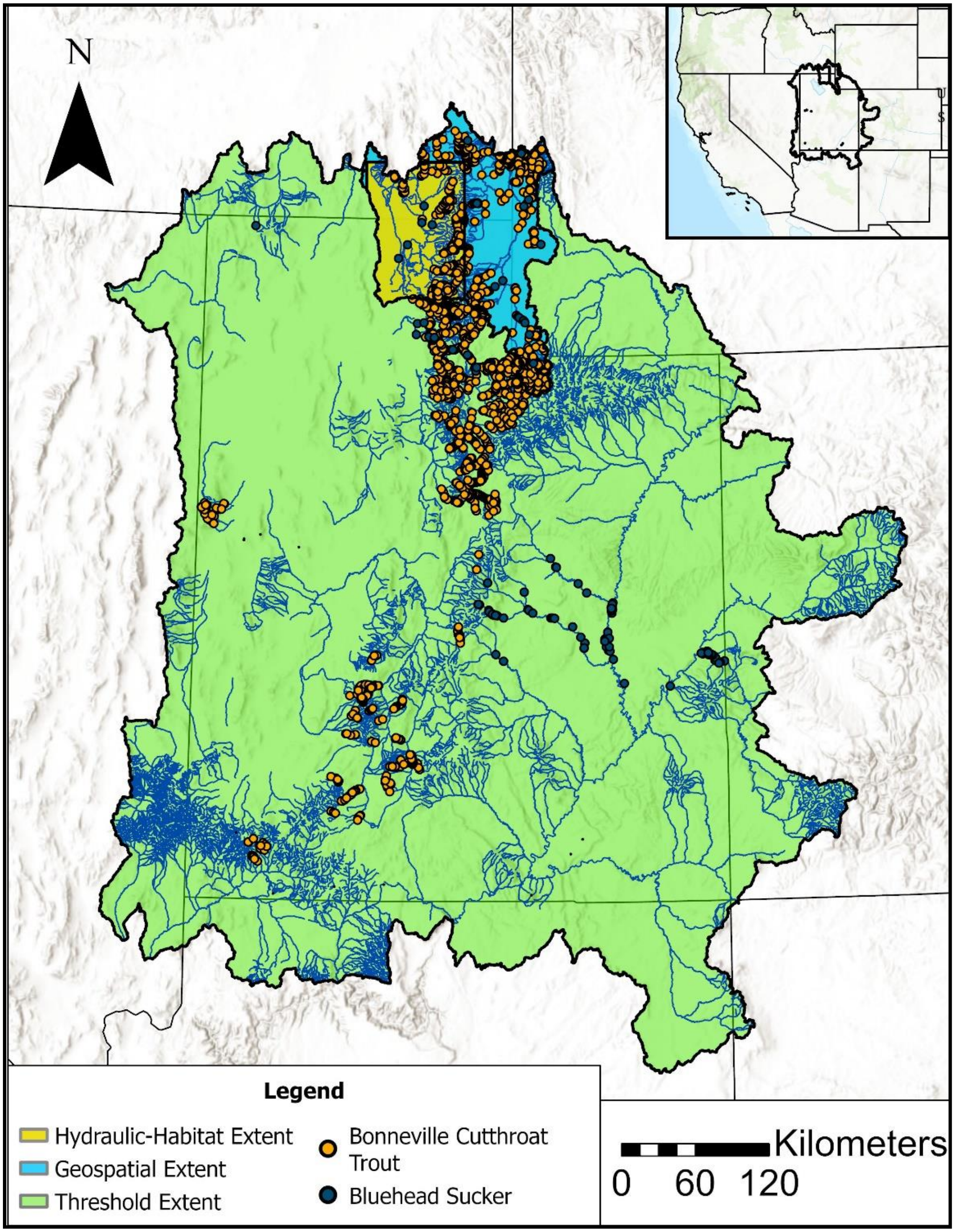

Fig. 9. Fish presence data with the extent of each model.

2.3.3.1 Method 1: Threshold-based Proportion of Correctly Classified Reaches (TPC)

In the most intuitive method of determining predictive accuracy, I reported model performance as the proportion of reaches with presence observations correctly classified as suitable habitat given a threshold across each model's full extent (TPC). A single fish presence observation in a reach predicted suitable habitat yields a correct classification, and a single fish presence observation in a reach predicted unsuitable habitat yields an incorrect classification. Incorrect classifications suggest models underpredict suitable habitat. TPC is represented by the following equation, where *i* represents a reach with a confirmed species presence, and *n* represents the total count of reaches with observed species presence:

$$TPC = \frac{\sum_i^n Binary\ Suitability\ Prediction_i}{n}$$

For the GEO and HYD models, a threshold was needed to reclassify continuous habitat suitability to binary suitability. For both the HYD and GEO models, I included a threshold of 0.5 to differentiate suitable versus unsuitable reaches, an ecologically intuitive boundary. For the GEO model, I calculated TPC at two additional thresholds commonly used in conservation applications of MaxEnt; the Minimum Training Presence (MTP) and 10th Percentile Minimum Training Presence (10PMTP) thresholds (Dorji et al., 2020; Liu et al., 2016; Radosavljevic and Anderson, 2014; Shcheglovitova and Anderson, 2013). These thresholds represent the lowest predicted suitability value among all occupied reaches, and the 10th percentile of suitability values among all occupied reaches, respectfully. MTP assumes habitat must be suitable wherever an individual is observed. 10PMTP recognizes species may occasionally be found temporarily inhabiting lesser quality areas, although these reaches may not be suitable for permanent occupation (Cecina Babich Morrow, 2019). To select another threshold for the HYD model, I conducted a sensitivity analysis of the suitability threshold to illustrate how it influenced habitat suitability. In the sensitivity analysis, the greatest drop in TPC at a step of 0.01 from 0-1 was 0.01 for both species, and this was used as the second threshold.

2.3.3.2 Method 2: Weighted Proportion of Correctly Classified Reaches (WPC)

For a second metric, I aimed to avoid choosing a threshold, so I calculated the proportion of the sum of continuous suitability predictions to total reaches with observed fish presences over each model's full extent, resulting in a weighted proportion of correctly classified reaches (WPC). WPC is represented by the following equation, where *i* represents a reach with a confirmed species presence, and *n* represents the total count of reaches with observed species presences:

$$WPC = \frac{\sum_{i}^{n} Continuous\ Suitability\ Prediction_i}{n}$$

By nature, this method almost always resulted in lower predictive accuracy estimates for the HYD and GEO models compared to threshold-based approaches like TPC unless a chosen threshold resulted in incorrect classifications. For example, should the HYD or GEO model have three reaches with species presence observations, each with habitat suitability predictions of 0.75, then the WPC suitability values would be 0.75, or 75%. In calculating TPC, if the threshold chosen was less than or equal to 0.75, then each of these reaches would have been classified fully correct, leading to a TPC of 1, or 100%. However, if the threshold used to determine suitable or unsuitable habitat had been greater than 0.75, each of these reaches would have been misclassified, leading to a TPC of 0, or 0%. Because THRESH model outputs were binary, WPC was identical to TPC from method 1 for the THRESH model.

2.3.3.3 Method 3: Adjusted Weighted Proportion of Correctly Classified Reaches (WPC_{Adj})

Since the THRESH model can potentially classify all reaches with a suitability of value of 1 and thus principally outcompete the other two models using the WPC method, I adjusted the WPC calculation by the proportion of a weighted unsuitable reach length over total reach length. This method favors predicting suitable habitat in reaches at and around occupied reaches, and not extraneously, and with that provides an indicator of both model precision and accuracy. This metric is unit-less and continuous between 0 and 1, with higher values indicating better model performance. This calculation is below, where *i* represents a given reach, and *n* represents the total number of reaches within a model's extent:

$$WPC_{Adj} = WPC * (1 - \frac{\sum_i^n Continuous\ Suitability\ Prediction_i * Reach\ Length_i}{\sum_i^n Reach\ Length_i})$$

If a model predicted all suitable habitat (suitability value of 1) in all reaches, then WPC would equal 1 as all presence observations would occur in suitable habitat. However, with zero reaches predicted unsuitable, the adjustment factor would equal 0, resulting in a WPC_{Adj} of 0. Should models with a WPC of 0.75, for example, reflect a gradient of suitability predictions across a modeled extent of 500 km equal to 100 km of suitable habitat, this would adjust WPC by a factor of 0.8, resulting in a WPC_{Adj} of 0.6. Should suitable habitat extent be even less, for example 50 km, a WPC of 0.75 would be adjusted by a factor of 0.9, resulting in a WPC_{Adj} of 0.675.

As predicted suitable habitat extent declined, the adjustment factor approached 1, resulting in a WPC_{Adj} closer in value to the non-adjusted WPC. However, models with a low WPC maintain low performance even at extremely low proportions of predicted suitable habitat. For example, should a model reflect a WPC of 0.1 but have an estimated suitable extent of only 1 km of the total 500 km, the adjustment factor would equal .998, resulting in a WPC_{Adj} of 0.0998. So, in combination a high WPC, WPC_{Adj} can approach its maximum of 1. However, this is near theoretically impossible as all reaches with presence observations would need to be correctly classified, with almost zero other areas predicted suitable (Fig. 10).

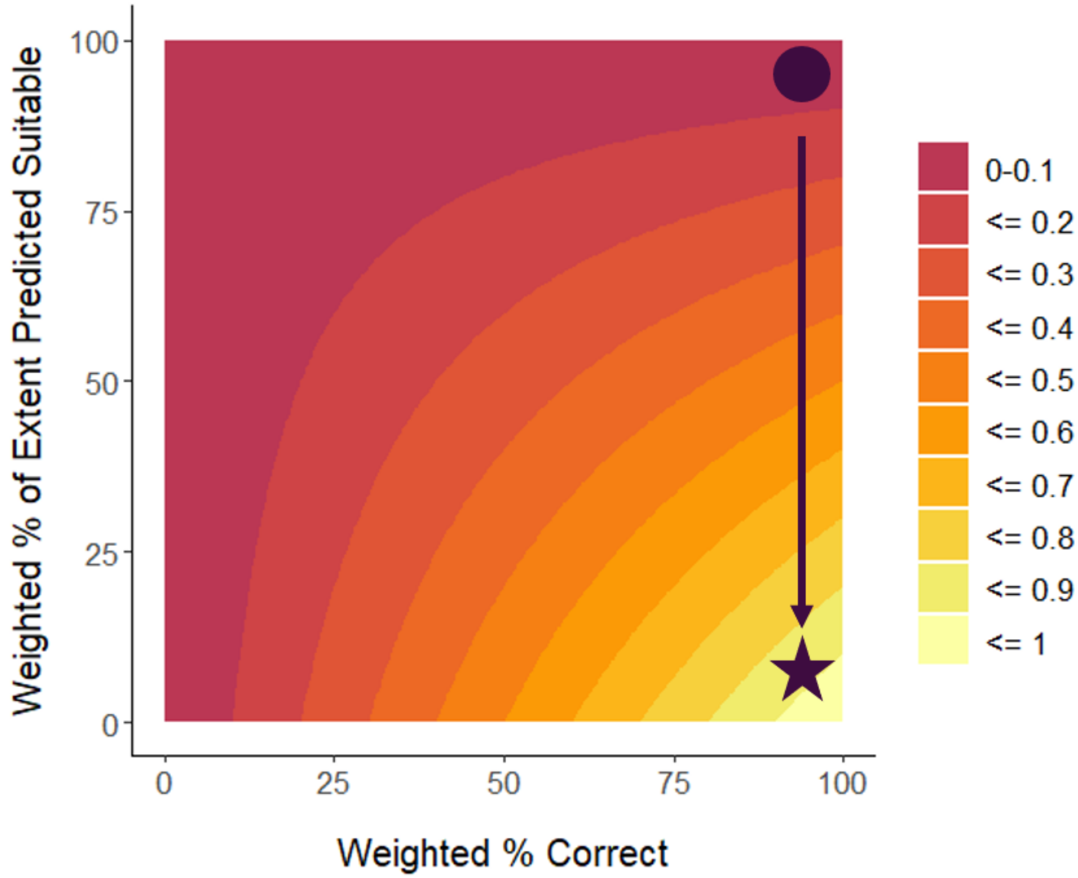

Fig. 10. Models with a high Weighted Proportion of Correctly Classified Reaches (WPC) but large total predicted suitable habitat extent (dot symbol) approach the optimal Adjusted WPC (WPC_{Adj}) as they predict less suitable area (star symbol).

RESULTS

3.1 Geospatial (GEO) Model Results

3.1.1 Variable Importance

Of the seven variables included in model runs for BCT, between 2-5 variables were retained in the final model fits to predict species distribution depending on the month and iteration. Total upstream catchment area was most frequently selected as an important variable across all months, and consistently had the highest Weighted Permutation Importance (WPI) of ~50% across all seasons (Fig. 11). Total upstream catchment area had a negative overall relationship with habitat suitability (Fig. 12); where lower upstream catchment areas were associated with higher habitat suitability values. Land use was the second most important variable across all seasons (Fig. 11), with BCT more often observed in forest-dominated HUC-12 basins (Fig. 13). In spring, average monthly temperature, and in fall, precipitation closely followed land use in WPI (Fig. 11). Summer months reflected greatest spread in variable importance, with proportion of WPI nearly equal among land use, road density, average monthly temperature, and precipitation (Fig. 11). Elevation was consistently found least influential throughout the year, and was not correlated with total upstream catchment area (Fig. 11). This is because in the Bear River Watershed, there are tributary reaches with low total upstream catchment area located at variable elevations all over the watershed. The overall relationships between the less important variables used in model fitting and habitat suitability are included in the Appendix (Figs. A1-A5).

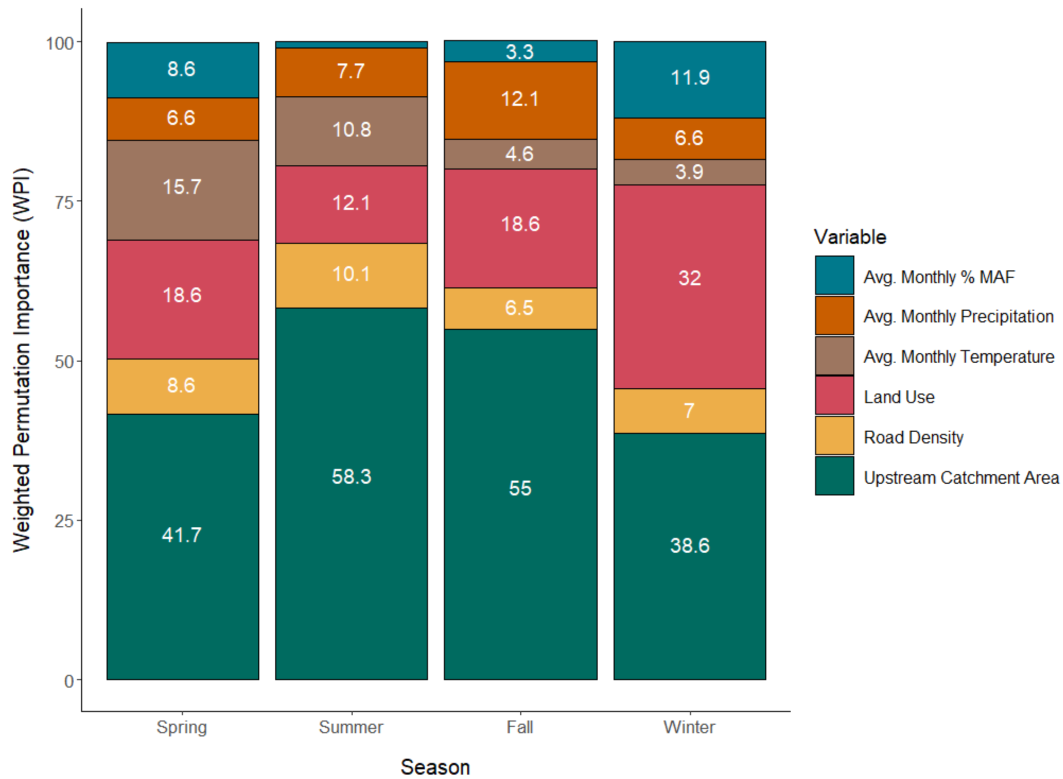

Fig. 11. Weighted Permutation Importance (WPI) of variables by season for Bonneville Cutthroat Trout (BCT) model runs.

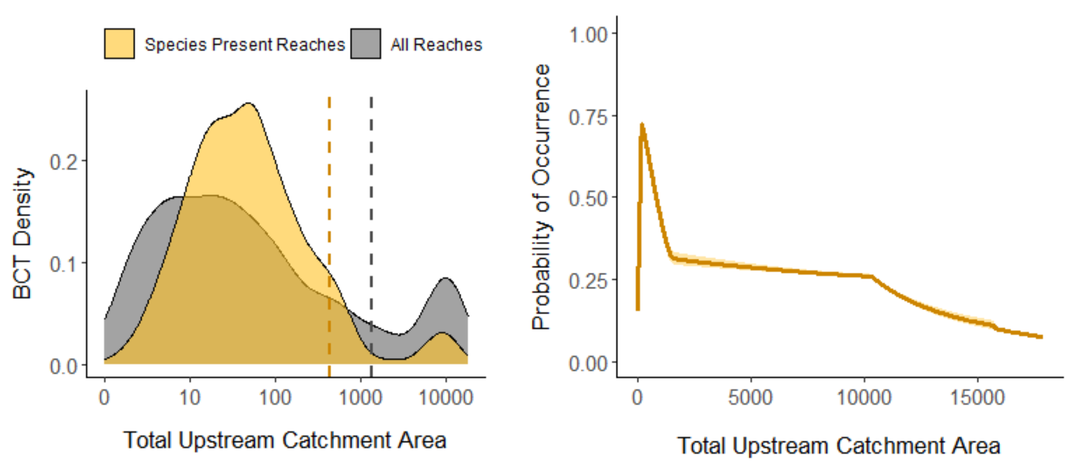

Fig. 12. Total upstream catchment area for Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different upstream areas across all months and ensemble iterations, with a 95% confidence interval (right).

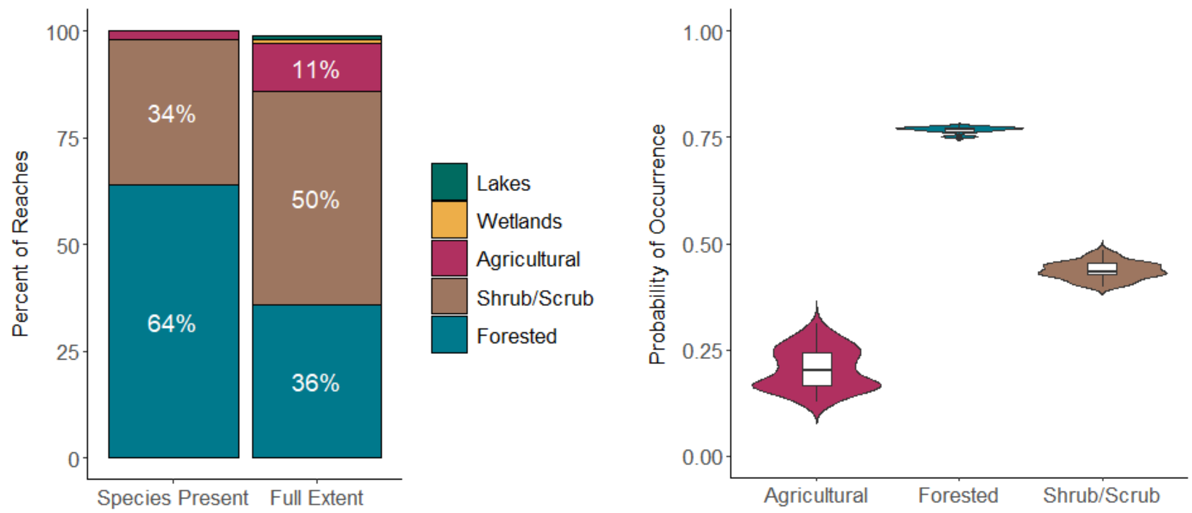

Fig. 13. Land use in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches (left) and MaxEnt probabilities of occurrence by land use category for all months and ensemble iterations.

Out of 1000 ensemble iterations for BHS, between 1-6 variables were retained in the final model fits to predict species distribution depending on the month and iteration. With only 19 species presence observations, ideally individual models would only be fit with 1-2 predictors (Peduzzi et al., 1996, 1995; Vittinghoff and McCulloch, 2007). However, it is impossible to ensure MaxEnt only uses a specific number of predictors in the final fit. With this, there were less obvious trends in selected important variables across iterations than in BCT model runs; with the exception of total upstream catchment area as consistently the most important predictor across all seasons (average WPI of about 25% for the full year) (Fig. 14). For BHS, total upstream catchment area had an overall horizontal relationship with habitat suitability (Fig. 15); where BHS were found to generally prefer reaches with higher total upstream catchment area than the study area mean. Beyond this, there was no clear second-best predictor across all seasons, with relatively even WPI among the remaining 6 variables (Fig. 14). These less definitive results are likely due to individual models being overfit to small datasets. The overall relationships between the less important variables used in model fitting and habitat suitability are included in the Appendix (Figs. A6-A11).

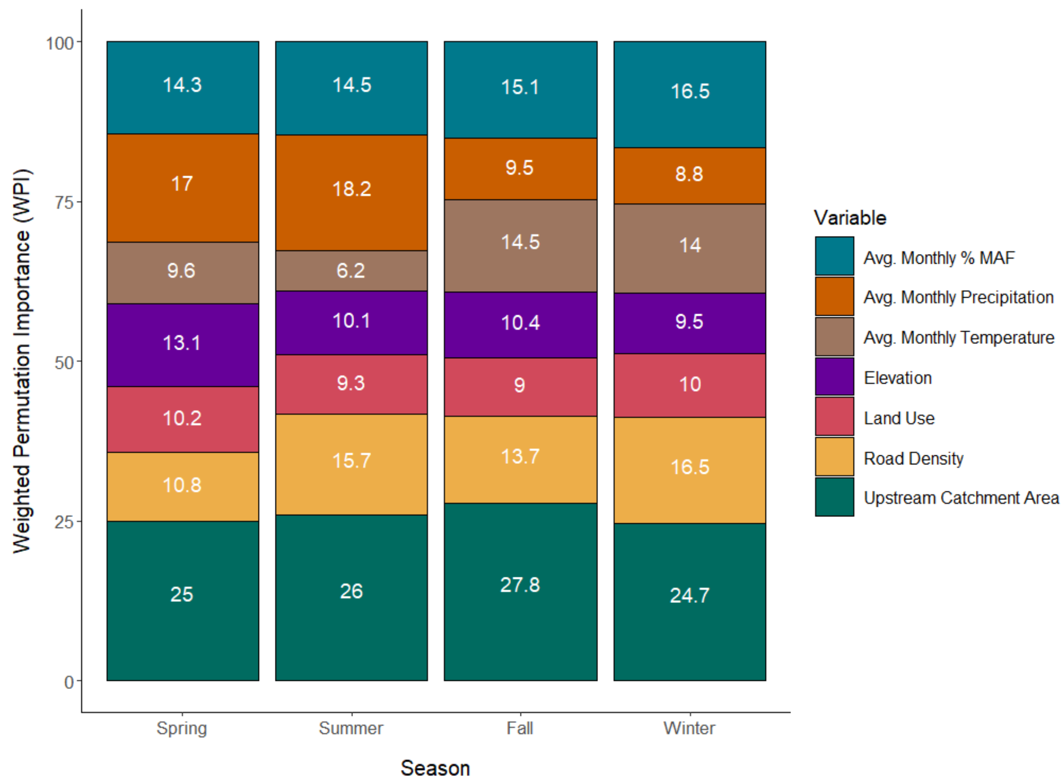

Fig. 14. Weighted Permutation Importance (WPI) of variables by season for Bluehead Sucker (BHS) model runs.

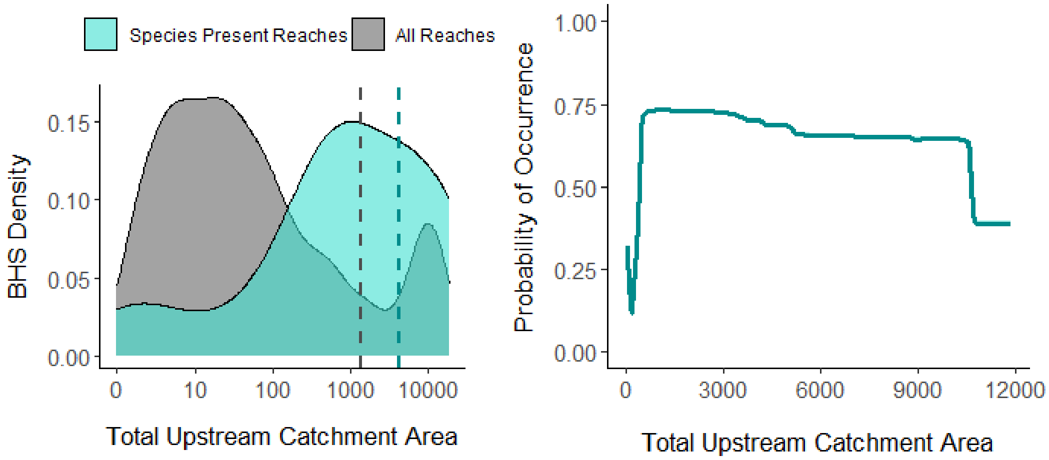

Fig. 15. Total upstream catchment area for Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different upstream areas across all months and ensemble iterations, with a 95% confidence interval (right).

3.1.2 Performance Criteria, Parameters, and Thresholds

Median monthly GEO model performance metrics, threshold values, and optimized MaxEnt parameters for BCT and BHS model runs are reported in Table 3 and 4; respectively columns 2-8, 9-10, and 11-12. Per Komac et al (2015), models with an AUC of 0.8 or greater, and a TSS of 0.6 or greater, reflect good to excellent model fit. SEDI ranges from -1 to 1, with 1 indicating perfect model fit, and 0 no better than random. The default regularization multiplier is 1, and the default iterations used to fit a given training dataset is 500 (Vignali et al., 2020b).

In BCT training and testing models, model optimizers increased both the regularization multiplier and iterations per fit from defaults. No months were classified as good to excellent fit when evaluated by AUC or TSS, though some were close. However, AUC and TSS performance values were consistent between training and test partitions.

When evaluated by SEDI, a threshold of 0.5 to differentiate suitable versus unsuitable habitat performed far better than lower thresholds (MTP and 10PMTP), though overall, model fit was mediocre. There were a few instances where SEDI could not be computed, which occurs when any of the following occurs: 1) there are 0 presence observations found in predicted suitable habitat, 2) there are 0 presence observations are found in predicted unsuitable habitat, 3) there are 0 instances of a lack of presence observations in predicted suitable habitat, 4) there are 0 instances of a lack of presence observations in predicted unsuitable habitat (Ferro and Stephenson, 2011). Further, high counts of predicted suitable reaches without presence observations can result in a very low SEDI. Thus, it is not surprising for a higher threshold (more exclusive) to lead to a higher SEDI, where there is less predicted suitable habitat overall.

Trained BHS models reflected excellent model fit when evaluated by AUC or TSS (Komac et al., 2016). Model performance declined considerably in all months, however, when models were evaluated using test partitions. SEDI model fit was very good regardless of threshold, though it could not be computed for any threshold in June. There were generally major differences between MTP and 10PMTP thresholds, with the MTP threshold consistently equal to 0 and the 10PMTP threshold usually equal to 0.63 . However, MTP and 10PMTP thresholds were occasionally close in value. With only 19 reaches with presence observations, the MTP and 10PMTP thresholds equal the two lowest predicted suitability values among occupied reaches. So, values equal to zero show that even with an ensemble approach, BHS models in all months did not accurately represent BHS habitat preference given conditions at all reaches with species observations. Sparse species distribution data hampered model fitting.

Table 3
Bonneville Cutthroat Trout (BCT) Model Fit Metrics (median values of 10 iterations).

Month	Train AUC	Test AUC	Train TSS	Test TSS	SEDI: MTP	SEDI: 10PMTP	SEDI: 0.5	MTP	10PMTP	Reg. Mult.	MaxEnt Iterations
Jan	0.75	0.74	0.39	0.38	0.14	0.14	0.56	0.04	0.3	2.75	4390
Feb	0.78	0.75	0.45	0.4	NA	NA	0.68	0.05	0.3	2.6	3740
Ma	0.78	0.74	0.42	0.38	0.46	0.46	0.65	0.03	0.31	2.7	4200
Apr	0.77	0.73	0.4	0.39	0.35	0.35	0.59	0.05	0.32	3.2	4610
May	0.79	0.75	0.44	0.4	0.33	0.33	0.64	0.02	0.26	2.1	4680
Jun	0.77	0.74	0.41	0.4	0.31	0.31	0.62	0.03	0.29	2.1	5060
Jul	0.78	0.76	0.42	0.4	0.45	0.45	0.63	0.05	0.31	2.95	4990
Aug	0.79	0.76	0.45	0.42	NA	NA	0.64	0.01	0.29	2.9	4420
Sep	0.77	0.74	0.41	0.38	0.21	0.21	0.64	0.03	0.25	2.25	4550
Oct	0.77	0.73	0.41	0.37	0.17	0.17	0.61	0.05	0.27	2.55	5930
Nov	0.78	0.77	0.44	0.42	0.15	0.15	0.66	0.03	0.27	1.85	4240
Dec	0.75	0.74	0.39	0.39	0.25	0.25	0.54	0.05	0.32	2.85	5670

Table 4

Bluehead Sucker (BHS) Model Fit Metrics (median values of 1000 iterations).

Month	Train AUC	Test AUC	Train TSS	Test TSS	SEDI: MTP	SEDI: 10PMTP	SEDI: 0.5	MTP	10PMTP	Reg. Mult.	MaxEnt Iterations
Jan	0.92	0.75	0.85	0.5	0.77	0.77	0.8	0	0.63	1	500
Feb	0.93	0.75	0.85	0.5	0.77	0.77	0.79	0	0.63	1	500
Ma	0.95	0.75	0.85	0.5	0.78	0.78	0.8	0	0.63	1	500
Apr	0.93	0.69	0.85	0.5	0.72	0.72	0.77	0	0.63	1	500
May	0.92	0.69	0.85	0.5	0.69	0.69	0.72	0	0	1	500
Jun	0.95	0.75	0.92	0.5	NA	NA	NA	0	0.63	1	500
Jul	0.92	0.63	0.85	0.5	0.67	0.67	0.72	0	0	1	500
Aug	0.94	0.72	0.85	0.5	0.76	0.76	0.7	0	0.07	1	500
Sep	0.93	0.75	0.85	0.5	0.75	0.75	0.78	0	0.63	1	500
Oct	0.92	0.75	0.85	0.5	0.7	0.7	0.73	0	0	1	500
Nov	0.95	0.75	0.92	0.5	0.76	0.76	0.79	0	0.63	1	500
Dec	0.94	0.75	0.92	0.5	0.79	0.79	0.81	0	0.63	1	500

3.2 Model Validation

3.2.1 Environmental Conditions

Overall, real-world conditions between June-October 2022 were not accurately represented by any of the three modeling methods compared. The HYD, THRESH, and GEO models all overestimated streamflow, and the THRESH and HYD models underestimated stream temperature. These results were generally expected due to comparing point measurements in a particularly hot, dry year to monthly averages over full reaches; and in the case of the HYD model, summarized over particularly large spatial extents.

Table 5
Environmental data validation using 2022 field observations and model input estimates. Color coding represents quality of fit: red = not satisfactory, orange = satisfactory, yellow = good, green = very good (D. N. Moriasi et al., 2015, 2007).

Model	Parameter	Adjusted R^2	PBIAS	NSE
HYD	Streamflow	0.221	4.6	-1.46
HYD	Average Depth	-0.026	18.3	-1.75
GEO/THRESH	Streamflow	0.357	30.9	-3.27
THRESH	Temperature	0.599	-27.5	-1.59

The HYD model used streamflow measurements from 2003 to predict channel depth and subsequently habitat suitability for BCT and BHS. Though streamflow was slightly higher in 2022 compared to 2003 based on ranked water years using streamflow data at Corinne, both streamflow and stream depth estimates should have been comparable to observed conditions in 2022 assuming channel morphology was unchanged.

Stream conditions in 2022 were not well represented by HYD model estimates (Table 5). Negative NSE values indicate using average observed streamflow measurements were less variable than model estimates. Adjusted R^2 values reflect a near random fit for both streamflow (Fig. 16) and channel depth (Fig. 17), which is likely attributed to effects of extreme outliers on small sample sizes. Positive PBIAS values indicated model predictions were greater than observed conditions; the opposite of what was expected given 2003 was an extremely dry water year.

Both the THRESH and GEO models used percent MAF as a predictor, both based on NHD monthly averages. Though the THRESH model's full extent far exceeds the Bear River Watershed, Table 5 presents results for just the Bear River Watershed to match the observation dataset. Overall, observed streamflow in 2022 in the Bear River Watershed was not well represented by NHD estimates (Fig. 18). NSE values were negative indicating average observed streamflow was less variable than modeled estimates, and large positive PBIAS value shows the THRESH and GEO models overestimated streamflow. Adjusted R^2 values also indicate a poor fit (Table 5). These findings, however, were not surprising because I expected modern conditions such as those observed in 2022 to reflect drier conditions than estimates from the NHD (1971-2000).

Stream temperature estimates were made for each reach of the THRESH model. The relationship between observed and modeled stream temperature reflected satisfactory model performance given adjusted R^2 (Table 5). However, NSE was negative, indicating poor model fit, and a large PBIAS estimate showed model estimates were often much lower than observed conditions (Fig. 19). This too is unsurprising, given observations were restricted only to the mainstem Bear River, and during one of the hottest years on record (National Drought Mitigation Center, 2023).

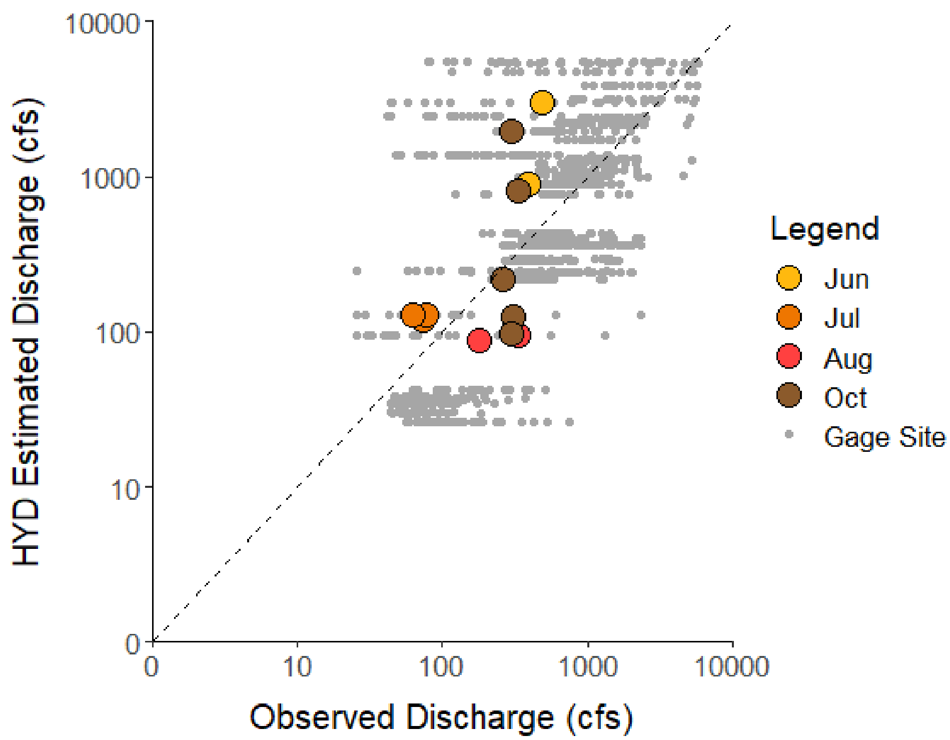

Fig. 16. Relationship between observed and modeled streamflow (discharge) in the Hydraulic-Habitat (HYD) model. Gage site observations (2001-2022) were included to visually compare with 2022 field observations. The dashed line indicates a 1:1 relationship.

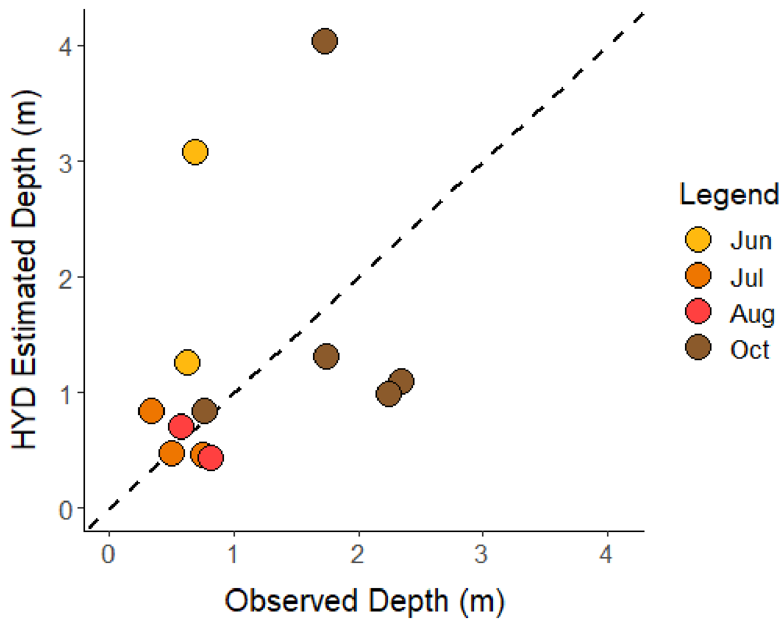

Fig. 17. Relationship between average observed and modeled channel depth in the Hydraulic-Habitat (HYD) model. The dashed line indicates a 1:1 relationship.

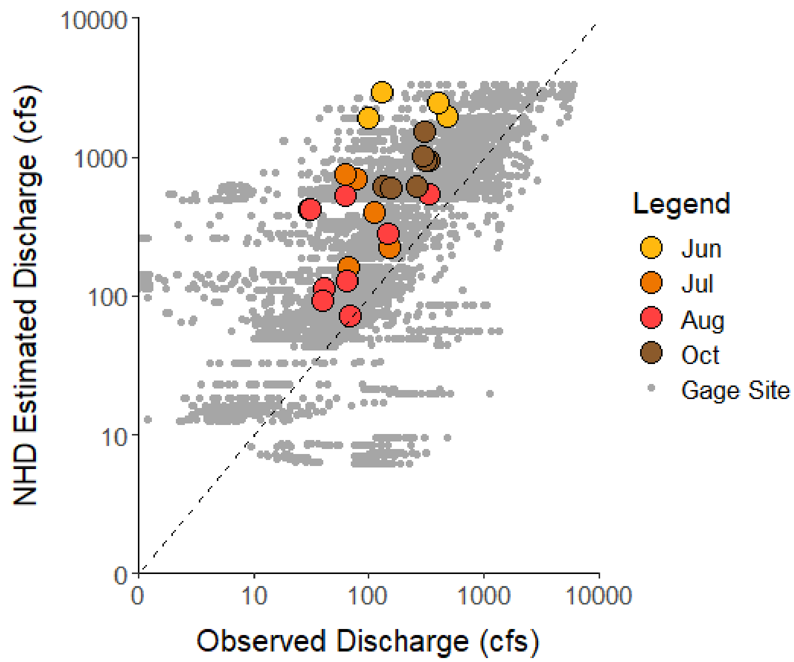

Fig. 18. Relationship between observed streamflow (discharge) versus average monthly National Hydrography Dataset (NHD) estimates used in the Geospatial (GEO) and Habitat Threshold (THRESH) models. Gage site observations (2001-2022) were included

to visually compare with 2022 field observations. The dashed line indicates a 1:1 relationship.

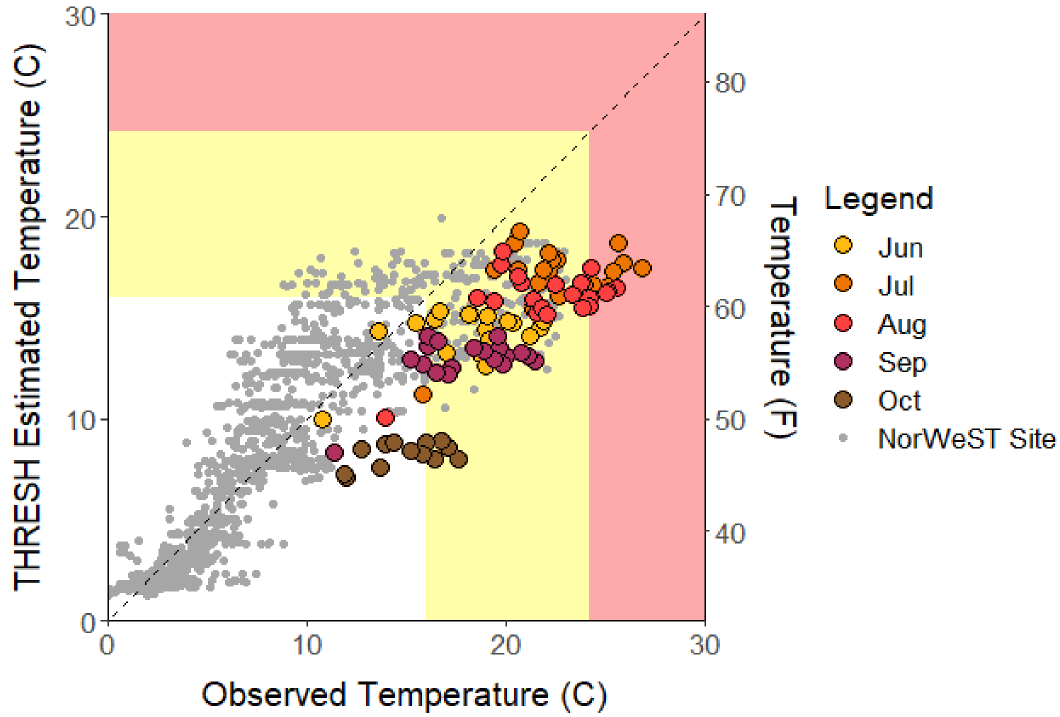

Fig. 19. Relationship between observed stream temperatures and Habitat Threshold (THRESH) model stream temperatures. NorWeST 2000-2018 measurements within the Bear River Watershed were included to visually compare with 2022 field observations. The dashed line indicates a 1:1 relationship. Yellow shaded areas represent temperatures considered stressful for Bonneville Cutthroat Trout (BCT), and red shaded areas represent temperatures that exceed the BCT lethal threshold.

Though impossible to capture by a performance metric, it is important to emphasize how models often simplify conditions, such as assuming a rectangular channel when true channel cross-sections are more complex (Fig. 20). Field measurements reflected a much wider reach, but with comparable average channel depth. Different channel morphologies can result in great differences in the relationship between streamflow and stream velocity.

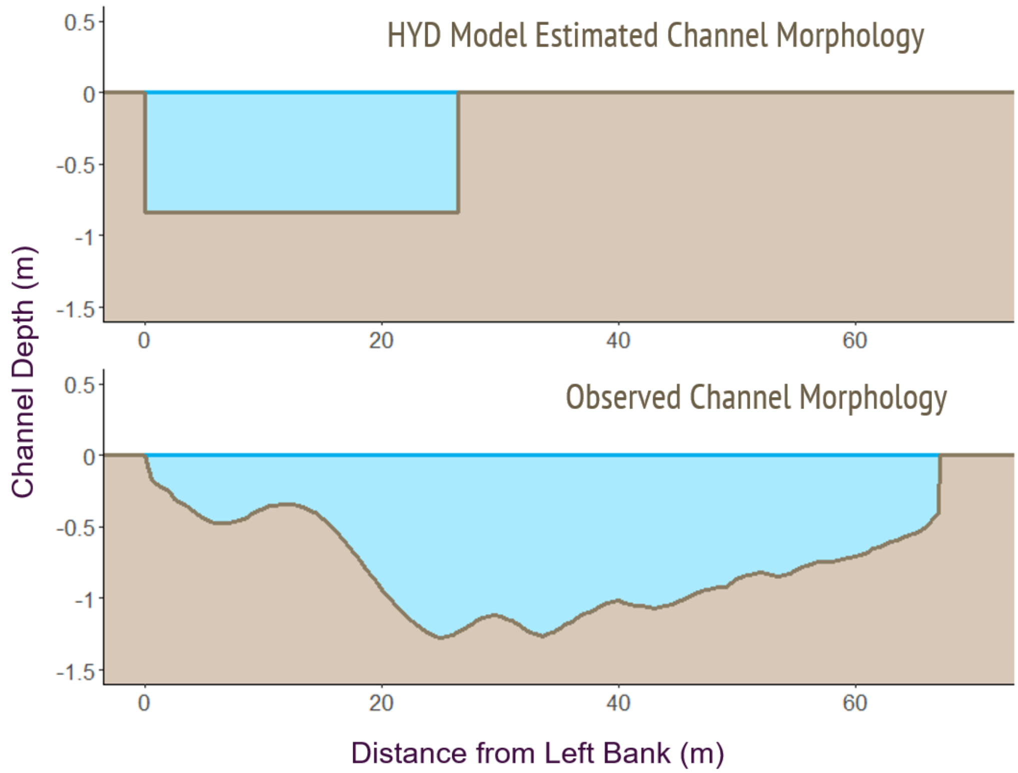

Fig. 20. Hydraulic-Habitat (HYD) modeled (top) and observed (bottom) channel morphology in the mainstem Bear River reach upstream of Oneida Reservoir, UT.

In a final notable point again not measurable by metrics suggested by Moriasi et al. (2015), I compared observed mainstem Bear River stream temperatures between June and October 2022 and HYD model thresholds (Fig. 21). In summer months, particularly July and August, observed temperatures in nearly all of the mainstem Bear River exceeded the HYD threshold for BCT temperature suitability (Alafifi and Rosenberg, 2020), as well as the BCT lethal limit throughout most of the modeled extent. I also observed occasional exceedance of the maximum expected temperature for the full extent of the HYD model, though this was independent of species thresholds, in downstream reaches toward Great Salt Lake in these months. Stream temperature observations along the mainstem Bear River in summer 2022 were consistently warmer than observations at NorWeST sites between 2000-2018 at gages both along the mainstem and in tributaries, as well. This exemplifies how outdated, static model conditions such as these can be inappropriate in a changing climate. Depending on the model developer's intent, exceedance of threshold such as this could have implications for determining whether or not reaches could be considered viable habitat, altering habitat quality estimates, and subsequently changing management implications.

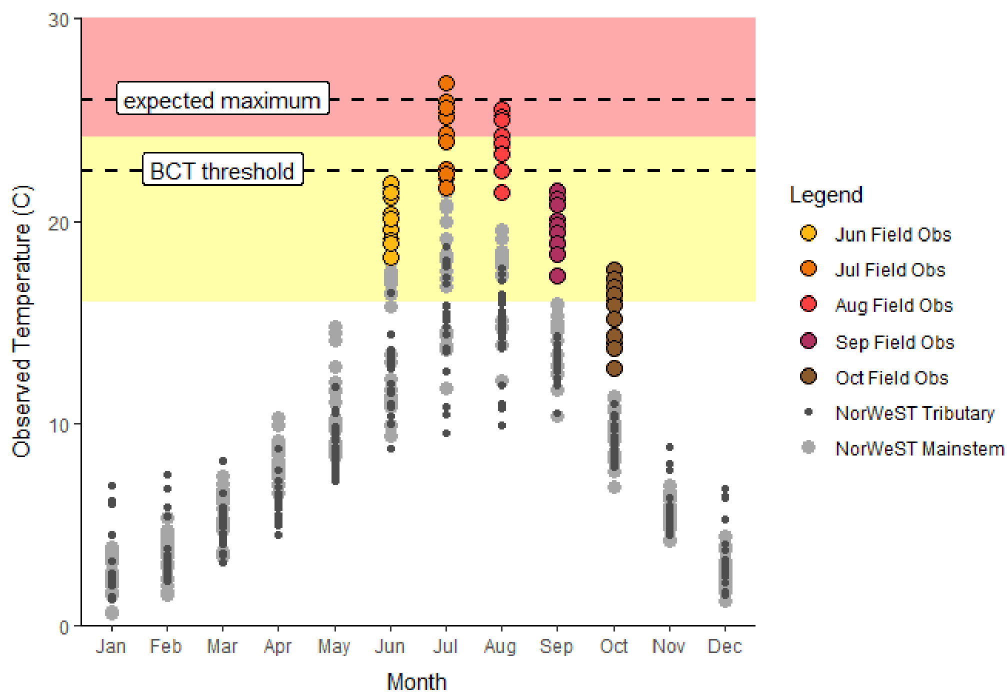

Fig. 21. Observed average monthly stream temperatures within the Hydraulic-Habitat (HYD) model extent from June-October 2022. NorWeST measurements within the Bear River Watershed show in-stream temperatures from 2000-2018, where small dark gray points are observations from tributary sites and larger light gray points are from mainstem sites. Yellow shaded area reflects the temperature range considered stressful for Bonneville Cutthroat Trout (BCT), where areas shaded in red reflect exceedance of the lethal threshold. The dashed line labeled "BCT threshold" is the HYD model temperature threshold. The dashed line labeled "expected maximum" is the maximum temperature expected within the HYD model extent.

3.2.2 Quantifying Model Accuracy

3.2.2.1 Threshold-based Proportion of Correctly Classified Reaches (TPC)

The THRESH and GEO models created with either MTP or 10PMTP thresholds were 100% accurate in correctly classifying occupied reaches as suitable over the full year (Table 6). The HYD model with binary suitability threshold determined by sensitivity analysis ranked third for predicting suitable BCT habitat, followed by the GEO model with a threshold of 0.5. The HYD model created with a threshold of 0.5 performed worst for predicting BCT habitat by this metric. These results show all three models perform well for predicting BCT habitat if thresholds are carefully selected, and threshold selection for models with continuous habitat suitability affects model performance.

Table 6

Threshold-based Proportion of Correctly Classified Reaches (TPC: expressed as a percentage) for Bonneville Cutthroat Trout (BCT) models, calculated using median predictions across model iterations for each month. SA means threshold determined by Sensitivity Analysis, and ME means MaxEnt determined thresholds (Minimum Training Presence (MTP) and 10th Percentile Minimum Training Presence (10PMTP)) results were combined because they were identical**.**

Model	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
	%	%	%	%	%	%	%	%	%	%	%	%
HYD (0.5)	67	44	0	0	0	0	0	0	50	33	33	56
HYD (SA)	83	89	83	83	86	88	83	83	83	83	83	89
GEO (0.5)	73	76	80	76	72	73	80	71	75	74	73	68
GEO (ME)	100	100	100	100	100	100	100	100	100	100	100	100
THRESH	100	100	100	100	100	100	100	100	100	100	100	100

All models created for BHS appeared to perform well and exhibited a narrower range in performance across model types (Table 7). This is likely because across all model types, most of the modeled extents were predicted suitable (regardless of threshold selected for the HYD and GEO models), resulting in most if not all reaches with presence observations correctly classified as suitable habitat. The HYD model with a threshold selected by sensitivity analysis correctly classified 100% of occupied reaches as suitable habitat for all months. The HYD model with a threshold of 0.5 was 100% accurate for all months except in winter, and the THRESH model was nearly perfectly accurate over the full year. All GEO models perform identically, except in August, where there was a slight decrease in performance for the model evaluated based on a threshold of 0.5. GEO models were consistently slightly less accurate than other model types, except where winter predictions made by the HYD model with a threshold of 0.5 were notably worse (though still overall quite good). Again, setting different thresholds within model type affected accuracy of predictions, if even slightly.

Table 7
Threshold-based Proportion of Correctly Classified Reaches (TPC: expressed as a percentage) for Bluehead Sucker (BHS) models, calculated using median predictions across model iterations for each month.

Model	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
	%	%	%	%	%	%	%	%	%	%	%	%
HYD (0.5)	75	75	100	100	100	100	100	100	100	100	100	75
HYD (SA)	100	100	100	100	100	100	100	100	100	100	100	100
GEO (0.5)	94	94	94	94	88	100	88	88	94	88	94	94
GEO (ME)	94	94	94	94	88	100	88	94	94	88	94	94
THRESH	97	97	99	99	99	99	99	99	99	99	99	99

3.2.2.2 Weighted Proportion of Correctly Classified Reaches (WPC)

WPC was calculated to assess model performance without introducing assumptions regarding threshold selection. For BCT models, the THRESH model correctly predicted occupied reaches as suitable with 100% accuracy in all months, the GEO model averaged 64% correct classifications, and the HYD model averaged 29% correct classifications for all months. WPC for the HYD model was especially poor between March and August (Table 8). For BHS models, the THRESH and HYD models performed similarly. The THRESH model averaged 99% correct classifications, and the HYD averaged 95% correct classifications. The GEO model performed worst with an average of 67% correctly classified occupied reaches in all months (Table 9).

Table 8
Weighted Proportion of Correctly Classified Reaches (WPC: expressed as a percentage) for Bonneville Cutthroat Trout (BCT) models, calculated using median predictions for Geospatial (GEO) models across iterations for each month.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	%	%	%	%	%	%	%	%	%	%	%	%
HYD	55	45	3	3	3	3	3	3	53	53	54	64
GEO	64	65	66	65	64	64	66	64	63	64	64	64
THRESH	100	100	100	100	100	100	100	100	100	100	100	100

Table 9
Weighted Proportion of Correctly Classified Reaches (WPC: expressed as a percentage) for Bluehead Sucker (BHS) models, calculated using median predictions for Geospatial (GEO) models across iterations for each month.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

	%	%	%	%	%	%	%	%	%	%	%	%
HYD	86	81	96	100	100	100	100	100	99	99	95	79
GEO	67	70	79	72	63	67	67	62	62	64	68	58
THRESH	97	97	99	99	99	99	99	99	99	99	99	99

3.2.2.3 Adjusted Weighted Proportion of Correctly Classified Reaches (WPC_{Adj})

Adjusted WPC was calculated to compare models without selecting thresholds and to normalize WPC based on the weighted proportion of modeled extent predicted suitable, because WPC was biased towards models that predicted suitable habitat almost everywhere. This WPC adjustment benefited models which exhibited greater discernment in predicting suitable habitat (*Figs. 22 and 23).* Values reported in tables are unitless, but are bounded between 0 and 1. Values closer to 0 or 1 reflect worse or better performance, respectively.

For BCT models, the GEO model averaged best performance throughout the year, though the HYD model performed best in winter months. The THRESH model performed worst by this metric given an annual average, though slightly better than the HYD model in summer months year (Table 10). The GEO model performed best among BHS models as well, with the THRESH and HYD models performing similarly in all but summer months where the HYD model performed substantially worse (Table 11).

Table 10
Adjusted Weighted Proportion of Correctly Classified Reaches (unitless) for Bonneville Cutthroat Trout (BCT) models, calculated using median predictions for Geospatial (GEO) models across iterations for each month.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

HYD	0.49	0.4	0.03	0.03	0.03	0.03	0.03	0.03	0.47	0.48	0.48	0.54
GEO	0.36	0.4	0.39	0.37	0.4	0.38	0.38	0.39	0.38	0.37	0.39	0.36
THRESH	0	0	0	0	0	0.09	0.15	0.03	0	0	0	0

Table 11
Adjusted Weighted Proportion of Correctly Classified Reaches (unitless) for Bluehead Sucker (BHS) models, calculated using median predictions for Geospatial (GEO) models across iterations for each month.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
HYD	0.25	0.31	0.13	0	0	0	0	0	0.22	0.17	0.2	0.26
GEO	0.53	0.54	0.6	0.54	0.49	0.53	0.51	0.48	0.49	0.5	0.53	0.47
THRESH	0.25	0.25	0.24	0.24	0.24	0.27	0.27	0.24	0.25	0.24	0.24	0.25

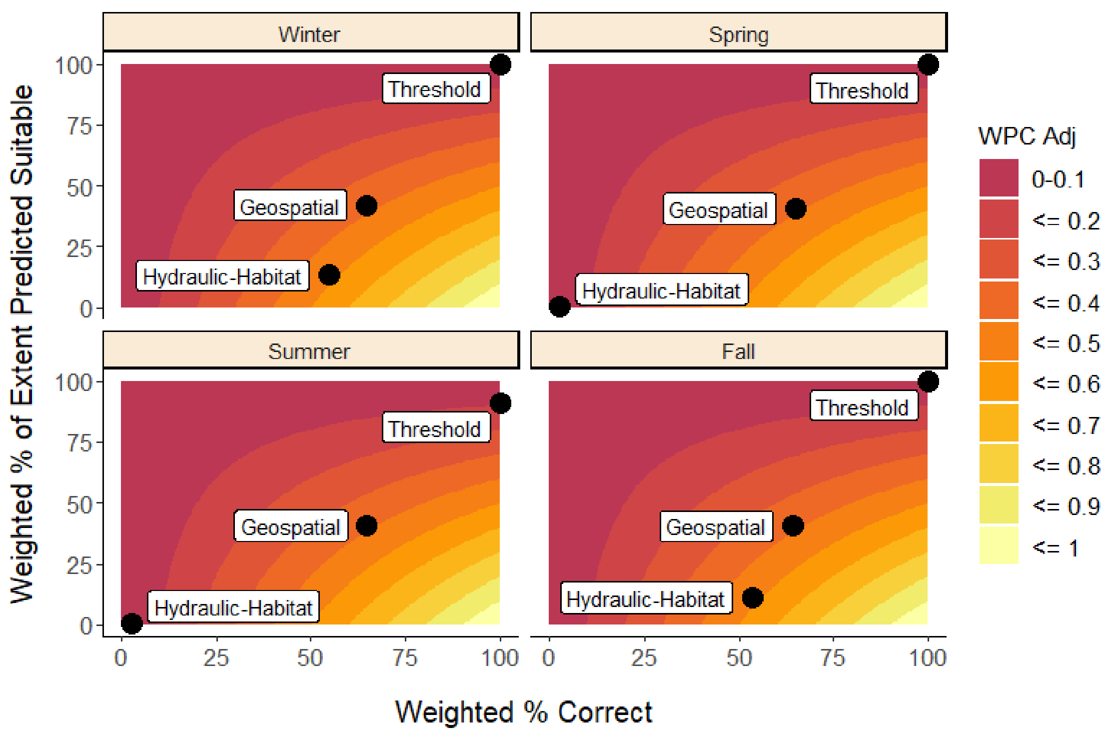

Fig. 22. Adjusted Weighted Proportion of Correctly Classified Reaches (WPC_{Adj}) as a function of Weighted Proportion of Correctly Classified Reaches (WPC) expressed as a percent, and percent of total habitat predicted suitable for Bonneville Cutthroat Trout (BCT) models.

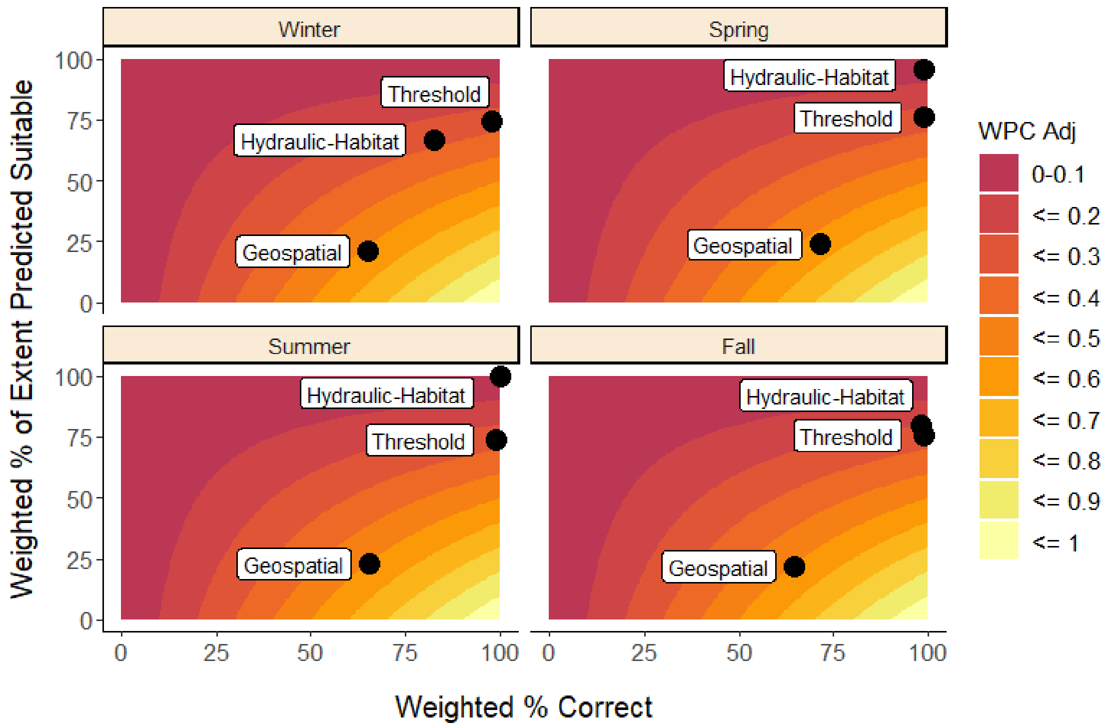

Fig. 23. Adjusted Weighted Proportion of Correctly Classified Reaches (WPC_{Adj}) as a function of Weighted Proportion of Correctly Classified Reaches (WPC) expressed as a percent, and percent of total habitat predicted suitable for Bluehead Sucker (BHS) models.

DISCUSSION

Model types compared differed in performance depending on threshold selection, model development choices, and performance metric selection. None of the models compared were particularly representative of real-world environmental conditions, though some provided more believable estimates of habitat availability, and better context for species preferences, than others. Overall, the THRESH and GEO models performed best for BCT, and the THRESH model performed best for BHS. Habitat model selection should be based on management goals, but generalized habitat models, such as the THRESH model, that incorporate both hydrologic and species-specific biological predictors may be best for use in conjunction with water management. However, considering model results across types may prove useful for managers looking to maximize both habitat model accuracy and precision. This additional precision could be gained through combining process-based models results, like the THRESH model, with empirical model results, like the GEO model.

*4.1 Geospatial (GEO) Model Habitat Predictors for Bonneville Cutthroat Trout (BCT) and Bluehead Sucker (BHS***)**

There has been a recent flux of studies comparing how accurately variables representing different spatial scales predict aquatic habitat suitability. These studies demonstrate that landscape-scale variables can adequately predict species distribution and may even more accurately represent species habitat preference than traditional microscale variables because landscape variables act as a surrogate, or proxy, for local conditions through stream hierarchical theory (Creque et al., 2005; Frissell et al., 1986; Maxwell et al., 1995; Richards et al., 1996). When microscale variables have been included in models, they have not significantly improved overall predictive accuracy (Falke, 2006; Kristensen et al., 2012; Meixler and Bain, 2012; Steen et al., 2008b).

In my ensemble application of MaxEnt, Total Upstream Catchment Area had highest WPI of landscape-scale predictors for BCT and BHS habitat suitability. As shown in *Figs. 12 and 15*, BCT prefer reaches with lower total upstream catchment area, while BHS prefer reaches with medium to high total upstream catchment area. Even though there are proportionally more reaches within the GEO model extent with lower total upstream catchment area, BCT were not found as often as would be randomly expected in reaches with greater total upstream catchment area (Fig. 12*).* BHS on the other hand, were found less often than would be expected in reaches with lower total upstream catchment area (Fig. 15).

For BCT models, land use consistently ranked as second most important throughout the year. Land use is often used as a predictor in fish habitat modeling, particularly in geospatial modeling (Dauwalter et al., 2011; Giacomazzo et al., 2020; Joy and Death, 2004; Meixler and Bain, 2012). Land use encapsulates many other habitat characteristics such as water quality and development. For example, streams located proximal or downstream of agricultural or grazing areas can have poor water quality due to sediment and nutrient inputs (Richards et al., 1996; Worthington et al., 2016), and development is an indicator of human influence (Argent et al., 2003; McKenna and Johnson, 2011; Van Sickle et al., 2004). Both land and water in the Bear River Watershed have been developed to meet human needs. These alterations to natural systems have harmed native species, particularly construction of in-stream barriers that fragment habitat, including dams, culverts, and road crossings (Budy et al., 2007; Dzara et al., 2019; Kraft et al., 2019; Walsworth and Budy, 2015; Worthington et al., 2016).

Based on the results of the GEO model, BCT preferred forested areas (Fig. 13). The Bear River Watershed has extensive agriculture and rangeland, as well as shrub-scrub dominated landscapes. In the Bear River Watershed, total upstream catchment area has a monotonically negative relationship with percent of canopy cover (Spearman's coefficient of -0.47), where reaches with less total upstream catchment area have greater percentage of forest cover. Percent of canopy cover also has a monotonically negative relationship with mean monthly air temperature, where reaches with less canopy cover have greater mean monthly air temperatures (Spearman's coefficient of -0.41). Air temperature and atmospheric conditions affect stream temperatures in warmer months (Goodrum, 2020; Goodrum and Null, 2022). While stream temperature is the single most important variable in determining habitat suitability for salmonids, reaches with greater canopy cover can also have more large woody debris (LWD). LWD creates habitat complexity, including opportunities for cover and foraging potential (Bocchiola, 2011; Mugodo et al., 2006; Wheaton et al., 2004).

Results suggest BCT prefer forested reaches with lower total upstream catchment area which are correlated with lower in-stream temperatures and great habitat complexity. This coincides with prior knowledge of BCT ecology (Goodrum and Null, 2022). However, in the Bear River Watershed, it is likely in-stream barriers have fragmented habitats, so BCT preference for lower total upstream catchment area may be an effect of fitting an empirical model to a fragmented river network rather than a natural preference of the species. Additionally, the lack of a timestamp for fish presence datasets likely affected the importance of predictors that had a temporal component.

There is little evidence to support or dispute the findings of BHS models because most BHS research has focused on the neighboring Colorado Basin, where BHS is soon to be described as a genetically distinct species (Bangs and Douglas, 2017; Dauwalter et al., 2011; Hopken et al., 2013; Unmack et al., 2014). While geospatial modeling approaches could be an excellent approach to further clarify habitat use and preference of BHS native to the Bear River Watershed, more comprehensive sampling records are needed. Small sample sizes could be attributed to either limited sampling efforts or true species rarity in the basin, but I suspect the BHS presence dataset could have included observations from distinct metapopulations with different local adaptations and habitat preferences (Bangs and Douglas, 2017; Thompson and Burnett, 2019; Webber et al., 2012). The results suggesting BHS prefer greater total upstream catchment area is likely a result of fitting models to remnant populations in the basin. Better understanding BHS habitat needs and preferences in the Bear River Watershed is a priority (Webber et al., 2012).

4.2 Model Comparison and Predictive Accuracy

Comparing models with different extents, outputs, underlying assumptions, validation datasets, and intended uses was complicated. The THRESH model predicted habitat to be suitable, or non-lethal, nearly everywhere in all months for both species. With this, it consistently performed well regardless of validation metric used, though was not always best. When calculating TPC, the GEO models based on either MTP or 10PTMP thresholds performed just as well as the THRESH model for BCT, and all models performed similarly for BHS. For BCT models, both the HYD and GEO models performed much better when evaluated at relatively low thresholds, but significantly dropped in performance when applying a higher threshold (0.5). In particular, the HYD model with a threshold of 0.5 performed markedly the worst for BCT. This is likely because at these lower thresholds, again, most reaches within the modeled extent were predicted to be suitable habitat. The effect of threshold selection on TPC was less pronounced for BHS models. When compared using WPC as the performance metric, unsurprisingly the THRESH model still performed best for BCT, and the THRESH and HYD models far outperformed the GEO model for BHS. This is unsurprising because most of the THRESH and HYD extents were predicted suitable for both species. With WPC_{Adj} as the performance metric, the THRESH and HYD models were heavily penalized for this (especially for BHS models), and the GEO model performed best.

As the THRESH model was created over the largest spatial extent, with coarse variables, and calibrated with the largest presence dataset, this approach is highly generalized and it is surprising for it to result in higher predictive accuracy than the HYD model in particular, which was developed for the smallest extent, with finer scale parameters and a smaller calibration dataset. This can be attributed to these fine resolution predictors being aggregated over the largest spatial extents. Even if scaled up, I would still not expect the HYD model to perform as well given its inaccurate, but specific, nature likely not fitting an even more diverse dataset (Wheaton et al., 2018). The GEO model, between the other two methods in regards to spatial extent, would be expected to be more generalized than the HYD model, but more specific than the THRESH model, and would be expected to perform well at reduced extents, but poorer if scaled up without being updated to population-scale datasets. I observed this effect, where there was usually a gradient in model performance relative to model generality and scale, except in cases where extremely low thresholds were chosen to calculate proportion of correct classifications.

4.2.1 Threshold Selection

As demonstrated by my first performance metric, TPC, threshold selection clearly plays a major role when evaluating predictive accuracy of binary habitat model predictions. Not all models may be intended to be used with thresholds. Though lower thresholds for models with continuous outputs lead to a greater number of correct classifications and apparent predictive accuracy, lowering thresholds may simply over-estimate suitable habitat or over-generalize a model that performs relatively well on its own.

It is common in geospatial modeling to select thresholds close to the estimated value of species prevalence (expressed as the proportion of occupied habitat to total suitable habitat; between 0 and 1), similar to those I reported (Benkendorf and Hawkins, In Review; Dorji et al., 2020; Liu et al., 2016). As demonstrated where estimated MTP and 10PMTP thresholds for my BHS GEO models approached or were equal to zero, poor model fit resulted in most or all of the modeled extent being classified as suitable habitat.

Habitat suitability in the HYD model was dependent on channel depth. The threshold to determine the line between suitable and unsuitable habitat selected by sensitivity analysis resulted in channel depths of .2 m for BCT and .02 m for BHS. Though BCT and BHS can survive in extremely shallow water, selection of a low threshold such as this may not indicate differences between high and low habitat quality, but temporarily survivable and outright lethal conditions. Similarly, habitat suitability in the THRESH model was dependent on known biological thresholds for stream temperature, gradient, and streamflow. These thresholds are considered to represent the lethal limits for species and are based on findings in laboratory studies (Johnstone and Rahel, 2003). Particularly in regards to lethal thermal limits, these thresholds may not reflect the true lethal limit for locally adapted populations, or individual variation. Additionally, summarizing conditions at the reach scale based on biological thresholds which may or may not be locally applicable could result in either vast over- or underprediction of habitat quality due to missing effects of spatial heterogeneity. Even should the mean temperature of a reach be thermally viable for occupation, should the only reasonable habitat available within that reach based on other conditions (velocity, for example) have inhospitable temperatures, then that model may overestimate habitat quality. Conversely, should the mean temperature of a reach be expected to exceed the thermal threshold for a species, it is likely there are still pockets of non-lethal conditions which would make excellent temporary refuge if even not perfect habitat.

In any of the models compared, threshold selection for determining habitat suitability resulted in suitable habitat predicted nearly everywhere. Because of this, in validation steps, models appeared to perform with near perfect predictive accuracy. In truth, it is unclear if model performance was a result of over-generalized suitability predictions or genuine accuracy (Escalante et al., 2013; Liu et al., 2016; Radosavljevic and Anderson, 2014). Regardless of appropriateness of threshold selection or calculated accuracy, nuance provided by continuous habitat suitability estimates is lost when converting predictions to binary estimates. Particularly when integrating ecological objectives into water management models, use of thresholds in habitat suitability modeling must be carefully considered, because over-estimations or over-generalizations of habitat suitability across large extents could unintentionally de-prioritize ecological objectives relative to other objectives if there is no perceived urgency.

4.2.2 Performance Metrics and Validation Data

Model validation has been emphasized in this study because it is important to evaluate how representative they are of the real world, as well as model performance. Without extensive validation, model accuracy can be misrepresented (Mouton et al., 2007), but without validation at all, there is no justification for a model's usefulness.

None of the models sufficiently represented real-world channel conditions as observed in summer of 2022. The HYD model assumed simple channel shapes and channel uniformity. As discussed, a rectangular channel does not accurately represent real-world channel conditions, and oversimplifies spatial heterogeneity of rivers with differing velocities and depths along a cross section. These differences provide meaningful habitat opportunities for fish. Two- and 3-dimensional models typically use measured data for channel shape, and provide more detail for lateral and longitudinal channel variation. The HYD model in this study was unique compared to other hydraulic-habitat models because reach resolution was quite coarse, with median reach length 24.3 km. Combined with outdated streamflow estimates, 1D hydraulic models with extremely large reaches result in inaccurate estimations of channel conditions over large spatial extents.

The THRESH and GEO models had coarse resolution as well. Most variables in the GEO model had 30 m resolution which could not be validated, and data were summarized at the HUC-12 scale. THRESH model temperature estimates were based on 5 km resolution data, and gradient estimates from a 10m Digital Elevation Model (DEM). It was beyond the scope of this project to validate stream gradient. While air temperature was an excellent proxy for stream temperature (Goodrum, 2020; Goodrum and Null, 2022), increasing frequency of hot years as a result of climate change can underestimate stream temperatures. Hydrologic variables in both models were based on the NHD, which has a resolution of 1 km, and is also a generalized and somewhat outdated approach for estimating channel conditions.

The relationship between my field observations in the Bear River Watershed in 2022 and average monthly NHD estimates could be an indicator for relationships between other ungaged streams and the NHD in the modern era. This is notable because while the NHD is commonly used in modeling efforts to represent streamflow, it may not be well suited in heavily regulated systems subject to dewatering of stream reaches or sites affected by climate warming or increased variability. Streamflow overestimates affect in-stream temperature estimates as well, which lead to inaccurate estimates of habitat quality for native fish. More extensive and potentially region-specific validations of this effect may be necessary before inclusion of these estimates as model predictors, particularly for drought periods in the American West.

When validating habitat suitability models, presence-absence datasets provide more context for the accuracy of model predictions. Species presence is not always synonymous with habitat suitability (Valavanis et al., 2008b), and rapidly changing habitat quality due to climate change, land development, or extreme events can lead to increased abundance in poor quality reaches (Brotons et al., 2004), or render preferred habitats inaccessible (Ward et al., 2009). Including true absence data can aid in determining suitable versus unsuitable conditions because absences can reflect low suitability (or accessibility), as well as provide insight to model overprediction (Brotons et al., 2004). Presence-absence datasets are uncommon, though, as determining true absence is difficult with migratory species where species occupation can be heavily subject to environmental stochasticity (Leroy et al., 2018; Ward et al., 2009).

The presence-only validation dataset I had was sparse and likely biased, given presence observations were almost certainly collected between late spring to early fall, and most sampling efforts are geared towards wadable tributary reaches. Knowing dates of observations would have improved modeling efforts and created a better understanding of how habitat preferences for BCT and BHS change throughout the year. When validating model performance with presence-only datasets, I recommend calculating proportion of correct classifications when model outputs are binary, or WPC when model outputs are continuous. As previously discussed, selection of a threshold to distinguish suitable from unsuitable conditions is key when calculating proportion of correct continuous classifications, so it is ideal to use this method with binary model results.

Conversely, WPC does not require a threshold and can provide clearer insights to compare model performance when outputs are continuous. However, it is difficult to differentiate effects of model output and model performance when WPC is used to compare models with both binary and continuous outputs. With large sample sizes (much greater than what I had available), Generalized Linear Models or Generalized Linear Mixed Effects Models with binary distributions can determine differences in model performance with presence-only data, but require selection of a threshold (Elith et al., 2006; Valavi et al., 2021). These methods can further differentiate between monthly habitat predictions, and provide a metric of statistical significance in regards to relative model performance.

Other methods I explored for determining habitat model performance, but do not recommend for presence-only validation, include commonly used metrics such as AUC, TSS, and SEDI, and my WPC_{Adj} method. Each of these methods move beyond simply indicating model accuracy to a validation dataset, and further provide inferences regarding model precision. However, all of these methods are either sensitive to sample size, outliers, or make assumptions about areas with unknown occupation. Because SEDI is a relatively new performance metric in presence-background modeling applications (Wunderlich et al., 2019), it was valuable to explore its practicality with my models. Despite SEDI's intended use with models created with small datasets, when evaluating my models, I discovered prevalence still affected accuracy, particularly when most of a modeled extent was predicted to be suitable habitat. As far as my WPC_{Adj} method, though this may be useful in differentiating model performance in instances where there is known to be a gradient in habitat suitability across a study extent, it may excessively penalize models representing study areas where all area is truly viable habitat. Large study areas with little unsuitable habitat, however, are increasingly unlikely given the extensive human-caused habitat alteration common to the American West in a changing climate, and penalizing overfit/over-generalized models may be less of a risk for fish conservation than overpredicting suitable habitat extent.

4.2.3 Non-native Species

Though the THRESH model was unique in including biologically relevant criteria for both species, none of the physical habitat models compared here incorporated locations of competing and predatory nonnative species (Boavida et al., 2014; Dauwalter et al., 2011; Freeman et al., 1997). As previously mentioned, there are many non-native fish species in the Bear River Watershed known to compete with and prey on native fish; including Brook, Rainbow, and Brown Trout, Largemouth and Smallmouth Bass, as well as many others. Brown Trout, notably, are known to coexist with BCT and BHS, but in some cases, high abundance of nonnative species could make otherwise suitable habitat completely unsuitable; particularly for BHS (Budy et al., 2008, 2007; Budy and Gaeta, 2017; Walsworth and Budy, 2015).

Based on predictions made by the GEO model (Fig. 24), Brown Trout occupy both unsuitable and suitable habitat for BCT and BHS. While there are instances where both BCT and BHS are observed in reaches also occupied by Brown Trout, Brown Trout presence may partially explain lack of BCT and BHS observations in otherwise predicted suitable habitat. BCT and BHS concentrations are highest in predicted suitable areas where Brown Trout were not found, which may be an outcome of competitive exclusion.

It would be advantageous to incorporate non-native fish distribution into habitat modeling efforts, and this would be most helpful with comprehensive presence-absence datasets. Brown Trout occupation locations shown here are from UDWR and TU, and such sparse observations may not accurately reflect species distribution. Exploring the relationship between continuous habitat suitability model predictions with Rainbow or Brown Trout abundance could be a promising future step. Areas with moderate to high suitability estimates that are unoccupied by native fish could stand as a proxy for greater non-native abundance, potentially explaining unoccupied suitable habitat (Souza et al., 2022; Stephens et al., 2015; Weber et al., 2017). Also, though Rainbow Trout are non-native, they are genetically similar to BCT, and managers and habitat modelers could use Rainbow Trout presence or presence-absence data to validate BCT habitat models and understand if predictions could be overly conservative..

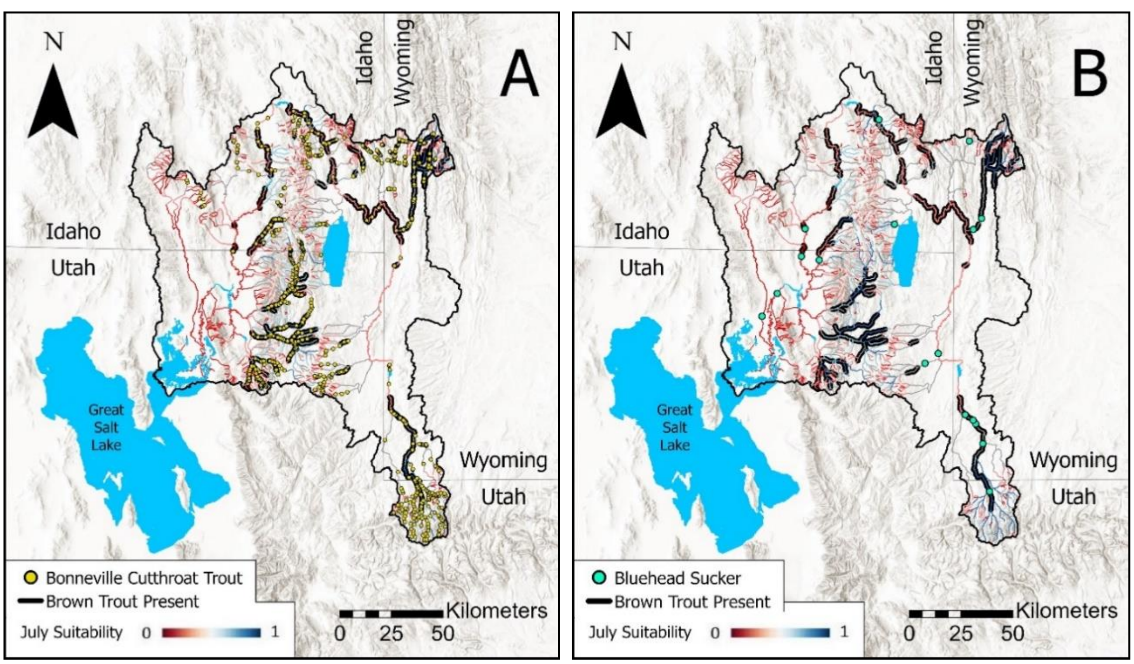

Fig. 24. Geospatial (GEO) model habitat suitability predictions for Bonneville Cutthroat Trout (BCT) (A) and Bluehead Sucker (BHS) (B) with reaches occupied by Brown Trout symbolized in black.

4.3 Management Implications

Fish and wildlife management plans are typically implemented by states, and can be inconsistent across neighboring state boundary lines. This makes managing species and incorporating ecological objectives into water management difficult, because watershed boundaries cross political boundaries. Political divisions like this are immaterial to migratory species such as fish (Berger et al., 2021; U.S. Fish and Wildlife Service, 2013). In transboundary watersheds, each state needs cohesive knowledge of native species distribution, status, and habitat preferences for water management models to effectively incorporate ecological objectives related to habitat conservation.

Simple and consistent methods for estimating species distribution or habitat suitability over large spatial extents can serve as a starting point for state agencies and biologists who must develop State Wildlife Action Plans (SWAP). It would be ideal for modeling efforts geared towards species within transboundary basins to either be implemented over entire basins, or for communication efforts be made for consistent modeling approaches across boundary lines. With uniformity, there could be easier incorporation of consistent ecological objectives into water management models.

Every modeling method has assumptions, limitations, and benefits that are affected by data quality and used to calibrate models, parameters, and underlying algorithms (Araújo et al., 2019; Valavi et al., 2021). Each habitat modeling method compared here has some unique limitations, and some are common to all. The choice of which method to use should be carefully weighed based on unique objectives behind modeling efforts and how well a method could meet those objectives.

Based on findings in this study, I recommend using the THRESH model approach for ease of implementation over large spatial extents and for different species, ability to be incorporated with water management models given inclusion of local hydrologic conditions, data availability, and meaningful incorporation of species-specific, biologically-relevant parameters. However, it important to recognize the univariate preference functions, often referred to as Habitat Suitability Indices, that are combined with hydraulic- or threshold-type models to estimate habitat suitability have been criticized for not meeting assumptions of independence among variables, and for misrepresenting habitat needs of populations locally adapted to stressful environmental conditions (Boavida et al., 2013; Mäki-Petäys et al., 1997; Mouton et al., 2007; Vezza et al., 2015). I emphasize that it is critical to validate and/or calibrate models to yield accurate model results.

When selecting habitat modeling methods to use in conjunction with water management modeling, it is important to consider differences among habitat models. Habitat models with binary outputs may be best for understanding where suitable habitat should be, and may not sufficiently represent the gradients between lethal and non-lethal habitats that are critical for protecting fragmented and/or threatened species. Thus, determining additional limiting influences on species occupation may be needed. Models with continuous outputs can yield more detail pertaining to these gradients, and this detail may be most advantageous across alternative management objectives. This is a difference between accuracy and precision, or representing the fundamental (historical or potential) versus realized (current) niche of a target species.

In this study, empirical model results that yielded continuous outputs were accurate and precise, but did not reflect hydrologic conditions as influential variables. Without a tie to hydrologic conditions, it would be difficult to effectively incorporate habitat model estimates into water management models because water management models track streamflow and water deliveries to competing objectives. For example, water management models simulate or optimize proposed management strategies like water conservation, water markets, or managed aquifer recharge, and understanding effects on streamflow, stream temperature, and aquatic habitat suitability is needed for holistic water management (Edwards and Null, 2019; Kirk et al., 2020). However, empirical models may be useful in combination with results of process-based modeling methods to increase precision of estimates, reinforce distribution estimates, clarify current limiting factors for a target species, or to direct conservation efforts to high-priority areas.

Should empirical models be integrated with other methods, it is important to remember that they can narrowly estimate suitable habitat extent, realistically representing the realized niche of restricted remnant populations rather than historical—or potential—habitat. Geospatial modeling approaches in particular, which incorporate algorithms to predict habitat suitability, require programming or GIS experience, and time and skill to interpret results. Performance of empirical models can be subject to algorithm selection, and precautions must be taken to avoid the negative effects of low prevalence, spatial bias, and overfitting on model estimates (Benkendorf and Hawkins, In

Review; Elith and Leathwick, 2009; Leroy et al., 2018; Olden et al., 2002; Phillips et al., 2017; Ryo et al., 2021; Sofaer et al., 2019; Vezza et al., 2015). However, suitable habitat from validated and accurate empirical models that falls within predicted suitable habitat from process-based methods could have high conservation importance, particularly for threatened or endangered species. For example, should fish and wildlife conservation managers be involved in selecting appropriate sites for dam construction, empirical model results could identify lesser quality current habitat areas that would not impede migration patterns or fragment threatened populations from higher quality habitat areas.

In regards to management implications for both BCT and BHS in the Bear River Watershed, I would expect inconsistency in action recommendations based on the model type chosen. If using the THRESH model to make decisions, suitable habitat is predicted everywhere for both species so no conservation actions are needed. If the GEO model were used to inform management, suitable habitat is estimated to be spotty, which could result in de-prioritizing conservation in reaches that are occupied by fish species of interest.

I recommend using results from the THRESH and GEO model to inform decision making for BCT in the Bear River Watershed, and using the THRESH model alone for BHS. While the GEO model was limited by fish presence datasets used in model fitting, habitat suitability estimates for BCT are believable and useful between May-October (the season observations likely came from). THRESH model estimates for BCT were also accurate. HYD model habitat estimates for BCT are likely inaccurate due to poor input data. Together, the THRESH and GEO models could provide context for habitat use and preference of BCT in the Bear River Watershed. For BHS, GEO model estimates were likely overfit and HYD model estimates were subject to inaccurate model inputs, so I recommend the THRESH model alone.

CONCLUSION

In this study, I demonstrated landscape-scale variables are adequate predictors of habitat suitability for aquatic species, and have provided support for existing understanding of BCT and BHS habitat use and preference in the Bear River Watershed. I have collected data to support previous studies which suggest parameters used in hydraulic-habitat models can be inaccurate. While it is not optimal to compare modeling methods across varying extents and with different intended uses, I have identified several methods of evaluating predictive performance of three commonly used aquatic habitat modeling methods with presence-only data. By these methods, I determined threshold and geospatial modeling approaches performed best for BCT, and the threshold approach performed best for BHS. I also emphasized when multiple approaches appear to yield high accuracy, differences between methods may only be better understood by validating model predictions with presence-absence data, which can provide insight on whether or not models overpredict.

There is more to model selection, however, than just considering predictive accuracy. Model choice should be highly dependent on management objectives. Models that provide distinction between suitable and unsuitable habitat, rather than those that predict habitat suitability everywhere, are most beneficial in determining areas of critical habitat, restoration potential, or understanding rationale for suspected species absence, while more generalized models may be most useful for understanding potential niche. However, should intended function of a habitat model be to assert needs of native fish in water management models, there may be distinctive differences in ease of incorporation between habitat model types. Based on results of this study, I would recommend modeling approaches similar to the habitat threshold model compared here. In an era of climatic change and increasing human demands on water resources, this should be a major consideration, as it is essential for water management models to incorporate native fish needs as ecological objectives.

REFERENCES

Ahmadi-Nedushan, B., St-Hilaire, A., Bérubé, M., Robichaud, É., Thiémonge, N., Bobée, B., 2006. A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment. River Research and Applications 22, 503–523. https://doi.org/10.1002/rra.918

Alafifi, A.H., 2018. INTEGRATED SYSTEMS MODELING TO IMPROVE WATERSHED HABITAT MANAGEMENT AND DECISION MAKING.

Alafifi, A.H., Rosenberg, D.E., 2020. Systems modeling to improve river, riparian, and wetland habitat quality and area. Environmental Modelling and Software 126. https://doi.org/10.1016/j.envsoft.2020.104643

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS): Assessing the accuracy of distribution models. Journal of Applied Ecology 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araújo, M.B., Anderson, R.P., Barbosa, A.M., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B., O'hara, R.B., Zimmermann, N.E., Rahbek, C., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv 5, 4858–4874.

Argent, D.G., Bishop, J.A., Stauffer, J.R., Carline, R.F., Myers, W.L., 2003. Predicting freshwater fish distributions using landscape-level variables. Fisheries Research 60, 17–32. https://doi.org/10.1016/S0165-7836(02)00076-0

Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M., Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. Fisheries Research 62, 143–170. https://doi.org/10.1016/S0165-7836(02)00160-1

Bangs, M.R., Douglas, M.R., 2017. Anthropogenic Impacts Facilitate Native Fish Hybridization in the Bonneville Basin of Western North America. Transactions of the American Fisheries Society 146, 16–21. https://doi.org/10.1080/00028487.2016.1235611

Bangs, M.R., Marlis, D., R., Chafin, T.K., Douglas, M.E., 2020. Gene flow and species delimitation in fishes of Western North America: Flannelmouth (Catostomus latipinnis) and Bluehead sucker (C. Pantosteus discobolus). Ecology and Evolution 10, 6477. https://doi.org/10.1002/ece3.6384

Barber, R.A., Ball, S.G., Roger, |, Morris, K.A., Gilbert, F., 2021. Target-group backgrounds prove effective at correcting sampling bias in Maxent models. Diversity and Distributions 28, 128–41. https://doi.org/10.1111/ddi.13442

Barker, J.R., Pasternack, G.B., Bratovich, P.M., Massa, D.A., Wyrick, J.R., Johnson, T.R., 2018. Kayak drifter surface velocity observation for 2D hydraulic model validation. River Res Applic 34, 124–134. https://doi.org/10.1002/rra.3238

Barry, S, Elith, J, Barry, Simon, Elith, Jane, 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43, 413–423. https://doi.org/10.1111/j.1365-2664.2006.01136.x

Benkendorf, D.J., Hawkins, C.P., In Review. Correcting for the effects of class imbalance improves the performance of machine-learning based species distribution models. Ecological Modelling.

Berger, A.M., Deroba, J.J., Bosley, K.M., Goethel, D.R., Langseth, B.J., Schueller, A.M., Hanselman, D.H., 2021. Incoherent dimensionality in fisheries management: consequences of misaligned stock assessment and population boundaries. ICES Journal of Marine Science 78, 155–171. https://doi.org/10.1093/icesjms/fsaa203

Boavida, I., Dias, V., Ferreira, M.T., Santos, J.M., 2014. Univariate functions versus fuzzy logic: Implications for fish habitat modeling. Ecological Engineering 71, 533–538. https://doi.org/10.1016/j.ecoleng.2014.07.073

Boavida, I., Santos, J.M., Katopodis, C., Ferreira, M.T., Pinheiro, A., 2013. Uncertainty in predicting the fish-response to two-dimensional habitat modeling using field data. River Research and Applications 29, 1164–1174. https://doi.org/10.1002/rra.2603

Bocchiola, D., 2011. Hydraulic characteristics and habitat suitability in presence of woody debris: A flume experiment. Advances in Water Resources 34, 1304–1319. https://doi.org/10.1016/j.advwatres.2011.06.011

Booker, D.J., 2016. Generalized models of riverine fish hydraulic habitat. Journal of Ecohydraulics 1, 31–49. https://doi.org/10.1080/24705357.2016.1229141

Brenden, T.O., Clark, R.D., Cooper, A.R., Seelbach, P.W., Wang, L., Aichele, S.S., Bissell, E.G., Stewart, J.S., 2006. A GIS Framework for Collecting, Managing, and Analyzing Multiscale Landscape Variables across Large Regions for River Conservation and Management. American Fisheries Society Symposium 48, 49–74.

Brenden, T.O., Wang, L., Clark, R.D., Seelbach, P.W., Lyons, J., 2007. Comparison between Model-Predicted and Field-Measured Stream Habitat Features for Evaluating Fish Assemblage-Habitat Relationships. Transactions of the American Fisheries Society 136, 580–592. https://doi.org/10.1577/t05-284.1

Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27, 437–448. https://doi.org/10.1111/j.0906-7590.2004.03764.x

Budy, P., Conner, M.M., Salant, N.L., Macfarlane, W.W., 2015. An occupancy-based quantification of the highly imperiled status of desert fishes of the southwestern United States. Conservation Biology 29, 1142–1152. https://doi.org/10.1111/cobi.12513

Budy, P., Dahle, K., Thiede, G.P., 2006. An evaluation of the fish community of Cutler Reservoir and the Bear River above the reservoir with consideration of the potential for future fisheries enhancement.

Budy, P., Gaeta, J.W., 2017. Brown Trout as an Invader: A Synthesis of Problems and Perspectives in North America, in: Lobón-Cerviá, J., Sanz, N. (Eds.), Brown Trout. John Wiley & Sons, Ltd, Chichester, UK, pp. 523–543. https://doi.org/10.1002/9781119268352.ch20

Budy, P., Thiede, G.P., McHugh, P., 2007. Quantification of the Vital Rates, Abundance, and Status of a Critical, Endemic Population of Bonneville Cutthroat Trout. North American Journal of Fisheries Management 27, 593–604. https://doi.org/10.1577/M06-085.1

Budy, P., Thiede, G.P., McHugh, P., Hansen, E.S., Wood, J., 2008. Exploring the relative influence of biotic interactions and environmental conditions on the abundance and distribution of exotic brown trout (*Salmo trutta*) in a high mountain stream. Ecology of Freshwater Fish 17, 554–566. https://doi.org/10.1111/j.1600-0633.2008.00306.x

Budy, P., Wood, S., Roper, B., 2012. A Study of the Spawning Ecology and Early Life History Survival of Bonneville Cutthroat Trout. North American Journal of Fisheries Management 32, 436–449. https://doi.org/10.1080/02755947.2012.675945

Buraeu of Reclamation, 2021. Columbia River Basin | Water Reliability in the West - 2021 SECURE Water Act Report.

Cecina Babich Morrow, 2019. Thresholding species distribution models [WWW Document]. URL https://babichmorrowc.github.io/post/2019-04-12-sdm-threshold/

Creque, S.M., Rutherford, E.S., Zorn, T.G., 2005. Use of GIS-Derived Landscape-Scale Habitat Features to Explain Spatial Patterns of Fish Density in Michigan Rivers. North American Journal of Fisheries Management 25, 1411–1425. https://doi.org/10.1577/m04-121.1

D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, T. L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50, 885–900. https://doi.org/10.13031/2013.23153

D. N. Moriasi, M. W. Gitau, N. Pai, P. Daggupati, 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Trans. ASABE 58, 1763–1785. https://doi.org/10.13031/trans.58.10715

Dauwalter, D.C., Wenger, S.J., Gelwicks, K.R., Fesenmyer, K.A., 2011. Land use associations with distributions of declining native fishes in the upper Colorado River basin. Transactions of the American Fisheries Society 140, 646–658. https://doi.org/10.1080/00028487.2011.587753

DeRose, R.J., Bekker, M.F., Wang, S.-Y., Buckley, B.M., Kjelgren, R.K., Bardsley, T., Rittenour, T.M., Allen, E.B., 2015. A millennium-length reconstruction of Bear River stream flow, Utah. Journal of Hydrology 529, 524–534. https://doi.org/10.1016/j.jhydrol.2015.01.014

Dobos, M.E., Corsi, M.P., Schill, D.J., DuPont, J.M., Quist, M.C., 2016. Influences of summer water temperatures on the movement, distribution, and resource use of fluvial westslope cutthroat trout in the south fork clearwater river Basin. North American Journal of Fisheries Management 36, 549–567. https://doi.org/10.1080/02755947.2016.1141124

Dorji, T., Linke, S., Sheldon, F., 2020. Optimal model selection for Maxent: a case of freshwater species distribution modelling in Bhutan, a data poor country. https://doi.org/10.22541/au.160551779.93380163/v1

Dzara, J.R., Neilson, B.T., Null, S.E., 2019. Quantifying thermal refugia connectivity by combining temperature modeling, distributed temperature sensing, and thermal infrared imaging. Hydrol. Earth Syst. Sci. 23, 2965–2982. https://doi.org/10.5194/hess-23-2965-2019

Edwards, E.C., Null, S.E., 2019. The cost of addressing saline lake level decline and the potential for water conservation markets. Science of The Total Environment 651, 435–442. https://doi.org/10.1016/j.scitotenv.2018.09.006

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M., E. Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., Leathwick, J.R., 2009. Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Escalante, T., Rodríguez-Tapia, G., Linaje, M., Illoldi-Rangel, P., González-López, R., 2013. Identification of areas of endemism from species distribution models: threshold selection and Nearctic mammals. TIP 16, 5–17. https://doi.org/10.1016/S1405-888X(13)72073-4

Falke, J., 2006. Fish-habitat relations across spatial scales in prairie streams, in: American Fisheries Society Symposium 48. pp. 265–285.

Fausch, K.D., Torgersen, C.E., Baxter, C.V., Li, H.W., 2002. Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes. BioScience 52, 483. https://doi.org/10.1641/0006-3568(2002)052[0483:LTRBTG]2.0.CO;2

Ferro, C.A.T., Stephenson, D.B., 2011. Extremal Dependence Indices: Improved Verification Measures for Deterministic Forecasts of Rare Binary Events. Weather and Forecasting 26, 699–713. https://doi.org/10.1175/WAF-D-10-05030.1

Ficklin, D.L., Abatzoglou, J.T., Robeson, S.M., Null, S.E., Knouft, J.H., 2018. Natural and managed watersheds show similar responses to recent climate change. Proc. Natl. Acad. Sci. U.S.A. 115, 8553–8557. https://doi.org/10.1073/pnas.1801026115

Fourcade, Y., Engler, J.O., Rödder, D., Secondi, J., 2014. Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. PLoS ONE 9, e97122. https://doi.org/10.1371/journal.pone.0097122

Freeman, M.C., Bowen, Z.H., Grange, J.H., 1997. Transferability of Habitat Suitability Criteria for Fishes in Warmwater Streams. North American Journal of Fisheries Management 20–31.

Frissell, C.A., Liss, W.J., Warren, C.E., Hurley, M.D., 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. Environmental Management 10, 199–214. https://doi.org/10.1007/BF01867358

Giacomazzo, M., Bertolo, A., Brodeur, P., Massicotte, P., Goyette, J.O., Magnan, P., 2020. Linking fisheries to land use: How anthropogenic inputs from the watershed shape fish habitat quality. Science of the Total Environment 717. https://doi.org/10.1016/j.scitotenv.2019.135377

Goodrum, G.C., 2020. Improving aquatic habitat representation in Utah using large spatial scale environmental datasets.

Goodrum, G.C., Null, S.E., 2022. Reduced complexity models for regional aquatic habitat suitability assessment. J American Water Resour Assoc 1752–1688.13077. https://doi.org/10.1111/1752-1688.13077

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hall, S.A., Adam, J.C., Yourek, M.A., Whittemore, A.M., Yorgey, G.G., Scarpare, F., Liu, M., McLarty, S., Asante-Sasu, C., McClure, S., Turk, J., Haller, D., Padowski, J., Deshar, R., Brady, M.P., Rajagopalan, K., Barber, M.E., Weber, R., Stockle, C.O., Goodspeed, H.L., Gustine, R.N., Kondal, A., Yoder, J., Deaver, B., Downes, M., Tarbutton, S., Callahan, M., M., Price, P. Roberts, T., Stephens, J., Valdez, 2022. 2021 Columbia River Basin Long Term Water Supply and Demand Forecast (No. 21-12– 006). Washington Department of Ecology, Olympia, WA.

Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2022. dismo: Species Distribution Modeling.

Hillyard, R.W., Keeley, E.R., 2012. Temperature-related changes in habitat quality and use by bonneville cutthroat trout in regulated and unregulated river segments.

Transactions of the American Fisheries Society 141, 1649–1663. https://doi.org/10.1080/00028487.2012.717517

Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., 2002. ECOLOGICAL-NICHE FACTOR ANALYSIS: HOW TO COMPUTE HABITAT-SUITABILITY MAPS WITHOUT ABSENCE DATA? Ecology 83, 2027–2036.

Hopken, M.W., Douglas, M.R., Douglas, M.E., 2013. Stream hierarchy defines riverscape genetics of a North American desert fish. Molecular Ecology 22, 956–971. https://doi.org/10.1111/mec.12156

Idaho Department of Fish and Game, 2016. Species of Greatest Conservation Need [WWW Document]. URL https://idfg.idaho.gov/species/taxa/list/sgcn

Jager, H.I., Smith, B.T., 2008. Sustainable reservoir operation: can we generate hydropower and preserve ecosystem values? River Res. Applic. 24, 340–352. https://doi.org/10.1002/rra.1069

Johnstone, H.C., 2000. Temperature tolerances and habitat conditions for Bonneville cutthroat trout in the Thomas Fork of the Bear River, Wyoming. Stationery Office.

Johnstone, H.C., Rahel, F.J., 2003. Assessing Temperature Tolerance of Bonneville Cutthroat Trout Based on Constant and Cycling Thermal Regimes. Transactions of the American Fisheries Society 132, 92–99.

Jorde, K., Schneider, M., Peter, A., Zoellner, F., 2001. Fuzzy based Models for the Evaluation of Fish Habitat Quality and Instream Flow Assessment, in: International Symposium on Environmental Hydraulics.

Jowett, I.G., Davey, A.J.H., 2007. A Comparison of Composite Habitat Suitability Indices and Generalized Additive Models of Invertebrate Abundance and Fish Presence–Habitat Availability. Transactions of the American Fisheries Society 136, 428–444. https://doi.org/10.1577/T06-104.1

Joy, M.K., Death, R.G., 2004. Predictive modelling and spatial mapping of freshwater fish and decapod assemblages using GIS and neural networks. Freshwater Biology 49, 1036–1052. https://doi.org/10.1111/j.1365-2427.2004.01248.x

Kearney, M., 2006. Habitat, environment and niche: what are we modelling? Oikos 115, 186–191. https://doi.org/10.1111/j.2006.0030-1299.14908.x

Kirk, R.W.V., Contor, B.A., Morrisett, C.N., Null, S.E., Loibman, A.S., 2020. Potential for Managed Aquifer Recharge to Enhance Fish Habitat in a Regulated River. Water 12, 673. https://doi.org/10.3390/w12030673

Knudby, A., Brenning, A., LeDrew, E., 2010. New approaches to modelling fish-habitat relationships. Ecological Modelling 221, 503–511. https://doi.org/10.1016/j.ecolmodel.2009.11.008

Komac, B., Esteban, P., Trapero, L., Caritg, R., 2016. Modelization of the Current and Future Habitat Suitability of Rhododendron ferrugineum Using Potential Snow Accumulation. PLoS ONE 11, e0147324. https://doi.org/10.1371/journal.pone.0147324

Kraft, M., Rosenberg, D.E., Null, S.E., 2019. Prioritizing Stream Barrier Removal to Maximize Connected Aquatic Habitat and Minimize Water Scarcity. Journal of the American Water Resources Association 55, 382–400. https://doi.org/10.1111/1752-1688.12718

Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., Macdonald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J.W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H., Wilting, A., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. Diversity Distrib. 19, 1366–1379. https://doi.org/10.1111/ddi.12096

Kristensen, E.A., Baattrup-Pedersen, A., Andersen, H.E., 2012. Prediction of stream fish assemblages from land use characteristics: implications for cost-effective design of monitoring programmes. Environ Monit Assess 184, 1435–1448. https://doi.org/10.1007/s10661-011-2052-4

Lamouroux, N., Capra, H., 2002. Simple predictions of instream habitat model outputs for target fish populations: Habitat modelling for fish populations. Freshwater Biology 47, 1543–1556. https://doi.org/10.1046/j.1365-2427.2002.00879.x

Lamouroux, N., Capra, H., Pouilly, M., 1998. Predicting habitat suitability for lotic fish: linking statistical hydraulic models with multivariate habitat use models. Regul. Rivers: Res. Mgmt. 14, 1–11. https://doi.org/10.1002/(SICI)1099-1646(199801/02)14:1<1::AID-RRR472>3.0.CO;2-D

Leroy, B., Delsol, R., Hugueny, B., Meynard, C.N., Barhoumi, C., Barbet-Massin, M., Bellard, C., 2018. Without quality presence-absence data, discrimination metrics such as TSS can be misleading measures of model performance. J Biogeogr 45, 1994–2002. https://doi.org/10.1111/jbi.13402

Li, W., Guo, Q., Elkan, C., 2011. Can we model the probability of presence of species without absence data? Ecography 34, 1096–1105. https://doi.org/10.1111/j.1600-0587.2011.06888.x

Li, X., Wang, Y., 2013. Applying various algorithms for species distribution modelling. Integrative Zoology 8, 124–135. https://doi.org/10.1111/1749-4877.12000

Liu, C., Newell, G., White, M., Correspondence, C., Liu, A., 2016. On the selection of thresholds for predicting species occurrence with presence-only data. Ecology and Evolution 6, 337–348. https://doi.org/10.1002/ece3.1878

Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeography 17, 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Loktefff, R.L., 2014. MOVEMENT AND HABITAT USE OF BONNEVILLE CUTTHROAT TROUT (ONCORHYNCHUS CLARKI UTAH): A CASE STUDY IN THE TEMPLE FORK WATERSHED.

Mäki-Petäys, A., Muotka, T., Huusko, A., Tikkanen, P., Kreivi, P., 1997. Seasonal changes in habitat use and preference by juvenile brown trout, Salmo trutta, in a northern boreal river. Canadian Journal of Fisheries and Aquatic Sciences 54, 520–530.

Maloney, B.C., 2017. Evaluating Habitat-Based Niche Requirements and Potential Recruitment Bottlenecks for Imperiled Bluehead Sucker (Catostomus discobolus).

Manel, S., Ceri Williams, H., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38, 921–931.

Maxwell, J.R., Edwards, C.J., Jensen, M.E., Paustian, S.J., Parrott, H., Hill, D.M., 1995. Framework of Aquatic Ecological Units in North America (Nearctic Zone). United States Department of Agriculture, Forest Service.

McKenna, J.E., Johnson, J.H., 2011. Landscape models of brook trout abundance and distribution in lotic habitat with field validation. North American Journal of Fisheries Management 31, 742–756. https://doi.org/10.1080/02755947.2011.593940

Meixler, M.S., 2021. A species-specific fish passage model based on hydraulic conditions and water temperature. Ecological Informatics 65. https://doi.org/10.1016/j.ecoinf.2021.101407

Meixler, M.S., Bain, M.B., 2012. A GIS framework for fish habitat prediction at the river basin scale. International Journal of Ecology. https://doi.org/10.1155/2012/146073

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Moir, H.J., Gibbins, C.N., Soulsby, C., Youngson, A.F., 2005. PHABSIM modelling of Atlantic salmon spawning habitat in an upland stream: testing the influence of habitat suitability indices on model output. River Res. Applic. 21, 1021–1034. https://doi.org/10.1002/rra.869

Mouton, A.M., Schneider, M., Depestele, J., Goethals, P.L.M., De Pauw, N., 2007. Fish habitat modelling as a tool for river management. Ecological Engineering 29, 305–315. https://doi.org/10.1016/j.ecoleng.2006.11.002

Mugodo, J., Kennard, M., Liston, P., Nichols, S., Linke, S., Norris, R.H., Lintermans, M., 2006. Local stream habitat variables predicted from catchment scale characteristics are useful for predicting fish distribution. Hydrobiologia 572, 59–70. https://doi.org/10.1007/s10750-006-0252-7

National Drought Mitigation Center, 2023. U.S. Drought Monitor [WWW Document]. URL https://droughtmonitor.unl.edu/ (accessed 4.10.23).

Null, S., Wurtsbaugh, W., 2020. Water Development, Consumptive Water Uses, and Great Salt Lake, in: Great Salt Lake Biology: A Terminal Lake in a Time of Change. Springer, Netherlands. https://doi.org/10.1007/978-3-030-40352-2

Null, S.E., Farshid, A., Goodrum, G., Gray, C.A., Lohani, S., Morrisett, C.N., Prudencio, L., Sor, R., 2020. A Meta-Analysis of Environmental Tradeoffs of Hydropower Dams in the Sekong, Sesan, and Srepok (3S) Rivers of the Lower Mekong Basin. Water 13, 63. https://doi.org/10.3390/w13010063

Null, S.E., Olivares, M.A., Cordera, F., Lund, J.R., 2021. Pareto Optimality and Compromise for Environmental Water Management. Water Resources Research 57. https://doi.org/10.1029/2020WR028296

Olden, J.D., Jackson, D.A., Peres-Neto, P.R., 2002. Predictive Models of Fish Species Distributions: A Note on Proper Validation and Chance Predictions. Transactions of the American Fisheries Society 131, 329–336. https://doi.org/10.1577/1548-8659(2002)131<0329:pmofsd>2.0.co;2

Olson, T., 2022. Section 401 Water Quality Certification No. DWQ-2022-07001.

Oregon State University, 2002. . Streamflow Evaluation for Watershed Restoration Planning and Design. URL https://streamflow.engr.oregonstate.edu/analysis/index.htm

Palialexis, A., Georgakarakos, S., Karakassis, I., Lika, K., Valavanis, V.D., 2011. Prediction of marine species distribution from presence-absence acoustic data: Comparing the fitting efficiency and the predictive capacity of conventional and novel distribution models. Hydrobiologia 670, 241–266. https://doi.org/10.1007/s10750-011-0673-9

Parasiewicz, P., Dunbar, M.J., 2001. Physical habitat modelling for fish - a developing approach. Large Rivers 12, 239–268.

Peduzzi, P., Concato, J., Feinstein, A.R., Holford, T.R., 1995. Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. Journal of Clinical Epidemiology 48, 1503–1510. https://doi.org/10.1016/0895-4356(95)00048-8

Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R., 1996. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology 49, 1373–1379. https://doi.org/10.1016/S0895-4356(96)00236-3

Peterson, A.T., Vieglais, D.A., 2001. Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem. BioScience 51, 363. https://doi.org/10.1641/0006-3568(2001)051[0363:PSIUEN]2.0.CO;2

Phillips, S., 2012. Inferring prevalence from presence-only data: a response to 'Can we model the probability of presence of species without absence data?' Ecography 35, 385–387. https://doi.org/10.1111/j.1600-0587.2011.07285.x

Phillips, S.J., 2017. A Brief Tutorial on Maxent [WWW Document]. URL http://biodiversityinformatics.amnh.org/open_source/maxent/

Phillips, S.J., Anderson, R.P., Dudik, M., Schapire, R.E., Blair, M.E., 2017. Opening the black box: an open-source release of MaxEnt. Ecography 40, 887–893. https://doi.org/10.1111/ecog.03049

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–252. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., Dudık, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 161–175.

Phillips, S.J., Elith, J., 2013. On estimating probability of presence from use-availability or presence-background data. Ecology 94, 1409–1419.

Ptacek, J.A., Rees, D.E., Miller, W.J., 2005. Bluehead Sucker (Catostomus discobolus): A Technical Conservation Assessment. USDA Forest Service, Rocky Mountain Region.

R Core Team, 2021. R: A language and environment for statistical computing.

Radosavljevic, A., Anderson, R.P., 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. J. Biogeogr. 41, 629–643. https://doi.org/10.1111/jbi.12227

Richards, C., Johnson, L.B., Host, G.E., 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53.

Rosenfeld, J.S., Campbell, K., Leung, E.S., Bernhardt, J., Post, J., 2011. Habitat effects on depth and velocity frequency distributions: Implications for modeling hydraulic variation and fish habitat suitability in streams. Geomorphology 130, 127–135. https://doi.org/10.1016/j.geomorph.2011.03.007

Ryo, M., Angelov, B., Mammola, S., Kass, J.M., Benito, B.M., Hartig, F., 2021. Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models. Ecography 44, 199–205. https://doi.org/10.1111/ecog.05360

Schmidt, H., Radinger, J., Teschlade, D., Stoll, S., 2020. The role of spatial units in modelling freshwater fish distributions: Comparing a subcatchment and river network approach using MaxEnt. Ecological Modelling 418, 108937. https://doi.org/10.1016/j.ecolmodel.2020.108937

Shcheglovitova, M., Anderson, R.P., 2013. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. Ecological Modelling 269, 9–17. https://doi.org/10.1016/j.ecolmodel.2013.08.011

Smith, B., 2022. Bluehead Sucker Distribution Data.

Smith, T.A., Kraft, C.E., 2005. Stream Fish Assemblages in Relation to Landscape Position and Local Habitat Variables. Transactions of the American Fisheries Society 134, 430–440. https://doi.org/10.1577/t03-051.1

Sofaer, H.R., Jarnevich, C.S., Pearse, I.S., Smyth, R.L., Auer, S., Cook, G.L., Edwards, T.C., Guala, G.F., Howard, T.G., Morisette, J.T., Hamilton, H., 2019. Development and Delivery of Species Distribution Models to Inform Decision-Making. BioScience 69, 544–557. https://doi.org/10.1093/biosci/biz045

Souza, A.C. de, Weber, M. de M., Prevedello, J.A., 2022. Protection status and density-dependent effects mediate the abundance-suitability relationship of a threatened species. Perspectives in Ecology and Conservation 20, 168–176. https://doi.org/10.1016/j.pecon.2022.03.002

Steen, P.J., Passino-Reader, D.R., Wiley, M.J., 2006. Modeling brook trout presence and absence from landscape variables using four different analytical methods, in: American Fisheries Society Symposium 48. pp. 513–531.

Steen, P.J., Zorn, T.G., Seelbach, P.W., Schaeffer, J.S., 2008a. Classification Tree Models for Predicting Distributions of Michigan Stream Fish from Landscape Variables. Transactions of the American Fisheries Society 137, 976–996. https://doi.org/10.1577/T07-119.1

Steen, P.J., Zorn, T.G., Seelbach, P.W., Schaeffer, J.S., 2008b. Classification Tree Models for Predicting Distributions of Michigan Stream Fish from Landscape Variables. Transactions of the American Fisheries Society 137, 976–996. https://doi.org/10.1577/T07-119.1

Stephens, P.A., Pettorelli, N., Barlow, J., Whittingham, M.J., Cadotte, M.W., 2015. Management by proxy? The use of indices in applied ecology. J Appl Ecol 52, 1–6. https://doi.org/10.1111/1365-2664.12383

Suding, K.N., Hobbs, R.J., 2009. Threshold models in restoration and conservation: a developing framework. Trends in Ecology & Evolution 24, 271–279. https://doi.org/10.1016/j.tree.2008.11.012

Taylor, A.T., Hafen, T., Holley, C.T., González, A., Long, J.M., 2020. Spatial sampling bias and model complexity in stream-based species distribution models: A case study of Paddlefish (*Polyodon spathula*) in the Arkansas River basin, USA. Ecol Evol 10, 705–717. https://doi.org/10.1002/ece3.5913

The Bear River Watershed Information System [WWW Document], 2017. URL https://bearriverinfo.org/

Thompson, P., 2016. Weber River Partnership Protects World-Class Fishery [WWW Document]. Utah Department of Environmental Quality: Water Quality. URL https://deq.utah.gov/communication/news/weber-river-partnership-trout-fishery

Thompson, Paul, 2015. Bonneville Cutthroat Trout and Bluehead Sucker in the Weber River: Endangered Species Act Implications?

Thompson, P.D., Burnett, P.C., 2019. The Weber River Partnership: How Fish Gained Relevance through a Recently Formed Watershed Group. American Fisheries Society Symposium 91, 565–588.

Toth, R.E., Baker, J.B., Bryner, C.L., Evans, J., Hinman, K.E., Kilpatrick, K.R., Seegmiller, K., 2005. Alternative Futures for the Bear River Watershed - Final Project Report No. 2005-1. College of Natural Resources, Utah State University, Logan, Utah.

Toth, R.E., Edwards, T.J., Perschon, A.L., White, D.C., 2009. Bear River Watershed: Its Role in Maintaining the Bear River Migratory Bird Refuge - Final Project Report No. 2010-1. U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah.

Unmack, P.J., Dowling, T.E., Laitinen, N.J., Secor, C.L., Mayden, R.L., 2014. Influence of Introgression and Geological Processes on Phylogenetic Relationships of Western North American Mountain Suckers (Pantosteus, Catostomidae). PLoS ONE 9, 90061. https://doi.org/10.1371/journal.pone.0090061

U.S. Fish and Wildlife Service, 2013. Land protection plan—Bear River Watershed Conservation Area. U.S. Department of the Interior, U.S. Fish and Wildlife Service, Regions 1 and 6, Lakewood, CO.

U.S. Geological Survey, 2016. National Water Information System data available on the World Wide Web (USGS Water Data for the Nation) [WWW Document]. URL https://waterdata.usgs.gov/monitoring-location/10126000/#parameterCode=00065&period=P7D (accessed 4.1.23).

Utah Water Science Center, 2018. Baseflow [WWW Document]. URL https://www.usgs.gov/centers/utah-water-science-center/science/baseflow (accessed 4.10.23).

Utah Wildlife Action Plan Joint Team, 2015. Utah Wildlife Action Plan: A plan for managing native wildlife species and their habitats to help prevent listing under the Endangered Species Act. (No. 15–14). Utah Division of Wildlife Resources, Salt Lake City, Utah, USA.

Vadas, R.L., Orth, D.J., 2001. Formulation of Habitat Suitability Models for Stream Fish Guilds: Do the Standard Methods Work? Transactions of the American Fisheries Society 130, 217–235.

Valavanis, V.D., Pierce, G.J., Zuur, A.F., Palialexis, A., Saveliev, A., Katara, I., Wang, J., 2008a. Modelling of essential fish habitat based on remote sensing, spatial analysis and GIS. Hydrobiologia 612, 5–20. https://doi.org/10.1007/s10750-008-9493-y

Valavanis, V.D., Pierce, G.J., Zuur, A.F., Palialexis, A., Saveliev, A., Katara, I., Wang, J., 2008b. Modelling of essential fish habitat based on remote sensing, spatial analysis and GIS. Hydrobiologia 612, 5–20. https://doi.org/10.1007/s10750-008-9493-y

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., 2021. Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. Ecological Monographs 92. https://doi.org/10.1002/ecm.1486

Van Sickle, J., Baker, J., Herlihy, A., Bayley, P., Gregory, S., Haggerty, P., Ashkenas, L., Li, J., 2004. PROJECTING THE BIOLOGICAL CONDITION OF STREAMS UNDER ALTERNATIVE SCENARIOS OF HUMAN LAND USE. Ecological Applications 368, 368–380.

Vezza, P., Muñoz-Mas, R., Martinez-Capel, F., Mouton, A., 2015. Random forests to evaluate biotic interactions in fish distribution models. Environmental Modelling and Software 67, 173–183. https://doi.org/10.1016/j.envsoft.2015.01.005

Vignali, S., Barras, A.G., Arlettaz, R., Braunisch, V., 2020a. *SDMtune* : An R package to tune and evaluate species distribution models. Ecology and Evolution 10, 11488–11506. https://doi.org/10.1002/ece3.6786

Vignali, S., Barras, A.G., Arlettaz, R., Braunisch, V., 2020b. *SDMtune* : An R package to tune and evaluate species distribution models. Ecology and Evolution 10, 11488–11506. https://doi.org/10.1002/ece3.6786

Vittinghoff, E., McCulloch, C.E., 2007. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. American Journal of Epidemiology 165, 710–718. https://doi.org/10.1093/aje/kwk052

Walsworth, T.E., Budy, P., 2015. Integrating Nonnative Species in Niche Models to Prioritize Native Fish Restoration Activity Locations along a Desert River Corridor. Transactions of the American Fisheries Society 144, 667–681. https://doi.org/10.1080/00028487.2015.1024333

Ward, G., Hastie, T., Barry, S., Elith, J., Leathwick, J.R., 2009. Presence-Only Data and the EM Algorithm. Biometrics 65, 554–563.

Webber, P.A., Thompson, P.D., Budy, P., 2012. Status and Structure of Two Populations of the Bluehead Sucker (*Catostomus discobolus*) in the Weber River, Utah. The Southwestern Naturalist 57, 267–276. https://doi.org/10.1894/0038-4909-57.3.267

Weber, M.M., Stevens, R.D., Diniz-Filho, J.A.F., Grelle, C.E.V., 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. Ecography 40, 817–828. https://doi.org/10.1111/ecog.02125

Western Division of the American Fisheries Society, 2011a. Bonneville cutthroat (Oncorhynchus clarkii utah) distribution and status by HUC8 [WWW Document]. Data Basin. URL https://databasin.org/datasets/b69c229ea9f24f6d9b191bdce1a13b4e/

Western Division of the American Fisheries Society, 2011b. Bluehead sucker (Catostomus discobolus) distribution and status by HUC8 [WWW Document]. Data Basin. URL https://databasin.org/datasets/ffaab9efcfb241fea1c441bc607c0577/

Wheaton, J.M., Bouwes, N., Mchugh, P., Saunders, C., Bangen, S., Bailey, P., Nahorniak, M., Wall, E., Jordan, C., 2018. Upscaling site-scale ecohydraulic models to inform salmonid population-level life cycle modeling and restoration actions - Lessons from the Columbia River Basin: Upscaling Ecohydraulic models. Earth Surf. Process. Landforms 43, 21–44. https://doi.org/10.1002/esp.4137

Wheaton, J.M., Pasternack, G.B., Merz, J.E., 2004. Use of habitat heterogeneity in salmonid spawning habitat rehabilitation design. Presented at the Fifth International Symposium on Ecohydraulics: Aquatic Habitats: Analysis and Restoration, Madrid, Spain.

White, S.M., Rahel, F.J., 2008. Complementation of Habitats for Bonneville Cutthroat Trout in Watersheds Influenced by Beavers, Livestock, and Drought. Transactions of the American Fisheries Society 137, 881–894. https://doi.org/10.1577/t06-207.1

Williams, A.P., Cook, E.R., Smerdon, J.E., Cook, B.I., Abatzoglou, J.T., Bolles, K., Baek, S.H., Badger, A.M., Livneh, B., 2020. Large contribution from anthropogenic warming to an emerging North American megadrought. Science 368, 314–318. https://doi.org/10.1126/science.aaz9600

Williams, R.D., Brasington, J., Hicks, M., Measures, R., Rennie, C.D., Vericat, D., 2013. Hydraulic validation of two-dimensional simulations of braided river flow with spatially continuous aDcp data. Water Resour. Res. 49, 5183–5205. https://doi.org/10.1002/wrcr.20391

Worthington, T.A., Zhang, T., Logue, D.R., Mittelstet, A.R., Brewer, S.K., 2016. Landscape and flow metrics affecting the distribution of a federally-threatened fish: Improving management, model fit, and model transferability. Ecological Modelling 342, 1–18. https://doi.org/10.1016/j.ecolmodel.2016.09.016

Wunderlich, R.F., Lin, Y.-P., Anthony, J., Petway, J.R., 2019. Two alternative evaluation metrics to replace the true skill statistic in the assessment of species distribution models. NC 35, 97–116. https://doi.org/10.3897/natureconservation.35.33918

Wurtsbaugh, W.A., Heredia, N.A., Laub, B.G., Meredith, C.S., Mohn, H.E., Null, S.E., Pluth, D.A., Roper, B.B., Carl Saunders, W., Stevens, D.K., Walker, R.H., Wheeler, K., 2014. Approaches for studying fish production: Do river and lake researchers have different perspectives? Canadian Journal of Fisheries and Aquatic Sciences 72, 149–160. https://doi.org/10.1139/cjfas-2014-0210

Wyoming Game and Fish Department, 2017. Wyoming State Wildlife Action Plan-2017 Bear River Basin. Wyoming Game and Fish Department.

Yi, Y., Cheng, X., Wieprecht, S., Tang, C., 2014. Comparison of habitat suitability models using different habitat suitability evaluation methods. Ecological Engineering 71, 335–345. https://doi.org/10.1016/j.ecoleng.2014.07.034

Yi, Y., Cheng, X., Yang, Z., Wieprecht, S., Zhang, S., Wu, Y., 2017. Evaluating the ecological influence of hydraulic projects: A review of aquatic habitat suitability models. Renewable and Sustainable Energy Reviews 68, 748–762. https://doi.org/10.1016/j.rser.2016.09.138

Zorn, T.G., Seelbach, P.W., Wiley, M.J., 2002. Distributions of Stream Fishes and their Relationship to Stream Size and Hydrology in Michigan’s Lower Peninsula. Transactions of the American Fisheries Society 131, 70–85. https://doi.org/10.1577/1548-8659(2002)131<0070:dosfat>2.0.co;2

APPENDIX: SUPPLEMENTAL FIGURES

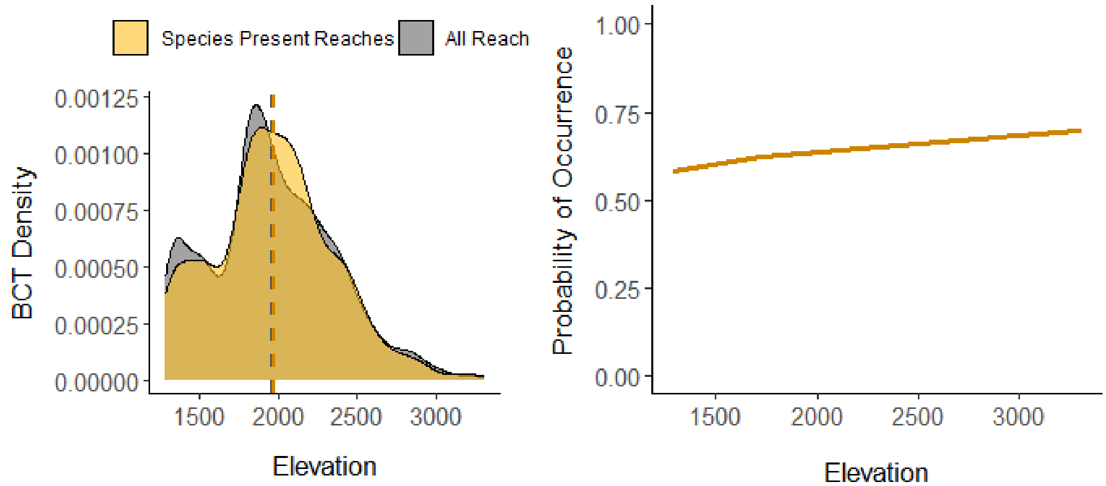

Fig. A1. Elevation in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different elevations across all months and ensemble iterations, with a 95% confidence interval (right).

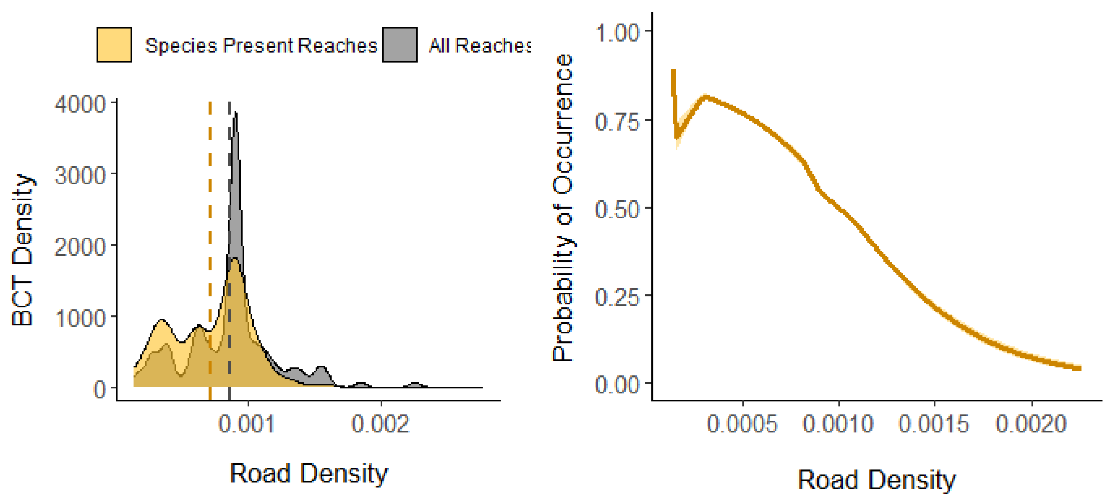

Fig. A2. Road density in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different road densities across all months and ensemble iterations, with a 95% confidence interval (right).

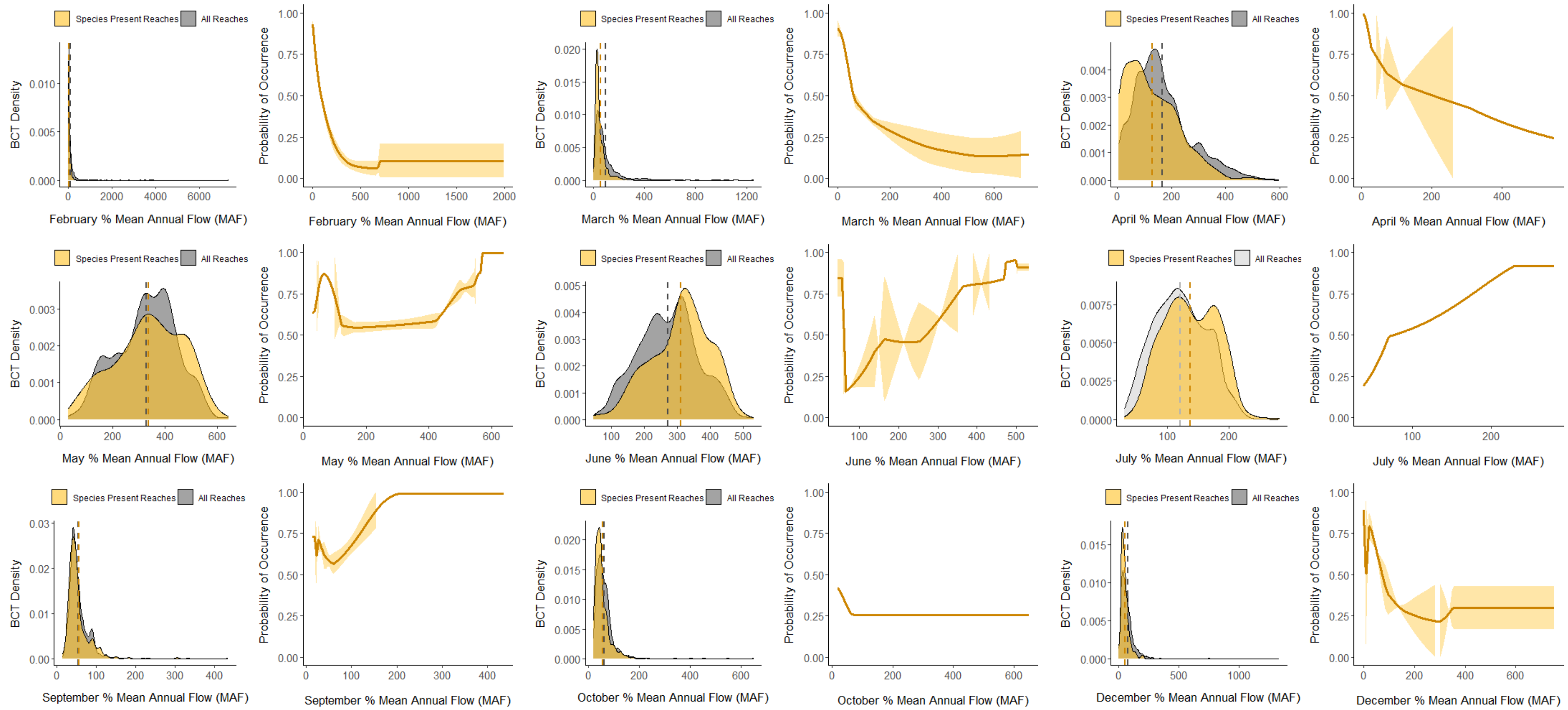

Fig. A3. Percent mean annual flow (calculated using monthly averages for streamflow) in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months selected for use in model fitting in at least one ensemble iteration.

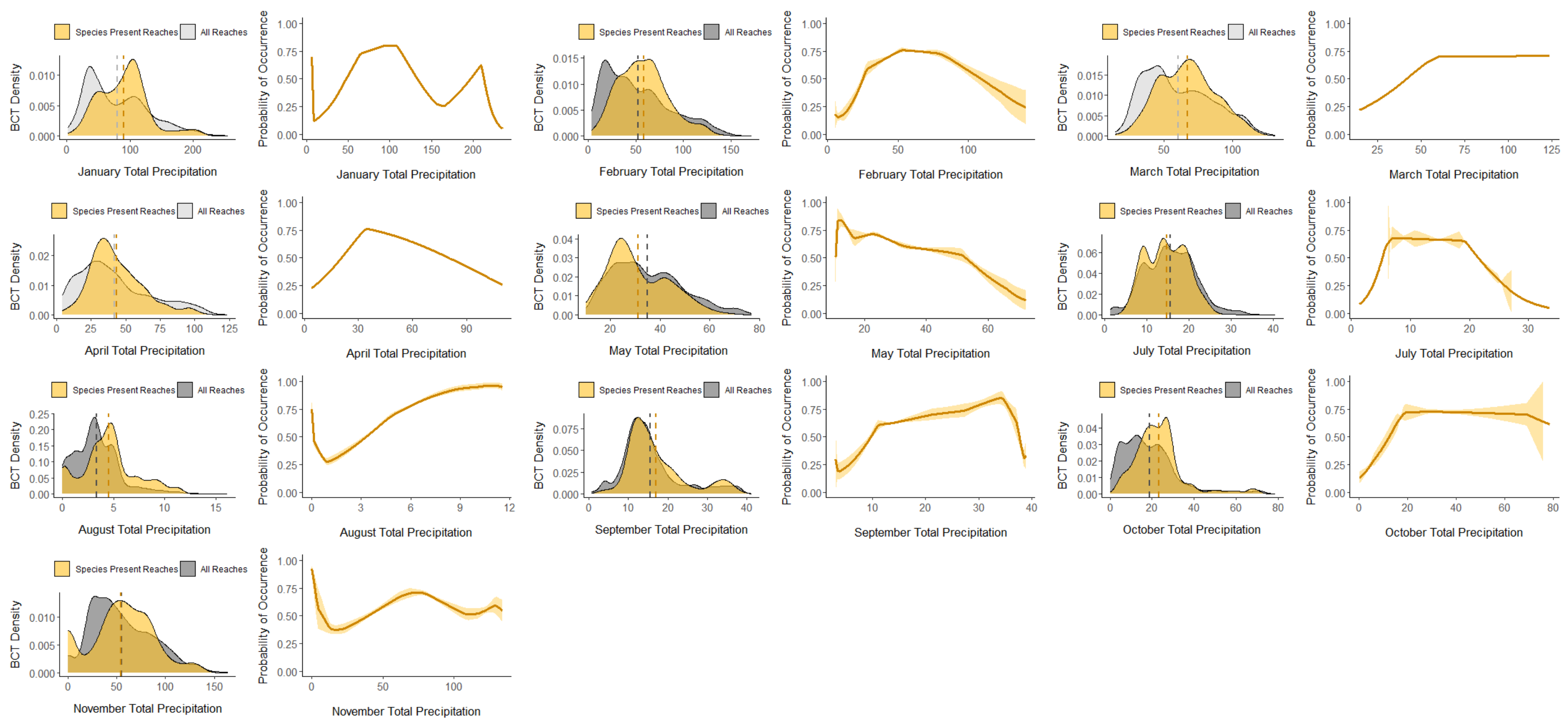

Fig. A4. Monthly precipitation totals in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months selected for use in model fitting in at least one ensemble iteration.

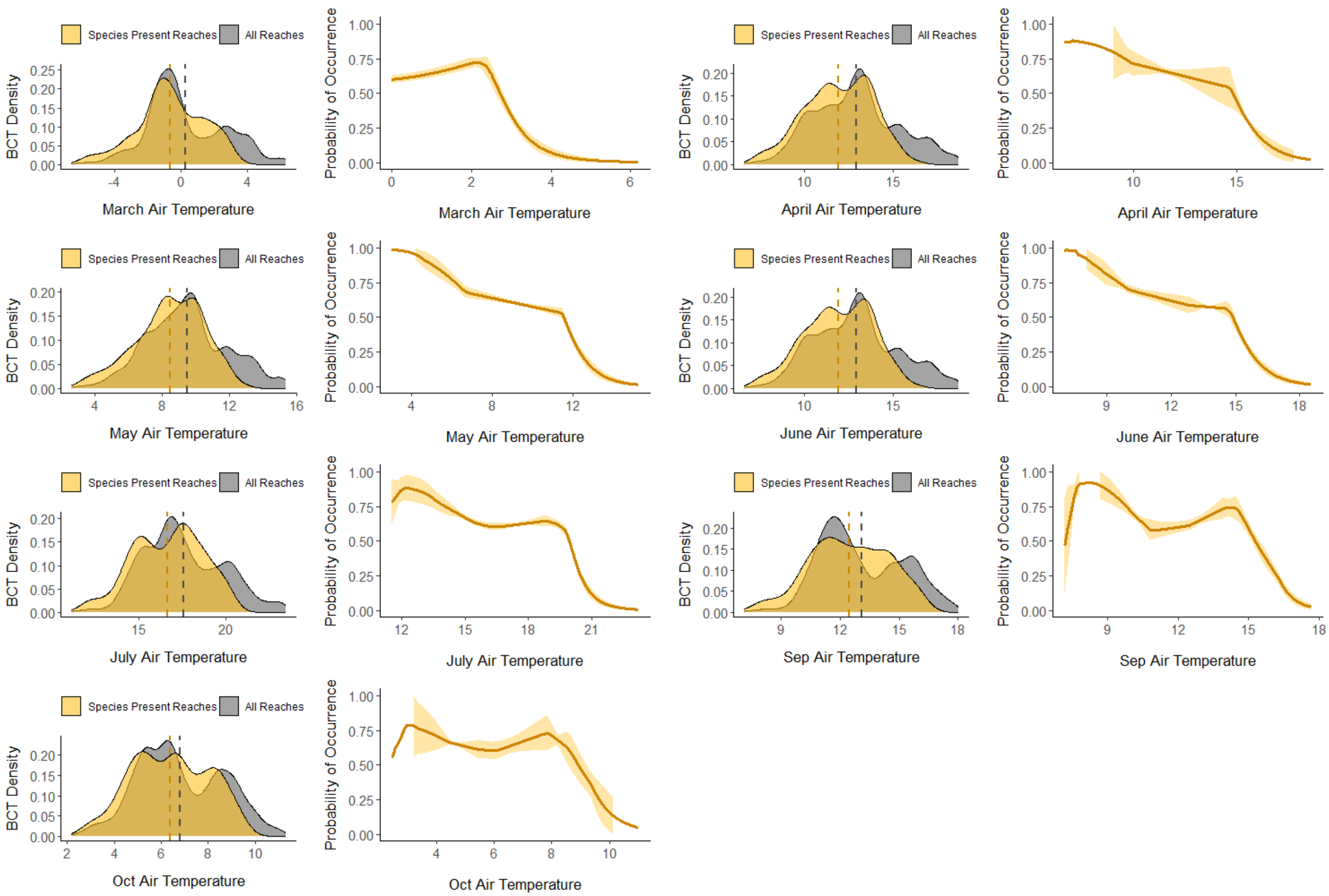

Fig. A5. Average monthly air temperature in Bonneville Cutthroat Trout (BCT) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months that were selected for use in model fitting in at least one ensemble iteration.

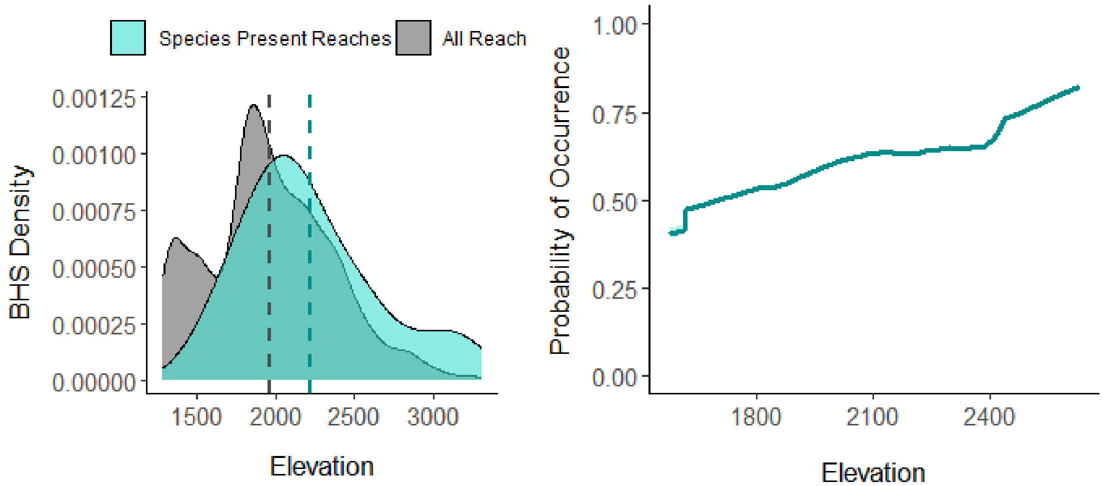

Fig. A6. Elevation in Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different elevations across all months and ensemble iterations, with a 95% confidence interval (right).

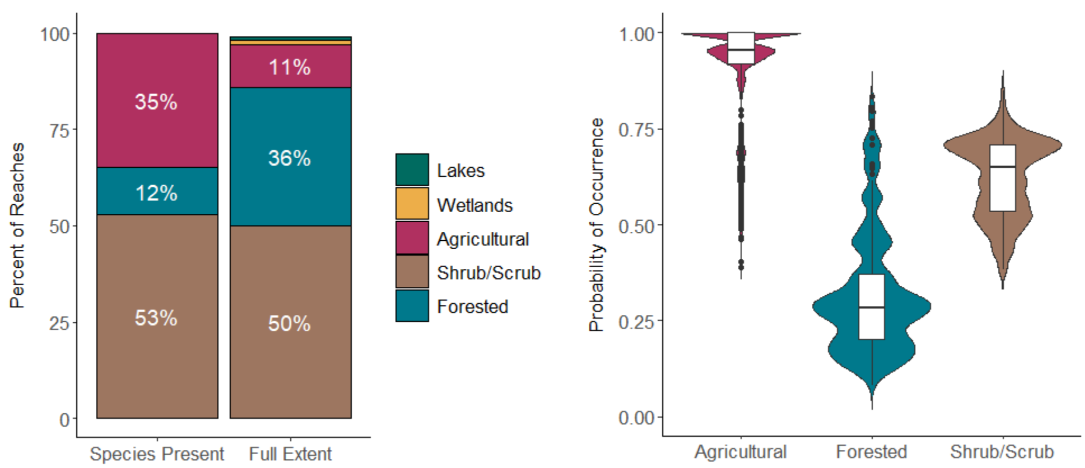

Fig. A7. Land use in Bluehead Sucker (BHS) occupied reaches compared to all reaches and MaxEnt probabilities of occurrence by land use category for all months and ensemble iterations.

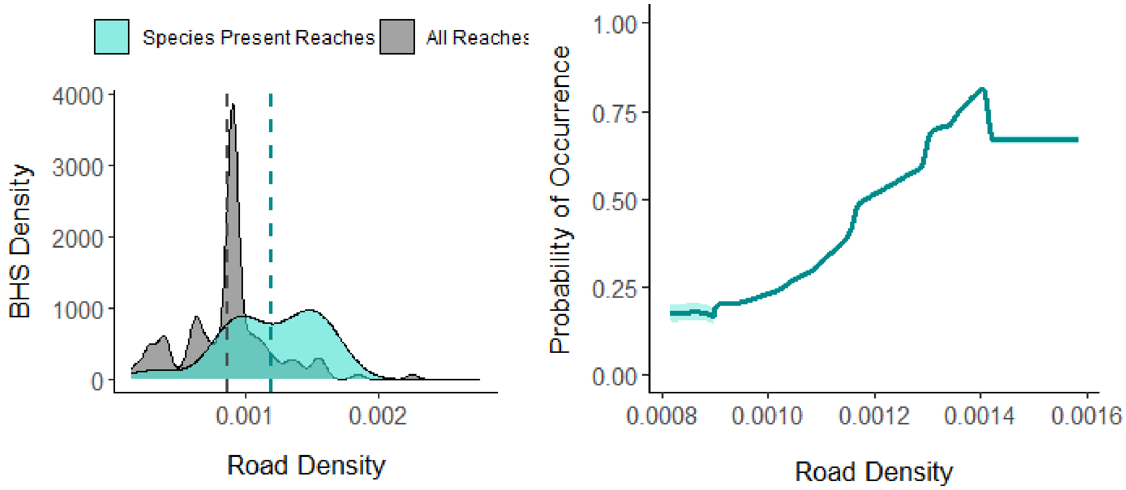

Fig. A8. Road density in Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different road densities across all months and ensemble iterations, with a 95% confidence interval (right).

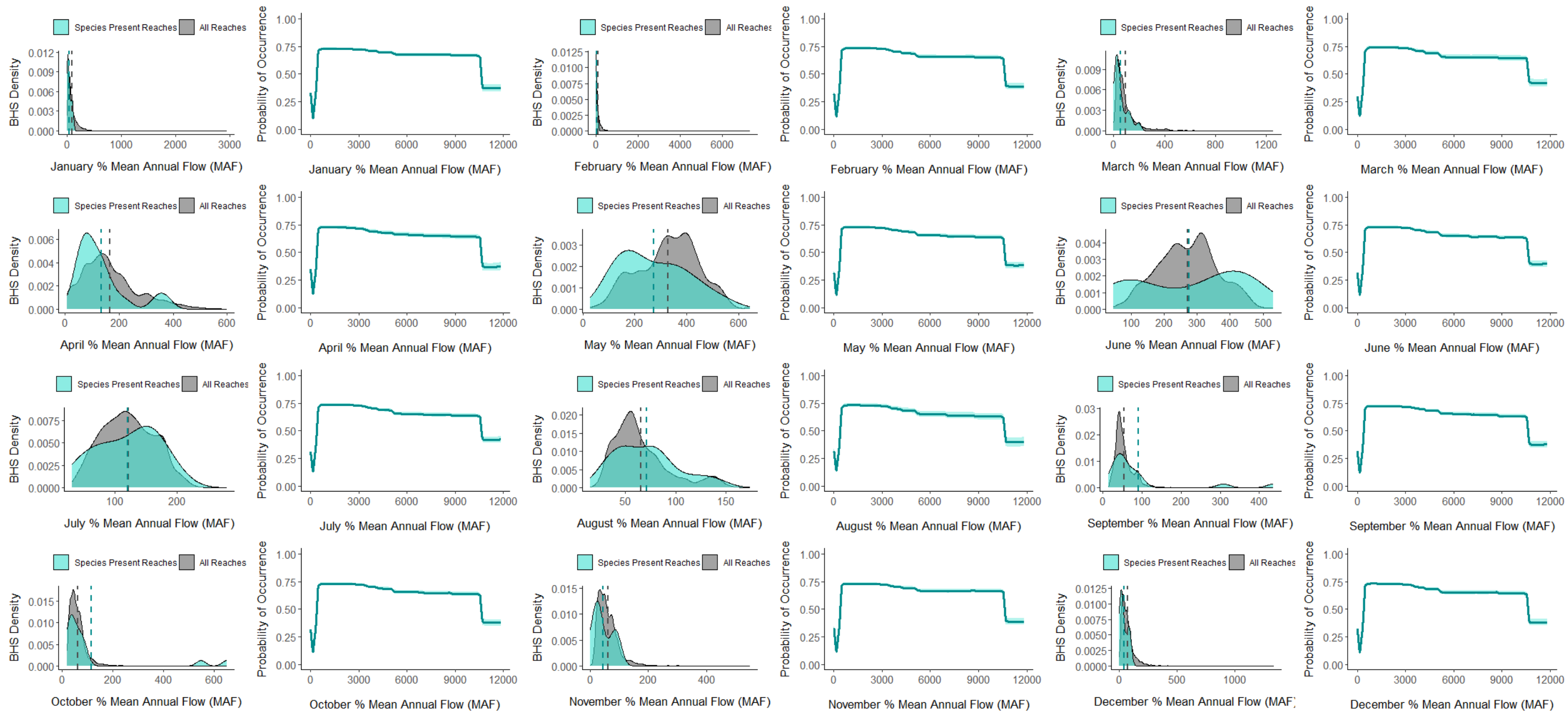

Fig. A9. Percent mean annual flow (calculated using monthly averages for streamflow) in Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months selected for use in model fitting in at least one ensemble iteration.

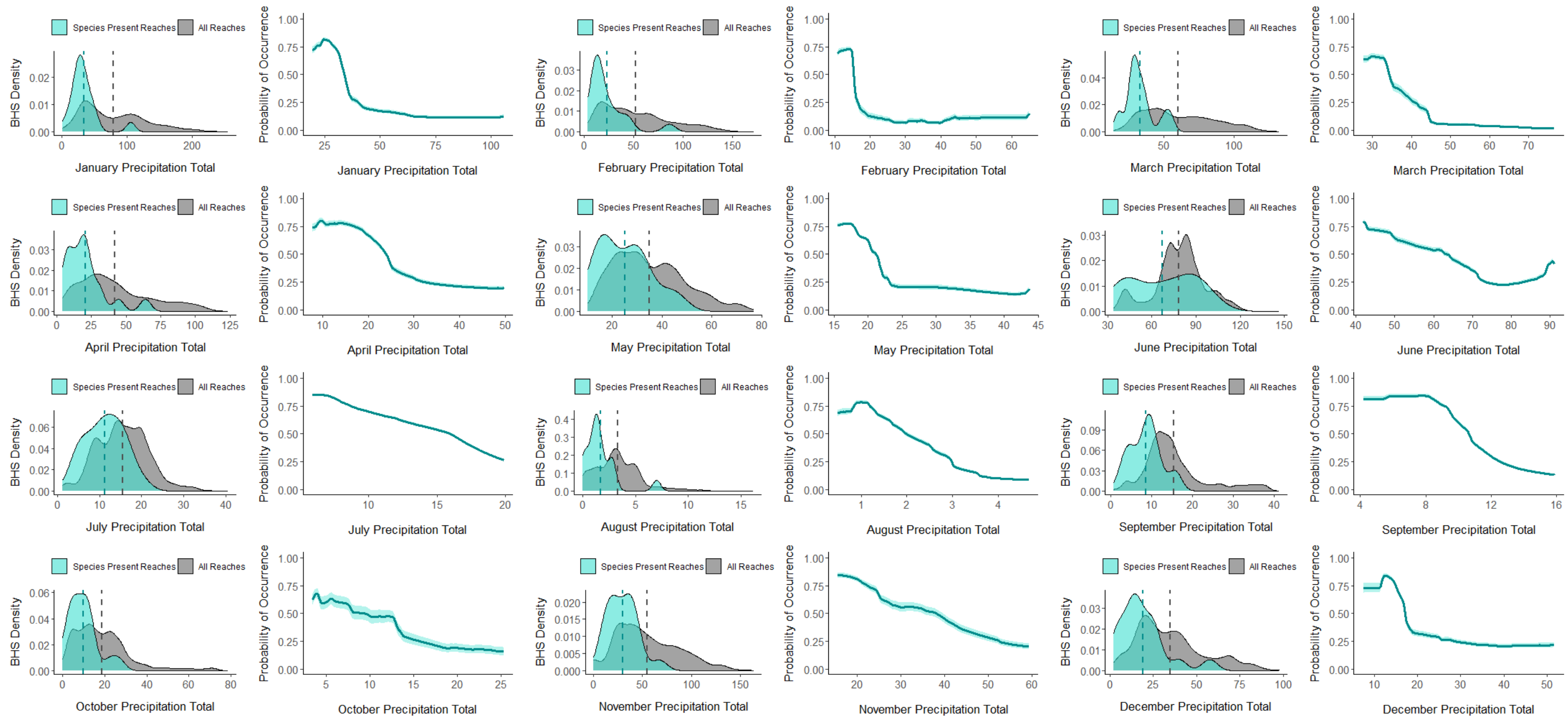

Fig. A10. Monthly precipitation totals in Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months selected for use in model fitting in at least one ensemble iteration.

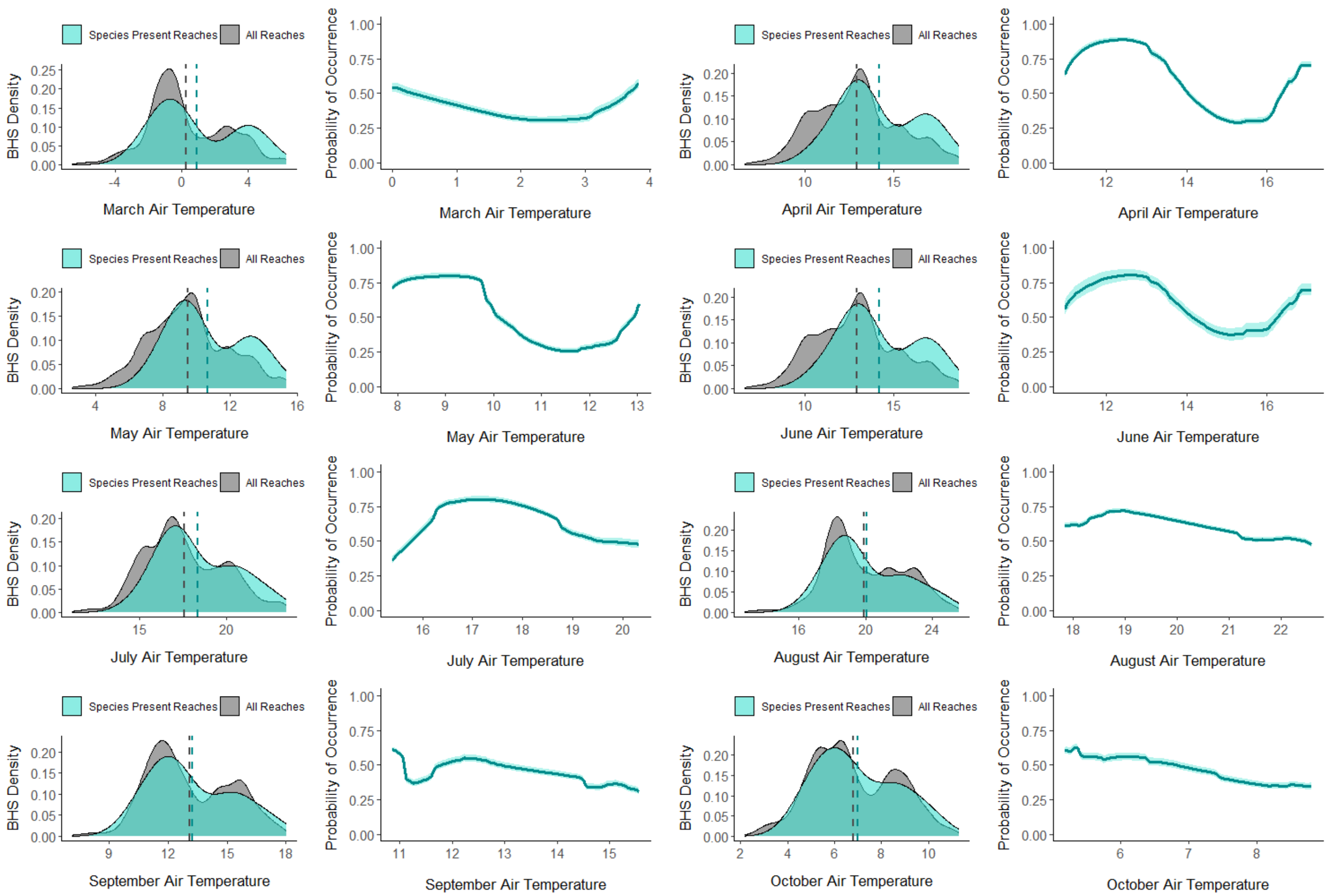

Fig. A11. Average monthly air temperature in Bluehead Sucker (BHS) occupied reaches compared to all reaches, with mean values for each category shown in dotted lines (left) and mean MaxEnt probability of occurrence at different values for each month, across all ensemble iterations with a 95% confidence interval (right). Plots included for all months that were selected for use in model fitting in at least one ensemble iteration.