Predicting the Potential Distribution of Two Threatened Stream Fish Species in Northeast Ohio

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Jeramy John Pinkerton, B.S.

Graduate Program in Evolution, Ecology, and Organismal Biology

The Ohio State University

2016

Thesis Committee:

Elizabeth A. Marschall, Advisor

Thomas Hetherington

Lauren Pintor
Abstract

Stream fish populations are at risk of population isolation due to habitat loss and fragmentation due to a multitude of anthropogenic impacts including dam construction, climate change, and development. Persistence of some stream trout populations has been shown to depend on connectivity with other parts of streams and with other populations. However, in some cases, native stream fish species have benefited from being isolated by barriers from advancing invasive species. Niche models have previously been used to aid in defining the distribution of stream fish populations and facilitating their protection. The goal of our study was to develop niche models using variables derived in GIS and modeled in Maxent to inform the conservation of two state-threatened fish species that reside in northeastern Ohio streams, the brook trout (*Salvelinus fontinalis*) and the bigmouth shiner (*Notropis dorsalis*). We used Maxent, a niche model that uses presence-only data, to model the distribution of these two threatened stream fish species. Brook trout may be negatively affected by isolation that prevents dispersal, while bigmouth shiners may be protected by isolation from competition with silverjaw minnows (*Notropis buccatus*). Because known brook trout populations in Ohio are rare, we used data from nearby Pennsylvania to inform the model of habitat suitability for brook trout in northeast Ohio. In a novel approach, we excluded streams in Pennsylvania that were outside the range of a given variable in northeast Ohio. We then developed an additional model that limited the Pennsylvania dataset to those reaches with habitat variables that...
fall within the range of variables for streams of the same order in northeast Ohio. Of the initial 11 variables that we included in each of our models, four were included in both of our final brook trout models and three were included in our final bigmouth shiner model. We obtained AUC scores of 0.80 and 0.81 respectively for our brook trout models, and 0.94 for our bigmouth shiner model. Model results were the same for both brook trout models. Second and third-order streams with relatively low gradients (between 0.03 and 0.13) and those associated with large amounts of forests near streams were predicted to be most likely to support brook trout. The likelihood of brook trout persistence also increased in streams with an increasing percentage of sandstone in the lithology below the stream. With the addition of a requirement of presence of spawning habitat, the brook trout model suggested 27 more potential brook trout streams in northeast Ohio. Relatively low elevation, high-order streams in relatively unforested landscapes were predicted to be most likely to support bigmouth shiner populations. One hundred and forty-six potential additional bigmouth shiner streams were identified in northeast Ohio. The results of our modeling can be used to aid in the conservation of both of these species in Ohio. Model results can be used to guide potential reintroduction efforts of brook trout in northeast Ohio. We also found that using a surrogate for a primary driver of habitat suitability, such as using stream order as a surrogate for stream temperature, should be used cautiously as the surrogate may affect habitat suitability differently in different areas.
Dedication

Dedicated to Susan Pinkerton, Alexa Carson, and Jasper Carson-Pinkerton
Acknowledgements

I would like to thank all of the folks at the Aquatic Ecology Laboratory for all of their support on the project. I would especially like to thank Reed Brodnik, Dave Glover, and Maura O’Brien. I would also like to thank Curt Wagner (Ohio DOW) and the Brook Trout Joint Venture for their support. I would especially like to thank, my advisor, Dr. Libby Marschall for her support and understanding as I worked to finish this thesis while working a full-time job.
Vita

May 2003 .......................................................Ewen-Trout Creek Jr./Sr. High

2007 .................................................................Lake Superior State University

2012 to 2015 ....................................................Graduate Teaching Associate,

The Ohio State University

2015 to Present ..................................................NR Specialist-Fisheries, MNDNR

Field of Study

Major Field: Evolution, Ecology, and Organismal Biology
# Table of Contents

Abstract ................................................................................................................................. ii
Dedication ............................................................................................................................... iv
Acknowledgments .................................................................................................................. v
Vita ........................................................................................................................................ vi
Table of Contents ................................................................................................................... vii
List of Table ........................................................................................................................... viii
List of Figures ......................................................................................................................... x
Introduction ........................................................................................................................... 1
Methods ................................................................................................................................. 8
Results ................................................................................................................................... 20
Discussion ............................................................................................................................. 23
Literature Cited ....................................................................................................................... 34
Appendix A: Tables and Figures .......................................................................................... 42
List of Tables

Table 1. Potential model variables, including a brief description, mean, and range in northeast Ohio and Pennsylvania, used for predicting brook trout and bigmouth shiner habitat suitability with the source of the data listed...............................................................42

Table 2. Potential model variables. The ‘x’ in the last four columns signifies whether or not a variable was used in a given Maxent model. Eleven variables were used in the initial brook trout model, and that number was reduced to five for the final brook trout model. Substrate type was not considered for the brook trout Maxent model as these data were available only for Ohio and not Pennsylvania. Substrate type was added to the results of the brook trout Maxent model and is denoted by an *. Eleven variables were also used in the initial bigmouth shiner model, and that number was reduced to three for the final bigmouth shiner model. ..............................................................................................................43

Table 3. Correlation matrix for all of the potential variables for the brook trout model in northeast Ohio and Pennsylvania. None of the correlations were greater in magnitude than the acceptable level of 0.6, with the strongest correlation being -0.48 between percent shale and percent sandstone. ..............................................................................................................44
Table 4. Correlation matrix for all of the potential variables for the bigmouth shiner model in northeast Ohio. None of the correlations were greater than our cutoff of ±0.6 with the largest correlation being -0.44 between percent forest and percent urban.

Table 5. NHD ID number and UTM coordinates of predicted brook trout streams in northeast Ohio using Maxent model that included all streams in Pennsylvania that fell within the range of habitats found in northeast Ohio and considering the presence of sand and gravel.
List of Figures

Figure 1. A map of Ohio with our primary study area in blue. The inset in the upper left corner shows the main stems of major watersheds of interest (black lines) and black dots representing streams in which brook trout occur and open circles representing streams where re-introduction was attempted but failed (Burt 2007). .......................................................... 47

Figure 2. The average receiver operating curve for the final brook trout model including all streams within the habitat range found in northeast Ohio with four variables over the 10 runs. This model has an average AUC of 0.80 and a standard deviation of 0.02. The 1:1 line shows what the prediction would be like if the distribution was due completely to random chance. ........................................................................................................ 48

Figure 3. The average receiver operating curve for the final brook trout model including only streams from the Pennsylvania dataset that with habitats that are within the range of their respective stream order in northeast Ohio with four variables over the 10 runs. This model has an average AUC of 0.81 and a standard deviation of 0.01. The 1:1 line shows what the prediction would be like if the distribution was due completely to random chance. ........................................................................................................ 49

Figure 4. Environmental response curves for the variables in the final brook trout Maxent model. The response curves are constructed using only the corresponding
variable. Each curve also shows its standard deviation in a different color shading. The black dots represent the given environmental value of each of the four known brook trout streams in northeast Ohio.

Figure 5. A map of the stream reaches in northeast Ohio with black dots representing streams in which brook trout occur, open circles representing streams where re-introduction was attempted but failed, red dots representing streams with a logistic score from Maxent greater than 0.75, and red triangles representing streams with a logistic score from Maxent greater than 0.75 and a sand and gravel substrate. The main stems of Black, Rocky, Chagrin, and Grand Rivers are denoted by the dark lines.

Figure 6. The average receiver operating curve for the final bigmouth shiner model with three variables over the 10 runs. This model has an average AUC of 0.934 and a standard deviation of 0.03. The 1:1 line shows what the prediction would be like if the distribution was due completely to random chance.

Figure 7. A map of the stream reaches in northeast Ohio with black dots representing streams in which bigmouth shiners occur and red dots representing streams with logistic score from Maxent greater than 0.75. The main stems of Black, Rocky, Chagrin, and Grand Rivers are denoted by the dark lines.

Figure 8. Environmental response curves for the variables in the final bigmouth shiner Maxent model. The response curves of constructed using only the corresponding variable. Each curve also shows its standard deviation in blue.
Introduction

Population isolation via habitat loss and fragmentation is a growing threat to species persistence across a wide variety of aquatic ecosystems (Gimmi et al. 2011; Jaeger et al. 2014; Liermann et al. 2012; Mora and Sale 2011; Roberts et al. 2013). Stream fish populations are particularly sensitive to habitat loss and fragmentation, and population isolation has been attributed to a range of human activities, including dam construction (Neves and Angermeier 1990), poorly designed road crossings (Perkin and Gido 2012), climate change (Palmer et al. 2009), and agricultural and residential development (Pess et al. 2002). A direct and immediate impact of such isolation is a reduction in sub-population connectivity. Connectivity is known to be vitally important for the persistence of stream fish populations. For example, brook trout (Salvelinus fontinalis) inhabiting streams have been shown to move regularly between the main stem and tributaries of a stream and to have higher reproductive success in tributaries than the main stem, regardless of the stream segment that the parents typically inhabit (Kanno et al. 2014). Genetic analysis of stream brook trout shows gene flow between subpopulations in the main stem and connected tributaries (Kelson et al. 2015). Also, during their first summer and winter, brown trout (Salmo trutta) have been shown to have high rates of downstream movement (Vøllestad et al. 2012) which can lead to individuals, and their genes, being lost from the upstream population if they move
downstream past a barrier (Garner et al. 2013). The previous examples show the importance of connectivity in stream trout populations. If the tributaries and main stems of streams are fragmented, the trout in the main stem may lose access to spawning habitat, with negative impacts on population growth and maintenance. Short stream fragments generally can support only small populations of fish, which are more susceptible to extirpation than large populations (Harvey and Railsback 2012; Roberts et al. 2013). Further, isolation and fragmentation can also prevent a species from recolonizing suitable habitat.

Persistence of trout populations in small streams has been found to be highly dependent on immigration (Morita and Yokota 2002), which is not possible when stream reaches are isolated. Dispersal between populations requires the absence of barriers over the entire length of stream connecting the populations, and, therefore, even a single obstacle easily isolates stream trout populations. Brook trout in Appalachian streams without barriers to movement show no genetic structuring, suggesting that individuals move freely between tributaries and the main stem of a stream (Aunins et al. 2015). Simulations of brown trout in a stream network containing a barrier showed rapidly decreasing populations above the barrier (Frank and Baret 2013). This ultimately resulted in a decrease in genetic diversity and eventual local extirpation above the barrier, both of which resulted from the loss of migrants into the above-barrier stream reach (Frank and Baret 2013). Field data show that as stream fragment size decreases, the genetic diversity of brook trout populations isolated in those fragments decreases, which leads to a higher probability of local extirpation (Whiteley et al. 2013). Similarly, white-
spotted charr (*Salvelinus leucomaenis*) populations in Japan that are isolated by dams often fail to persist over long periods of time (Morita and Yokota 2002). Even when isolated populations are located in suitable habitat, stochastic events can cause small populations to go extinct (Allendorf and Ryman 2002; Foley 1994; Gilpin and Soulé 1986; Hilderbrand 2003; Mangel and Tier 1994; Mills and Smouse 1994). This is particularly true for small stream segments in which stochastic events such as debris flows and stream freezing have the potential to cause great perturbations to the system and populations therein (Roberts et al. 2013).

The impacts of habitat fragmentation are not always negative, however, and can help to protect fish populations in some cases. Natural barriers such as waterfalls have been shown to protect native fish in New Zealand from invading brown trout (Townsend and Crowl 1991). Artificial barriers have been used to protect Colorado River cutthroat trout (*Oncorhynchus clarki pleuriticus*) populations from invasion by non-native trout (Thompson and Rahel 1998). There are many trade-offs to be considered in decisions to enhance, maintain, or reduce isolation as a part of a population conservation plan (Fausch et al. 2009; Peterson et al. 2008). While barriers may protect against invading species, there is some doubt that these unnaturally isolated populations can persist in the long term (Novinger and Rahel 2003).

To understand the link between habitat on a landscape level and stream fish distribution, previous researchers have used niche models. Niche models can be used to assess suitability of sites for establishment or re-establishment of populations (Cianfrani et al. 2010; Guisan et al. 2006). As such, niche models have become an important tool
for informing strategies for conservation. Niche models use habitat variables and organism location records to estimate the potential distribution of a study organism in a given area. Habitat variables can include local site characteristics, landscape variables, and climatic information. These variables, when input in geographic information systems (GIS), can easily be used in niche models. Niche models have seen widespread use in stream fish conservation. Niche models incorporating landscape variables in GIS have been used to model and map fish distributions, such as westslope cutthroat trout (Oncorynchus clarkii lewisii) in the western United States (Shepard et al. 2005), brook trout in Pennsylvania (Kocovsky and Carline 2006), and brook trout, brown trout, and rainbow trout (Oncorynchus mykiss) near the southern edge of their ranges in the United States (DeRolph et al. 2015). The potential distribution of the invasive red shiner (Cyprinella lutrensis) was mapped using GIS and maximum entropy (Maxent) models, a type of niche model, across the United States (Poulos et al. 2012). Landscape-level habitat variables used in a variety of different types of niche modes were successful in predicting the distribution of multiple stream fish species in the Great Plains, United States (Oakes et al. 2005). Neural network models, a type of niche model, using climatic, landscape, and stream-reach variables were used to predict the distribution of fallfish (Semotilus corporalis) and to determine the optimal predicted habitat of fallfish in the Lake Ontario and the upper St. Lawrence River watersheds in New York (McKenna et al. 2012). A GIS model was used to predict areas where mountain suckers (Catostomus platyrhynchus) have a high probability of occurrence in the Black Hills of South Dakota such that sites suitable for translocation could be identified (Dauwalter and Rahel 2008).
Similarly, potential habitat was modeled in northern Michigan to identify streams that may be suitable for arctic grayling (*Thymallus arcticus*) reintroduction (Tingley 2010). In these applications, habitat variables at multiple spatial scales were used to describe suitable habitat to predict whether the target fish species would be able to inhabit particular stream reaches (Dauwalter and Rahel 2008; Tingley 2010).

Maximum entropy (Maxent) models, one type of niche model, have been used to model the potential habitat of threatened stream fish species. For example, goldline darter (*Percina aurolineata*) habitat was modeled in an urbanized area of Georgia (Albanese et al. 2013) and blackside dace (*Phoxinus cumberlandensis*) habitat was modeled in Kentucky (Laing et al. 2013). Maxent is a machine learning program that uses presence-only records and measures the relative entropy between the probability density estimated from the presence data and the probability density estimated from the landscape (Elith et al. 2011; Phillips et al. 2006). A full description of Maxent methods can be found in Phillips et al. (2006) and Elith et al. (2011). One of the benefits of using Maxent methods is that it requires only presence, and not absence, data. Many streams have not been surveyed, so lack of stream fish records does not necessarily imply absence of stream fish.

The goal of this study was to develop niche models using variables derived in GIS and modeled in Maxent to inform the conservation of two state-threatened fish species that reside in fragmented habitats in small northeastern Ohio streams, the brook trout and the bigmouth shiner (*Notropis dorsalis*).
Brook trout populations in Ohio are both rare and isolated, with potential suitable habitat separated by long stretches of naturally uninhabitable stream area. One goal of our study was to identify potential brook trout habitat and recommend reintroduction sites to assist in colonization – via propagation and reintroduction - between fragmented habitat sites. A recent study used hierarchical models to predict stream reaches in the eastern United States that support the occurrence of brook trout in their native range (DeWeber and Wagner 2015). A neural network model using watershed and stream reach characteristics including current-day mean air temperature, prior 7-d mean air temperature, network area, network forest cover, network mean aspect, network mean base flow index, and riparian forest cover within the local catchment was used to predict stream reach temperature (DeWeber and Wagner 2015). Stream reach temperature was not predicted for northeast Ohio due to lack of data. This study found that high predicted stream temperature, high soil permeability, and location in urban and agricultural landscapes all negatively affected the likelihood of the reach supporting brook trout (DeWeber and Wagner 2015). This study also found an interaction between agricultural development and predicted stream reach temperature, where predicted stream temperatures increased as agricultural development increased (DeWeber and Wagner 2015). This previous study did not include data for a large portion of northeast Ohio and used a stream dataset at a spatial scale that was too coarse to include streams as small as required to study brook trout in northeast Ohio.

In contrast to the negative effect of fragmentation to Ohio brook trout populations, Trautman (1981) suggested that fragmentation may be important for the
persistence of bigmouth shiners in Ohio streams, as barriers to dispersal prevent silverjaw minnows (*Notropis buccatus*) from colonizing upstream sites and displacing bigmouth shiners. Thus, the second goal of our study was to determine if there may be streams in northeast Ohio with unknown bigmouth shiner populations or suitable bigmouth shiner habitat. To our knowledge there has not been a study investigating the presence of bigmouth shiner in Ohio.
Methods

Study Area

Our study area encompasses northeast Ohio, United States, and specifically the stream systems of the Chagrin River, Black River, Rocky River, and Grand River watersheds, all of which are in the Lake Erie drainage (Figure 1). We confined our study to these watersheds because they either currently support, or are thought to have historically supported, native brook trout populations, and the bigmouth shiner has historically been found in the Black River and Rocky River watersheds. Surficial substrate ranges from fertile till (farm land) to bedrock, and the elevation ranges from about 173-402 meters. Air temperature in this area reaches a monthly average high of 22.3° C in July to a monthly average low of -3.1° C in January (NOAA, Akron/Canton weather station, 1981-2010). There are a variety of land uses in this area including undeveloped deciduous forest, agricultural development, urban, and residential development.

Study species

Brook trout

Brook trout are listed as threatened in Ohio by the Ohio Department of Natural Resources (ODNR) and, before 1972, were considered to have been extirpated from the
state. In 1972, populations of brook trout were found in Spring Brook and Woodie Brook in the Chagrin River watershed in northeast Ohio (Burt 2007). During the 1990’s, these populations were analyzed and found to be of a native strain that is genetically distinct from populations in neighboring states and those commonly stocked in Ohio streams, suggesting that these populations are from a remnant population (Burt 2007). In 1993, the population in Woodie Brook was extirpated via habitat destruction caused by residential development (Burt 2007). Between 1992 and 2000, the ODNR surveyed a subset of streams in northeast Ohio for additional brook trout populations, but none were found (Burt 2007). Despite not finding additional populations, the search was successful in finding 15 streams in which the investigators determined that brook trout could possibly survive, based on metrics such as mid-summer water temperature, substrate type, the Ohio Environmental Protection Agency’s (OEPA) Qualitative Habitat Evaluation Index (QHEI, Rankin 1989), and fish community (Burt 2007). The streams were found via word of mouth and were the only streams thought to be suitable for brook trout reintroduction after examining about 200 streams (Curt Wagner, ODNR, personal communication). Reintroduction was attempted by taking gametes from fish in Spring Brook and using the offspring produced in a hatchery as brood stock to produce fish for reintroduction (Burt 2007). These 15 streams were stocked for three consecutive years and the populations were monitored post-stocking to assess whether the populations were able to persist (Burt 2007). After the reintroduction efforts, five of the reintroduced sites have been able to maintain strong populations via natural reproduction, four additional
sites hold small populations, two streams had brook trout spawn at least once but not persist, and four streams showed no signs of natural reproduction (Burt 2007).

Given the small number of streams in Ohio that are currently known to support the native strain of brook trout, it is important to protect these areas and any other areas that may potentially provide the habitat necessary to support these fish. The populations are located only in headwater streams that have significant groundwater inputs (Burt 2007). The second-order and larger streams that these headwater streams connect with apparently are too warm to support brook trout. Any additional unknown stream reaches, if found, could give the ODNR more areas for brook trout reintroductions with the hopes of adding more naturally reproducing populations. By increasing the number of brook trout and the number of streams containing brook trout, we can increase the probability of regional persistence for brook trout in Ohio. To aid in this, we developed a strategy to systematically look for potential brook trout streams in Northeast Ohio.

**Bigmouth shiner**

The bigmouth shiner may also be at risk due to loss of habitat, but for reasons that differ from brook trout. Habitat destruction and inter-specific competition with the silverjaw minnow (*Notropis dorsalis*) are believed to be the main causes of decline in the bigmouth shiner population in Ohio (Smith et al. 1973). The silverjaw minnow is native to most of Ohio and thought to be naturally expanding its range into the area inhabited by the bigmouth shiner. Although we have no current data on changes in the range of silverjaw minnows, in 1981, it was believed to be expanding its range into area inhabited
by the bigmouth shiner (Trautman 1981). The goal of the bigmouth shiner study is to determine if there are other streams in northeast Ohio with bigmouth shiner populations or suitable bigmouth shiner habitat.

**Presence Records**

Species locations were obtained from multiple sources. Locations for all streams in which brook trout are found in northeast Ohio were obtained from the Ohio Division of Wildlife. Given the small number of brook trout streams in northeast Ohio, we included locations from Pennsylvania to increase our sample size and improve our model performance. Locations where brook trout were sampled in Pennsylvania between 2004 and 2013 were obtained from the Pennsylvania Fish and Boat Commission. Since many streams were sampled more than once during this period, we included every stream reach that had at least one record of brook trout during the ten-year time span.

The GPS coordinates of brook trout and bigmouth shiner sampling events were assigned to the nearest stream reach if they fell within 50 meters of a stream reach in the National Hydrography Dataset (NHD24k, obtained from the U.S. Geological Survey). If the GPS coordinates recorded for a sampling point did not fall within 50 meters of a stream reach in NHD24k dataset we did not include it in our analysis, as there may have been an error in recording the sampling location or the sample may have been collected in a stream that is not included in our dataset. Any stream reach that had at least one presence record for a given species during the dates listed above was considered to have the species for our modeling. In all, we had four stream reaches in northeast Ohio and
821 stream reaches in Pennsylvania that contained brook trout for a total of 825 stream reaches with brook trout for our study. There were 42 stream reaches in northeast Ohio with recorded bigmouth shiner presence.

*Habitat variables*

Nine variables were used to characterize the habitat found in streams in northeast Ohio and Pennsylvania, encompassing geological, morphological, biological, and anthropogenic features that may affect the distribution of brook trout or bigmouth shiners (Table 1). In this section, we briefly describe each variable and why it is important to our study goals. The range and source of each habitat variable can be found in Table 1.

Anthropogenic effects on streams can be characterized in part by land-use and number of road crossings. Land-use in our model was characterized by the percent of a given land-use (developed, forest, agriculture, wetland/herbaceous or open water) within a 10-m buffer around a stream reach. Altered landscapes can negatively affect fish habitat by degrading available habitat. Specifically, brook trout presence is positively correlated with percent forest and negatively correlated with percent agriculture (Hudy et al. 2008). Forests also shade streams from solar radiation, helping to maintain cool water temperatures in the summer (Johnson and Jones 2000; Moore et al. 2005). Separating forest types between shrub-dominated and tree-dominated landscapes was not necessary because there is no significant difference in the mean weekly temperature between these forest types (Cross et al. 2013). Road crossings can both positively and negatively affect fish distributions. As the density of road crossings increase, the persistence of brook
trout decreases (Hudy et al. 2008). Also, poorly designed road crossings can negatively affect brook trout habitat quality by allowing sand to enter the stream (Lachance et al. 2008). However, in some instances, poorly designed road crossings may benefit fish species persistence by preventing invading species from advancing upstream into areas inhabited by the original species.

The general morphology of streams can be used to help predict fish habitat types. We used Strahler stream order (Strahler 1954; Strahler 1957), slope, and elevation to characterize stream reaches. Strahler stream order is calculated by labeling the most upstream stream in a network as 1st-order. When two 1st-order streams come together, they form a 2nd-order stream and so on (Strahler 1954; Strahler 1957). Stream order increases only if two streams of the same order merge (Strahler 1954; Strahler 1957). In Ohio, brook trout are generally thought to be found only in 1st-order streams (Burt 2007), as larger streams are thought to be too warm for persistence (Burt 2007). Bigmouth shiners are generally found in larger streams, which tend to have a slower current speed. Elevation can affect the temperature of streams, and slope can affect the habitat types available in a stream reach, and thus, the fish species that inhabit these reaches. Strahler stream order was calculated using the ArcHydro package in ArcGIS on the NHD24k streams. Slope and elevation were calculated using the zonal statistics tool in ArcGIS.

Geology can impact the substrate of the stream, water chemistry, and ground water inputs to streams. We looked at surface substrate and bedrock types in and under streams. Stream substrate can be determined by looking at the composition of soil types in a landscape. We used five-meter buffers around streams to characterize the percent of
six different substrate categories for streams in northeast Ohio. This buffer allowed us to characterize the substrate of the stream. Substrate was categorized as being composed primarily of fine sediment, sand and gravel, or till with different distinctions for secondary components. Substrate type is thought to be an important determinant of brook trout (Burt 2007) and bigmouth shiner (Trautman 1981) occurrence in streams in Ohio. Sharon sandstone and conglomerate is thought to be important for brook trout reproduction in Ohio, as the conglomerate, which is typically small quartz, provides spawning gravel for brook trout (Curt Wagner, ODNR, personal communication). Also, fine particulate matter, such as silt, may be detrimental to brook trout populations, as fine particles can cover spawning gravel and fill in deeper pools which are essential for the survival of larger fish (Alexander and Hansen 1986). Bigmouth shiners are found in streams with sandy bottoms and also are negatively affected by silt (Trautman 1981). We found this substrate type data only for Ohio, so we could use it to help characterize only northeast Ohio streams, and it could not be used in our initial brook trout Maxent model that included data from Pennsylvania. Lithology of the bedrock under streams can impact the water chemistry of streams and can potentially impact groundwater inputs from deep ground sources. We characterized the percent of seven different lithological categories in northeast Ohio and Pennsylvania streams: silt, sand, sand stone with conglomerate, conglomerate, shale, limestone, and hard (metamorphic and igneous) rock types. For both substrate type and lithology, we characterized stream reaches by percent of each type under a 5-m buffer around each reach.
As noted previously, the presence of silverjaw minnows may prevent bigmouth shiners from inhabiting a stream reach through competitive exclusion. We categorized stream reaches as having silverjaw minnows if there was a record of them being sampled at least once between July 2006 and October 2013 in a given stream reach. There were 220 stream reaches with recorded silverjaw minnow presence. We also modeled the potential distribution of silverjaw minnows using Maxent and the same initial suite of variables as the bigmouth shiner model to determine the extent of overlap in the potential distribution of bigmouth shiners and silverjaw minnows.

Some variables in Pennsylvania have a much greater range than the same variables in northeast Ohio (Table 1). Since our interest is modeling species distributions in northeast Ohio, we removed all streams from our Pennsylvania dataset that had ranges outside of those found in northeast Ohio. Streams in Pennsylvania with an elevation, slope, stream order, or number of road crossings higher than those found in northeast Ohio were removed from our data set.

Species distribution modeling

To model the potential distribution of brook trout, bigmouth shiner, and silverjaw minnow, we used samples with data (SWD) in Maxent version 3.3.3k following the methods described by Elith et al. (2011). Maxent is a presence only model that is advantageous in part, because presence data can be combined from samples using a variety of different methods and museums samples in which the collection method may not be known (Phillips et al. 2006). The SWD frame work allows the modeler to run the
Maxent model with presence point, in this case a stream reach, having its attributes associated with the point. Some categories were removed from land cover, substrate, and lithology to prevent the values for each category from summing to one. Categories were removed if they were found only in Pennsylvania or were rare and were not present and/or variables in streams that supported either brook trout or bigmouth shiner. The categories used in the model can be found in Table 2. We used streams that are within the same watershed as a stream with our species of interest for our background data, so that our background stream reaches came from watersheds that support our study species. The default settings for regularization parameters were used in the study. These settings are well suited for a wide range of presence-only datasets because they were tuned on a wide range of species, taxa groups, number of occurrence records and species prevalence (Phillips and Dudík 2008).

A receiver operator curve (ROC) is used to assess the performance of a Maxent model. More specifically, the area under the curve (AUC) is commonly used (Byers et al. 2009; Elith et al. 2011; Albanese 2013). The AUC value characterizes the performance of a model as a single number by plotting the true positive rate (sensitivity), which represents absence of omission error against the false positive rate, (1-specificity) also known as commission error (Elith et al. 2011). A value of 0.5 suggests that the model performs no better than selecting random data. The AUC ranges from 0 to 1. According to Swets (1988), values less than 0.7 indicate low accuracy, 0.7-0.9 can be useful, and greater than 0.9 exhibit high accuracy.
Logistic scores that are used to predict the likelihood of each individual stream supporting a given fish species are also produced by Maxent models (Phillips and Dudik 2008). These logistic scores are transformations of the raw data output by Maxent that make the output much more intuitive (Elith et al. 2011). We took the average of these logistic scores for each stream reach to determine a final logistic prediction for each individual reach. We also looked at environmental response curves of predicted values from the model versus habitat variables, produced by Maxent. Environmental response curves are produced using only a single corresponding habitat variable.

Variables were systematically removed from our models to increase model performance. First, we used Pearson correlation to determine if any of our variables were highly correlated (Pearson correlation > 0.6, Byers et al. 2013), in which case we removed one of the variables. Jackknifing plots were used to determine which of two correlated variables should be removed. Jackknifing assesses the change in gains in model fit with and without each variable. The variable that explained the model results best on its own and caused the largest decrease in performance when removed from the model was included in the model. We also used jackknifing and percent contribution to determine if variables could be removed without significantly decreasing the predictive power of our models. Percent contribution represents the contribution of a variable to the final model (i.e. the higher the percent contribution the more variance a variable explains). If a variable has a very low percent contribution, it may be able to be removed from the model without negatively impacting the AUC of the model. To summarize, we removed variables in two ways. We removed variables 1) due to correlation and 2) by
analyzing the jackknifing plots and percent contribution to see if a variable could be removed without negatively impacting the model. We ran the model after each individual variable was removed to analyze model performance and to determine if we should remove another variable. Each model run consisted of a cross validation in which we ran the model ten times for each species, leaving out a different one-tenth of the presence data each time in each replicate (Elith et al. 2011).

We included streams, with variables within the range of those in northeast Ohio from the Pennsylvania dataset to increase the number of presence points for brook trout in our study, as there are only seven streams with brook trout in Ohio. The larger number of data points gives us more confidence in our model predictions. We used the resulting model to make predictions about brook trout in northeast Ohio, including one additional habitat variable available only for Ohio streams. The additional habitat variable, stream sand and gravel, is believed to be important for brook trout persistence in Ohio (Amey, OEPA, unpublished data; Pira et al., Geauga County Metro Parks, unpublished data). Streams were considered to have sand and gravel substrate if any portion of the substrate consisted of sand and gravel. Similar to Byers et al. (2013), we used a two-part modeling process: development of a Maxent model to study the distribution of a species over a large geographic area, followed by use of local habitat variables to fine-tune the model results for a specific area.

Because we were concerned that a specific stream order in Pennsylvania streams may be associated with a different range of habitat variables than streams of the same order in Ohio, we developed a second model with a more limited data set. For a given
stream order, we included only those Pennsylvania streams that had slopes, elevations, and number of road crossings within the range of those variables at that stream order in Ohio streams. This reduced the data set by 29%.

Finally, streams that had a Maxent logistic score above 0.75 and had sand and gravel substrate were examined to determine the presence of in-stream ponds. Ponds of this nature could add warm water to streams and make them unsuitable for brook trout habitation. We used Google Earth to examine streams for the presence of obvious ponds.
Results

Brook trout

Our initial brook trout model using all streams with habitat variables within the range of those found in Ohio consisted of 11 variables modeled in Maxent for northeast Ohio and Pennsylvania (Table 2). This initial model had an area-under-the-curve (AUC) score of 0.82. No pairwise correlations between these variables were greater than our maximum acceptable correlation of ±0.6 (Table 3). The strongest correlations were between percent forest and percent agriculture (Correlation = -0.41) and between percent shale and percent sandstone (correlation = -0.49). Analysis of the jackknife plots and percent contribution allowed for seven variables to be removed from the model. Our final brook trout model contained four variables including slope, stream order, percent forest, and percent sandstone, (Table 2) and had an AUC score of 0.80 (Figure 2).

The second brook trout model, limiting the Pennsylvania data to only those streams with habitat variables within the range of Ohio streams, within in any given stream order, started with the same initial suite of variables and had similar results to our first brook trout model. Our final suite of variables for this second model was the same as those of our previous brook trout model: slope, stream order, percent forest and percent sandstone (Table 2). The AUC score for this model was 0.81 (Figure 3).

Logistic scores suggest that a number of streams in our study area have the potential to support brook trout. Based on the final brook trout Maxent model
considering all stream in Pennsylvania with habitats within the northeast Ohio habitats, 755 streams within the range of variables in northeast Ohio, in Pennsylvania and northeast Ohio, have at least a 75% likelihood of being able to support brook trout. The final brook trout model using the stream-order matched data from Pennsylvania found 609 streams in Pennsylvania and northeast Ohio with a logistic score over 0.75. The two models had identical response plots with one minor exception. Stream slopes up to 0.3 are included in the dataset that includes all streams within the range of northeast Ohio variables, while slope only goes up to 0.28 meters per meter in the dataset that considers only Pennsylvania streams with variables within the range of streams with the same stream order in northeast Ohio. The probability of a stream reach being able to support brook trout increases with increasing percent forest and percent sandstone (Figure 4). Stream reaches with relatively low slopes and low (2\textsuperscript{nd} or 3\textsuperscript{rd}) stream order have a high likelihood of being able to support brook trout (Figure 3). Percent forest was the most important variable in our model followed by stream order. Given the similarities between the two models we focused on the logistic output from the model that included all streams in the Pennsylvania dataset that where within the range of variables found in northeast Ohio. Focusing on northeast Ohio, there are 39 streams with a logistic score over 0.75 for the model (Figure 5). When substrate was added to the model output, we found 27 streams with a logistic score greater than 0.75 that also contained sand and gravel (Table 5, Figure 5). None of the four known brook trout streams that were used in our study had a logistic score over 0.75. None of the 27 streams identified as potential brook trout streams in northeast Ohio had ponds visible using Google Earth.
Bigmouth Shiner

The initial bigmouth shiner model contained 11 variables (Table 2). The mean AUC for this model was 0.94. Similar to the brook trout model, there were no variables that were correlated by more than ±0.6 (Table 4). The highest correlations were negative correlations between percent forest and percent urban (-0.44) and percent forest and percent agriculture (-0.36). Analysis of the jackknife plots and percent contribution allowed for eight variables to be removed from the model. The final model for predicting the potential distribution of bigmouth shiner contained three variables including stream order, elevation, and percent forest (Table 2) and had a mean AUC of 0.94 (Figure 6). This model identified 146 streams in our study area with a logistic score over 0.75 (Figure 7). The probability of a stream reach being able to support bigmouth shiner appears to peak near the lower end the elevation and percent forest gradient and is highest in 4th- and 5th-order streams (Figure 8). Stream order was the most important variable in this model, followed by elevation. Ten of the 42 known bigmouth shiner streams had a logistic score of 0.75 or greater and an additional 20 streams were between 0.5 and 0.75. The average logistic score for the 42 known bigmouth shiner streams was 0.58.

The silverjaw minnow model found that 136 streams had a logistic score over 0.75. Only 13 of the potential silverjaw minnow streams overlapped with the potential distribution of the bigmouth shiner.
Discussion

Brook trout

In this study we used what we believe to be a novel approach for using a landscape model to examine habitat suitability for a stream fish species in a region with only a small number of known populations. We used two subsets of streams from a nearby region, with many recorded populations, to describe suitable habitat, and used this to predict locations of suitable habitat in the original region of interest. Because streams in northeast Ohio have a smaller range of habitat types than those in Pennsylvania, we filtered the Pennsylvania dataset to include data only from streams with habitats within the range of habitat found in northeast Ohio. We then filtered the Pennsylvania dataset further by sorting Pennsylvania streams by order and removing streams that were outside of the range of habitats of a particular stream order found in northeast Ohio to form a dataset for a second brook trout model. Below, we address the results of these models as well as caveats to consider in using data from different regions to describe habitat suitability.

Northeast Ohio has only a few known native strain brook trout populations. All but one of these populations are in the Chagrin River basin. Our model including all Pennsylvania streams that fall within the range of habitats present in northeast Ohio predicted 27 additional sites that are suitable for brook trout. There are two primary clusters of new potential sites, one of which is near the Chagrin River and known brook trout locations. The other cluster is west of the Chagrin River. There are also a few
potential streams that are located on the west side of the region. These results can be used to guide the conservation efforts of brook trout in northeast Ohio.

The majority of streams identified by our brook trout model appear to be isolated by distance and potentially unsuitable brook trout habitat from other streams that are suitable for brook trout. Populations may be able to adapt to these isolated conditions, as evidenced by the persistence of the isolated brook trout populations currently residing in northeast Ohio. Shifts in demographic rates in response to isolation can increase the probability of persistence of isolated stream fish populations without the use of supplementation (Harvey and Railsback 2012; Letcher et al. 2007). Viability in fish populations found above barriers has been found to be enhanced by altered life-histories such as adaptations to isolation, including high early survival rates and reproduction at young ages and small sizes, and by selection against emigration (Harvey and Railsback 2012; Letcher et al. 2007). Population resilience to destructive stochastic events is expected to increase due to high early survival rates and early age at first reproduction (Winemiller 2005). For example, simulated isolated cutthroat trout (*Oncorhynchus clarki*) populations show higher egg to age-one survival than populations simulated without barriers (Harvey and Railsback 2012). White-spotted charr isolated by barriers showed increased growth rates and exhibited a resident lifestyle more frequently than the population that inhabited the river downstream of the barrier, which still had access to individuals returning from the ocean (Morita et al. 2000). Generally, reproduction at younger ages and smaller sizes and higher early-life survival have been found to make fish populations more resilient to perturbations such as hydraulic alteration and siltation as long as there are refugia (Winemiller 2005).
Understanding the landscape characteristics that affect the distribution of threatened or endangered fish species is important for the preservation of these species. We can use information about habitat suitability, as determined by landscape characteristics to inform discovery of populations of threatened species in previously unsurveyed areas and to inform selection of reintroduction sites. A large-scale, comprehensive field study to search for these areas is not feasible in many cases. Our results identified northeast Ohio streams that are predicted to be suitable to support populations of two threatened stream fish species. We identified 27 more potential brook trout streams, which, if a field study found them suitable, could almost quadruple the number of streams that brook trout currently inhabit.

Similar to other studies, we found that percent forest near the stream is a strong predictor of brook trout habitat (Hudy et al. 2008; Stranko et al. 2008). Forested streams can reduce stream warming by shading the stream (Vannote et al. 1980), and can help to reduce erosion (Zaimes et al. 2004) which could contribute to siltation that fills in deep pool habitat for adults and covers spawning gravel that could reduce successful reproduction and egg survival (Alexander and Hansen 1986). Forest cover measured as percent forest within a stream buffer (this study), percent canopy cover over a stream (northeast Ohio brook trout reintroduction study; Amey, OEPA, unpublished data; Pira et al., Geauga County Metro Parks, unpublished data), and percent forest in the watershed (Hudy et al. 2008; Stranko et al. 2008) all were strongly correlated with suitable brook trout habitat. Similarly, DeWeber and Wagner (2015) found that increasing agriculture and urban or residential development (which is negatively correlated with forest cover) negatively impacted the probability of a stream reach supporting brook trout. Current
and potential future land cover and land use should be carefully considered when thinking about potential brook trout re-introductions. Urban development had temporarily claimed one northeast Ohio brook trout population (Burt 2007) and has been shown to negatively impact brook trout populations in Maryland (Stranko et al. 2008). Streams that are or can be protected from development should garner consideration for brook trout re-introductions to avoid the negative impacts of development.

Predicted brook trout streams were also typically low-order streams (2\textsuperscript{nd} and 3\textsuperscript{rd} order) with a relatively low slope. Low-order streams with a gentle slope have also been shown to be important for another threatened stream fish species, blackside dace (Liang et al 2012). Because low-order streams are small, when riparian forests are present the streams can be completely shaded. This, and the fact that low-order streams often are heavily influenced by groundwater inputs (Vannote et al. 1980), results in low-order streams typically having cool water in the summer. A previous study concluded that brook trout could persist in northeast Ohio only in headwater streams (Burt 2007). Our models predict that 2\textsuperscript{nd}- and 3\textsuperscript{rd}-order streams may be most suitable for brook trout. While most of the 27 streams in Ohio that are suggested by our model to be able to support brook trout are classified as 2\textsuperscript{nd}- and 3\textsuperscript{rd}-order streams, 15 and 11 streams respectively, one was 4\textsuperscript{th}-order. These higher order streams may have characteristics suitable for brook trout, but we also must consider that the model results may be a reflection of the differences in stream orders between Pennsylvania and Ohio. In Pennsylvania, brook trout are known to inhabit 6\textsuperscript{th}-order streams. Stream order is not the primary driver of habitat suitability in brook trout; stream order is used because it is correlated with the primary drivers and it is easy to assess from maps. Within a region,
stream order is strongly correlated with water temperature, slope, and substrate, for example, but the exact correlation differs among regions. In Pennsylvania, first-order streams tend to have higher elevation and steeper slopes than first-order streams in Ohio. Thus, stream order represents a different suite of habitat conditions in each of these two regions, which has been shown in other studies as well (Dauwalter and Rahel 2008).

The Appalachian Mountains directly affect streams in Pennsylvania, but not Ohio, with elevations being up to 500 m higher in Pennsylvania than Ohio. While we removed high elevation streams from the Pennsylvania data set, the lower elevation streams that we included were higher order than the streams at the same elevation in Ohio. Also, these higher order streams in Pennsylvania may be cooler than similarly ordered streams in Ohio. The low-order, high-elevation streams that merge to form these higher order streams in Pennsylvania, may be cooler because they are found in higher elevations, and thus cooler air temperatures (Rahel and Nibbelink 1999).

In attempt to tease apart some of the confounding variables that affect stream order, we also modeled brook trout distribution using streams of a given order from Pennsylvania that had other habitat variables only within the range found in that particular order in northeast Ohio. The results of this model were essentially the same as the results of our previous brook trout model. Mainly 2nd and 3rd order streams still were found to be most suitable for brook trout. This suggests that something other than the correlation of these habitat variables with stream order is driving the differences in where brook trout are found in Pennsylvania and northeast Ohio. We do not have the data to test whether the correlation between water temperature and stream order is driving the differences between sites. We conclude that future studies should avoid using stream
order, or other surrogate measures of primary drivers of habitat suitability, when combining data from two regions.

Similar to our results, brook trout streams in other areas also typically have relatively low gradients. Trout streams in the upper Midwest are typically low gradient (Alexander and Hansen 1986). Brook trout are typically found in low to moderate gradients in the Western United States, with slopes less than 7% (Fausch 1989, from Adams et al. 2000). Approximately 5% of the known brook trout streams in our combined study area had a slope of over 10% and 75% of the known brook trout streams in our combined study had a slope less than 5%. Steep slopes may inhibit the movement of small brook trout (Adams et al. 2000). Brook trout less than 95 millimeters in length moved less often in streams with a slope greater than 6% than in those with a slope less than 3% (Adams et al. 2000).

Lithology and presence of sand and gravel in the substrate influenced predicted habitat suitability in our study. Streams with a higher percent of sandstone as the primary lithology had a higher likelihood of being able to support brook trout than streams with little or no sandstone. While lithology has been shown to influence stream fish distribution in other studies (e.g., trout in New Zealand are associated with volcanic ash, Jowett 1990), we know of no other stream fish species associated with a sandstone lithology. We are unsure as to why increasing the percent of sandstone increases the likelihood of a stream being able to support brook trout.

Sand and gravel substrate (data available from Ohio only) were added to the original Maxent model (data from Ohio and Pennsylvania), as gravel is believed to be an important feature in determining brook trout persistence in northeast Ohio (Amey,
OEPA, unpublished data; Pira et al., Geauga County Metro Parks, unpublished data). Gravel provides appropriate substrate for nests, offering adequate space for flowing water to oxygenate eggs and remove metabolic waste (Coble 1961). The addition of sand and gravel substrate reduced the number of potentially suitable brook trout streams from 39 to 27 in northeast Ohio. This low number is unsurprising given the low number of streams that are currently known to support brook trout. These streams are spread out across the region (Figure 5), which may be why they were not identified in previous attempts to locate potential brook trout streams.

Percent forest near the stream, stream order, slope, and sand and gravel substrate all together drive a stream’s ability to support brook trout. High percent forest near streams and low stream order interact to keep water temperatures cold. Slope has been used as a surrogate for physical habitat in streams (Isaak and Hubert 2000). Here, low slope may allow for the formation of pools that could provide habitat for adult fish to seek sanctuary from predators. As previously stated, the gravel in the sand and gravel substrate could provide areas for brook trout reproduction in northeast Ohio streams. These variables together demonstrate that a stream has the habitat necessary to support brook trout of all ages and thus, possibly, population persistence.

**Bigmouth shiners**

In contrast to brook trout, bigmouth shiners are predicted to be found in relatively unforested landscapes. Our model suggested that the likelihood of persistence increased as the percent forest near the stream decreased to a point. Similar to this, a niche model constructed for the bigmouth shiner across the central United States found that, as the
percent of row crops in a watershed increased, the likelihood of bigmouth shiner persistence increased (Bouska et al. 2015). It is not surprising that bigmouth shiners are successful in the Great Plains (Nico and Fuller 2016), given that they do well in agricultural landscapes.

Stream order and elevation also affect the distribution of bigmouth shiners in northeast Ohio. Streams with a 4th- or 5th-order and streams at relatively low elevations favored bigmouth shiners. High order streams tend to be deeper and wider than low-order streams (Platts 1979) and tend to have a lower gradient (Whiteside and McNatt 1972, Platts 1979). Bigmouth shiners are typically found in large, slow flowing streams (Trautman 1981), which supports our model results that high order streams are more suitable for bigmouth shiners. In this region, high order streams are likely to be found at lower elevations than high order streams in northeast Ohio because low order streams merge and form high order streams as they flow toward Lake Erie.

The presence of silverjaw minnow was not important in determining the presence of bigmouth shiners with our model. We predicted that silverjaw minnows were replacing bigmouth shiners in some streams in northeast Ohio, based on observations by Trautman (1981). Six of 42 streams in our study that are known to support bigmouth shiner also supported silverjaw minnow. Also, there are over 21,000 stream reaches in our northeast Ohio study area, of which only 220 contain known populations of silverjaw minnows. There may not be enough streams in northeast Ohio with silverjaw minnows to detect a relationship with bigmouth shiners. Alternatively, studies have shown that biotic interactions are harder to detect as spatial scale increases (Pearson and Dawson 2003), so this type of modeling may not be suitable for observing this interaction.
Another way to look at the interaction potential interaction between bigmouth shiners and silverjaw minnows was to model the distribution of silverjaw minnows to determine the potential overlap between these species. Only 13 of 136 potential silverjaw minnow streams with a logistic score or 0.75 or above overlapped with the potential distribution of the bigmouth shiner. This suggests that the silverjaw minnows do not likely exclude bigmouth shiners.

Ultimately, we identified 146 more potential bigmouth shiner streams, which is approximately three times the number of currently known streams that support bigmouth shiners.

Limitations of the models

Our results are limited by the fact that we could not include stream temperature in the model, and stream temperature has been shown to be an important predictor of brook trout presence (e.g. Bozek and Hubert 1992; Wehrly et al. 2007). Data on stream temperature are rare except at gauging stations in large streams, but these data do not capture the important differences in stream temperature across stream orders. Air temperature has been used as a surrogate for water temperature in other studies (Wenger et al. 2011), but these models are not as accurate in low-order streams in which temperature may be heavily influenced by ground-water input. Stream temperature is believed to strongly affect habitat suitability for brook trout (MacCrimmon and Campbell 1969), and groundwater input is known to affect stream temperature in low-order streams; our current model is limited by the lack of a model for estimating groundwater inputs.
In combining data between two regions, it is important to work with the variables that are the primary drivers of habitat suitability, such as temperature, rather than variables that are not absolute measures of primary drivers (e.g., stream order). Because many low-order streams in Pennsylvania have higher elevation and slope than low-order streams in Ohio, stream order may be associated with a different suite of habitat characteristics between the two systems. Our model may not be highly accurate in modeling the potential distribution of brook trout in northeast Ohio given the fact that we were unable to use all primary drivers of habitat suitability. In fact, the model predicted that none of the four known northeast Ohio brook trout streams used in the model would be suitable for brook trout, which may suggest that this model is not ideal for predicting the distribution of brook trout in northeast Ohio. We do feel, however, that it would be worth investigating the habitat suitability of the predicted trout streams in northeast Ohio and that the results of this model may still be able to aid in the conservation on brook trout in this region. These results are still derived from regional brook trout data and the small number of streams that are suggested to support brook trout in northeast Ohio would be easily investigated.

Conclusions

Landscape models developed in this study suggest that a number of streams in northeast Ohio may be able to support populations of brook trout or bigmouth shiners. Our novel approach for creating a brook trout model found 27 more streams in northeast Ohio that could support brook trout. The bigmouth shiner model found 146 streams in northeast Ohio that may support bigmouth shiner populations.
We showed that using presence-only data, which can be found in museums and is collected by management agencies across the country, and landscape variables that can be derived from data readily available on-line, we were able to provide predictions on the distribution of two threatened stream fish species in northeast Ohio. However, caution should be used when considering using surrogates to primary drivers when including samples from areas outside of the area of interest. Surrogates can mean different things in different areas. This technique can become an important tool in the conservation of freshwater fishes, which have the highest extinction rates of vertebrates worldwide (Burkhead 2012).
Literature Cited


Burt, A. 2007. Brook trout reintroduction: Lake Erie Drainage, NE Ohio. D. O. W. Ohio Department of Natural Resources, editor, Columbus, OH.


population genetics. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(9): 1513-1524.


### Appendix A: Tables and Figures

#### Tables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean, Range-Ohio</th>
<th>Mean, Range-PA</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>18, 0-100</td>
<td>11, 0-100</td>
<td>Percent of landuse of a given type for a 10 meter buffer around each stream reach</td>
<td>USDA-NSAA</td>
</tr>
<tr>
<td>Forest</td>
<td>45, 0-100</td>
<td>53, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>16, 0-100</td>
<td>14, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wetland/herbaceous</td>
<td>12, 0-100</td>
<td>14, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open water</td>
<td>9, 0-100</td>
<td>8, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Road crossings</td>
<td>1, 0-22</td>
<td>0, 0-46</td>
<td>Number of roads that cross a stream reach</td>
<td>PAMAP, Ohio Department of Transportation</td>
</tr>
<tr>
<td>Stream Order</td>
<td>2, 1-6</td>
<td>2, 1-9</td>
<td>Strahler stream order</td>
<td>NHD</td>
</tr>
<tr>
<td>Slope</td>
<td>0.01, 0-0.03</td>
<td>0.028, 0-1.12</td>
<td>Slope of stream reach (m/m)</td>
<td>DEM</td>
</tr>
<tr>
<td>Elevation</td>
<td>279, 173-402</td>
<td>337, 0-903</td>
<td>Mean elevation of stream reach (meters)</td>
<td>DEM</td>
</tr>
<tr>
<td>Lithology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siltstone</td>
<td>0, 0-94</td>
<td>17, 0-100</td>
<td>Percent of area with a given type of primary bedrock within a 5 meter buffer of a stream reach</td>
<td>USGS-MRP</td>
</tr>
<tr>
<td>Sandstone</td>
<td>17, 0-100</td>
<td>35, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conglomerate</td>
<td>0, 0-0</td>
<td>0, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandstone and conglomerate</td>
<td>0, 0-0</td>
<td>8, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shale</td>
<td>82, 0-100</td>
<td>25, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard (metamorphic/igneous)</td>
<td>0, 0-0</td>
<td>4, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lime</td>
<td>0.00, 0-44</td>
<td>7.29, 0-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Notropis buccatus</em> presence *</td>
<td>N/A</td>
<td>N/A</td>
<td>presence of silverjaw minnow (1=yes, 0=no)</td>
<td>OSUM-Fish Division</td>
</tr>
<tr>
<td>Substrate Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine</td>
<td>8, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine/sand and gravel</td>
<td>19, 0-100</td>
<td>N/A</td>
<td>percent of area with a given substrate type within a 5 meter buffer of a stream reach</td>
<td>Ohio Geological Survey</td>
</tr>
<tr>
<td>Sand and gravel/fine</td>
<td>4, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sand and gravel/till</td>
<td>0, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Till</td>
<td>27, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Till/sand and gravel</td>
<td>42, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td>0, 0-100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Bigmouth shiner model only*

Table 1. Potential model variables, including a brief description, mean, and range in northeast Ohio and Pennsylvania, used for predicting brook trout and bigmouth shiner habitat suitability with the source of the data listed. Source abbreviations: USDA-NSAA, United States Department of Agriculture, National Agriculture Statistics Service; NHD, National Hydrography Dataset; DEM, Digital Elevation Model; USGS-MRP, U.S. Geological Survey’s Mineral Resources Program; OSUM Fish Division, The Ohio State Museum of Biological Diversity, Fish Division.
Table 2. Potential model variables. An ‘x’ signifies a variable used in a given Maxent model. Eleven variables were used in the initial brook trout model, and that number was reduced to five for the final brook trout model. Substrate type was not considered for the brook trout Maxent model as these data were available only for Ohio and not Pennsylvania. Substrate type was added to the results of the brook trout Maxent model and is denoted by an *. Eleven variables were also used in the initial bigmouth shiner model, and that number was reduced to three for the final bigmouth shiner model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Initial brook trout model</th>
<th>Final brook trout model</th>
<th>Initial bigmouth shiner model</th>
<th>Final bigmouth shiner model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Forest</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Agriculture</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Wetland/herbaceous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open water</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Road crossings</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Stream Order</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Slope</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Elevation</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Lithology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siltstone</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sandstone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conglomerate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandstone and conglomerate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard (metamorphic/igneous)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lime</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Notropis buccatus</em> presence</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Substrate Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine/sand and gravel</td>
<td>N/A</td>
<td>X’</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Sand and gravel/fine</td>
<td>N/A</td>
<td>X’</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Sand and gravel/till</td>
<td>N/A</td>
<td>X’</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Till</td>
<td>N/A</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Till/sand and gravel</td>
<td>N/A</td>
<td>X’</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

*Added to Maxent results for northeast Ohio
Table 3. Correlation matrix for all of the potential variables for the brook trout model in northeast Ohio and Pennsylvania. None of the correlations were greater in magnitude than the acceptable level of 0.6, with the strongest correlation being -0.48 between percent shale and percent sandstone.
Table 4. Correlation matrix for all of the potential variables for the bigmouth shiner model in northeast Ohio. None of the correlations were greater than our cutoff of ±0.6 with the largest correlation being -0.44 between percent forest and percent urban.
<table>
<thead>
<tr>
<th>NHD ID Number</th>
<th>UTM</th>
<th>Easting</th>
<th>Northing</th>
</tr>
</thead>
<tbody>
<tr>
<td>128581044</td>
<td>17</td>
<td>465129.19</td>
<td>4586068.46</td>
</tr>
<tr>
<td>128581588</td>
<td>17</td>
<td>469575.40</td>
<td>4602529.50</td>
</tr>
<tr>
<td>128581535</td>
<td>17</td>
<td>477080.60</td>
<td>4606276.09</td>
</tr>
<tr>
<td>128584219</td>
<td>17</td>
<td>467640.65</td>
<td>4602103.74</td>
</tr>
<tr>
<td>128581572</td>
<td>17</td>
<td>475219.79</td>
<td>4603851.42</td>
</tr>
<tr>
<td>128580755</td>
<td>17</td>
<td>467424.50</td>
<td>4594088.77</td>
</tr>
<tr>
<td>128584246</td>
<td>17</td>
<td>466465.74</td>
<td>4599150.42</td>
</tr>
<tr>
<td>123062083</td>
<td>17</td>
<td>453208.49</td>
<td>4565833.58</td>
</tr>
<tr>
<td>123061983</td>
<td>17</td>
<td>453678.82</td>
<td>4565468.37</td>
</tr>
<tr>
<td>123059937</td>
<td>17</td>
<td>476035.45</td>
<td>4555544.38</td>
</tr>
<tr>
<td>123061417</td>
<td>17</td>
<td>454811.54</td>
<td>4569676.68</td>
</tr>
<tr>
<td>123061272</td>
<td>17</td>
<td>450597.12</td>
<td>4571473.85</td>
</tr>
<tr>
<td>123062776</td>
<td>17</td>
<td>453749.64</td>
<td>4580755.19</td>
</tr>
<tr>
<td>150440886</td>
<td>17</td>
<td>490207.28</td>
<td>4616531.48</td>
</tr>
<tr>
<td>150440876</td>
<td>17</td>
<td>490870.43</td>
<td>4613703.86</td>
</tr>
<tr>
<td>155389259</td>
<td>17</td>
<td>366062.41</td>
<td>4564356.21</td>
</tr>
<tr>
<td>123062236</td>
<td>17</td>
<td>450463.52</td>
<td>4558861.31</td>
</tr>
<tr>
<td>150440919</td>
<td>17</td>
<td>484068.31</td>
<td>4607881.39</td>
</tr>
<tr>
<td>151057786</td>
<td>17</td>
<td>397062.72</td>
<td>4560253.20</td>
</tr>
<tr>
<td>123061329</td>
<td>17</td>
<td>451783.99</td>
<td>4570216.04</td>
</tr>
<tr>
<td>123057494</td>
<td>17</td>
<td>447505.07</td>
<td>4580769.12</td>
</tr>
<tr>
<td>155388346</td>
<td>17</td>
<td>380428.51</td>
<td>4545770.14</td>
</tr>
<tr>
<td>150440925</td>
<td>17</td>
<td>486176.33</td>
<td>4611372.62</td>
</tr>
<tr>
<td>123062194</td>
<td>17</td>
<td>449369.30</td>
<td>4556946.79</td>
</tr>
<tr>
<td>123062239</td>
<td>17</td>
<td>450563.21</td>
<td>4563463.15</td>
</tr>
<tr>
<td>150440934</td>
<td>17</td>
<td>485634.98</td>
<td>4613484.23</td>
</tr>
<tr>
<td>155388521</td>
<td>17</td>
<td>382069.43</td>
<td>4561358.33</td>
</tr>
</tbody>
</table>

Table 5. NHD ID number and UTM coordinates of predicted brook trout streams in northeast Ohio using Maxent model that included all streams in Pennsylvania that fell within the range of habitats found in northeast Ohio and considering the presence of sand and gravel.
FIGURES

Figure 1. A map of Ohio with our primary study area in blue. The inset in the upper left corner shows the main stems of major watersheds of interest (black lines) and black dots representing streams in which brook trout occur and open circles representing streams where re-introduction was attempted but failed (Burt 2007).
Figure 2. The average receiver operating curve for the final brook trout model including all streams within the habitat range found in northeast Ohio with four variables over the 10 runs. This model has an average AUC of 0.80 and a standard deviation of 0.02. The 1:1 line shows what the prediction would be like if the distribution was due completely to random chance.
Figure 3. The average receiver operating curve for the final brook trout model including only streams from the Pennsylvania dataset that with habitats that are within the range of their respective stream order in northeast Ohio with four variables over the 10 runs. This model has an average AUC of 0.81 and a standard deviation of 0.011. The 1:1 line shows what the prediction would be like if the distribution was due completely to random chance.
Figure 4. Environmental response curves for the variables in the final brook trout Maxent model. The response curves are constructed using only the corresponding variable. Each curve also shows its standard deviation in a different color shading. The black dots represent the given environmental value of each of the four known brook trout streams in northeast Ohio.
Figure 5. A map of the stream reaches in northeast Ohio with black dots representing streams in which brook trout have been reported, open circles representing streams where re-introduction was attempted but failed, red dots representing streams with a logistic score from Maxent greater than 0.75, and red triangles representing streams with a logistic score from Maxent greater than 0.75 and a sand and gravel substrate. The main stems of Black, Rocky, Chagrin, and Grand Rivers are denoted by the dark lines.
Figure 6. The average receiver operating curve for the final bigmouth shiner model with three variables over the 10 runs. This model has an average AUC of 0.938 and a standard deviation of 0.034. The 1:1 line shows what the prediction would be like if the distribution was due completely too random chance.
Figure 7. A map of the stream reaches in northeast Ohio with black dots representing streams in which bigmouth shiners have been reported and red dots representing streams with logistic score from Maxent greater than 0.75. The main stems of Black, Rocky, Chagrin, and Grand Rivers are denoted by the dark lines.
Figure 8. Environmental response curves for the variables in the final bigmouth shiner Maxent model. The response curves of constructed using only the corresponding variable. Each curve also shows its standard deviation in blue color.