HYDROGEOEMORPHOLOGICAL FACTORS
INFLUENCING THE DISTRIBUTION OF CICHLID NESTING SITES
IN THE BLADEN RIVER, BELIZE

by

EMILY A. BUEGE

SARAH J. PRASKIEVICZ, COMMITTEE CHAIR
PETER C. ESSELMAN
MICHAEL K. STEINBERG

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Geography
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2017
Cichlids are among the most diversified families of fish in Central America and thus represent an important component of aquatic biodiversity. Cichlid nesting site selection is important for assemblages that face an uncertain future due to potential invasion by alien species and environmental change. We determined what hydrogeomorphological factors correlate to the preferred nesting sites of four native cichlid species in the Bladen River, Belize. The study was conducted in the portion of the river draining the Bladen Nature Reserve, an area relatively untouched by recent anthropogenic influences. We developed innovative two-dimensional species distribution models to assess the relative influences of physical variables on the distribution of occupied nesting sites. We recorded the locations of nesting sites of four native cichlid species and collected hydrological and geomorphic data through the study reaches, including flow velocity, depth, sediment type, and fish cover. Nest locations and physical habitat data were used to construct spatially explicit habitat models using the maximum entropy approach (Maxent). Models were tested against external data collected in a validation reach and by cross validation in the training reach. The models provided some evidence that physical habitat variables influence the distribution of the nesting sites. We found that two species were most heavily influenced by water depth (*A. spilurus* and *V. maculicauda*), one was influenced by sediment type (*C. salvini*), and one was most responsive to water velocity (*T. meeki*). The distribution of nests for the smaller-bodied species tended to be more closely related to the habitat parameters we measured. This study revealed that nesting sites within this cichlid
community are more predictable than random and are at least partially governed by physical habitat controls. Understanding the physical controls governing cichlid nesting locations can aid in detecting suitable areas for conservation and predicting impacts by climate change and the encroaching logging, mining, and agricultural industries.
DEDICATION

I dedicate this thesis to Scott and Ann Buege, who have always encouraged my fascination with the natural world. I wouldn’t be where I am today without the use of the kitchen cups and spoons to catch and examine tadpoles and trout fry in the backyard.
ACKNOWLEDGEMENTS

We express gratitude to the staff of the Belize Foundation for Research and Environmental Education (BFREE) for providing food, lodging, and company during data collection. Assistance from Jacob Marlin and Heather Barrett of BFREE was imperative to the planning and execution of fieldwork. For access to the aerial imagery, we thank National Geographic, the University of California San Diego, Alexis Thomas of the Geoplan Center at the University of Florida, James Rotenberg of the Department of Environmental Science at University of North Carolina Wilmington, and BFREE. Thank you to the Belize Fisheries Department for granting the required permits to work in the Bladen Nature Reserve and the Ya’axche Conservation Trust for supporting this work. We appreciate field hands Melito Bustamente and Melvin Choc for their hard work and attention to detail. We thank Mike Johnson of UC Santa Barbara for help with data transformation and Dr. David Keellings of the University of Alabama for his assistance in R programming.

Funding was provided by the University of Alabama Research Grants Committee (to SP), University of Alabama graduate student funding sources (to EB), American Association of Geographers Water Resources Specialty Group (to EB), and the Edward H. Moseley Endowed Scholarship (to EB).
CONTENTS
ABSTRACT................................................................................................................... ii
DEDICATION................................................................................................................... iv
ACKNOWLEDGEMENTS.............................................................................................. v
LIST OF TABLES............................................................................................................ vii
LIST OF FIGURES........................................................................................................ viii
CHAPTER 1: INTRODUCTION......................................................................................... 1
CHAPTER 2: MATERIALS AND METHODS .................................................................... 8
  2.1 Study Site ............................................................................................................. 8
  2.2 Biological Sampling ............................................................................................ 9
  2.3 Habitat Sampling ................................................................................................ 10
  2.4 Species Distribution Modeling .......................................................................... 10
  2.5 Model Validation ............................................................................................... 12
CHAPTER 3: RESULTS................................................................................................. 14
  3.1 Biological Data ................................................................................................. 14
  3.2 Model Evaluation .............................................................................................. 16
CHAPTER 4: DISCUSSION........................................................................................... 19
CHAPTER 5: CONCLUSIONS....................................................................................... 26
REFERENCES............................................................................................................... 27
APPENDIX.................................................................................................................. 33
LIST OF TABLES

Table 1: Summary of habitat characteristics for nesting locations by species ..................... 15
Table 2: Nearest neighbor contingency table........................................................................ 16
Table 3: Summary of Model AUC outputs............................................................................. 16
LIST OF FIGURES

Figure 1: Map of study location ................................................................. 34
Figure 2: Field sampling photos ............................................................ 35
Figure 3: Model input layers for MaxEnt .................................................. 36
Figure 4: Observed locations of cichlid nests .......................................... 37
Figure 5: MaxEnt presence prediction maps by species ............................. 38
Figure 6: MaxEnt predicted presence for cichlids in validation reach ........ 39
Figure 7: Variable contribution graphs generated by MaxEnt .................... 40
Figure 8: MaxEnt model outputs for cichlid community ......................... 41
CHAPTER 1
INTRODUCTION

The relationship between an organism and its physical environment is hypothesized to be a driver of adaptive radiation, or the evolutionary response to fill ecological niche vacancies (Stearns, 1976). The family Cichlidae is a group of fish well known for species diversification to fill niches in lake and river environments across the Americas, Africa, and Asia. With an estimated 3,000 species worldwide, cichlids are not only the most speciose family of fishes, but of all vertebrates (Kocher, 2004). In Central and South America, nearly 600 species of cichlid have been identified, making them the third most diverse group of fishes in the Neotropics (Ries et al., 2003). In contrast with African cichlids, which underwent rapid adaptive radiation from common ancestors in the past 10 million years (Kocher, 2004), neotropical cichlids have a relatively long evolutionary history (López-Fernández et al., 2010). Because of the unusually rapid evolutionary timescale of the diversification of African cichlids, they have attracted the majority of the attention when it comes to studying cichlid phylogeny (Kocher, 2004; López-Fernández et al., 2013). However, the vast ecomorphological diversity in cichlid species across both groups is largely driven by adaptive radiation (Winemiller et al., 1995; Kocher, 2004; López-Fernández et al., 2013). Today this family exhibits a wide range of morphological and behavioral features. The growing body of work linking Neotropical cichlid species with their physical habitat has focused primarily on trophic morphology, as their jaw and gut morphologies are highly variable across species. This area of study helps link habitat selection with trophic activities (e.g. Cochran-Biederman and Winemiller, 2010), but there is an apparent lack of
literature that examines the relationship between habitat selection and other vital life history traits, such as reproduction.

Despite being one of the most widely-distributed families of fish in the world, most cichlid species share a life history trait in that they provide parental care to their young (McKay, 1977). Most are considered to be $k$-strategy species in that they invest a high amount of energy into relatively few offspring (Barletta et al., 2010). All American cichlid species north of Panama exhibit life history strategies that involve laying eggs on substrates and both parents defending the young through the larval stage (Barlow 1974, as cited by McKay, 1977). This mobility-limiting trait is typically combined with a tendency to exhibit territoriality by aggressively defending the nesting site (McKay, 1977).

Similar to knowledge of cichlid evolution, the existing body of work on cichlid reproductive biology primarily pertains to African species (McKay et al., 2010). Further, in situ studies regarding Neotropical cichlid reproduction have focused on lacustrine over riverine populations (e.g. McKay et al., 2010; Vivas & McKay, 2001; McKay, 1977). One reason for this could be that riverine fish communities in the tropics are not primarily composed of cichlid species as lakes often are (Winemiller et al., 1995). Reproductive behaviors in rivers may not necessarily align with findings in these lacustrine studies because the physical controls in rivers are more dynamic and because river fish assemblages have generally undergone natural selection for a longer period of time than lake fish assemblages (Lowe-McConnell, 1969).

Relatively undisturbed native cichlid populations are necessary to study the family in the context of evolutionary ecology, because responses to evolutionary drivers can be obscured by responses to anthropogenic stressors. Presently, some Neotropical cichlid assemblages in protected waters boast capacity for quality in situ studies (eg. Cochran-Biederman & Winemiller,
2010), but the threat imposed by land cover change and non-indigenous species, such as African
tilapia, may weaken the potential to uncover natural evolutionary relationships. Perhaps the
tilapia’s impact on scientific research is less of a concern than their proven destructive potential.
Throughout their nonnative range, tilapias have competed with other tropical fish species for
space and other resources (Canonico et al., 2005). Canonico et al. (2005) summarize that tilapia
are exceptionally successful in reproducing, as they are aggressive enough to push native species
out of prime breeding habitat and can reproduce year-round in the right conditions. In Lake
Victoria, native species declined after the offspring of introduced tilapia outcompeted the native
young in nursery areas (Lowe-McConnell, 1975). Additionally, because tilapia feed on multiple
trophic levels, large populations can eliminate food sources and breeding habitat for native fish
(Canonico et al., 2005). Although the tilapia invasion is widespread, its influence on native
species is under-studied, particularly in Mesoamerica (Esselman et al., 2013).

Body morphology may be a predictor of habitat selection; previous studies have shown a
link between preferred habitat and ecomorphological characteristics (Winemiller et al., 1995;
Gatz, 1979; Poff, 1995). Specifically, deeper-bodied fish have been known to seek environments
with slower-moving water (Gatz, 1979; Poff, 1995), whereas fish with more streamlined body
shapes have greater maneuverability, and therefore can withstand higher water velocities
(Videler & Wardle, 1991). Fish are often divided into functional groups based on their
ecomorphologies called guilds, which could be based on habitat use (Vadas & Orth, 1997) or life
history styles (Balon, 1975). Trophic specializations in fish allow for the development of guilds
based on feeding strategies as well (Kassam et al. 2003). The diversity of trophic morphologies
in cichlids is reflected by body types specialized for the microhabitats that need to be accessed in
order to feed (Mérigoux & Ponton, 1998; Kassam et al. 2003; Winemiller et al., 1995).
In the field of landscape ecology, rivers are often viewed as single homogenous components of a greater, primarily terrestrial environment (Wiens, 2002). Aquatic ecologists, however, tend to adhere to the idea that riverine systems are composed of a mosaic of heterogeneous habitat patches (Wiens, 2002). What defines the boundaries of an individual patch will vary by spatial scale, which can be addressed using a hierarchical framework, such as that described in Frissell et al. (1986). It has been postulated that aquatic biota assemblages are influenced largely by the abiotic controls present at a small scale, but small-scale conditions can only operate within the constraints of the large-scale system (Poff, 1997; Frissell et al., 1986). Ultimately, the interactions between ecological processes and spatial patterns can and should be examined across a variety of spatial scales (Poff, 1997; Ward et al., 2002; Allan & Castillo, 2007).

Spatial patterns in the distribution of species have been shown to heavily influence population biology (Fonteyn & Mahall, 1978). Spatial partitioning via competitive interactions is a topic ecologists easily come to question, particularly when the space is limited by confines such as the edges of a river channel. Understanding species distributions is essential in determining habitat suitability, predicting future conditions in a changing world, and many other aspects of environmental and biological sciences (Franklin, 2009). Poff (1997) describes the interplay between habitat “filters,” or the physical constraints acting on species distribution at various spatial scales, and the functional traits of the organisms present. The functional traits of an organism must be subject to natural selection, and the abiotic factors present in the organism’s environment compose the foundation on which natural selection can operate (Moyle & Cech, 1994). On a finer scale, biotic interactions (i.e. spatial competition, predation) may also play a role in community structure and distribution, but the microhabitat-scale is also where the
individual exists and its functional traits are actually employed. In the present study, we approach the riverine landscape at a fine scale, which can reveal in detail how biota respond to physical habitat variables (Weins, 2002).

Water depth, flow, and substrate are important microhabitat characteristics for stream fish assemblages (Gorman & Karr, 1978). Previous studies have linked fish assemblages with hydrologic variables, such as current velocity (Poff & Allan, 1995), and other physical controls like preferred substrates for nesting (Henley et al., 2000). More recently, Bey & Mazeika Sullivan (2015) showed that water velocity, channel dimensions, and sediment size were all predictors of biota in Big Darby Creek of southern Ohio, USA. However, like most of the existing work done on physical controls of aquatic species, this study was conducted in a midlatitude setting. Although it is plausible that these same habitat characteristics are predictors of cichlid nesting sites in systems such as the Bladen River, tropical montane rivers may vary in hydrogeomorphological properties from their temperate counterparts (Wohl, 2005). For example, tropical rivers lack influence by glaciation and tend to have more frequent and higher rates of precipitation per unit area (Wohl, 2005). Accordingly, tropical rivers generally demonstrate highly dynamic responses to changes in external drivers, which may require that aquatic species be adapted to a range of physical habitat conditions to withstand frequent disturbance.

Although previous studies have examined stream habitat suitability using statistical methods (Ahmadi-Nedushan et al., 2006; Lamouroux et al., 1998; Esselman & Allan, 2011), few have accounted for horizontal variation within the channel. Instead, most studies predict habitat suitability using inter-confluence river reaches as basic analytical units. Species distribution models (SDMs) can be employed in aquatic environments such as lakes and oceans (e.g. Poulos
et al. 2016; Sequeria et al., 2012; Olden & Jackson, 2002), but become complicated when applied to stream environments, where organisms are highly dependent on the arrangement of distinct longitudinally-distributed habitat units (Liang et al., 2013). This arrangement is so longitudinally-derived that riverine SDMs are often unidirectional (e.g. Elith et al., 2011) or detect suitable habitat patches throughout a wetland or watershed instead of being confined to the river channel (e.g. Joy and Death, 2004; Liang et al., 2013). Although this method of predictive modeling is suitable for pinpointing hotspots of vagile individuals or groups of organisms, it ignores the lateral use of space in the channel, which could vary enough to influence nesting site selection. This is the first study to explore two-dimensional Maximum Entropy (Maxent) species distribution modeling in a riverine environment. It is also important to note that species distribution modeling in tropical environments is an under-studied area, despite having high potential for management and conservation implications (Phillips et al., 2006; Anderson et al., 2002).

In this study, we use the Maxent approach to model the distribution of nesting sites of four native cichlid species in the Bladen River, *Archocentrus spilurus* (blue-eyed cichlid), *Cichlasoma salvini* (yellow belly cichlid), *Vieja maculicauda* (black belt cichlid), and *Thorichthys meeki* (fire mouth cichlid). We hypothesized that the distribution of nesting sites would relate to the physical variables we recorded. Further, we hypothesized that the smaller-bodied *A. spilurus*, *C. salvini*, and *T. meeki* would nest in deeper portions of the channel and would be able to withstand faster water velocities. We also hypothesized that *V. maculicauda* would have a higher affinity for slower-moving, shallow areas of the channel as a result of the larger, thicker-bodied morphology of this species. Additionally, little data regarding the physical parameters of rivers in Belize exist (Esselman & Boles, 2001), and the present study serves to
build on the existing information, further extend the current understanding of the physical controls governing fish assemblages, and produce an innovative application of 2-dimensional SDMs in a riverine environment.
CHAPTER 2
MATERIALS AND METHODS

2.1 Study Site

Approximately 36.5% of Belize’s terrestrial territory is part of a protected area (Meerman, 2005). The Bladen Nature Reserve (average elevation of 45 meters, Cochran-Biederman & Winemiller, 2010) of the Toledo district of southern Belize is a protected area of dense and essentially pristine rainforest. The reserve is a biodiversity hotspot for a multitude of species including 250 avian, 93 mammalian, and 92 reptilian and amphibian species (Iremonger & Sayer, 1994). Comparatively, the diversity of fish species is rather low with 18 species recorded, six of them being cichlids (Walker & Walker, 2006). Aquatic biodiversity has been shown to decrease with increasing distance from the coast, perhaps in part due to physical barriers such as waterfalls (Esselman et al., 2006) and because conditions generally become less variable with increasing stream order (Moyle & Cech, 1994). The reserve is almost completely surrounded by other national protected lands including Chiquibul National Park, Cockscomb Basin Wildlife Sanctuary, and three forest reserves. This region has remained undisturbed in recorded human history, although unexcavated Mayan ruins can be found throughout the reserve (BFREE Head Ranger, Sipriano Canti, personal communication, May 5, 2016) probably dating back to the Classic Mayan Era 300-900 AD.

The Bladen River (Figure 1) at the study site drains approximately 400 km² of the southeastern part of the Maya Mountains in the Bladen Nature Reserve (Day, 1996). It is one of the three major tributaries to the sixth-order Monkey River, all originating in a protected inland area,
and then flowing through a developed coastal plain area (Esselman et al., 2006). Between 330 cm and 406 cm of precipitation falls in the Bladen Nature Reserve annually on average (Walker & Walker, 2006). The Upper Bladen River is a good study location for this research because recent anthropogenic influences on the physical habitat controls are minimal, offering insight into the natural processes governing cichlid nesting site distribution.

The focal reach within the Bladen River extended from just above the Belize Foundation for Research and Environmental Education field site upriver 700 meters. The data from this main study reach were used to train the models, and a second, 250-meter reach was sampled beginning at the start of the first run. The data from the second reach functioned as validation data for the models. Data was collected during base flow conditions between 28 April and 15 May 2016 during the seasonal dry period, when water levels are low and consistent. The study reaches are situated at the upstream extent of the outwash plain where the river leaves the mountain valley.

The main, 700-meter reach (henceforth the “calibration reach”) contained, from upstream to downstream, a deep pool, a riffle, a back pool, and a run. The 250-meter upstream reach (henceforth the “validation reach”) was chosen to reflect the channel habitat units present in the calibration reach and also contains a pool, a riffle, a back pool, and a run.

2.2 Biological Sampling

To note the cichlid nest locations, the lead author (the observer) snorkeled three equally spaced longitudinal passes parallel to shore in a downstream direction. When the observer encountered a nesting pair of cichlids, a target was held over the water, directly above the nest (Error! Reference source not found.A). A recorder then used a digital compass and laser range finder to record the distance and bearing of the target from one of 36 known locations on
shore that had been previously surveyed using a handheld GPS in the calibration reach, and another 13 surveyed locations in the validation reach. Sampling for the calibration reach was replicated three times over three consecutive days, with data from the third day used to build the model.

2.3 Habitat Sampling

Sampling of physical habitat variables occurred after all biological sampling was completed to avoid disturbing cichlid nests. Transects were spaced 20 meters apart. At each transect, depth was measured using a stadia rod (Error! Reference source not found.) every two meters across the channel throughout the calibration reach and at five equally spaced intervals across transects in the validation reach. The sample densities were different between reaches because of another concurrent data collection project in the calibration reach. Irregular changes in depth that would not be reflected by interpolation were recorded using a compass and laser range finder from known locations on shore. Velocity in both reaches was measured at five equally spaced intervals across the channel using a Marsh-McBirney® digital current meter, and was recorded at approximately half the depth of the water. Substrate was documented every two meters across the channel by visually classifying the dominant particle sizes available for nesting (detritus, very fine, fine, medium, large, and extra-large/bedrock).

2.4 Species Distribution Analysis and Modeling

Density estimation functions that have been in development over the last century provide ecologists with powerful tools to analyze spatial pattern (Haase, 1995). Spatial segregation is a topic ecologists naturally come to question – are species more likely to be near conspecifics or do they seek to be near individuals of other species? The null hypothesis in this context is that individuals are “randomly labeled,” or that an individual is equally likely to be near an individual
of its own species as it is to another species (Pielou, 1961). This has been assessed for many different species, but rarely more than two species at a time. Dixon (2002) explains the value in using nearest-neighbor contingency tables for assessing multiple species, although the “species” could be substituted for things like age class, sex, or any other categorical variable. The table is based largely on a measurement of segregation, denoted in the ‘dixon’ R package as variable “S” (de la Cruz Rot, 2015). This is the degree to which the expected number of nearest neighbors of the other species does not match the observed number. To determine what relationships exist between species, we constructed a nearest neighbor contingency table using package ‘dixon’ in R Studio 3.3.2 (de la Cruz Rot, 2015). By using a nearest neighbor analysis tool (ArcMap 10.2), we assessed whether the spatial pattern of the overall community was randomly structured or not.

Maximum entropy modeling (MaxEnt) is a widely used method for modeling the geographic distributions of species where presence-only (PO) data is available. The underlying assumption behind MaxEnt is that the situation in which the species is most uniformly distributed (exhibiting maximum entropy) will be the most probable distribution for the species of interest, so long as underlying physical constraints are considered (Phillips et al., 2006). Extensive literature exists regarding the statistical theory behind MaxEnt (Phillips et al., 2006; Merow et al., 2013). We ran our models in R Studio 3.3.2 using the dismo package (Hijmans et al., 2011).

Four variables were used as predictors in the models: water depth (Error! Reference source not found.A), water velocity (Error! Reference source not found.B), dominant substrate (Error! Reference source not found.C), and fish cover (Error! Reference source not found.D). Preparation of the features was done in ArcMap 10.2. The two continuous variable
layers, depth and velocity, were interpolated using the kriging tool in the Spatial Analyst Toolbox. Substrate patches were delineated by creating Thiessen polygons for each data point, and then dissolving like polygons. Fish cover was hand-delineated using field notes, photographs and 10-centimeter resolution aerial imagery that was collected in December 2015, the beginning of the same seasonal dry period in which the biological and physical data were collected (Jamie Rotenberg, UNC Wilmington). Four main types of fish cover were recorded: small woody debris, large woody debris, extra-large boulders with overhang, and overhanging vegetation. These four fish cover classes were eventually dissolved into two classes – submerged cover (all woody debris and boulders) and canopy cover.

MaxEnt requires the use of background or ‘pseudo-absence’ points from which the method discriminates species presence locations. These were randomly generated in ArcMap 10.2. Pseudo-absences can negatively affect model output if they appear at the same location of a presence point. We used 0.5-meter exclusion buffers around each presence location and removed points that fell inside. We were left with 1,985 background points for the calibration reach and 995 background points in the validation reach.

2.5 Model Validation

Because of the exploratory nature of this study, we chose to use several methods of evaluating model performance. First, we used a five-fold cross validation technique, in which the presence locations for each species are divided into five groups (e.g. Hijmans & Elith, 2017). Twenty percent of the presence points were withheld, and the remaining 80% were used to construct the model. This was repeated five times, once for each fold of the data to achieve the Area under the Receiver Operated Curve (AUC). Plotting the frequency of true positives against the false positives allows the Receiver Operated Curve to be drawn, and the area under the curve
to be calculated (Phillips et al., 2006). The advantage of using AUC values over other model assessment metrics is that it is independent of arbitrary thresholds (Phillips et al., 2006). AUC average and standard deviation were calculated for ten replicate runs. Second, we used the third day of data collection in the calibration reach to build the model, then used the data collected on the first and second days for validation. Finally, we used cichlid nesting locations collected in the validation reach to evaluate the model. All model evaluation methods were repeated for each of the four species. We also combined presence locations of all four species and built a model using the entire cichlid assemblage.

Using AUC as the primary method of evaluation has been criticized (Jimenez-Valverde, 2012). However, one of the strongest arguments against the use of AUC is that increasing the spatial extent of the background points tends to inflate the fit of the model (Hijmans & Elith, 2017; Jimenez-Valverde, 2012). Because our models are confined to the shape of the channel and our environmental layers are limited to the study extent, inflation of AUC values is not an issue.

Package ‘dismo’ also has an option to produce variable contribution information after the model is trained. We used the contribution information to determine to what extent depth, water velocity, substrate, and fish cover influence the distribution of cichlid nesting locations. Based on knowledge of body morphology of the four study species and morphology-habitat linkages, these models also aid in drawing conclusions about how body shape may influence nesting site selection.
3.1 Biological Data

Four species of cichlid were observed nesting within the study region during the time of sampling: *Archocentrus spilurus*, *Cichlasoma salvini*, *Vieja maculicauda*, and *Thorichthys meeki* (Figure 4). Two other native cichlid species, *Astatheros robertsoni* (false fire mouth cichlid) and *Petenia splendida* (bay snook), and one nonnative species, *Oreochromis spp.* (African tilapia), were also present during sampling, but were not observed nesting. This is the first published record of *Oreochromis spp.* in the Upper Bladen River (n = 5). Sample sizes of nesting pairs as used for model calibration were as follows: *A. spilurus* – 26, *C. salvini* – 52, *V. maculicauda* – 64, and *T. meeki* – 15. Data from the remaining two days were used as an additional method of validation. The validation reach was sampled once, and sample sizes were: *A. spilurus* – 39, *C. salvini* – 26, *V. maculicauda* – 21, and *T. meeki* – 3.

Across four days of biological sampling, a total of 719 cichlid nesting locations were documented (nest density = 0.012 nests/m²). The nearest neighbor analysis in ArcMap 10.2 showed clustered distributions (p = 0.005). In the calibration reach, expected mean distance between nests was 4.6 meters and observed mean distance was 4.1. Summary statistics for physical habitat data at nesting locations can be found in Table 1.

The nearest neighbor contingency table (}
Table 2) showed that there was significant clustering within two species, *V. maculicauda* and *C. Salvini* at the p < 0.05 significance level, and *A. spilurus* and *T. meeki* both exhibited clustering at the p < 0.10 significance level. All
interspecific observed nearest neighbor values were higher than the respective expected number. Interestingly, all statistically significant intraspecific nearest neighbors had observed values lower than expected and involved *C. salvini*, indicating that there was a tendency for these species to nest farther apart from one another than expected. These interactions were between *V. maculicauda* and *C. Salvini* and between *T. meeki* and *C. Salvini*.

**Table 1:** Summary of habitat characteristics for nesting locations by species. Values were extracted from MaxEnt model inputs. Standard deviations are found in parentheses.

<table>
<thead>
<tr>
<th>Species</th>
<th>Avg. Depth (m)</th>
<th>Avg. Velocity (m/s)</th>
<th>Dominant Substrate</th>
<th>Canopy Cover (%)</th>
<th>Submerged Cover (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. spilurus</em></td>
<td>1.2 (0.5)</td>
<td>0.06 (0.03)</td>
<td>1 (27.0%)</td>
<td>1.2</td>
<td>7.7</td>
</tr>
<tr>
<td><em>C. salvini</em></td>
<td>0.9 (0.4)</td>
<td>0.06 (0.06)</td>
<td>2 (32.7%)</td>
<td>25.0</td>
<td>5.8</td>
</tr>
<tr>
<td><em>V. maculicauda</em></td>
<td>1.0 (.6)</td>
<td>0.06 (0.06)</td>
<td>2 (37.5%)</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td><em>T. meeki</em></td>
<td>0.9 (0.4)</td>
<td>0.03 (0.009)</td>
<td>2 (40.0%)</td>
<td>0</td>
<td>4.4</td>
</tr>
</tbody>
</table>
Table 2: Nearest neighbor contingency table. Species codes are as follows: BB = *V. maculicauda*, BE = *A. spilurus*, FM = *T. meeki*, and YB = *C. Salvini*.

<table>
<thead>
<tr>
<th>Species From To</th>
<th>Obs. Count</th>
<th>Exp. Count</th>
<th>S</th>
<th>Z</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB BB</td>
<td>42</td>
<td>28.1</td>
<td>0.29</td>
<td>2.8</td>
<td>0.005</td>
</tr>
<tr>
<td>BB BE</td>
<td>19</td>
<td>19.2</td>
<td>-0.010</td>
<td>-0.050</td>
<td>0.96</td>
</tr>
<tr>
<td>BB FM</td>
<td>8</td>
<td>9.9</td>
<td>-0.10</td>
<td>-0.67</td>
<td>0.50</td>
</tr>
<tr>
<td>BB YB</td>
<td>17</td>
<td>28.8</td>
<td>-0.31</td>
<td>-2.69</td>
<td>0.007</td>
</tr>
<tr>
<td>BE BB</td>
<td>17</td>
<td>19.2</td>
<td>-0.08</td>
<td>-0.61</td>
<td>0.54</td>
</tr>
<tr>
<td>BE BE</td>
<td>19</td>
<td>12.8</td>
<td>0.24</td>
<td>1.69</td>
<td>0.091</td>
</tr>
<tr>
<td>BE FM</td>
<td>6</td>
<td>6.7</td>
<td>-0.050</td>
<td>-0.29</td>
<td>0.77</td>
</tr>
<tr>
<td>BE YB</td>
<td>16</td>
<td>19.4</td>
<td>-0.12</td>
<td>-0.95</td>
<td>0.34</td>
</tr>
<tr>
<td>FM BB</td>
<td>9</td>
<td>9.9</td>
<td>-0.060</td>
<td>-0.36</td>
<td>0.72</td>
</tr>
<tr>
<td>FM BE</td>
<td>10</td>
<td>6.7</td>
<td>0.24</td>
<td>1.46</td>
<td>0.14</td>
</tr>
<tr>
<td>FM FM</td>
<td>7</td>
<td>3.4</td>
<td>0.38</td>
<td>1.74</td>
<td>0.082</td>
</tr>
<tr>
<td>FM YB</td>
<td>4</td>
<td>10</td>
<td>-0.51</td>
<td>-2.34</td>
<td>0.020</td>
</tr>
<tr>
<td>YB BB</td>
<td>19</td>
<td>28.8</td>
<td>-0.25</td>
<td>-2.23</td>
<td>0.026</td>
</tr>
<tr>
<td>YB BE</td>
<td>22</td>
<td>19.4</td>
<td>0.070</td>
<td>0.68</td>
<td>0.50</td>
</tr>
<tr>
<td>YB FM</td>
<td>3</td>
<td>10</td>
<td>-0.56</td>
<td>-2.45</td>
<td>0.014</td>
</tr>
<tr>
<td>YB YB</td>
<td>43</td>
<td>28.8</td>
<td>0.30</td>
<td>2.85</td>
<td>0.044</td>
</tr>
</tbody>
</table>

3.2 Model Evaluation

Table 3: Summary of Model AUC outputs.

<table>
<thead>
<tr>
<th>Species</th>
<th>K-fold Average AUC</th>
<th>K-fold Std. Dev.</th>
<th>Day 1 AUC</th>
<th>Day 2 AUC</th>
<th>Validation AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. spilurus</em></td>
<td>0.672</td>
<td>0.068</td>
<td>0.627</td>
<td>0.626</td>
<td>0.657</td>
</tr>
<tr>
<td><em>C. salvini</em></td>
<td>0.641</td>
<td>0.069</td>
<td>0.579</td>
<td>0.580</td>
<td>0.533</td>
</tr>
<tr>
<td><em>V. maculicauda</em></td>
<td>0.544</td>
<td>0.067</td>
<td>0.538</td>
<td>0.593</td>
<td>0.290</td>
</tr>
<tr>
<td><em>T. meeki</em></td>
<td>0.658</td>
<td>0.16</td>
<td>0.678</td>
<td>0.602</td>
<td>0.435</td>
</tr>
<tr>
<td>Community</td>
<td>0.630</td>
<td>0.046</td>
<td>0.630</td>
<td>0.610</td>
<td>0.407</td>
</tr>
</tbody>
</table>
A. spilurus

Of the four cichlid nesting locations we modeled, A. spilurus consistently produced the highest AUC values (Table 3). An AUC value of 0.5 is the same as chance, and as the model’s predictions become more accurate, the AUC will increase. The highest value achievable is 1.0, which would indicate a perfect prediction (Hijmans & Elith, 2017). The k-folding evaluation method showed the highest average AUC for ten replicate runs. Day 2 validation data produced the highest AUC across species for that day, but Day 1 validation output was the second highest by species and below the overall community AUC. The independently-validated method produced the highest AUC of all species by 0.124. The dismo package’s variable contribution output for A. spilurus (Figure 7A) showed that depth was the most important variable in determining nesting location, contributing approximately 70% of the prediction. The prediction map for A. spilurus in the main study reach is seen in Figure 5A and with the validation reach included in Figure 6A.

C. salvini

K-fold validation of the model built for C. salvini nesting locations produced an AUC better than random. AUC decreased with the other validation techniques, though still remained better than random in the alternate day validation trials and marginally better than random when validated using the independent reach. Figure 7B shows that unlike all other species, substrate is the most important factor in predicting C. salvini nesting locations (contributing about 50%). Prediction maps for C. salvini are located in Figure 5B and Figure 6B.
**V. maculicauda**

The SDM for *V. maculicauda* produced the lowest AUC with the k-fold validation technique across the four species (Table 1). Day 1 of alternate day validation decreased the AUC (AUC = 0.538), but the AUC improved using day 2 data. Evaluation using the independent reach produced the lowest AUC of the species by 0.145. Depth was revealed to be the most important predictor variable, contributing approximately 45% (Figure 7C). Figure 5C shows the MaxEnt presence prediction map, and Figure 6C visualizes the predictions incorporating the validation reach.

**T. meeki**

*T. meeki* had the second highest AUC when evaluated using the k-fold method (Table 3), however we note that this species had the highest standard deviation, nearly twice that of the other species. Alternate day validation resulted in AUCs that remained above 0.6. More than 50% of the variable contribution for *T. meeki* can be attributed to velocity, making it the only species with velocity as the most important variable (Figure 7D.) This is also the only species to exhibit fish cover as the second most important explanatory variable, contributing 30% to the model. *T. meeki* has the lowest proportion of nests associated with fish cover, with no nests found under canopy cover and few nests near submerged cover (Table 1). Visualizations of the predicted presence of nesting locations for this species can be found in Figure 5D and Figure 6D.

**Community**

AUC values for k-fold validation and alternate day validation were better than random (Table 3). AUC dropped below 0.5 when evaluated using the independent reach. At the family level, the most important predictor variable was depth (Figure 8A). Figure 8B and Figure 8C depict the predicted locations of cichlid nests.
CHAPTER 4

DISCUSSION

Our results support the assertion that within reach scale nest site selection by the species studied was non-randomly structured and moderately related to abiotic gradients. In particular, the nesting locations of smaller-bodied species *A. spilurus*, *C. salvini*, and *T. meeki*, which produced AUC values better than random, appeared to be at least partially governed by the variables we measured. Models of *V. maculicauda* nest location models performed more poorly than the other species, indicating that this species may be more difficult to predict, or that sample sizes were insufficient to adequately capture its relationships to measured habitat variables. Alternatively, *V. maculicauda* may be more randomly distributed or governed by other, unmeasured variables or interactions. Outside of a reproductive context, food availability and predation risk influence the spatial distribution of fish (Hugie & Dill, 1994) and may be reasons for the seemingly more random distribution of *V. maculicauda*. These topics are discussed in greater detail below.

Understanding the role geomorphic and hydrologic controls play in spatial partitioning for cichlid species on a local scale could offer insights into how changes in river flow regime and other local habitat characteristics affect cichlid assemblages. This is important for all rivers in the face of global climate change, as inter-annual precipitation is predicted to vary more as Earth’s climate warms, especially in the tropics (Trenberth, 2011). In the case of the Bladen River, it is imperative to understand the role of hydrogeomorphological factors in case agricultural, mining and logging activities currently present in the coastal plains region encroach
on the upper portions of the watershed (Esselman et al., 2006). Invasion of African tilapia will also likely influence spatial distributions of native cichlid nesting locations as a biotic influence. Understanding how physical habitat characteristics may limit the possible extent of displaced nesting pairs could have management implications, such prioritizing sites for the conservation of Neotropical cichlids. At the time of sampling, we observed several African tilapia individuals occupying the area, but did not observe nesting behaviors or locate nesting cavities. Models by Esselman et al. (2013) projected the possible upriver extent of the tilapia to be multiple kilometers upstream of our study site, though there were no known individuals present in the Upper Bladen River at the time the models were generated. Their presence indicates that the models and the invasion are indeed threats to native cichlids in the area.

MaxEnt models are in ecological niche theory (Phillips et al., 2006), which suggests that a combination of variables suitable for an organism to live and reproduce dictates the space it will occupy (Hutchinson, 1957). However, the realized niche, or the actual, observed distribution of a species, is often influenced by biotic factors as well (Brown et al., 1991). Because realized niches are the result of a combination of both biotic and abiotic controls, it is difficult to disentangle the contributing factors. For example, the nearest neighbor contingency table suggests that conspecifics tend to cluster and that certain species tend to avoid others. On one hand, the interpretation could be that conspecifics only appear to cluster because they seek similar habitat conditions. On the other hand, these four cichlid species may experience some degree of niche overlap. In our case, we note that life history strategies are similar, and partitioning of space is inevitably driven by inter- and intraspecific competition. Central American cichlids have been observed expelling competitors from territories (McKaye, 1977), indicating that certain territories possess more ideal characteristics than others.
Spatial competition could also involve selecting sites with necessary food resources. It is generally considered that tropical riverine fishes display a high degree of morphological character displacement, so resources are partitioned in such a way that little competition exists (Moyle & Cech, 1994). Limited documentation on the feeding habits of these species while defending their nests exists, however, so the specific requirements for food resources while brooding remain unknown. Underwater video camera traps (GoPro Hero 4 attached to a dive weight) anecdotally recorded *C. salvini* feeding while still providing parental care, whereas we observed *A. spilurus* aggressively defending their broods instead of foraging. Cichlid species richness for this particular assemblage is relatively low – only six species were present, whereas some assemblages in the neotropics host more than 20 species (Cochran-Biederman & Winemiller, 2010). In addition to exhibiting low species diversity, Cochran-Biederman & Winemiller (2010) documented that this particular cichlid assemblage exhibits a high degree of dietary niche segregation, so differences in the requirements for an ideal nest site may explain some of the differences in the variable contribution between species.

Another biotic factor that could affect nesting site selection is predation, which is likely the cause of death for the majority of tropical fishes (Lowe-McConnell, 1975). In nature, habitat selection is partially driven by predation because of the balance between richness of resources and predation risk in habitat patches (Hugie & Dill, 1994). Reproductive behaviors and their predation risks have been studied in fish and generally conclude that vigilance-reducing behaviors (such as nest guarding) increase the risk of predation (e.g. Magnhagen & Vestergaard, 1991; Brick, 1998). During data collection, we observed predation on our study species from both above (several species of kingfisher) and within the river channel (Neotropical river otter). Undoubtedly, unobserved predators were also present. The natural assumption is that nesting
pairs of cichlids would choose nesting locations near fish cover to reduce predation risk. However, the largest cichlid species in this assemblage, *P. splendida*, is primarily a piscivore (Cochran-Biederman & Winemiller, 2010), and was most commonly seen near woody debris during data collection. Their affinity for fish cover could have some influence over why fish cover contributed to the models less than we anticipated. Magnhagen & Vestergaard (1991) suggest that some cooperatively breeding cichlids may utilize the safety-in-numbers survival tactic. Although cichlids in the Bladen do not breed cooperatively outside of their mate, that survival strategy might have some implication for the clustered result of the nearest neighbor analysis.

Tradeoffs between low-risk, low-competition, and low-energy cost nesting locations likely exist. To fully understand what factors influence cichlid nesting sites in this river, additional research on the sources and consequences of biotic interactions needs to be done. Further, we did not examine the success of the broods produced by the nesting cichlids we observed. McKaye (1977) reports observations of cichlids spawning in various habitat conditions, and notes that nests in deeper portions of Lake Jiloa did not experience the same success rate as pairs nesting at shallower depths. Our criterion for denoting a nest site was the presence of a mating pair, and we did not specify stage of spawning or age of the fry. Doing so might offer insight into the quality of the nesting sites as well as address issues with mobility; in some species, older fry can be somewhat mobile, and territory may change (Meral, 1973). We observed this to some degree in *V. maculicauda* and *C. salvini*.

Maxent modeling on a fine scale within reaches presents clear tradeoffs compared to modeling catchment scale variation using reaches as minimum mapping units. The model takes
into account fine details in the differences between presence and pseudo-absence locations, but perhaps at the cost of the model’s applicability to a broader geographical region (Elith et al., 2010), and the ability to capture wide gradients that could account for more variation in species distribution patterns. Use of the independent reach for model validation purposes was intended to test the model’s applicability to a range of habitats outside of the training reach, however the ability to apply Maxent models outside of the geographic regions from which they were generated has been questioned (Baldwin, 2009). In the present study, the calibration reach may function as a transitional zone between the Bladen’s constrained, montane headwaters and its open alluvial floodplain. A two-tailed t-test revealed that the water depths were significantly different between the calibration and validation reaches \((p < 0.0001)\), indicating that the two reaches may not be well-suited to compare. Additionally, the community composition between the two reaches is significantly different (Chi-square value = 8.97, \(p = 0.03\)), with a notable increase in breeding pairs of *A. spilurus* in the upstream reach (22% of community downstream and 47% of community upstream). The community model’s poor performance when evaluated against the independent reach data may also be evidence of larger-scale geomorphic differences between the two reaches. The complications in finding independent, comparable river reaches in this particular area of the Bladen lends support to the reliability of the validation methods we used within the same area. Using the first and second days of data and withholding a subset of testing data from day three were valuable methods in evaluating the robustness of the model within the same environment.

We also acknowledge the limitations and uncertainty involved with the spatial and temporal scale used in this study. Maxent modeling has not been attempted on this fine of a scale before, so placing it into context is difficult. The advantage to using this fine scale is that
exhaustive sampling can be done in a relatively short amount of time, which is helpful when field days are cut short due to heavy, unpredictable tropical rainstorms. The high degree of variability in model output for *T. meeki* could be attributed to the relatively small sample size compared to the other species (n=15), but the number of presence locations is not the smallest used to build models in the literature (Wisz *et al.*, 2008). In other studies, as few as four presence locations have been used to construct Maxent models (Pearson *et al.*, 2007). Pearson *et al.* (2007) concludes that the main application of models built with small sample sizes should be to identify similarities in the environment across presence locations and not necessarily predict the species distribution across a greater extent than was biologically sampled. Because our goal was to determine what microhabitat characteristics were most important in nest site selection and not to project the distribution, the relatively small sample size, particularly in *T. meeki*, should not be problematic.

There are other implicit temporal limitations to the present study as well. Tropical montane watersheds typically encounter high variation in annual precipitation (Wohl, 2005), and our data was collected during the seasonal low flow period. Changes in water depth and velocity, which were shown to be the most important contributing factors in three of the study species and in the overall cichlid community, could affect habitat suitability and accessibility throughout the channel. Additionally, our data represent only a snapshot of the multi-month spawning events by cichlids across the dry season. Because of this, our models should only be applied to low-flow conditions, and additional research is necessary to determine how variations in flow affect nesting site distributions.

Although direct human impacts on the native cichlid assemblage within the Bladen Nature Reserve are likely low, understanding their life history strategies in an undisturbed
environment is important for conservation. Some fish populations are being exploited throughout the unprotected waters of southern Belize, including the larger cichlid species *P. splendida* and *V. maculicauda* (Walker & Walker, 2006). Conserving habitat for these species in the Upper Bladen River may prove useful for replenishment of fish elsewhere (Walker & Walker, 2006), but understanding the factors that influence site selection may aid in identification of other ideal sites for conservation or in future restoration of degraded habitat.
CHAPTER 5
CONCLUSIONS

In this exploratory implementation of maximum entropy modeling in a two-dimensional river environment, we presented some evidence that cichlid nesting site distributions in the Upper Bladen River are related to physical habitat controls. Our models predicted nesting sites of *A. spilurus*, *C. salvini*, and *T. meeki* better than a random distribution, but the distribution of *V. maculicauda* nests was harder to predict. *A. spilurus* was most influenced by water depth, *C. salvini* by substrate type, and *T. meeki* by water velocity. We can loosely accept our hypotheses regarding *A. spilurus*, *C. salvini*, and *T. meeki*, as their distributions were weakly related to the parameters we measured. Despite our AUC values for these species predicting nesting locations with better-than-random success, the influence of these variables on nest site distribution was lower than expected. We reject our hypothesis that *V. maculicauda* nesting sites can be predicted by water depth, velocity, and substrate. Our results partially supported our hypotheses linking body size and nesting site selection, as *A. spilurus* and *T. meeki* preferred the deeper and faster portions of the channel, but *C. salvini* and *V. maculicauda* were not strongly associated with water depth and velocity. Factors driving the spatial distribution of nesting locations are likely a combination of abiotic and biotic, and further research is required to offer more insight into the spatial niches of these cichlid species.
REFERENCES


Hijmans, R. J., & Elith, J. *Species distribution modeling in R*. Retrieved from https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf


Figure 1: Map of study location. A: Country of Belize, black box denotes the Bladen and Monkey Rivers. B: Bladen River, black box shows focal region.
Figure 2: Field sampling photos. A: Observer noting nest location by holding target above nest. B: Depth sampling using stadia rod and velocity sampling using Marsh-McBirney velocity meter.
Figure 3: Model input layers for MaxEnt. A: Depth (m), B: Substrate, C: Velocity (m/s), D: Fish Cover
Figure 4: Observed locations of cichlid nests in calibration reach on day 3 of biological sampling.
Figure 5: MaxEnt presence prediction maps by species. Green dots indicate the known location of breeding pair. A: *A. spilurus*, B: *C. salvini*, C: *V. maculicauda*, and D: *T. meeki*.
Figure 6: MaxEnt predicted presence for cichlids in validation reach. Red dots indicate the known location of a nest. A: *A. spilurus*, B: *C. salvini*, C: *V. maculicauda*, and D: *T. meeki*.
Figure 7: Variable contribution graphs generated by MaxEnt for the calibration reach by species. A: *A. spilurus*, B: *C. salvini*, C: *V. maculicauda*, and D: *T. meeki*.
Figure 8: MaxEnt model outputs for cichlid community. A: Variable contribution. B: Predicted cichlid nesting locations in the validation reach. Red dots indicate known location of a nesting pair. C: Predicted cichlid nesting locations in calibration reach. Green dots indicate known location of a nesting pair.